



Élisabeth Issaieva, Isabelle Nizet
& Lucie Mottier Lopez

Pluralité de contextes, pluralité d'évaluations en éducation : quelles interactions et quels enjeux ?

33^e Colloque de l'ADMEE-Europe, avril 2022 en Guadeloupe

**Pluralité de contextes, pluralité
d'évaluations en éducation :
quelles interactions et quels
enjeux ?**

Portées par l'ADMEE-Europe et l'ADMEE-Canada, les Presses de l'ADMEE publient des ouvrages portant sur l'évaluation en éducation et formation. Leur vocation est de nourrir le débat et d'enrichir les pratiques dans la communauté internationale francophone des chercheurs et praticiens de ces domaines.

Les Presses de l'ADMEE publient des ouvrages scientifiques permettant de présenter les travaux de recherche issus des colloques internationaux, des activités des réseaux scientifiques thématiques, et plus largement de toute problématique portant sur l'évaluation en éducation et formation. Elles publient également des livres plus spécialement destinés aux praticiens, qu'il s'agisse d'analyse de pratiques, de conception d'outils et de dispositifs d'évaluation ou de formation.

www.presses-admee.com

Pluralité de contextes, pluralité d'évaluations en éducation : quelles interactions et quels enjeux ?

Elisabeth Issaieva, Lucie Mottier Lopez et Isabelle Nizet



© 2025 Les Presses de l'ADMEE

Éditeur : ADMEE Europe

c/o IRDP, Faubourg de l'Hôpital 43

Case postale 556, CH-2002 Neuchâtel, SUISSE

Impression : Books on Demand, Norderstedt, Allemagne

Conception couverture : Borsu Olivier - IFRES ULiège

ISBN : 978-2-9701536-5-8

Dépôt légal : Décembre 2025

A toutes les personnes et services institutionnels qui se sont fortement investis dans la réalisation du 33ème colloque de l'ADMEE-Europe, à l'origine de cet ouvrage. Notre pensée particulière va aux membres de nos familles et aux proches partis prématurément pendant la pandémie, ainsi qu'aux collègues qui nous ont quittés dans les dernières années. Leurs valeurs n'ont jamais cessé de nous porter alors qu'un soutien collectif remarquable permettait à chacune et chacun d'entre nous de persévérer dans des conditions très complexes et exigeantes. Merci, Mèsi (en créole guadeloupéen).

Table des matières

Introduction : 'Pluralité de contextes, pluralité d'évaluations en éducation : quelles interactions et quels enjeux ?' 12

I.	Multi-dimensionnalité et complexité du processus évaluatif	12
II.	Pluralité de contextes évaluatifs, pluralité d'interactions et d'enjeux	13
III.	Diversification et transformations des contextes et des enjeux évaluatifs	14
IV.	Et face à des bouleversements contextuels inattendus et soudains, quels cheminements ?	15
V.	Présentation de la structure de l'ouvrage	17
VI.	Bibliographie	22

Chapitre 1. Mesures, démesures, demi-mesures, contre-mesures, sur-mesure... 25

I.	Avant-propos	25
II.	La mesure créole du monde	26
III.	Mesures et démesures de la colonisation	27
IV.	Demi-mesures : données chiffrées et statistiques	29
V.	Contre-mesures	30
VI.	En guise de conclusion, partir vers l'école du sur-mesure	32
VII.	Bibliographie	32

Chapitre 2. Les normes de l'évaluation portées par les chefs d'établissement et les inspecteurs 34

I.	Introduction	34
II.	Le cadre de la recherche	36
III.	Méthodologie	40
IV.	Résultats	42
V.	Synthèse et perspectives	46
VI.	Bibliographie	48

Chapitre 3. L'évaluation formative et l'apprentissage du langage : liens d'interdépendance entre la régulation et la micro-culture de la classe 52

I.	Introduction	52
II.	Cadre théorique	54
III.	Méthodologie	58
IV.	Résultats	61
V.	Conclusion	66
VI.	Bibliographie	67

Chapitre 4. La planification stratégique : pluralités de contexte, pluralités de conceptions **77**

I.	Introduction	77
II.	Les plans stratégiques : une réflexion sur la complexité et la diversité	78
III.	Le plan stratégique et ses composantes	79
IV.	Une typologie des conceptions des universités en vue d'en évaluer leur performance	81
V.	Méthodologie	85
VI.	Résultats	87
VII.	Discussion et conclusion	90
VIII.	Bibliographie	91

Chapitre 5. L'évaluation aujourd'hui et demain. Quelles ruptures contextuelles ? Quelles ruptures culturelles ? **95**

I.	Introduction	95
II.	Continuité en évaluation : un système séculaire	96
III.	Une continuité sur fond de dissensions	98
IV.	Du chronosystème au microsystème de l'évaluation	99
V.	Et pourtant... l'évaluation évolue	101
VI.	À la recherche de la valeur pédagogique de l'évaluation	103
VIII.	L'urgence de l'équité et ses limites : une rupture culturelle axiologique	105
IX.	L'enjeu du rapport à la vérité en évaluation : une rupture culturelle épistémologique	106
X.	Invitation à une certaine émancipation culturelle	108
XI.	Bibliographie	109

Chapitre 6. Expérimenter une évaluation collaborative en formation à l'enseignement. Une interface entre les contextes évaluatifs ? **114**

I.	Introduction	114
II.	Problématisation et cadrage conceptuel	115
III.	Contexte de l'étude	117
IV.	Objectifs de l'étude	119
V.	Méthodologie	119
VI.	Présentation des résultats	121
VII.	Discussion conclusive	126
VIII.	Bibliographie	129

Chapitre 7. Co-construire un référentiel de compétences pour évaluer des articles scientifiques **133**

I.	Introduction	133
II.	Cadre théorique de notre étude	136
	L'objet « référentiel »	136
	Le paradigme de l'évaluation située	136

	La référentialisation	137
	Une enquête collective	138
	Questionnement de recherche	139
III.	Contexte de l'étude et méthodologie	140
IV.	Résultats	142
V.	Discussion conclusive	149
VI.	Bibliographie	150

Chapitre 8. Évaluer la qualité d'un dispositif de formation en ligne **153**

I.	Contexte et problématique	153
II.	Cadre théorique	155
III.	Méthodologie	160
IV.	Résultats	165
V.	Discussion	170
VI.	Conclusion	172
VII.	Bibliographie	173

Chapitre 9. Évaluation des contextes et contextes d'évaluation dans la recherche-intervention **176**

I.	Introduction	176
II.	Une démarche de recherche participative : la recherche-intervention (RI)	177
III.	L'évaluation de la RI	185
IV.	Conclusion : Vers un tiers-espace de l'évaluation de la RI (TEERI)	190
V.	Bibliographie	191

Chapitre 10. La « complexité » dans l'évaluation des compétences : des défis méthodologiques et contextuels **193**

I.	Introduction et problématisation	193
II.	Cadrage Théorique	195
III.	Méthodologie	197
IV.	Résultats	202
V.	Discussion	212
VI.	Conclusion	215
VII.	Bibliographie	216

Chapitre 11. L'évaluation au prisme de la contextualisation : l'exemple du système éducatif français **220**

II.	Les outre-mer français comme terrain d'investigation	221
III.	De la modélisation des contextes à la transposition didactique	222
IV.	Contexte territorial et diplômes nationaux : l'exemple du Brevet de technicien supérieur (BTS)	224
V.	Une évaluation des établissements d'enseignement supérieur contextualisée ? Le cas de l'université des Antilles	227

VI.	Discussion et conclusion : vers une évaluation plus contextualisée ?	230
VII.	Bibliographie	233

Conclusion : 'De quels contextes parle-t-on pour problématiser l'évaluation ?'

237

I.	Introduction	237
II.	De quel(s) contexte(s) parle-t-on ?	238
III.	Perspective située de l'activité sociale et humaine	243
IV.	L'évaluation située, le choix d'un ancrage épistémologique	245
V.	Bibliographie	249

Les contributeurs

251

Introduction : 'Pluralité de contextes, pluralité d'évaluations en éducation : quelles interactions et quels enjeux ?'

Elisabeth Issaieva, Lucie Mottier Lopez et Isabelle Nizet

I. Multi-dimensionnalité et complexité du processus évaluatif

Tout processus évaluatif possède des facettes multiples (définir les objets, les attentes, les objectifs, les référents, les référés, les critères ; collecter et confronter des informations), complexes (interpréter, attribuer des significations, prendre des décisions) et peut avoir des effets variés (De Ketele, 2006, 2010 ; Figari, 1994 ; Hadji, 1997). Les saisir demande de caractériser les contextes (et les niveaux contextuels) dans lesquels le processus évaluatif s'inscrit, les fonctions remplies et les démarches poursuivies (De Ketele, 2006, 2010). Les fonctions et les démarches de l'évaluation ont fait l'objet de nombreux travaux et de thèmes de colloques de l'ADMEE-Europe depuis la fondation de l'association en 1986. Par ailleurs, la question de l'importance des contextes évaluatifs (de leurs impacts et de leurs relations) pour cerner l'évaluation comme objet de recherche, pratique évaluative ou outil méthodologique a été mise en avant par plusieurs chercheurs et chercheuses (De Ketele, 2006 ; Figari et Achouche, 2001 ; Mottier Lopez, 2009 ; Detroz et al., 2017 ; Gremion et al., 2019) mais son investigation de manière empirique et systémique a débuté récemment, surtout dans ses aspects multidimensionnels et interactifs.

II. Pluralité de contextes évaluatifs, pluralité d'interactions et d'enjeux

De Ketele (2010) insiste sur l'importance de préciser le niveau contextuel (micro, méso ou macro) au sein duquel une pratique évaluative donnée s'inscrit et de s'interroger sur les problématiques que cela peut soulever, certaines étant toujours actuelles. Par exemple, au niveau micro-contextuel, comprenant les caractéristiques de l'enseignement et de l'apprentissage en classe ou en groupe de formation, une interrogation pointée depuis longue date et toujours d'actualité concerne la manière dont les enseignants et enseignantes identifient les obstacles et les besoins d'apprentissages des élèves et la pertinence des actions proposées. Au niveau méso-contextuel désignant le plan institutionnel (école, établissement, institut de formation), la question de la construction d'une culture commune évaluative et d'une harmonisation des pratiques se pose actuellement en termes par exemple de collaboration et de concertation engendrant des décisions collectives. Au niveau macro-contextuel qui renvoie au système éducatif ou de formation, la prise en compte de l'évolution des enjeux sociétaux pour concevoir et ajuster des curriculums pertinents constitue une problématique récurrente. Ces quelques exemples sont révélateurs d'enjeux diversifiés, traduisant la complexité des interactions qui existent au sein d'un même niveau contextuel. Et qu'en est-il des relations entre les niveaux contextuels (micro, méso, macro) ? Peut-on parler de multi-contextualité ? Cette question a aussi attiré l'attention des chercheurs et des chercheuses. Les premiers travaux se sont inscrits dans un mouvement plutôt descendant, étudiant davantage l'influence du niveau macro-contextuel sur les niveaux méso et micro-contextuel, de manière unidirectionnelle. Une perspective interactionniste et complexe a ensuite émergé, avec les travaux de Mottier Lopez (2008) qui se sont attachés à dévoiler des relations entre des contextes variés, situés à des niveaux différents, vus comme constitutifs d'un processus dynamique d'influences réciproques. Par exemple, une analyse de l'élaboration d'un dispositif d'évaluation par des formateurs et formatrices dans un institut de formation (niveau méso-contextuel) montre que celui-ci ne subit pas passivement une influence descendante mais qu'il s'élabore en interaction et par réajustements successifs avec le niveau micro-contextuel, tout en tenant compte des pratiques évaluatives situées (Mottier Lopez et Van Nieuwenhoven, 2009). Dans une autre recherche, Mottier Lopez (2013) met en évidence comment des enseignant-es négocient des cultures d'évaluation en prenant en considération à la fois des référents institués et

standardisés et des référents informels ancrés dans des savoirs d'expérience. Malgré la portée de ces travaux dans une perspective interactionniste, des réflexions et des investigations empiriques soutenues apparaissent nécessaires afin de mieux cerner la complexité et la pluralité des relations entre contextes au regard d'enjeux évaluatifs. A ce titre, une problématisation centrale a émergé autour des interactions et des tensions existant entre des évaluations de niveaux contextuels différents tels que le pilotage des systèmes et la régulation des apprentissages. Cette problématique a inspiré les débats du 20ème colloque de l'ADMEE-Europe en 2008 (Mottier Lopez et Crahay, 2009). Six ans plus tard, en 2014, l'interrogation sur les enjeux évaluatifs dans des contextes socio-culturels, pédagogiques et didactiques variés est mise en lumière lors du 27ème colloque de l'ADMEE-Europe (Detroz et al., 2017). En 2019, le 31ème colloque de l'ADMEE-Europe (Gremion et al., 2019) soutient la poursuite d'un ensemble de questionnements autour des tensions et synergies possibles entre des paramètres contextuels à l'échelle individuelle, institutionnelle et nationale. La prise en compte des contextes pluriels dans l'évaluation a fait l'objet d'un bilan sur l'évolution des activités de l'ADMEE-Europe par Allal et al. (2017). Un numéro thématique de la revue Contextes et didactiques (Issaieva et Mottier Lopez, 2017) consacré à cette problématique a réuni des articles traitant de la nature des relations entre certains paramètres contextuels au regard d'enjeux évaluatifs ou de leur modélisation (Mottier Lopez et Dechamboux, 2017).

III. Diversification et transformations des contextes et des enjeux évaluatifs

Une autre raison pour laquelle l'analyse de l'évaluation à la lumière de paramètres contextuels multiples reste d'une grande actualité est liée à la diversification et l'étendue des objets, des enjeux et des contextes évaluatifs, ainsi qu'à leur élargissement à différents champs de l'éducation et de la formation (Detroz et al., 2017). Dans nombre de sociétés, des changements d'ordre politique, socioéconomique, culturel n'ont cessé de s'opérer ces dernières années provoquant l'instabilité et la transformation des contextes en créant des tensions productives ou contre-productives, des controverses ainsi que des effets contrastés. Ainsi, au cours des dix dernières années, des crises économiques successives ou imbriquées, des conflits socio-politiques et culturels menant à des flux migratoires en contexte de mondialisation croissante ont provoqué des changements, de nature variable, locale ou globale, plus ou moins anticipés et marquants.

L'hétérogénéisation de contextes multiculturels et plurilingues a été renforcée et leur diversification s'est accrue. À cet égard, plusieurs systèmes éducatifs ont mis en valeur l'éducation interculturelle et se sont engagés dans la voie de l'inclusion sociale et scolaire. Malgré une insistance sur des valeurs interculturelles et inclusives fortes, les transformations réelles des pratiques en faveur de plus grandes validité, équité et qualité, peinent à se réaliser (Hadji, 2019). En évaluation, la prise en compte de la diversité et de l'inclusion reste un défi, car elle suppose que l'on considère les potentiels, les obstacles, et les besoins d'apprentissages de chaque élève avec comme corollaire la double contrainte d'une potentielle réduction des exigences d'enseignement et du maintien de la valeur des diplômes pour des élèves en grande difficulté ; cela suppose également de rendre l'enseignement et l'évaluation souples, flexibles sans perdre de vue leurs finalités éducatives multiples que sont l'accès aux savoirs transversaux ou disciplinaires, la socialisation, l'épanouissement et l'émancipation. Face à cette multitude de paramètres, des professionnel·les de l'éducation et de la formation se demandent de plus en plus comment articuler les besoins d'apprentissage individuels et collectifs dans des situations très diversifiées ? Ces dilemmes peuvent-ils réduire leur pouvoir d'agir, ou bien mener à des tensions ingérables aux retentissements potentiellement problématiques sur les enjeux éducatifs et évaluatifs et donc aussi sur les publics concernés ? Compte tenu de ces réflexions, on peut se demander à quelles conditions la diversité et la richesse des contextes peuvent générer ou conduire à des synergies constructives et formatrices ? Tenter de répondre à ces questions ne signifie pas s'enfermer dans un seul modèle explicatif ou privilégier certains modèles au détriment d'autres. Bien au contraire, cela veut dire qu'il est particulièrement important de définir précisément les concepts et paramètres contextuels rentrant en jeu dans une situation évaluative donnée et de comprendre comment s'opérationnalisent les mécanismes de leurs interactions au regard d'enjeux évaluatifs visés.

IV. Et face à des bouleversements contextuels inattendus et soudains, quels cheminements ?

En plus de ces transformations, nos sociétés sont confrontées à des bouleversements contextuels inattendus et soudains, comme le montre la crise sanitaire provoquée par la pandémie du Covid-19. Elle a déstabilisé nos systèmes politiques, socio-économiques et éducatifs en provoquant des ruptures contextuelles. Elle incite à se questionner encore plus sur le

réel pouvoir d'action des systèmes éducatifs et des acteurs et actrices scolaires et sur les effets que cela peut provoquer sur les personnes concernées.

Dans ces conditions d'instabilité, d'incertitude importante et de rupture contextuelle, de nombreuses interrogations émergent. Sommes-nous en mesure de continuer à enseigner et à évaluer face à des contraintes de nature différente mais exceptionnelle ? Comment les systèmes et les pratiques ont-elles dû s'adapter par rapport à l'enseignement à distance en contexte de crise sanitaire ? Eu égard à quelles ressources ? Face à quels obstacles ? Et si l'éducation et la formation ont pu résister, plus ou moins, au séisme provoqué par la crise sanitaire (Weiss et Issaieva, 2022), quelles sont les conséquences et les leçons à tirer à plus long terme ? Et que retenir globalement de cela pour le futur ?

Ce sont ces questionnements et réflexions qui ont propulsé l'idée de l'organisation du 33ème colloque de l'ADMEE-Europe autour du thème « Pluralité de contextes, pluralité d'évaluations en éducation : quelles interactions et quels enjeux ? ». Ayant eu lieu en avril 2022, pour la première fois en contexte caribéen et en plein contexte de bouleversement lié à la crise sanitaire, ce colloque a pu interroger ce thème de manière très investie scientifiquement et humainement. Les nombreuses communications présentées en présentiel mais aussi à distance ont permis de réunir physiquement et virtuellement des chercheurs et chercheuses ainsi que des professionnel·les de l'éducation et de la formation provenant de plusieurs continents pour débattre avec ferveur les enjeux transversaux suivants :

- Identifier l'impact des différents contextes pluriels au niveau macro (par exemple : contexte économique, historique, socioculturel ; politiques et cultures éducatives, pilotage d'un système), méso (par exemple : établissements, équipes disciplinaires, etc.) et micro (par exemple : classe, discipline, contenu, acteur·rices) sur l'évaluation et inversement ;
- Modéliser les relations et interactions entre les contextes et les différentes formes d'évaluations ;
- Rendre compte des éventuelles tensions et transformations liées à la relation ou à l'interaction entre contextes au regard d'enjeux évaluatifs ;
- Comprendre les mécanismes pouvant expliquer l'impact réciproque entre contextes au regard d'enjeux évaluatifs.

Ces enjeux se reflètent fortement dans les trois axes thématiques que nous avons proposés comme fil conducteur des travaux et des débats lors du colloque. Ce sont ces mêmes enjeux et axes thématiques qui constituent la trame de structuration du présent ouvrage.

V. Présentation de la structure de l'ouvrage

L'ouvrage s'articule autour des trois axes thématiques traités lors du 33ème colloque de l'ADMEE-Europe.

Les textes qui composent cet ouvrage sont issus des conférences, de travaux de recherche scientifiques et d'analyses réflexives de pratiques présentés durant le colloque. La richesse de ces contributions témoigne de la diversité des contextes générateurs d'idées pertinentes à notre problématique.

Axe 1 - L'agir évaluatif en situation scolaire et en formation universitaire ou professionnelle : entre des contextes pluriels ?

Cet axe thématique vise à appréhender l'agir évaluatif des enseignant-es et des formateur-rices dans différents ordres et niveaux de formation, allant de l'enseignement primaire à l'enseignement supérieur, ou à la formation professionnelle. Il réunit les quatre premiers chapitres de l'ouvrage. L'agir évaluatif peut être saisi à travers des cadres théoriques, spécifiques et plusieurs questions peuvent être abordées pour mettre en lumière des actions, des représentations, des croyances en tenant compte des contextes et de la diversité des référents (pensées auto-référenciées, éléments du micro, méso ou macrocontexte, etc.). Il s'agit, par exemple, d'expliquer le processus sous-jacent des interactions entre contextes évaluatifs pluriels convoqués par le corps enseignant, mais aussi d'identifier les éléments qui sont réellement mobilisés dans l'agir évaluatif (Mottier Lopez, 2017). Quelles sont les contraintes qui pèsent et lesquelles sont anticipées ? Sont-elles révélatrices de préoccupations (de validité, faisabilité, utilité de l'évaluation, etc.) complémentaires ou contradictoires (Tourmen et Mayen, 2012) ? Comment des décalages éventuels entre des contextes nationaux ou locaux sont-ils gérés et au prix de quels compromis ? Comment l'agir évaluatif interagit-il avec l'agir auto-évaluatif des élèves ou des étudiant-es (Hadji, 2012) ? Comment identifier les besoins éducatifs particuliers des élèves et de quels paramètres contextuels tenir compte ou pas ? Comment différencier au mieux, sans

stigmatiser, et réguler les activités des élèves en vue de soutenir les progrès ? Telles sont les questions qui seraient soulevées dans les chapitres réunis autour de cet axe thématique 1 de l'ouvrage.

L'ouvrage débute avec le chapitre 1, de Lambert Félix Prudent, tiré de sa conférence « Mesures, démesures, demi-mesures, contre-mesures... Evaluer dans une société créole ». En partant de l'hypothèse que la non prise en compte des éléments contextuels linguistiques et culturels dans le processus évaluatif mène à des décalages importants, l'auteur propose une analyse sociolinguistique pour comprendre les failles et les possibles de l'agir évaluatif dans un contexte antillais complexe, plus précisément au sein de l'école guadeloupéenne.

Le chapitre 2, proposé par Yann Mercier-Brunel, a pour titre « Les normes de l'évaluation portées par les chefs d'établissement et les inspecteurs ». Ce chapitre expose les résultats d'une recherche qui vise à comprendre comment les cadres scolaires (inspecteurs et chefs d'établissement) et les enseignant-es s'emparent des réformes en matière d'évaluation en France et en quoi leurs représentations constituent des soutiens ou des obstacles pour l'évolution des pratiques évaluatives.

Avec le troisième chapitre, intitulé « La relation entre l'évaluation formative, le langage et le contexte scolaire pour soutenir l'apprentissage en classe », Janet Cerda examine l'agir évaluatif dans un autre contexte, multiculturel, à savoir le contexte états-unien. A partir d'une étude de cas qualitative illustrant les interrelations entre l'évaluation formative, le langage, l'autorégulation et la co-régulation, une réflexion est portée sur les apports et les limites de ces processus et de leurs interactions.

Le quatrième et dernier chapitre reflétant l'axe thématique 1 du colloque s'intitule « La planification stratégique des universités canadiennes : pluralité de conceptions, pluralité de contextes ». Il est proposé par Catherine Larouche, Ashley Byrne, Denis Savard, Camille Larouche et Lucie Héon qui exposent les résultats d'une recherche visant à analyser le processus de planification stratégique des universités canadiennes comme révélateur des divers changements contextuels d'ordre socio-économique et politique.

Axe 2 - Former à l'évaluation à la croisée de contextes multiples : vers une culture plurielle ?

Dans des périodes de réformes successives des systèmes éducatifs, aux enjeux plus ou moins antagonistes, oscillant entre des modèles de

formations « traditionnelles à dominante instructives et applicatives » (Boudjaoui et Clénet, 2011) ou à l'inverse plus axées sur le développement professionnel, l'interrogation sur la construction des savoirs professionnels en évaluation (Jorro, 2009 ; Nizet, 2013, 2014) est importante. En prenant en considération les quatre registres de l'agir évaluatif (culturel, identitaire, fonctionnel et symbolique), on peut s'interroger sur leurs interactions à la lumière de la diversité des contextes, et des interrogations complémentaires peuvent être envisagées. Par exemple, quels sont les référents privilégiés actuellement selon les programmes dans des contextes différents ? Quels sont leurs limites et leurs failles ? Comment ces référents sont-ils concrètement déployés dans les formations ? Comment sont-ils interprétés par les formés ? Quels sont les « transferts » possibles vers le terrain ? Quelle est la prégnance des cultures évaluatives personnelles et collectives sur la construction de savoirs professionnels en évaluation (Nizet, 2013, 2014 ; Nizet et Leroux, 2015) ?

La suite de l'ouvrage cible ces questionnements. Ainsi, le chapitre 5 d'Isabelle Nizet, intitulé « L'évaluation aujourd'hui et demain : Quelles ruptures contextuelles/ruptures culturelles ? » fait suite à sa conférence donnée lors du colloque. L'autrice présente une analyse fine de nombreuses ruptures contextuelles générées ces derniers temps et les discontinuités et difficultés que cela provoque en matière évaluative. A travers cette analyse, une piste intéressante est développée pour la formation en mettant l'accent sur une authentique « évaluation culturelle » afin de préserver le pouvoir d'agir des tous les acteurs et actrices.

Le chapitre 6, proposé par Gonzague Yerly est intitulé « Expérimenter une évaluation collaborative en formation à l'enseignement. Une interface entre les contextes évaluatifs ? ». Dans ce texte, l'auteur analyse l'expérience de futur-es enseignant-es de l'Université de Fribourg, en Suisse, afin de cerner leurs vécus d'une évaluation collaborative, notamment des feedbacks entre pairs et les perspectives qu'ils envisagent (ou non) pour leur propre enseignement et évaluation dans les classes, à la suite de cette expérience.

L'axe thématique 2 se termine avec le chapitre 7, intitulé « Co-construire un référentiel de compétences pour évaluer des articles scientifiques », proposé par Lionel Dechamboux, Lucie Mottier Lopez, Céline Girardet et Natalia Constenla Martinez. Comme son titre l'indique, il présente une démarche de co-construction d'un référentiel de compétences en matière d'évaluation d'articles scientifiques à l'Université de Genève en Suisse. Cette démarche collective a été conduite dans le cadre d'un

accompagnement doctoral à l'écriture et à l'évaluation scientifiques. Les résultats révèlent comment le processus de référentialisation collective s'inscrit dans des vécus biographiques individuels dans des contextes culturels différents.

Axe 3- L'évaluation de dispositifs et de politiques de formation : comment conjuguer des sources multiples ?

L'axe thématique 3 du colloque tentait de saisir la complexité de l'évaluation des dispositifs de formation qui s'articule théoriquement entre plusieurs finalités : estimer la productivité mais aussi assurer le suivi et la régulation de l'ensemble du processus de diagnostic, d'analyse et de conception ; diagnostiquer les points forts et faibles, les points de vigilances (dans les principales étapes de l'ingénierie, de l'analyse des besoins et la conception ou l'appropriation du cahier des charges à l'appropriation des objectifs de formation). Compte tenu de cette complexité, ainsi que de la diversité des acteur·rices (externes, internes) parties prenantes, la multitude des démarches évaluatives possiblement impliquées peut être interrogée : sont-elles convergentes ou contradictoires ? Quels sont les référents convoqués ? A partir de quelle référentialisation (Figari et Remaud, 2014) ? A quelles difficultés et contraintes contextuelles les démarches se heurtent-elles ? Prennent-elles en compte le même type de réalités, voire les différents aspects des réalités concernées par le dispositif sous-jacent ?

Le chapitre 8 s'inscrit dans cet axe thématique 3 du colloque. Il interroge l'articulation de démarches plurielles requises pour évaluer la qualité d'un dispositif de formation en ligne, développé dans le cadre d'un programme de formation à l'enseignement de la lecture en Belgique francophone. Ce chapitre est proposé par Olivier Leyh, Charlotte Dejaegher, Marine André, Patricia Schillings et s'intitule « Évaluer la qualité d'un dispositif de formation en ligne ».

Le chapitre 9 qui suit est proposé par Jean-François Marcel à l'issue de sa conférence donnée lors du colloque. Il a pour titre « Evaluation des contextes et contextes d'évaluation dans la recherche-intervention. Vers la théorisation d'un tiers-espace de l'évaluation ». Ce chapitre apporte des éléments de réponse aux questions soulevées par l'évaluation d'une démarche de recherche-intervention mettant en relation les contextes hétérogènes suivants : le contexte académique (dont relève la recherche) et un contexte socio-professionnel en France (dont relève l'intervention). Les interrogations portent aussi sur les articulations entre ces deux contextes :

leurs visées, leurs fonctionnements, leurs temporalités, leurs modes de validation, etc.

L'ouvrage se poursuit avec le chapitre 10, proposé par Fadi El Hage et Joe Ephrem et qui s'intitule : « La « complexité » dans l'évaluation des compétences : des défis méthodologiques et contextuels. Cas du Baccalauréat arabe international (IAB) ». Les auteurs s'intéressent à un dispositif de formation dans un contexte multiculturel, multiconfessionnel et complexe qui est celui du Liban. Le dispositif de formation qui fait l'objet de l'analyse est le Baccalauréat arabe international (IAB). Les auteurs présentent une recherche menée pour concevoir un outil d'évaluation quantitative des compétences dans le cadre de ce programme. Les résultats montrent comment le contexte socio-politique influence le développement de certaines compétences chez les étudiants et étudiantes.

L'axe thématique 3 se termine avec le chapitre 11, d'Antoine Delcroix dans la continuité de sa conférence « L'évaluation au prisme de la contextualisation : l'exemple du système éducatif français. Vers une théorie de l'évaluation contextualisée ? ». Ce chapitre porte une réflexion sur le système d'enseignement supérieur français et les démarches d'assurance qualité. Il part de la considération que les démarches d'assurance-qualité mènent souvent à des processus normalisés et standardisés, conçus de manière relativement rigide, aussi bien par les critères d'évaluation retenus que par l'infrastructure technique associée. Ceci pourrait être vécu par les actrices et acteurs locaux comme inapproprié, voire injuste et provoquer une défiance à l'égard de l'évaluation. Le texte propose aussi un débat plus général sur l'articulation entre le contexte local et global et tente de contribuer au questionnement sur la nécessité d'une théorie de l'évaluation contextualisée.

L'ouvrage s'achève avec le chapitre conclusif « De quels contextes parle-t-on pour problématiser l'évaluation ? » écrit par Lucie Mottier Lopez à la suite de sa conférence conclusive en tant que présidente du comité scientifique du 33ème colloque de l'ADMEE-Europe. A partir des trois axes thématiques du colloque et des communications présentées, l'autrice problématise des enjeux de « polycontextualité » dans une perspective d'évaluation située élargie. La réflexion et l'analyse s'intéressent notamment aux phénomènes spécifiques que l'intégration de différents contextes et leurs mises en relation plus ou moins constructives engagent dans les pratiques et processus d'évaluation en éducation et en formation.

VI. Bibliographie

- Allal, L., Dauvisis, M.-C. et De Ketele, J.-M. (2017). L'ADMEE-Europe, née à Dijon en 1986 : développements et perspectives. *Evaluer. Journal international de Recherche en Education et Formation*, 3(3), 107-137. <https://doi.org/10.48782/d0fw5y51>
- Boudjaoui, M. et Clénet, J. (2011). Les enjeux d'une pédagogie de l'alternance dans la formation des enseignants. Les apports d'expériences françaises de formation et de formateurs d'adultes. Dans P. Maubant, J. Clénet et D. Poisson (dir.), *Débats sur la professionnalisation des enseignants. Les apports de la formation des adultes* (p. 13-56). Presses de l'Université du Québec.
- De Ketele, J.-M. (2006). La recherche en évaluation : propos synthétiques et prospectifs. *Mesure et évaluation en éducation*, 29(1), 99-118.
- De Ketele, J. (2010). Ne pas se tromper d'évaluation. *Revue française de linguistique appliquée*, XV(1), 25-37. <https://doi.org/10.3917/rfla.151.0025>
- Detroz, P., Crahay, M. et Fagnant, A. (2017). *L'évaluation à la lumière des contextes et des disciplines*. De Boeck. <https://doi.org/10.3917/dbu.detro.2017.01>
- Figari, G. (1994). *Évaluer : quel référentiel ?* De Boeck.
- Figari, G. et Achouche, M. (2001). *L'activité évaluative réinterrogée : regards scolaires et socioprofessionnels*. De Boeck. <https://doi.org/10.3917/dbu.figar.2001.01>
- Figari, G. et Remaud, D. (2014). *Méthodologie d'évaluation en éducation et formation, ou l'enquête évaluative*. De Boeck.
- Gremion, C., Sylvestre, E. et Younes, N. (2019). Actes du 31ème Colloque scientifique international de l'ADMEE-Europe : *Entre normalisation, contrôle et développement formatif. Évaluations sources de synergies ?* Université de Lausanne, Suisse.
- Hadji, C. (1997). *L'évaluation démystifiée : mettre l'évaluation scolaire au service des apprentissages*. ESF.
- Hadji, C. (2012). *Faut-il avoir peur de l'évaluation ?* De Boeck.
- Hadji, C. (2019). Pour une évaluation humaniste, *Contextes et didactiques*, 13. <https://doi.org/10.4000/ced.811>

- Issaieva, E. et Mottier Lopez, L. (2017). Éditorial : contextes et évaluations. Quelles interactions ? *Contextes et didactiques*, (9), 6-10.
- Jorro, A. (dir.). (2009). *La reconnaissance professionnelle en éducation : évaluer, valoriser, légitimer*. University of Ottawa Press.
- Mottier Lopez, L. (2008). *Apprentissage situé : la microculture de classe en mathématiques*. Peter Lang.
- Mottier Lopez, L. (2009). L'évaluation en éducation : des tensions aux controverses. Dans L. Mottier Lopez et M. Crahay (dir.), *Évaluations en tension : entre la régulation des apprentissages et le pilotage des systèmes* (p. 7-25). De Boeck.
- Mottier Lopez, L. (2013). Pour une culture partagée de l'évaluation. Du retour à des pratiques d'évaluation traditionnelles vers la construction d'un nouveau rapport entre la recherche, le terrain et le politique. *Revue Éducation Canada*. <http://www.ceaa-ce.ca/education-canada/article/pour-une-culture-partagée-de-l'évaluation>
- Mottier Lopez, L. (2017). Une modélisation pour appréhender la référentialisation dans l'évaluation des apprentissages des élèves. Dans P. Detroz, M. Crahay et A. Fagnant (dir.), *L'évaluation à la lumière des contextes et des disciplines* (p. 169-192). De Boeck.
- Mottier Lopez, L. et Crahay, M. (2009). *Evaluations en tension : entre la régulation des apprentissages et le pilotage des systèmes*. De Boeck.
- Mottier Lopez, L. et Van Nieuwenhoven, C. (2009). Co-élaboration d'une pratique d'évaluation entre formateurs : quelles relations de régulation au sein de la communauté de pratique ? Dans L. Mottier Lopez et M. Crahay (dir.), *Évaluations en tension : entre la régulation des apprentissages et le pilotage des systèmes* (p. 181-199). De Boeck.
- Mottier Lopez, L. et Dechamboux, L. (2017). D'un référentiel d'évaluation fixe à une co-constitution référentielle dynamique, ce que nous apprend le jugement situé de l'enseignant. *Contextes et didactiques*, 9, 12-29.
<https://journals.openedition.org/ced/706?lang=en>
- Nizet, I. (2013). Construction de savoirs professionnels en évaluation : Étayages culturels d'enseignants formés en ligne. Dans *Actes du*

25ème colloque de l'ADMEE-Europe : Évaluation et autoévaluation, quels espaces de formation ? Fribourg, Suisse.

- Nizet, I. (2014). Formation des enseignants du secteur de l'éducation des adultes à l'évaluation des compétences : défis et perspectives. *Revue des sciences de l'éducation*, 40(2), 373-396.
- Nizet, I. et Leroux, J. L. (2015). La construction de savoirs conceptuels en évaluation : enjeux pour un développement professionnel en contexte de formation continue. *Evaluer. Journal international de Recherche en Education et Formation*, 1(2), 15-29. <https://journal.admee.org/index.php/ejiref/article/view/45>
- Tourmen, C. et Mayen, P. (2012). Les évaluateurs, des épistémologues en actes ? *Raisons Educatives*, 16, 63-75.
- Weiss, P-O. et Issaieva, E. (2022). L'École a-t-elle résisté au séisme du COVID-19 ? Contextes et didactiques. <https://doi.org/10.4000/ced.3800>

Chapitre 1. Mesures, démesures, demi-mesures, contre-mesures, sur-mesure...

Évaluer dans une société créole.

Lambert Félix Prudent

I. Avant-propos

Mon propos est de sensibiliser une assemblée scientifique à l'originalité de pratiquer une évaluation de l'enseignement en milieu créole. Pour dresser ma problématique, je vais parler du fondement de cette société, la créolisation, de l'installation de l'école française en Guadeloupe et en Martinique, du contact des cultures en cours et de la difficulté pour une métropole de mesurer les mœurs, les pratiques de son ancienne colonie. Dans le fond, je pose la question de la valeur des mesures, chiffrées, statistiques et autres qui sont prises par des institutions et des agents (les enseignants) en revenant à la valeur de la loi, de la norme décrétée à Paris et promulguée valable à Pointe à Pitre, à Fort-de-France et dans les collectivités d'Outre-mer de la Caraïbe.

Maintenant, j'insisterai sur la dimension linguistique. Aujourd'hui deux langues se répartissent l'énonciation des Antillais. L'école se trouve au cœur de ce bilinguisme-biculturalisme, traitant la question en monolinguisme de fait. Dépassant le concept de diglossie, l'une, la plus forte, domine l'autre, la plus faible. Calvet en 1974 avait inventé le concept de glottophagie pour caractériser ce type de rapport inégal. Les choses sont un peu plus compliquées. Et nous allons devoir passer à une approche plus dynamique pour comprendre le contact des langues et la difficulté de les cantonner dans des systèmes bien établis et organiser l'évaluation en Guadeloupe et en Martinique. La plus faible mange aussi des aspects de la

plus forte, et la population parle aussi bien du créole francisé que du français créolisé, ce qui laisse l'évaluateur désespéré.

Comment fonctionne le système ? Des dispositifs baptisés eux-mêmes d'évaluation existent. Chaque année le système français demande à des maîtres d'utiliser des méthodes d'évaluation pour mieux bâtir leur pédagogie au préélémentaire, à l'entrée du CP et de la sixième, les maîtres reçoivent un test national qu'on leur demande d'appliquer à leurs élèves. Je vais tenter avec vous d'examiner ce que ce test peut mesurer, et vous montrer qu'en fait les évaluateurs ne mesurent les performances que d'une partie de la classe.

II. La mesure créole du monde

« ... Je suis arrivé aux Indes, j'y ai trouvé quantité d'îles peuplées de gens sans nombre et comment d'elles toutes j'ai pris possession pour Leurs Altesses et sans que nul n'y contredit.

A la première que j'ai rencontrée, j'ai donné le nom de San Salvador, en hommage à Sa Haute Majesté. Les Indiens appellent cette île Guanahani. J'ai nommé la deuxième île Santa Maria de la Concepcion, la troisième Fernandina, la quatrième Isabela, la cinquième Juana et ainsi à chacune d'elles un nom nouveau ... », (Colomb, Lettre à Santangel, 1493)

En arrivant aux Antilles en 1492, Christophe Colomb note le caractère accueillant des habitants indigènes et présume de la bonne fortune de l'aventure coloniale séculaire de ses successeurs : « Ils sont tendres et gentils, ils feront de bons serviteurs ». Dans sa vision du monde, (il a lu Marco Polo et il a fait le voyage pour prouver qu'une route maritime était praticable), il est arrivé aux Indes et ces peuples qui vivent dans les bois seront désormais des *Sauvages* et des *Indiens*. Par l'acte de nomination, il établit une nouvelle *mesure* du monde et de ses habitants. Chaque territoire qu'il visite est peuplé de gens qui chassent, qui pêchent, qui cultivent le manioc pour en faire la cassave, et qui dorment dans des carbets où ils accrochent leurs hamacs. En même temps qu'il recueille les mots amérindiens, il redessine une carte du monde où les îles sont affublées d'un nom chrétien, les hommes entrent dans de nouvelles catégories et il croit même entendre un rossignol lorsque le jour se lève. Ce baptême inaugural préfigure toute l'entreprise coloniale qui permettra à des découvreurs, des conquistadores, des prêtres, des savants de dire le réel

dans un nouvel idiome. La démarche coloniale fonde du neuf à partir de la science (certains disent la sagesse) des anciens. Assoiffés d'or et de richesses diverses, les conquistadores apportent avec eux la technique, la religion et la certitude qu'ils sont puissants face à des êtres inférieurs. Lorsque les Européens lèveront les yeux, ils prendront conscience qu'ils ont fondé un nouveau monde, ils inventeront un mot pour le dire, c'est le mot **criollo** qui donne **créole** et **créolisation** en français. Les premiers créoles sont blancs et ils font savoir au reste du monde qu'ils ont agrandi l'espace, qu'ils lui donnent une nouvelle mesure. Ce patron s'élargit aux animaux, aux objets, à la langue et même aux esclaves qu'ils font venir d'Afrique : un bœuf d'origine française qui est lâché dans les prairies haïtiennes est déclaré créole dès sa deuxième génération. D'un meuble, d'une maison, de la cuisine, des vêtements, de la manière de soigner, de la musique et des danses venues d'Europe et s'acclimatant au contexte, de tous ces composants culturels, on dit qu'ils créolisent.

Nous sommes dans un archipel qui était habité depuis au moins un millénaire et demi quand des navigateurs européens sont arrivés et sont venus s'installer avec un certain nombre d'intentions, la principale étant celle de faire du profit. Les premiers habitants déclarés Amérindiens des îles seront balayés, certains historiens parlent d'ethnocide. L'originalité du monde créole caribéen, c'est l'arrivée progressive de centaines, de milliers, de dizaines de milliers de travailleurs achetés à des marchands d'esclaves. L'arrivée de ces Africains dans nos territoires a façonné une civilisation nouvelle, on a inventé un droit particulier, la religion a été utilisée comme un adjuvant spécifique et autour de l'exploitation de la canne à sucre un mode de vie, une économie, une langue, une société créole sont apparues.

III. Mesures et démesures de la colonisation

Juste une fois lancé le processus de colonisation, les nouveaux maîtres du pays découvrent la démesure : contrairement au continent, les Antilles n'ont pas de minerai précieux, l'Europe découvre l'odeur d'une feuille qui se fume et le goût doux de la confiture. La soif de profit oriente l'économie vers l'exploitation de ces deux agro-industries. Et pour cultiver le tabac et la canne, il faut de la main d'œuvre. Les Amérindiens s'y refusent, les engagés blancs n'y suffiront pas, on élargit la Traite des Africains qui existait depuis un siècle ou deux à petite échelle. A quoi donc peuvent bien servir ces îles qui n'ont pas d'or ? Il faudra trouver le tabac, le sucre, et quelques produits agricoles dont l'Europe deviendra friande pour que la rentabilité des colonies soit établie. Cette rentabilité nous vaudra l'arrivée

pendant deux siècles d'une population servile qui n'amoindrira jamais le regard évaluatif.

En Guadeloupe et en Martinique, les poids et mesures que la Révolution française et sa fringale scientifique avaient animés pendant un siècle (Guedj, 2000) n'arrivent qu'au compte-gouttes. Les Bernini-Montbrand, Ludwig, Pouillet et Telchid (2014) et Confiant (2022), dictionnaires modernes, ont pourtant du mal à garder trace des unités de l'époque. Ne valent que la **livre**, le **pot** et sa moitié à vue d'œil pour les farines et les grains (**demi-pot**), la **muse**, la **chopine**, la **roquille**, le **quart**, le **quartaut**, le **muid** et la **barrique** pour les liquides, le **pouce**, le **pied**, la **toise** et la **lieue** pour les distances. La question de la monnaie (**livre**, **tournoi**, **gourde**, **sol**, **sou**) mérite un traitement à part. Deux considérations doivent être apportées tout de suite. La première, linguistique, est que ces mots sont français de source et que leurs traductions créoles restent proches de leur origine. La seconde, que ces notions sont approximatives ou au moins relatives en France ; on peut poser qu'elles le sont encore plus aux colonies. En réalité, aux Antilles, on évalue plus qu'on ne mesure. On échange, on brocante plus qu'on ne compte et qu'on n'achète.

Pour l'école, les maîtres blancs ont estimé que les esclaves et les affranchis peuvent s'en passer.

« ... Une expérience déplorable a prouvé que l'abus des lumières est souvent le principe des révolutions, et que l'ignorance est un bien nécessaire pour les hommes enchaînés par la violence ou flétris par les préjugés. Ce serait donc une imprudence bien dangereuse de tolérer plus longtemps dans les colonies des écoles pour les nègres et pour les gens de couleur. Qu'iraient-ils apprendre dans ces établissements ? Ils n'y puiseront que les connaissances supérieures qui font de l'homme éclairé le premier esclave de la loi et de leur intelligence, enorgueillie d'une instruction imparfaite et grossière, leur représentera sans cesse le régime colonial comme le code de la tyrannie et de l'oppression... »

« ... J'ai donc jugé nécessaire et je vous ordonne expressément..., de faire fermer les écoles publiques où sont admis les nègres et les gens de couleur. » (Le Capitaine Général de la Martinique Villaret-Joyeuse, le 10 novembre 1802).

Pour ce qui concerne notre propos, l'école publique ne date que de la fin du XIX^{ème} siècle. Durant toute la période esclavagiste, l'éducation ne se faisait que par les religieux et il ne s'agissait pas d'apprendre à lire et à écrire aux enfants des esclaves. Dès lors que l'Abolition de l'esclavage est prononcée, la question de l'appétit des Nègres pour le savoir est avérée. Si la plus grande partie continue à travailler aux champs dans les *ti-bandes*, quelques-uns s'investissent à l'école et réussissent. Cet appétit des nègres

pour l'école est perçu comme dangereux par certains observateurs et prometteur pour d'autres. Et là commence notre débat : l'école qui s'installe aux colonies, est une école à la française : il faut apprendre le savoir de la Métropole, lire écrire compter, l'histoire de notre glorieuse nation (« Nos ancêtres les Gaulois », Jeanne d'Arc, les quatre grands fleuves et les quatre grandes chaînes de montagne, les poids et mesures) même lorsque quelques chapitres auront traité à la grandeur des Outremers, il faudra en accepter le point de vue de Paris.

Il faudra attendre l'assimilationnisme, la curiosité pour « l'intelligence du Noir » et les proverbes créoles de Schoelcher (1842), l'efficacité grammaticale de Thomas (1869) et les exigences du socialisme local pour que les poids et mesures de la Métropole soient validés aux colonies. Or cette affaire, comme bien d'autres, résonne comme une victoire à la Pyrrhus. Au début, les enseignants sont contents de voir ouvrir des classes, les élèves adhèrent à l'école, des colonisés pensent qu'ils sortent de « la nuit de l'obscurité » et on leur montre, chiffres à l'appui, qu'ils ont raison d'être fidèles à cette institution. Au nom du progressisme, les plus doués des colonies occupent des fonctions prestigieuses, ils deviennent peu à peu avocats, médecins, hauts fonctionnaires, ils partent dans d'autres colonies en Afrique noire ou en Algérie. Mais progressivement une certaine hiérarchie de valeurs voit le jour. Jusqu'aux années 1980, l'école antillaise produit des individus voire des élites, mais de moins en moins de cohortes de bons, voire de moyens sortent du lot. Au discours anti-scolaire entendu au début de la scolarisation dans la bouche des dirigeants et des *békés*, et non écouté par la population générale, fait suite un *laïus* anti-scolaire de cette population qui pourrait avoir intérêt à se servir de l'école comme ascenseur social. Une autre forme d'opposition à l'école apparaît, fin du XX^{ème} siècle. L'échec, le retard, le rattrapage, l'illettrisme, l'abandon, la désertion sont les nouveaux termes de l'évaluation. De ce point de vue, la Départementalisation n'est pas une réussite. La belle image de l'école à la sortie de l'esclavage et plus encore de la colonie devenue part indiscutable de la France est flétrie. L'école produit l'échec scolaire.

IV. Demi-mesures : données chiffrées et statistiques

Puisqu'il est question d'évaluer dans ce colloque, puisqu'il faut citer des chiffres dans cette assemblée, j'ai consulté l'Insee et j'en citerai quelques-uns, en vrac pour montrer que ces sociétés sont spéciales. On pourrait

lancer un SOS crise. Au hasard, 78 % des enfants naissent hors mariage en Martinique, 40 % des gens vivent au-dessous du seuil de pauvreté monétaire. Si la moyenne nationale est de 11 cardiologues pour 100 000 habitants et de 20 psychiatres pour la même population, en Guadeloupe le nombre de cardiologues est de 7 et de psychiatres de 10. Je vous passe le montant des salaires et leur répartition globale, et je vous rappelle que vous avez devant vous les deux départements les moins peuplés et les plus vieillissants de France. Chaque année une quantité de jeunes Guadeloupéens et Martiniquais s'en vont en « Métropole » pour trouver un emploi et « la vie ». Les évaluations intermédiaires, le baccalauréat, et les réussites aux examens et concours font état de pourcentages très ressemblants aux moyennes nationales. Mais on parle rarement en ces occasions d'abandon scolaire et de changement de voie vers des formations pas ou peu diplômantes. Quand j'étais plus jeune, je m'occupais d'autres chiffres : les mariages, l'accroissement démographique, les décès, le taux de suicides... sur lesquels il y a tant à dire. Mais mon exposé, faute de place, n'abordera pas cette dimension.

Demi-mesure parce qu'on ne peut pas tout mesurer : l'intelligence, la compétence linguistique, le milieu social... Parlons d'acculturation ou de déculturation. Sautons l'épisode de « nos ancêtres les Gaulois », l'adhésion au récit national : écoutez le maître vous raconter les premières victoires de Jeanne d'Arc, ou Charles Martel écrasant les arabes à Poitiers, on s'identifie à nos héros, mais plus on grandit plus il y a malaise.

Ces notions de **retard** et d'**échec** entraînent le concept de **rattrapage** et la mesure de ce manque est exigée de manière de plus en plus précise pour que l'on puisse établir les marques d'un retour à niveau. Très à la mode dans le discours des politiciens, des administratifs et des commentateurs politiques l'exigence de la mesure a fait apparaître celle de l'excès, la démesure et celle des faux semblants. D'où la problématique de mon titre, mesurer certes, mais mesurer depuis un point de référence centré.

V. Contre-mesures

Qu'est ce qui est pertinent ? Le contexte ! Ils apprennent où et quand ? Quelle langue convient ? Ils sont créolophones, parlons leur donc créole ! Ils sont pour la plupart Antillais, Noirs ou Indiens de phénotype, cessons de leur parler comme à des Blancs. Essayons de tirer des poncifs coloniaux ou racistes dans la langue : qu'appelle-t-on **tête de nègre**, et est-ce que le mot est supportable pour un Antillais noir ? Les plus fins observateurs de

l'école des années 1960 lisent Fanon (1952), entendent une critique, élaborent des propositions et esquissent des réformes. Un livre de lectures locales, d'un non-Antillais Corzani (1969), *De sel et d'azur*, sonne la tentative d'antillaniser l'école. Des noms annoncent une prise de conscience : Nainsouta et l'Acra (l'Académie Créole Antillaise), Lauriette, surnommé Papa Yaya et l'AGEP (l'Association Générale de l'Enseignement), Bébel-Gisler (1989) et Bouadoubout (école où l'on n'utilise que le créole), Pouillet et Telchid, le Collège de Capesterre-Belle-Eau et le premier Dictionnaire de langue minorée de la Guadeloupe, pour finir Bernabé et le Gerec (Groupe de Recherches en Espace Créole travailleur de l'université enfin au diapason). On constate la diglossie langue basse *vs* langue haute, on entame une réflexion sur l'équilibrage des langues. Pour l'école suivent une série d'analyses sur le bilinguisme contrarié, sur la nécessité de prendre en compte la langue maternelle, la langue première, sur l'invalidité des instruments de mesure de l'école française aux Antilles.

Ceux de 2022 seront très à l'aise pour dire comment l'école n'est plus cette institution pleine de valeur. Si l'investissement « garderie » est plus fort, le viatique vers le succès n'est pas garanti. Des établissements, des professeurs, des matières gardent une image positive mais le simple examen du discours des étudiants qui viennent nantis d'une licence (leurs prédécesseurs avaient un certificat d'étude ou un brevet il y a cinquante ans) s'asseoir sur les bancs de cet immeuble qui s'appelle INSPE ne manifeste pas grand-chose des sentiments qui animaient leurs aînés ! Difficulté de mesurer la compétence en français d'élèves français de nationalité mais guadeloupéens et martiniquais de naissance, insérés dans un milieu créole qui n'a pas complètement intégré le français. Avant de pousser le principe que les deux langues sont distinctes et que les Antillais sont aussi des créolophones, qu'il suffirait de parler créole, on se doit de considérer que le type de bilinguisme local est plus que jamais l'opposé de la somme de deux monolinguisms. L'acquisition de la norme française est vérifiée dans une certaine proportion mais on ne sait pas laquelle. La permissivité de l'accent antillais, d'un certain nombre de formes phoniques, morphologiques, syntaxiques et la tolérance sémantique n'est pas établie. Avec des maîtres qui sont eux-mêmes en délicatesse avec la norme, demander qu'on juge de la qualité de la forme de la parole des élèves relève de la gageure.

Il faudrait au moins travailler à former les maîtres pour les rendre conscients de cette situation polylectale. Arrêter d'opposer l'oubli du créole à la préconisation du créole langue maternelle, arriver à l'enseignement du bilinguisme, et pas dans une co-officialité hypocrite

mais dans une pédagogie intelligente de la cohabitation des langues et des normes. Travailler à la découverte et à l'exploitation de l'*interlecte* plutôt qu'à la valorisation des unités linguistiques.

VI. En guise de conclusion, partir vers l'école du sur-mesure

Les communautés créoles offrent un terrain rêvé pour la saisie des langues et cultures en contact et constituent de ce fait un défi pour la pédagogie. Pourquoi certains maîtres et certains élèves accèdent à la filière, acquièrent le code haut, se concilient la norme dominante et d'autres n'entrent pas dans le système et, tant que la langue française n'est pas contrôlée, ils resteront des locuteurs approximatifs. Terrain linguistique mouvant, les langues défient autant le pédagogue qui veut enseigner le standard que celui qui se résout à enseigner le code minoré. Tout se passe comme si les Guadeloupéens et les Martiniquais se séparaient dans la pratique sociale sur une ligne frontière entre créolophones et francophones. Alors qu'en réalité, ils communiquent ensemble.

La pédagogie des langues doit tenir compte de la variation et la grammaire doit se faire contrastive. Plus que jamais, le créole et le français doivent devenir langues partenaires au sein du système éducatif si l'on veut mesurer des choses mesurables.

VII. Bibliographie

- Bébel-Gisler, D. (1989). *Le défi culturel guadeloupéen. Devenir ce que nous sommes*. Editions Caribéennes.
- Bernini-Montbrand, D., Ludwig, R., Pouillet, H. et Telchid, S. (2014). *Dictionnaire créole – français avec un lexique français – créole et un abrégé de grammaire*. Orphie.
- Calvet, L-J. (1974). *Linguistique et colonialisme*. Payot.
- Colomb, Ch. (1493/1979). *La découverte de l'Amérique*. La découverte.
- Confiant, R. (2022). *Dictionnaire créole martiniquais – français*. Orphie.
- Corzani, J. (1969). *De sel et d'azur. Textes d'explication antillo-guyanais*. Hachette.

Fanon, F. (1952). *Peaux noires masques blancs*. Seuil.

Guedj, D. (2000). *Le mètre du monde*. Seuil.

Schoelcher, V. (1842). *Des colonies françaises, abolition immédiate de l'esclavage*.
Edition Desormeaux.

Thomas, J. J. (1869/1969). *The theory and practice of Creole grammar*. New
Beacon Books.

Chapitre 2. Les normes de l'évaluation portées par les chefs d'établissement et les inspecteurs

Cas de la réforme du socle commun au collège en France

Yann Mercier-Brunel

I. Introduction

En France, l'évaluation scolaire fait l'objet d'un intérêt politique régulier, avec des réformes venant bousculer certaines pratiques des enseignants mais également leur demander parfois un changement réel de paradigme (De Ketele, 2016), suivant une démarche ministérielle essentiellement descendante. L'introduction du socle commun dans le système éducatif français ne fait pas exception : il s'agit d'une perspective de démocratisation des savoirs, avec l'idée qu'un ensemble minimal de connaissances doit être détenu par la totalité des élèves (rapport Thélot, 2004), tout en cherchant à garantir l'insertion des acteurs dans les contextes professionnels et sociaux en mutation (CNET, 2014).

Au niveau de l'école primaire (grades 1 à 5) et du collège français (grades 6 à 9), le socle commun de connaissances, de compétences et de culture (S4C) a contraint les enseignants¹ à évaluer un certain nombre de compétences de leurs élèves. Si la culture du primaire y était préparée depuis plusieurs années, au collège le changement a été plus important et

¹ Par souci de lisibilité, notamment dans l'explication des missions et rôles de chacun, le masculin générique sera employé. Pour autant, chacun des titres concerne dans la réalité des hommes et des femmes ainsi que des personnes non binaires, et doit être considéré comme les intégrant et les respectant tous également.

les cadres, supérieurs hiérarchiques des enseignants (chefs d'établissement et inspecteurs), ont été particulièrement sollicités pour accompagner cette évolution importante des pratiques évaluatives.

Notre étude porte sur la mise en œuvre du S4C au collège. Nous nous intéressons particulièrement à la façon dont les cadres se représentent l'évaluation scolaire et les demandes institutionnelles, et la façon dont ils ont accompagné la mise en œuvre de cette réforme. Cette dernière est censée amener les enseignants à modifier leurs pratiques évaluatives en plaçant au centre de leur attention le développement des compétences des élèves. Or, selon plusieurs travaux de recherche (par exemple Issaieva et al., 2011), on trouve une inadéquation entre la façon dont les enseignants pensent évaluer leurs élèves et la façon dont les élèves se sentent évalués, l'évaluation formative suivant les premiers étant perçue comme certificative par les seconds. Par ailleurs, pour beaucoup d'enseignants, le S4C est perçu comme une injonction à renoncer à une évaluation centrée sur une note chiffrée ou à articuler de façon plus ou moins cohérente des notes et niveaux de compétences (Mercier-Brunel, 2023). Les cadres de l'Education nationale constituent alors pour l'Institution le moyen de faire appliquer la réforme et d'assurer une certaine forme d'harmonisation des pratiques évaluatives sur le terrain.

Mais ces cadres partagent-ils eux-mêmes des normes professionnelles concernant les pratiques évaluatives et les évolutions à opérer ? En effet, dans l'Education nationale, la plupart des cadres sont d'anciens enseignants. Si ces normes sont unanimes, en quoi constituent-elles alors un levier pour accompagner les enseignants, et si elles sont diverses, en quoi diffèrent-elles et pourquoi ? La question est cruciale, car les normes professionnelles constituent à la fois un élément d'engagement dans sa communauté professionnelle (Fournier, 2015) et une caractérisation de la profession d'enseignant aux yeux des élèves et des familles (Wittorski, 2015) ; elles peuvent également être le support d'une créativité pédagogique des pratiques évaluatives (Mercier-Brunel et Crocé-Spinelli, 2021) ou devenir des « épreuves de professionnalité » (Jorro et Van Nieuwenhoven, 2019). Notre question de recherche porte ainsi sur les normes professionnelles des cadres de l'Education nationale, et la façon dont elles se positionnent entre les normes imposées par l'Institution et celles auxquelles adhèrent les enseignants.

Nos données reposent sur des entretiens conduits entre 2017 et 2019, auprès de 34 cadres : 23 chefs d'établissement et adjoints, 11 inspecteurs du second degré. A partir de leur vision de l'intérêt de la réforme du S4C et de leur place dans son accompagnement, il s'est agi de mettre au jour

leurs normes professionnelles en lien avec l'évaluation. Nous cherchons à identifier les normes qui fondent leur culture et leur posture relatives à l'évaluation scolaire (Jorro, 2016) et les points de convergence et de divergence des cadres et des enseignants (Mercier-Brunel, 2023).

II. Le cadre de la recherche

Contexte de la réforme

En 2000, lors du Conseil de Lisbonne, l'Union européenne pose les bases d'une éducation fondée sur des compétences que les élèves doivent développer. Se développent alors de nombreux standards et référentiels, basés sur la notion de compétences-clefs, et différentes procédures d'évaluation (CNESCO, 2014), au premier rang desquelles l'évaluation internationale PISA qui repose sur des tests auprès d'élèves de 15 ans.

C'est dans ce contexte que, au niveau de l'école et du collège français, le socle commun de connaissances et de compétences (S3C) introduit par le décret du 11 juillet 2006 impose une évaluation par compétences. A partir de 2013, le socle commun devient « de connaissances, de compétences et de culture » (S4C) et, mieux articulée aux programmes d'enseignement. L'évaluation de cinq « domaines » de compétences opérée par les enseignants doit apparaître à la fin de chaque trimestre dans le « Livret scolaire unique » (LSU) et être communiquée aux parents. Cette évaluation compte désormais pour la moitié des points pour obtenir le Diplôme national du brevet (DNB), le premier diplôme national que les élèves tentent d'obtenir en fin de scolarité obligatoire. Les cadres de proximité, supérieurs hiérarchiques des enseignants (chefs d'établissement) et les cadres académiques² (les Inspecteurs d'académie – Inspecteurs pédagogiques régionaux, les IA-IPR) ont été sollicités pour accompagner cette évolution importante des pratiques évaluatives. Or, le rôle des cadres de l'Education nationale dans l'accompagnement des réformes est un sujet rarement évoqué en recherche. Ils sont pourtant à la fois les porteurs des intentions de l'Institution, les garants de la mise en œuvre des réformes, les accompagnateurs des enseignants sur le plan organisationnel mais

² En France, une académie est une zone issue d'un découpage géographique du territoire. Une région contient d'une à trois académies. La France métropolitaine est découpée en 25 académies, et il existe 5 académies d'outremer.

également des contrôleurs de la mise en œuvre effective des réformes dans une perspective d'efficacité (Barrère, 2009).

Il semble alors nécessaire de s'interroger sur la façon dont les cadres appréhendent cette complexité, dans un contexte français qui se caractérise 1/ par une certaine culture de l'évaluation dans le second degré (Merle, 2018), fondée sur une logique de performance et de comparaison, 2/ une forte résistance aux réformes en lien avec l'évaluation et la sélection au lycée (Prost, 2013), 3/ une formation continue des enseignants peu développée (Mons et al., 2021) et 4/ un sentiment de déprofessionnalisation des enseignants lié à une conception, un pilotage et un contrôle « top-down » des réformes (Roquet et Wittorski, 2013).

Les Inspecteurs d'académie – Inspecteurs pédagogiques régionaux (IA-IPR)

L'éducation en France est très centralisée ce qui se traduit par des programmes nationaux devant s'appliquer dans tout le pays. Ainsi, au collège, le corps des IA-IPR rend compte au recteur du respect des injonctions nationales par les enseignants dans leurs classes (Condette, 2017). Ce sont ces IA-IPR qui exercent sur les enseignants une autorité dite « pédagogique », les évaluant formellement lors d'entretiens individuels et les accompagnant via des réunions d'équipes. Ces inspecteurs demeurent au service d'un pilotage académique des pratiques pédagogiques (Le Guern et Themines, 2017).

La position des IA-IPR vis-à-vis de l'évaluation est complexe à identifier. Ces derniers se caractérisent par un haut niveau de maîtrise disciplinaire (ils sont titulaires du concours de recrutement des enseignants du second degré le plus élevé), ont souvent été repérés pour la qualité de leurs pratiques pédagogiques lorsqu'ils étaient enseignants, et ont suivi une formation institutionnelle (bien que relativement sommaire) commune avec les chefs d'établissement. Buisson-Fenet (2015) met en lumière à ce sujet la formation longuement tâtonnante de ces inspecteurs, qui constituent un groupe « quelque peu confidentiel » (Albanet, 2016, p. 93) faisant l'objet de peu de travaux de recherche sur ses logiques et représentations. Les principales données sur les IA-IPR sont fournies par l'Institution elle-même, notamment grâce à des rapports de l'Inspection générale, sollicités par l'Institution et publiés uniquement en cas d'accord ministériel (ce qui peut conduire à en interroger l'impartialité ou l'exhaustivité).

Par ailleurs, les missions des IA-IPR font régulièrement l'objet d'évolutions, généralement dans le sens d'une augmentation de leurs tâches administratives au détriment du temps consacré à l'accompagnement des enseignants (Pons, 2016). Or, leur rôle d'inspection auprès des enseignants joue un rôle important dans la façon dont le métier d'enseignant s'exerce, particulièrement sur le plan de l'accompagnement des élèves (Heurdiere et Prost, 2021).

Les chefs d'établissement

Le chef d'établissement est « le représentant de l'État, placé sous l'autorité du recteur et de l'inspecteur d'académie, directeur des services départementaux de l'éducation nationale »³. Il est ainsi garant de la cohérence d'une politique nationale et académique dans laquelle s'inscrivent ses objectifs et son action éducative, portée par le projet de l'établissement. En plus de la gestion purement administrative de l'établissement et de la conduite des partenariats extérieurs, une de ses principales missions consiste à conduire une politique pédagogique et éducative au service de la réussite des élèves, en y associant entre autres l'ensemble des enseignants et des parents d'élèves. Il doit notamment participer à l'évaluation des enseignants.

Les enseignants du second degré constituent le principal vivier de recrutement des chefs d'établissement (environ 60% des lauréats, contre 20% de Conseillers principaux d'éducation et 20% de professeurs des écoles).

Si les chefs d'établissement s'impliquent plutôt dans la mise en œuvre des réformes, Buisson-Fenet et Pons (2019) identifient plusieurs types de relations que les chefs d'établissement entretiennent avec les enseignants à ce sujet, du conflit politisé lorsqu'ils s'opposent à la collaboration lorsque les demandes relayées par les chefs d'établissement modifient profondément les pratiques évaluatives des enseignants. On trouve finalement deux grandes catégories de postures professionnelles des chefs d'établissement : un type administratif et comptable (pour l'amélioration des résultats) et un type stratégique et pédagogique (réflexion avec les enseignants sur des évolutions pertinentes des pratiques pédagogiques).

³ Chaque académie est dirigée par un recteur sous l'autorité directe du ministre. Les départements qui composent une académie ont à leur tête un inspecteur d'académie, directeur des services départementaux de l'éducation nationale (IA-DASEN), qui dépend du recteur.

Certains chefs revendiquent ainsi leur rôle de « premiers pédagogues de l'établissement ».

Ainsi, les chefs d'établissement comme les IA-IPR doivent, dans leurs missions, accompagner les enseignants dans la mise en œuvre de la réforme du socle commun pour que « l'évaluation [permette] à chaque élève d'identifier ses acquis et ses difficultés afin de pouvoir progresser » (Ministère de l'Éducation nationale, de l'enseignement supérieur et de la recherche, 2016) : les IA-IPR sur le plan de l'évolution des pratiques pédagogiques et les chefs d'établissement au niveau de la concertation des équipes et du remplissage du LSU pour chaque élève.

La question des normes

Si la question des normes en sciences humaines et sociales n'est pas nouvelle, elle a toutefois été renouvelée dans le domaine de l'éducation (Prairat, 2012). La réflexion sur la norme suppose de s'extraire de la considération de l'individu pour embrasser un groupe social dont les membres orienteraient leurs actions selon des valeurs et des formes communes. Dans la sociologie moderne, des approches structuralistes (Durkheim, 1912) et pragmatiques (Weber, 1922) posent comme constitutives de la norme des règles comportementales socialement structurantes bien que coupées du motif rationnel de leur apparition (allant au-delà des simples représentations sociales pour intégrer les comportements générés). Ainsi, cette acception de la norme appliquée aux pratiques évaluatives des enseignants conduirait à identifier certains comportements de ces derniers, lorsqu'ils évaluent leurs élèves, hérités de la culture du métier sans forcément démontrer de liens de pertinence avec les contextes et les objectifs actuels, mais garantissant une forme de cohérence à l'école, tant pour eux que pour leurs élèves.

De nombreux sociologues de l'éducation (Dubet, 2016 ; Merle, 2018) fustigent par exemple, tout en soulignant leur force normative, les pratiques évaluatives des enseignants reposant sur la norme d'une évaluation juste parce qu'identique pour tous les élèves, qu'ils soient en réussite ou en difficulté, qu'ils maîtrisent les critères scolaires implicites (notamment en matière d'évaluation des performances) ou qu'ils ne les maîtrisent pas. Ces pratiques enseignantes illustrent ainsi l'idée de l'illusion de l'objectivité de la note, une note objective maintes fois battue en brèche par plusieurs auteurs depuis Noizet et Caverni en 1983 jusqu'à Merle en 2018.

D'autres éléments de la norme sociale sont à trouver chez Foucault (1975) qui distingue la construction sociale (par *normation* du réel) de son application pour contraindre le réel (normalisation des situations).

Cette approche de la norme, très rapidement synthétisée ici (on trouvera questionnée et discutée avantageusement la normalisation notamment dans Verhoeven (2018) et Mottier Lopez (2020)), conduit à formuler deux principes sur lesquels il convient de conserver une attitude critique : 1/ la norme n'est souvent pas conscientisée formellement par l'acteur comme un motif d'agir possiblement irrationnel du point de vue de l'objectif poursuivi (Waldenfels, 2005) ; 2/ la posture *normative* (discours prescriptif conduisant à appliquer et imposer une norme dans un processus de normalisation) prend généralement le pas sur une posture *normatrice* (discours normateur) consistant à créer des normes locales (proche de la « normation » selon Foucault, 1975). Or, la création de ces normes locales est potentiellement génératrice de progrès des élèves lorsqu'elles sont négociées avec eux (Mottier Lopez, 2020). Leur évolution au cours du processus évaluatif est même un élément important quand on observe les enseignants (Mottier Lopez et Dechamboux, 2017). La présente étude n'analyse toutefois pas directement le discours des enseignants (pour cela, voir Mercier-Brunel, 2023) mais celui des cadres chargés d'accompagner les enseignants dans la mise en œuvre d'une réforme de l'évaluation (celle du S4C) qui incite à questionner certaines normes de l'évaluation scolaire. Nous tentons de déterminer si les cadres interrogés relaient les mêmes normes que les enseignants, eux-mêmes dans une posture normative, ou s'ils tentent d'introduire d'autres normes de l'évaluation, plus en phase avec le S4C et la demande institutionnelle, dans une posture normatrice ?

III. Méthodologie

Mode de recueil

Le master Management des organisations scolaires (MADOS) qui s'adresse aux cadres éducatifs, est porté par un consortium de six universités françaises, et bénéficie d'une réputation de rigueur et de solidité reconnue par le Ministère de l'Éducation nationale, tant sur le plan des contenus que pour l'expertise de ses intervenants. Il se déroule en deux ans et s'adresse principalement à des chefs d'établissement (ou adjoints) et à des inspecteurs en poste, dans un cadre de formation continue.

L'étude présentée ici s'appuie sur un ensemble d'entretiens conduits durant les années de la mise en place de la réforme (avant la crise sanitaire

de 2020, qui l'a évidemment perturbée), soit des entretiens recueillis entre 2017 et 2019. Notre corpus est constitué de 23 chefs d'établissement et adjoints⁴ et de 11 inspecteurs (n= 34).

Population interrogée

La population interrogée est constituée de personnes accessibles pour les enquêteurs, d'où une forme d'opportunité dans le corpus, celui-ci n'ayant pas vocation à être représentatif mais n'ayant pas non plus été retenu pour des caractéristiques spécifiques (ni pro ni anti S4C). Les entretiens ont été conduits par des étudiants du master MADOS formés à la conduite d'entretiens semi-directifs et validés par un chercheur qui a ensuite mobilisé une analyse croisant une approche pragmatique du discours et une analyse linguistique énonciative (Kerbrat-Orecchioni, 2006 et 2008) suivant une méthode déjà utilisée par des recherches précédentes (voir par exemple Mercier-Brunel, 2019)⁵.

L'échantillon est ainsi composé :

- 11 IA-IPR (7 femmes et 4 hommes : 4 en Mathématiques, 3 en Lettres, 2 en Espagnol, 1 en Histoire-géographie et 1 en Physique). Nous soulignons l'absence d'IA-IPR en Education physique et sportive et en Arts (Arts plastiques et Education musicale), qui auraient pu apporter un autre regard, celui de disciplines historiquement plus fortement investies dans l'évaluation par compétences ;
- 23 chefs d'établissement en collège : 20 principales et principaux (12 hommes et 8 femmes) et 3 adjoints et adjointes (2 femmes et 1 homme).

⁴ Les chefs d'établissement adjoints ont passé le même concours de la fonction publique que les chefs d'établissement. Ils sont affectés comme adjoints pour différentes raisons : parce qu'ils ne souhaitent pas porter les responsabilités administratives d'un chef d'établissement, parce qu'ils préfèrent parfaire leurs compétences avant de prendre un poste de chef, parce qu'ils n'ont pas assez d'ancienneté pour prétendre au type d'établissement qu'ils visent, etc. La responsabilité de l'accompagnement des enseignants varie d'un établissement à l'autre, revenant parfois au chef, parfois à son adjoint, parfois étant répartie.

⁵ Nous remercions Nathalie Mons, responsable du master, ainsi que les étudiants du master MADOS des promotions 2017 à 2020.

Méthode d'analyse

Nous avons interrogé les IA-IPR et les chefs d'établissement à partir de questions ouvertes sur ce qu'avait modifié la réforme du S4C dans les pratiques évaluatives des enseignants, sur l'intérêt de cette dernière, et sur les obstacles et leviers qu'ils avaient identifiés pour faire évoluer ces pratiques. Les résultats des entretiens ont été synthétisés puis analysés suivant trois critères mentionnés avant les discussions.

IV. Résultats

Résultats des entretiens des IA-IPR

Sur les 11 IA-IPR interrogés, deux d'entre eux (en Histoire-géographie et en Lettres) nous ont indiqué que la question du S4C ne se posait pas, puisque que cette forme d'évaluation était déjà mise en œuvre dans leur académie (deux académies différentes) sans difficulté. Deux éléments viennent cependant nuancer leurs affirmations : d'une part certains chefs d'établissement des deux académies ont un discours très différent, d'autre part, au fur et à mesure des entretiens, les discours se nuancent, évoquant une disparité sur le territoire et l'existence de « professeurs réfractaires [qui] n'ont pas su voir les enjeux et la précision de l'évaluation, meilleure que la note ». On note donc un discours institutionnel, conforme au message politique de réussite de la réforme véhiculé par le ministère, qui se nuance voire se contredit au fur et à mesure des échanges (« de toute façon, il ne s'agit pas de favoriser car c'est la loi, on est obligé de le faire »). Cette observation met en lumière la difficulté d'interroger des cadres institutionnels dans le contexte d'une recherche, les questions pouvant générer un conflit de loyauté vis-à-vis de leur chaîne hiérarchique.

Un second constat est que trois des IA-IPR, au fur et à mesure de l'entretien, ont remis en question l'intérêt de la réforme : deux IA-IPR de Mathématiques sur le fond de leur discipline (« honnêtement, je ne pense pas que [les compétences] permettrai[en]t aux élèves de mieux appréhender les mathématiques. On est une discipline où on est encore avec énormément beaucoup [sic] de savoir et de savoir-faire à faire passer. Je pense que le sens il est ailleurs » ; « à vouloir trop développer de la compétence en mathématiques, on s'est écarté des savoirs fondamentaux »), et une IA-IPR de Lettres pour des questions de cohérence avec le reste du système (« il est difficile d'aller contre la culture de la note [qui] demeure forte partout, la poursuite des études, la transition

avec le lycée, les attendus des examens »). Leurs normes semblent très proches de celles mobilisées par un bon nombre d'enseignants, et on retrouve des discours similaires à ceux-ci dans la recherche menée (Mercier-Brunel, 2023).

En fait, seulement deux inspectrices (l'une en Lettres, l'autre en Espagnol) montrent qu'elles adhèrent à la réforme et qu'elles se sont appropriées son sens réel - une évaluation qui puisse faire passer les élèves d'un but de performance à un but de compétence – (Pintrich, 2000), explicitant son intérêt dans la construction des compétences par les élèves. Ces deux inspectrices intègrent la question du temps, jugeant prématuré la recherche d'amélioration des indicateurs dès la deuxième ou la troisième année de la réforme, reprochant à certains personnels d'opposer à la réforme l'absence de résultats probants rapides. Elles intègrent également la question de l'oral et d'une évaluation de la compétence autrement qu'au travers d'une performance réussie.

Les quatre autres IA-IPR semblent s'inscrire dans une analyse compréhensive vis-à-vis des enseignants dans leur mise en œuvre de la réforme, soulignant les obstacles (« les enseignants sont perdus dans les différentes formes d'évaluation », « c'est difficile de faire travailler les équipes et certains Perdir⁶ ne veulent pas mettre le feu ») et l'absence de capacité normatrice de l'Institution (« l'approche institutionnelle descendante empêche que les enseignants construisent une autre façon d'enseigner par compétence », « les formations sont inefficaces si les enseignants ne changent pas de posture », etc.). S'ils identifient différentes normes des enseignants qui constituent des obstacles au S4C et sur lesquels ils ont un point vue critique (« les enseignants entrent trop par les grilles et ne cherchent pas à comprendre la logique », « les professeurs se rattachent à une note qui finalement satisfait tout le monde »), il est difficile de savoir ce qui relève de leur loyauté vis-à-vis de l'Institution et de leurs propres normes vis-à-vis de la réforme du S4C.

Précisons toutefois que cinq IA-IPR soulignent la volatilité des priorités ministérielles, qui les a conduits à privilégier d'autres accompagnements que celui de l'évaluation du S4C et les a mobilisés sur ces autres chantiers au détriment de la continuité du S4C, largement non abouti. Ils ne contestent pas la réforme, mais la façon dont elle est pérennisée. Un discours normateur volatil face à des normes anciennes solidement ancrées dans les pratiques enseignantes (par exemple la note chiffrée)

⁶ Terme signifiant « personnel de direction », employé pour désigner essentiellement les chefs d'établissement et leurs adjoints.

expliquerait l'incapacité de faire évoluer les normes sous-tendant les pratiques évaluatives, et le discours réservé d'IA-IPR ne se sentant pas en mesure d'imposer des règles changeantes.

En guise de synthèse, nous retiendrons que tous les IA-IPR n'adhèrent pas de la même façon à la mise en œuvre de la réforme : pour certains elle a réellement du sens pour la réussite des élèves, certains l'accompagnent parce que c'est leur mission et tentent d'analyser et de dépasser les obstacles observés, d'autres enfin semblent peu convaincus (sans présumer de leur implication sur le terrain). En fait, il semble difficile de distinguer les normes institutionnelles que ces inspecteurs relaient dans leur discours (allant d'une communication positive à une réserve prudente) des normes auxquelles ils adhèrent. Plus gênant, les IA-IPR peuvent-ils imposer un discours normateur quand ils savent que celui-ci ne sera plus forcément défendu par l'Institution, au risque de perdre leur légitimité voire leur crédibilité ?

Résultats des entretiens des chefs d'établissement

Le premier constat est que les chefs d'établissement ont souvent des propos nuancés vis-à-vis des enseignants, et se réfèrent nettement plus que les IA-IPR à des situations précises. Ils identifient ainsi quasiment toujours des enseignants avancés sur l'évaluation du S4C et d'autres moins (« il y a une très grande hétérogénéité des professeurs, une grande différence entre des professeurs cantonnés dans une évaluation "certification" et d'autres plus experts »), avec un propos plus dynamique (« les choses changent, le regard porté sur l'élève évolue, [mais] il reste encore beaucoup à faire »).

Le second constat est que leur discours est moins réservé, et plus normateur en ce sens qu'ils cherchent à renouveler les normes des enseignants à partir des normes institutionnelles tout en étant plus normatif dans les faits (c'est-à-dire dans l'application de normes institutionnelles). On relève ainsi une réelle capacité à identifier les normes qui devraient évoluer (« le cours est resté magistral, "descendant", sans prise en compte des acquis des élèves, ni de leurs besoins »), tout en conservant une légitimité reposant sur l'application des normes imposées par l'Institution (donc dans une posture normative). Deux d'entre eux indiquent clairement, malgré leur parcours d'enseignant et leurs compétences, ne pas se sentir légitimes pour intervenir dans les pratiques pédagogiques des enseignants (« je n'ai pas l'impression d'avoir mon mot à dire sur les choix opérés » ; « le personnel de direction est compétent sur

les questions traitées, mais cela ne veut pas dire qu'il sait faire, il n'est plus prof »).

Ces évocations des normes des enseignants (qui permettent d'identifier leurs propres normes) s'observent chez 17 chefs qui analysent les postures des enseignants : « c'est compliqué d'évaluer par compétences, la certification ne demande pas une grande expertise » ; « la posture de contrôle de l'enseignant qui reste majoritaire et que les professeurs ont encore du mal à abandonner » ; « on doit faire face à des équipes attentistes et passives face à l'ampleur de la réforme ». Ils soulignent la difficulté des enseignants à articuler les programmes et le S4C, la « question du sens et de l'apport de l'évaluation par compétences que les enseignants ne perçoivent pas », notamment parce « qu'ils considèrent qu'ils n'ont pas besoin d'être formés puisque ça marchait bien avant ». Sans présumer que ce discours soit effectivement normateur (en ce sens qu'il soit effectivement relayé auprès des enseignants), les normes des chefs sont bien plus clairement exprimées que celles des IA-IPR.

On note à ce propos que 12 d'entre eux reprochent aux IA-IPR ce manque de réflexion dans le relais de normes claires auprès des équipes (« ce dont les enseignants auraient besoin, c'est de temps d'échange institutionnels sur l'évaluation par compétences et la pédagogie différenciée »). Neuf d'entre eux indiquent d'ailleurs s'être sentis seuls face à leurs équipes pour porter un discours normateur, certains ayant dû jouer le rôle d'arbitre entre les enseignants ou ayant dû rattraper certains effets des interventions du corps d'inspection : « la formation proposée [par l'IA-IPR] autour de l'évaluation par compétences m'a valu une levée de boucliers de la part de la salle des profs » ; « son intervention a été contre-productive ».

Enfin, les chefs d'établissement mobilisent assez souvent (11 sur 23) le contexte comme élément favorable ou défavorable : « dans un collège de centre-ville avec des gens qui sont là depuis plus de 20 ans, ils n'ont plus envie de travailler autant qu'un jeune professeur » ; « le fait d'avoir des équipes plutôt jeunes en éducation prioritaire facilite le changement de posture ». Ils relèvent parfois l'obstacle constitué par certains enseignants : « la deuxième année, quelques-uns se sont portés volontaires puis, ont reculé sous la pression de leurs collègues ». Ils soulignent également le contexte institutionnel peu favorable, et le revirement ministériel sur la réforme du collège qui, pour l'un d'eux, a « anéanti toute initiative ».

En guise de synthèse, on peut relever que la plupart des chefs d'établissement interrogés adhèrent à la réforme du S4C, en relayant des normes claires reposant sur la primauté de l'apprentissage de l'élève dans les pratiques évaluatives. Bien qu'ils soient majoritairement d'anciens

enseignants, ils portent un regard critique sur une évaluation-sanction reposant sur une note chiffrée auxquels beaucoup d'enseignants adhèrent (Merle, 2018) et souligne la difficulté des enseignants (voire le manque d'implication professionnelle ou l'opposition idéologique) à comprendre l'intérêt de cette réforme et à faire évoluer leur pratique dans l'intérêt de leurs élèves. Le discours est moins réservé que celui des IA-IPR, et les valeurs de soutien à la réussite des élèves, plus souvent exprimées.

Il est important de noter que les chefs d'établissement sont personnellement évalués en fonction de la réussite des élèves de leur établissement, ce qui n'est pas le cas des enseignants, et que cela est susceptible d'intervenir dans leur positionnement.

Notre question de recherche nous a alors conduit à analyser ces entretiens à partir de trois critères :

- les normes qui sous-tendent les pratiques évaluatives des enseignants mentionnées dans les entretiens ;
- le discours porté sur ces normes, d'adhésion ou de contestation ;
- la posture normative ou normatrice des cadres, c'est-à-dire le relais de normes existantes (des enseignants ou de l'Institution) ou la construction de nouvelles normes dans l'esprit de la réforme.

V. Synthèse et perspectives

Il semble d'emblée que les cadres du second degré, IA-IPR comme chefs d'établissement, montrent un recul certain sur les pratiques évaluatives scolaires en lien avec le S4C. Ainsi, moins impliqués dans le lien entre l'évaluation et la gestion quotidienne de la classe, leur position leur permet sans doute d'échapper plus facilement aux normes d'une évaluation emblématique, sanctionnant le (manque de) travail ou l'attitude (in-)appropriée des élèves en classe (Bressoux et Pansu, 2005) pour la mettre au service de l'apprentissage des élèves. De même, leur vision large d'un ensemble de classes, d'élèves et d'enseignants, leur permet plus facilement d'observer ce qui fonctionne et ce qui fonctionne moins dans l'enseignement, sans se sentir personnellement remis en cause puisque ce n'est pas eux qui mettent en œuvre cet enseignement (mais bien les enseignants).

Ainsi, ils portent également un discours distancié sur des normes sous-tendant régulièrement l'évaluation scolaire (Mercier-Brunel, 2023), comme le recours principal à la note chiffrée (rétroaction pseudo-objective donc peu contestable de mesure d'un acquis), la monopolisation du pouvoir évaluatif (c'est à l'enseignant seul qu'il appartiendrait de définir le niveau d'acquisition d'un élève), ou la primauté de l'épreuve de performance pour évaluer (on n'évalue bien qu'en soumettant l'ensemble des élèves à une même tâche en situation contrôlée). Notre étude montre qu'ici les normes de l'évaluation scolaire des cadres semblent alors différentes de celles des enseignants.

Nuançons toutefois notre propos. Nous trouvons dans le corps des IA-IPR, des inspecteurs qui adhèrent aux normes des enseignants (Mercier-Brunel, 2023) et non à celles de la réforme. Leur identité professionnelle étant davantage liée à une discipline scolaire que celle des chefs d'établissement, il semble qu'une forme d'identification avec les enseignants peut perdurer pour certains d'entre eux, et que la prise de recul ne soit pas toujours flagrante. À l'inverse, c'est au sein de ce corps professionnel que nous retrouvons le discours le plus élaboré, avec des IA-IPR ayant plus fortement et le plus finement identifié les enjeux en matière de but de compétences des élèves, avec des arguments didactiques précis (par exemple la place de l'oral pour évaluer précisément ce qu'un élève a compris).

Du côté des chefs d'établissement interrogés, nous avons constaté des normes sur l'évaluation scolaire bien différentes de celles des enseignants, avec l'intérêt et la progression des élèves comme principe sous-jacent des normes mobilisées. Toutefois, que ce soit parce qu'ils ne s'y autorisent pas ou parce qu'ils sont désormais éloignés du fonctionnement de la classe, leur proposition et leur analyse se limite à l'intérêt de travailler en équipe (sans aller au-delà d'un objectif d'harmonisation des pratiques entre enseignants), et vont jusqu'à réinterroger le modèle d'enseignement et la place faite aux besoins de l'élève (mais apparemment pas jusqu'à imposer un discours normateur auprès des enseignants pour faire évoluer leurs pratiques évaluatives).

Ainsi, les résultats de cette étude pourraient conduire à penser que les IA-IPR sont, au départ, mieux outillés pédagogiquement pour porter et imposer de nouvelles normes évaluatives qui vont dans le sens des préconisations de la recherche (et de l'esprit du S4C), en adoptant une posture normatrice, mais que *in fine* ce sont les chefs d'établissement, peut-être parce qu'ils sont davantage en proximité des élèves et de leurs

familles, qui portent ce discours face aux enseignants, sans toutefois avoir les moyens de l'imposer.

Ainsi, le discours normateur que l'on est en droit d'attendre de l'Institution face à une demande de changement de paradigme de l'évaluation scolaire (qui est bien ce qu'elle demande avec le S4C au regard de la formulation des textes réglementaires) se trouve davantage porté par des chefs d'établissement qui ne sont pas en mesure de l'imposer à l'intérieur de la classe (que ce soit par manque de légitimité ou par manque de compétence didactique), que par l'ensemble des IA-IPR (que ce soit par manque de maîtrise des enjeux, par manque d'adhésion ou par manque de temps).

Plusieurs questions se posent alors selon nous : est-il souhaitable que ce soient les chefs d'établissement qui portent et imposent ces nouvelles normes de l'évaluation scolaires ? Si oui, comment les outiller pour le faire ? Ne doit-on pas renforcer le rôle des IA-IPR dans le relais de ce discours, qu'ils pourront actualiser à la discipline dont ils sont spécialistes ? Mais comment les mobiliser réellement, à la fois en leur permettant de maîtriser les enjeux réels de la réforme du S4C et en leur donnant le temps de s'y consacrer ? N'y a-t-il pas une troisième voie à envisager, celle d'un modèle organisé de coopération entre IA-IPR et chef d'établissement (non soumis uniquement aux affinités personnelles), au-delà de l'évaluation réglementaire des enseignants lors des rendez-vous de carrière ?

VI. Bibliographie

- Albanel, X. (2016). La face cachée de l'inspection : retour sur le travail réel des inspecteurs pédagogiques régionaux. *Administration & Éducation*, 149, 93-97.
- Barrère, A. (2009). Les directions d'établissement scolaire à l'épreuve de l'évaluation locale. *Carrefours de l'éducation*, 28, 199-214.
- Bressoux, P. et Pansu, P. (2005). *Quand les enseignants jugent leurs élèves*. Presses universitaires de France.
- Buisson-Fenet, H. (2015). Une culture commune à défaut de grand corps ? *Recherche et formation*, 78, 19-34.
- Buisson-Fenet, H. et Pons, X. (2019). Responsable, mais pas redevable ? Gouvernance par les résultats et relations d'accountability dans

- les établissements scolaires en France. *Education et Sociétés*, 43, 41-56.
- CNESCO (2014). *L'évaluation des élèves par les enseignants dans la classe et les établissements : réglementation et pratiques*. CNESCO.
- Condette, J. (2017). *Les personnels d'inspection. Contrôler, évaluer, conseiller les enseignants retour sur une histoire France-Europe, XVIIe-XXe siècle*. Presses universitaires de Rennes.
- De Ketele J-M. (2016). L'évaluation et ses nouvelles tendances, sources de dilemmes. *Education Permanente* 208(3), 19-32.
- Dubet, F. (2016). Égalité des chances scolaires : le paradoxe français. *Après-demain*, 2(2), 14-16.
- Durkheim, E (1912/2013). *Les formes élémentaires de la vie religieuse*. Presses universitaires de France.
- Foucault, M. (1975). *Surveiller et punir*. Gallimard.
- Fournier, S. (2015). Les Pratiques managériales dans les EPLE et implication des enseignants. *Gestion et management public*, 2(2), 27-48.
- Heurdier, L. et Prost, A. (2021). *Les politiques de l'éducation en France*. La Documentation française.
- Issaieva, E., Pini, G. et Crahay, M. (2011). Positionnements des enseignants et des élèves du primaire face à l'évaluation : une convergence existe-t-elle ? *Revue française de pédagogie*, 176, 5-26.
- Jorro, A. et Van Nieuwenhoven, C. (2019). Le clair-obscur de l'activité évaluative. Dans A. Jorro et N. Droyer (dir.), *L'évaluation, levier pour l'enseignement et la formation* (p. 33-44). De Boeck.
- Jorro, A. (2016). Se former à l'activité évaluative. *Éducation Permanente*, 208, 53-64.
- Kerbrat-Orecchioni, C. (2006). *L'Énonciation*. Armand Colin.
- Kerbrat-Orecchioni, C. (2008). *Les actes de langage dans le discours. Théorie et fonctionnement*. Armand Colin.
- Le Guern A.-L. et Themines, J.-F. (2017). Monde en partage ou espace de domination : la spatialité des objets ordinaires dans les salles de classe. Dans Y. Bonny, N. Bautès et V. Gouëset (dir.), *L'espace en*

partage. Approche interdisciplinaire de la dimension spatiale des rapports sociaux (p. 253-271). Presses Universitaires de Rennes.

- Mercier-Brunel, Y. (2019). Les feedbacks des enseignants pour soutenir les processus d'autorégulation lors de corrections collectives. *Evaluer - Journal international de recherche en éducation et formation*, 5(2), 47-64.
- Mercier-Brunel, Y. (2023). Place des normes de l'évaluation des enseignants dans la mise en place du socle commun. *e-JIREF Journal international de recherche en éducation et formation*, 8(3), 29-47.
- Mercier-Brunel, Y. et Crocé-Spinelli, H. (2021). Créativité intégrée des pratiques évaluatives : entre routines et opportunités. *Le sujet dans la cité, HS 11*, 141-154.
- Merle, P. (2018). *Les pratiques d'évaluation scolaire. Historique, difficultés, perspectives*. Presses universitaires de France.
- Ministère de l'Education nationale, de l'enseignement supérieur et de la recherche (2016). Circulaire de rentrée 2016. *Bulletin officiel de l'Education nationale*, 15, 3-17.
- Mons N., Chesné J.-F. et Piedfer-Quênay L. (2021). *Comment améliorer les politiques de formation continue et de développement professionnel des personnels d'éducation ?* CNESCO-Cnam.
- Mottier Lopez, L. (2020). Processus de normalisation dans l'évaluation des apprentissages des élèves. Dans N. Younès, C. Gremion et E. Sylvestre (dir.), *Évaluations, sources de synergie* (p. 128-156). Presses de l'ADMEE.
- Mottier Lopez, L. et Dechamboux, L. (2017). D'un référentiel d'évaluation fixe à une co-constitution référentielle dynamique, ce que nous apprend le jugement situé de l'enseignant. *Contextes et didactiques*, 9, 12-29.
- Noizet G et Caverni J.-P. (1983). Les procédures d'évaluation ont-elles leur part de responsabilité dans l'échec scolaire ? *Revue française de pédagogie*, 62, 7-14.
- Pintrich, P. (2000). The role of goal orientation in self-regulated learning. Dans M. Boekaerts, P. Pintrich et M. Zeidner (dir.), *Handbook of self-regulation* (p. 451-502). Academic Press.

- Pons, X. (2016). Travail d'inspection et régulation du système scolaire français. *Administration & Éducation*, 149, 23-30.
- Prairat, E. (2012). Concepts et notions. *Les Sciences de l'éducation - Pour l'Ère nouvelle*, 1(1-2), 145-163.
- Prost, A. (2013). *Du changement dans l'École. Les Réformes de l'éducation de 1936 à nos jours*. Seuil.
- Roquet, P. et Wittorski R. (2013). La déprofessionnalisation : une idée neuve ? *Recherche et formation*, 72, 9-14.
- Thélot, C. (2004). *Pour la réussite de tous les élèves. Rapport de la commission du débat national sur l'avenir de l'école*. La documentation française.
- Verhoeven, M. (2018). Normes et professionnalité enseignante : entre performance et réflexivité. *Recherche et formation*, 88, 105-117.
- Waldenfels, B. (2005). Normalité et normativité : Entre phénoménologie et structuralisme. *Revue de métaphysique et de morale*, 1(1), 57-67.
- Weber, M. (1922). *Economie et société*. Agora.
- Wittorski, R. (2015). Questions posées à la professionnalisation aux métiers de l'humain. Dans J.-Y. Bodergat et P. Buznic-Bourgeacq (dir.), *Des professionnalités sous tension : Quelles (re)constructions dans les métiers de l'humain ?* (p. 29-41). De Boeck.

Chapitre 3. L'évaluation formative et l'apprentissage du langage : liens d'interdépendance entre la régulation et la micro-culture de la classe

Janet Cerda

I. Introduction

L'évaluation formative est une pratique d'enseignement et d'apprentissage itérative, utilisée par les enseignants et les élèves pour guider et améliorer les situations d'enseignement et d'apprentissage. C'est grâce à ce processus que l'enseignant et l'apprenant rassemblent des preuves (traces) pour démontrer comment les connaissances et les compétences des apprenants sont développées de manière indépendante, avec les pratiques de l'enseignant et en situation d'apprentissage avec ses pairs. Une caractéristique essentielle de l'évaluation formative réside dans les actions de l'enseignant et de l'apprenant à recueillir des preuves (traces) de l'apprentissage (Heritage, 2016). Ces actions comprennent la régulation de l'apprentissage, comme l'obtention et l'utilisation d'un retour d'information ou d'un suivi. En apprentissage du langage, cela peut consister par exemple à se baser sur des informations provenant du matériel et des outils d'apprentissages (tableaux de classe, cahiers des élèves) ou bien des éléments linguistiques (le langage à l'école) (Bailey et Heritage, 2008) déployés lors de la communication des connaissances relatives au domaine de l'enseignement du langage.

L'apprentissage (auto)régulé et l'apprentissage du langage sont des éléments constitutifs de l'évaluation formative dans les environnements

d'apprentissage et ils sont influencés par la micro-culture de la classe (Cerdea et al., 2020).

Une micro-culture de classe est composée d'une dimension sociale, qui comprend les normes sociales, les normes relatives au domaine de contenu enseigné et la pratique en classe ainsi que d'une dimension psychologique, qui comprend les croyances, les interprétations ou les connaissances spécifiques d'un domaine d'enseignement (par exemple, les mathématiques) ou une discipline (par exemple, l'éducation) (Cobb et Yackel, 1996 ; Mottier Lopez et Allal, 2007 ; Cerdeja et al., 2020). Pour ce qui concerne la dimension sociale de la micro-culture de la classe il est important de préciser que les normes sociales sont établies conjointement par les membres des communautés scolaires. Les normes relatives au contenu d'enseignement sont des actes de communication spécifiques à un domaine d'apprentissage ou à une discipline, et les pratiques de classe relatives au contenu sont des activités normatives ou des conceptualisations qui n'ont plus besoin d'être justifiées parce qu'elles sont considérées comme des vérités par l'enseignant et les élèves. Par exemple, dans les classes de mathématiques de l'école primaire, il y a un moment où les élèves apprennent que « 0 » est un nombre entier. Lorsque ce concept n'a plus besoin d'être expliqué ou justifié, il devient une pratique de la classe de mathématiques.

Les micro-cultures des classes linguistiquement diverses présentent un intérêt particulier pour la recherche exposée dans ce chapitre. Les élèves bi/multilingues apportent en classe des expériences et des antécédents bi/multilingues uniques qui sont façonnés par leur communauté et le paysage socioculturel et socio-politique plus large (Cerdeja, 2023). De ce point de vue, les classes linguistiquement diverses sont super-diverses (Romo et al., 2018). Compte tenu de cette super-diversité, certains éducateurs coordonnent des espaces de translangage dans leurs classes bilingues afin de tirer parti des répertoires linguistiques des élèves (c'est-à-dire les compétences et les connaissances d'une ou de plusieurs langues et variétés) lorsqu'ils produisent, interprètent et naviguent dans les pratiques discursives sociales et académiques (Poza, 2018).

De plus, les micro-cultures de classe qui disposent d'apprenants bi/multilingues, un groupe d'élèves souvent considéré comme défavorisé aux États-Unis (Olsen, 2010), des opportunités de s'engager dans des stratégies d'apprentissage régulées, soutiennent les élèves dans l'acquisition de capital académique/scolaire (Selvester et Summers, 2012). Les élèves bi/multilingues comprennent les élèves qui apprennent l'anglais comme nouvelle langue ou ceux dont c'est la langue maternelle

tout en apprenant une langue maternelle ou une langue reçue en héritage à l'école. De même, White et Bembenuddy (2014) ont avancé l'idée d'une pédagogie auto-réglée et culturellement proactive qui implique les enseignants et les élèves, ensemble, en fixant des objectifs et des stratégies et en s'engageant dans une auto-réflexion afin de promouvoir l'équité et l'inclusion. Cette approche correspond à la « pédagogie culturellement durable » qui renforce le besoin du développement continu des ressources culturelles et linguistiques des élèves.

Par conséquent, prenant en compte l'importance de la micro-culture dans la construction des composantes de l'évaluation formative, ce chapitre explore comment la relation d'interdépendance entre (a) l'apprentissage régulé, (b) l'apprentissage d'une langue et, (c) le contexte d'apprentissage soutient les pratiques d'évaluation formative.

II. Cadre théorique

La construction nécessaire des éléments composant le processus d'évaluation formative

L'évaluation formative fournit aux enseignants des opportunités de faciliter l'apprentissage régulé des apprenants bi/multilingues. L'apprentissage auto-réglé est un processus cognitif dynamique interne et est assisté par la modélisation et les réponses/retours externes (Schunk et Usher, 2013). C'est un processus par lequel les apprenants fixent des objectifs, exécutent, réfléchissent et adaptent l'apprentissage en vue de l'améliorer. En revanche, l'apprentissage co-réglé fait référence à l'influence commune de l'auto-régulation de l'élève et de la régulation d'autres sources (par exemple, les enseignants, les pairs) sur l'apprentissage de l'élève. L'apprentissage co-réglé est un processus dans lequel un apprenant (élève) développe ses compétences d'auto-régulation grâce à un accompagnement des enseignants ou des pairs au sein de la zone proximale de développement de Vygotsky (Hadwin et Oshige, 2011).

Souvent, la régulation des comportements des enseignants et des élèves constitue également des actions d'évaluation. Les actions d'évaluation formative de l'enseignant, telles que des questions approfondies, de la modélisation et des pauses afin de donner du temps aux élèves de digérer les nouvelles informations, sont toutes des exemples de co-régulation, car elles soutiennent les apprenants bi/multilingues dans l'acquisition d'aspects de stratégies d'auto-régulation et de langage dans la construction du contenu savant disciplinaire (Cerdeja et al., 2020). Les

actions d'évaluation formative des élèves qui sont également des comportements auto-régulés comprennent du feedback suscitant ouvertement des enseignants et des pairs et générant leur propre feedback en auto-surveillant leur apprentissage et en utilisant des techniques de questionnement (exemple : Quel est l'objectif de la tâche ? De quelles manières mon travail atteint l'objectif de la tâche ?).

C'est durant ces moments d'évaluation formative que le langage et les processus de régulation s'entremêlent (Cerde et al., 2020). D'une part, les compétences d'auto-régulation et la co-régulation mettent les apprenants bilingues/multilingues dans une position où ils apprennent de nouvelles compétences langagières plus efficacement (Bailey et Heritage, 2018). D'autre part, une plus grande compétence dans la langue permet aux apprenants bilingues/multilingues de participer dans les interactions formatives qui promeuvent les compétences d'auto-régulation et la co-régulation. De plus, c'est à travers les pratiques d'évaluation formative que les élèves apprennent à gérer les stratégies d'auto-régulation, en observant, en sollicitant les autres et en expérimentant avec les autres. En matière d'évaluation formative, les stratégies d'auto-régulation soutiennent la construction de sens des élèves, qui lorsqu'elles sont communiquées, fournissent aux enseignants la preuve de l'apprentissage de l'élève.

En outre, les apprenants font de la construction de sens du savoir linguistique et disciplinaire lorsqu'il y a des opportunités d'évaluation formative afin de négocier et réguler l'apprentissage (Cerde et al., 2020). En s'impliquant dans des pratiques d'évaluation formative, les apprenants bilingues/multilingues deviennent compétents dans la construction de distinctions entre les éléments linguistiques et leur but. Les éléments linguistiques incluent : (1) des termes et des expressions de tous les jours, (2) des usages spécialisés de langage tels que le discours mathématique, (3) des usages de langue du domaine général qui recoupent plusieurs domaines d'étude spécifiques ou de disciplines, et (4) un langage scolaire (Bailey et Heritage, 2008) qui utilise le langage pas directement en relation avec les disciplines mais toujours importants pour les interactions scolaires telles que respecter les consignes (Cerde et al., 2020). Les apprenants bilingues/multilingues deviennent également compétents dans la reconnaissance des outils matériels tels que des brouillons écrits dans différentes langues, des tableaux de classe et des cahiers qui soutiennent le partage d'idées verbalement en écrivant ou en dessinant.

Par ailleurs, au cours de l'activité communicative, les ressources langagières fonctionnent comme un assemblage de ressources sémiotiques

(exemple : visuelles, linguistiques, tactiles) et culturelles aussi bien que des facilités qui intercèdent avec les langues, les histoires, les cultures, les identités et les expériences des individus (Pennycook, 2017). Cerda et ses collègues (2020) ont trouvé que les assemblages au cours de l'activité communicative dans des environnements de classe, soutiennent les processus d'apprentissage de langue régulés, ce qui complète les pratiques d'évaluation formative.

Micro-culture de classe

Normes de classe sociales et le processus d'appropriation régulatrice commune

Le lien entre la micro-culture de classe et l'apprentissage régulé peut se trouver au sein des normes d'une salle de classe. Les normes sociales de classe reflètent des comportements collectifs typiques de classe qui sont établis conjointement par des membres de la communauté de classe. Les élèves, au sein de l'activité organisée par l'enseignant ou dans une activité de collaboration d'élèves, peuvent, ensemble, modifier les normes de classe, les pratiques et les conceptualisations grâce à des gestes tacites ou explicites, si la culture de classe favorise ses échanges/ va-et-vient (Cerda et al., 2020). C'est au cours de cette modification conjointe que les processus d'apprentissage régulés entrent en jeu car dans la co-régulation, aucun participant n'est un interactant dominant. Par conséquent, la gestion de la co-régulation par un enseignant, un élève ou la classe peut changer entre participants dans un processus de propriété régulatrice conjointe afin de déterminer le meilleur moyen d'avancer dans l'apprentissage (Heritage, 2016). Toutefois, la modification conjointe des normes de classe, des pratiques et de réflexion sont dépendantes de la culture de classe qui encourage les opportunités des élèves à échanger de cette façon (par exemple Forman et al., 2017 ; Cerda et al., 2020).

La micro-culture de classe est liée à l'activité de communication

À partir des perspectives socio-culturelles concernant l'apprentissage d'une langue, je maintiens que les aspects des dimensions sociales et psychologiques dans la micro-culture de classe sont liés à l'activité sociale où l'enseignant et les élèves participent dans des actes de communication — l'écriture, le faire, le dire et l'être — afin de s'exprimer sur du contenu disciplinaire en utilisant leurs compétences langagières (Cerda et al., 2020). De cette perspective, les élèves créent du sens de la langue et du contenu lorsqu'ils participent à des pratiques sociales ; et les pratiques sociales sont modelées et représentées par le langage social ou le

pragmatisme (Gee, 2015) ainsi que par la micro-culture de classe (Mottier Lopez et Allal, 2007).

Conception d'une auto-évaluation et son impact sur la micro-culture de classe et les pratiques d'évaluation formative

De manière générale, l'auto-évaluation est un type d'évaluation formative conçue pour suivre la compréhension des standards d'apprentissage de l'élève (Panadero et al., 2016). Cela induit des comportements d'apprentissage régulés : les élèves évaluent leur apprentissage et les enseignants utilisent les évaluations des élèves pour suivre leur compréhension. Par exemple, un enseignant peut demander à un élève d'écrire une marche à suivre et/ou stratégie qu'il utiliserait pour résoudre un problème en mathématiques, puis lui demander d'évaluer son travail en expliquant s'il changerait des choses à la procédure et/ou stratégie utilisée. Ensuite, l'enseignant suivrait la compréhension de l'élève en lisant les commentaires de ce dernier et enfin, donnerait son avis afin de le guider à l'étape suivante.

De plus, la manière dont l'auto-évaluation est mise en place en classe dépendra des dimensions sociales et/ou psychologiques de la micro-culture de classe (Cerdeja et al., 2020). Lorsque les enseignants conçoivent une auto-évaluation, ils sont influencés par leur croyances et interprétations d'un contenu spécifique d'étude (par exemple, les mathématiques) et des cadres ou des philosophies spécifiques de leur discipline (éducation) (Cerdeja et al., 2019). Ainsi, si les enseignants croient que l'enseignement devrait être collectif, la conception de leur auto-évaluation se reflètera sur ces perspectives. Par exemple, Cerdeja et ses collègues (2019) ont constaté que les enseignants qui travaillent dans une école primaire bilingue concevaient une auto-évaluation collective afin de : (1) jauger la compréhension des élèves de leur progrès en lecture bilingue/multilingue et (2) recueillir des informations sur l'identité de lecteur bilingue/multilingue des élèves, en lien avec leurs progrès en lecture et leur communauté (par exemple, s'engager dans la lecture en espagnol afin de maintenir la communication avec la famille, les gens de leur quartier et les personnes qui partagent leur appartenance ethnique). Cette auto-évaluation basée sur la communauté, en tant qu'évaluation formative, a fourni aux enseignants, aux élèves et aux parents un retour d'information sur les progrès scolaires et le développement de l'identité des élèves, ainsi que sur les perspectives des élèves concernant la communauté et l'utilisation de la langue.

Pour illustrer la manière dont les relations entre (a) l'apprentissage auto-régulé, (b) l'apprentissage de la langue et (c) le contexte d'apprentissage

soutiennent les pratiques d'évaluation formative, j'ai analysé les données issues d'une recherche collaborative dans une école primaire linguistiquement diversifiée.

III. Méthodologie

Contexte de la recherche : école concernée et modèle de programme bilingue

La recherche s'est réalisée au sein d'une école publique proposant des enseignements allant de niveau de la maternelle à la terminale. L'école est affiliée à une université et située dans un quartier urbain très pauvre du sud de la Californie. La plupart des élèves inscrits à l'école primaire ont une origine latino-américaine (80%), ils apprennent l'anglais comme nouvelle langue (75%) et bénéficient d'un repas gratuit ou à prix réduit (91%), ce qui est un indicateur de faibles revenus. Dans le cadre de ce chapitre, j'ai ciblé le programme bilingue espagnol de l'école primaire, dans lequel les enseignants utilisent l'espagnol et l'anglais pour enseigner le contenu et aider les élèves à développer des compétences en anglais tout en maintenant et continuant à développer des compétences dans la langue familiale, l'espagnol.

Le programme bilingue fonctionne de la manière suivante : 90% de l'enseignement est dispensé en espagnol et 10% en anglais dans les classes de maternelle première année ; chaque niveau augmentant ensuite le pourcentage d'enseignement en anglais jusqu'à atteindre 50% en cinquième année. Toutefois, les enseignants ne restreignent pas l'utilisation de la langue. Les enseignants et les élèves ont recours à des pratiques linguistiques fluides utilisant les deux langues de manière interchangeable et stratégique tout au long de la journée scolaire. En outre, les élèves sont regroupés dans des classes multi-âges - Den 1 (grades K-1), Den 2 (grades 2-3) et Den 3 (grades 4-5).

Activités collaboratives et partenariat

M. Alvaro (pseudonyme) est un enseignant et collaborateur dans cette recherche. La collaboration s'est mise en place autour d'un besoin des collègues de M. Alvaro de créer un outil d'évaluation formative afin de faciliter l'analyse des productions des élèves de lecture bilingue et soutenir le développement de l'apprentissage régulé des élèves.

À l'époque M. Alvaro occupait de nombreuses fonctions de direction. L'une de ses responsabilités consistait à organiser les traces (productions) de lecture des élèves recueillies par les enseignants chaque automne et chaque printemps. Les enseignants du programme bilingue utilisaient ces données pour guider l'enseignement et suivre les progrès des élèves en lecture bilingue, espagnol et anglais. En conséquence, M. Alvaro et moi-même avons été chargés de concevoir une auto-évaluation que nous avons baptisée Reader Identity Self-Assessment (RISA) afin de fournir aux enseignants et aux élèves des données formatives permettant d'orienter l'enseignement et l'apprentissage.

Pendant cinq mois, M. Alvaro et moi-même avons conçu et testé RISA. À cette époque, j'ai également observé ses pratiques en classe et je l'ai interrogé. L'année suivante, M. Alvaro et ses collègues de classe ont piloté RISA dans leurs classes pendant deux ans avant d'étendre l'outil aux classes de premier cycle de l'école primaire.

La classe de M. Alvaro

M. Alvaro est d'origine latino-américaine, fils d'immigrés mexicains. Il a grandi dans le sud de la Californie, dans le quartier où il enseigne actuellement. Au moment de l'étude, il enseignait dans une classe bilingue de quatrième/cinquième année et était enseignant depuis treize ans dont dix à l'école communautaire partenaire. La classe multi-âges de M. Alvaro, en quatrième et cinquième année était composée de 33 élèves (élèves de quatrième $n=20$; élèves de cinquième $n=13$) ; 28 de ces 33 élèves apprenaient l'anglais comme nouvelle langue (voir tableau 1). La plupart des élèves se sont déclarés bilingues (42%, $n=14$) ou multilingues (par exemple, parlant des langues supplémentaires telles que l'amérindien, le coréen) (36%, $n=12$). Lorsque les élèves ont été interrogés sur la langue qu'ils utilisent à l'école, ils ont déclaré utiliser davantage l'anglais et se sentir plus à l'aise dans cette langue. Les données provenant d'une enquête scolaire et des dossiers d'élèves ont été utilisées pour décrire la composition linguistique de la classe (voir la figure 1).

Données recueillies

Les données recueillies pour cet article s'inscrivent dans le cadre d'un effort plus large de recherche collaborative avec les enseignants de l'école et des chercheurs universitaires. Davantage d'informations sur le partenariat et la recherche collaborative sont présentés dans le texte de Cerda (2023). Les données ont été initialement collectées sur une période de cinq mois, de

janvier 2015 à juin 2015. J'ai effectué quatre observations de l'enseignement des mathématiques et de l'écriture, tout en enregistrant les leçons et en rédigeant des notes de terrain. Les enregistrements audio ont été transcrits et intégrés aux notes de terrain correspondantes. J'ai mené trois entretiens semi-structurés approfondis. Les entretiens ont porté sur le parcours de M. Alvaro, son expérience de l'enseignement et sa réflexion. Chaque entretien a duré entre 60 et 90 minutes. D'autres entretiens de suivi ont été menés par téléphone et par courrier électronique.

La RISA a été administrée à l'automne et au printemps à tous les élèves des classes Den 3 (y compris la classe de M. Alvaro) après avoir été testés dans la classe de M. Alvaro l'année précédente. À l'automne, les élèves ont représenté graphiquement leurs niveaux de lecture en anglais et en espagnol (données de la maternelle au CM2). Ils ont ensuite analysé leurs diagrammes à barres et tiré des conclusions sur les facteurs susceptibles d'avoir influé sur leur progression en lecture bilingue (voir figure 2). Les élèves ont fixé des objectifs de lecture, préparé un plan et décrit les pratiques de lecture nécessaires pour atteindre leurs objectifs. Au printemps, les élèves ont mis à jour leurs diagrammes longitudinaux en ajoutant leurs niveaux de lecture du printemps dans les deux langues. Puis, les élèves ont réfléchi à leurs progrès en décrivant les raisons pour lesquelles ils ont atteint (ou non) leurs objectifs de lecture et comment ils ont surmonté les obstacles.

Fiabilité

Tous les enseignants partenaires lisent les différentes itérations d'analyses pour s'assurer qu'on faisait correctement l'état des lieux des expériences d'enseignement et les pratiques de classe. Une pratique qui nous a permis de vérifier les biais (Merriam et Tisdell, 2015).

Analyse de la codification

Les codes déductifs de la taxonomie de Bailey et Heritage (2018) (voir tableau 2) et techniques de codification inductive ouverte (Merriam et Tisdell, 2015) étaient appliqués aux données. Les codes déductifs et inductifs faisaient tous deux partie du recueil de codes sous l'appellation « unité de sens » (la phrase ou le paragraphe) dans laquelle on les trouvait pour éviter de perdre des données contextuelles (voir Bazeley, 2014), voir figure 3 pour apercevoir un schéma de codification final.

IV. Résultats

La micro-culture de classe de M. Alvaro

Dans la classe de M. Alvaro, les élèves sont encouragés à évaluer les environnements d'apprentissage ou les activités à leur disposition et à utiliser les informations recueillies pour choisir les activités qu'ils estimaient les plus propices à l'apprentissage. Par exemple, au milieu d'une leçon, M. Alvaro dirige un petit groupe sur le tapis (en espagnol, en anglais ou dans les deux langues) pour aider les élèves qui ont besoin d'une pratique guidée par l'enseignant, tandis que d'autres travaillent en groupes, seuls ou en binômes, avec l'aide d'un enseignant. Une élève nommée Sylvia a quitté le groupe de son propre chef, s'est assise à sa place et a continué à travailler sur la fraction de feuille (écrite en espagnol) en utilisant des feutres de couleurs différentes. Tranquillement, les élèves du groupe ont continué à rejoindre et à quitter le groupe comme bon leur semblait (Cerdea et al., 2020, p.11).

Ce va-et-vient était une norme sociale construite à la fois par M. Alvaro et par ses élèves. Essentiellement, les élèves ont appris à tout d'abord s'auto-évaluer ou évaluer leur compréhension, à évaluer les options d'activités et à choisir les activités qui pourraient les aider à s'acquitter d'une tâche avec ou sans la présence de leur enseignant. Il est important de noter que l'acte d'évaluation ou d'auto-évaluation, qui est considéré comme une action d'apprentissage auto-régulé, engage les élèves dans des comportements d'évaluation formative. En outre, si les dimensions sociales de la micro-culture de la classe ont soutenu ce va-et-vient, les dimensions psychologiques, en particulier les convictions de M. Alvaro sur l'enseignement, en ont constitué la base. Il a expliqué : « bien que la classe ne soit qu'un « environnement social », elle a le potentiel « d'aider à transformer » la vie des élèves et, à mesure que « les élèves grandissent, ils emportent avec eux ces expériences d'apprentissage et les appliquent ». Il est donc important pour lui que la classe soit moins dirigée par l'enseignant afin que les élèves puissent participer à la composition, à l'évaluation et à la négociation des règles, normes et des structures qu'ils ont établies ensemble et planifier les prochaines étapes (Cerdea et al., 2020, p.10).

En outre, M. Alvaro a noté que les élèves initiaient et facilitaient souvent un discours évaluatif de classe. Il s'est rappelé : « l'an dernier, durant le cours de mathématiques, mes élèves et moi avons une conversation sur le nombre zéro et ce qu'il représentait. Certains de mes élèves m'ont dit : « je

ne suis pas vraiment sûr », « je pense qu'il est positif » ou « je pense qu'il est négatif ». Je ne plaisante pas ; ils ont littéralement divisé la classe en 3. Ils ont dit : « Tous les gens là-bas, à gauche, sont d'accord, les gens au milieu sont des gens qui ne savent pas trop où aller, et les gens à droite ne sont pas d'accord ». Ils ont commencé à avoir une conversation plus approfondie sur le zéro en tant que classe alors qu'ils se trouvaient dans ces groupes récemment constitués. À la suite de cette discussion, les élèves ont créé un protocole. Chaque fois qu'ils sentaient que la classe avait une conversation difficile ou animée, ils savent qu'ils devaient se séparer pour former ces groupes (Cerdea et al., 2020, p.11).

En tant qu'exemple d'appropriation conjointe de la régulation, cette activité d'évaluation en classe entière montre comment la gestion de la co-régulation est passée entre les mains de chaque participant lorsque la classe a déterminé une norme sociale qui soutiendrait davantage l'apprentissage de chacun. En outre, cette norme a créé un espace qui a suscité un dialogue en classe sur des sujets importants pour les élèves, ce qui a donné naissance à des pratiques de contenu de classe telles que « 0 » est le seul nombre entier qui n'est ni positif ni négatif.

Cependant, la détermination des élèves à créer une norme de classe qui soit à la fois une pratique d'évaluation formative et un échafaudage pour l'apprentissage régulé n'aurait peut-être pas été facile à atteindre sans le soutien des élèves et de M. Alvaro. Ainsi, si la dimension sociale de la micro-culture de la classe fait partie intégrante de la conception de cette activité d'évaluation en classe entière, la dimension psychologique a été tout aussi essentielle. Par exemple, M. Alvaro a encouragé les élèves à participer au discours en classe parce qu'il pensait qu'il s'agissait d'une pratique de justice sociale. Il a partagé son point de vue : « la justice sociale, c'est l'accès... la discussion... l'apport de points de vue différents, [et] la possibilité de parler les uns avec les autres » (Cerdea et al., 2020, p. 10).

Les actions d'évaluation formative de M. Alvaro : demandes, questions et temps d'attente

M. Alvaro a utilisé des caractéristiques linguistiques (exemple : demandes, questions) et la communication non-verbale (exemple : temps d'attente) comme éléments d'évaluation formative pour co-réguler la compréhension des élèves durant le cours en classe entière pendant que les élèves travaillaient sur les compétences linguistiques orales. Par exemple, il demandait souvent aux élèves de s'auto-contrôler : « J'aimerais que vous réfléchissiez à la façon dont ce que vous avez fait maintenant

vous a aidé à mieux comprendre le problème d'hier » (Cerde et al., 2020, p.14). Parfois, si les élèves travaillaient avec des partenaires, il leur demandait d'évaluer leur partenariat : « En quoi le fait de travailler en partenariat vous a-t-il aidé à travailler ? En quoi cela a-t-il aidé l'autre personne ? Qu'avez-vous appris ? Partagez maintenant avec votre partenaire » (Cerde et al., 2020, p.14). Ces échanges montrent comment M. Alvaro a utilisé les demandes et les questions, en tant qu'évaluation formative pour promouvoir la régulation des élèves (contrôle, évaluation, motivation), améliorer les compétences linguistiques orales en anglais et développer des conceptualisations disciplinaires.

En outre, M. Alvaro a utilisé le « temps d'attente » (ou la pause) pendant la discussion en classe entière comme stratégie de régulation et comme pratique d'évaluation formative pour soutenir l'utilisation de la langue orale par les élèves, encourager leur participation et leur fournir des informations sur les compétences linguistiques des élèves et leurs connaissances du domaine. Il a expliqué que le fait de prolonger le silence permettait aux élèves de « rassembler leurs pensées [afin qu'ils puissent] exprimer ce qu'ils voulaient partager » (Cerde et al., 2020, p. 13).

Par exemple, lors de l'observation d'une activité de discussion dynamique (sur la base de méthode aquarium (fishbowl method), qui est une tâche communicative impliquant un petit groupe d'élèves modélisant un processus ou une compétence pour le reste de la classe, M. Alvaro a commencé une discussion en classe en demandant (Cerde et al., 2020, p. 13) : « Qu'est-ce que vous vouliez changer ? Y a-t-il des choses que vous auriez ajoutées ? » Au début, les élèves étaient silencieux, mais au bout de quelques secondes, voire d'une minute, ils se sont mis à parler. Dans ce cas, M. Alvaro a utilisé le « temps d'attente » (1) comme une stratégie de co-régulation pour aider les élèves à contrôler les stratégies de communication efficaces, (2) comme une évaluation formative pour transformer la discussion en classe sur les stratégies d'écriture en opportunités pour les élèves de créer et de sélectionner des objectifs d'écriture personnalisés, et (3) comme un processus de régulation linguistique culturellement proactif pour encourager la participation des élèves qui développaient un langage disciplinaire à la fois dans leur langue maternelle et dans leur nouvelle langue.

Des assemblages sémiotiques qui améliorent le retour d'information sur le partenariat et permettent la pratique d'évaluation formative

M. Alvaro est convaincu qu'offrir un « accès » aux élèves signifie leur apprendre à pratiquer le « choix ». Il explique : « En écriture, les élèves

choisissent des sujets, en mathématiques, le choix porte sur les séries de nombres et les stratégies que les élèves veulent utiliser » (Cerde et al., 2020, p. 10). Par conséquent, lorsque les élèves ont travaillé en partenariat pour accomplir des tâches en classe, ils ont puisé dans une myriade de ressources sémiotiques qu'ils avaient apportées, recherchées et/ou rassemblées avec leur partenaire, tout en recevant les encouragements de M. Alvaro à cet effet. Les élèves ont exploité ces ressources pour donner un sens au contenu disciplinaire et, ce faisant, se sont mutuellement fournis un retour d'information et les étapes suivantes, ce qui constitue des exemples de pratiques d'évaluation formative. A titre illustratif, quelques autres exemples de pratiques d'évaluation formative : pendant l'atelier d'écriture, M. Alvaro a demandé (en anglais) aux élèves de trouver un partenaire pour les aider à réviser leurs essais persuasifs en espagnol ou en anglais ; Daniel (élève de CM1) a demandé à Dario (élève de CM2) s'il pouvait lui donner son avis sur son essai persuasif en espagnol ; après que Daniel a lu à haute voix son essai en espagnol, Dario (en espagnol) l'a aidé à identifier l'opinion, les raisons et les détails ; Dario a remarqué que les détails de Daniel ne correspondaient pas à ses raisons, et lui a donc posé des questions en espagnol pour essayer de comprendre ce qu'il essayait de faire passer comme informations. Dario lui a expliqué en espagnol : « Daniel, imagine que ton dessin est la cancha et que les pelotes qui pataugent vers la cancha sont les détails. L'opinion est le jeu. [Daniel, imagine que ton paragraphe est le but et que les ballons de football que tu envoies dans le but sont les détails. L'opinion est le jeu] ». Ensemble (en espagnol), ils ont revu chaque paragraphe et corrigé les détails qui ne correspondaient pas (Cerde et al., 2020, p. 12).

Dans cet exemple, le pair le plus expert, Dario, (1) a écouté activement Daniel lire son essai en espagnol, (2) a évalué le travail de Daniel, (3) a trouvé une erreur, puis (4) a choisi d'utiliser une métaphore comme stratégie d'apprentissage pour enseigner et améliorer la compréhension de Daniel. Plus précisément, Dario a partagé verbalement une référence culturelle en espagnol (c'est-à-dire les règles du football et le vocabulaire de cette situation) pour transférer la connaissance de cette référence à la situation comparable de l'écriture académique. Ainsi, Dario a combiné diverses ressources sémiotiques (auditives, textuelles, culturelles, environnementales) et a co-régulé le langage pour soutenir la compréhension du contenu (savoir comment structurer un essai persuasif). Parallèlement, Daniel a suivi ses propres progrès en travaillant avec Dario pour réviser son travail. Cet exemple montre comment la manifestation des croyances de M. Alvaro dans la dimension psychologique de la micro-culture de la classe, et les normes de partenariat

reflétant la dimension sociale, ont permis aux élèves d'utiliser diverses ressources sémiotiques pour contrôler eux-mêmes l'achèvement d'une tâche en espagnol ou en anglais, ainsi que pour demander de l'aide et un retour d'information à un camarade. Les stratégies d'apprentissage régulé utilisées par Dario et Daniel - contrôle, évaluation, fourniture/réception de commentaires - sont autant d'exemples de pratiques d'évaluation formative.

Comment M. Alvaro et ses collègues enseignants ont utilisé une auto-évaluation en classe pour améliorer les comportements d'apprentissage régulé et linguistique des élèves ?

Après que les enseignants de Den 3 (quatrième et cinquième années) ont testé la grille d'auto-évaluation (Reader Identity Self-Assessment (RISA) dans leurs classes, nous avons analysé les réponses des élèves afin d'évaluer la précision des questions d'évaluation pour susciter la réflexion des élèves. Les enseignants ont constaté que les élèves avaient du mal à rédiger des réflexions détaillées. La plupart des réflexions des élèves portaient sur leur niveau de lecture. Par exemple, les élèves de cinquième année écrivaient qu'ils avaient progressé de cinq niveaux de lecture depuis la troisième année. Par la suite, les enseignants ont évalué leur programme de lecture et leurs pratiques pédagogiques. Ils ont conclu que la co-régulation (co-modélisation) de la langue et de la connaissance du contenu étaient nécessaires pour soutenir la réflexion des élèves. Ils ont donc commencé à enseigner explicitement la lecture, les pratiques de réflexion et la fixation d'objectifs dans le cadre de leur programme d'enseignement de la lecture. Les élèves ont bénéficié d'une aide à l'utilisation des carnets de lecture et ont appris à utiliser les données de ces carnets pour créer des plans de lecture concrets et mesurables (voir figure 4).

Les enseignants ont également apporté des modifications à la RISA. Ils ont remplacé les questions qui invitaient les élèves à comparer leurs progrès en lecture dans les deux langues, telles que « En quoi votre progression en lecture en espagnol est-elle différente de votre progression en lecture en anglais ? » et « Qu'avez-vous appris sur vous-même en tant que lecteur d'espagnol et d'anglais ? ». Les enseignants ont estimé que le fait d'opposer leurs progrès en lecture en espagnol et en anglais pouvait nuire à l'identité bi/multilingue des élèves. Ils pensent que ce type de questions pourrait pousser les élèves à considérer leurs progrès dans l'une ou les deux langues de manière négative, ce qui aurait un impact sur leur estime de soi et leur efficacité personnelle. Les enseignants craignaient que les élèves ne

se découragent. Les questions initiales ont donc été remplacées par les questions et les messages-guides suivants :

1. « Mes données de lecture me montrent... »
2. « Pour moi, cela signifie... »
3. « Êtes-vous un lecteur bilingue ou multilingue ? »
4. « Pourquoi lisez-vous dans deux ou plusieurs langues ? »
5. « Je veux être le type de lecteur bilingue ou multilingue qui... »
6. « Comment vous sentez-vous par rapport à la lecture en deux ou plusieurs langues ? »

Les enseignants étaient satisfaits des résultats. Ils ont estimé que les nouvelles questions de réflexion encadraient l'analyse par les élèves des graphiques à barres (progression de la lecture bilingue) et les poussaient à considérer leur lecture bilingue / multilingue comme un atout (voir figure 5).

Les parents ont également apprécié d'entendre leurs enfants parler de leur progression en lecture bilingue et de leurs objectifs lors des réunions parents-élèves-enseignants. Un enseignant a fait part de son expérience : « Les parents ont vraiment apprécié de voir la progression de leurs [enfants] [...] Ils ont pensé que c'était tellement une chose à faire à l'âge adulte et ont dit : « Je n'ai pas appris à faire des graphiques avant d'être plus âgé ». D'autres sont allés plus loin et ont dit : « Personne ne m'a jamais demandé de me considérer comme un lecteur ».

V. Conclusion

Dans l'ensemble, les résultats de cette recherche collaborative ont révélé l'importance de la (co)-régulation dans l'apprentissage de la langue pour la mise en œuvre des pratiques d'évaluation formative par les élèves et les enseignants. Dans certains cas, la régulation soutenait spécifiquement l'apprentissage de la langue. Dans d'autres cas, la langue a été utilisée pour soutenir les processus de régulation qui améliorent l'apprentissage du contenu disciplinaire et de la langue. Quelle que soit leur orientation, les actions de régulation et d'apprentissage de la langue suscitent ensemble des pratiques d'évaluation formative interactives qui guident et améliorent l'enseignement et l'apprentissage en montrant comment les

connaissances et les compétences de l'apprenant sont développées de manière indépendante, avec l'enseignant et avec ses pairs.

Plus important encore, les pratiques d'évaluation formative décrites dans ce chapitre étaient des éléments constitutifs de la micro-culture de classe. En d'autres termes, les comportements d'apprentissage régulé et linguistique essentiels aux pratiques d'évaluation formative décrites étaient des normes sociales modélisées par l'enseignant et les élèves. En outre, l'enseignant avait des croyances et des conceptualisations sur l'enseignement et l'apprentissage qui étaient propices aux comportements d'apprentissage régulé et linguistique, et donc également aux pratiques d'évaluation formative. En outre, les outils d'auto-évaluation communautaires et les pratiques d'évaluation formative favorisent l'apprentissage des élèves. Toutefois, cette conclusion est transférable à condition que des élèves soient formés et encouragés à adopter des comportements d'apprentissage régulés et linguistiques.

VI. Bibliographie

- Bailey, A. L. et Heritage, H. M. (2008). *Formative assessment for literacy, grades K-6: Building reading and academic language skills across the curriculum*. Corwin Press.
- Bailey, A. L. et Heritage, M. (2018). *Self-regulation in learning: The role of language and formative assessment*. Harvard Education Press.
- Bazeley, P. (2014). *Qualitative data analysis: Practical strategies*. Sage Publications.
- Buckingham Shum, S. et Deakin Crick, R. (2012). Learning Dispositions and Transferable Competencies: Pedagogy, Modelling and Learning Analytics. Dans *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge* (New York: Association for Computing Machinery [ACM], 92-101.
- Canagarajah, S. (2017). Translingual practice as spatial repertoires: Expanding the paradigm beyond structuralist orientations. *Applied Linguistics*, 39(1), 31-54.
- Cerda, J., Bailey, A. L. et Heritage, M. (2020). Contexts for self-and co-regulated learning in a dual-language elementary school classroom. *Language and Education*, 34(5), 407-424.

- Cerda, J., Garcia, N., Jimenez, R. et Kim, Q. (2019). *The power of self-assessment: Using community-based measures to advance biliteracy*. UCLA Center for Community Schooling Research Report 1-RR19. <https://communityschooling.gseis.ucla.edu/the-power-of-self-assessment/>
- Cerda, J. (2023). *Examining the impact of sociopolitical and school contexts on teachers' and students' multilingual and psychosocial development*. University of California Los Angeles.
- Cerda, J., Darriet, C., Quartz, K. H., Kim, Q., McNaughton, I et Jiménez, R. (under review). Building on community assets: A school's journey developing a TK-12 community language approach.
- Cobb, P. et Yackel, E. (1996). Constructivist, emergent, and sociocultural perspectives in the context of developmental research. *Educational Psychologist*, 31(3-4), 175-190.
- Forman, E. A., Ramirez-Del Toro, V., Brown, L et Passmore, C. (2017). Discursive strategies that foster an epistemic community for argument in a biology classroom. *Learning and Instruction*, 48, 32-39.
- Gee, J. P. (2015). *Social linguistics and literacies: Ideology in discourse*. Routledge.
- Hadwin A. F. et Oshige, M. (2011). Self-regulation, co-regulation, and socially shared regulation: exploring perspectives of social in self-regulated learning theory. *Teachers College Record*, 113(2), 240-264.
- Heritage, M. (2016). Assessment for learning: Co-regulation in and as student-teacher interaction. Dans D. Laveault et L. Allal (dir.), *Assessment for learning: Meeting the challenge of implementation, The enabling power of assessment* (p. 327-343). Springer International.
- Merriam, S. B. et Tisdell, E. J. (2015). *Qualitative research: A guide to design and implementation*. John Wiley & Sons.
- Mottier Lopez, L et Allal, L. (2007). Sociomathematical norms and the regulation of problem solving in classroom microcultures. *International Journal of Educational Research*, 46, 252-265.
- Olsen, L. (2010). *Reparable harm: Fulfilling the unkept promise of educational opportunity for California's long-term English learners*. Californians Together.

- Panadero, E., Brown, G. T. et Strijbos, J. W. (2016). The future of student self-assessment: A review of known unknowns and potential directions. *Educational Psychology Review*, 28(4), 803-830.
- Paris, D. (2012). Culturally sustaining pedagogy: A needed change in stance, terminology, and practice. *Educational Researcher*, 41(3), 93-97.
- Pennycook, A. (2017). Translanguaging and semiotic assemblages. *International Journal of Multilingualism*, 14(3), 269-282.
- Poza, L. E. (2018) The language of sciencia: translanguaging and learning in a bilingual science classroom. *International Journal of Bilingual Education and Bilingualism*, 21(1), 1-19.
- Romo, H. D., Thomas, K. J. A. et García Eugene E. (2018). Changing demographics of dual language learners and English learners: Implications for school success. *Social Policy Report*, 31(2), 1-35. <https://doi.org/10.1002/j.2379-3988.2018.tb00028.x>
- Selvester, P. M. et Summers, D. G. (2012). *Socially responsible literacy: Teaching adolescents for purpose and power*. Teachers College.
- Schunk, D. H. et Usher, E. L. (2013). Barry J. Zimmerman's theory of self-regulated learning. Dans H. Bembenutty, T. J. Cleary et A. Kitsantas (dir.), *Applications of self-regulated learning across diverse disciplines: A tribute to Barry J. Zimmerman*, (p. 1-28). Information Age.
- White, M. C. et Bembenutty, H. (2014). Teachers as culturally proactive agents through cycles of self-regulation. Dans S. J. Farenga (Chair), *Implications of diversity toward the preparation of teachers for urban schools and communities*. Symposium conducted at the biannual Department of Secondary Education and Youth Services Research Symposium, Queens, NY.
- Zumbrum, S., Tadlock, J. et Roberts, E.D. (2011). *Encouraging self-regulation in the classroom: A review of the literature*. Metropolitan Educational Research Consortium (MERC), Commonwealth University.

Tableau 1 : Description démographique des élèves (N = 33)

	<i>n</i>	%
Niveau scolaire		
Élèves de 4e	20	60.61%
Élèves de 5e	13	39.39%
Ethnie		
Latino-américains	32	96.97%
Caucasiens	1	3.03%
Genre		
Féminin	17	51.52%
Masculin	16	48.48%
Elèves à haut potentiel	6	18.18%
Education inclusive	5	15.15%
Classification de développement de la langue anglaise		
Anglais comme nouvelle langue	12	36.36%
Reclassé à bonne maîtrise de l'anglais	16	48.49%
Bonne maîtrise de l'anglais	5	15.15%

Note. Classification du développement de la langue anglaise établie par l'école à partir des résultats du California English Language Development Test (test de développement de la langue anglaise de Californie).

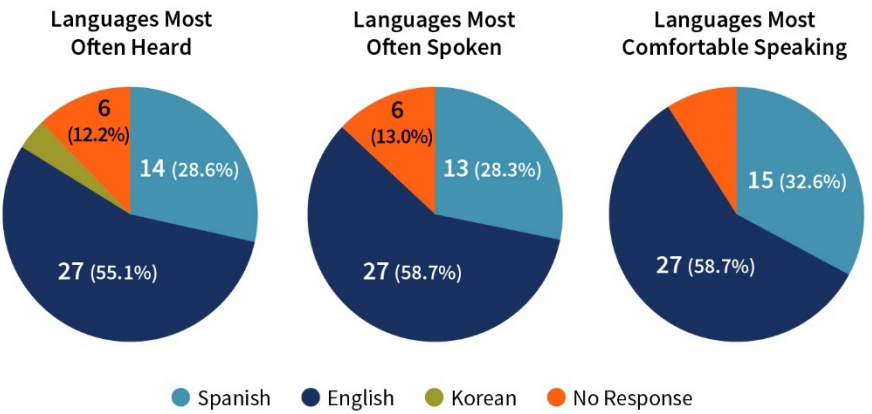
Tableau 2 : Stratégies et comportements de régulation et taxonomie de dispositions

Stratégies	Comportement et disposition
Mise en place des objectifs	Baser les objectifs à court-terme sur les preuves et retours ; mettre en place et prioriser les objectifs et sous-objectifs individuels et de groupe
Planification	Établir des stratégies dans le but d'atteindre des objectifs individuels ou de groupe (2 élèves ou plus) ; déterminer combien de temps et quelles ressources sont nécessaires pour atteindre ces buts.
Motivation	Sans l'utilisation d'une ou plusieurs stratégies pour garder leur propre apprentissage ou celui de leurs pairs sur la bonne voie afin d'atteindre leurs objectifs.
Contrôle de l'attention	Assumer des tâches et supprimer les distractions de l'environnement ; choisir les conditions qui rendent l'apprentissage seul ou en groupe plus facile.
Utilisation flexible de stratégies	Mettre en place de multiples stratégies d'apprentissage tout au long des tâches et ajuster ces stratégies au moment opportun afin de consolider les progrès seul ou en groupe ; utiliser des preuves pour s'adapter et inventer des stratégies d'apprentissage seul ou en groupe.
Contrôle	Contrôler leur progrès vers les objectifs d'apprentissage ou ceux de leurs pairs et faire des ajustements pour consolider le progrès.
Recherche d'aide et retours rétroactifs	Demander de l'aide et du soutien aux adultes et aux pairs ; demander des informations (exemple : bibliothèques, internet, contacts) ; solliciter du retour sur le progrès de la part des professeurs et des pairs.

Évaluation	Évaluer leur propre apprentissage et celui de leurs pairs indépendamment des enseignants, et ajuster pour de futures tâches similaires
------------	--

Note. Tiré de Cerda et al. (2020). La taxonomie a été adaptée à l'origine par Bailey et Heritage, (2018), à partir d'informations trouvées dans Buckingham Shum et Deakin Crick, (2012) et Zumbum et al., (2011).

Figure 1 : Auto-évaluation de l'utilisation des langues par les élèves et aisance dans l'utilisation des langues à l'école (N = 33)



Note. Données auto-déclarées par les élèves dans le cadre de l'enquête sur l'expérience linguistique ; les élèves ont répondu aux questions suivantes : « Écrivez la ou les langues que vous entendez le plus à l'école », « Écrivez la ou les langues que vous parlez le plus à l'école » et « Écrivez la ou les langues dans laquelle vous vous sentez le plus à l'aise de parler à l'école ».

Figure 2 : Élève montrant ses diagrammes en barres



Note. Image tirée de Cerda, Garcia, Jimenez et Kim (2019)

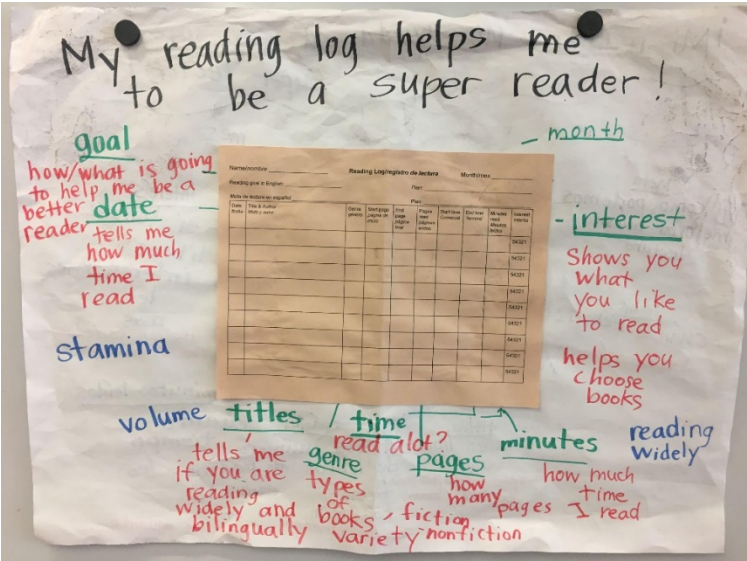
Figure 3 : Schéma final de codification

Pratiques d'apprentissage d'une langue	Activités de communication	Tâches de classe entière **	
		Tâches de groupe **	
		Tâches de partenariat **	Tutorat par élèves d'âges différents **
			Partenaires linguistiques **
		Comportement d'évaluation formative	Questionnement *, (Heritage, 2016)
			“Temps d’attente” **
		Modélisation du langage académique	Langage scolaire navigationnel *, (Bailey et Heritage, 2008)
			Langage disciplinaire **
		Assemblages sémiotiques *, (Canagarajah, 2017 ; Pennycook, 2017)	
Répétition**			
Pratiques d'apprentissage régulées	Stratégies de régulation, * (Bailey et Heritage, 2018)	Définition d'un objectif *	
		Planning *	
		Motivation *	
		Contrôle de l'attention *	
		Utilisation flexible des stratégies *	
		Contrôle *	
		Recherche d'aide et rétroaction *	
		Évaluation (Auto-évaluation) *	
Contexte	Culture de classe	Environnement de l'attention**	
		Environnement flexible **	

		Gestion de la régulation *, (Heritage, 2016)	
		Normes de classe, * (Cobb et Yackel, 1996)	Normes sociales de classe
			Normes socio- disciplinaires de classe
			Pratiques disciplinaires de classe*

Note. * = Codes déductifs,
** = Codes inductifs. Adapté de Cerda, Bailey et Heritage (2020).

Figure 4 : Le professeur a créé une affiche du carnet de lecture



Note. Image issue de Cerda, Garcia, Jimenez et Kim (2019)

Figure 5 : Les réponses de Fabricio (pseudonyme) au RISA

1ère année pilote - Printemps	2e année pilote - Printemps
<p>Dans quelle mesure votre progression en lecture en espagnol est-elle différente de votre progression en lecture en anglais ?</p> <p>Niveaux différents.</p>	<p>Mes données en lecture me montrent que...</p> <p>Que mon niveau en lecture s’améliore de plus en plus.</p>
<p>Dans quelle mesure votre progression en lecture en espagnol est-elle similaire à votre progression en lecture en anglais ?</p> <p>Parce que c’est de la lecture dans les deux cas.</p>	<p>Selon moi, cela signifie...</p> <p>Que je me suis entraîné à lire.</p>
<p>Qu’avez-vous appris sur vous-même en tant que lecteur d’espagnol et d’anglais ?</p> <p>Que la lecture est un plaisir.</p>	<p>Etes-vous un lecteur bilingue ou multilingue ?</p> <p>Lecteur bilingue.</p>
<p>Je veux être le genre de lecteur bilingue/multilingue qui...</p> <p>Lit 4 livres par jour ou parfois 6 livres par jour.</p>	<p>Pourquoi lisez-vous en deux langues ou plus ?</p> <p>Parce que ma mère et mon père veulent apprendre à parler anglais, c’est pour cela que je le parle.</p>
<p>Comment vous sentez-vous par rapport à la lecture en deux ou plusieurs langues ?</p> <p>Je suis fier de moi.</p>	

Chapitre 4. La planification stratégique : pluralités de contexte, pluralités de conceptions

Une analyse typologique des visions stratégiques des universités canadiennes

Catherine Larouche, Ashley Byrne, Denis Savard, Camille Larouche, Lucie Héon

I. Introduction

Cette étude s'inscrit dans l'axe 1 du 33ème colloque international annuel de l'ADMEE-Europe de 2022, tel que présenté dans l'introduction de l'ouvrage. Elle vise à identifier les différentes conceptions présentes dans les énoncés de vision des plans stratégiques d'universités canadiennes. Cette étude permet de faire ressortir la diversité des conceptions prédominantes en fonction de réalités contextuelles régionales et leur évolution sur deux temps de planification. Identifier les conceptions dominantes est un premier pas vers la sélection des modes d'évaluation et des indicateurs de performance qui correspondent mieux à la spécificité des universités. L'enseignement supérieur a changé plus que jamais dans les trois dernières décennies. Les universités subissent des transformations liées à l'émergence de la société du savoir, au développement démographique, aux effets de la mondialisation, à la compétition accrue, au développement de technologies de l'information et à la fluctuation des conditions économiques (Dobbins et al., 2011). La promotion de l'idéologie néolibérale, l'émergence d'un quasi-marché en enseignement supérieur et l'implantation de réformes basées sur le New Public Management constituent des courants majeurs qui orientent les réponses des universités à ces transformations sociales et économiques (Bruckmann, 2015 ;

Reed, 2002 ; Schuetze et Bruneau, 2004). Dans ce contexte de fluctuations économiques, de concurrence accrue et de changements technologiques rapides, les universités deviennent de plus en plus imputables de leurs actions avec l'augmentation d'activités de planification et de production de rapports que cela implique (Bessant et al., 2015) et ce qu'elles doivent accomplir tout en conciliant des attentes divergentes. Les universités doivent ainsi satisfaire aux exigences générales des gouvernements tout en respectant leurs propres spécificités historiques et culturelles ; s'éloigner d'un élitisme plus traditionnel (Musselin et Teixeira, 2014) afin de répondre à une demande accrue pour un système de masse avec son lot de « charges supplémentaires sur des ressources déjà étirées » (Caspersen et Frolich, 2014 ; Trow, 1974), entre autres. Face aux pressions qu'elles subissent, aux changements qui les traversent, aux dangers de dérive qui les guettent, à la diversité des pratiques qui s'installent, il est important de dresser le portrait des universités en fonction des conceptions qui y prédominent et de vérifier l'état de la situation à cet égard afin de constater l'évolution et la diversité de la réponse stratégique des universités à leurs environnements complexes. Dans cette perspective, nous émettons l'hypothèse que l'identification de conceptions dominantes, perceptible à travers les plans stratégiques de 117 universités canadiennes est un premier pas vers la reconnaissance du caractère contextuel et spécifique de leurs énoncés de vision, ce qui devrait être pris en compte dans les mécanismes d'évaluation institutionnelle et politique de leur performance.

II. Les plans stratégiques : une réflexion sur la complexité et la diversité

Afin de comprendre la réponse des universités aux changements, il importe d'identifier les diverses conceptions qui les habitent, de documenter les logiques d'actions qui mènent à leur planification et de dresser le portrait des changements qui les traversent. Notre projet vise à réaliser une analyse typologique (Larouche et al., 2012) des plans stratégiques et de leur processus d'élaboration comme révélateur des changements intervenus dans les universités canadiennes. Identifier les différentes conceptions présentes dans les plans stratégiques permet de repérer les tendances, de vérifier l'alignement stratégique, de mieux guider les actions à mettre en place pour atteindre leur vision et leurs objectifs et de choisir les indicateurs de suivi en lien avec leurs intentions.

Au Canada, particulièrement au Québec, en Ontario et en Colombie-Britannique, on note une tendance croissante à étudier les plans stratégiques des universités, cependant, ces études se limitent à une ou deux dimensions d'intérêt (par exemple, les valeurs institutionnelles, les réponses aux changements climatiques, l'internationalisation, la gouvernance...). Les portraits par province sont souvent effectués par études de cas ou études multi-cas touchant la planification stratégique d'une faculté typique, d'une université représentative ou, en de rares situations, portant sur plusieurs universités (Albon et al., 2016 ; Bégin-Caouette et al., 2018 ; Bouchard et al., 2015 ; Buzzelli et Allison, 2017 ; Eastman et al., 2019 ; Menna et al., 2016 ; Mueller, 2015 ; Viczko et Tascón, 2016). Jusqu'à maintenant, aucun portrait n'inclut une analyse de l'ensemble des documents stratégiques des établissements universitaires d'une province (hors les publications par la présente équipe sur le Québec) ou de l'ensemble des provinces (Buzzelli et Allison, 2017 ; Liinamaa, 2018 ; Menna et al., 2016).

Les résultats présentés dans ce chapitre s'inscrivent dans le cadre d'un projet de recherche triennal qui vise à analyser le processus de planification stratégique des universités canadiennes comme révélateur de changements en produisant une analyse typologique longitudinale et transversale des plans stratégiques des universités canadiennes (Larouche et al., 2012) et en identifiant les logiques d'action prédominantes privilégiées par les acteurs de la planification (Masengesho et al., 2009). Ce chapitre présente une analyse typologique longitudinale et transversale des énoncés de vision des plans stratégiques des universités canadiennes par régions canadiennes. Il tire son originalité du fait qu'il permet d'identifier les conceptions dominantes en fonction des différents contextes régionaux canadiens et de discerner les changements de conceptions entre deux exercices de planification stratégiques (variant de trois à cinq ans).

III. Le plan stratégique et ses composantes

La planification stratégique est décrite comme « a process of anticipating change, identifying new opportunities, and executing strategy » (Voorhees, 2008, p. 79) avec l'amélioration comme élément implicite du processus (Dooris et al, 2002 ; Mueller, 2015). Un plan stratégique est un document utilisé par l'organisation pour communiquer ses objectifs, les actions nécessaires pour atteindre ces objectifs et tout autre élément crucial développé pendant le processus de planification (Bieler et McKenzie, 2017,

p. 1160). Selon Masengesho et ses collaborateurs (2009) : « L'action de planifier de même que l'action de recherche sur la planification reposent sur une vision, sur une manière de connaître et d'appréhender le monde en lien avec des modes d'action et de transformation de ce monde » (p. 56). La vision est définie selon Mazouz et Leclerc (2011) comme une représentation d'une situation future désirée qui comprend un ensemble de croyances et un but général à atteindre ; la vision guide l'organisation dans la gestion du changement souhaité. La planification stratégique prend de plus en plus d'importance dans l'évaluation de la performance des établissements. Les plans comprennent dans bien des cas des cibles souvent numériques, auxquelles les établissements comparent leurs résultats. Les plans stratégiques s'inscrivent dans un courant rationaliste classique de la planification (Mintzberg, 1994). Ce processus comprend des étapes interreliées comme la définition d'une vision, l'énonciation d'une mission, de valeurs, la fixation d'orientations, d'objectifs, le choix de moyens et d'indicateurs.

Selon Elizabeth Buckner, professeure adjointe spécialisée dans la privatisation de l'enseignement supérieur à l'Ontario Institute for Studies in Education de l'Université de Toronto (citée dans MacDonald, 2019), les universités ont dû faire preuve d'entrepreneuriat et utiliser des pratiques de gestion d'entreprise pour démontrer leur capacité à gérer efficacement les ressources. Elle est d'avis que les plans stratégiques sont aussi symboliques et « témoignent, comme le feraient des artéfacts, de la manière dont les universités expriment leurs valeurs » (MacDonald, 2019).

Pour les praticiens, le plan stratégique constitue un guide qui oriente les décisions et les actions futures de l'organisation (Masengesho et al., 2009, p. 57). La planification stratégique entame un processus qui engage des parties prenantes variées, incluant le personnel académique et professionnel, les étudiants et les autres représentants de la communauté (He et Oxendine, 2019 ; Hinton, 2012 ; Paris, 2003).

Aujourd'hui, la présence d'un plan stratégique institutionnel est la norme pour les établissements postsecondaires afin de démontrer leur capacité à s'orienter d'une manière efficace et compétitive selon leur contexte. La production des plans stratégiques n'est pas une obligation imposée par les gouvernements dans toutes les provinces. Les plans stratégiques des universités canadiennes incluent des énoncés généraux (des énoncés larges sur les buts institutionnels) ainsi que des énoncés plus spécifiques tels que la mission, les buts, et les objectifs. La structure du plan est généralement composée de cinq parties : une description de la mission et de la vision de l'organisation ; le contexte dans lequel évolue l'organisation

ainsi que les principaux enjeux auxquels elle fait face ; les orientations stratégiques, les axes d'intervention et les objectifs retenus ; les résultats visés (cibles) au terme de la période couverte par le plan ; et les indicateurs de performance utilisés pour mesurer l'atteinte des résultats.

IV. Une typologie des conceptions des universités en vue d'en évaluer leur performance

Ce projet d'analyse typologique ne vise pas à évaluer l'efficacité des plans stratégiques. Il s'inscrit plutôt dans une perspective interprétative du sens des énoncés de vision, de mission, de valeurs, d'objectifs et de moyens dans le but de déceler les conceptions qui leur sont sous-jacentes et à tracer leur évolution au cours de deux exercices de planification subséquents (précédent : temps 1 ; et actuel : temps 2). Le plan stratégique est ainsi vu comme un construit émergent des interprétations que les différents acteurs impliqués se font des objectifs poursuivis (Weick, 1995 ; Larouche, 1995). La planification n'est pas considérée comme un processus neutre, mais comme étant soumise à la pression d'acteurs présentant différents intérêts.

Afin de tenir compte de la complexité et de la diversité des universités canadiennes, les énoncés de visions des plans stratégiques ont été analysés selon la typologie de Larouche et al. (2012) (voir tableau 1). Cette approche repose sur les postulats qu'à l'intérieur des établissements coexistent une pluralité de conceptions qui varient dans le temps, dans l'espace, selon des rythmes d'alternance et de prédominance, et en fonction des instances et des individus. La typologie a été développée par un processus d'anasynthèse, en s'appuyant sur une recension exhaustive des écrits sur des modèles de mesure de performance des universités et des conceptions qui leur sont sous-jacentes en vue de faire ressortir une typologie préliminaire. Cette typologie préliminaire a ensuite été validée par des entretiens avec des experts dans le domaine (Larouche et al., 2012, 2016). Elle a pour but de diversifier les indicateurs de performance en fonction des différentes conceptions des universités.

La typologie de Larouche et al., utilisée comme outil d'analyse des planifications stratégiques, permet de caractériser les conceptions présentes dans les documents analysés. Elle comporte sept types de conception : académique, service public, entrepreneuriale, marché, politique, milieu de vie, apprenante. Elles se différencient par leurs caractéristiques présentes dans sept dimensions : valeurs et principes, système de gouvernance, stratégies, recherche, enseignement, modes

d'évaluation et critères de performance. Pour les fins de cette analyse, les dimensions de la typologie ne sont pas utilisées. Les types de conceptions sont décrits sommairement dans les paragraphes suivants.

Conception académique : l'université concrétise une idée civilisatrice qui vise la formation des citoyens à un idéal d'humanité, de culture et de vie commune. Cette conception valorise la recherche de la vérité et le développement d'un savoir critique (Freitag, 2002 ; Larouche et al., 2012).

Conception service public : l'éducation, l'enseignement supérieur et la formation universitaire constituent un bien personnel et social (Larouche et al., 2012 ; MEQ, 2000), l'Université est un bien commun faisant partie du patrimoine collectif (Fédération québécoise des professeures et professeurs d'université, 1997).

Conception entrepreneuriale : l'université vise à exploiter les ressources et les possibilités de développement au-delà des moyens disponibles (Clark, 1998) : elle s'adapte à un univers complexe et incertain, mais le fait à ses propres conditions et selon ce qu'elle offre (Larouche et al., 2012).

Conception de marché : l'université est vue comme ayant une valeur marchande : elle répond à la demande dans un univers concurrentiel (Larouche et al., 2012 ; Porter, 1993).

Conception politique : l'université recherche la légitimité de ses actions auprès des parties prenantes (Morin et al., 1994 ; Quinn et Rohrbaugh, 1983) : il s'agit ici d'obtenir le soutien et la confiance des acteurs de son environnement (Larouche et al., 2012).

Conception milieu de vie : s'appuie sur un partage des traditions, des valeurs et des habitudes propres à la vie de l'organisation (Larouche et al., 2012 ; Mintzberg, 1994) assurant la pleine réalisation des capacités de la personne.

Conception apprenante : l'université est une organisation apprenante (Duke, 1992 ; James, 2000 ; Mead, 1995 ; Ramsden et Martin, 1996 ; Senge, 1990) : elle crée, acquiert, transfère des connaissances et modifie sa conduite pour refléter des connaissances et des perspectives nouvelles (Dill, 2003 ; Garvin, 1993 ; Larouche et al., 2012).

Une université peut incarner plusieurs ou toutes les conceptions simultanément ; elles ne sont pas mutuellement exclusives. Les traits caractéristiques ont été délibérément accentués dans les paragraphes précédents afin de les différencier clairement. Une première analyse des plans stratégiques d'universités québécoises a démontré l'applicabilité de la typologie de Larouche dans l'analyse des processus de planification

(Larouche et al., 2016). Fort de ces résultats, l'enjeu est maintenant de procéder à une analyse documentaire des planifications stratégiques des universités canadiennes permettant ainsi de mieux comprendre leur évolution et d'identifier les tendances sous-jacentes guidant les planificateurs. Le tableau 1 présente la typologie de Larouche et al. (2012) en détail.

Tableau 1 : Typologie des conceptions des universités en vue d'évaluer leur performance (Larouche et al., 2012)

Conceptions	Valeurs et principes	Système de gouvernance	Stratégies	Enseignement	Recherche	Évaluation de la performance	Critères d'évaluation de la performance
Académique	<ul style="list-style-type: none"> Recherche de la vérité et développement du savoir Liberté académique Fonction critique Rigueur scientifique Disciplinarité 	<ul style="list-style-type: none"> Autonomie institutionnelle Gestion collégiale 	<ul style="list-style-type: none"> Recherche de l'excellence Sélection et élitisme Méritocratie Maintien de la tradition Assurance du capital de réputation 	<ul style="list-style-type: none"> Disciplinarité Transmission des connaissances (savoir) Formation à la culture générale et scientifique à tous les niveaux Formation à la recherche aux cycles supérieurs 	<ul style="list-style-type: none"> Recherche libre Recherche fondamentale et appliquée 	<ul style="list-style-type: none"> Évaluation par les pairs avec vérification externe Reconnaissance internationale Agrément 	<ul style="list-style-type: none"> Excellence académique Avancement des connaissances Qualité des programmes Notoriété en recherche Rayonnement sur le monde Prestige et distinctions
Service public	<ul style="list-style-type: none"> Réponse aux besoins de la société Accès à l'université Démocratisation Normes explicites 	<ul style="list-style-type: none"> Démocratisation de l'université Présence active de l'État et de la société Coûts assumés principalement par la société 	<ul style="list-style-type: none"> Plan d'action Débat démocratique 	<ul style="list-style-type: none"> Offre de formation en fonction des besoins diversifiés de la société 	<ul style="list-style-type: none"> Recherche qui répond aux besoins de la société Valorisation de la diversité des approches et des modes de diffusion Recherche développement 	<ul style="list-style-type: none"> Imputabilité Reddition de compte à l'État et à la population Audit institutionnel 	<ul style="list-style-type: none"> Accès et réussite du plus grand nombre Équité Atteinte des objectifs sociétaux Services à la collectivité Développement régional Intégration citoyenne et professionnelle
Marché	<ul style="list-style-type: none"> Concurrence Rentabilité Niche stratégique 	<ul style="list-style-type: none"> Système client Lois de l'offre et de la demande avec prédominance à la demande 	<ul style="list-style-type: none"> Position stratégique sur le marché Alliances tactiques Marketing Branding (pouvoir de la marque) 	<ul style="list-style-type: none"> Commercialisation de l'enseignement Enseignement sur mesure Développement des connaissances techniques et industrielles 	<ul style="list-style-type: none"> Recherche orientée vers les besoins du marché Commercialisation de la recherche Recherche commanditée 	<ul style="list-style-type: none"> Palmarès Veille concurrentielle Profit 	<ul style="list-style-type: none"> Recrutement de clientèles Satisfaction de la clientèle Intégration des étudiants sur le marché du travail et taux de diplomation Expansion et diversification du marché Compétitivité Adaptabilité Rentabilité
Entrepreneuriale	<ul style="list-style-type: none"> Autonomie Innovation Goût du risque Pérennité institutionnelle 	<ul style="list-style-type: none"> Système producteur (offre) Leader Équipe de gestion Équipe de recherche Diversification des sources de financement 	<ul style="list-style-type: none"> Vision Mobilisation des ressources Changement délibéré 	<ul style="list-style-type: none"> Développement des compétences et des créneaux en lien avec la vision 	<ul style="list-style-type: none"> Recherche-développement Innovation et transfert technologique Brevet d'invention Interdisciplinarité et résolution de problème 	<ul style="list-style-type: none"> Concrétisation des finalités Évaluation des forces et des faiblesses, des opportunités et des menaces 	<ul style="list-style-type: none"> Capacité d'être proactif Innovation Capacité de pilotage Diversité des sources de financement Partenariat

Politique	<ul style="list-style-type: none"> Processus démocratique Reconnaissance des intérêts divergents Respect de la diversité des composantes de la société Changement social 	<ul style="list-style-type: none"> Diversification et participation des parties prenantes Répartition des lieux de pouvoir Rapport de force 	<ul style="list-style-type: none"> Processus de négociation Recherche du compromis Consultation Alliances stratégiques 	<ul style="list-style-type: none"> Participation des parties prenantes à l'élaboration des programmes et des curriculums 	<ul style="list-style-type: none"> Recherche stratégique Participation des parties prenantes aux orientations et au développement des projets de recherche 	<ul style="list-style-type: none"> Imputabilité et reddition de compte envers les parties prenantes Opinion publique 	<ul style="list-style-type: none"> Satisfaction des parties prenantes Satisfaction des employeurs Satisfaction des organismes régulateurs Satisfaction des associations étudiantes, professionnelles et syndicales Satisfaction des communautés internes et externes
Milieu de vie	<ul style="list-style-type: none"> Respect et épanouissement de la personne Amélioration de la condition humaine Bien-être au travail Solidarité intergénérationnelle 	<ul style="list-style-type: none"> Leadership orienté vers le développement de la personne Développement durable Gestion écologique et humaniste 	<ul style="list-style-type: none"> Processus humain Soutien au développement de carrière et au ressourcement professionnel 	<ul style="list-style-type: none"> Développement intégral de la personne Enseignement personnalisé Petits groupes Valorisation des activités parascolaires 	<ul style="list-style-type: none"> Valorisation de la diversité de la recherche Recherche orientée vers les humanités Recherche libre 	<ul style="list-style-type: none"> Climat d'étude et de travail Culture organisationnelle Évaluation institutionnelle 	<ul style="list-style-type: none"> Mobilisation au travail Sentiment d'appartenance Disponibilité des espaces Santé et sécurité au travail Climat d'étude et de travail Développement des ressources humaines Éthique Esthétisme
Apprenante	<ul style="list-style-type: none"> Multiplication des lieux de savoir et d'information Communication Utilisation des TIC Ouverture 	<ul style="list-style-type: none"> Gestion participative Communauté universitaire Gestion décentralisée et démocratique 	<ul style="list-style-type: none"> Processus d'apprentissage Développement organisationnel Diffusion du savoir Dissémination du savoir Communautés de pratique et d'apprentissage Ouverture (différents schèmes de pensée) Décloisonnement 	<ul style="list-style-type: none"> Rencontre des savoirs entre le milieu universitaire et les autres milieux Processus d'enseignement Enseignement différencié Reconnaissance des acquis et valorisation des compétences Pédagogie par projet Communauté de pratique Transdisciplinarité 	<ul style="list-style-type: none"> Recherche transdisciplinaire Réseaux de recherche Consortium de recherche Implication des partenaires dans la programmation de recherche Recherche collaborative et recherche-action 	<ul style="list-style-type: none"> Tableau de bord de gestion Évaluation mixte impliquant des acteurs externes 	<ul style="list-style-type: none"> Système d'information adéquat Mécanisme de réflexion et d'amélioration Capacité de s'autoévaluer Développement et extension des réseaux Apprentissage de l'organisation

V. Méthodologie

L'analyse typologique des plans stratégiques s'effectue à l'aide d'une analyse documentaire qui permet d'obtenir une coupe tant longitudinale

que transversale et qui favorise l'observation des processus et leur évolution (Cellard, 1997). Elle facilite l'identification des tendances dans le monde universitaire, la prise de conscience de ces tendances, et elle apporte un meilleur éclairage quant aux décisions futures.

Échantillon et collecte de données

Les documents stratégiques de 117 universités (sur les 146 au Canada), ont été colligés sur deux cycles de planification, ce qui se traduit en 192 documents stratégiques pour une période de couverture de 1990 jusqu'à 2030. Les documents stratégiques incluent les plans stratégiques et, pour certaines universités, en l'absence des plans stratégiques, des ententes de mandats gouvernementaux ou des plans académiques qui portent un regard sur l'institution au complet, qui mentionnent la mission, la vision et les valeurs et qui ont été développés à des fins stratégiques.

Afin d'effectuer une analyse longitudinale des visions, les analyses se sont intéressées aux universités pour lesquelles au moins deux documents stratégiques consécutifs ont été recueillis : l'actuel plan stratégique (temps 2) et le précédent (temps 1) selon la périodicité retenue. Ainsi, les analyses ont porté sur 132 documents provenant de 66 universités, ce qui représente, par région : 18 documents des provinces atlantiques (9 des 23 universités), 42 documents de l'Ontario (21 des 43 universités), 32 documents du Québec (16 des 21 universités) et 40 des provinces à l'Ouest de l'Ontario (20 des 59 universités).

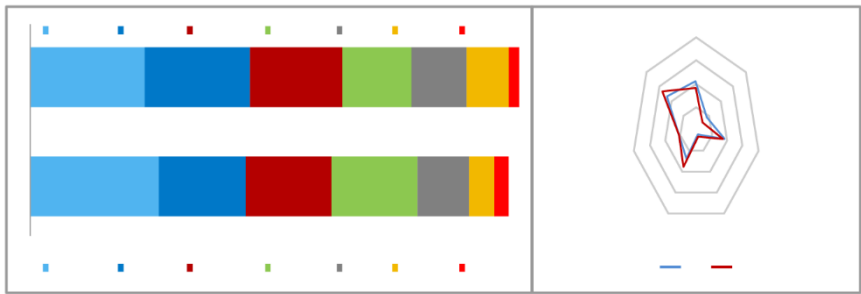
Analyses

Le contenu des documents stratégiques a été codé selon les composantes d'un plan (document) stratégique (mission, vision, valeurs, orientations, priorités, axes, enjeux, buts, objectifs, moyens, indicateurs et cibles) et segmenté en unités de sens, selon la méthode de Morin et Chartier cités par Leray (2011). Par la suite, les documents ont été codés en fonction des sept conceptions et des sept dimensions de la typologie de Larouche et al. (2012). Les segments ont été codés, validés doublement par trois codeurs différents pour garantir la fiabilité inter-codeur. Tous les codages ont été réalisés avec le logiciel QDA Miner de Provalis Research. Les occurrences ont ensuite été compilées pour l'ensemble des universités (portrait global), par grande région canadienne (1. Atlantique ; 2. Québec ; 3. Ontario et 4. Ouest), par province et par université. Nous présentons les résultats de la composante de la vision, en comparant les quatre régions du Canada.

VI. Résultats

En regardant la comparaison des deux moments de recueil de l'ensemble des universités canadiennes (Figure 1), on constate que les résultats entre le temps 1 et le temps 2 sont assez stables. Les universités canadiennes dans l'ensemble incarnent principalement les conceptions service public, académique, entrepreneurial et milieu de vie. Les résultats laissent présager de légers changements de fréquence des conceptions sur ce profil global canadien : une légère augmentation des conceptions milieu de vie (3,8 %), service public (3,5 %), et marché (0,9 %) et une légère diminution de la présence des conceptions apprenante (-3,5 %), académique (-3,3 %), entrepreneuriale (-0,7 %) et politique (-0,4 %).

Figure 1 : Pourcentages d'occurrences des conceptions dans les énoncés de vision des plans stratégiques pour l'ensemble du Canada



Occurrences t1 = 310 ; t2 = 323

Pour les deux provinces où l'on retrouve le plus d'universités, le Québec (figure 2) et l'Ontario (figure 3), on peut voir une différence au niveau des conceptions dominantes. Au Québec, les trois conceptions prédominantes sont le service public, le milieu de vie et l'académique, alors qu'en Ontario la prédominance va aux conceptions service public, entrepreneuriale et académique. Au Québec, entre les temps 1 et 2, on note une légère augmentation pour les conceptions politique (5,0 %), académique (3,1 %), milieu de vie (2,6 %), et une légère diminution pour les conceptions entrepreneuriale (-5,0 %), apprenante (-2,8 %), service public (-2,5 %) et marché (-0,7 %). En ce qui a trait à l'Ontario, on observe une légère augmentation pour les conceptions politique (4,4 %), milieu de vie (3,5 %)

et marché (1,6 %) et une diminution pour les conceptions apprenante (-7,4 %), académique (-1,4 %) et service public (-0,9 %). Le pourcentage des occurrences pour la conception entrepreneuriale n'a pas changé entre le temps 1 et le temps 2.

Figure 2 : Pourcentages d'occurrences des conceptions dans les énoncés de vision des plans des stratégies pour le Québec

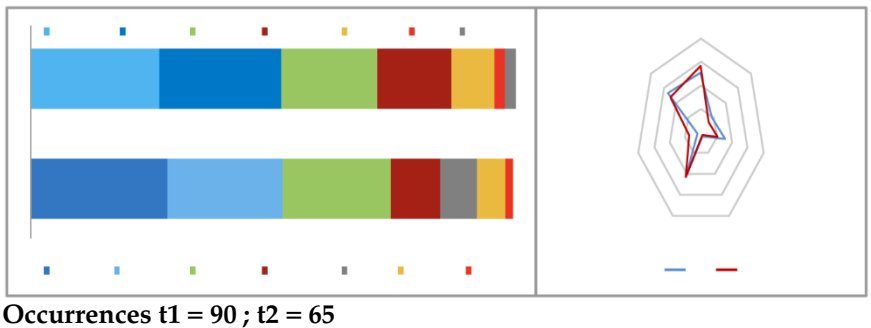
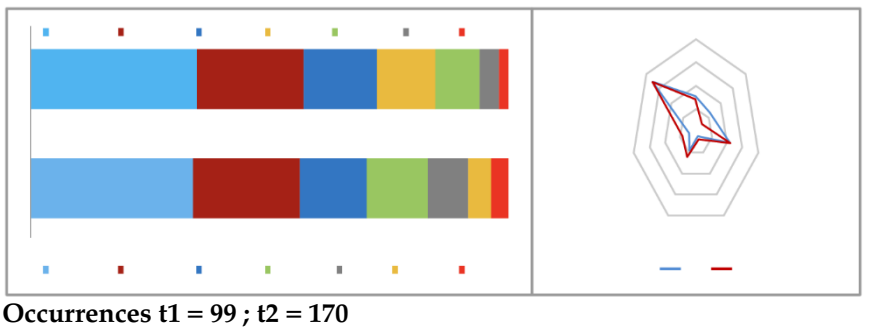


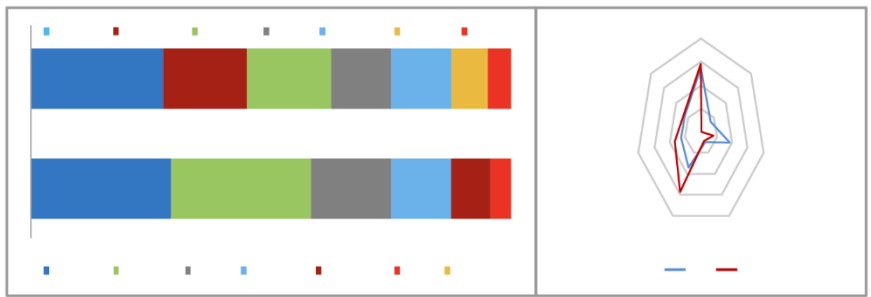
Figure 3 : Pourcentages d'occurrences des conceptions dans les énoncés de vision des plans stratégiques pour l'Ontario



Pour les deux autres régions qui regroupent les provinces atlantiques et de l'Ouest canadien, on observe des résultats plus variés. Pour la région Atlantique, les conceptions milieu de vie (11,7 %), politique (4,2 %) et académique (1,7 %) ont vu une légère augmentation entre le temps 1 et le

temps 2 (figure 4). Les conceptions entrepreneuriale (-9,2 %) et marché (-0,8 %) ont connu une légère diminution. La conception apprenante est passée de 7,5 % à néant (0,0 %) alors que les pourcentages associés à la conception service public n'ont pas changé.

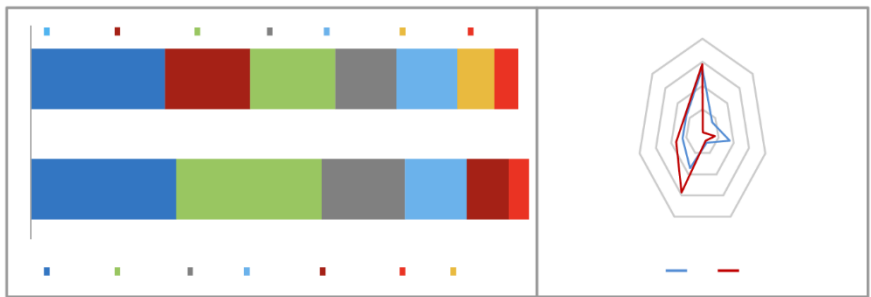
Figure 4 : Fréquence de l'occurrence des conceptions dans les énoncés de vision des plans stratégiques pour l'Atlantique



Occurrences t1 = 41 ; t2 = 23

Pour la région de l'Ouest (figure 5), il y a des augmentations pour les conceptions milieu de vie (10,1 %), service public (5,6 %), apprenante (2,6 %) et marché (1,7 %) ainsi que des diminutions pour les conceptions politique (-11,3 %), académique (-7,0 %) et entrepreneuriale (-1,7 %).

Figure 5 : Fréquence de l'occurrence des conceptions des énoncés de vision dans les plans stratégiques pour l'ouest



Occurrences t1 = 80 ; t2 = 65

L'échantillon plus petit pour l'Atlantique et l'Ouest (9 cas et 20 cas respectivement) appelle à la prudence dans les conclusions à tirer relativement à ces régions, alors il faudra attendre d'autres analyses avant de porter des conclusions sur ces régions.

VII. Discussion et conclusion

Cette recherche permet aux planificateurs et aux responsables de l'élaboration de politiques, de mieux situer leurs enjeux de planification et leurs évolutions à l'aide de la typologie des conceptions, et de développer des stratégies et des outils d'évaluation des performances des universités en fonction des conceptions révélées par ce portrait. Les résultats confirment la diversité des universités, reflétée par la différenciation claire des quatre régions. Chacune des régions ayant des environnements distincts, on peut présumer que la diversité qui existe dans les visions est une représentation de la réponse stratégique des universités à leurs environnements complexes. Cependant, les résultats au niveau régional n'expliquent pas au premier abord les nuances du profil canadien agrégé. Chaque région présente une combinaison distincte de conceptions. Pour certaines régions, comme c'est le cas au Québec et en Ontario, les conceptions dominantes restent dominantes au fil du temps. Dans l'Atlantique, la conception dominante reste la même tandis que toutes les autres changent d'importance, cédant la place à une dominance partagée de l'académique et du milieu de vie. Dans le cas de la région de l'Ouest, l'incohérence entre les deux temps est remarquable. Cependant, aucun test n'a été effectué pour déterminer la signification de ces comparaisons et la taille d'effet de ces statistiques. Pour cette raison, nous ne sommes pas en mesure de confirmer si les analyses du changement de moyen entre le temps 1 et le temps 2 sont statistiquement significatives, ni si ces résultats sont représentatifs de l'ensemble de la population universitaire régionale ou canadienne. Ces analyses préliminaires ont néanmoins confirmé la diversité des conceptions des universités canadiennes à l'égard des visions présentes dans leurs documents stratégiques.

Afin d'approfondir la compréhension de la diversité des conceptions, une prochaine étape consistera à identifier les logiques d'action prédominantes privilégiées par les acteurs de la planification en fonction du contexte et de stratégies identifiées à l'aide d'études de cas de huit universités canadiennes.

VIII. Bibliographie

- Albon, S. P., Iqbal, I. et Pearson, M. L. (2016). Strategic Planning in an Educational Development Centre: Motivation, Management, and Messiness. *Collected Essays on Learning and Teaching*, 9, 207.
- Bégin-Caouette, O., Trottier, C., Eastman, J., Jones, G. A., Noumi, C. et Li, S. X. (2019). Analyse de la gouvernance systémique des universités au Québec et comparaisons avec quatre autres provinces canadiennes. *Canadian Journal of Higher Education*, 48(3), 1–22. <https://doi.org/10.7202/1057126ar>
- Bieler, A. et McKenzie, M. (2017). Strategic Planning for Sustainability in Canadian Higher Education. *Sustainability*, 9(2), 161.
- Bessant, S. E. F., Robinson, Z. P. et Ormerod, R. M. (2015). Neoliberal-ism, new public management and the sustainable development agenda of higher education: history, contradictions and synergies. *Environmental Education Research*, 21(3), 417–432.
- Bouchard, I., Boucher, M.-P., Ferland, M., Lalande, C., Lavergne, N., Ma-rois, S. et Bilodeau, A. (2015). *L'internationalisation de l'enseignement supérieur : Enjeux et défis pour le Québec*. Fédération nationale des enseignantes et des enseignants du Québec.
- Bruckmann, S. (2015). Shifting Boundaries in Universities Governance Models: The Case of External Stakeholders. *Higher Education Research in the 21st Century*, 8, 163–186.
- Buzzelli, M. et Allison, D. J. (2017). Proposed Strategic Mandates for Ontario Universities: An Organizational Theory Perspective. *Canadian Journal of Higher Education*, 47(3), 170–191.
- Caspersen, J. et Frølich, N. (2014, forthcoming). Læringsutbytte som styringsredskap for ledelsen i høyere utdanning. Dans N. Frølich, E. Hovdhaugen et L. I. Terum (dir.), *Kvalitet, kapasitet og relevans*. Cappelen Damm Akademisk.
- Cellard, A. (1997). L'analyse documentaire. Dans J., Poupart, L.-H. Groulx, J.P. Deslauriers, A. Laferrière et R. Mayer, (dir.), *La recherche qualitative : enjeux épistémologiques et méthodologiques* (p. 251-262). Gaétan Morin éditeur.
- Clark, B.R. (1998). *Creating entrepreneurial universities: organizational pathways of transformation*. Pergamon Press.

- Dill, D. (2003). Le paradoxe de la qualité académique : Implications pour les universités et les politiques publiques. *Revue des sciences de l'éducation*, 29(2), 337-352.
- Dobbins, M., Knill, C. et Vogtle, E. M. (2011). An analytical framework for the cross-country comparison of higher education governance. *High Educ*, 19.
- Dooris, M. J., Kelley, J. M. et Trainer, J.F. (2002). Strategic planning in higher education. *New Directions in Higher Education*, 116, 5–11.
- Duke, C. (1992). *The learning University: Toward a new paradigm ?* Society for research into Higher Education et Open University Press.
- Eastman, J. A., Jones, G. A., Begin-Caouette, O., Li, S. X., Noumi, C. et Trotter, C. (2019). Provincial Oversight and University Autonomy in Canada: Findings of a Comparative Study of Canadian University Governance. *Canadian Journal of Higher Education*, 48(3), 65-81.
- Freitag, M. (2002). L'université aujourd'hui : Les enjeux du maintien de sa mission institutionnelle d'orientation de la société. Dans G. Gagné (dir.), *Main basse sur l'éducation*. Éditions Nota bene.
- Garvin, D. (1993). Building a learning organization. *Harvard Business Review*, 71(4), 78- 84
- He, Y. et Oxendine, S. D. (2019). Leading positive change in higher education through appreciative inquiry: A phenomenological exploration of the strategic planning process. *Journal of Higher Education Policy and Management*, 41(2), 219–232.
- Hinton, K. E. (2012). *A Practical Guide to Strategic Planning in Higher Education*. The Society for College and University Planning, (p. 59).
- James, R. (2000). L'assurance qualité et le difficile problème de la gestion du savoir organisationnel en milieu universitaire. *Gestion de l'enseignement supérieur*, 12(3), 45-64.
- Larouche H. (1995). From decision to action in organizations: decision-making as social Representation. *Organisation Science*, (6), 62-75.
- Larouche, C., Savard, D., Héon, L. et Moisset, J.-J. (2012). Typologie des conceptions des universités en vue d'en évaluer la performance : Rendre compte de la diversité pour en saisir la complexité. *Canadian Journal of Higher Education*, 42(3), 45–64. <https://doi.org/10.47678/cjhe.v42i3.2429>

- Larouche, C., Savard, D. et Richard, J.-F. (2016). Applicabilité et utilité de la typologie de Larouche et al. (2012) comme outil de gouvernance, de planification et de gestion. Rapport de recherche. [Document inédit].
- Leray, C. (2011) *L'analyse de contenu : De la théorie à la pratique-La méthode Morin Chartier*. Les presses de l'Université du Québec.
- Liinamaa, S. (2018). Negotiating a « Radically Ambiguous World »: Planning for the Future of Research at the Art and Design University. *International Journal of Art & Design Education*, 37(3), 426-437.
- McDonald, M. (2019, 8 mai). Les plans stratégiques aident les universités à se raconter. Affaires universitaires. <https://www.affairesuniversitaires.ca/articles-de-fond/article/les-plans-strategiques-aident-les-universites-a-se-raconter/>
- Masengesho, K., Bonami, M. et De Ketele, J.-M. (2009). Modèle d'analyse des conceptions et des pratiques de planification. *Revue française de gestion*, 5(195), 55-83.
- Mazouz, B. et Leclerc, J. (2011). *La gestion intégrée par résultats. Concevoir et gérer autrement la performance dans l'Administration publique*. Presses de l'Université du Québec.
- Mead, P. (1995). Utilising the university as a learning organization to facilitate quality improvement. *Quality in Higher education*, 1(2), 111-121.
- Menna, A., Catalfamo, H. et Girolamo, C. (2016). Applied Entrepreneurship Policy: Ontario's Colleges in the Age of Globalization. *Educational Planning*, 23(2), 59-74.
- Ministère de l'Éducation du Québec (MEQ). (2000). *Politique québécoise à l'égard des universités : Pour mieux assurer notre avenir collectif*. Gouvernement du Québec.
- Mintzberg H. (1994). *Grandeur et décadence de la planification stratégique*. Dunod.
- Morin, E.M., Savoie, A. et Beaudin, G. (1994). *L'efficacité de l'organisation : Théories, représentations et mesures*. Gaëtan Morin.
- Mueller, R. A. (2015). Do Values Drive the Plan ? Investigating the Nature and Role of Organizational Values in University Strategic

- Planning. *Tertiary Education and Management*, 21(1), 41-55.
doi:10.1080/13583883.2014.998270
- Musselin, C. et Teixeira, P. N. (2014). *Reforming Higher Education: Public Policy Design and Implementation* (Vol. 41). Springer Netherlands.
<https://doi.org/10.1007/978-94-007-7028-7>
- Paris, K. A. (2003). *Strategic Planning in the University*. University of Wisconsin – Madison.
- Porter, M.E. (1993). *L'avantage concurrentiel des nations*. InterÉditions.
- Quinn, R. E. et Rohrbaugh, J. (1983). A spatial model of effectiveness criteria: towards a competing values approach to organizational analysis. *Management Science*, 29(3), 363.
- Ramsden, P. et Martin, E. (1996). Recognition of good university teaching: Policies from an Australian study. *Studies in Higher Education*, 16(2), 129–150.
- Reed, M. I. (2002). New Managerialism, Professional Power and Organisational Governance in UK Universities: A Review and Assessment. Dans A. Amaral, G. A. Jones et B. Karseth (dir.), *Governing Higher Education: National Perspectives on Institutional Governance* (p. 163–185). Springer Netherlands.
https://doi.org/10.1007/978-94-015-9946-7_9
- Schuetze, H. G. et Bruneau, W. (2004). Less State, More Market: University Reform in Canada and Abroad. *Canadian Journal of Higher Education*, 34(3), 1–12.
- Senge, P. (1990) *The Fifth Discipline : The Age and Practice of the Learning Organization*. Century Business.
- Trow, M. (1974). *Problems in the transition from elite to mass higher education. Policies for Higher Education*. OECD.
- Viczko, M. et Tascón, C. I. (2016). Performing Internationalization of Higher Education in Canadian National Policy. *Canadian Journal of Higher Education*, 46(2), 1-18.
- Voorhees, R. A. (2008). Applying mixed methods techniques in strategic planning. *New Directions for Institutional Research*, (137), 5–13.
doi:10.1002/ir.235
- Weick, K. E. (1995). *Sensemaking in organizations*. Sage Publications.

Chapitre 5. L'évaluation aujourd'hui et demain. Quelles ruptures contextuelles ? Quelles ruptures culturelles ?

Conférence plénière

Isabelle Nizet

I. Introduction

Les dernières années ont placé la planète au cœur d'événements générateurs de ruptures à tous niveaux. En éducation : menace sur la sécurité physique des lieux d'éducation, discontinuité des services éducatifs et de formation, déplacement spatio-temporel du travail scolaire, du travail enseignant, de l'apprentissage et bien sûr, de l'évaluation des apprentissages. L'éducation, dans ses fonctions sociales, a été confrontée à des réalités socio-sanitaires, économiques, voire écologiques menaçant les conditions de sa viabilité à court terme, du moins dans les formes que nous connaissons. L'évaluation a émergé de cette réalité comme objet de débats : fardeau psychologique insoutenable dans un contexte de santé mentale précaire des apprenants et des enseignants pour certains, dernier rempart pour assurer une valeur minimale aux diplômes en situation de crise pour d'autres. Objet d'accommodements innombrables, l'évaluation, comme ensemble de pratiques professionnelles, sociales et culturelles a alors vu vaciller plusieurs de ses principes fondateurs : une perspective exclusivement linéaire du temps et son impact sur la notion de réussite ou d'échec, l'inscription des individus dans des trajectoires de développement balisées par la passation d'épreuves et des échéances normées, la nécessité d'arbitrages évaluatifs garantissant une stabilité de la valeur symbolique ou marchande des diplômes. Bref, ce qui nous paraissait aller de soi ne va

plus de soi. Dans ce texte, il s'agit bien de réfléchir à la discontinuité de ces principes, voire à leur remise en question pour aborder la thématique de la pluralité des contextes, comme l'a proposé la 33^{ème} édition du colloque annuel de l'Admée Europe. Ainsi à la réalité de ces ruptures contextuelles, nous accolerons celle de ruptures culturelles liées aux enjeux évaluatifs. Nous explorerons aussi comment en contexte de discontinuité et de transition, une authentique éducation culturelle en évaluation peut représenter une piste pour conserver un pouvoir d'agir pour tous les acteurs.

II. Continuité en évaluation : un système séculaire

En premier lieu, nous avons choisi de repérer un certain nombre de continuités qui marquent l'évaluation depuis de très nombreuses années à travers une relecture socio-historique et critique des tendances qui l'animent.

Historiquement, l'évaluation scolaire est une réalité séculaire, du moins dans son expression formalisée puisque son origine est, en Europe, étroitement liée à l'instauration du système d'instruction des Jésuites dès le XV^e siècle. Si la pédagogie ignatienne peut être considérée comme un progrès puisqu'elle vise une transition entre un système d'instruction élitiste fondé sur le préceptorat réservé aux classes supérieures et une scolarisation étendue et gratuite (Maulini, 1996 ; 2017), c'est aussi elle qui introduira progressivement l'idée de classer et noter les élèves. L'introduction d'une éducation collective - associée au modèle de la classe comme un espace fermé - établit dès la fin du XVI^e siècle un esprit d'uniformisation et d'émulation dans laquelle ce que l'on appelle aujourd'hui l'évaluation individuelle prend place, mettant en œuvre un système fondé sur des programmes d'apprentissage. L'évaluation devient alors objet d'une forme de ritualisation selon un modèle méritocratique. Ainsi dès cette époque, dans le système jésuitique, le classement des élèves est effectué en fonction d'une double logique de compétition : la classe est divisée en deux camps, et au sein de ces deux camps, les élèves sont placés dans des « décuries » (terme militaire romain) hiérarchisées, au sein desquelles ils sont aussi en compétition pour y obtenir un meilleur rang et progresser vers la décurie supérieure ; le jugement sur l'élève détermine donc sa place dans une « hiérarchie de groupes » (Durkheim, 1938, cité par Maulini, 1996, p. 4). Par la suite, lors de la sécularisation de l'instruction,

les états reprendront ce système de classement pour l'adapter et le transformer progressivement, selon les pays, en appréciation chiffrée (et parfois lettrée dans les pays anglo saxons) en pointage sur 10, 20, 100 ou parfois 6. On voit ensuite apparaître au XIX^e siècle une organisation scolaire structurée pour répondre à un important besoin de scolarisation de masses, privilégiant un classement des élèves fondé sur les chiffres et devant répondre à la nécessité d'homogénéiser les classes pour faciliter le travail du maître. C'est ainsi qu'apparaissent les concepts de « moyenne » et de « redoublement » requis pour assurer une gestion systématique reposant sur trois facteurs qui influencent désormais les pratiques évaluatives : l'optimisation des ressources éducatives, l'encadrement des masses et la sélection des élites (Maulini, 1996). La forme scolaire contemporaine reproduit encore largement cette structure, agissant comme toile de fond dans les débats sur l'évaluation.

Le domaine de la mesure et évaluation s'est développé autour d'un courant psychométrique très influent centré sur la mesure dite scientifique (Frenette et Hébert, 2021). Historiquement, c'est bien le versant quantitatif du domaine de la mesure qui a d'abord été développé surtout à partir des années 1930, ce qui s'est traduit par « l'appréciation métrique des comportements des élèves de leurs caractéristiques psychologiques et bien sûr de leurs apprentissages » (Morrissette, 2021, p.55). La fonction de « sanction », généralement exprimée de manière chiffrée et quantitative, répond à un besoin social de certification, de sélection ou de classement. La perspective sociocritique de Morrissette nous permet de considérer l'évaluation comme une activité sociale mue et portée par « des contextes politiques et socioculturels qui contribuent à en déterminer la portée et la signification » (Morrissette, 2021, p. 69). De manière générale, ce regard sociocritique invite à considérer l'évaluation scolaire comme favorisant le renouvellement d'une forme de méritocratie (Brown et Duru-Bellat, 2010) fondée sur un esprit quantophrénique défini en sociologie comme un excès d'usage mathématique pathologique consistant à vouloir traduire systématiquement les phénomènes sociaux et humains en langage mathématique (Sorokin, 1959, évoqué par Morrissette, 2021, p. 55). Cette perspective a contribué à la légitimité et à la crédibilité sociale de l'évaluation scolaire (Hadj, 2021), méritocratie et quantophrénie semblant donc indissolublement liées. Cette évolution historique laisse bien sûr des traces profondes dans les systèmes évaluatifs contemporains. Bien que les effets de cette double tendance soient sporadiquement questionnés à petite échelle, dans plusieurs pays, sa stabilité n'est cependant ni politiquement, ni socialement remise en cause ; ce qui la pare de l'illusion d'une

permanence intouchable, malgré sa relativité sur une échelle de temps long.

III. Une continuité sur fond de dissensions

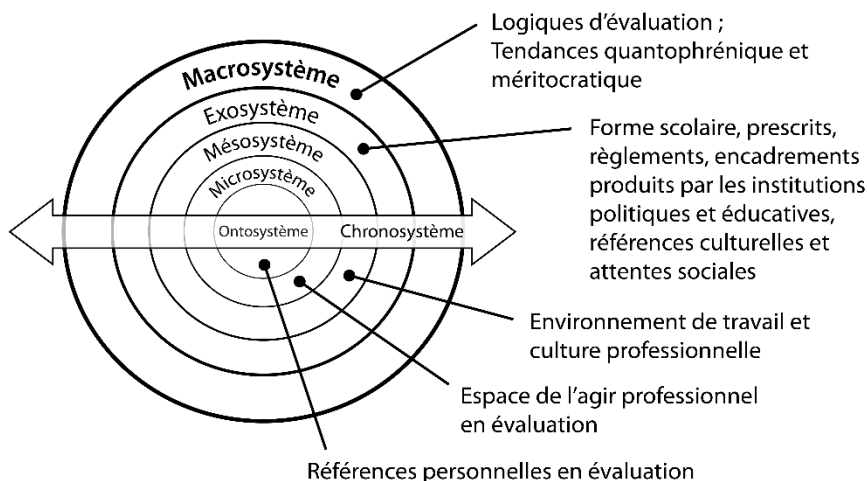
C'est ici qu'il nous semble important de revisiter les propositions de Vial (2009), selon lequel l'agir évaluatif individuel s'ancre dans une dynamique culturelle de référenciation (au sens d'apports de références culturelles diversifiées), ce que démontre bien évidemment la perspective historique que nous venons d'aborder. Ainsi, si nous paraphrasons cet auteur, chaque personne évaluatrice possède et développe un état d'esprit, des attitudes permettant l'utilisation de postures dont l'intelligibilité repose sur le partage de références communicables. Ces références ne sont pas toujours portées à la conscience de la personne évaluatrice, à moins qu'on ne lui demande de les expliciter et qu'elle soit formée à les communiquer ; elles constituent cependant des ancrages culturels très influents sous forme de modes de pensées et de logiques évaluatives socialement partagées et très souvent polarisées. Selon Vial (2009), deux logiques évaluatives orientent la culture évaluative individuelle et collective dans nos contextes : la première qu'il nomme « logique du contrôle » s'apparente à une conception plutôt mécaniciste du monde dans laquelle prévaudrait le souci d'une évaluation dont la fonction est de stabiliser, de fixer, d'ordonner, voire de rigidifier (Nizet, 2015). La seconde qu'il dénomme la « logique du Reste » (c'est-à-dire tout ce qui échappe au contrôle) s'apparente à une conception plus biologiste du monde dans laquelle prévaudrait le souci d'une évaluation plutôt écologique prenant en compte la réalité évolutive qui caractérise le vivant (Nizet, 2015). Contrairement à ce que l'on pourrait croire, ces deux logiques « ne sont pas deux pôles qui se repoussent et s'excluent [...] » (Vial, 2009, p. 76-77), car les tenants de ces deux logiques, poussés par le poids de leurs convictions, à force de vouloir s'en tenir à une seule logique, dérivent vers l'autre logique. Par exemple, des démarches d'évaluation formative qui tentent d'éviter à tout prix une logique de contrôle pourraient très bien dériver vers celle-ci de manière insidieuse en visant la maîtrise complète de savoirs morcelés (Nizet, 2015).

IV. Du chronosystème au microsystème de l'évaluation

Dans notre conférence, nous avons d'emblée proposé de lire la réalité contemporaine de l'évaluation dans nos contextes nationaux à la lumière du cadre de Bronfenbrenner (1993), présenté comme une structure écosystémique à cinq niveaux mutuellement interdépendants et orientant le développement humain. Bien que ce cadre ne soit pas initialement destiné à l'analyse de notre sujet, y référer nous a semblé utile pour proposer une clé de description des dimensions à interroger lorsqu'on s'intéresse à l'évaluation pédagogique dans sa continuité sociohistorique et pour situer les ruptures culturelles dont elle serait l'objet. Nous décrirons plus loin pourquoi nous le qualifions plutôt de modèle socio-écosystémique.

Dans la perspective proposée (voir Figure 1), c'est au niveau du macrosystème que se situeraient les deux logiques évaluatives, les tendances quantophréniques et la dynamique méritocratique. En effet, ce macrosystème se centre sur les valeurs sociales et culturelles ayant un pouvoir d'influence sur les comportements et attitudes de l'individu. Je prends ici le parti de nommer en vrac des tendances générales associées à ce macrosystème, comme un ensemble de valeurs et de finalités associées à l'évaluation scolaire et qui ont été abondamment décrites : performance, compétition, reddition de compte, régulation sociale, relation de pouvoir, etc. Mais aussi, malgré tout, des valeurs de démocratisation et d'innovation en tension récurrente.

Figure 1 : Modèle écosystémique des dimensions de l'évaluation (inspiré de Bronfenbrenner, 1993)



Cette influence s'exerce à travers deux sous-systèmes : l'exosystème constitué de facteurs influençant plus ou moins directement les pratiques évaluatives de l'individu à travers des structures sociales et organisationnelles (par exemple la forme scolaire, les prescrits ministériels, les règlements, les encadrements produits par les institutions politiques et éducatives, les références culturelles et les attentes sociales sur lesquelles elles s'appuient à leur insu), et le mésosystème constitué des lieux dans lesquels l'individu investit un temps significatif (par exemple son environnement de travail et les cultures professionnelles qui s'y enracinent). L'individu interagit avec un microsystème constitué de son environnement immédiat, espace proche de son agir professionnel avec ses contraintes et ses possibles, et son ontosystème : l'ensemble de ses références culturelles personnelles en évaluation, ses habitudes, son expérience, etc.

Enfin, le chronosystème représente à la fois la dimension temporelle de l'environnement humain de l'individu et de ses variables sociales et culturelles. Le chronosystème de l'évaluation est, quant à lui, étroitement relié à une forme scolaire marquée par un espace-temps stable et uniforme, fondé sur une conception linéaire du temps et close de l'espace d'apprentissage (Vincent et al., 2012), dans un cheminement balisé par des

épreuves et des moments de diplomation à hauts enjeux. La validité et la fiabilité de l'évaluation reposent encore largement sur la reproduction de conditions comparables qui ne sont possibles que dans ce chronosystème contraignant. Le chronosystème traverse tous les espaces, du macro à l'ontosystème. Le bouleversement de ce chronosystème et de l'espace scolaire semble être le principal effet des périodes de discontinuité pédagogique imposées par la situation pandémique, dévoilant ainsi la fragilité d'une organisation spatiotemporelle sur laquelle repose les fondements pragmatiques de l'évaluation scolaire et qui détermine largement sa réalité contextuelle.

De ce tour d'horizon, il nous semble qu'il faut retenir l'étroite intrication existante entre la dimension culturelle et contextuelle de l'évaluation, et cela à tous les niveaux d'un modèle *socio-écosystémiques*. De cette manière, nous rejoignons les propos de Ofir (2018) pour qui la coévolution de la culture et du contexte influence les dispositions sociales face à l'évaluation, autant que nos conceptions personnelles de celle-ci. Si nous insistons sur ce point, c'est entre autres parce que les débats actuels entourant l'évaluation scolaire dans un contexte « post pandémique » continuent de démontrer deux choses : 1) une méconnaissance sociale et politique des mécanismes culturels qui fondent les pratiques évaluatives et 2) la polarisation chronique des débats autour d'expériences personnelles évaluatives valorisantes ou traumatisantes par les individus. Ces deux tendances freinent la possibilité d'ouvrir un débat social et politique pacifié sur les questions d'évaluation.

V. Et pourtant... l'évaluation évolue

En second lieu, nous avons tenté de comprendre en quoi les mouvements qui animent l'évaluation constituent des ruptures culturelles praxéologiques, axiologiques et épistémologiques par rapport à ce modèle culturellement dominant.

Pour revenir à la proposition de Vial (2012), les découvertes des dernières décennies semblent avoir contribué à créer une tension récurrente et un équilibre incertain dû au côtoiement de deux fonctions à la fois antagonistes et complémentaires de l'évaluation : une fonction de sanction ou de reconnaissance (généralement associée à une « logique de contrôle » et liée aux fonctions sommative et certificative de l'évaluation) et une fonction d'aide à l'apprentissage (davantage associée à une « logique du

Reste »). Cette tension se manifeste dans un contexte temporel long, marqué par une prédominance historique de l'évaluation-mesure scientifiquement et socialement valorisée et ayant percolé de manière prégnante dans le monde de l'éducation.

C'est sur ce fond de logique de contrôle, exacerbée par l'esprit quantophrénique, qu'ont émergé un ensemble de nouvelles perspectives en évaluation, telles que celle de l'évaluation formative dans les années 1970, des techniques d'évaluation critériée dans les années 1980 et 1990, l'émergence d'objet d'apprentissage et d'évaluation complexes et du jugement professionnel en évaluation dans les politiques scolaires au cours des années 2000. On peut leur ajouter plus récemment des mécanismes de médiation et de relation différentielle avec les référentiels d'évaluation inscrits au sein même de processus évaluatifs dont la plus-value pédagogique devient centrale (Mottier Lopez, 2021).

L'évaluation scolaire comme objet de recherches n'a donc pas cessé d'évoluer. Ces tensions sont connues, décrites, analysées et comprises de manière récurrente depuis des décennies ; elles polarisent les opinions autour de questions mettant en évidence deux extrêmes. D'une part, la nécessité de l'évaluation garante de la reconnaissance du mérite, de la progression scolaire comme marque sociale symbolique de l'éducabilité de toute personne, valeur fondamentale que les systèmes éducatifs dans une diversité de contextes doivent inlassablement défendre et promouvoir. Et d'autre part, la réalité de l'évaluation comme outil stigmatisant symboliquement l'incapacité d'une société de reconnaître à toute personne une valeur à laquelle elle a fondamentalement droit, quelle que soit cette valeur et les manifestations de sa reconnaissance. La viabilité de ce processus paradoxal profondément bivalent, reste hanté par des intentions profondément non-démocratiques, dans la mesure où - si l'évaluation reconnaît réussite, succès, progression, persévérance - elle produit aussi malheureusement des effets collatéraux tels que l'exclusion, la marginalisation, la sélection, le décrochage et l'abandon scolaire. Dans l'évolution de la recherche en évaluation, nous proposons de reconnaître ce que Morrisette (2021) nommerait des signes d'un « redressement démocratique » de l'évaluation, celle-ci étant davantage pensée dans une perspective « resocialisante » (Morrisette, 2009).

VI. À la recherche de la valeur pédagogique de l'évaluation

Il semble qu'un premier mouvement de resocialisation de l'évaluation ait émergé par la valorisation de sa finalité pédagogique, il y a de cela plusieurs décennies. Cette valorisation n'a été possible qu'à partir du moment où nous avons accepté de douter du fait que les actes évaluatifs en contexte d'enseignement-apprentissage ne seraient que la transposition à un niveau micro de démarches évaluatives fondées sur des principes docimologiques, comme le mentionnait Allal en 2012. L'émergence du concept d'évaluation formative dans le contexte d'évaluation de programmes est transposée à l'évaluation des apprentissages dans les années 1960-70 ; dès le début des années 1980 l'évaluation formative s'inscrit dans une théorisation sous l'impulsion de Allal, sa fonction régulatrice des apprentissages étant mise en valeur (Allal et Mottier Lopez, 2007), ce qui fera d'elle un objet de prescription dans les premiers programmes par compétences établis dans les années 2000, notamment au Québec. Par la suite, les travaux de Black et Wiliam (2009) ont été déterminants pour la dissémination de ce concept dans une perspective pragmatique, grâce à la mise à jour d'une implantation à grande échelle et de production de travaux de recherche largement diffusés.

Cette valorisation est concomitante à l'affirmation d'une incontournable subjectivité en évaluation, déjà évoquée par certains chercheurs (Cardinet, 1987 ; Gerard, 2002 ; Hadji, 1992). Cette remise en cause a permis d'ouvrir de nombreuses voies jusque-là non explorées et qui continuent de nourrir nos recherches. En voici un aperçu non exhaustif.

La valorisation du sujet évaluant comme partie prenante d'un processus de validation négociée de savoirs produits par plusieurs personnes évaluatrices sur un même objet (Morrissette et al., 2012) ; la réflexion sur les référents et les référentiels et sur le rapport qu'entretiennent les personnes évaluatrices à ces référentiels dans une perspective de micro-culture de classe, de justice et de justesse du jugement évaluatif et de ses modalités (Mottier Lopez, 2021) ; la valorisation du sujet évalué comme être singulier dans le processus évaluatif avec les recherches sur l'évaluation durable (Boud et Soler, 2016) sur la validité de conséquences ou validité sociale, sur l'auto-évaluation et toutes les variations de la co-évaluation, sont autant de tendances manifestant une volonté sociale de faire contrepoids à un système qui génère un nombre croissant de problématiques révélant les limites de la viabilité d'une perspective

quantophrénique appliquée à toutes les sphères de l'évaluation. En 2018, le Conseil supérieur de l'éducation du Québec publiait un rapport intitulé *Évaluer pour que ça compte vraiment*. Il y dénonçait les effets chroniques d'une prédominance de l'évaluation centrée directement ou indirectement sur le classement des apprenants de l'école primaire à l'Université, notamment le développement d'un rapport négatif et potentiellement toxique à l'évaluation. Ce rapport négatif limite chez les personnes enseignantes autant que chez les apprenants la percolation des découvertes de plusieurs décennies de recherche en évaluation, notamment celle qui porte sur sa fonction pédagogique.

Le questionnement sur l'objet d'évaluation, sur le processus lui-même et sur son produit font partie de cette valorisation de la fonction pédagogique de l'évaluation. La complexification des objets d'apprentissage, notamment avec l'apparition des *curricula* par compétences, entraîne l'émergence de techniques de recueil et d'interprétation plus complexes elles aussi, telle que la technique des grilles critériées, qui reste pour beaucoup d'enseignants un terrain encore largement inexploré. Dans cette lignée innovatrice, l'émergence du concept de jugement professionnel en évaluation ouvre, dès les années 2000, une brèche dans le caractère nécessairement objectif de l'évaluation. La production de connaissances sur la performance d'un élève reposant sur l'expression d'un jugement encadré par des critères portant sur des facettes d'une production complexe et des échelles permettant de situer la production selon des degrés variables d'adéquation avec un « attendu » idéal (Scallon, 2004). La valeur d'une production n'est plus « intrinsèque » et fondée sur une vérité binaire : elle devient relative et fondée sur une interprétation multidimensionnelle dans laquelle une délibération professionnelle est requise.

En soi, cette évolution historique constitue sur un temps long un continuum de ruptures culturelles dans le contexte que nous avons décrit plus haut. Bien que sa portée pragmatique soit reconnue, elle peine encore à jouer un rôle dans l'imaginaire social lié à l'évaluation.

VIII. L'urgence de l'équité et ses limites : une rupture culturelle axiologique

L'évaluation scolaire dans ses formes actuelles est interpellée par la question cruciale de l'inclusion d'une diversité d'élèves dans un chronosystème qui n'a pas été pensé pour eux, loin s'en faut. La thématique de l'inclusion donne à penser que la resocialisation de l'évaluation dont nous parlons implique un glissement radical de la question de l'égalité vers celle de l'équité.

Un des fondements éthiques de l'équité dans un système démocratique est de permettre à des individus d'exister comme personne - et comme personne unique - dans un vivre ensemble qui fait plus que tolérer les différences : il les inclut, les promeut, les défend et les valorise. Ainsi, dans une perspective dynamique et individuelle, les personnes enseignantes comprennent que les élèves éprouvant des enjeux d'inclusion scolaire, ont besoin d'une évaluation bienveillante dans laquelle leur valeur est reconnue. Mais cette réalité est en rupture avec un des principes fondamentaux de la validité de l'évaluation qui présuppose l'existence d'une égalité de conditions pour tous les évalués. L'uniformité de l'espace-temps dans laquelle doivent être placés les individus pour que la décision évaluative qui attribue une valeur chiffrable à leur performance soit considérée comme légitime et crédible est donc remise en cause. Les accommodements appliqués en fonction des principes largement admis d'équité et d'inclusion en évaluation visent essentiellement les conditions spatio-temporelles de l'évaluation par le biais de divers accommodements : par exemple, octroyer plus de temps à un élève pour passer un examen et attribuer des locaux réservés pour des élèves stressés ou ayant des troubles neurocognitifs.

On commence aussi à considérer que dans la poursuite d'un idéal d'égalité à tout crin, l'évaluation devient une source potentielle de souffrances et de traumatismes (Conseil supérieur de l'éducation, 2018 ; Prokofieva et al., 2017). La question de l'anxiété de performance et du stress des élèves face à l'évaluation est bien documentée et on doit continuer de s'interroger sur la valeur des mécanismes par lesquels l'appréciation des progrès et la réussite individuelle demeurent assujetties à des conditions psychologiques reconnues comme contre-productives, déterminant un rapport toxique à l'évaluation elle-même. La brèche ouverte par la rupture contextuelle spatio-temporelle de la pandémie semble avoir à cet égard révélé les limites du chronosystème de l'évaluation, imposant le report

d'épreuves nationales, voire leur annulation. Elle a aussi forcé une réflexion sur les garanties d'équité pour les élèves les plus vulnérables ainsi que des aménagements relatifs aux conséquences de l'évaluation sur le parcours des élèves. En l'absence d'épreuves nationales, il a fallu diversifier les indicateurs de réussite, et abandonner temporairement toute visée comparatiste dans l'analyse des résultats, tout en autorisant la mise en place d'expériences alternatives (Yerly et Laveault, 2020). La différenciation et l'inclusion ont acquis un caractère d'urgence avec la rupture pandémique et le questionnement sur le maintien de conditions spatio-temporelles uniformes pour tous les apprenants pour une évaluation valide et fiable aux échelons nationaux ne fait que s'exacerber dans le contexte post pandémique.

Dans cette perspective, une resocialisation de l'évaluation devrait tenir compte non seulement des élèves, mais de chaque élève, celui-ci ayant maintenant une voix qui lui est propre (Bain, 2010). L'introduction d'une évaluation bienveillante, ou du « bien évaluer » (Hadj, 2021), constitue certainement un antidote potentiel permettant de donner sens à des gestes évaluatifs prescrits dans un cadre quantophrénique incompatible avec les besoins de reconnaissance de l'individu.

IX. L'enjeu du rapport à la vérité en évaluation : une rupture culturelle épistémologique

Un des arguments fréquemment mis en avant dans les débats publics sur l'évaluation est celui de la valeur de « vérité » absolue que l'on peut attribuer aux chiffres en évaluation. La production de cette vérité est un enjeu essentiel lorsqu'on pense à une resocialisation de l'évaluation dans une perspective démocratique. En effet, une autre caractéristique de la démocratie est qu'elle permet aux individus et aux groupes d'entretenir un lien souple avec la vérité (Boudou, 2015 ; Dupont, 2022).

En soi, l'évaluation serait une démarche épistémologique fondée sur l'exercice du jugement, puisqu'elle crée de la connaissance à partir de l'interprétation de traces diverses en vue d'une prise de décision et d'une communication. Un des enjeux majeurs de cette démarche nous semble être celui de la gestion des inférences nécessaires à l'exercice de ce jugement et à la prise de décision. Ces inférences créent un espace dans lequel des référents expérientiels, théoriques, normés ou pas sont mobilisés, faisant de la personne évaluatrice un épistémologue « en acte »

au sens où l'indiquent Tourmen et Mayen (2012). Les évaluateurs mobilisent un « ensemble de représentations opératoires non nécessairement conscientes ou verbalisées qui guident leur rapport à la certitude, aux objets évalués et à leur propre processus psychique de jugement » (p. 74).

En contexte quantophrénique, la production de « résultats » sous forme quantitative prédomine. L'évaluation produit une information chiffrée relative à un niveau de performance, à partir de traces recueillies dans des conditions comparables, considérées comme fiables du seul fait de l'homogénéité de ces conditions. Or, selon Desrosières (2008), la quantification des objets ne se limite pas à en objectiver les qualités, mais aussi à les construire, cette construction répondant à des fondements idéologiques et des processus sociaux auxquels il semble toujours difficile de résister. Par exemple, malgré l'évolution des productions scolaires ancrées dans des approches par compétences, on continue d'observer une tendance tenace à « atomiser » les objets d'évaluation puisqu'il est - épistémologiquement parlant - beaucoup plus simple de produire un résultat de jugement à partir de mécanismes comptables de sommation et de soustraction ou à l'aide de traitements statistiques de base tels que le calcul d'une moyenne ou d'un écart type. C'est pourquoi probablement, un questionnement récurrent est posé face à l'évaluation critériée mettant en jeu un processus inférentiel complexe, qui faute de « traçage » adéquat demeure obscur, faiblement légitime et donc - en contexte quantophrénique - toujours épistémologiquement douteux.

En contexte pandémique, la valeur de cette vérité chiffrée a bien souvent dû être reconnue comme relative, voire caduque, les conditions de validité n'ayant généralement pas été réunies, ou l'ayant été à coup de processus mathématiques et statistiques aléatoires, autant au niveau micro par la personne enseignante qu'au niveau macro au niveau politique et institutionnel. Par contre, l'imputabilité liée à l'expression d'un jugement vrai, en dehors de l'espace clos du vrai ou faux, de l'exact ou de l'inexact, ramène à la personne enseignante elle-même la responsabilité épistémologique de créer une connaissance valide, crédible, légitime, intelligible et communicable sur ce qu'elle évalue. C'est pourquoi l'expérience pandémique pourrait avoir ouvert un espace de resocialisation épistémologique de l'évaluation en favorisant une évaluation de proximité produite dans la transparence, du moins dans certains contextes scolaires. L'évaluation au service de l'apprentissage, au bénéfice d'un élève, ne peut fonctionner que sur base d'une transparence des critères, leur co-définition avec l'enseignant, la mise en relativité du

réfèrent que la délibération convoque, le partage de ces référents, leur évolution, la médiation dans l'exercice du jugement, l'*empowerment* de l'apprenant qui gère une évaluation durable dont il pilote le processus, se l'approprie, devenant l'évaluateur (Nizet et., al, 2023). Ce redressement favorise aussi une évaluation qui fait sens pour l'élève : crédible pour que la délibération et donc la production du verdict, l'énonciation de la proximité ou la distance avec l'attendu soit la plus juste possible, la plus contextualisée possible. On doit alors considérer que le jugement évaluatif produit une vérité pédagogique fondée sur un travail de référentialisation dont le caractère collaboratif est de plus en plus valorisé (Mottier Lopez, 2023).

X. Invitation à une certaine émancipation culturelle

Jusqu'à présent mon propos a tenté de démontrer que sur un fond d'histoire commune, nos systèmes d'évaluation scolaires sont ancrés dans une vision relativement monolithique, tout en constatant que l'émergence des sciences de l'éducation a produit un florilège de mécanisme de rééquilibration de cette tendance dont plusieurs peinent encore à percoler dans une conscience sociopolitique. Depuis des décennies, des ruptures culturelles avec cette perspective se manifestent par l'émergence de problématiques sociales chroniques comme le décrochage scolaire, la nécessité de l'inclusion et de l'équité et plus récemment la question de l'anxiété scolaire.

La rupture pandémique a exacerbé ces réalités, du fait de l'inadaptation temporaire d'un chronosystème dont la légitimité de l'évaluation scolaire dépend. Beaucoup d'acteurs de l'éducation y ont été confrontés (Yerly et Laveault, 2020), puisant dans leur expérience et leur créativité pour mettre en application une multitude d'alternatives bien souvent nourries des découvertes de plusieurs décennies de recherche en évaluation (Béland et al., 2020 ; Borges et Tardif, 2020 ; Detroz et al., 2020 ; Issaieva et al. 2020 ; Yerly et Issaieva, 2021). Par exemple, la valorisation de l'évaluation critériée et de l'évaluation pour l'apprentissage et comme apprentissage, les tâches complexes comme solution au plagiat en ligne, l'importance de l'évaluation collaborative, la nécessité de l'alignement pédagogique, la diversification des preuves d'apprentissage, la triangulation, l'autoévaluation, l'évaluation par les pairs, la différenciation, l'évaluation dynamique et personnalisée, voire même l'abandon de la notation, voici

autant de pistes qui résultent de la recherche, qui ont été théorisées, explicitées, mais apparaissent encore trop souvent comme périphériques et insuffisamment légitimées socialement pour entraîner une refonte structurelle des socio et chronosystèmes de l'évaluation.

En corollaire de ces approches, on a vu se développer des formes scolaires alternatives : le temps scolaire se déplaçant en fonction de la disponibilité des enseignants ou disparaissant temporairement ; l'espace scolaire s'est virtualisé et individualisé remettant ainsi en cause une forme scolaire séculaire (Ria et Rayou, 2020).

L'évaluation en ligne, pas nouvelle elle non plus, devient objet de toutes les attentions, incontournable nécessité. On la pare pêle-mêle de tous les attraits et de tous les défis : l'inclusion, l'automatisation, la surveillance avec les mécanismes de *proctoring*, la régulation formative, l'équité, la persévérance. Mais sa sujétion aux contraintes d'un chronosystème dont elle pourrait s'émanciper, empêche d'en faire un véritable levier dans un certain nombre de cas. Elle ne peut à elle seule réduire l'anxiété des élèves ou la crainte du plagiat chez les enseignants.

Si nos propos ont permis de rendre compte des efforts consentis pour maintenir les conditions de viabilité des chrono et macrosystèmes de l'évaluation en temps de pandémie, ils tentent aussi de rendre compte des gestes d'émancipation dont se sont prévalus certains acteurs au niveau de leurs méso, micro et onto-systèmes en évaluation.

Bref, les ruptures contextuelles sont-elles réellement porteuses de ruptures culturelles durables en évaluation ? Nous pensons que oui, et nous pensons aussi que des ruptures culturelles durables seront nécessaires et qu'elles proviendront très certainement de microsystèmes émancipés et collectivement actifs pour autant que la liberté leur en soit donnée.

XI. Bibliographie

Allal, L. (2012). Les exigences inconciliables des activités évaluatives et leurs paradigmes de référence. Dans L. Mottier Lopez (dir.), *Modélisations de l'évaluation en éducation : Questionnements épistémologiques* (p. 181-194). De Boeck Supérieur. doi.org/10.3917/dbu.lopez.2012.01.0181

- Allal, L. et Mottier Lopez, L. (2007). *Régulation des apprentissages en situation scolaire et en formation*. De Boeck.
- Bain, J. (2010). Integrating student voice: assessment for empowerment. *Practitioner Research in Higher Education*, 4(1), 14-29. https://research.gold.ac.uk/4529/1/Integrating_student_voice-assessment_for_empowerment.pdf
- Béland, S., Bureau, J. S. et Peters, M. (2020). Plagier en temps de pandémie. *Évaluer. Journal international de recherche en éducation et formation*, Numéro Hors-série, 1, 35-40.
- Black, P. et Wiliam, D. (2009). "Developing the theory of formative assessment". *Educational Assessment, Evaluation and Accountability*. 21(1), 5–31. <http://doi:10.1007/s11092-008-9068-5>. S2CID 55955635
- Borges, C. et Tardif, M. (2020). Faire école en temps de pandémie. Formation et profession. *Revue scientifique internationale en éducation*, 28(4), numéro hors-série. <https://formation-profession.org/numeros/view/29>.
- Boud, D. et Soler, R. (2016). Sustainable assessment revisited. *Assessment & Evaluation in Higher Education*, 41(3), 400-413. doi:10.1080/02602938.2015.1018133
- Brown, P. et Duru-Bellat, M. (2010). La méritocratie scolaire : Un modèle de justice à l'épreuve du marché. *Sociologie*, 1(1), 161-175.
- Boudou, B. (2015). *La crise de la vérité comme condition démocratique*. Dans Y.Couture, S.Vibert et M.Chevrier (dir.), *Démocratie et modernité : La pensée politique française contemporaine* [en ligne]. Presses universitaires de Rennes. <http://doi.org/10.4000/books.pur.73121>.
- Bronfenbrenner, U. (1993). Ecological models of human development. Dans M. Gauvain et M. Cole (dir.), *Reading on the development of children* (2e éd., p. 37-43). https://impactofspecialneeds.weebly.com/uploads/3/4/1/9/3419723/ecological_models_of_human_development.pdf
- Cardinet, J. (1987). L'objectivité de l'évaluation. *Recherches, Institut romand de recherches et de documentations pédagogiques*. https://www.irdp.ch/data/documents/manifestations/cardinet/objectivite_evaluation.pdf
- Conseil supérieur de l'éducation (2018). *Évaluer pour que ça compte vraiment, Rapport sur l'état et les besoins de l'éducation 2016-2018*. Le Conseil. <https://www.cse.gouv.qc.ca/wp-content/uploads/2020/01/50-0508-RF-evaluer-compte-vraiment-REBE-16-18.pdf>
- Desrosières, A. (2008). La statistique outil de gouvernement et outil de preuve. Dans A. Desrosières (dir.), *L'argument statistique. Tome 1*,

- pour une sociologie historique de la quantification* (p. 7-20). Presses de l'École des Mines.
- Detroz, P., Tessaro, W., Younès, N. (2020). Évaluer en temps de pandémie. *Journal international de recherche en éducation et formation*, Numéro Hors-série. URL : <https://journal.admee.org/index.php/ejiref/issue/view/21>.
- Dupont M. (2022, 8 juillet). La vérité, pilier de la démocratie. *Le Monde*, 3. <https://mediascitoyens-diois.info/2022/07/la-verite-pilier-ebroule-de-la-democratie/>
- Frenette É. et Hébert, M.-H. (2021). L'évaluation d'aujourd'hui et du futur, Garante de son passé ? Dans C. Barosso da Costa, I. Nizet et D. Leduc (dir.), *40 ans de mesure et d'évaluation* (p. 31-51). Presses de l'Université du Québec.
- Gerard, F.-M. (2002). *L'indispensable subjectivité de l'évaluation*. Antipodes, 156, 26-34.
- Hadji, C. (1992). *L'évaluation des actions éducatives*. Presses Universitaires de France.
- Hadji, C. (2021). *Le défi d'une évaluation à visage humain. Dépasser les limites de la société de performance*. ESF. Sciences humaines.
- Issaieva, E., Odacre, E., Lollia, M. et Joseph-Theodore, M. (2020). Enseigner et apprendre en situation de pandémie : caractéristiques et effets sur les enseignants et les élèves. *Formation et profession*, 28(4), 1-12. <https://doi.org/10.18162/fp.2020.702>
- Maulini, O. (1996). *Qui a eu cette idée folle un jour d'inventer [les notes à l'école ? Petite histoire de l'évaluation chiffrée à l'usage de celles et ceux qui désirent s'en passer (et des autres)]*. <https://www.unige.ch/fapse/SSE/teachers/maulini/publ-9601.pdf>
- Maulini, O. (2017). Mesurer ou classer ? Histoire et avenir de la notation, chiffrée ou non. Dans M. Neumayer (dir.), *Évaluer sans noter : éduquer sans exclure* (p. 22-32). Chronique sociale.
- Morrisette, J. (2009). Une re-socialisation de l'évaluation. *Revue canadienne des jeunes chercheurs en éducation/Canadian journal for New Scholars in Education*, 1(2).
- Morrisette, J. (2021). Un pluralisme disciplinaire et méthodologique comme perspective d'avenir du domaine. Dans C. Barosso da Costa, I. Nizet et D. Leduc (dir.), *40 ans de mesure et d'évaluation*, (p. 53-75). Presses de l'Université du Québec.
- Morrisette, J., Mottier Lopez, L. et Tessaro, W. (2012). La production de savoirs négociés dans deux recherches collaboratives sur l'évaluation formative. Dans L. Mottier Lopez et G. Figari (dir.), *Modélisations de l'évaluation en éducation : Questionnements épistémologiques*

- (p. 27-43). De Boeck Supérieur. <https://doi.org/10.3917/dbu.lopez.2012.01.0027>
- Mottier Lopez, L. (2021). Évaluation située élargie : objectivation à partir de l'évaluation formative élargie, l'apprentissage situé et la référentialisation. Dans C. Barroso da Costa, I. Nizet et D. Leduc (dir.), *40 ans de mesure et d'évaluation* (p. 119-138). Presses Universitaires du Québec.
- Mottier Lopez, L. (2023). Rapports différentiels dans la référentialisation de l'évaluation : des expériences démocratiques ? Essai théorique. *La Revue LEeE*, 7. <https://doi.org/10.48325/rleee.007.01>
- Nizet, I. (2015). Les pratiques évaluatives au quotidien : enjeux culturels et identitaires. Dans J.-L. Leroux (dir.), *Évaluer les compétences au collégial et à l'université*, (p. 111-134). Chenelière éducation.
- Nizet, I., Lépine, M., Denis, I. et Hétu, S. (2023). Oser une évaluation partagée : étude de cas dans une école québécoise. *La Revue LEeE*, 7. <https://doi.org/10.48325/rleee.007.05>
- Ofir, Z. (2018). *Evaluation for transformation*. <https://zendaofir.com/the-nice-framework-part-1/>
- Prokofieva, V., Brandt-Pomares, P., Velay, J.-L., Hérold, J.-F. et Kostromina, S. (2017). Stress de l'évaluation scolaire : un nouveau regard sur un problème ancien, *Recherches & éducations*, 18. <https://doi.org/10.4000/rechercheseducations.4657>
- Ria, L. et Rayou, P. (2020). La forme scolaire en confinement : enseignants et parents à l'épreuve de l'enseignement à distance, *Formation et profession*, 28(4 hors-série), 1-11. <http://dx.doi.org/10.18162/fp.2020.675>
- Scallon, G. (2004). *L'évaluation des apprentissages dans une approche par compétences*. Renouveau pédagogique.
- Sorokin, P. (1959), *Tendances et déboires de la sociologie américaine* (traduit de l'Américain par C. Arnavon). Aubier.
- Tourmen, C. et Mayen, P. (2012). Les évaluateurs, des épistémologues en actes ? Dans L. Mottier-Lopez et G. Figari (dir.) *Modélisations de l'évaluation en éducation : Questionnements épistémologiques* (p. 63-77). De Boeck Supérieur. <https://doi.org/10.3917/dbu.lopez.2012.01.0027>
- Vial, M. (2009). *Se former pour évaluer. Se donner une problématique et élaborer des concepts*. De Boeck Université.
- Vial, M. (2012). *Se repérer dans les modèles de l'évaluation. Méthodes, dispositifs, outils*. De Boeck.
- Vincent, G., Courtebras, B. et Reuter, Y. (2012). La forme scolaire : débats et mises au point : Entretien de Guy Vincent avec Bernard Courtebras et Yves Reuter. *Recherches en didactiques*, 13, 109-135. <https://doi.org/10.3917/rdid.013.0109>

- Yerly, G. et Issaieva, E. (2021). (Re)penser l'évaluation des apprentissages au postsecondaire en temps de crise : défis à relever et occasions à saisir en période de COVID-19. *Revue internationale des technologies en pédagogie universitaire*, 18(1), 89-101. <https://doi.org/10.18162/ritpu-2021-v18n1-09>.
- Yerly, G. et Laveault, D. (2020). Évaluer les apprentissages en contexte de pandémie : aller au-delà de la notation pour soutenir la réussite de tous les élèves. *Formation et profession*, 28(4), 1-12. <https://doi.org/10.18162/fp.2020.676>

Chapitre 6. Expérimenter une évaluation collaborative en formation à l'enseignement. Une interface entre les contextes évaluatifs ?

Gonzague Yerly

I. Introduction

Le présent chapitre⁷ vise à contribuer aux questionnements de l'axe 2 présentés dans l'introduction de cet ouvrage en interrogeant les interactions entre les différents contextes évaluatifs dans le cadre de la formation des enseignants⁸ à l'évaluation des apprentissages. Il interroge le potentiel de l'expérimentation par des étudiants d'une « évaluation collaborative » (Allal et Mottier Lopez, 2014) comme interface entre les contextes individuels (celui de l'enseignant), collectifs (celui d'une équipe d'enseignants) et institutionnels (celui d'un établissement, d'une juridiction scolaire).

L'étude porte sur l'analyse du vécu d'étudiants dans un cours portant sur l'évaluation des apprentissages à l'Université de Fribourg (Suisse), dans un programme de formation à l'enseignement au secondaire post-

⁷ Cette contribution est issue d'une communication orale présentée dans le cadre du symposium « Formation professionnalisante à l'évaluation des apprentissages en questions : analyse de dispositifs interrogés à la lumière de différents contextes », coordonné par Lucie Mottier Lopez et Gonzague Yerly. Les autres contributeurs-trices étaient, par ordre de passage : Françoise Pasche Gossin et Jean-Claude Desboeufs (HEP Béjune), Isabelle Monnard (HEP FR), Céline Girardet (UniGE), Yann Mercier-Brunel (U. Orléans), Isabelle Nizet (U. Sherbrooke) et Marie-Hélène Hébert (Teluq).

⁸ Dans ce texte, le masculin est utilisé comme générique.

obligatoire (élèves de 15 à 19 ans). Lors de ce cours, les étudiants avaient à créer et à analyser, en équipe, un outil ou une démarche d'évaluation des apprentissages des élèves. Pour ce projet, ils ont également pu expérimenter le feedback entre pairs. L'analyse porte sur les perceptions des étudiants quant à leur vécu lors de ce projet collaboratif. En quoi est-il pertinent de faire vivre une expérience d'évaluation collaborative aux étudiants en formation à l'enseignement ? Un transfert existe-t-il entre leur expérience en tant que formés et leur pratique professionnelle future ?

Le texte présente tout d'abord une problématisation et un cadrage théorique portant sur les différents éléments conceptuels mobilisés. Par la suite, le cours servant de contexte à l'expérience est décrit. Enfin, les sections suivantes présentent la méthodologie, les résultats et une discussion conclusive.

II. Problématisation et cadrage conceptuel

Le travail collectif des enseignants pour l'évaluation des apprentissages

Les autorités scolaires comptent aujourd'hui fortement sur le travail collectif des enseignants comme levier de développement des pratiques évaluatives, avec des conséquences diverses (Laveault et Yerly, 2017 ; Yerly, 2021). Le travail collectif peut prendre des formes bien différentes, dont vont dépendre les effets sur les pratiques (par ex. collaboration, coopération, collégialité contrainte, etc.) (Marcel et al., 2007 ; Vangrieken et al., 2015). Parmi les pratiques d'« évaluation collaborative » (Allal et Mottier Lopez, 2014), qui touchent la collaboration entre élèves mais aussi entre professionnels, les dispositifs de « modération sociale » visent à engager collectivement les enseignants dans la construction d'une compréhension commune des standards institutionnels et à atteindre des consensus dans les pratiques évaluatives (Adie et al., 2012 ; Laveault et Yerly, 2017 ; Mottier Lopez et al., 2012). La modération sociale porte donc sur les résultats d'évaluation mais peut également toucher d'autres étapes du processus évaluatif, dont l'élaboration d'outils ou de démarches évaluatives (Girardet et Mottier Lopez, 2023).

Une « évaluation collaborative » (Allal et Mottier Lopez, 2014), aussi bien entre les élèves qu'entre enseignants, est un processus complexe qui demande de trouver des consensus entre enseignants. Comme toute collaboration, il nécessite des « négociations » (Germier et Marcel, 2016) qui dépendent de facteurs très divers : historiques, culturels, épistémologiques, contextuels, politiques, etc. Les échanges entre

enseignants ne sont pas toujours équilibrés ou mutuellement bénéfiques. Les enseignants doivent faire cohabiter les enjeux, les croyances et les pratiques des contextes évaluatifs individuels (celui de l'enseignant), collectif (celui de l'équipe d'une discipline scolaire) et institutionnel (celui d'un établissement, d'une juridiction scolaire).

L'isomorphisme comme stratégie de formation

Surtout au secondaire, où les pratiques évaluatives du terrain restent encore plutôt individuelles et « traditionnelles » (Goffin et Fagnant, 2017), implanter une évaluation davantage collaborative représente un important défi pour les gestionnaires des systèmes et pour les enseignants en fonction. Les défis existent aussi pour la formation des futurs enseignants au sein des universités et hautes écoles, où l'évaluation dite « traditionnelle » prédomine également (Allal, 2013 ; Romainville, 2013). En formation à l'enseignement, et notamment dans le domaine de l'évaluation des apprentissages, les représentations des étudiants évoluent durant la formation. Toutefois, il a été démontré que les représentations initiales réapparaissent souvent à l'entrée sur le terrain (Luisoni et Monnard, 2015). Afin de contrer ce phénomène, pour rendre l'apprentissage plus durable, davantage en profondeur (Fischer et al., 2020), une stratégie clé est de faire vivre aux apprenants des situations d'évaluation qui sont en adéquation avec ce que l'on veut leur faire apprendre. Cette modalité augmente la probabilité que les étudiants transfèrent les pratiques souhaitées dans leurs classes. Il s'agit d'un principe d'isomorphisme ou d'homologie (Goffin et Fagnant, 2017). L'évaluation des apprentissages constitue donc non seulement un objet de savoirs, mais également un moyen pour apprendre et pour faire apprendre de manière durable (Mottier Lopez, 2021).

Le feedback par les pairs

Le feedback par les pairs est une manière d'impliquer les étudiants dans le processus évaluatif. Il est en outre un moyen de développer les compétences de collaboration (Girardet, 2020). L'effet positif de certains types de feedbacks, lorsqu'ils suivent certaines conditions, a été largement démontré (Hattie, 2017). Surtout, le feedback doit être compris, accepté et utilisé par son destinataire (Carless et Boud, 2018). En formation à l'enseignement, au niveau universitaire où les effectifs sont souvent importants, le feedback par les pairs représente une stratégie qui permet un feedback régulier, immédiat et individualisé (Girardet, 2020). Le

feedback entre pairs permet également de faire vivre le feedback aux étudiants, à la fois dans le rôle de celui qui élabore et communique le feedback, mais aussi de celui qui le reçoit. Le feedback par les pairs donne une grande part de responsabilité aux apprenants dans le processus, leur permet de s'approprier les attentes et représente des occasions de les expérimenter. Selon Sadler (2010), cité par Girardet (2020), « trois éléments doivent être travaillés avec les étudiants afin qu'ils puissent réellement bénéficier de l'évaluation entre pairs : la compréhension de la tâche, la compréhension de ce qu'est une évaluation de qualité, et la compréhension des critères d'évaluation » (p. 2).

III. Contexte de l'étude

Description du cours

L'étude porte sur un dispositif de formation à l'évaluation des apprentissages, dont l'auteur était responsable à l'Université de Fribourg (Suisse)⁹. Ce cours s'inscrivait dans un programme de formation de futurs enseignants du secondaire post-obligatoire, le Diplôme d'enseignement pour les écoles de maturité (DEEM) de l'Université de Fribourg (section francophone) qui se déroule sur un an, en parallèle entre stage sur le terrain et cours à l'Université sur l'ensemble de l'année scolaire. Les étudiants se destinent à l'enseignement d'une ou, plus majoritairement, de deux disciplines scolaires pour des élèves de 15 à 20 ans. Il s'agit d'un programme de formation post-Master, les étudiants ayant obtenu préalablement un Master dans une ou deux disciplines enseignables. Le cours qui est l'objet de l'analyse était intitulé « Evaluation des apprentissages au secondaire II », il se déroulait au milieu de cette année de formation (janvier-mars). Ce cours était doté de 2 ECTS, réparties entre 5 séances de cours de 4 heures chacune en présentiel, de travail et de lectures personnelles et d'un projet collaboratif. Le cours regroupait en moyenne une cinquantaine d'étudiants aux profils particulièrement hétérogènes, jusqu'à 21 disciplines scolaires pouvaient être représentées. Ce cours visait principalement à faire développer des compétences professionnelles en évaluation des apprentissages inscrites dans le plan d'études du DEEM, mais également des compétences plus larges touchant

⁹ La description du cours est rédigée au passé car l'édition du cours en question était la dernière sous cette forme, l'auteur ayant changé d'institution par la suite. Ce cours est mis en œuvre dans son institution actuelle.

notamment l'entrée « praticien réflexif ». Le dispositif répondait ainsi à la nécessité de former les étudiants à acquérir un bagage de connaissances et de premières compétences professionnelles en évaluation des apprentissages, mais également de développer des compétences de collaboration.

Intentions du cours

Le cours a été construit dans l'intention d'orienter les pratiques professionnelles des futurs enseignants vers le paradigme de l'évaluation pour apprendre (De Ketele, 2012), mobilisant plus particulièrement le concept de l'Évaluation-soutien d'apprentissage (Laveault et Allal, 2016). Avec l'intention de toucher en profondeur les croyances et les compétences des enseignants de manière durable (Mottier Lopez, 2021), le dispositif a été conçu afin de faire vivre les principales caractéristiques de l'EsA tout au long du cours. Ce dispositif de formation a été expérimenté sous cette forme pendant quatre années (2018-2022) par l'auteur.

Les étudiants étaient impliqués tout au long du cours dans le processus d'évaluation qui allait servir à leur progression mais également à en faire le bilan certificatif. Parmi d'autres, la tâche principale des étudiants (projet collaboratif) était d'élaborer en équipe et tout au long du cours un outil ou une démarche d'évaluation inscrite dans l'une de leurs disciplines scolaires (ou interdisciplinaire) et en parallèle d'en rédiger l'analyse¹⁰ qui en montre le bien-fondé, sur la base d'éléments conceptuels travaillés en cours. C'est d'ailleurs uniquement la qualité de l'analyse qui était évaluée dans une visée certificative par le formateur. Les étudiants étaient impliqués dans le processus évaluatif en étant notamment amenés à coconstruire la grille d'évaluation du projet avec le professeur (les critères étaient présentés dès le début du cours par le professeur) et à apporter des feedbacks aux collègues de leur équipe. La constitution des équipes était laissée libre aux étudiants, avec un nombre indicatif de 3 à 4 étudiants. Le choix des étudiants se portait habituellement sur la constitution de groupes par rapport à une discipline scolaire. L'analyse d'outils et de démarches d'évaluation étaient travaillées en cours.

En conclusion, ce dispositif entendait faire saisir aux étudiants l'importance de placer l'apprenant comme acteur principal du processus d'évaluation afin de lui permettre de développer à son tour des

¹⁰ Au niveau taxonomique (Anderson et Krathwohl, 2002) on pourrait associer ce projet aux niveaux cognitifs « analyser, évaluer et créer ».

compétences d'(auto)évaluation et non comme un simple consommateur ou collaborateur (*assessment as learning*) (Earl, 2013). L'intention était *in fine* que les futurs enseignants, une fois sur le terrain, fassent vivre, à leur tour, de telles démarches à leurs élèves, les impliquent dans le processus évaluatif et ainsi les rendent (aussi) maîtres de l'évaluation des apprentissages.

IV. Objectifs de l'étude

Comme décrit dans la section précédente, l'une des intentions du cours « Evaluation des apprentissages au secondaire II » était de faire vivre aux étudiants, futurs enseignants, une évaluation collaborative. Notre étude vise à analyser l'expérience des étudiants afin d'interroger le potentiel d'une telle activité comme interface entre les contextes évaluatifs individuels, collectifs et institutionnels. Les objectifs spécifiques sont d'analyser :

- a) leur vécu d'une évaluation collaborative ;
- b) leur vécu du feedback entre pairs ;
- c) les perspectives qu'ils envisagent (ou non) pour leur propre enseignement dans les classes, suite à cette expérience.

V. Méthodologie

Approche de recherche

L'approche de cette étude est compréhensive et la méthode qualitative (Miles et Huberman, 2010). De ce fait, les analyses visent à décrire, puis à interpréter les perceptions d'un nombre restreint d'étudiants dans un contexte particulier. Aussi, nous sommes bien conscients du caractère peu sophistiqué de l'étude (une seule prise de données auto-rapportée) et des biais inhérents à la situation de formation (désirabilité sociale).

Participants

Les résultats de notre analyse reposent sur le témoignage des étudiants qui ont participé au cours « Evaluation des apprentissages au secondaire II » à l'Université de Fribourg (Suisse) lors de l'année académique 2021-2022.

Il s'agit des réponses de 55 étudiants (35 femmes, 20 hommes) sur les 56 que comptait le cours. Tous les participants visaient l'obtention du diplôme d'enseignement dans les écoles de maturité (DEEM) et se destinaient à enseigner une ou, majoritairement (41 sur 55), deux disciplines scolaires dans des écoles de maturité gymnasiale, écoles de culture générale ou écoles de formation professionnelle. Cette année-ci, 20 disciplines étaient représentées au total dans le groupe¹¹.

Outil de recueil des données et passation

Le vécu des étudiants a été investigué par un questionnaire en ligne, via la plateforme googleforms. Il comprenait huit questions ouvertes qui touchaient aux objets de notre étude (collaboration, feedback par les pairs, perspectives professionnelles) :

- Q1 - De manière générale, comment avez-vous vécu personnellement le fait de devoir travailler en équipe pour ce travail ?*
- Q2 - Comment s'est passé ce travail « en équipe » ? Selon vous, peut-on le qualifier de véritable « collaboration » ? En quoi ?*
- Q3 - Durant ce travail, il vous a été demandé d'apporter des feedbacks à vos collègues, et donc d'en recevoir. Comment avez-vous vécu ces situations ?*
- Q4 - Et comment les autres personnes de votre groupe ont-elles réagi aux feedbacks ?*
- Q5 - Comment avez-vous procédé pour élaborer et communiquer ces feedbacks ?*
- Q6 - Merci de transmettre ici un exemple de feedback écrit (ou transcrit) que vous avez communiqué dans le cadre de ce travail.*
- Q7 - Finalement, que retenez-vous de cette expérience (collaboration + feedback) pour votre future carrière dans les classes ?*
- Q8 - Éventuellement, autres commentaires.*

Les réponses au questionnaire étaient anonymes. Aucune information personnelle n'était demandée. Le questionnaire a été envoyé et retourné à la

¹¹ Anglais, biologie, chimie, droit, économie, espagnol LE, français, français LE, géographie, histoire, informatique, italien, italien LE, mathématiques, musique, philosophie, physique, psychopédagogie, sciences des religions, éducation physique et sportive.

suite des séances de cours et de la reddition du projet collaboratif, avant le feedback du formateur sur le projet.

Analyse des données

Une analyse catégorielle de contenu (Mucchielli, 2006) a été menée pour traiter les réponses au questionnaire. Les unités de sens ont tout d'abord été classées dans une analyse descriptive avec des catégories construites *a priori*. Il s'agissait de catégories globales (collaboration, feedback, perspectives pour la profession, autres) dont les deux premières ont été détaillées en sous-catégories (conditions, effets, autres). Les unités de sens classées ainsi ont ensuite été analysées pour ressortir des éléments d'interprétations pour chaque catégories et sous-catégories.

VI. Présentation des résultats

Dans cette section, une synthèse des résultats est organisée en fonction des catégories, représentant nos trois objectifs et objets : collaboration, feedback entre pairs et perspectives pour la profession.

Vécu des étudiants d'une évaluation collaborative

De manière majoritaire, les étudiants se sont organisés en suivant la structure proposée par le formateur. Dans un premier temps, chaque participant cherchait un outil ou une démarche d'évaluation préexistante sur le terrain, auprès d'un enseignant-formateur, d'un collègue, etc. Dans un deuxième temps, les membres de l'équipe se regroupaient pour sélectionner l'outil ou la démarche qui allait être analysé et amélioré. Certaines équipes décidaient de créer un nouvel outil ou une nouvelle démarche, en repartant de zéro. Un fonctionnement très majoritaire, voire unique, était ensuite de se séparer la tâche, selon les angles d'analyse (par ex. niveaux taxonomiques ; qualités docimologiques ; qualité du feedback ; implication de l'élève). Ensuite, les étudiants remettaient ensemble les différentes parties et élaboraient une conclusion commune.

Pour certains étudiants, le fait de travailler en équipe a permis de créer une ambiance de travail conviviale et motivante. Si les étudiants ont pour la plupart travaillé avec des personnes qu'ils connaissaient déjà - étant en cours dans la-les même-s didactique-s spécifique-s - le projet collaboratif a

également permis de renforcer les liens ou de créer de nouveaux liens avec des personnes qu'ils ne fréquentaient pas avant l'expérience. De nombreux étudiants ont relevé que la démarche collective a permis d'aller plus loin que s'ils avaient dû travailler seul. Les discussions et débats lors des phases collectives (choix de l'outil, écriture de parties communes) et des feedbacks ont permis, selon eux, de rehausser la qualité de leur travail. Pour certains étudiants, l'expérience a été l'occasion de développer leurs compétences de collaboration, par ex. s'organiser, écouter les autres, oser donner son point de vue, mais aussi trouver des personnes « fiables ».

Selon les témoignages recueillis, l'expérience a également permis aux étudiants de développer une compréhension commune des différents éléments liés à l'évaluation des apprentissages. La confrontation des idées, mais aussi de certaines croyances et valeurs au sein des groupes a permis des débats, des explications entre pairs (parfois une sorte de tutorat). Le travail en équipe a aussi permis de s'apercevoir que parfois certaines positions n'étaient pas identiques. L'expérience a demandé de trouver des consensus, de trouver des solutions acceptables pour tous les membres du groupe.

Les réponses au questionnaire permettent de relever les conditions qui, selon les étudiants interrogés, favorisent ou empêchent une « véritable collaboration » :

- Une structure qui permet/demande de la collaboration. Selon certains étudiants, la structure du travail était un facteur clé pour favoriser une véritable collaboration. En effet, ces étudiants ont réalisé de nombreux travaux de groupe durant leur formation. Toutefois, certains relèvent que le projet en question nécessitait et permettait la collaboration car il comportait une importante « interdépendance positive » (Buchs, 2017). Le projet comportait un but commun qui ne pouvait être atteint que par l'engagement de plusieurs personnes ayant des expertises différentes. En effet, le travail était structuré et présenté selon des étapes communes (choix de l'outil ou de la démarche ; choix des angles d'analyse ; feedbacks ; conclusion commune) et des étapes plus individuelles (recherche de documents ; parties d'analyse spécifique). Aussi, le niveau de complexité demandé - analyser, évaluer, voire créer, selon la taxonomie d'Anderson et Krathwohl (2001) - exige une réflexion qui demande de débattre à plusieurs. Enfin, ce sont également des éléments liés à l'évaluation du travail qui ont renforcé le but commun : le résultat final (la note) était commun à l'ensemble du groupe ; l'un des critères de validation était la cohérence du travail.

- Organisation efficace du groupe. Les étudiants témoignent de l'importance de l'organisation au sein du groupe. La formation DEEM étant très chargée, les agendas des personnes très divers selon leurs stages, cette organisation était, selon eux, des plus importantes. Certains étudiants témoignent d'une planification initiale importante, d'autres de l'usage d'outils numériques collaboratifs pour mener leur projet à bien.
- Compétences personnelles de collaboration et intérêt pour le travail en équipe. Bien que le projet demandé visait à les développer, son bon fonctionnement a reposé également sur les compétences individuelles de collaboration déjà présentes. Les étudiants du DEEM ont en effet l'habitude de fonctionner en équipe pour des travaux universitaires. Toutefois, les compétences n'étaient pas toutes initialement développées chez chacun, et tous les étudiants n'apprécient pas les travaux de groupe.
- Difficultés d'organisation. Le programme de DEEM étant chargé, entre cours et stage dans des établissements divers, certains étudiants ont relevé la difficulté de trouver des temps et des espaces de collaboration. Les outils numériques ont pu être une solution pour se réunir plus facilement.
- Lassitude face aux travaux de groupes. Un point soulevé par certains étudiants est celui de la trop nombreuse récurrence de travaux de groupes dans le cursus de formation, ce qui peut freiner leur enthousiasme à travailler à plusieurs.
- Difficulté de s'adapter aux autres. Selon certains étudiants interrogés, travailler avec des collègues n'est pas si aisé. Il est pour certains coûteux de devoir s'adapter au rythme de travail, aux compétences, à l'investissement, aux attentes des collègues. Aussi, certains étudiants ont relevé que les « visions de l'évaluation » étaient parfois très (trop) différentes pour créer quelque chose de commun.

Vécu des étudiants du feedback par les pairs

Les feedbacks ont surtout été donnés sur les parties individuelles, par ex. les parties spécifiques d'analyse des outils ou démarches d'évaluation. Le fonctionnement majoritaire des groupes, voire unique, était une relecture individuelle des parties des autres membres du groupe, puis une ou des discussions communes pour débattre puis valider les choix. Les feedbacks ont été donnés par oral et/ou par écrit, en présence et/ou à distance, le plus souvent à l'aide d'outils numériques (commentaires dans un document partagé online, messages par téléphone, mails, visioconférence).

De nombreuses réponses montrent l'effort qui a été fait pour donner des feedbacks compréhensibles et émotionnellement acceptables. Quelques étudiants font le lien dans leur réponse avec le cours sur le feedback. Il est à souligner qu'un seul groupe soulève, dans le questionnaire, avoir utilisé la grille d'évaluation avec les critères.

L'expérience des feedbacks entre pairs a été vécue par les 55 étudiants comme positive, voire très positive. De nombreux adjectifs positifs ressortent des réponses : positif, fructueux, constructif, formateur, enrichissant. Il a permis selon eux, une amélioration significative du travail. Le feedback entre pairs a également permis de créer des controverses et des débats sur certains concepts liés à l'évaluation, par ex. sur la question de la note. Ces controverses n'ont pas été toutes porteuses de consensus, mais le vecteur de réflexions formatrices. Une certaine crainte de blesser et d'être blessé par les feedbacks est tout de même soulevée dans différentes réponses, surtout au début de l'expérience vécue.

Les réponses au questionnaire permettent de relever les conditions qui, selon les étudiants interrogés, favorisent l'effet positif des feedbacks par les pairs :

- des feedbacks « constructifs ». La manière d'élaborer et de communiquer un feedback à un pair semble, pour les étudiants interrogés, une condition essentielle pour que le feedback soit pris en compte. Celui-ci doit respecter certains codes, surtout au niveau émotionnel : être franc, mais dans des formes acceptables.
- une certaine habitude de recevoir des feedbacks. Certains étudiants écrivent avoir l'habitude de recevoir des feedbacks, par exemple dans leur stage. Ils disent même en être friand pour pouvoir s'améliorer.
- plus aisé dans une relation déjà établie. Recevoir et donner du feedback dans un groupe constitué de personnes connues auparavant semble un élément facilitateur pour certains étudiants.
- un objectif commun, une validation commune. Le fait d'être engagé dans un projet commun, avec des enjeux communs de validation semble également être un moteur dans l'intérêt et la prise en compte des feedbacks afin de pouvoir améliorer la qualité du travail.

Perspectives que les étudiants envisagent (ou non) pour leur propre enseignement dans les classes, à la suite de cette expérience

Les réponses des étudiants montrent très fortement que l'expérience vécue a renforcé chez eux l'idée déjà partagée que la collaboration entre enseignants est essentielle pour un enseignement de qualité. La nécessité de la remise en question mutuelle et l'apport de la diversité des compétences semblent être une évidence pour les étudiants. En cela, le projet collaboratif aura permis de renforcer certaines certitudes mais aussi d'en découvrir l'importance spécifiquement pour l'évaluation des apprentissages : par ex. nécessité de croiser les points de vue afin que les éléments soient très clairs pour tous, permet d'éviter certains biais. L'expérience aura permis à certains étudiants de prendre confiance en leur compétences d'évaluateur et d'oser partager leurs pratiques évaluatives.

Certains étudiants relèvent que l'expérience aura permis de renforcer l'idée que le feedback est un puissant moyen pour s'améliorer en tant qu'enseignant. Ils relèvent la nécessité de tenir une posture de praticien réflexif qui est prêt à recevoir des feedbacks de manière positive.

Pour de nombreux étudiants, l'expérience aura également permis de confirmer l'importance de faire travailler les élèves en groupe afin de les responsabiliser et de les impliquer davantage dans leurs apprentissages. Certains enseignants mentionnent le temps important que nécessite un tel travail collaboratif, mais que cela en vaut la peine. Selon certains témoignages, l'expérience vécue aura été un bon exemple de structure d'un travail collaboratif. Aussi, certains étudiants relèvent l'importance de la formation des groupes et de la nécessité de veiller à la gestion des émotions.

Certaines réponses montrent que l'expérience aura permis de montrer ou de renforcer l'importance du feedback entre les apprenants. Entre pairs, l'écoute et la compréhension peuvent être meilleures, les conseils également. Pour ce faire, des étudiants soulèvent l'importance de la composition des groupes. Si le groupe ne fonctionne pas, le feedback peut également devenir contreproductif. Certaines réponses montrent également comme apprentissage que le feedback ne se fait pas qu'à la fin du processus mais tout au long de celui-ci. Certains témoignages soulignent l'importance de l'autoévaluation et de l'implication des élèves dans le processus évaluatif. Enfin, certains témoignages soulèvent tout de même l'aspect chronophage d'un feedback individualisé et de l'importance de sélectionner les bons moments et des méthodes efficaces.

VII. Discussion conclusive

L'un des buts du cours « Evaluation des apprentissages au secondaire II » était de faire vivre aux étudiants, futurs enseignants, certaines caractéristiques d'une évaluation-soutien d'apprentissage (Laveault et Allal, 2016), avec une emphase sur l'implication des apprenants dans le processus évaluatif. Dans l'article, l'analyse portait spécifiquement sur leur vécu de la collaboration et du feedback par les pairs dans le cadre d'un projet en équipe d'analyse et de création d'un outil ou d'une démarche d'évaluation, projet fil-rouge du cours. Les objectifs de l'étude étaient d'interroger le potentiel d'une telle activité comme interface entre les contextes évaluatifs individuels, collectifs et institutionnels.

Apports d'une évaluation collaborative dans la formation à l'enseignement

L'analyse des réponses des 55 étudiants de la cohorte interrogée a pu montrer que les étudiants ont apprécié l'expérience collaborative et en tirent très majoritairement des constats positifs, autant pour le développement des compétences liées à l'évaluation des apprentissages que de celles liées à la collaboration. Cette expérience semble importante pour préparer les étudiants à entrer dans une profession où le travail collectif devient une norme. En outre, une telle préparation semble d'autant plus importante dans la formation au secondaire, où l'enseignement est davantage réputé hermétique à la collaboration (Borgès et Lessard, 2007). L'expérience semble également nécessaire dans le champ spécifique de l'évaluation des apprentissages. Ce domaine de la profession reste en effet souvent considéré comme étant privé, comme une pratique plutôt individuelle (Allal et Mottier Lopez, 2014), mais tend - pour le meilleur comme pour le pire - à devenir une pratique davantage collective (Yerly, 2021). Les données recueillies, certes des pratiques auto rapportées, montrent que les étudiants ont très majoritairement mis en place une véritable « collaboration », voire même une « coopération » où l'intensité du partage et de la réflexion commune est accrue (Marcel et al., 2007). L'expérience aura surtout permis de faire vivre aux étudiants des situations de « modération sociale » (Girardet et Mottier Lopez, 2023 ; Laveault et Yerly, 2017). Selon les témoignages recueillis, cette collaboration a été rendue possible grâce à une structure claire et une nécessaire interdépendance positive (Buchs, 2017), qui donne du sens au travail collectif. Certes, celui-ci a été ici imposé et certains étudiants auraient bien fait l'économie d'aspects chronophages et énergivores du

travail collectif. Toutefois, les données recueillies montrent que les étudiants relèvent les apports de la collaboration pour favoriser la qualité de leur production et qu'ils en emportent certains apprentissages pour leur futur professionnel, pour eux-mêmes mais aussi pour leurs élèves, mais dans une moindre mesure.

L'expérience collaborative vécue positivement par les étudiants a fortement été incarnée dans le feedback par les pairs. Les réponses des étudiants montrent qu'ils ont apprécié donner et recevoir des feedbacks dans le cadre de leur groupe. Ils considèrent le feedback par les pairs comme une très grande plus-value pour la qualité des travaux. Le feedback a permis d'impliquer les étudiants dans le processus évaluatif, de les placer dans le rôle d'évaluateur (*assessment as learning*) (Earl, 2013). Il aura permis de faire vivre aux étudiants un concept qui leur a été enseigné lors du cours. Le témoignage des étudiants montre la réflexion sur les manières d'élaborer et de communiquer des feedbacks (Hattie, 2017), et en fonction de la situation et de la personne qui va recevoir le feedback (Carless et Boud, 2018). L'expérience démontre l'importance de former les futurs enseignants à l'art du feedback en cours, mais également (voire surtout) de leur faire vivre émotionnellement une telle situation. Toutefois, les analyses montrent que l'expérience vécue n'a pas été forcément transférée aussi directement que supposé vers le feedback entre élèves dans les classes.

L'évaluation collaborative, interface entre les contextes ?

Faire vivre aux étudiants l'expérience d'une évaluation collaborative (Allal et Mottier Lopez, 2014) avait comme intention de faire comprendre aux étudiants que l'évaluation est multicontextuelle. L'évaluation des apprentissages fait le lien entre les perspectives des différents acteurs ou groupes d'acteurs (enseignants, élèves, équipe d'enseignants, parents, institution), mais aussi entre les divers espaces (formation, stage sur le terrain, pratique professionnelle future). Au-delà de la collaboration ou du feedback en soi, c'est l'importance de prendre ces diverses perspectives en compte (Issaieva et al., 2011 ; Yerly et Berger, 2022) et d'impliquer les apprenants dans le processus évaluatif qui était visé. Toutefois, cette dimension n'apparaît pas toujours clairement dans les témoignages recueillis. Ceux-ci sont surtout focalisés sur le vécu en tant que futurs enseignants et ne font que peu référence au fait de donner de telles occasions aux élèves dans les classes. Les témoignages soulignent surtout la collaboration et le feedback entre professionnels, et moins comme

stratégie d'enseignement dans la classe et/ou activités d'apprentissage dévolues aux élèves.

L'expérience d'évaluation collaborative en question dans cette étude aura donc surtout permis aux étudiants, futurs enseignants, de confronter leurs propres représentations, croyances et pratiques à celles de leurs collègues. La perspective des élèves a quant à elle été moins présente dans les témoignages, alors que ce transfert, par isomorphisme (Goffin et Fagnant, 2017), était également visé dans le cours. Aussi, lorsque ce transfert a été évoqué dans les témoignages, certains éléments de contextes sont apparus comme des obstacles : le manque de temps dans les classes, le nombre important d'élèves. Ces éléments montrent une certaine difficulté à transférer des pratiques vécues en formation vers le terrain. Ceci démontre que l'isomorphisme doit être pensé non seulement au niveau des contenus, mais également par rapport aux conditions contextuelles. Les expériences vécues par les étudiants universitaires ne peuvent être reproduites dans une classe du secondaire sans adaptation.

Certes, le dispositif a permis de faire vivre les éléments visés aux étudiants, cependant pour renforcer la transposition de ce vécu à leur enseignement et aux actions des élèves, il aurait été nécessaire de 1) mieux expliciter cette attente et 2) de faire des retours métacognitifs sur l'expérience vécue. Un retour individuel leur a été offert par le questionnaire, qui aura permis pour certains de conscientiser l'importance du transfert, mais il est très clair qu'un travail plus approfondi aurait été nécessaire pour pouvoir transposer davantage ce vécu à une implication en classe, avec les élèves. Cette explicitation peut se faire tout au long de l'expérience. Elle peut se faire grâce aux travaux de recherche empiriques qui démontrent l'importance et les effets positifs des éléments travaillés. Dans cette expérience, le formateur aurait pu souligner encore les plus-values de la collaboration et du feedback par les pairs pour l'apprentissage, mais aussi de manière plus globale l'importance d'impliquer les élèves dans le processus évaluatif. Cette explicitation peut également se faire de manière plus pragmatique. La présente étude met en lumière l'importance d'explicitier et de discuter avec les étudiants les aménagements nécessaires pour une transposition dans un autre contexte. Après l'expérience, il aurait été nécessaire qu'un débriefing amène les étudiants à conscientiser davantage ce qu'ils ont vécu, mais aussi d'identifier les manières de transférer leurs observations et stratégies pour des acteurs différents et dans un contexte différent.

Finalement, si ce n'était pas l'objet de l'analyse, il faut relever que les travaux des étudiants se sont montrés d'année en année de très bonne

qualité. Ceci peut témoigner d'un apprentissage réussi en termes d'évaluation, soutenu par l'aspect collectif et le feedback par les pairs. Mais l'intention du dispositif était aussi de développer des compétences plus globales, telles que la collaboration et la réflexivité afin de bousculer certaines croyances quant à l'évaluation de manière durable. Car au-delà des apprentissages spécifiques, faire vivre aux étudiants un tel dispositif vise un apprentissage en profondeur et à utiliser l'évaluation comme un moyen pour apprendre et pour faire apprendre (Earl, 2013 ; Mottier Lopez, 2021). Toutefois, le développement de ces compétences n'a pas été observé et nécessiterait une étude beaucoup plus sophistiquée et longitudinale. Elle serait surtout à observer de manière située, au moment de l'insertion des nouveaux enseignants dans le contexte de leur établissement. Enfin, un tel projet de formation ne peut se faire sur un seul cours et sur un laps de temps aussi réduit. Faire vivre une évaluation des apprentissages qui soutient les apprentissages (Laveault et Allal, 2016) et qui vise le développement de compétences de manière durable (Mottier Lopez, 2021) nécessite l'adhésion de toute une communauté à cette culture et l'implication des différents acteurs dans leurs contextes propres.

VIII. Bibliographie

- Adie, L. E., Klenowski, V. et Wyatt-Smith, C. (2012). Towards an understanding of teacher judgement in the context of social moderation. *Educational Review*, 64(2), 223-240. <https://doi.org/10.1080/00131911.2011.598919>
- Allal, L. (2013). Évaluation : Un pont entre enseignement et apprentissage à l'université. Dans M. Romainville (dir.), *Évaluation et enseignement supérieur* (p. 19-40). De Boeck Supérieur. <https://doi.org/10.3917/dbu.romai.2013.01.0019>
- Allal, L. et Mottier Lopez, L. (2014). Teachers' Professional Judgment in the Context of Collaborative Assessment Practice. Dans C. Wyatt-Smith, V. Klenowski et P. Colbert (dir.), *Designing Assessment for Quality Learning* (p. 151-165). Springer Netherlands. https://doi.org/10.1007/978-94-007-5902-2_10
- Anderson, L.W. et Krathwohl, D.R. (2001). A taxonomy for Learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives. Addison Wesley Longman.

- Borgès, C. et Lessard, C. (2007). Qu'arrive-t-il quand la collaboration enseignante devient une norme ? Dans C. Borgès et C. Lessard (dir.), *Coordonner, collaborer, coopérer* (p. 61-74). De Boeck Supérieur. <https://doi.org/10.3917/dbu.marce.2007.01.0061>
- Buchs, C. (2017). Apprendre ensemble : Des pistes pour structurer les interactions en classe. Peter Lang. <https://orfee.hepl.ch/handle/20.500.12162/5562>
- Carless, D. et Boud, D. (2018). The development of student feedback literacy : Enabling uptake of feedback. *Assessment & Evaluation in Higher Education*, 43(8), 1315-1325. <https://doi.org/10.1080/02602938.2018.1463354>
- De Ketele, J.-M. (2012). À la recherche de paradigmes unificateurs ou de paradigmes intégrateurs ? Dans L. Mottier Lopez (dir.), *Modélisations de l'évaluation en éducation* (p. 195-210). De Boeck Supérieur. <https://doi.org/10.3917/dbu.lopez.2012.01.0195>
- Earl, L. M. (2013). *Assessment as Learning : Using Classroom Assessment to Maximize Student Learning*. Corwin Press.
- Fischer, L., Girardet, C., Mottier Lopez, L., Romainville, M. et Tessaro, W. (2020). Le portfolio dans l'enseignement supérieur : Perceptions des étudiants et intentions des enseignants. *e-JIREF*, 6(1), 3-30.
- Germier, C. et Marcel, J.-F. (2016). La négociation des accords. Approche ethnographique des pratiques de deux collectifs d'enseignants. *Les dossiers des sciences de l'éducation*, 35, 73-96. <https://doi.org/10.4000/dse.1271>
- Girardet, C. (2020). Lorsque les étudiant-es produisent et reçoivent des feedbacks entre pairs à l'université : Etude des perceptions d'un dispositif d'évaluation continue facilité par le numérique. *La Revue LEeE*, 2, Article 2. <https://doi.org/10.48325/rleee.002.04>
- Girardet, C. et Mottier Lopez, L. (2023). Modération sociale entre étudiant-es à l'université : Apprendre pour soi et apprendre ensemble dans une visée démocratique. *La Revue LEeE*, 7, Article 7. <https://doi.org/10.48325/rleee.007.04>
- Goffin, C. et Fagnant, A. (2017). Faire vivre une expérience d'évaluation mutuelle en formation initiale : Quelle transposition vers les pratiques d'enseignement projetées ? Contextes et didactiques. *Revue*

semestrielle en sciences de l'éducation, 9, Article 9.
<https://doi.org/10.4000/ced.734>

Hattie, J. (2017). *L'apprentissage visible pour les enseignants : Connaître son impact pour maximiser le rendement des élèves*. PUQ.

Issaieva, E., Pini, G. et Crahay, M. (2011). Positionnements des enseignants et des élèves du primaire face à l'évaluation : Une convergence existe-t-elle ? *Revue française de pédagogie. Recherches en éducation*, 176, 5-26. <https://doi.org/10.4000/rfp.3362>

Laveault, D. et Allal, L. (2016). Implementing Assessment for Learning : Theoretical and Practical Issues. Dans D. Laveault et L. Allal (dir.), *Assessment for Learning : Meeting the Challenge of Implementation* (p. 1-18). Springer International Publishing.
https://doi.org/10.1007/978-3-319-39211-0_1

Laveault, D. et Yerly, G. (2017). Modération statistique et modération sociale des résultats scolaires : Approches opposées ou complémentaires ? *Mesure et évaluation en éducation*, 40(2), 91-123.
<https://doi.org/10.7202/1043569ar>

Leduc, L., Mohr, A., Marichal, E., Deum, M. et Detroz, P. (2017). Viser le développement des pratiques d'Assessment for Learning (AfL) dans le contexte de la 1^{re} année à l'université : Cadre théorique, dispositif et initiatives pédagogiques résultant du projet « Feedbacks 1^{er} bac ». Dans P. Detroz (dir.), *L'évaluation à la lumière des contextes et des disciplines* (p. 221-255). De Boeck Supérieur.
<https://doi.org/10.3917/dbu.detro.2017.01.0221>

Luisoni, M. et Monnard, I. (2015). L'évolution des conceptions de l'évaluation scolaire chez les futurs enseignants en formation initiale. Dans P.-F. Coen et L. M. Bélair (dir.), *Évaluation et autoévaluation* (p. 175-196). De Boeck Supérieur.
<https://doi.org/10.3917/dbu.ceon.2015.02.0175>

Marcel, J.-F., Dupriez, V. et Périsset Bagnoud, D. (2007). Introduction. Le métier d'enseignant : Nouvelles pratiques, nouvelles recherches. Dans J.-F. Marcel, V. Dupriez, D. Périsset Bagnoud et M. Tardif (dir.), *Coordonner, collaborer, coopérer* (p. 7-17). De Boeck Supérieur.
<https://doi.org/10.3917/dbu.marce.2007.01.0007>

Miles, M. et Huberman, A. M. (2010). *Analyse des données qualitatives*. De Boeck Université.

- Mottier Lopez, L. (2021). Une évaluation continue pour apprendre durablement, une évaluation à visée inclusive. *Revue suisse de pédagogie spécialisée*, 11(4), 9-16.
- Mottier Lopez, L., Tessaro, W., Dechamboux, L. et Villabona, F. M. (2012). La modération sociale : Un dispositif soutenant l'émergence de savoirs négociés sur l'évaluation certificative des apprentissages des élèves. *Questions Vives. Recherches en éducation*, 6(18), 159 -175. <https://doi.org/10.4000/questionsvives.1235>
- Mucchielli, R. (2006). *L'analyse de contenu des documents et des communications*. ESF Editeur.
- Romainville, M. (2013). Évaluation et enseignement supérieur : Un couple maudit, au bord du divorce ? Dans M. Romainville (dir.), *Évaluation et enseignement supérieur* (p. 273-322). De Boeck Supérieur. <https://doi.org/10.3917/dbu.romai.2013.01.0273>
- Sadler, D. R. (2010). Beyond feedback: developing student capability in complex appraisal. *Assessment & Evaluation in Higher Education*, 35, 535-550. <https://doi.org/10.1080/0260293090354101>
- Vangrieken, K., Dochy, F., Raes, E. et Kyndt, E. (2015). Teacher collaboration : A systematic review. *Educational Research Review*, 15, 17-40. <https://doi.org/10.1016/j.edurev.2015.04.002>
- Yerly, G. (2021). La collaboration des enseignants autour de l'évaluation des apprentissages comme moyen de piloter le système ? Le cas de la politique « évaluation commune » dans le secondaire supérieur en Suisse. *Les Dossiers des Sciences de L'Éducation*, 45, 77-97.
- Yerly, G. et Berger, J.-L. (2022). Des pratiques d'évaluation constructives à l'université ? Sous quelles conditions un examen peut-il soutenir les processus d'autorégulation de l'apprentissage. *La Revue LEeE*, 6. <https://doi.org/10.48325/rleee.006.06>

Chapitre 7. Co-construire un référentiel de compétences pour évaluer des articles scientifiques

Expérience d'une enquête collective lors d'un accompagnement doctoral

Lionel Dechamboux, Lucie Mottier Lopez, Céline Girardet, Natalia Constenla Martinez

I. Introduction

Ce chapitre relate une expérience de co-construction d'un référentiel de compétences centré sur une activité du métier de chercheur en éducation : l'évaluation d'articles scientifiques. Afin de faciliter la compréhension de nos lecteurs, nous donnons à voir immédiatement le résultat de cette démarche par le biais d'un extrait de ce référentiel. Ce référentiel est disponible dans sa totalité dans Mottier Lopez et al. (2022).

Figure 1 : Extrait du référentiel de compétences à évaluer un article scientifique

« Domaines »	« Compétences »	« Détails »
Savoirs (ressources internes/ externes)	Savoirs éditoriaux propres à la publication scientifique	<ul style="list-style-type: none">• Structure des textes scientifiques• Composantes des « grandes parties » des textes• Normes bibliographiques• Règles éditoriales et typographiques• La langue
	Savoirs épistémologiques en jeu	<ul style="list-style-type: none">• Paradigmes• Courants de pensées• Hypothèses fondatrices
	Savoirs théoriques en jeu	<ul style="list-style-type: none">• Textes fondateurs du domaine• Champs conceptuels• Actualité des références citées
	Savoirs méthodologiques en jeu	<ul style="list-style-type: none">• Caractéristiques des démarches de recherche• Outils de recherche• Types d'analyses et interprétation
	Savoirs liés à l'éthique de la recherche	<p>Intégrité scientifique</p> <ul style="list-style-type: none">• Plagiat• Diffamation• Protection des personnes (anonymat)
Évaluer - comprendre	Se décentrer pour comprendre le point de vue exprimé (posture)	<ul style="list-style-type: none">• Par rapport à son propre point de vue, adopter une posture qui vise à comprendre le point de vue de l'autre• Partir du principe que les auteur·rices sont des personnes compétentes, mais qu'elles peuvent être remises en question
	Appréhender la globalité tout en examinant la précision du texte	<p>Pour comprendre le sens du texte :</p> <ul style="list-style-type: none">• Se faire une idée globale du texte• Être attentif·ve aux parties et aux détails (contexte, concepts mobilisés, méthode, etc.)• Mettre en rapport le global et le spécifique
	Extraire l'intention scientifique du texte	<ul style="list-style-type: none">• Repérer le type de texte et son intention d'écriture• S'assurer qu'il y ait une cohérence entre les intentions déclarées et ce que l'auteur·rice a rédigé concrètement

La démarche a impliqué des chercheurs confirmés et débutants. Elle s'est inscrite dans le cadre d'un accompagnement doctoral à l'écriture et à l'évaluation scientifiques (Mottier Lopez et Girardet, 2021 ; Mottier Lopez et al., 2022).¹² Les participants devaient se mettre d'accord sur les énoncés du référentiel, avec le postulat que cette activité soutiendrait leur développement professionnel en suscitant les interrogations suivantes : quelles sont les compétences en jeu quand il s'agit d'évaluer / quand je dois évaluer un article en sciences de l'éducation ? Comment décrire ces compétences ? A quel niveau de granularité ? Dans quelle mesure chacun adhère-t-il aux énoncés proposés¹³ ? Etc.

Au-delà du processus de co-construction que nous prendrons soin de décrire et d'analyser, en nous appuyant sur un exemple significatif qui a fait l'objet de débats longs et intenses, nous montrerons que ce processus est nourri par des références à des contextes vécus par les participants, contribuant aux négociations de signification en jeu.

Le chapitre commence par clarifier la façon dont nous appréhendons la question du référentiel et la référentialisation au regard du paradigme de l'évaluation située dans lequel nous inscrivons nos recherches. Nous précisons ensuite le processus d'enquête au sens de Dewey (1938/1993), car nos analyses montrent que c'est bel et bien une enquête que le collectif a poursuivie pour aboutir à la construction d'un référentiel commun. Après la description contextuelle et méthodologique de notre étude, nos analyses sont présentées. La discussion finale revient sur quelques enjeux des contextes dans les processus de référentialisation au sens où nous les définissons.

¹² Dans le cadre du groupe de recherche « Évaluer, Réguler, et Différencier pour apprendre » (EReD) de l'Université de Genève : <https://www.unige.ch/fapse/ered/accueil>

¹³ La co-construction de ce référentiel a donc servi de médiation à ce travail sur l'écriture et l'évaluation scientifiques lors de ces séminaires doctoraux mais a été aussi, en tant que produit fini, mis à la disposition d'autres chercheurs sur la plateforme Uscope (lien dans la note 8).

II. Cadre théorique de notre étude

L'objet « référentiel »

Un référentiel peut être classiquement considéré comme « un construit social qui clarifie les normes d'une activité... Il est ce par rapport à quoi un jugement ou un sens est donné à un objet ou à une action » (Cros et Raisky, 2010, p. 107). Ici, l'activité est celle d'évaluer un article scientifique (dans le domaine des sciences de l'éducation). Concrètement, nous pouvons nous représenter un référentiel, dans une première approche, comme un document (numérique) présentant des énoncés reconnus, décrivant les compétences nécessaires¹⁴ pour effectuer la pratique sociale cible (évaluer un article scientifique). Nous considérons que ces *énoncés* peuvent, potentiellement, servir de *référents*, s'ils sont effectivement mobilisés dans l'activité évaluative. Le paradigme dans lequel nous pensons l'évaluation nous permet de préciser cette « prudence ».

Le paradigme de l'évaluation située

Un référent, à la suite de Barbier (1985), est « 'ce par rapport à quoi' le jugement de valeur est porté ou est susceptible d'être porté » (p. 72). Cette définition conduit à entrevoir les référents comme « normes de jugement, ... critères d'appréciation, ... conduites attendues, ... points de repères » (p. 73). Le référent est associé au référé qui désigne « 'ce à partir de quoi' un jugement de valeur est porté ou est susceptible d'être porté » (p. 64). Le référé est la portion de la réalité confrontée au référent. Ces deux notions sont cruciales dans le champ de l'évaluation. De Ketele (2012), à la recherche d'éléments unificateurs, considère que c'est le point de convergence de toute évaluation : « il s'agit chaque fois de la confrontation d'un référé à un référent » (p. 206), en tant que caractéristiques fondamentales, élémentaires.

Il existe plusieurs manières d'appréhender l'évaluation des travaux scientifiques pour déterminer la réalité que ces notions décrivent. Nous considérons que, dans ce contexte, l'évaluation est une opération impliquant l'interaction d'un acteur (individuel ou collectif) et d'une situation, dans la lignée des travaux situationnistes qui invitent « à conceptualiser (et observer) l'évaluation en interaction avec les contextes, les situations et les dispositifs / outils tels qu'ils sont effectivement mis en

¹⁴ Sans s'attarder sur le débat autour de la notion de compétence.

œuvre » (Mottier Lopez, 2021, p. 5). Ainsi, l'évaluation est foncièrement située : non seulement elle ne peut se penser en dehors des situations dans lesquelles elle se déroule mais, bien plus, elle porte la marque de ces situations. Celles-ci ne sont pas « neutres » ; elles sont construites par les acteurs qui les vivent, les expérimentent. En accord avec Lave (1988), chaque situation est différente pour chaque évaluateur, en raison de l'histoire singulière qu'il rapatrie dans ladite situation, en convoquant d'autres situations antérieurement vécues ou projetées. Dans cette perspective, la confrontation référent-référé devient indissociable des contextes socio-institutionnels mobilisés, des acteurs en présence et de leur histoire. Bien plus qu'une mesure standardisée ou standardisable, l'évaluation équivaut à une construction de signification située.

Dans cette perspective, un référentiel (censé contenir les référents, ce que nous questionnons) recèle un paradoxe puisqu'il est supposé transiter sans modification d'une situation évaluative à une autre. Autrement dit, ses énoncés devraient être à la fois d'une nature générale et stable et en même temps s'accorder avec toutes les situations dans lesquelles leur usage serait requis. Ce paradoxe a alerté de longue date plusieurs chercheurs qui ont développé pour en rendre compte la notion de référentialisation.

La référentialisation

Figari fut le premier, en 1994, dans le champ de recherche sur l'évaluation, à proposer la notion de référentialisation. D'ordre méthodologique, elle vise à dépasser le caractère désincarné des référentiels, pour désigner la « démarche qui consiste à entreprendre une recherche de références pertinentes » (p. 48) servant à modéliser l'objet à évaluer et l'activité évaluative elle-même. Elle permet d'adapter un référentiel préexistant à un contexte donné, voire d'en élaborer un ad hoc.

Parallèlement à ces travaux, nous avons développé la notion de multiréférentialité permettant, entre autres, de pointer le fait qu'au cours de leurs évaluations, les acteurs ne limitent pas les référents qu'ils mobilisent à ceux, par exemple, prescrits par l'institution scolaire (par ex., Mottier Lopez, 2017, 2023 ; Mottier Lopez et Dechamboux, 2019). Les sources de référents sont hétérogènes, inscrites dans une situation particulière et mettant en relation cette situation et des contextes distants, toujours singuliers. Nos travaux empiriques mettent alors en évidence que la référentialisation est plus qu'une méthode : elle s'actualise au cœur de chaque évaluation de manière formelle ou informelle. Dechamboux (2018) décline cette notion dans un multiréférentiel (image instantanée d'un

référentiel toujours mouvant) et dans la multiréférentialisation décrivant ce processus en acte. Observant cette référentialisation en acte et en relation avec la sélection des référés lors d'évaluations, Mottier Lopez et Dechamboux (2019) parlent de *co-constitution référentielle dynamique* soulignant la relation indissociable, dialectique et toujours située entre référents, référés et contextes.

La conception d'un référentiel figé, autorisant la sélection de référés toujours identiques, semble avoir fait long feu. C'est pourquoi, nous considérons que le référentiel, objet matériel élaboré en dehors d'une évaluation, est constitué d'*énoncés* en tant que repères potentiels pour une évaluation (Mottier Lopez, 2017). Ils ne deviendront des référents que s'ils sont effectivement saisis dans l'activité évaluative, au cœur du processus de co-constitution référentielle dynamique. Dans la suite du texte, c'est donc délibérément que nous parlons d'*énoncés* du référentiel et non directement de référents.

Une enquête collective

La co-construction du référentiel qui nous intéresse est fondée sur des débats nourris. Il est apparu que le flux des interactions, en lien avec les expériences évaluatives situées et les biographies personnelles (Mottier Lopez et al., 2022), répondait à une logique d'enquête au sens de Dewey (1938/1993). Pour ce dernier « l'enquête est la transformation contrôlée ou dirigée d'une situation indéterminée en une situation qui est si déterminée en ses distinctions et relations constitutives qu'elle convertit les éléments de la situation originelle en un tout unifié » (p. 169). Elle se déclenche lorsque l'individu est face à un problème qui vient troubler ses habitudes. Elle devient le moyen de rétablir un équilibre dans une nouvelle situation, différente, mais vue comme satisfaisante, apaisée. Par l'enquête, le trouble initial est dépassé en passant par un processus dont les différentes phases sont toujours présentes mais dont l'ampleur et la complexité peuvent considérablement varier.

L'enquête débute par le constat de l'existence d'une situation indéterminée¹⁵. Le futur enquêteur ne sait pas précisément ce qui ne convient pas mais ressent une gêne. Dès celle-ci identifiée, la situation devient problématique : le problème se construit en distinguant notamment ce qui, dans la situation, est connu, maîtrisé, de ce qui ne l'est

¹⁵ Indéterminée, déterminée, problématique sont les termes utilisés par Dewey pour décrire les différents « états » de la situation lors d'une enquête.

pas. Dès lors, l'enquêteur peut formuler des hypothèses dont il peut observer les conséquences par de nouvelles observations, par des expérimentations, etc. Le succès de l'enquête se signe par l'atteinte d'une situation à nouveau déterminée. Il ne faut cependant pas imaginer cette détermination comme totale, objective. Elle n'est que temporaire, contextuelle. L'acteur met en adéquation les moyens et les fins en vue de manière satisfaisante en cet instant pour qu'il puisse poursuivre son activité, sans la gêne initiale. En cas d'échec, la situation reste encore indéterminée mais cette indétermination se réduit, car une hypothèse est désormais éliminée ; la situation progresse malgré tout. Un ou plusieurs cycles d'enquête peuvent permettre d'aboutir à la détermination désirée. Quel que soit son déroulement, la fin de l'enquête n'est jamais un retour à la situation initiale.

La vie offre de nombreuses occasions de mener des enquêtes non seulement individuelles mais également collectives : c'est le cas dans notre projet de co-construire un référentiel commun faisant sens pour toutes et tous. Comme observé par Zask (2004), l'enquête collective devient une inter-objectivation à l'issue de laquelle le public¹⁶ doit pouvoir se saisir de l'objet qui lui permettra non pas une expérience identique mais convergente. Nous considérons dans une perspective située qu'elle reconnaît alors les *différentiels* de chacun en tant que versant personnel du référentiel (Vial, 2012). Agissant dans l'évaluation, les différentiels sont singuliers au regard de la relation dialectique entre l'évaluateur agissant et les situations et contextes co-constitutifs (Mottier Lopez, 2023).

Questionnement de recherche

Notre questionnement s'est voulu très ouvert : qu'est-ce qui se joue dans la co-construction d'un référentiel au fil d'activités évaluatives et de confrontations de significations situées, subjectives et inter-subjectives au sein de ce collectif de chercheurs ? Ou encore, en quoi cette activité de co-construction s'apparente-t-elle à un processus d'enquête dont il s'agira de repérer et d'analyser les différentes étapes ?

¹⁶ Dans notre étude, le public est donc constitué des participants mais également de ceux qui se saisiraient de ce référentiel une fois celui-ci achevé.

III. Contexte de l'étude et méthodologie

Avec des variations selon les disponibilités de chacun et l'entrée dans l'accompagnement doctoral, le groupe¹⁷ était composé de :

- Lucie : professeure ordinaire, directrice des thèses des doctorants et jeunes chercheurs, thèse soutenue en 2005 ;
- Céline : maîtresse assistante de l'équipe de recherche, thèse soutenue en 2017 ;
- Lionel : chargé d'enseignement dans l'équipe de recherche, thèse soutenue en 2018 ;
- Fernando : enseignant-chercheur, thèse soutenue en 2019 ;
- Valérie : doctorante au début de la démarche puis qui a soutenu sa thèse durant le processus en 2021 ;
- Huit doctorants à différents états d'avancement de leur thèse (que certains ont d'ailleurs soutenue depuis la fin de la démarche) : Elisabeth, Francisco, Murielle, Sébastien, Olivier, Stéphanie, Natalia et Pala.

La démarche s'est déroulée entre novembre 2019 et juillet 2021.¹⁸ Elle a articulé des séminaires de formation et des activités réelles et fictives d'écriture et d'évaluation scientifiques. Un outil numérique (plateforme Uscope¹⁹) a été utilisé, permettant la construction collaborative du référentiel et l'observation de son évolution au fil du temps. Cet outil a également permis de collaborer hors réunion par des conversations numériques liées aux entrées du référentiel. Le tableau 1 présente le déroulement en trois phases des réunions et des activités réalisées par les participants dans diverses modalités (P = en présentiel ; D = à distance ; GC = groupe complet ; SG = sous-groupe ; I = individuellement). Pour les lecteurs intéressés, la démarche et la version finale du référentiel sont détaillées dans Mottier Lopez et al. (2022).

¹⁷ Dont l'anonymat n'a pas été souhaité par eux comme montré par le numéro 3 de La Revue LEE dans lequel on retrouve leurs noms : <https://revue.lee.on-line/index.php/info/issue/view/9>

¹⁸ Mottier Lopez et Girardet (2021) présentent une objectivation de cet accompagnement doctoral.

¹⁹ Nous remercions Patrick Favre pour la mise à disposition de cet outil. <https://uchronic.ch/uscope/>

Tableau 1 : Déroulement des séances et activités (adapté de Mottier Lopez et al., 2022)

Phases	Dates	Modalités	Contenus
Initiation	15.11.19	P GC	<ul style="list-style-type: none"> Présentation de l'outil de co-construction de référentiels Uscope en présence de son concepteur et introduction
	*07.02.20	P GC	<ul style="list-style-type: none"> Échanges sur les compétences pour évaluer un article Début de co-construction collective du référentiel À la fin du séminaire, chacun-e a élaboré sa propre version
	20.02.20	D SG	<ul style="list-style-type: none"> Élaboration d'une première version du référentiel sur la base des premiers échanges et des versions individuelles
20.02.2020 : mise en ligne d'une première version du référentiel sur la plateforme Uscope			
Expéri- mentation	*25.06.20 03.07.20	D SG	<ul style="list-style-type: none"> Modifications apportées au référentiel en ligne sur la base d'un exercice d'évaluation d'un article scientifique
	24.09.20	D GC	<ul style="list-style-type: none"> Relance de la démarche Uscope. 1 heure pour reprendre contact après l'été et planifier la suite
	22.10.20 23.10.20	D GC	<ul style="list-style-type: none"> Modifications apportées au référentiel en ligne sur la base d'un exercice d'évaluation d'un article scientifique
	27.11.20	D GC	<ul style="list-style-type: none"> Échanges sur les expériences d'évaluation scientifique
	02.2020 - 03.2021	D I	<ul style="list-style-type: none"> Évaluation authentique d'un article dans une revue Accompagnée si besoin de la directrice de thèse
Finalisa- tion	*12.03.21	D GC	<ul style="list-style-type: none"> Suite de la co-élaboration du référentiel Début de résolution des nœuds existants

	*07.05.21	D GC	<ul style="list-style-type: none"> • Suite de la co-élaboration du référentiel • Discussion sur la schématisation conceptuelle du référentiel
	*26.05.21	D SG	<ul style="list-style-type: none"> • Reprise en sous-groupe du référentiel pour appliquer les décisions prises lors de la dernière réunion et identifier les derniers nœuds à traiter avec le groupe complet.
	*18.06.21	D GC	<ul style="list-style-type: none"> • Finalisation du référentiel de compétences Uscope • Amorce d'un projet d'article collectif
06.07.2021 : Publication du référentiel sur Uscope			
14.07.2022 : Publication de l'article collectif (Mottier Lopez et al., 2022)			

Note. * = séances enregistrées

Les données analysées dans ce chapitre sont issues de six séances enregistrées (* dans le tableau 1). Elles ont fait l'objet d'une transcription intégrale puis d'une analyse qualitative de contenu (Bardin, 2013). Elles ont été retenues parce qu'elles ont donné lieu à un débat central entre les participants. Un tableau d'analyse a été élaboré contenant les passages de transcriptions concernés par ce débat, les arguments énoncés et l'inférence d'obstacles²⁰ à la co-construction du référentiel. Les analyses ont ensuite été interprétées à l'aune de la notion d'enquête (Dewey, 1938/1993).

IV. Résultats

L'exemple significatif présenté ci-dessous concerne un débat récurrent à propos des termes *identifier* et *maîtriser* comme verbe initiant le libellé des compétences du référentiel. Nous restituons ci-dessous quelques éléments du cheminement collectif à partir de la logique de l'enquête.

²⁰ La longueur à disposition pour l'écriture de ce texte ne nous permettra pas d'en faire état ici.

Une situation indéterminée

Après avoir adhéré au défi de définir et de structurer un référentiel de compétences pour évaluer des textes scientifiques, une préoccupation (parmi d'autres) émerge :

Olivier : moi je me souviens que j'ai un commentaire que j'avais fait sur un texte... je me suis dit mais comment est-ce qu'on peut se permettre de dire ça...

Céline : ouais... on va plus ou moins avoir besoin d'aller rechercher les ressources nécessaires pour comprendre. (Z1-Tdp 390, 487)

A partir d'une expérience évaluative exprimée par Olivier, quelque chose de problématique est pressenti. Des premiers mots sont posés, mais aucune question ou solution n'est encore clairement identifiée. L'enquête démarre.

Une situation problématique

Nous pouvons cependant repérer un mouvement, un ensemble de prises de parole qui aboutissent à une première formulation de ce qui semble être à résoudre. Ce passage, très important, correspond dans l'enquête à la transition de la *situation indéterminée* à la *situation problématique*.

A cette occasion, les participants affinent leur questionnement en essayant de délimiter l'étendue du savoir à maîtriser pour pouvoir évaluer un article. Nous voici donc en présence de l'origine du questionnement : pour évaluer un article scientifique, doit-on maîtriser tous les savoirs ou des savoirs spécifiques (par exemple épistémologiques comme le dit Olivier) ?

Olivier : ... ça me demande tellement de compétences et tellement de connaissances que finalement... et puis comme ça, ça nous permet d'être, d'être sur non pas maîtriser les savoirs et connaître les savoirs il y a juste **maîtriser** les savoirs épistémologiques en jeu... (Z1-Tdp 524)

Au regard de l'autoévaluation de ses compétences et connaissances dans une situation précédemment vécue, Olivier introduit le terme *maîtriser*. Prémisses du débat qui suivra sur le choix des termes entre *identifier* et *maîtriser* pour libeller l'énoncé d'une compétence.

Des tentatives d'organisation de la situation

Suite à la première séance, Olivier rédige un commentaire sur la plateforme, visible par tous :

Olivier : je n'ai jamais été à l'aise avec la formulation maîtriser... parce que je ne vois pas concrètement à quoi elle renvoie. Ici, est-ce que « connaître les méthodologies, les outils associés... » ne suffirait pas ? (Commentaire déposé sur la plateforme Uscope)

Le débat reprend en réunion et se polarise très clairement autour de ce terme « maîtriser » avec des interventions qui, par exemple, essaient d'exemplifier le problème :

Céline : et surtout que dans *maîtriser les savoirs* à chaque fois on a rajouté *en jeu* parce que c'est **propre** à, ça dépend de l'activité ! Et l'article qu'on est en train d'évaluer, parce qu'on n'a pas besoin de maîtriser les savoirs en analyse corrélationnelle si c'est un article avec une méthodologie qualitative. (Z2-Tdp 284)

La dernière citation est ici fidèle au questionnement initial (l'étendue des savoirs à posséder) mais très vite, le débat va s'orienter différemment lorsqu'Olivier propose de préférer *identifier* à *maîtriser*.

Olivier : c'est pas forcément de les connaître, mais peut-être plutôt de les identifier. Et du coup, quand on est dans *identifier* alors peut-être là c'est un peu plus **juste** de faire le lien entre, soit une connaissance qu'on a déjà à l'intérieur de nous, soit la capacité à aller **rechercher** (les ressources nécessaires) ... dans *identifier* je trouve qu'on est plus **clairement** sur une lecture d'article, que dans *maîtriser* ou *connaître*. (Z2-Tdp 292)

Cette reformulation, qui réorganise la situation problématique et le questionnement, paraît acceptée par les participants par souci notamment de réalisme :

Stéphanie : oui je suis d'accord... parce que quand on évalue **chaque** article... on peut pas **tout** savoir **tout** chercher... alors identifier... **oui**, c'est plus **global**...

Francisco : c'est plus faisable, en tout cas. (Z2-Tdp 300, 301)

Mais si nous analysons cette reformulation *a posteriori*, nous pouvons déceler une sorte d'*erreur logique*. En effet, les participants cherchant une solution quant à l'*étendue* des savoirs à posséder (tous les savoirs *vs* un type de savoir), discutent désormais sur la *profondeur* du savoir (reconnaître, identifier *vs* maîtriser), autrement dit, le degré de maîtrise d'un savoir donné. Mais cela n'est pas relevé dans les interactions et c'est finalement dans cette direction que l'enquête se poursuit.

Les discussions s'engagent sur la base de diverses expériences convoquées explicitement, y compris hors de l'évaluation scientifique (évaluation en classe, notamment). Certains reprochent au terme *maîtriser* un manque de précision à l'instar du terme *comprendre*, par exemple. Apparemment plus facilement appréhendable dans leur propre activité évaluative, *identifier*, pour eux, c'est reconnaître. Cela leur paraît une compétence plus accessible (que maîtriser les savoirs) quand ils sont confrontés à l'activité réelle d'évaluateur à propos des savoirs en jeu dans un article scientifique, quel que soit leur type (savoirs éditoriaux, savoirs méthodologiques, etc.). Cette discussion est soutenue par le recours de certains à la taxonomie de Bloom. Mais à la mobilisation de cet outil théorique est opposée la confrontation avec l'activité réelle de l'évaluateur à propos de différents savoirs pouvant être en jeu dans un article. Un exemple, celui des savoirs éditoriaux, servira de test aux différentes conceptions. Pour certains, il suffit de comparer les modalités éditoriales utilisées avec celles prescrites par la revue (tenants de l'identification) alors que pour d'autres, il faut maîtriser les savoirs éditoriaux prescrits pour évaluer ceux utilisés (tenants de la maîtrise). Ainsi, chaque argument procède implicitement ou explicitement d'un contexte, d'une expérience vécue ou projetée.

À l'issue de la deuxième séance, une participante s'exprime ainsi :

Elisabeth : Je me sens plus à l'aise avec *identifier* qu'avec *maîtriser* d'une manière générale car la compétence me semble... plus accessible pour moi énoncée de la sorte. (Commentaire déposé sur la plateforme Uscope)

Le terme *maîtriser* ne correspond pas à la manière dont elle s'auto-évalue face aux exigences perçues de la pratique sociale. Mais Céline objecte après s'être mise en situation :

Céline : Du coup moi en m'auto-évaluant avec le terme *identifier* au lieu de *maîtriser*, j'ai eu l'impression qu'il manquait quelque chose. Par exemple je peux identifier qu'ils parlent de savoirs sur des approches ergonomiques, mais à quoi cela m'avance de l'avoir identifié ? Pour évaluer leurs propos, j'ai quand-même besoin d'une certaine maîtrise de ces éléments. (Commentaire déposé sur la plateforme Uscope)

Mais *identifier* remporte, à l'exception donc d'une personne, les suffrages même si son appréhension reste sujette à questionnement :

Olivier : Pour ma part, *maîtriser* n'a jamais exprimé rien de concret, alors qu'*identifier* ça m'apparaît concret. J'ai l'impression de pouvoir dire ça je l'ai vu, ça non. Mais d'une certaine façon, pour *identifier*, je suis bien obligé

de *maîtriser* un minimum, sinon je ne suis pas capable d'*identifier*.
(Commentaire déposé sur la plateforme Uscope)

Au cours de la troisième séance, si le terme *identifier* convient presque à tout le monde, il est pour certains une étape transitoire. Reste à *évaluer*, à *apprécier* ce qui aura été identifié :

Murielle : bon, t'identifies l'élément puis après tu passes à un autre stade pour évaluer si c'est adéquat ou pas. (Z3-Tdp 49)

D'autres débats se développent en parallèle, notamment celui de la finalité du référentiel : est-il censé représenter l'activité de tout évaluateur y compris expert, ou est-il un outil de formation ? L'objectif est rappelé et clôt cette interrogation : il s'agit d'élaborer un référentiel de compétences utile pour sa propre activité en développement (pour soi) et susceptible de circonscrire une activité constitutive du métier de chercheur en éducation (pour la communauté concernée).

Un autre élément émerge, mettant en cause la structure de la plateforme numérique qui accueille le référentiel, ne permettant pas suffisamment de restituer la dynamique de l'activité.

Murielle : mais moi en fait j'ai un petit souci parce qu'il me semble que l'aspect de l'évaluation n'apparaît pas dans ce tableau, parce qu'on va identifier par exemple les parties d'un article mais en fait ça va être un **va-et-vient** dans notre tête entre, on **sait**, on a les connaissances qu'un article scientifique doit avoir telle et telle partie et puis on fait un va-et-vient entre est-ce qu'y'a cette partie est-ce qu'y'a cette partie et c'est pas juste identifier. La partie dynamique qui se fait dans notre tête elle est pas là [rire] et moi ça me pose un souci quoi ! (Z3-Tdp 58)

Si cet axe de questionnement n'est pas alors plus approfondi à cet instant, il convient de le garder en tête pour notre analyse puisqu'il s'agira, nous le verrons, d'un germe de ce qui constituera la porte de sortie du débat.

En attendant, le débat évolue. Certes, il s'agirait d'*identifier* pour ensuite *apprécier*. Mais pour *identifier* quelque chose, faut-il le connaître préalablement ? Et si oui, ne faudrait-il donc pas avoir une certaine *maîtrise* de ce savoir... *identifier*, *connaître* et *maîtriser* apparaissent très proches. C'est autour de cette question que les échanges se poursuivent :

Murielle : ... mais ça me dérange un petit peu parce que pour moi c'est vraiment **préalable** à l'identification. On doit maîtriser un certain nombre de concepts. On doit maîtriser ce que c'est qu'un article scientifique et qui comporte un certain nombre de parties. On doit maîtriser un certain nombre de **choses** pour pouvoir ensuite identifier ces choses-là dans l'article de l'autre, enfin de l'article qu'on évalue. (Z3-Tdp 119)

Le débat se fait profondément épistémologique : commence-t-on par repérer un signe, une information disponible, qui évoque quelque chose sur le plan cognitif et situationnel ? Ou la connaissance antécédente de certains signes, incluant les contextes de leur reconnaissance, est-elle la condition nécessaire pour les voir émerger dans l'article à évaluer ? Ces deux visions s'affrontent, y compris dans la manière dont chacun conçoit et découpe sa propre activité d'évaluateur. Le terme *maîtriser* pose problème tout autant qu'il paraît nécessaire :

Olivier : ... je suis d'accord d'un côté avec Murielle pour dire ben y'a déjà le niveau qui permet *d'identifier*, puis après y'aura les niveaux qui permettront *d'apprécier*, de *comprendre*... mais c'est vrai que le niveau *maîtriser* je sais pas où il m'emmène. (Z3-Tdp 129)

Un consensus est alors tenté : retenir les deux termes, *identifier* et *maîtriser*. Mais cette proposition n'est également pas retenue, car elle modifierait le niveau de formulation généralement adopté dans les référentiels (un seul verbe d'action). On observe que l'organisation de la situation, en l'état, reste insatisfaisante et ne permet pas d'avancer vers le choix d'un objet commun (un libellé accepté par tous), permettant de rétablir une situation apaisée. Toutes les expérimentations, les observations, les hypothèses échouent avec cette manière de structurer la situation.

Le débat revient alors sur la structure du référentiel et son incapacité à rendre compte de la dynamique de l'activité à laquelle tiennent les participants. Une nouvelle organisation de la situation problématique est recherchée. Finalement, il n'est plus question d'une phase préalable pendant laquelle l'évaluateur identifierait dans l'article des savoirs qu'il maîtrise ou non avant de pouvoir les évaluer. Ils sont à repérer et à apprécier de manière dynamique tout au long de l'activité. Des remarques sont formulées sur l'épistémologie sous-jacente à l'outil numérique et à la notion de compétence. Apparaît également dans la discussion la notion de *ressource* qui s'y rattache. Face à l'obstacle à la fois structurel et conceptuel, le groupe décide alors de rédiger un texte introductif au référentiel pour expliquer sa position. La troisième séance s'achève sur cet accord et cette nouvelle organisation de la situation.

Une situation à nouveau déterminée

La quatrième séance reprend avec l'idée de ce texte introductif. Couplé à la remise en question de l'outil (véhiculant une certaine vision de l'activité évaluative et son propre contexte), le terme de *ressource*, désigné de manière quasi synonyme à celui de savoir ou de connaissance, ouvre soudainement à une nouvelle compréhension de la situation problématique. Toutes les hypothèses avaient jusqu'alors abouti à des impasses. Cette nouvelle formulation apporte une issue et, ce faisant, offre une nouvelle organisation de la situation problématique :

Céline : ... oui, bien sûr, on a des connaissances préalables qui sont nécessaires à l'activité... si la ressource n'est pas entièrement déjà dans ma tête, je vais aller chercher une ressource externe... Du coup, je pense que dans nos ressources, il faut bien qu'on spécifie qu'on est sur les ressources y compris les ressources internes à nous. (Z4-Tdp 238)

Cette notion de *ressource* permet aux participants de se délester de l'idée qu'il s'agit de posséder « tous » les savoirs, mais qu'ils pouvaient tout aussi bien accomplir l'évaluation en les cherchant à l'extérieur, hors d'eux-mêmes. Pour le dire dans les termes de l'enquête, cette nouvelle organisation de la situation permet de trouver une issue débouchant sur un objet commun accepté par toutes et par tous, un moyen compatible avec les fins en vue. Elle permet, dans sa forme, de rétablir la continuité de l'expérience. Lors de la sixième séance, les participants conviennent d'introduire le référentiel contraint par la plateforme numérique non seulement par un texte mais également par un schéma où les ressources sont présentées de manière transversale et en spécifiant qu'elles peuvent être internes ou externes.

Nous avons pointé plus haut la dérive des discussions par rapport à la question d'enquête initiale : elles portaient en effet sur la profondeur de la connaissance (identifier *vs* maîtriser) alors que la problématique initiale était liée à son étendue. Or la situation s'apaise lorsque le choix est fait de désigner des ressources internes et externes. Nous comprenons ainsi le retour à une situation déterminée signant la fin de l'enquête. En dépit de ces détours, la conclusion répond bien à la question initiale. En plus de son caractère logique, elle est d'ailleurs plus compatible épistémologiquement avec l'idée d'une distribution des ressources entre un acteur et son environnement dans les approches situées qui caractérisent l'épistémologie de ces participants.

V. Discussion conclusive

Nombre d’auteurs notent une difficulté à élaborer un consensus lors de la co-construction d’un référentiel (Parent et al., 2006), notamment parce que les acteurs impliqués ont du mal à s’extraire de leurs expériences individuelles (Houot et Lavielle-Gutnik, 2018). Cette difficulté peut se comprendre à la lueur de l’importance des contextes d’où proviennent ces expériences antécédentes. De tels aspects ont été observés dans nos analyses, tout en reconnaissant la nécessité de partir de sa propre expérience pour élaborer les énoncés qui seront ensuite intégrés au référentiel. Nous avons donné à voir un exemple de la difficulté des participants à s’accorder sur l’emploi de certains termes. Cette difficulté est constatée également par de nombreux auteurs (par ex., Pingoud, 2004 ; Pochet et al., 2021).

Les choix débattus et opérés par les participants relèvent d’une décision située, qui leur paraît acceptable à un moment donné, relativement à leurs expériences passées dans divers contextes : en tant que professionnels dans une classe ou déjà en tant que chercheurs, en tant qu’étudiants doctorants, au regard d’expériences évaluatives précédentes et/ou réalisées tout au long de l’accompagnement doctoral, par rapport à la projection des exigences de la pratique sociale « évaluer des articles scientifiques ». Non seulement le référentiel co-élaboré doit prendre sens par rapport aux expériences objectivées dans différents lieux (voir tableau 1), de façon subjective et intersubjective, mais également pour servir à leurs expériences futures au-delà du contexte spécifique de l’accompagnement doctoral. L’ambition du projet était en plus de rendre public, et donc légitime, le référentiel élaboré pour la communauté des sciences de l’éducation. Pour le dire autrement, l’élaboration d’un référentiel générique se détermine sous l’influence d’anciennes référentialisations, et le référentiel ainsi élaboré est vu comme pouvant potentiellement être utilisé pour des référentialisations ultérieures (au sens de Mottier Lopez). Les participants sont en effet conscients que leurs propositions sont susceptibles d’être questionnées, remises en cause, modifiées.

L’angle d’analyse choisi pour cet article, c’est-à-dire autour des débats qui peuvent avoir lieu à propos des termes utilisés dans un référentiel, montre que derrière un même mot, chacun des acteurs habite ce même mot avec un sens qui peut différer, parfois grandement, un sens qui provient des différents contextes vécus par chacun, qui légitime l’existence de son « différentiel » (Mottier Lopez, 2023). C’est un constat auquel le groupe est

finallement parvenu, se traduisant dans le référentiel par un espace explicitement dévolu à l'expression de ce différentiel, comme montré dans Mottier Lopez et al. (2022). Le problème n'est donc pas seulement un problème de granularité de la description mais également, et plus profondément, de sens. Choisir un terme plutôt qu'un autre (identifier *vs* maîtriser par exemple) pourrait avoir peu de sens si le débat reste en surface. Aller au-delà de ce choix amène à questionner la différence de signifié associé à chaque signifiant, en raison d'une construction de sens différente, qui s'ancre dans des expériences différentes, des contextes particuliers. Ce constat mérite réflexion pour de futures élaborations de référentiels. Comment prendre en considération cette donnée incontournable pour qu'elle ne constitue pas un obstacle trop important et éviter que le débat ne tourne en rond, comme nous avons parfois pu le constater dans notre étude ? Faudrait-il encourager les participants à expliciter non seulement le sens attribué aux termes problématiques mais aussi leur suggérer de mettre à jour les expériences, les contextes qui ont permis de construire ce sens ? Cet effort « méta » aurait-il la vertu, par les prises de conscience occasionnées, de permettre aux participants de réaliser l'origine des divergences et, par là-même, de les dépasser, de co-crée du sens nouveau ? Cela reste à éprouver.

VI. Bibliographie

- Barbier, J.-M. (1985). *L'évaluation en formation*. Presses Universitaires de France.
- Bardin, L. (2013). *L'analyse de contenu*. Presses Universitaires de France.
- Cros, F. et Raïsky, C. (2010). Référentiel. *Recherche et formation*, 64.
- Dechamboux, L. (2018). *L'enquête formative dans le paradigme de l'évaluation située : une modélisation énaïve de l'activité évaluative d'enseignantes lors de séances de production écrite au CP* [Thèse de doctorat]. Université de Genève, Suisse. Archive ouverte UNIGE.
- De Ketele, J.-M. (2012). A la recherche de paradigmes intégrateurs ou de paradigmes unificateurs. Dans L. Mottier Lopez et G. Figari (dir.), *Modélisations de l'évaluation en éducation – questionnements épistémologiques* (p. 195-210). De Boeck.
- Dewey, J. (1938/1993). *Logique, la théorie de l'enquête*. Presses Universitaires de France.
- Figari, G. (1994). *Évaluer : quel référentiel ?* De Boeck Université.

- Houot, I. et Lavielle-Gutnik, N. (2018). Nouveau référentiel de compétences dans les écoles de la deuxième chance : vers de nouvelles transactions. *Savoirs*, 46, 91-110.
- Lave, J. (1988). *Cognition in practice: Mind, mathematics and culture in everyday life*. Cambridge University Press.
- Mottier Lopez, L. (2017). Une modélisation pour appréhender la référentialisation dans l'évaluation des apprentissages des élèves. Dans P. Detroz, M. Crahay et A. Fagnant (dir.), *L'évaluation à la lumière des contextes et des disciplines* (p. 169-192). De Boeck.
- Mottier Lopez, L. (2021). Évaluation située élargie : objectivation à partir de l'évaluation formative élargie, l'apprentissage situé et la référentialisation. Dans C. Barroso da Costa, I. Nizet et D. Leduc (dir.), *Quarante ans de mesure et d'évaluation* (p. 119-138). Presses Universitaires du Québec.
- Mottier Lopez, L. (2023). Rapports différentiels dans la référentialisation de l'évaluation : des expériences démocratiques ? *La Revue LEeE*, 7.
- Mottier Lopez, L. et Dechamboux, L. (2019). Co-construire le référentiel de l'évaluation formative pour soutenir un processus de co-régulation dans la microculture de classe. *Évaluer. Journal international de recherche en éducation et formation*, 5(2), 87-111.
- Mottier Lopez, L. et Girardet, C. (2021). Accompagnement doctoral à l'écriture et à l'évaluation scientifiques : objectivation du dispositif expérimenté autour des articles du numéro 3 de La Revue LEeE. *La Revue LEeE*, 3.
- Mottier Lopez, L., Girardet, C., Bendela, P., Constenla Martinez, N., Elisme Pierre, E., Gibert, S., Lyu, X., Roth, M., Sauret, O. et Seguel Tapia, F. (2022). Quand la co-construction d'un référentiel devient un moyen d'apprendre : expériences de doctorantes et doctorants, le cas d'un référentiel pour évaluer des textes scientifiques. *La Revue LEeE*, 3.
- Parent, F., Lemenu, D., Lejeune, C., Komba, D., Baulana, R., Kahombo, G. et Coppieters, Y. (2006). Référentiel de compétences infirmières en RD du Congo dans le cadre du renforcement des ressources humaines en santé. *Santé Publique*, 18(3), 459-473.
- Pingoud, K. (2004). Élaboration d'un référentiel de compétences pour exercer la profession d'infirmière en milieu scolaire. *Santé Publique*, 2(16), 225-238.
- Pochet, B., Durieux, N., Collette, C., Decoster, S., Dieudonné, M. et Vandendput, S. L., (2021). Le référentiel « cinq piliers de la maîtrise de

l'information scientifique » : élaboration et construction de l'outil. *Revue internationale de pédagogie de l'enseignement supérieur* 37(2), 1-11.

Vial, M. (2012). *Se repérer dans les modèles de l'évaluation : méthodes, dispositifs, outils*. De Boeck.

Zask, J. (2004). L'enquête sociale comme inter-objectivation. Dans B. Kar-senti et L. Quéré (dir.), *La croyance et l'enquête. Aux sources du pragmatisme* (p. 141-163). De Boeck.

Chapitre 8. Évaluer la qualité d'un dispositif de formation en ligne

Un processus complexe qui articule validité sociale, validité scientifique et productivité

Olivier Leyh, Charlotte Dejaegher, Marine André et Patricia Schillings

I. Contexte et problématique

Entre 2017 et 2020, notre équipe de recherche a été chargée de mesurer les effets du programme PARLER (Parler, Apprendre, Réfléchir, Lire Ensemble pour Réussir) (Zorman et al., 2015) sur les performances des élèves issus de plusieurs écoles de Belgique francophone sélectionnées pour cette étude (Dejaegher et al., 2022).

Le programme PARLER, développé par Michel Zorman et une équipe de chercheurs de l'Université de Grenoble, entend prévenir les difficultés d'entrée dans le monde de l'écrit en travaillant, dès la troisième année maternelle (GS) et jusqu'en deuxième année primaire (CE1), les habiletés qui sous-tendent l'apprentissage de la lecture : la conscience phonologique, la maîtrise du code, la fluence, la catégorisation (vocabulaire) et la compréhension de la langue (Zorman et al., 2015).

Plusieurs principes organisateurs, basés sur des recommandations issues de la recherche, régissent ce programme : la formation de petits groupes d'élèves de même niveau pour une habileté donnée, l'adoption de pratiques de différenciation qui amènent les enseignants à consacrer davantage de temps aux élèves en difficulté, le recours à un enseignement structuré, basé sur l'explicitation des enjeux et des démarches d'apprentissage ou encore la mise en place de séances régulières, courtes et intensives (Zorman et al., 2015).

Dès le début de notre projet, alors que les enseignants impliqués dans l'étude découvraient le programme qu'ils allaient être chargés de mettre en œuvre, nous avons remarqué qu'il existait un décalage parfois important entre les principes organisateurs (ou recommandations de la recherche) cités plus haut et les croyances, les connaissances et les pratiques habituelles de ces enseignants (Dejaegher et al., 2023). Dès lors, nous avons compris que notre rôle ne se limiterait pas seulement à mesurer l'impact de la mise en œuvre du programme PARLER dans les classes de ces professionnels. Il s'agirait d'abord de les accompagner et de les soutenir dans l'appropriation des pratiques pédagogiques innovantes liées à cette implantation.

Notre réflexion s'est ensuite poursuivie pour envisager la mise en œuvre du programme PARLER par d'autres équipes pédagogiques que celles impliquées dans notre étude. Puisqu'un accompagnement important de la part des chercheurs semblait nécessaire pour aider les enseignants à adopter les pratiques recommandées, comment, en dehors du contexte de la recherche, répondre aux demandes de professionnels désireux d'utiliser le programme PARLER dans leurs établissements lorsqu'ils seraient confrontés aux mêmes difficultés que leurs pairs ? Comment leur permettre d'apporter à leurs pratiques les changements nécessaires ? Proposer un accompagnement personnalisé à chaque école serait difficile, voire impossible, notre équipe de chercheurs ne pouvant à elle seule consacrer suffisamment de temps à cette mission et répondre au nombre grandissant de sollicitations.

Face à ces constats, nous avons formulé le projet de concevoir un dispositif de formation en ligne poursuivant un objectif double : soutenir l'implantation du programme PARLER sur le terrain et véhiculer auprès des praticiens des savoirs scientifiques susceptibles de participer à leur développement professionnel (Kervyn et Goigoux, 2021).

À notre rôle, déjà double, de chercheurs et d'accompagnateurs, s'est ainsi ajouté celui de concepteurs d'un dispositif de formation (Dejaegher et al. 2023).

II. Cadre théorique

Définir le dispositif de formation, une étape importante pour envisager sa conception et penser l'évaluation de sa qualité

La concrétisation de ce projet nécessitait d'abord que nous développions une vision commune du dispositif à concevoir. Nous nous sommes donc posé la question suivante : qu'entend-on par « dispositif de formation » ? Cette question, simple en apparence, soulève pourtant des points de vue différents sur la manière de définir le dispositif, d'envisager sa conception et d'en évaluer la qualité.

Plusieurs définitions du dispositif de formation se côtoient dans la littérature scientifique. L'une d'elles consiste à dire qu'il s'agit d'un « ensemble de moyens (...) agencés en vue de faciliter un processus d'apprentissage » (Blandin, 2002, cité par Bourdet et Leroux, 2009, p.14). Cette définition sous-entend que le dispositif est, en amont de son usage, « 'disposé', installé et prêt à servir » (Bourdet et Leroux, 2009, p.16), ce qui implique une approche de la conception relativement centrée sur les artefacts qui composent le dispositif et sur les apprentissages qu'ils permettent de faciliter.

La qualité du dispositif repose alors principalement sur la pertinence de l'anticipation de l'action des utilisateurs (la manière dont ils vont utiliser les artefacts mis à leur disposition) et sur les résultats produits à la suite de son usage (Bonamy et Charlier, 2003 ; Gerard, 2001 ; Kirkpatrick et Kirkpatrick, 2006). Dans cette perspective, l'évaluation de la qualité prend généralement place à la fin du processus de conception : elle intervient dans une comparaison des « usages observés avec ce qui a été modélisé au moment de la conception » (Bourdet et Leroux, 2009, p.15) et donne éventuellement lieu à des adaptations du dispositif.

Il existe cependant une autre manière de penser le dispositif de formation, du côté de laquelle nous nous sommes rangés. Dans une perspective instrumentale (Béguin et Cerf, 2004 ; Bobillier-Chaumon, 2016 ; Rabardel, 1995 ; Rabardel, 1997), il peut être considéré que les usages que font les utilisateurs du dispositif lui donnent justement son existence propre. Il devient alors un ensemble d'artefacts considéré dans une situation où les utilisateurs, par l'utilisation qu'ils en font, lui donnent une certaine signification. Vu sous cet angle, le dispositif n'est pas un donné, un « figé », mais le fruit d'une élaboration par les sujets.

Définir le dispositif de formation dans une perspective instrumentale implique de dépasser l'approche technocentrique de la conception et de lui préférer une approche anthropocentrique qui accorde aux sujets-utilisateurs l'importance qui leur revient (Béguin et Cerf, 2004 ; Charlier et al., 2006, cités par Cèbe et Goigoux, 2007). Ainsi, la conception d'un dispositif de formation ne peut être menée indépendamment de ses potentiels utilisateurs : ceux-ci peuvent tout autant l'adopter ou le rejeter, le rendre puissant ou inutile et parallèlement, développer des schèmes d'utilisation qui seront propices ou non à l'amélioration de leurs pratiques.

Cette façon de définir le dispositif de formation influence aussi les composantes qui sous-tendent sa qualité et, de ce fait, les critères à observer pour l'évaluer ainsi que la temporalité de cette évaluation.

Validité sociale

Selon Kervyn (2020) la qualité d'une ressource de formation repose également sur sa validité sociale, son adéquation aux besoins exprimés par les différents acteurs du système. Les exigences du métier et les caractéristiques propres des professionnels doivent en effet être étudiées et prises en compte si l'on souhaite favoriser l'utilisation du dispositif et l'appropriation des savoirs scientifiques qu'il véhicule (Dejaegher et al., 2023 ; Kervyn, 2020 ; Kervyn et Goigoux, 2021).

Évaluer la validité sociale d'un dispositif de formation en ligne demande dès lors que l'on vérifie qu'il présente certaines dimensions régulièrement mises en avant dans le monde de l'ergonomie : est-il utilisable, utile et acceptable ? Selon Tricot et al. (2003), ces dimensions peuvent être définies de la manière suivante. L'utilisabilité d'un dispositif fait référence à sa maniabilité, à la facilité avec laquelle il peut « se mettre à la main » de son utilisateur. L'utilité concerne l'alignement des caractéristiques du dispositif (les contenus qu'il expose et les démarches d'apprentissage auxquelles il recourt) aux objectifs d'apprentissage poursuivis et la pertinence de ces objectifs. Enfin, l'acceptabilité indique dans quelle mesure le dispositif est compatible avec les connaissances, les croyances et les pratiques habituelles des utilisateurs ou encore avec leurs valeurs et leurs styles professionnels (Clot et Faïta, 2000).

L'évaluation de ces dimensions repose sur l'observation d'indicateurs tels que ceux repris dans le tableau 1 (Renaud, 2020).

Tableau 1 : Liste de 15 indicateurs permettant de juger de l'utilisabilité, de l'utilité et de l'acceptabilité d'un outil didactique (Renaud, 2020).

Dimensions	Indicateurs
Utilisabilité	1 Facilité pour comprendre l'outil
	2 Confort d'utilisation (simple, maniable)
	3 Charge de travail (durée de la préparation, charge cognitive, ressources attentionnelles)
	4 Flexibilité (modifiable et/ou adaptable par le professeur)
	5 Ajustement (ajusté et/ou ajustable au public d'élèves)
Utilité	6 Pertinence des objectifs poursuivis et/ou redéfinis
	7 Pertinence de la nature et de l'ordre des tâches proposées et des choix de supports
	8 Pertinence de la temporalité de l'enseignement (durée des tâches, séances, séquences)
	9 Apports de l'outil par comparaison avec les autres outils ou techniques disponibles
	10 Constats d'intérêt, d'attention et de motivation des élèves
	11 Constats de progrès des élèves
Acceptabilité	12 Compatibilité avec l'éthique et les valeurs de l'enseignant
	13 Compatibilité avec la prescription (programme, horaires, méthodes)
	14 Compatibilité avec le style pédagogique de l'enseignant, ses démarches habituelles et son organisation pédagogique
	15 Favorable au développement professionnel de l'enseignant (ses propres apprentissages, son intérêt personnel, sa reconnaissance ou valorisation, le sens de son travail)

Ces indicateurs, s'ils concernent plus spécifiquement la conception d'outils didactiques à destination d'enseignants, nous semblent adaptables pour évaluer les dimensions qui nous intéressent dans le contexte de la conception d'un dispositif de formation. Plus loin, nous présenterons la manière dont nous avons adapté ces indicateurs.

Validité scientifique

Une autre composante de la qualité d'un dispositif de formation tient évidemment dans la validité scientifique des contenus qu'il expose, l'enjeu étant de sélectionner les connaissances scientifiques qui représentent pour les destinataires un levier pour leur développement professionnel (Kervyn, 2020 ; Kervyn et Goigoux, 2021) et de se questionner sur la manière de rendre ces connaissances intelligibles pour eux.

Évaluer la validité scientifique revient alors à s'assurer que les contenus présentés durant la formation participent au développement professionnel des enseignants dans une direction qui n'entre pas « en contradiction avec des connaissances scientifiques majeures ou consensuelles » (Kervyn, 2020, p. 2).

Productivité

Enfin, comme nous l'avons déjà mentionné plus haut, la qualité d'un dispositif repose également sur son caractère productif (Kirkpatrick et Kirkpatrick, 2006). Analyser les produits de l'action de formation permet, entre autres, de juger de son efficacité (Bonamy et Charlier, 2003 ; Gerard, 2001). Il s'agit d'observer si les apprentissages réalisés par les enseignants correspondent aux objectifs fixés (efficacité interne), si leurs pratiques tendent progressivement vers celles qui font l'objet de recommandations et si ces changements ont un impact sur les performances de leurs élèves (efficacité externe).

Placer l'évaluation au cœur du processus de conception

Il apparaît ainsi que la qualité d'une ressource de formation repose sur de multiples composantes interdépendantes : validité sociale, validité scientifique et productivité. Le processus visant à évaluer cette qualité est corollairement un processus complexe qui, d'une part, porte sur l'ensemble de ces objets et qui, d'autre part, fait intervenir une pluralité d'acteurs à la fois internes et externes.

Dès lors, comment mener à bien l'évaluation de la qualité du dispositif ? À quel(s) moment(s) du processus de conception recueillir des informations relatives aux différents indicateurs témoignant de la qualité ? À quels

outils de collecte de données recourir pour les questionner ? Comment instaurer un dialogue entre concepteurs et utilisateurs dans le processus d'évaluation ?

Ces questions nous ont amenés à réfléchir non seulement aux caractéristiques finales du dispositif de formation que nous voulions construire, mais avant tout à la façon dont nous pouvions intégrer l'évaluation de sa qualité *dans* la démarche de conception elle-même. En effet, il nous est apparu évident que la temporalité de la mise en œuvre de l'évaluation devait elle aussi être redéfinie : elle ne pouvait être vue comme la dernière étape du processus de conception, mais devait se retrouver inscrite au cœur de ce processus.

Les principes de la démarche de conception continuée dans l'usage, principalement théorisée par Goigoux (Cèbe et Goigoux, 2007 ; Goigoux, 2017), ont fait écho à nos considérations. Dans ce type de démarche, les utilisateurs (dans le cas qui nous concerne, une majorité d'enseignants) sont investis d'un rôle de co-concepteurs (et donc de co-évaluateurs). Le travail de conception n'a pas seulement lieu en amont de l'usage, dans le cercle fermé des chercheurs, mais il prend place *dans* l'usage et intègre les démarches d'évaluation de la qualité du dispositif selon un processus itératif et participatif impliquant les acteurs de terrain (Leblanc, 2004). Plusieurs boucles de conception-évaluation-régulation s'enchaînent alors.

De manière plus précise, la démarche comporte trois phases auxquelles l'évaluation est intégrée (voir figure 1). La première phase concerne la formalisation d'un premier prototype du dispositif basé, entre autres, sur la construction d'un « modèle de l'utilisateur » (Béguin et Cerf, 2004 ; Rabardel et Pastré, 2005) et sur l'état des connaissances scientifiques dans le domaine concerné. Le rôle des chercheurs consiste à identifier le mieux possible l'activité et les besoins des utilisateurs, puis à construire un premier prototype qui adhère au maximum à la modélisation construite et qui soit un compromis acceptable entre les pratiques ordinaires des acteurs et celles qu'ils pourraient adopter avec l'aide du dispositif (Kervyn et Goigoux, 2021). Dès cette première phase, validité sociale et validité scientifique sont ainsi soumises à l'évaluation.

Dans la seconde phase, appelée « co-conception » (Cèbe et Goigoux, 2007), les caractères itératif et participatif de la démarche apparaissent (Goigoux, 2017). Le prototype du dispositif est soumis aux professionnels, qui sont invités à l'utiliser et à en évaluer l'utilisabilité, l'utilité et l'acceptabilité. Leurs commentaires au sujet de ce prototype sont ensuite recueillis et pris en compte pour en produire une nouvelle version, améliorée mais néan-

moins toujours valide sur le plan scientifique. La démarche est alors répétée plusieurs fois. Ce dialogue entre chercheurs et praticiens favorise, au fil des boucles d'évaluation-régulation, l'élaboration de prototypes de plus en plus proches des besoins de ces derniers, scientifiquement pertinents et finalement prêts à être diffusés et validés à plus large échelle.

Enfin, la troisième phase concerne le déploiement du prototype final, son introduction devant un public plus large (celui de l'ensemble des utilisateurs potentiels) et l'évaluation de son caractère productif, de son efficacité.

Figure 1 : Démarche de conception continuée dans l'usage (Cèbe et Goigoux, 2007 ; Goigoux, 2017).



III. Méthodologie

Pour la conception du dispositif de formation en ligne qui nous concerne, nous avons donc décidé de recourir à une démarche de conception continuée dans l'usage. Plusieurs étapes d'élaboration se sont succédé, chacune caractérisée par la construction d'un prototype du dispositif soumis à l'évaluation.

Étapes de la démarche de conception

Le prototype initial

Dans un premier temps, nous avons, par l'analyse de l'activité des professionnels utilisant le programme PARLER (et par comparaison de l'activité

d'enseignants chevronnés à celle de novices), identifié et synthétisé les besoins de formation (Dejaegher et al., 2023). Au départ de cette synthèse, nous avons sélectionné les contenus à présenter (en nous appuyant sur des connaissances scientifiques étayées) et construit une scénarisation pédagogique (articulation des diverses activités de formation) formalisée dans un premier prototype du dispositif de formation.

Ce prototype, existant alors uniquement sur le papier, a été présenté à un groupe de travail interprofessionnel composé de 21 membres exerçant des fonctions diverses dans le système éducatif belge francophone (enseignants, formateurs d'enseignants, conseillers au soutien et à l'accompagnement...). Lors des réunions de ce groupe de travail, nous avons principalement réfléchi à la *validité scientifique* de ce premier prototype et à son *utilité*. Les contenus sélectionnés étaient-ils pertinents sur le plan scientifique ? Répondaient-ils à de véritables besoins de formation ? Proposaient-ils des réponses satisfaisantes ?

Le deuxième prototype

L'étape du second prototype s'est caractérisée par le basculement vers le support numérique. La plateforme de formation a été « mise en ligne » et nous avons, par des laboratoires cognitifs, étudié l'*utilisabilité* de l'environnement numérique conçu : était-il suffisamment maniable, intuitif et facile à prendre en main ? Durant ces laboratoires, cinq sujets ont été observés alors qu'ils utilisaient la plateforme afin d'identifier les éventuelles difficultés de navigation qu'ils rencontraient.

Parallèlement, une phase de test de la plateforme en situation réelle a été menée : 52 sujets (que nous appellerons co-concepteurs) (enseignants, directeurs d'écoles, conseillers au soutien et à l'accompagnement et autres professionnels de l'éducation) ont utilisé la plateforme pour se former au programme PARLER. Systématiquement, ils ont documenté leur expérience d'utilisation dans un outil de récolte prévu à cet effet. Leurs retours nous ont permis de juger de l'*utilisabilité*, mais surtout de l'*utilité* et de l'*acceptabilité* du dispositif : était-il véritablement cohérent avec les besoins des praticiens et favorable à leur développement professionnel ?

Le troisième prototype

Au départ d'un troisième prototype de la plateforme de formation, l'expérimentation a été répliquée avec un groupe de 28 sujets (dont certains étaient déjà présents lors du test précédent). Les dimensions évaluées pour le deuxième prototype ont été à nouveau soumises à l'évaluation et nous avons pu identifier dans quelle mesure les adaptations apportées répondaient efficacement aux observations soulevées précédemment.

Le prototype final

Parvenu à sa version définitive, une autre composante de la qualité du dispositif de formation doit à présent être soumise à l'évaluation : il s'agit de sa capacité à produire les effets souhaités sur des enseignants et, par la suite, sur leurs élèves (Cèbe et Goigoux, 2007 ; Goigoux, 2017). Par-là, nous souhaitons observer si les contenus de formation, portant particulièrement sur les recommandations de la recherche en matière d'enseignement de la lecture, ont bel et bien été intégrés par les enseignants qui suivent la formation. Nous verrons également si ces contenus ont une influence sur leurs croyances, leurs connaissances et leurs pratiques d'enseignement dans ce domaine et si des effets sur leurs élèves peuvent être observés. Cette étape du projet implique l'élaboration et l'exécution d'une méthodologie orientée vers l'évaluation du caractère productif du dispositif (Bonamy et Charlier, 2003 ; Kirkpatrick et Kirkpatrick, 2006) : la conduite d'observations en classe, d'entretiens réflexifs, de tests destinés à mesurer l'évolution des compétences des élèves...

Cette dernière étape étant toujours en cours de réalisation, nous nous contenterons de rendre compte de trois moments d'évaluation entrant dans la deuxième phase de la démarche de conception continuée dans l'usage (voir figure 1) et portant principalement sur la validité sociale du dispositif. Pour mieux les décrire, tous les outils auxquels nous avons eu recours pour collecter les informations nécessaires à la mise en œuvre du processus évaluatif (voir tableau 2) sont listés ci-après.

Outils de collecte de données

Les réunions du groupe de travail interprofessionnel. Ces réunions, principalement orientées vers l'évaluation de la *validité scientifique* du dispositif et de son *utilité*, portaient la plupart du temps sur la conception de storyboards, de capsules vidéo et d'activités de formation. Un procès-verbal rédigé pour chacune de ces réunions a permis de garder trace des réflexions et des échanges.

Les laboratoires cognitifs. Cet outil, utilisé pour étudier les processus mentaux de sujets accomplissant une tâche spécifique (Zucker et al., 2004), a été utilisé pour évaluer l'*utilisabilité* de la plateforme de formation mise en ligne. Positionnés devant leur ordinateur, les participants ont dû réaliser différentes tâches sur la plateforme et verbaliser leur raisonnement et les actions qu'ils menaient pour les effectuer. Nous avons ainsi pu observer la manière dont ils naviguaient dans l'interface numérique, comment ils procédaient pour retrouver des informations spécifiques et dans quelle mesure ils comprenaient la structure et le fonctionnement de la plateforme. Des enregistrements vocaux et des enregistrements d'écran nous ont permis de conserver les données issues de ces laboratoires.

Les journaux de bord. Durant les deux phases de test en situation réelle (deuxième et troisième prototypes), les co-concepteurs ont été invités à consigner leur expérience d'utilisation de la plateforme dans des « journaux de bord » construits dans le but de les amener à poser un regard réflexif sur cette expérience. Ces documents reprenaient une série de questions portant particulièrement sur l'*utilisabilité* et l'*utilité* de l'outil et des activités de formation.

Les travaux intégrateurs. Parallèlement aux journaux de bord, il a été demandé aux co-concepteurs de réaliser, en fin de parcours de formation, un rapport intégrant quatre axes de réflexion : les découvertes et les apprentissages réalisés, le transfert de ces apprentissages dans leur pratique professionnelle, leur relation avec la formation en ligne et la synthèse de leurs commentaires à propos des activités de formation (régulièrement consignées dans leurs journaux de bord). Au-delà de l'*utilisabilité* et de l'*utilité* du dispositif de formation, cet outil nous a permis de collecter des données relatives à son *acceptabilité*.

Tableau 2 : Outils de collecte de données utilisés pour évaluer les différentes composantes de la qualité du dispositif de formation (aux différentes étapes de sa conception).

	VALIDITÉ SOCIALE			VALIDITÉ SCIENTIFIQUE	PRODUCTIVITÉ
	Utilisabilité	Utilité	Acceptabilité		
Prototype initial (version papier)		Réunions du groupe de travail interprofessionnel		Réunions du groupe de travail interprofessionnel	
Deuxième prototype (version en ligne)	<ul style="list-style-type: none">• Laboratoires cognitifs• Journaux de bord• Travaux intégrateurs	<ul style="list-style-type: none">• Réunions du groupe de travail interprofessionnel• Journaux debord• Travaux intégrateurs	Travaux intégrateurs	Réunions du groupe de travail interprofessionnel	
Troisième prototype (version en ligne)	<ul style="list-style-type: none">• Journaux de bord• Travaux intégrateurs	<ul style="list-style-type: none">• Réunions du groupe de travail interprofessionnel• Journaux debord• Travaux intégrateurs	Travaux intégrateurs	Réunions du groupe de travail interprofessionnel	
Prototype final (version en ligne)					(Évaluation en cours)

Outil d'analyse des données

Les données récoltées via les outils de collecte listés plus haut ont été systématiquement rassemblées dans un tableau de synthèse. Pour analyser et utiliser ces informations en vue d'améliorer les prototypes de la plateforme de formation, nous avons emprunté à Renaud (2020) la liste

de 15 indicateurs relatifs aux dimensions d'utilisabilité, d'utilité et d'acceptabilité et les avons adaptés au contexte de la formation.

Ces indicateurs nous ont aidés à identifier, prototype après prototype, les caractéristiques du dispositif qu'il était nécessaire d'améliorer pour augmenter sa validité sociale. Parallèlement, tout au long du processus, une attention constante a été portée à la validité scientifique du dispositif, validité dont les membres du groupe de travail interprofessionnel (dont nous faisons partie) se sont montrés garants.

Ci-dessous, nous présentons de manière succincte et en guise d'exemples quelques résultats issus des différents moments d'évaluation.

IV. Résultats

Conception et évaluation du prototype initial

Ayant identifié les besoins de formation auxquels nous souhaitons répondre (notamment par une analyse approfondie des difficultés rencontrées par plusieurs professionnels mettant en œuvre le programme PARLER dans leurs classes), nous avons conçu un premier prototype du dispositif composé de six modules de formation. Les deux premiers modules présentaient les fondements du programme PARLER et ses principes organisateurs. Les quatre modules suivants étaient chacun consacrés à une habileté langagière ciblée par le programme. Soumis à l'évaluation lors des réunions du groupe de travail interprofessionnel, ce prototype initial présentait plusieurs faiblesses.

Tableau 3 : Évaluation de l'utilité du prototype initial du dispositif de formation (adapté de Renaud (2020)).

Dimensions	Indicateurs	
Utilité	6 Pertinence des objectifs poursuivis et/ou redéfinis	●
	7 Pertinence de la nature et de l'ordre des tâches proposées et des choix de supports	●
	8 Pertinence de la temporalité de l'enseignement (durée des tâches, séances, séquences)	●

D’abord, les storyboards de certaines capsules vidéo théoriques présentaient des contenus inaccessibles pour le public visé (voir tableau 3, indicateurs 6 ● et 7 ●), en ce sens qu’ils ne répondaient pas suffisamment aux préoccupations premières des enseignants : identifier précisément les habiletés considérées comme les piliers de la lecture, comprendre les raisons de les enseigner et identifier les gestes professionnels permettant d’y parvenir. Certaines activités de formation n’étaient pas suffisamment fonctionnelles pour la classe et les contenus des différents modules n’étaient pas assez spécifiques aux thématiques censées être abordées par chacun d’eux. Les capsules vidéo paraissaient également beaucoup trop longues (voir tableau 3, indicateur 8 ●).

Nous avons donc procédé aux changements nécessaires et intégré ces observations dans un deuxième prototype du dispositif.

Conception et évaluation du deuxième prototype

Pour l’évaluation de la qualité de ce prototype, les différentes activités de formation et leurs contenus ont été mis en ligne. Le passage vers l’environnement numérique nous a permis de mener à bien les laboratoires cognitifs nécessaires à l’obtention d’informations relatives à l’utilisabilité du dispositif. Ces laboratoires nous ont notamment permis d’identifier différents problèmes liés à l’ergonomie de l’interface.

Figure 2 : Exemple de donnée issue d'un laboratoire cognitif.

Martin pointe le manque d'éléments visuels permettant de comprendre que le menu de droite et la barre de progression présentent des informations parallèles. Au premier abord, il n'a pas compris que le menu de droite est une autre manière de visualiser la barre de progression. « Ce n'est pas intuitif, il faut le savoir », dit-il.

Proposition d'amélioration : remettre les titres des séquences sur les différentes parties de la barre de progression, même si l'on ne verra pas la différence dans les niveaux de titres.



Parallèlement à la mise en œuvre de ces laboratoires, les journaux de bord et les travaux intégrateurs réalisés par les co-concepteurs de la plateforme nous ont amenés à identifier d'autres problèmes dans le prototype.

Tableau 4 : Évaluation de l'utilisabilité, de l'utilité et de l'acceptabilité du deuxième prototype du dispositif de formation (adapté de Renaud (2020)).

Dimensions	Indicateurs	
Utilisabilité	1 Facilité pour comprendre l'outil	●
	2 Confort d'utilisation (simple, maniable)	●
	3 Charge de travail (durée de la préparation, charge cognitive, ressources attentionnelles)	●
	4 Flexibilité (modifiable et/ou adaptable par le professeur)	●
	5 Ajustement (ajusté et/ou ajustable au public d'élèves)	●
Utilité	6 Pertinence des objectifs poursuivis et/ou redéfinis	●
	7 Pertinence de la nature et de l'ordre des tâches proposées et des choix de supports	●
	8 Pertinence de la temporalité de l'enseignement (durée des tâches, séances, séquences)	●
	9 Apports de l'outil par comparaison avec les autres outils ou techniques disponibles	●
	10 Constats d'intérêt, d'attention et de motivation des élèves	●
	11 Constats de progrès des élèves	●
Acceptabilité	12 Compatibilité avec l'éthique et les valeurs de l'enseignant	●
	13 Compatibilité avec la prescription (programme, horaires, méthodes)	●
	14 Compatibilité avec le style pédagogique de l'enseignant, ses démarches habituelles et son organisation pédagogique	●
	15 Favorable au développement professionnel de l'enseignant (ses propres apprentissages, son intérêt personnel, sa reconnaissance ou valorisation, le sens de son travail)	●

Pour les modules 3 et 4, certaines capsules vidéo théoriques restaient trop complexes et trop longues (voir tableau 4, indicateurs 7 ● et 8 ●). Les activités de formation destinées à soutenir le transfert des apprentissages réalisés par les apprenants dans leurs classes ne permettaient pas de déterminer suffisamment facilement les gestes pédagogiques essentiels à mobiliser pour mener à bien les ateliers du programme (voir tableau 4, indicateur 6 ●). Enfin, la découverte du troisième module représentait pour les usagers une charge de travail trop importante et disproportionnée en regard des autres modules (voir tableau 4, indicateur 3 ●).

Un travail d'adaptation des modules concernés a donc eu lieu.

Conception et évaluation du troisième prototype

À cette étape, les données récoltées pour l'évaluation de ce prototype nous ont indiqué que les modifications intégrées à cette version de la plateforme participaient à l'amélioration globale de sa qualité. En effet, à ce stade, l'ensemble des critères d'utilisabilité et d'utilité semblaient être rencontrés.

Figure 3 : Exemples de données issues des journaux de bord et travaux intégrateurs.

« À ma grande surprise, accéder à la plateforme est un jeu d'enfant et à la portée de tous. Et l'utilisation est très facile, il suffit de suivre les instructions. Je ne rencontre aucune difficulté lors de la navigation au sein des différents modules. C'est simple d'utilisation. L'enchaînement des activités de la formation suit une logique assez juste. »

« Toutes les activités reprises dans les différents modules me permettent d'atteindre les objectifs fixés. »

« Ce module m'a fait prendre conscience qu'il était nécessaire de travailler, dans un premier temps, la compréhension séparément du décodage afin de soulager les élèves.[...] Les enseignants - et moi aussi avant d'ailleurs - ont trop tendance à attirer l'attention sur le résultat obtenu alors que l'essentiel, c'est la démarche mise en œuvre pour arriver à l'objectif. »

« J'ai osé mettre en place de nouveaux processus structurels engageant l'autonomie des enfants. »

Tableau 5 : Évaluation de l'utilisabilité, de l'utilité et de l'acceptabilité du troisième prototype du dispositif de formation (adapté de Renaud (2020)).

Dimensions	Indicateurs	
Utilisabilité	1 Facilité pour comprendre l'outil	●
	2 Confort d'utilisation (simple, maniable)	●
	3 Charge de travail (durée de la préparation, charge cognitive, ressources attentionnelles)	●
	4 Flexibilité (modifiable et/ou adaptable par le professeur)	●
	5 Ajustement (ajusté et/ou ajustable au public d'élèves)	●
Utilité	6 Pertinence des objectifs poursuivis et/ou redéfinis	●
	7 Pertinence de la nature et de l'ordre des tâches proposées et des choix de supports	●
	8 Pertinence de la temporalité de l'enseignement (durée des tâches, séances, séquences)	●
	9 Apports de l'outil par comparaison avec les autres outils ou techniques disponibles	●
	10 Constats d'intérêt, d'attention et de motivation des élèves	●
	11 Constats de progrès des élèves	●
Acceptabilité	12 Compatibilité avec l'éthique et les valeurs de l'enseignant	●
	13 Compatibilité avec la prescription (programme, horaires, méthodes)	●
	14 Compatibilité avec le style pédagogique de l'enseignant, ses démarches habituelles et son organisation pédagogique	●
	15 Favorable au développement professionnel de l'enseignant (ses propres apprentissages, son intérêt personnel, sa reconnaissance ou valorisation, le sens de son travail)	●

V. Discussion

Si l'observation du tableau 5 semble indiquer que nous sommes parvenus à concevoir un dispositif qui rencontre les critères de qualité attendus en termes de validité sociale, certaines réflexions méritent encore d'être approfondies.

L'évaluation de la qualité du dispositif de formation nous a conduit à recourir à des outils de collecte de données pluriels faisant émerger de très

nombreuses informations. Articuler de manière cohérente les données qui en découlait est alors devenu une préoccupation majeure de notre travail. Comment gérer la diversité des informations récoltées, les messages qu'elles véhiculaient et les points de vue qu'elles représentaient ? Que faire quand ces informations étaient en tension ? Quelle valeur accorder aux différentes composantes évaluées dans la prise de décisions relatives à la conception du dispositif ? La validité sociale devait-elle l'emporter sur la validité scientifique ? À l'inverse, fallait-il toujours se ranger derrière l'exigence de la recherche au risque de voir les utilisateurs délaisser le dispositif de formation qui leur était proposé ? En d'autres termes, comment, en tant que concepteurs, concilier notre rôle de chercheurs, chargés de véhiculer des recommandations scientifiques, et celui d'accompagnateurs, occupés à soutenir la mise en œuvre de ces recommandations dans les classes ?

Pour tenter de répondre à ces questions, revenons d'abord sur la notion de validité scientifique. Comme le souligne Kervyn (2020, p. 2), « il est (...) nécessaire qu'aux yeux des chercheurs, les ressources produites (...) ne soient pas en contradiction avec des connaissances scientifiques majeures ou consensuelles ». Il nous a donc incombé, en tant que chercheurs, de nous porter garants de cette validité dans la succession des prototypes soumis à l'évaluation. Garantir la validité scientifique a parfois été difficile, notamment lorsque nous cherchions à améliorer l'acceptabilité du dispositif de formation. Parmi les indicateurs utilisés pour évaluer celle-ci, il convenait de nous demander s'il était vraiment nécessaire, voire possible, de chercher la comptabilité avec les styles pédagogiques des formés (voir tableau 5, indicateur 14 ●), ceux-là étant, par définition, multiples et personnels (Clot et Faïta, 2000). Postulant que certaines pratiques recommandées par la recherche, si elles semblaient de prime abord en tension avec ces styles professionnels, pouvaient justement être à l'origine de trajectoires de développement professionnel (Dejaegher et al., 2023), nous avons parfois décidé de « privilégier » la cohérence scientifique du dispositif, invitant les acteurs de terrain à partager avec nous la volonté de respecter cette rigueur et à « développer une culture scientifique commune » (Kervyn, 2020, p. 3). Concrètement, nous avons alors veillé à fournir aux utilisateurs des aides supplémentaires leur permettant de dépasser ces tensions : témoignages, activités d'analyse réflexive, ressources scientifiques vulgarisées...

Enfin, notons qu'une attention particulière portée aux critères de flexibilité et d'ajustement (voir tableau 5, indicateurs 4 et 5) pourrait permettre d'aboutir à un dispositif capable de gérer l'hétérogénéité des destinataires, notamment par la possibilité offerte aux apprenants d'emprunter

différentes voies dans le parcours de formation, de consulter des ressources supplémentaires ou d'en ignorer certaines, de tisser des liens clairs avec leurs pratiques personnelles, d'interagir avec les autres formés...

VI. Conclusion

Nous l'avons vu, la définition que l'on donne du dispositif de formation influence la manière d'envisager sa conception et les composantes qui en sous-tendent la qualité. Le processus d'évaluation de celle-ci s'en trouve alors lui-même impacté, tant dans sa forme (quelles composantes observer, quels outils de collecte de données utiliser et quels acteurs faire intervenir ?) que dans sa temporalité (à quel(s) moment(s) du processus de conception faire intervenir l'évaluation : en fin de processus, en cours de processus ?).

Dans notre cas, nous avons choisi d'envisager le dispositif de formation non comme un « donné » mais comme le fruit d'une élaboration par ses utilisateurs (Béguin et Cerf, 2004 ; Bobillier-Chaumon, 2016 ; Rabardel, 1995 ; Rabardel, 1997). Choisir cette définition implique de reconsidérer plusieurs composantes pour définir la qualité du dispositif : sa validité sur le plan scientifique, sa validité sociale (Kervyn, 2020) et son caractère productif (Bonamy et Charlier, 2003 ; Kirkpatrick et Kirkpatrick, 2006). Ces considérations nous ont menés à envisager le processus évaluatif de la qualité comme un enchaînement complexe de démarches poursuivant de multiples finalités : garantir la cohérence scientifique des contenus du dispositif, assurer les conditions d'une rencontre réussie avec les utilisateurs et analyser les produits de la formation en termes d'efficacité. Il était également nécessaire que nous fassions de l'évaluation un « lieu » de partage entre les différents acteurs impliqués : les concepteurs et les utilisateurs.

Ainsi, nous avons cherché à construire un dispositif qui :

- permettrait de véhiculer des savoirs scientifiques utiles pour les praticiens et les rendrait abordables pour ceux-ci (Kervyn, 2020 ; Kervyn et Goigoux, 2021) ;
- serait utilisable, utile et acceptable (Tricot et al., 2003) ;
- produirait sur ses destinataires les effets escomptés (atteinte des objectifs fixés, intégration de recommandations scientifiques dans

leurs pratiques et développement efficace des compétences de leurs élèves).

Le recours à une démarche de conception continuée dans l'usage (Goigoux, 2017 ; Leblanc, 2004) nous a aidés à mener à bien ce processus évaluatif complexe. S'inscrivant dans l'usage, la temporalité de cette démarche particulière nous a permis de rompre avec un schéma classique dans lequel l'évaluation de la qualité du dispositif succède aux étapes de son élaboration. Au contraire, par cette démarche, nous avons pu intégrer l'évaluation dans le processus même de conception et développer la collaboration entre les chercheurs et les praticiens, favorisant la construction de prototypes aussi pertinents que possible pour les espaces de la recherche et les contextes de terrain.

Nous nous apprêtons désormais à entrer dans une phase d'évaluation à plus grande échelle du caractère productif de notre dispositif. Quels effets a-t-il sur les apprentissages d'enseignants qui s'en emparent ? Les aide-t-il à s'approprier les pratiques recommandées pour mettre efficacement en œuvre le programme PARLER ? Contribue-t-il finalement à l'amélioration des performances de leurs élèves ?

Bien que placée en fin de démarche, il nous paraît déjà évident que cette facette de l'évaluation ne clôturera pas le processus de conception mais qu'elle sera un instrument de plus pour nous aider à élaborer un dispositif de formation de qualité.

VII. Bibliographie

- Béguin, P. et Cerf, M. (2004). Formes et enjeux de l'analyse de l'activité pour la conception des systèmes de travail. *Activités*, 1(1), 54-71. doi : 10.4000/activites.1156
- Bobillier-Chaumon, M.-É. (2016). L'acceptation située des technologies dans et par l'activité : premiers étayages pour une clinique de l'usage. *Psychologie du Travail et des Organisations*, 22(1), 4-21. doi : 10.1016/j.pto.2016.01.001
- Bonamy, J. et Charlier, B. (2003). Un dispositif efficace ? Dans B. Charlier et D. Peraya (dir.), *Technologie et innovation en pédagogie : Dispositifs innovants de formation pour l'enseignement supérieur* (p. 181-195). De Boeck.

- Bourdet, J.-F. et Leroux, P. (2009). Dispositifs de formation en ligne : de leur analyse à leur appropriation. *Distances et savoirs*, 1(7), 11-29. doi : 10670/1.g9iyqi
- Cèbe, S. et Goigoux, R. (2007). Concevoir un instrument didactique pour améliorer l'enseignement de la compréhension de textes. *Repères, Recherches en Didactique du Français Langue Maternelle*, 35, 185-208. doi : 10.3406/reper.2007.2756
- Charlier, B., Deschryver, N. et Peraya, D. (2006). Apprendre en présence et à distance : Une définition des dispositifs hybrides. *Distances et savoirs*, 4(4), 469-496. doi : 10.3166/ds.4.469-496
- Clot, Y. et Faïta, D. (2000). Genres et styles en analyse du travail : concepts et méthodes. *Travailler*, 4(4), 7-43.
- Dejaegher, C., André, M., Leyh, O. et Schillings, P. (2023). Concilier fiabilité scientifique et pertinence sociale : retour sur l'expérimentation belge du programme P.A.R.L.E.R. Dans B. Kervyn, M. Lebrun, V. Marmy Cusin et C. Scheepers (dir.), *Questionner l'articulation entre théories et pratiques en didactique du français* (p.129-150). Presses universitaires de Namur.
- Dejaegher, C., Lafontaine, D., André, M., Leyh, O., Rappe, J. et Schillings, P. (2022). Expérimentation du programme PARLER en Belgique francophone. *Revue Française de Pédagogie*, 216, 53-69. doi : 10.4000/rfp.12056
- Gerard, F.-M. (2001). L'évaluation de la qualité des systèmes de formation. *Mesure et évaluation en éducation*, 24(2-3), 53-77. doi : 10.7202/1091170ar
- Goigoux, R. (2017). Associer chercheurs et praticiens à la conception d'outils didactiques ou de dispositifs innovants pour améliorer l'enseignement. *Éducation et Didactique*, 11(3), 135-142. doi : 10.4000/educationdidactique.2872
- Kervyn, B. (2020). De l'utilité de la recherche collaborative pour produire des ressources de formation robustes. *Recherches en didactique des langues et des cultures*, 17(2), 1-15. doi : 10.4000/rdlc.7339
- Kervyn, B. et Goigoux, R. (2021). Produire des ressources didactiques : une modalité originale de vulgarisation scientifique tournée vers le développement professionnel et créatrice de nouveaux savoirs scientifiques. *Repères, Recherches en Didactique du Français Langue Maternelle*, 63, 185-210. doi : 10.4000/reperes.4253

- Kirkpatrick, D. L. et Kirkpatrick, J. D. (2006). *Evaluating Training Program: The Four Levels*. Berrett-Koehler.
- Leblanc, S. (octobre 2004). Conception et évaluation de dispositifs de formation innovants à partir de l'analyse des usages in situ [Actes de congrès]. *Technologies de l'Information et de la Connaissance dans l'Enseignement Supérieur et de l'Industrie*, Compiègne, France.
- Rabardel, P. (1995). *Qu'est-ce qu'un instrument ? Appropriation, conceptualisation, mises en situation*. Centre National de Documentation Pédagogique.
- Rabardel, P. (1997). *Les hommes et les technologies : Approche cognitive des instruments contemporains*. Armand Colin.
- Rabardel, P. et Pastré, P. (2005). *Modèles du sujet pour la conception*. Octarès Éditions.
- Renaud, J. (2020). Évaluer l'utilisabilité, l'utilité et l'acceptabilité d'un outil didactique au cours du processus de conception continue dans l'usage : Cas d'un outil pour l'enseignement de la lecture de textes documentaires numériques. *Éducation et didactique*, 14(2), 65-84. doi : 10.4000/educationdidactique.6756
- Tricot, A., Plégat-Soutjis, F., Camps, J.-F., Amiel, A., Lutz, G. et Morcillo, A. (15, 16 et 17 avril 2003). Utilité, utilisabilité, acceptabilité : interpréter les relations entre trois dimensions de l'évaluation des EIAH [Actes de congrès]. *Environnements Informatiques pour l'Apprentissage Humain*, Strasbourg, France.
- Zorman, M., Bressoux, P., Bianco, M., Lequette, C., Pouget, G. et Pourchet, M. (2015). « PARLER » : un dispositif pour prévenir les difficultés scolaires. *Revue Française de Pédagogie*, 4, 57-76. doi : 10.4000/rfp.4890
- Zucker, S., Sassman, C. et Case, B. J. (2004). *Cognitive Labs* [Rapport technique]. Pearson.

Chapitre 9. Évaluation des contextes et contextes d'évaluation dans la recherche-intervention

Vers la théorisation d'un tiers-espace de l'évaluation

Jean-François Marcel

I. Introduction

Ce chapitre investit le thème « Pluralité des contextes, pluralité des évaluations » à partir de l'évaluation d'une démarche de recherche participative (la « recherche-intervention ») qui se déploie dans différents contextes, investit plusieurs niveaux de contextes et mobilise plusieurs modèles d'évaluation. Cette démarche met au jour un certain nombre d'enjeux pour l'évaluation (par la proposition d'un tiers-espace de l'évaluation) pour les sciences de l'éducation et de la formation (en lien avec des contextes, mobilisés et investis). Ce texte se donne pour objectif un partage de réflexions sur l'évaluation (en particulier au travers des contextes) d'une démarche de recherche participative en SEF.

Pour ce faire, il se structure en deux parties. La première présente la recherche-intervention (RI), elle la définit, la caractérise, puis met une focale sur son principe fondateur : la participation. La seconde partie cible l'évaluation de la RI sur la base du choix d'une approche écologique. Dans le cadre de ce chantier en cours sont présentés trois outils d'évaluation. En s'appuyant sur l'ensemble du chapitre, la conclusion amorce une proposition de théorisation d'un tiers-espace de l'évaluation de la RI.

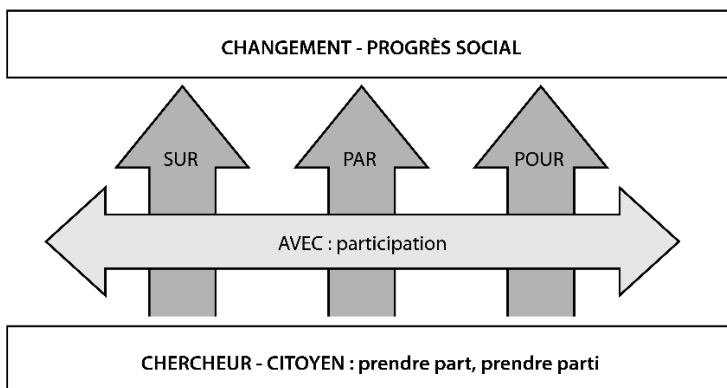
II. Une démarche de recherche participative : la recherche-intervention (RI)

Pour débiter, nous présenterons la RI comme une démarche d'accompagnement du changement au sein d'organisations sociales et professionnelles dans le champ de l'enseignement, la formation, l'animation, le secteur social, médico-social, etc. (Marcel et Broussal, 2022).

Définir la RI

Il s'agit d'une démarche de recherche qui poursuit trois visées interdépendantes : une visée heuristique (une RI « sur ») qui cible la production de savoirs scientifiques, une visée transformative (une RI « pour ») qui assume de contribuer au changement en partant d'une demande de l'organisation et une troisième, critique (une RI « par ») qui est vigilante aux conditions de l'émancipation de tous les participants. La posture du chercheur-citoyen (Marcel et Nunez Moscoso, 2012) porte cet engagement dans le champ social (il « prend part » et assume sa contribution au changement). Elle est basée sur une dialectique entre intériorité (induite par le contexte de l'intervention) et extériorité (maintenue par la visée généralisante de la recherche), dialectique que matérialise le trait d'union.

La RI est régie par le principe fondateur de la participation (le « avec ») de tous les acteurs à toutes les phases du projet, sans nier l'hétérogénéité de leurs contributions, mais en respectant la légitimité de tous les savoirs mis en jeu. Elle est donc caractérisée par la convocation simultanée et nécessaire du « sur », du « pour », du « par » et du « avec » comme le résume la figure suivante.

Figure 1 : La démarche de RI

Pour rendre plus concrète cette définition, nous listons ci-dessous quelques contextes au sein desquels nous avons déployé des RI :

- Pour l'enseignement : un établissement agricole (élaboration d'un « projet vie scolaire »), un réseau de lycées professionnels (projet visant le « pilotage intermédiaire » de l'établissement), un collège (dans le cadre d'un déménagement en lien avec sa rénovation),
- Pour la formation : l'école des sapeurs-pompiers (projet visant l'évaluation de leur offre de formation, puis de dispositifs de formation en situation de travail), l'INSPE (projet visant la formation des tuteurs de stage des Conseillers Principaux d'Éducation),
- Pour le médico-social : des établissements et services d'aide par le travail (des ESAT, fabrication de ressources professionnelles pour « outiller » les moniteurs d'ateliers), une association (co-élaboration d'un observatoire de la participation), des établissements d'hébergement pour personnes âgées dépendantes EHPAD (évaluation de l'intervention de clowns auprès des résidents ou, plus récemment, projet visant l'inclusion des EHPAD), une entreprise spécialisée dans l'autisme (soutien aux professionnels à partir de dispositifs d'analyse des pratiques),
- Pour les collectivités territoriales : une Municipalité (co-construction d'un observatoire du « vivre sa jeunesse »), une Métropole (dispositifs de prévention relatifs à la santé et la sécurité au travail des agents de

collecte des déchets), un centre culturel (accompagnement d'expositions dans le cadre d'une rénovation du bâtiment),

- Pour la culture : un groupe de poètes occitans (co-élaboration sur la base d'un dialogue poésie-recherche autour de 6 concepts sur le Pays d'Oc).

Caractérisation de la RI

Pour préciser cette définition, nous développerons trois caractéristiques fortes de la RI : le principe de son démarrage (avec la stabilisation de la commande), ses trois visées, les savoirs produits par l'ensemble de la démarche et les trois ancrages qui fondent sa théorisation.

Demande et commande

Pour retracer la chronologie d'une RI, la première étape, fondatrice, est celle du « passage » entre la demande (unilatérale) et la commande (partenariale). En effet, elle débute par une rencontre entre une Unité de Recherche et une « entreprise » (établissement, association, collectivité territoriale, etc.) porteuse d'une demande de changement (souvent assez vague, elle peut être, pour partie, suscitée). S'en suivent d'autres rencontres, basées sur des négociations (portant sur des volets divers : les attendus, les contraintes, les ressources, les finances, les temporalités, la faisabilité, etc.) permettant d'aboutir à la construction d'un accord pour un engagement dans un projet partagé. C'est la formalisation de ce projet qui permet une contractualisation désignée par le terme de commande.

La commande, qui constitue en fait le premier résultat de la RI puisqu'elle concrétise son existence, est très importante. D'une part, elle explicite la contribution de la recherche par rapport au projet partagé et inscrit donc ainsi la légitimité de l'Unité de Recherche dans la RI. Elle permet ensuite d'amorcer deux problématisations interdépendantes : une problématisation que nous qualifierons de socioprofessionnelle (qui correspond à l'effectivité du changement de la RI « pour ») et la problématisation scientifique « traditionnelle » (qui renvoie à la production de savoir de la RI « sur »). Enfin, la commande constitue le cadre contractuel de la RI. En cela, elle explicite les attendus et les conditions de sa mise en œuvre (instances, modalité de participation), elle protège les acteurs (et sert de référence en cas de dysfonctionnement), mais est dotée d'une certaine plasticité qui permet d'adapter le projet aux contingences de la mise en œuvre.

Les trois visées ²¹

La RI poursuit deux visées principales non hiérarchisées. La visée praxéologique concerne la production de savoirs contextualisés et des ressources « pour » l'action (et « avec » les acteurs, en se basant sur leurs connaissances) en lien avec les attendus principaux de la demande traduits dans la commande. La visée heuristique concerne la production de savoirs scientifiques « sur », l'objet étudié en effectuant une « montée en généralité » à partir des savoirs contextualisés, co-élaborés en fonction de l'avancée du projet. Cette visée est en lien avec les repères posés par la commande alors que, généralement, ils sont moins présents dans la phase de demande.

Ces deux visées sont interdépendantes mais leur articulation requiert toutefois trois conditions importantes. La première est qu'il n'y ait ni substitution ni confusion entre une logique de connaissance et une logique d'action (chaque visée préserve une autonomie relative) ; la seconde est qu'il est important de prévenir toute instrumentalisation de l'une par l'autre en préservant rigoureusement les deux visées (connaître et transformer) ce qui peut se traduire par l'élaboration de dispositifs prévoyant des espace-temps de dissociation et/ou d'association. La troisième condition relève d'une double vigilance, celle de mobiliser le principe de participation et de préserver la visée critique (nous allons y revenir). Ces modalités d'articulation spécifiques justifient le lien entre les deux sphères de la recherche et de l'intervention, traduite, dans le nom de RI par le trait d'union.

La RI poursuit une troisième visée, en liens étroits avec les deux précédentes, la visée critique qui se décline en deux objectifs. Le premier est de contribuer, au travers d'un projet de changement partagé par un collectif hétérogène, à l'émancipation de tous les participants. Dans le prolongement de Freire (voir Marcel, 2020b ou Aussel, 2017), l'émancipation est entendue ici comme un « déplacement » entre la place à laquelle on se trouve assigné et celle à laquelle on aspire, que cette place relève du champ professionnel, social, personnel, etc. Si ce premier volet pose les fondements du « prendre part » au progrès social, le second volet cible plutôt les conditions du « prendre part », en termes d'exigence et de vigilance critique. Il investit la sphère de la responsabilité du chercheur : il

²¹ Un schéma récapitulatif est présenté à la suite de la partie évaluation.

s'agit d'assumer notre contribution au changement, en tant que chercheurs, ce qui requiert une vigilance déontologique et éthique.

Les savoirs produits

En France, plus particulièrement, les recherches participatives sont régulièrement accusées de ne pas produire de savoirs scientifiques. Notre proposition de RI s'inscrit radicalement en faux par rapport à cette suspicion (voir Marcel 2020a).

En effet, à partir des connaissances des professionnels et des méthodologies de la recherche à dominante qualitative, la RI permet d'abord la co-élaboration de savoirs contextualisés à l'organisation au sein de laquelle est déployée la RI. Ces savoirs contextualisés sont à la fois utiles pour connaître (car dotés d'une forte validité scientifique interne) et utiles pour agir, pour élaborer des ressources (car dotés d'une validité socio-professionnelle). Dans un second temps, la RI prend en charge l'étape allant des savoirs contextualisés vers des savoirs scientifiques. Elle assume une montée en généralité en respectant les exigences de la validité scientifique externe (pour les démarches qualitatives qui visent la généralisation des processus étudiés plutôt que celle des résultats).

Les trois ancrages de la RI

Notre théorisation de la RI s'appuie sur trois catégories d'ancrages que nous décrivons ici rapidement.

Le premier ancrage est de nature philosophique et mobilise les travaux de Freire sur l'émancipation que nous avons déjà évoqués mais surtout sur son concept d'incomplétude de l'être humain, ses aspirations à « être plus » et à dépasser les « situations limites ». Cet ancrage est étayé par l'apport des épistémologies du Sud qui rejettent un savoir académique à prétention dominante et revendiquent la légitimité des savoirs profanes, professionnels ou traditionnels (de Sousa Santos, 2011). Enfin, le concept de décoïncidence²², emprunté à Jullien (voir Marcel, 2023) donne toute sa

²² Jullien définit la décoïncidence « par écart d'avec l'adéquation précédente et s'ouvrant par insatisfaction de ce que la normalité a muré – d'où sa fissure du dedans d'elle-même qui, à son insu, s'étiolait – se découvrent de nouveaux possibles, jusqu'ici inenvisagés » (Jullien, 2020, p. 49). Il précise « c'est du fait de décoïncider – de défaire continuellement la coïncidence acquise – que procède la possibilité

signification au trait d'union de recherche-intervention. Il justifie l'interrelation entre les deux sphères comme génératrice de leur potentialité d'invention.

Le deuxième ancrage est de nature épistémologique. Face à la puissance de deux traditions philosophiques, l'une différencialiste posant comme indépassables la séparation et le clivage entre la connaissance et l'action et l'autre, anti-différencialiste, insistant au contraire sur leurs interconnexions permanentes, sur les passages fréquents entre les deux (Robert et Marcel, 2019), la RI s'adosse à l'ouverture d'une tierce voie. Elle défend même la double hypothèse d'un rapprochement possible entre recherche et action (ici l'intervention), mais aussi d'un enrichissement réciproque de chacune de ces sphères par leur mise en relation et par le travail consenti pour leur articulation.

Le troisième ancrage est de nature politique, entendue comme dimension politique de la recherche. Elle concerne la place de la recherche dans la société, sa contribution au débat social et même au progrès social. La RI défend la nécessité d'occuper pleinement une place non surplombante, c'est à dire une recherche au sein des territoires et auprès de leurs acteurs plutôt qu'une recherche déconnectée des préoccupations sociales, comme isolée dans une tour d'ivoire. Elle défend une recherche citoyenne (qui correspond au « prendre part »²³) et nécessairement engagée, ce « prendre parti »²⁴ étant inhérent à l'accompagnement du changement. Elle se rapproche en cela du courant interdisciplinaire des sciences participatives et du « manifeste pour une recherche scientifique responsable » (Robert et Marcel, 2019). De plus, l'élaboration de cette place contribue à une meilleure reconnaissance sociale (et sans doute scientifique) de la recherche en sciences de l'éducation et de la formation (Marcel, 2019b).

La participation et les contextes mobilisés

La RI repose sur le principe fondateur de la participation, le fait de chercher « avec ». Or, la participation ne se décrète pas, elle se construit et,

de vivre, c'est à dire de la vie en tant qu'elle est vivante » (Jullien, 2020, p. 45).

²³ Nous pouvons établir un lien ici avec les travaux de Zask (2011), qui décline le concept de participation en trois moments : prendre part (s'engager), apporter une part (contribuer) et recevoir une part (reconnaissance).

²⁴ Il retrouve le « apporter une part » de Zask.

dans le cadre de la RI, elle se traduit dans la co-élaboration d'un tiers-espace socio-scientifique (TESS)

Le TESS

L'espace du TESS (Marcel, 2019a) correspond au trait d'union²⁵ de recherche-intervention. Le TESS est qualifié de socio-scientifique à deux niveaux interdépendants, d'une part au travers des liens que la RI met « au travail », entre science et société et entre recherche et transformation, et, d'autre part, au travers des contextes, professionnels et de recherche, que la RI relie par ses dynamiques partenariale et participative.

Le TESS est institué par la signature de la commande qui en pose le cadre. Il porte, garantit, opérationnalise, concrétise et fait vivre la RI. Sur la base de la commande, le TESS est co-élaboré selon trois cercles qui correspondent à la fois à trois étapes chronologiques dans le déroulement de la RI et à trois modalités de participation se déployant au sein de contextes de plus en plus larges et ouverts. Par ailleurs, cette co-élaboration mobilise la dialectique entre la dimension coercitive (les contraintes liées aux repères du cadre issu de la commande et de la demande) et la dimension émancipatrice (portée par les processus d'appropriation et d'initiatives des acteurs concernés) qui caractérise tout dispositif.

Les trois cercles de participation

Parmi les trois cercles, schématisés dans le 2.1, le premier est sans doute le plus important. C'est celui des acteurs et des instances directement engagés dans la mise en œuvre et le suivi du projet. Il correspond au niveau micro systémique et participe d'une socialisation intrinsèque. Il est composé de deux instances qui sont instituées par la commande : le Comité de Pilotage (CoPil) (il est chargé du suivi de la commande et de son adaptation) et le Groupe de Travail (GTra) (chargé de la mise en œuvre collaborative du projet)²⁶. Sa fonction est d'installer et de lancer la RI.

²⁵ Le trait d'union préserve l'autonomie relative des deux termes mais, simultanément, génère un troisième terme qui les dépasse.

²⁶ Notons que deux représentants du Groupe de Travail participent au Comité de Pilotage.

Le deuxième cercle investit l'organisation sociale et professionnelle (que, sur le modèle du dispositif CIFRE²⁷, nous qualifions pour simplifier « d'entreprise ») au sein de laquelle se déploie la RI. Il s'ouvre au niveau méso systémique et développe une dynamique de socialisation à l'interface de l'intrinsèque du groupe engagé dans le projet et de l'extrinsèque de ses environnements. C'est le cercle de la valorisation des ressources pour le changement (les livrables identifiés par la commande) au sein de « l'entreprise » (toujours dans le sens des dispositifs CIFRE) et de ses réseaux (aussi bien locaux que régionaux, voire nationaux) selon des modalités diverses (des journées d'études, des capsules vidéo, des expositions, des articles professionnels, etc.). Cette valorisation s'appuie sur la mobilisation de l'Unité de Recherche (au-delà des chercheurs impliqués) avec l'objectif de faire connaître et de faire reconnaître la RI.

Le troisième cercle s'ouvre plus largement aux secteurs professionnels et académiques et au niveau macro systémique. C'est le cercle de la socialisation extrinsèque de la RI qui correspond à une autonomie relative des deux partenaires dans une forme de dissociation toutefois circonscrite par le collectif de départ. Ainsi, nous avons un travail scientifique, par des stratégies de publication et de diffusion à destination de la communauté scientifique et une démarche d'industrialisation²⁸ et d'implémentation des ressources sur l'ensemble du secteur de « l'entreprise ». La fonction de ce troisième cercle est de laisser se déployer la RI, ses retombées et ses prolongements. Nous pourrions dire qu'elle vit sa vie dans différents contextes et continue de contribuer aux changements.

²⁷ <https://www.enseignementsup-recherche.gouv.fr/fr/les-cifre-46510>

²⁸ Entendu comme un processus permettant à une ressource produite pour un contexte limité d'être déployée sur un contexte beaucoup plus large, en se basant sur la différenciation entre fabrication artisanale et fabrication industrielle.

III. L'évaluation de la RI

Afin de prendre en compte la complexité des niveaux de la RI et en particulier les trois cercles de participation dans l'évaluation de la RI, nous mobilisons une approche qualifiée d'écologique par ses promoteurs (Marcel et Bedin, 2018).

Une approche écologique

Cette approche (Vial, 2012 ; Younes, 2020 ; Younes et al., 2020) mobilise et met en synergie trois modèles de l'évaluation, brièvement présentés ici :

- le modèle de la mesure, une évaluation centrée sur les effets et qui s'attache à quantifier l'efficacité du dispositif,
- le modèle des valeurs, une évaluation centrée sur les processus et qui s'attache à qualifier le vécu et le sens des acteurs,
- le modèle de la gestion, une évaluation centrée sur les processus et leurs effets et qui s'attache à l'appréciation des transformations à deux niveaux, celui des acteurs (leur développement professionnel, leur émancipation, etc.) et celui du contexte (de nouvelles organisations du travail, de nouveaux dispositifs, de nouvelles pratiques, etc.).

Dans ce prolongement, nous pouvons schématiser le design de la RI de la façon suivante :

Figure 2 : Le design de la RI

CHANGEMENTS					
Contextes professionnels, académiques, sociétaux					
Cercle 3 <i>(autonomie relative)</i>	Diffusion <i>(contexte académique)</i>	Action politique <i>(contexte sociétal)</i>	Implémentation <i>(contexte professionnel)</i>	Modèle de la mesure <i>(efficacité)</i>	
	Restitution finale aux commanditaires <i>(livrables, démarche)</i>				
Cercle 2 <i>(dynamique mixte)</i>	Valorisations <i>(outils, démarches, acteurs)</i>			Modèle de la gestion <i>(reconnaissance)</i>	
	Restitution intermédiaire aux commanditaires <i>(livrables, démarche)</i>				
Cercle 1 <i>(dynamique intrinsèque)</i>	Savoirs contextualisés <i>(sur et pour l'action)</i>		Outils pour l'action		Modèle des valeurs <i>(en commun)</i>
TESS « avec »	Visée heuristique « sur »	Visée critique « par »	Visée praxéologique « pour »		Co-évaluation
COMMANDE (contractualisation)					
Propédeutique du TESS					

Les chantiers ouverts

Nous venons de circonscrire le périmètre d'une approche qualifiée d'écologique, ce qui se traduit à présent par l'invitation à outiller les neuf cases du tableau ci-dessous, envisagée comme une feuille de route pour nos travaux futurs :

Figure 3 : Panorama des chantiers ouverts

Cercle 3				Modèle de la mesure (efficacité)
Cercle 2				Modèle de la gestion (reconnaissance)
Cercle 1				Modèle des valeurs (en commun)
TESS « avec »	Visée heuristique « sur »	Visée critique « par »	Visée praxéologique « pour »	Co-évaluation

Des outils d'évaluation

Nous avons initié ce chantier et nous citerons brièvement deux outils d'évaluation avant de nous attarder plus longuement sur un troisième.

Participation et émancipation

Tout d'abord, dans l'objectif de se départir d'une approche incantatoire²⁹ de la participation, nous avons construit une grille d'évaluation du « degré » de participation par rapport à neuf critères correspondant à une opérationnalisation des différents volets de la RI : élaboration du projet,

²⁹ Qui se limite à un discours enjôleur.

définition des objectifs, construction du dispositif, structuration du collectif, volet recherche, volet intervention, négociations et régulation avec les commanditaires, restitution et évaluation du projet (Marcel, 2019a).

Nous avons également stabilisé un outil d'évaluation de l'émancipation professionnelle³⁰ (Marcel et Broussal, 2017, 2021, 2022). Il comprend les marqueurs de déplacement (sentiment d'efficacité professionnelle, autorisation, pouvoir d'agir), les indicateurs d'une nouvelle légitimité, la contribution des modalités collectives à l'accompagnement du déplacement et la caractérisation du vécu de ce déplacement.

Le récit du changement

En nous appuyant³¹ sur Ricœur défendant l'analogie entre actions et textes, nous avons emprunté à Todorov la notion « d'aspect du récit » qui correspond à la façon dont l'histoire est perçue par le narrateur. Il distingue trois points de vue : le récit « avec » (le narrateur en sait autant que son personnage, il est le « porte-parole » du personnage), le récit « du dehors » (le narrateur en sait moins que son personnage (il se réduit à ce qu'il voit et entend) et le récit « par derrière » (le narrateur en sait davantage que son personnage puisqu'il a connaissance de deux points de vue différents) qui correspond à l'articulation des deux récits précédents.

Pour la structuration générale du récit, nous avons repris le schéma narratif quinaire en le transposant à la RI :

- 1) La situation initiale : le diagnostic, l'appropriation du projet,

³⁰ Rappelons que nous entendons par émancipation une modification du rapport aux autres et au monde (qui n'est pas réductible à l'affranchissement d'une domination) qui se traduit par un « déplacement » de chacun des acteurs, grâce aux dynamiques collectives d'un projet partagé (qui n'est pas réductible à une lutte contre un oppresseur). Ce déplacement consiste, pour chaque acteur à s'extraire du positionnement auquel il se trouvait assigné au départ et à conquérir, simultanément, un positionnement plus conforme à ses aspirations. Dans le cas de l'émancipation professionnelle, ce positionnement se cantonne au champ professionnel.

³¹ L'ensemble de la démarche est théorisé et référencé dans Marcel, (2014).

- 2) L'événement perturbateur (ou déclencheur) : la commande, les instances,
- 3) Les différentes péripéties (opposants et adjuvants) : le déroulement, la mise en œuvre,
- 4) L'événement de résolution (ou de dénouement) : les restitutions, les livrables,
- 5) La situation finale : les transformations, les effets de la RI.

Le récit de chacune des 5 phases est élaboré en 3 étapes :

Figure 4 : Les trois étapes de l'élaboration des récits

Le récit « avec »	Le récit du « dehors »
À partir des points de vue des acteurs de l'organisation (G'Tra)	A partir d'éléments collectés (observations, documents, traces, données chiffrées, etc.)
Le récit « par derrière »	
Récit reconstruit à partir des deux récits précédents	

En s'appuyant sur la réflexivité que permet l'écrit, la mise en récit constitue un outil de co-évaluation très intéressant, par exemple pour la mise au jour et la structuration du « en commun », pour la transversalité et la globalité de l'exercice (co-construction du sens de la RI), pour son caractère longitudinal (une sorte de fil rouge) ou pour les perspectives de diffusion importantes en termes de visibilité et de reconnaissance. Nous devons toutefois reconnaître que cet exercice peut s'avérer contraignant et nécessite une vigilance soutenue pour maintenir sa dynamique.

IV. Conclusion : Vers un tiers-espace de l'évaluation de la RI (TEERI)

Pour conclure ce chapitre et dans le prolongement des réflexions initiées relatives à la diversité des contextes mobilisés pour approfondir la question de l'évaluation de la RI, en tant que démarche et en tant que dispositif, nous proposerons une amorce de théorisation TEERI (tiers-espace de l'évaluation de la RI) à partir de dix caractéristiques :

1. Le TEERI serait un sous-ensemble du TESS, il renforcerait la symbolique du trait d'union à propos des interrelations entre les deux sphères,
2. Il s'intéresserait à la démarche de RI, envisagée comme un dispositif, pas seulement à l'avancée ou aux résultats du projet,
3. Il s'agirait d'une instance nouvelle et mixte, englobant Gtra et CoPil, pensée et installée par la commande,
4. Il impliquerait la participation par une pleine démarche de co-évaluation,
5. Il constituerait l'espace privilégié de travail de la référentialisation (Figari et Tourmen, 2006), de l'élaboration et de l'explicitation du référent, en lien avec le projet,
6. Il aurait une fonction longitudinale de régulation et d'ajustement à la fois en lien avec la plasticité de la commande et en lien avec la « mise en débat » de l'interprétation des écarts entre référent et référé (envisagés en interrelations),
7. Il accueillerait la construction des outils d'évaluation (pour investir empiriquement le référé avant d'interpréter les éléments collectés). Ces outils couvriraient les périmètres des trois modèles, aussi bien la démarche que l'ensemble des processus et des produits de la RI,
8. Il aurait une fonction de facilitation des passages du cercle 1 au cercle 2, puis du cercle 2 au cercle 3 (en assumant la mémoire du cercle 1),
9. De par son hétérogénéité, il offrirait un espace privilégié à la distanciation et à une parole critique,
10. Il contribuerait à renforcer la lisibilité (tant interne qu'externe) de la démarche et éclairerait différemment ses enjeux (tant pour

l'entreprise que pour l'Unité de Recherche, voire pour la discipline des sciences de l'éducation et de la formation.

V. Bibliographie

- Aussel, L. (2017). Emancipation et dispositif : analyse des voies alternatives. Dans J.-F. Marcel et D. Broussal (dir.), *Emancipation et recherche en éducation. Conditions de la rencontre entre science et militance* (p. 193-216). Editions du Croquant.
- De Sousa Santos, B. (2011). Épistémologies du Sud, *Etudes rurales*, 187. <https://journals.openedition.org/etudesrurales/9351>
- Figari, G. et Tourmen, C. (2006). La référentialisation : une façon de modéliser l'évaluation de programme, entre théorie et pratique. Vers une comparaison des approches au Québec et en France. *Mesure et évaluation en éducation*, 29(3), 5–25.
- Jullien, F. (2020). *Politique de la décoïncidence*. Editions de L'observatoire.
- Marcel, J.-F. (2014), Les pratiques enseignantes mises en récit. *Recherches en éducation*, 19. <https://journals.openedition.org/ree/8319?lang=en>
- Marcel, J.-F. (2019a). Intervention, participation et évaluation dans la recherche en éducation. *La Revue LEeE*, 1. <https://revue.leeeonline/index.php/info/article/view/40>
- Marcel, J.-F. (2019b). D'aujourd'hui à demain : une discipline sous l'égide de Janus. *Les Sciences de l'éducation – Pour l'ère nouvelle*, 52(2), 29-46.
- Marcel, J.-F. (2020a). Fonctions de la recherche et participation : une épistémo-compatibilité dans le cas de la recherche-intervention. *Questions Vives*, 33. <https://journals.openedition.org/questionsvives/4691>
- Marcel, J.-F. (2020b). Quand la recherche-intervention puise son inspiration dans la pensée de Freire. *Les Dossiers des Sciences de l'éducation*, 44, 101-122.
- Marcel, J.-F. (2023). Une « vraie » place dans la recherche-intervention. *Recherches en éducation*, 51. <https://journals.openedition.org/ree/11631>

- Marcel, J.-F. et Bedin, V. (2018). Contribution à l'élaboration d'un dispositif d'évaluation de la Recherche-Intervention. *Phronesis*, 7(1), 79–91.
- Marcel, J.-F. et Broussal, D. (2017). *Emancipation et recherche en éducation*. Editions du Croquant.
- Marcel, J.-F. et Broussal, D. (2021). Evaluer l'émancipation professionnelle des enseignants. Dans C. Gremion et C. Power (dir.), *Comment évaluer la professionnalité émergente ?* (p. 107-127). De Boeck.
- Marcel, J.-F. et Broussal, D. (2022). La recherche-intervention : accompagner le changement. Dans B. Albero et J. Thiévenaz (dir.), *Traité de méthodologie de la recherche en Sciences de l'Education et de la Formation* (p. 247-258). Editions Raison et Passions.
- Marcel, J.-F. et Nunez Moscoso, J. (2012). La figura del investigador-ciudadano: hacia un (re)encuentro con el ethos de la investigación en educación. *Revista Estudios Cooperativos*, 17(1-2), 101-121.
- Robert, A-D et Marcel, J-F (2019). Des relations entre recherches en éducation et engagements militants : éléments pour un débat. Dans J-F Marcel, V. Bordes et L. Lescouarch (dir.), *Recherches en éducation et engagements militants. Vers une tierce approche* (p. 25-42). Presses Universitaires du Midi.
- Younès, N. (2020). *Vers une évaluation écologique dans l'enseignement supérieur : dispositifs et transformations en jeu*. [habilitation à diriger des recherches]. Université de Lorraine.
- Younès, N., Gremion, C. et Sylvestre, E. (2020). *Évaluations, sources de synergies ?* Les Presses de l'ADMEE.
- Vial, M. (2012) *Se repérer dans les modèles de l'évaluation. Méthodes – Dispositifs – Outils*. De Boeck.
- Zask, J. (2011). *Participer. Essai sur les formes démocratiques de la participation*. Editions Le Bord de l'eau.

Chapitre 10. La « complexité » dans l'évaluation des compétences : des défis méthodologiques et contextuels

Cas du Baccalauréat arabe international (IAB)

Fadi el-Hage et Joe Ephrem

I. Introduction et problématisation

Cette contribution s'appuie sur la conférence qui a été proposée dans le cadre du 33^{ème} colloque de l'ADMEE-Europe. L'objet de cette conférence consistait à présenter le développement d'un questionnaire d'évaluation des compétences dans le cadre du programme IAB (Baccalauréat arabe international) au Liban, en tenant compte de la complexité du contexte socio-politique et économique libanais et des enjeux de l'évaluation des compétences. Ce texte expose le questionnaire développé et les résultats d'une recherche sur les compétences acquises par des lycéens libanais suivant ce programme, ainsi que les défis à relever.

L'importance du développement de ce questionnaire et de son utilisation dans le cadre de l'évaluation des compétences s'inscrit dans un souci d'améliorer les compétences des élèves suivant le programme de l'IAB et dans une préoccupation plus générale de contribuer au développement de l'approche par compétences au sein du système éducatif libanais tout en gardant à l'esprit sa diversité.

Pour mieux comprendre les enjeux de cette démarche, il est important de rappeler quelques caractéristiques actuelles du système éducatif libanais (programmes, évaluation). Ce système s'inscrit dans un contexte multiculturel (avec trois langues dominantes : l'arabe, le français et l'anglais) et multiconfessionnel (dix-huit confessions présentes). Du point

de vue structurel, la place de l'école privée a été longtemps très importante. A titre d'exemple, environ 65% des élèves ont poursuivi leur scolarité dans des écoles privées. Depuis les 1990, des réformes éducatives sont propulsées en vue de moderniser le système éducatif mettant en avant le principe d'équité et de l'égalité des chances, et l'instauration de l'approche par compétences. Mais ces changements s'entament dans des conditions particulières et non sans difficultés en raison de la guerre civile, la succession des crises socio-politiques et économiques incessantes et de la présence et de l'impact de multiples idéologies contradictoires (International Center for Transitional Justice, 2014). C'est dans ces conditions que dans les années 2000, le programme de l'IAB (l'International Arab Baccalaureate) met l'accent sur le développement d'un ensemble de compétences, appelées les compétences du XXIème siècle. Ces compétences s'articulent autour de quatre axes thématiques/dimensions : la dimension cognitive et apprendre à savoir ; la dimension instrumentale (l'employabilité et apprendre à faire) ; le développement personnel (la dimension individuelle et apprendre à être) ; la dimension de citoyenneté active (apprendre à vivre ensemble) (Fisser et Thijs, 2015 ; National Research Council, 2010 ; OECD, 2019). Face à cette approche holistique par compétences, il devient aussi important de penser à l'évaluation des compétences visées. C'est dans ce cadre que nous avons, en se référant à différentes grilles internationales (Soland et al., 2013 ; Chu et al., 2017), élaboré une grille multidimensionnelle de compétences. Sur cette base, nous avons aussi construit un questionnaire permettant une évaluation quantitative des compétences et enfin nous avons évalué les compétences du XXIème siècle chez des élèves poursuivant le programme de l'IAB (l'International Arab Baccalaureate). Par ailleurs, compte tenu des résultats d'une étude menée par Abla et al. (2015) montrant que des élèves libanais entretiennent des relations tendues avec leurs pairs non libanais, nous avons formulé l'hypothèse que le développement des compétences sociales est lié aux autres dimensions de compétences (cognitives, instrumentales, personnelles). Avant de présenter la grille et le questionnaire développés ainsi que les principaux résultats, nous apporterons des éclairages sur nos choix théoriques et définitions conceptuelles.

II. Cadrage Théorique

La compétence comme objet curriculaire complexe

Une compétence se définit comme la capacité de mobiliser un savoir-agir complexe dans un contexte spécifique. En effet, elle se caractérise par son caractère intégrateur, combinatoire, développemental, contextuel et évolutif (El Hage, 2005 ; Manson 2008 ; Morin et Le Moigne, 1999 ; Tardif, 2006). Pour résoudre des situations complexes à l'intérieur d'une famille de situations (Le Boterf, 2002), il faut mobiliser des ressources internes et externes. Les ressources internes concernent divers domaines (cognitif, affectif, comportemental et sensorimoteur), différents types de connaissances (déclaratives, procédurales) et différentes origines de savoirs (codifiés et expérientiels). Les ressources externes peuvent concerner des collègues dont on sollicite l'avis ou l'expertise, des bases de données scientifiques, des références, les technologies de l'information et les logiciels, des organisations logistiques, etc. Bref, tout ce qui ne relève pas de l'individu lui-même, concerné par les compétences à développer.

La complexité de l'évaluation des compétences

L'évaluation des compétences implique la prise en compte d'un réseau intégré et fonctionnel, formé de composantes cognitives, affectives, sociales et sensorimotrices, susceptibles d'être mobilisées en actions finalisées face à une famille de situations (Allal, 2003). L'évaluation des compétences se base sur la prise en compte d'un ensemble de critères et des micro-critères évalués qui correspond à l'émergence globale et complexe de la compétence à évaluer. Le risque de l'évaluation des compétences est de tomber dans un réductionnisme qui peut être synthétisé en trois postulats que nous suggérons :

Le postulat de la « décomposabilité » : toute compétence est décomposable en composantes élémentaires et la somme des évaluations isolées de chacune de ces composantes fournit un indicateur pertinent de développement de la compétence, indépendamment des interactions entre chacune des composantes (le tout est assumé comme étant strictement égal à la somme des parties).

Le postulat de la « décontextualisation » : suppose la définition de chaque composante d'une compétence, en admettant qu'elle peut se manifester indépendamment du contexte dans lequel elle prend place.

Le postulat du « transfert automatique » : L'étudiant est capable de faire l'intégration et le transfert des compétences et des savoirs acquis dans les différents contextes où ces derniers seront sollicités, de façon spontanée. L'apprentissage de l'intégration des savoirs et des compétences dans des situations complexes et de leurs transferts est également maîtrisé. (Gerard, 2005).

En plus de ce réductionnisme illustré par ces trois postulats, il existe un autre risque qui est le fait d'adopter une approche simpliste (au lieu d'une approche systémique et/ou interactive) dans l'élaboration des référentiels des compétences. Une approche simpliste consiste à ne pas tenir compte de l'ensemble des compétences visées et/ou de leur complexité, multi-dimensionnalité et interdépendance (Crahay, 2006). Par exemple, dans un contexte professionnel, lorsqu'une organisation souhaite s'inscrire dans une formation par compétences, elle peut commencer par définir un « **référentiel métier** » qui décrit et organise toutes les situations auxquelles les travailleurs seront confrontés dans leur pratique professionnelle. On peut élaborer ensuite un « **référentiel de compétences** » qui permet de dégager toutes les compétences nécessaires pour résoudre les situations décrites dans le référentiel métier. Puis, on peut définir d'une part un « **référentiel de formation** » qui permettra de savoir comment on va agir pour que les travailleurs maîtrisent effectivement les compétences nécessaires et d'autre part un « **référentiel d'évaluation** » qui permettra l'évaluation des compétences réelles des travailleurs afin d'élaborer des plans individuels de formation (Parent et al., 2010).

La question de la complexité, la multi-dimensionnalité et l'interdépendance des compétences est à prendre en considération non seulement sur le plan théorique, mais aussi dans la réflexion sur l'évaluation des compétences. Différentes approches d'évaluation critériée descriptive (qualitative) et de notation (quantitative) ont commencé à se développer afin de répondre aux besoins d'évaluer le plus adéquatement possible des compétences. Les approches d'évaluation critériée descriptive (qualitative) ont souvent marqué l'évaluation des compétences mobilisées dans des « tâches complexes ». L'évaluation par notation (quantitative) devient de plus en plus rare, et l'importance des approches d'évaluation mixtes est soulevée afin de mieux saisir la complexité, la multi-dimensionnalité et l'interdépendance des compétences (Jorro, 2006).

Pour ce que concerne les défis liés à l'évaluation des compétences du XXIème siècle ceux-ci résident dans le manque de compréhension complète de la nature et du développement des compétences, de leur caractère multidimensionnel et de la difficulté de distinguer les

connaissances des compétences (Care et Griffin, 2018). Un autre problème lié à l'évaluation des compétences du XXIème siècle est la création et l'utilisation des outils de mesure quantitatifs pour évaluer et recueillir des indices de performances sur des activités spécifiques des apprenants liées aux compétences en question (Bourn, 2018).

III. Méthodologie

Dans le but d'évaluer les compétences du XXIème siècle des élèves (notamment 774 élèves en première et seconde du lycée au sein de 4 établissements privés et 4 établissements publics), nous avons opté pour une combinaison d'évaluation critériée descriptive (qualitative) et de notation (quantitative). Nous avons utilisé l'évaluation critériée descriptive pour construire la grille d'évaluation des compétences et la notation pour évaluer ces compétences selon des critères bien définis. Quatre objectifs opérationnels ont été fixés :

- 1- Construire une grille synthétique des compétences du XXIème siècle,
- 2- Construire un questionnaire permettant une évaluation des compétences,
- 3- Évaluer les compétences du XXIème siècle chez les élèves,
- 4- Déterminer la présence d'associations éventuelles entre les compétences de la dimension sociale d'une part, et les compétences de la dimension sociale et les compétences des autres trois dimensions d'autre part.

Construction d'une grille synthétique des compétences du XXIème siècle

Étapes de conception de la grille

Pour construire la grille des compétences du XXIème nous avons adopté l'approche qualitative en nous référant premièrement à la liste des compétences couvertes dans la version 2016 du Skills-based test de l'International Arab Baccalaureate (EducationalRC.org/IAB, 2020 – aujourd'hui iab.academicassessment.com), deuxièmement, au rapport le plus récent de l'UNICEF, publié en 2017, intitulé « Reimagining Life Skills and Citizenship Education in the Middle East and North

Africa – A Four-Dimensional and Systems Approach to 21st Century Skills », et troisièmement à des travaux (Soland et al., (2013) et des grilles de compétences préexistantes (« Twenty-First Century Skills and Global Education Roadmaps » de Chu et al., (2017), et « Towards Defining 21st Century Competencies for Ontario – 21st Century Competencies » de La fonction publique de l'Ontario, (2016).

Pour développer la grille, nous avons réalisé une analyse de contenu documentaire, en suivant l'approche méthodologique décrite par Bardin (2003). Nous avons ensuite construit une grille qui englobe la majorité des compétences et en adaptant les dimensions de l'apprentissage aux quatre piliers du « Profile Shaping Education » (le cadre pédagogique de l'International Arab Baccalaureate), et enfin en intégrant les compétences couvertes dans l'« IAB Skills-Based test » sous les quatre axes thématiques (piliers) et dimensions de l'apprentissage (voir tableau 1) : la dimension cognitive et apprendre à savoir ; la dimension instrumentale (l'employabilité et apprendre à faire) ; le développement personnel (la dimension individuelle et apprendre à être) ; la dimension de citoyenneté active (apprendre à vivre ensemble). Ces quatre piliers sont inspirés des principes de Halloun, qui structure l'apprentissage autour des « 4 P » : paradigmatique (apprendre à savoir), productive (apprendre à faire), proactive (apprendre à être) et « principe » (apprendre à vivre ensemble) (Halloun, 2011).

Chacune des dimensions se décline ensuite en différents indicateurs (voir tableau 1) : descriptifs de compétences précises. Ces indicateurs (descriptifs de compétences) sont utilisés comme un guide pour rédiger les questions et les catégoriser en fonction de la compétence mesurée. De plus, ils ont été utilisés comme une référence de contrôle qualité pour refuser ou accepter une question.

Tableau 1 : Grille de compétences du XXIème siècle

Piliers : Axes	Dimensions tirées du PSE et « IAB skills-based tests »	Compétences du XXIème siècle
Apprentissage (Paradigmatique)	Dimension Cognitive (Apprendre à savoir)	<ul style="list-style-type: none"> • Littératie numérique - Culture numérique (Digital Literacy) (TIC, médias...) • Créativité/Innovation • Curiosité • Pensée critique • Résolution de problèmes • Métacognition
Employabilité (Productive)	Dimension Instrumentale (Apprendre à faire)	<ul style="list-style-type: none"> • Coopération / Collaboration/Travail d'équipe • Négociation • Prise de décision
Autonomisation Personnelle (Pro-active)	Dimension Individuelle (Apprendre à être)	<ul style="list-style-type: none"> • Protection de soi • Autogestion • Résilience • Communication (Intra- et Inter-) • Flexibilité et adaptabilité • Développement personnel
Citoyenneté active (Principe)	Dimension Sociale (Apprendre à vivre ensemble)	<ul style="list-style-type: none"> • Interculturalité/Citoyenneté mondiale • Respect de la diversité • Empathie • Participation/environnement social • Équité • Droits de l'Homme • Valeurs et justice sociale (éthique)

Construction d'un questionnaire permettant une évaluation des compétences

Étapes de conception du questionnaire

Pour construire le questionnaire, nous avons adopté l'approche qualitative en nous référant à la grille que nous avons construite pour générer une liste de questions organisée par compétence et par dimension. Nous avons complété selon les quatre étapes suivantes : 1) travailler chaque dimension séparément, 2) rédaction de x questions par dimension/compétences, puis 3) validation interne des questions en se référant à une liste de critères bien définis, et enfin 4) validation externe.

Le questionnaire a été soumis au test statistique « coefficient alpha de Cronbach » utilisé notamment en psychométrie pour mesurer la cohérence interne (ou la fiabilité) des questions posées lors d'un test (valeur = 0,5).

La validation du questionnaire a ensuite été faite par un expert externe à l'équipe de travail et par un test-pilote administré auprès de 50 élèves. Le questionnaire est constitué de 41 questions dont 39 questions fermées et 2 questions ouvertes, réparties comme l'indique le tableau suivant :

Tableau 2 : Répartition des questions sur les dimensions

Dimensions	Nombre de questions fermées	Nombre de questions ouvertes
La dimension cognitive	11	1
La dimension instrumentale	5	--
La dimension individuelle	11	--
La dimension sociale	12	1
Total	39	2

Recueil de données et échantillon. Analyse de données

Afin de mener à bien cette évaluation des compétences du XXI^{ème} siècle chez les élèves de seconde et de première scientifique, inscrits à l'IAB, nous avons mis en place une collecte de données durant le premier trimestre de l'année scolaire 2019-2020, couvrant un échantillon de 774 élèves dans huit écoles réparties entre secteurs public (571 élèves) et privé (203 élèves). Le questionnaire a été distribué et supervisé pendant les heures de cours en classe. Les étudiants ont répondu à 32 questions à choix multiples, où ils devaient choisir leurs actions dans des situations de la vie réelle. Ils ont également répondu à 5 questions sur une échelle de Likert (de 1 à 5) pour exprimer leurs perceptions concernant des situations réelles, ainsi qu'à 4 questions ouvertes pour résoudre des problèmes et partager leurs opinions.

Analyse de données

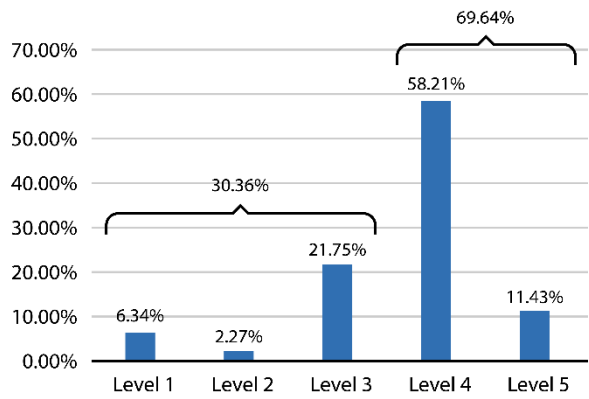
Pour évaluer les compétences du XXI^{ème} siècle chez les élèves de seconde et de première scientifique, inscrits à l'IAB, nous avons adopté l'approche quantitative en analysant les réponses des élèves sur la version finale du questionnaire (Skills-Based test). Nous avons classé toutes les compétences mesurées (par le biais de questions fermées de type QCM et ouvertes) dans les différentes dimensions sous trois niveaux : niveau 1 : Non acquis, niveau 2 : Acquis, et niveau 3 : Hautement acquis. En plus, un score est calculé pour chaque dimension (groupe de compétences) en additionnant les points correspondant à toutes les questions incluses dans la dimension ; le nombre calculé est ensuite converti sur 20. Selon le score calculé dans chaque dimension, un codage a été appliqué afin de classer les résultats : niveau 1 : Très faible, pour un score compris entre 0 et 5 sur 20 ; niveau 2 : Faible, pour un score compris entre 6 et 9 sur 20 ; niveau 3 : Modéré, pour un score compris entre 10 et 13 sur 20 ; niveau 4 : Bon, pour un score compris entre 14 et 16 sur 20 ; niveau 5 : Excellent, pour un score compris entre 17 et 20 sur 20.

Ensuite, nous avons utilisé le logiciel SPSS pour l'analyse statistique descriptive des réponses. Nous avons calculé « Chi Square » pour déterminer la présence d'association entre les compétences de la dimension sociale d'une part, et entre les compétences de la dimension sociale et les compétences des trois autres dimensions d'autre part. L'association est présente lorsque la valeur « Chi square » est supérieure à 0.4. Dans la partie suivante, nous allons présenter les résultats obtenus suivant les objectifs opérationnels.

IV. Résultats

Les compétences du XXIème siècle, développées chez les élèves du cycle secondaire

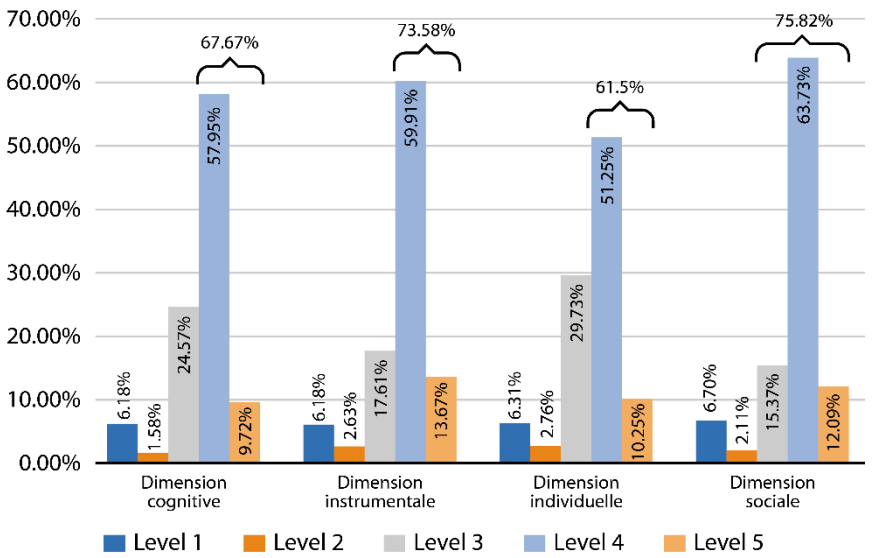
Figure 3 : Distribution des élèves en termes de niveau d'acquisition globale des compétences relatives aux quatre dimensions



Comme l'illustre la figure 3, pour ce qui concerne l'acquisition du niveau global des compétences relatives aux quatre dimensions (cognitive, instrumentale, individuelle et sociale), 30,36% des élèves ont été classés dans les niveaux 1, 2 et 3 correspondants à un faible niveau de développement des compétences et 69,64% des élèves ont été classés dans les niveaux 4 et 5 correspondants à un haut niveau de développement des compétences.

La figure ci-dessous montre les résultats obtenus pour chaque dimension.

Figure 4 : Distribution des élèves en termes de niveau d'acquisition des compétences pour chaque dimension



Les résultats relatifs à chaque compétence de la dimension individuelle dans ces histogrammes ne montrent pas tous le même niveau d'acquisition. Les tableaux ci-dessous (3, 4, 5 et 6) montrent la distribution des élèves en termes de niveau d'acquisition pour chacune des compétences des différentes dimensions.

Tableau 3 : Répartition des élèves en termes de niveau d'acquisition par compétence de la dimension cognitive

			1.00	2.00	3.00	Total
Dimension Cognitive	Littératie numérique - Culture numérique	Count	56	623	71	750
		%	7%	83%	9%	100%
	Créativité	Count	402	313	35	750
		%	54%	42%	5%	100%
	Curiosité	Count	67	236	447	750
		%	9%	31%	60%	100%
	Pensée critique	Count	48	327	375	750
		%	6%	44%	50%	100%
	Résolution de problèmes	Count	97	184	469	750
		%	13%	25%	63%	100%
	Métacognition	Count	48	278	424	750
		%	6%	37%	57%	100%

Tableau 4 : Répartition des élèves en termes de niveau d'acquisition par compétence de la dimension instrumentale

			1.00	2.00	3.00	Total
Dimension Instrumentale	Coopération	Count	131	497	122	750
		%	17%	66%	16%	100%
	Négociation	Count	47	213	490	750
		%	6%	28%	65%	100%
	Prise de décision	Count	69	429	252	750
		%	9%	57%	34%	100%

Tableau 5 : Répartition des élèves en termes de niveau d'acquisition par compétence de la dimension individuelle

			1.00	2.00	3.00	Total
Dimension Individuelle	Protection de soi	Count	47	248	455	750
		%	6%	33%	61%	100%
	Autogestion	Count	49	499	202	750
		%	7%	67%	27%	100%
	Résilience	Count	48	329	373	750
		%	6%	44%	50%	100%
	Communication (Intra- et Inter-)	Count	51	397	302	750
		%	7%	53%	40%	100%
	Flexibilité et adaptabilité	Count	50	252	448	750
		%	7%	34%	60%	100%
	Développement personnel	Count	356	233	161	750
		%	47%	31%	21%	100%

Tableau 6 : Répartition des élèves en termes de niveau d'acquisition par compétence de la dimension sociale

			1.00	2.00	3.00	Total
Dimension Sociale	Interculturalité / Citoyenneté mondiale	Count	53	489	208	750
		%	7%	65%	28%	100%
	Respect de la diversité	Count	62	433	255	750
		%	8%	58%	34%	100%
	Empathie	Count	61	410	279	750
		%	8%	55%	37%	100%
	Participation / environnement social	Count	52	260	438	750
		%	7%	35%	58%	100%
	Équité	Count	113	120	517	750
		%	15%	16%	69%	100%
	Droits de l'Homme	Count	90	494	166	750
		%	12%	66%	22%	100%
	Valeurs et justice sociale (éthique)	Count	401	316	33	750
		%	53%	42%	4%	100%

En se référant aux tableaux (3, 4, 5 et 6) ci-dessus, le pourcentage le moins élevé d'élèves classés au niveau 3 d'acquisition de compétences de la dimension cognitive concerne la créativité avec 5% de la population, suivi de 9% pour la littératie numérique, 50% pour la pensée critique, 57% pour la métacognition, 60% pour la curiosité et 63% pour la compétence de résolution de problèmes. Le pourcentage le moins élevé d'élèves classés au niveau 3 d'acquisition de compétences de la dimension instrumentale concerne la compétence de coopération/collaboration/travail d'équipe avec 16% de la population, suivi de 35% pour la compétence de prise de décision et de 65% pour la négociation. Le pourcentage le moins élevé d'élèves classés au niveau 3 d'acquisition de compétences de la dimension individuelle concerne la compétence du développement personnel avec 21% de la population, suivi de 27% pour l'autogestion, 40% pour la communication, 50% pour la compétence de résilience, 60% pour la compétence de flexibilité et d'adaptabilité et 61% pour la compétence de protection de soi (autoprotection). Le pourcentage le moins élevé d'élèves classés au niveau 3 d'acquisition de compétences de la dimension sociale concerne les valeurs et la justice sociale avec 4% de la population, suivi de 22% pour les droits de l'homme, 28% pour la citoyenneté mondiale, 34% pour le respect de la diversité, 37% pour l'empathie, 58% pour la participation et 69% pour la compétence d'équité. Les figures ci-dessous montrent la distribution des élèves dans chaque école et leur répartition selon les niveaux entre bons et excellents.

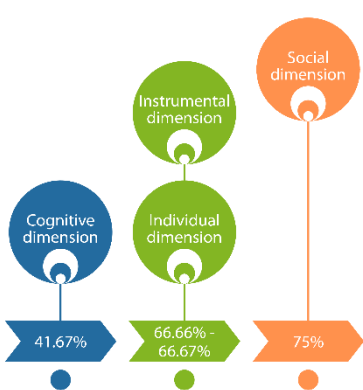


Figure 5 : Patriarcal - privée - musulmans

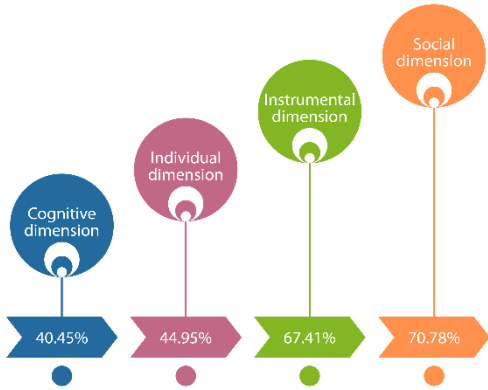


Figure 6 : Global International School (GIS) - privée - musulmans

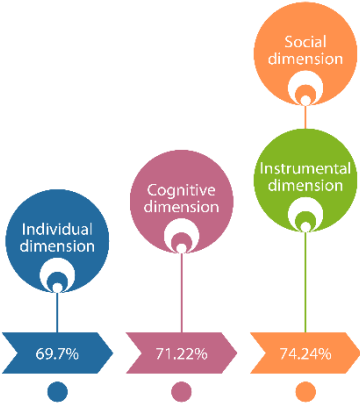


Figure 7 : NEIGB - privée - musulmans

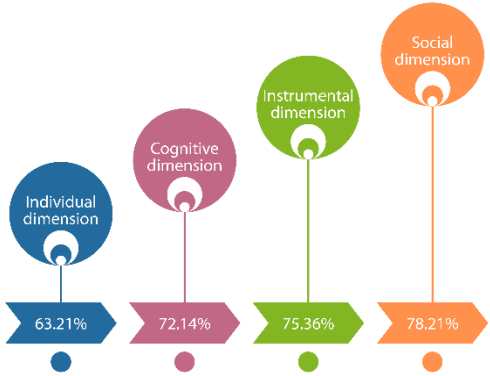


Figure 8 : Nazih el Bizri - publique - musulmans (sunrites)

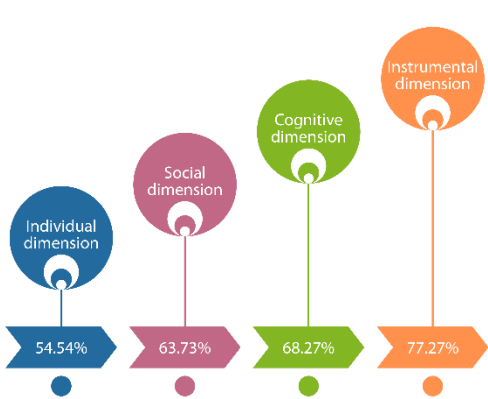


Figure 9 : Gebran Andraous Tueini (GAT) - publique - musulmans (Chiites)

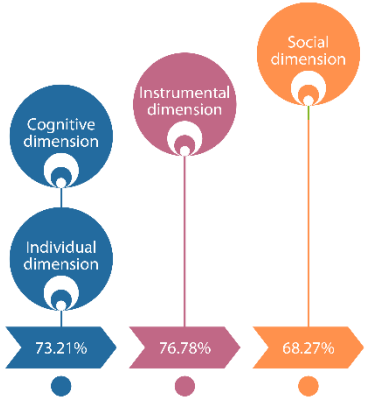


Figure 10 : Hemmana - publique - Chrétiens

Ces résultats indiquent que le contexte social influence plus que l'appartenance religieuse sur le développement d'une dimension de compétences donnée.

L'influence croisée du développement des compétences

Pour mieux comprendre comment aider les élèves inscrits au programme de l'IAB à améliorer leurs compétences acquises tout en gardant l'esprit de diversité du système éducatif au Liban.

Tableau 7 : Degré d'association entre le développement de chaque compétence de la dimension sociale l'une avec l'autre.

		Interculturalité/ Citoyenneté mondiale	Respect de la diversité	Empathie	Participation/ environnement social	Équité	Droits de l'Homme	Valeurs et justice sociale
Dimension Sociale	Interculturalité/ Citoyenneté mondiale		0.563	0.537	0.487		0.438	
	Respect de la diversité	0.563		0.585	0.524		0.45	
	Empathie	0.537	0.585		0.448		0.422	
	Participation/ environnement social	0.487	0.524	0.448		0.419	0.471	
	Équité				0.419			
	Droits de l'Homme	0.438	0.45	0.422	0.471			
	Valeurs et justice sociale							

L'analyse de « l'association modérée » étudiée entre les compétences de la dimension sociale montre que le développement des compétences liées à l'ouverture à l'autre progresse simultanément. Ce groupe de compétences englobe : interculturalité, citoyenneté mondiale, respect de la diversité, empathie, participation sociale et la compétence des droits de l'Homme. De plus, l'analyse montre que l'équité et la compétence liée aux valeurs et à la justice sociale ne sont pas associées aux autres compétences de la dimension sociale.

Tableau 8 : Degré d'association entre le développement des compétences de la dimension sociale et le développement des compétences de la dimension cognitive

		Interculturalité/ Citoyenneté mondiale	Respect de la diversité	Empathie	Participation/ environnement social	Équité	Droits de l'Homme	Valeurs et justice sociale
Dimension Cognitive	Culture numérique							
	Créativité							
	Curiosité	0.592	0.631	0.579	0.501		0.457	
	Pensée critique	0.446	0.433	0.433	0.513		0.436	
	Résolution de problèmes	0.412	0.464	0.404	0.45			
	Métacognition	0.43	0.442		0.512		0.434	

L'analyse de « l'association modérée » étudiée entre les compétences liées à la dimension cognitive montre que les élèves qui ont des compétences d'ouverture à l'autre développent aussi des compétences métacognitives importantes (curiosité, pensée critique, résolution de problèmes et métacognition) ... L'ouverture à l'autre semble être une condition fondamentale pour le développement des capacités réflexives et de l'intelligence multiple. La fermeture sur soi et le cloisonnement bloquent le cerveau : les neurosciences et la plasticité cérébrale sont facilitées par la tolérance et l'ouverture à l'autre. Par ailleurs, les compétences d'équité et de valeurs et celles de la justice sociale ne sont pas associées aux compétences de dimensions cognitives. La littératie numérique et la créativité ne sont pas associées à toutes les compétences de la dimension sociale.

Tableau 9 : Degré d'association entre le développement des compétences de la dimension sociale et le développement des compétences de la dimension instrumentale

		Interculturalité/ Citoyenneté mondiale	Respect de la diversité	Empathie	Participation/ environnement social	Équité	Droits de l'Homme	Valeurs et justice sociale
Dimension Cognitive	Coopération	0.435	0.434	0.415	0.529		0.456	
	Négociation							
	Prise de décision	0.569	0.618	0.575	0.474		0.449	

L'analyse de « l'association modérée » étudiée entre les compétences liées à la dimension instrumentale montre que les élèves qui ont des compétences d'ouverture à l'autre développent aussi des compétences de coopération et de prise de décision. De plus, la négociation n'est pas associée à toutes les compétences de la dimension sociale. Ceci montre que les élèves ont une attitude positive à l'égard du travail en groupe et travaillent à convaincre un collègue de leur propre point de vue. Cependant, leurs choix d'arguments ne sont pas liés à un rationalisme et ne peuvent pas bien justifier leurs décisions. En outre, les compétences d'équité et de valeurs et celles de la justice sociale ne sont pas associées aux compétences de la dimension instrumentale.

Tableau 10 : Degré d'association entre le développement des compétences de la dimension sociale et le développement des compétences de la dimension individuelle

		Interculturalité/ Citoyenneté mondiale	Respect de la diversité	Empathie	Participation/ environnement social	Équité	Droits de l'Homme	Valeurs et justice sociale
Dimension Individuelle	Protection de soi	0.461	0.467	0.42	0.555	0.387	0.460	0.137
	Autogestion	0.445	0.462	0.418	0.551	0.404	0.471	0.157
	Résilience	0.379	0.383	0.332	0.454	0.344	0.399	0.116
	Communication	0.275	0.287	0.260	0.368	0.273	0.330	0.105
	Flexibilité et adaptabilité	0.420	0.432	0.389	0.512	0.371	0.440	0.142
	Développement personnel	0.179	0.172	0.164	0.223	0.133	0.173	0.024
	Protection de soi	0.461	0.467	0.42	0.555	0.387	0.460	0.137

L'analyse de « l'association modérée » étudiée entre les compétences liées à la dimension individuelle montre que les élèves qui ont des compétences d'ouverture à l'autre développent aussi des compétences de protection de soi, d'autogestion, ainsi que de flexibilité et d'adaptabilité. De plus, les résultats montrent que le fait d'être exposé à différentes cultures, engagements sociaux et aux droits de l'Homme les aide à développer des compétences de protection de soi et d'autogestion pour être conscients de ce qui est « bon » pour eux. Cependant, les résultats montrent que les compétences de communication et de développement personnel ne sont pas associées à toutes les compétences de la dimension sociale. Ceci serait un indicateur que la communication reste dans le cadre d'un engagement social et non pas dans un cadre d'échange d'information. Par ailleurs, les compétences de valeurs et de justice sociale ne sont pas associées aux compétences de la dimension instrumentale.

Ainsi, en se référant à la totalité des tests et leurs résultats, nous pouvons conclure que les compétences de la dimension sociale, spécifiquement celles de l'ouverture à autrui, influencent étroitement sur les compétences de différentes dimensions essentielles à la vie quotidienne. Leur intégration dans le programme national libanais nous semble donc indispensable.

V. Discussion

Compétences développées chez les élèves suivant le curriculum libanais

Les résultats montrent que les élèves qui suivent le programme libanais ont eu la chance de développer les compétences sociales plus que les autres dimensions. La raison pourrait être liée au décret publié en 2012 par le ministère de l'Éducation et de l'Enseignement supérieur et stipulant que chaque élève du secondaire doit « travailler » sur un projet relié au travail social afin de motiver les élèves à acquérir des compétences sociales et celles du XXI^{ème} siècle en plus de les mieux ancrer dans la société.

De plus, selon les résultats, la compétence la plus développée dans la dimension sociale est « l'équité » et cela s'inscrit dans le plan de rétablissement du secteur éducatif qui a été signé en 1994 après la guerre civile libanaise en 1975. Ce plan avait pour objectif de soutenir l'équité et l'égalité des chances d'apprendre et d'être éduqué pour chaque citoyen. Cependant, le faible pourcentage d'acquisition de compétences liées aux valeurs et à la justice est dû à l'absence de réformes visant à renforcer la cohésion sociale sur lesquelles le gouvernement successif des années 90 a

tenté de travailler et a éventuellement échoué. Cela est également clair dans Smaira (2014) citant que la transition du Liban après la guerre a été imparfaite, partielle et largement inefficace. Malgré les déclarations du gouvernement libanais concernant son attachement au droit international, Le Liban n'a fait aucune tentative sérieuse pour se conformer à ses obligations juridiques internationales de poursuivre les responsables de violations graves des droits de l'homme ou de répondre aux droits et aux besoins des victimes, et il n'a pas non plus abordé la culture d'impunité qui a envahi la société libanaise. Cela a également été confirmé par l'étude menée par Abla et al (2015) où les résultats indiquent que les élèves libanais montrent des relations tendues avec les élèves non libanais.

Pour la dimension cognitive, la compétence la plus développée est la résolution de problèmes et la compétence la moins développée est la créativité. Ceci pourrait être expliqué par le fait que le curriculum libanais, même s'il évoque l'importance du développement des compétences et notamment la résolution de problèmes, présente deux obstacles majeurs : d'une part lors de la transposition didactique, les manuels scolaires ne sont pas conçus selon une approche par compétence et d'autre part les interventions éducatives et les méthodes d'enseignement utilisées restent traditionnelles et magistrales dans la plupart des cas. Quant à la créativité, elle reste très marginalisée et très peu travaillée en classe. Ce curriculum qui date de 1997 se concentre surtout davantage sur les connaissances plutôt que sur les compétences à développer. Il vise surtout à préparer les élèves à réussir l'examen national à la fin du cycle secondaire (Gharib et Hamdan Saade, 2017).

Pour la dimension instrumentale, la compétence la plus développée est la négociation et la compétence la moins développée est la coopération. Cela reflète très évidemment la scène politique et religieuse au Liban, mais aussi le contexte social. Chaque parti politique ou religieux déploie beaucoup d'efforts pour soutenir ses convictions et convertir son adversaire à rejoindre ses groupes, mais des efforts limités sont faits pour trouver un terrain d'entente et coopérer pour réaliser et compléter des projets nationaux. Comme le mentionne Akkary (2014), le gouvernement et les décisions prises au niveau national affectent la politique et les réformes de l'éducation. Le conflit entre les différents partis politiques se répercute également sur le système éducatif. De plus, ces résultats coïncident avec ceux de Naddaf (2014) stipulant que toutes les écoles privées religieuses veulent que les élèves respectent strictement les normes mono-culturelles de l'école, et que la majorité des enseignants n'avaient aucune formation sur la gestion de la diversité religieuse en classe.

Pour la dimension individuelle, la compétence la plus développée est l'autoprotection et la compétence la moins développée est le développement personnel. Ces résultats sont probablement dus au faible niveau de sécurité auquel le Liban est confronté depuis quelques années. Les bouleversements, les explosions et les assassinats de politiciens et journalistes sont tous incrustés dans l'inconscient des jeunes. Cela coïncide avec Iskandar (2021) qui explique la diminution de la motivation des jeunes en raison aussi des opportunités limitées disponibles dans le pays.

La dimension sociale comme référence pour le développement d'autres dimensions des compétences

Concernant l'association testée entre les compétences de la dimension sociale, les résultats (Tableau 4.7) montrent que sept compétences peuvent être classées sous deux catégories. La première catégorie de compétences en lien avec l'ouverture à l'autre et l'inclusion, englobe les compétences suivantes : interculturalité, citoyenneté mondiale, respect de la diversité, empathie, participation sociale et la compétence des droits de l'Homme. En contrepartie, la deuxième catégorie de compétences en lien avec la vie sociale inclut l'équité et les compétences de valeurs et de justice sociale. Nous pouvons clairement distinguer que les compétences de la première catégorie sont liées au comportement de l'élève dans des activités avec ses pairs, alors que les compétences de la deuxième catégorie sont liées à la connaissance des lois et des règles de la vie sociale. Ce résultat coïncide avec Mancha et Ahmad (2016) qui ont étudié l'amélioration des compétences sociales en référence aux activités parascolaires. Dans leur étude, ils ont défini les compétences sociales par les sous-compétences suivantes : participation, collaboration, travail d'équipe et gestion des conflits. Le centre de l'étude était le comportement de l'élève. Les résultats de cette étude ont été énumérés en termes de compétences sociales comme une unique entité et non pas en termes de sous-compétences. Les résultats ont montré que les élèves qui participent plus à des activités parascolaires ont de meilleures compétences sociales.

L'association mobilisée entre les compétences d'ouverture à l'autre et les compétences cognitives suivantes - curiosité, pensée critique, résolution de problèmes et métacognition - serait la conséquence des multiples situations dans lesquelles un individu est placé lorsqu'il interagit avec des amis et des membres de sa société ; ce qui coïncide avec l'étude menée par Soto-Icaza et al. en 2015. Au niveau du contexte libanais, les élèves qui interagissent avec des élèves de différentes écoles, religions et affiliations politiques auraient une meilleure capacité réflexive que les élèves qui

n'interagissent qu'avec des membres de leur propre communauté. Les résultats coïncident également avec Daraee et al., (2016) notant que les élèves talentueux avaient des scores plus élevés au niveau des compétences sociales appropriées et ont un excès de confiance en soi.

Une association entre les compétences d'ouverture à l'autre et les compétences suivantes - coopération et prise de décision - est notée (tableau 4.9). Ceci coïncide avec les résultats de la recherche menée par Mancha et Ahmad (2016) qui ont défini ces compétences comme étant des sous-compétences de la dimension sociale. Ceci serait également lié aux interactions et aux activités sociales des élèves avec d'autres. Plus leur cercle d'interaction sociale est élargi, plus ils améliorent leurs compétences instrumentales.

De plus, une association entre les compétences d'ouverture à autrui et les compétences suivantes - les compétences de protection de soi, d'autogestion, ainsi que de flexibilité et d'adaptabilité - est à souligner (tableau 4.10). Des relations « saines » favorisent non seulement les compétences sociales, mais améliorent également les compétences personnelles et sont bénéfiques pour les deux éléments d'une relation. Les relations sans abus et sans intimidation (bullying) aident les élèves à évoluer et à se soutenir mutuellement pour saisir les opportunités et apprendre des expériences. Cela coïncide avec l'étude menée en 2001 par O'Moore et Kirkham. Le résultat de cette étude montre que les élèves du cycle primaire qui ont été impliqués dans l'intimidation en tant que victimes, intimidateurs ou les deux, avaient une estime de soi significativement plus faible que les enfants qui n'avaient ni intimidé, ni été intimidés.

VI. Conclusion

L'évaluation des compétences du XXI^{ème} siècle, développées chez les élèves, reste complexe dans le domaine de l'éducation et en recherche. Notre étude s'est focalisée sur l'élaboration d'une grille critériée des compétences du XXI^{ème} selon quatre dimensions d'apprentissage : la dimension cognitive, la dimension instrumentale, la dimension individuelle et la dimension sociale et sur la construction d'un instrument de mesure quantitatif de ces compétences. Cet outil de mesure (le questionnaire) a permis d'évaluer les compétences mobilisées par les élèves. Les résultats obtenus ont montré que les élèves libanais ont surtout développé leurs compétences en termes de résolution de problèmes, de

flexibilité, d'équité et de justice sociale au détriment d'autres types de compétences évaluées. Des associations entre les compétences d'ouverture à autrui et des compétences cognitives, de coopération et de négociation ont été significatives. Ces compétences sont nécessaires pour préparer les élèves à être des citoyens capables de s'investir dans le monde de demain qui demande une flexibilité et une gestion des incertitudes tenant compte du contexte et des changements possibles.

En conclusion, nous avons pu déduire que le développement de certaines compétences chez les élèves (notamment les compétences sociales) pourrait contribuer au développement d'autres compétences (instrumentales, personnelles et cognitives). Le développement des compétences chez les élèves n'est donc pas un processus linéaire et simple. Il dépend, entre autres, du contexte et des liens à construire entre différents types de compétences. Dans un contexte multiculturel et multiconfessionnel, comme celui du Liban, ce processus devient encore plus complexe. Le questionnaire critérié construit constitue une contribution simplifiant ce processus d'évaluation des compétences, même si cette évaluation nécessite la mobilisation de ces compétences dans des situations plus complexes.

VII. Bibliographie

- Abla, Z., Atoui, M. et al. (2015). *Education et diversité au Liban : Étude sur les relations interpersonnelles des élèves*. Université Saint-Joseph Press.
- Akkary, R. (2014). Facing the challenges of educational reform in the Arab world. *Journal of Educational Change*. 15. 10.1007/s10833-013-9225-6.
- Allal, L. (2003). *L'acquisition et l'évaluation des compétences dans le contexte du curriculum*. *Puzzle - Bulletin du CIFEN (Centre interfacultaire de formation des enseignants, Université de Liège)*, 13, 5-8.
- Bardin, L. (2003). *L'analyse de contenu*. Presses Universitaires de France.
- Bourn, D. (2018). *Understanding Global Skills for 21st Century Professions*. Palgrave. Macmillan. London, UK. ISBN 978-3-319-97654-9
- Care E et Griffin P. (2018). *Assessment and Teaching of 21st Century Skills Research and Applications*. Springer. New York. NY. ISBN 978-3-319-65366-2.

- Chu, S., Reynolds, R., Tavares, N., Notari, M. et Lee, C. (2017). Twenty-First Century Skills and Global Education Roadmaps. *21st century skills development through inquiry-based learning: From theory to practice*, 17-32. 10.1007/978-981-10-2481-8_2.
- Crahay, M. (2006). Dangers, incertitudes et incomplétude de la logique de la compétence en éducation, *Revue Française de Pédagogie*, n° 154, octobre 2006, 97-110.
- Daraee, M., Salehi, K. et Fakhr, M. (2016). *Comparison of Social Skills between Students in Ordinary and Talented Schools. European Proceedings of Social and Behavioural Sciences* (p. 512-521). Future Academy. <https://doi.org/10.15405/epsbs.2016.11.52>
- EducationalRC.org. (2020). *International Arab Baccalaureate*. <https://www.educationalrc.org/IAB>
- El Hage, F. (2005). *Le morcellement des connaissances en physiologie : du constat à la remédiation. Intégration du paradigme de la complexité dans l'étude de la construction des liens entre différents concepts enseignés en physiologie, aux niveaux des pratiques enseignantes et des productions des élèves*. [thèse de Doctorat en didactique de la biologie.] Université de Montpellier 2 et Université Saint-Joseph.
- Fisser, P. et Thijs, A. (2015). *Integration of 21st century skills into the curriculum of primary and secondary education*. National Institute for Curriculum Development the Netherlands. presented at the Society for Information Technology & Teacher Education International Conference, Mar 5, 2015 in Las Vegas, NV, United States.
- Gerard, M.F. (2005). L'évaluation des compétences par des situations complexes, *Actes du Colloque de l'Admee-Europe*, IUFM Champagne-Ardenne, Reims, 24-26 octobre 2005.
- Gharib Y., et Hamdan Saadé N., (2017). L'enseignement supérieur au Liban : un paysage complexe. *Revue internationale d'éducation de Sèvres*. URL : <http://journals.openedition.org/ries/6029> ; DOI : <https://doi.org/10.4000/ries.6029>
- Hadar, L. L., Ergas, O., Alpert, B. et Ariav, T. (2020). Rethinking teacher education in a VUCA world: student teachers' social-emotional competencies during the Covid-19 crisis. *European Journal of Teacher Education*, 43(4), 573-586.
- Halloun, I. (2011). Profile Shaping Education: A paradigm shift in education to empower students for success in modern life. *Proceedings*

of the 11th International History, Philosophy and Science Teaching Group biennial conference, (p. 337-343). IHPST.

Institut National de Santé Publique du Québec. (2010). *Référentiel de valeurs pour soutenir l'analyse éthique des actions en santé publique*. 24. https://www.inspq.qc.ca/pdf/publications/2010_Referentiel_Valeurs_Analyse_Ethique.pdf

International Arab Baccalaureate. (2020). *International Arab Baccalaureate (IAB)*. <https://educationalRC.org/IAB>

International Center for Transitional Justice (2014). *Failing to Deal with the Past: What Cost to Lebanon*. www.ictj.org

Iskandar, M. (2021). *Lebanon's Six Years of Darkness 2014 – 2020*. M.I. Associates.

Jorro, A. (2006). Devenir ami critique. Avec quelles compétences et quels gestes professionnels. *Mesure et évaluation en éducation*. 29. 10.7202/1086966ar.

La Fonction publique de l'Ontario (2016). *Définir les compétences du 21^e siècle pour l'Ontario compétences du 21^e siècle - document de réflexion*. Ontario. Canada. ISBN 978-1-4606-5841-3

Le Boterf, G. (2002). Développer la compétence des professionnels. Construire les parcours de professionnalisation. (4^e éd.). *Editions d'Organisation*.

Mancha, S. A. et Ahmad, A. (2016). Co-curricular activities and its effect on social skills. *International conference on education and regional development*, 774-781.

Mason, M., 2008. Complexity Theory and the Philosophy of Education. *Educational Philosophy and Theory* 40 (1), 4-18.

Morin, E. et Le Moigne, J.-L. (1999). *L'intelligence de la complexité*. L'Harmattan.

Naddaf C. (2014). *An Explorative Study on the Approach of Private Schools in Lebanon: Preparing Students for a Society with Religious Diversity*. [submitted Thesis to the Department of Education of the Faculty of Arts and Sciences at the American University of Beirut]. Lebanon.

- National Research Council. (2010). *Exploring the Intersection of Science Education and 21st Century Skills: A Workshop Summary*. The National Academies Press. <https://doi.org/10.17226/12771>
- O'Moore, M. et Kirkham, C. (2001). Self-esteem and its relationship to bullying behaviour. *Aggressive Behavior*, 27(4), 269–283. <https://doi.org/10.1002/ab.1010>
- OECD. (2019). *Les compétences du XXI^e siècle : Enseigner et évaluer des compétences transversales*. Éditions OCDE.
- Parent, F., Baulana, R., Coppieters, Y., Kahombo, G., d'Hoop, E., Lemenu, D., Garant, M. et de Ketele, J.-M. (2010). Mieux gérer la cohérence pour renforcer les ressources humaines en santé : Paradigmes et méthodes pour une intégration efficace des pratiques professionnelles et communautaires en formation. *Pédagogie Médicale*, 11(2), 111–125. <https://doi.org/10.1051/pmed/2010014>
- Smaira, D. (2014). *Changing Security: Theoretical and Practical Discussions. The Case of Lebanon*. [thèse de doctorat], Durham University.
- Soland L., James S. et Hamilton B., (2013). *Measuring 21st-century competencies: Guidance for educators*. Asia Society
- Soto-Icaza P., Aboitiz F. et Billeke P. (2015). Development of social skills in children: neural and behavioral evidence for the elaboration of cognitive models. *Front. Neurosci.* 9:333. DOI: 10.3389/fnins.2015.00333
- Tardif, J. (2006) L'évaluation des compétences : De la nécessité de documenter un parcours de formation. (Assessment of Skills: About the Need to Document a Training Path.). Chenelière Education.
- UNICEF. (2017). *Reimagining Life Skills and Citizenship Education in the Middle East and North Africa. A Four-Dimensional and Systems Approach to 21st Century Skills: Conceptual and Programmatic Framework*. UNICEF MENA Regional Office United.

Chapitre 11. L'évaluation au prisme de la contextualisation : l'exemple du système éducatif français

Vers une théorie de l'évaluation contextualisée ?

Antoine Delcroix

I. Introduction

Dans les pays présentant simultanément des contextes géographiques et sociétaux variés et un système éducatif centralisé, la question de l'évaluation se pose de manière particulière en raison des articulations entre les différents niveaux d'intervention : celui de l'État, souvent le principal prescripteur, celui des divisions territoriales disposant parfois de peu de marges dans l'exécution des décisions nationales, et celui des acteurs de terrain, pris entre le respect des dispositions réglementaires et le besoin d'adapter leurs interventions et les évaluations correspondantes aux caractéristiques de leur lieu d'exercice.

Cette question ne peut faire l'économie de réinterroger les finalités assignées au système éducatif par le pouvoir politique et, en particulier, les connaissances, compétences et valeurs que ce système a pour mission de transmettre. Il s'agit donc d'insérer la question de l'évaluation dans une perspective plus globale, conforme à l'approche de la transposition de Verret (1975), qui interroge l'ensemble de la chaîne de transposition, depuis les savoirs et les valeurs jusqu'aux évaluations que le système met en place. Sur ce dernier point, il est autant question de mesurer les acquisitions des élèves que de vérifier l'accomplissement par les institutions (établissements scolaires, universités...) de leur mission. En bref, il s'agit de s'assurer, pour l'État, que le système éducatif, par les transpositions qui s'y opèrent, joue socialement le rôle qui lui est confié.

Nous présentons deux exemples pris dans le système éducatif français, un prototype d'organisation très centralisée s'adressant à un pays aux contextes très divers, notamment par ses outre-mer. Le premier est celui des examens permettant l'obtention de diplômes nationaux. Les enseignants sont confrontés à la nécessité de préparer au mieux leurs étudiants à ces épreuves nationales, tout en assurant une connaissance suffisante des caractéristiques propres au territoire de formation qui sera fréquemment le lieu de travail des lauréats. Le second porte sur l'évaluation d'une université d'outre-mer par l'autorité en charge de ce processus. Les démarches d'assurance-qualité conduisent souvent à des processus normalisés et standardisés, par les critères d'évaluation retenus et par l'infrastructure technique associée. Ainsi, les acteurs locaux peuvent les vivre comme inappropriés, voire injustes, ce qui peut provoquer une défiance à leur égard.

Ces exemples s'inscrivent dans un débat plus général sur l'articulation entre le local et le global, maintenant bien documenté dans le champ de la contextualisation didactique. Notre espoir est de contribuer au questionnement sur la nécessité d'une théorie de l'évaluation contextualisée, qui peut trouver ses bases dans les différentes contributions du 33ème colloque de l'ADMEE-Europe, et viendrait en complément des courants portant sur l'évaluation située.

II. Les outre-mer français comme terrain d'investigation

Sans entrer dans une définition géographique ou politique des outre-mer européens, nous soulignons cependant qu'une dialectique centre-périphérie s'y exprime de manière variée en fonction du degré d'autonomie dont chaque territoire dispose vis-à-vis du pays européen dont il dépend ou a dépendu. Des aspirations vers une capacité renforcée de s'administrer librement ou d'indépendance totale s'expriment. Ainsi, des évolutions contrastées se dessinent, comme le détachement de la Barbade de la Couronne britannique en 2021 ou la réalisation de référendums d'autodétermination en Nouvelle-Calédonie, le dernier dans des conditions contestées. Les statuts variés de ces territoires entraînent des conséquences sur leurs systèmes d'enseignement scolaire et universitaire : ils peuvent parfois être organisés comme dans l'ancienne puissance coloniale et administrés localement par une autorité décentralisée ou, à l'inverse, davantage dépendre du niveau territorial,

avec certains niveaux régis localement et d'autres non. En général, ils conservent des traits d'organisation issus de l'ancienne puissance coloniale.

La majeure partie des travaux dont est issue cette contribution a été effectuée en Guadeloupe où le système est organisé comme dans l'Hexagone, la représentation de l'État étant assurée, en matière éducative, par la Région académique. Les possibilités d'adaptation des programmes restent assez marginales et concentrées sur certaines disciplines (langue et culture régionale : créole, histoire-géographie, sciences de la vie et de la terre). Nous évoquons également, pour une des investigations dont il est rendu compte ci-après, la Polynésie française qui relève plutôt du deuxième type d'organisation décrite ci-dessus, par la présence d'un gouvernement local doté d'un ministère de l'éducation (Robo, 2021). Cependant, les possibilités ainsi offertes sont encore peu exploitées, les programmes scolaires y restant très proches de ceux de l'Hexagone et l'enseignement supérieur échappant largement au gouvernement local.

III. De la modélisation des contextes à la transposition didactique

Recenser et décrire les différentes modélisations du contexte est un travail quasiment impossible (Forissier, 2019), tant elles sont nombreuses. Nous nous limitons donc à quelques modèles qui nous semblent correspondre au mieux aux phénomènes en jeu dans les pratiques professionnelles et institutionnelles en lien avec l'évaluation.

Les modèles binaires opposent contextes interne et externe. Le contexte interne se situe à une échelle micro. Pour certains auteurs, il est « celui de l'établissement, de la classe, de la discipline, des différents acteurs de la relation pédagogique, du savoir enseigné, souvent lié aux aspects institutionnels, politiques et éducatifs » (Anciaux et al., 2013, p. 11). Pour d'autres, il est placé à l'échelle de l'individu, se rapprochant alors des conceptions (Giordan et De Vecchi, 1987), du *déjà-là* (Dorville, 2005) ou du *déjà-su* (Verret, 1975). Le contexte externe, « relatif aux aspects sociologiques, écologiques et culturels de l'environnement » (Anciaux et al., 2013, p. 11), se situe à l'échelle d'une division territoriale du système éducatif ou du système dans son ensemble. D'autres auteurs, comme Duranti et Goodwin (1992), ajoutent l'idée d'interactions, dans une situation d'apprentissage, entre un *focal event* – l'événement focal, à

rapprocher du centre de la focale de Blanchet (2016) – et un *field of action* – le champ d'action où se déroule l'action concernée.

Ces oppositions assez binaires se montrent souvent efficaces, mais peuvent s'avérer insuffisantes pour décrire l'ensemble des paramètres pertinents éclairant le centre de la focale. De nombreuses modélisations ont donc été développées dont Bazire et Brézillon (2005) et Forissier (2019) livrent des méta-analyses. Nous mentionnons simplement qu'on gagne parfois à remplacer l'opposition interne *versus* externe par un découpage plus fin du contexte, par exemple en termes de micro, méso, et macro-contextes.

Les phénomènes de contextualisation s'inscrivent en outre au sein des mécanismes de transposition didactique (Delcroix et al., 2013), que Verret (1975) décrit comme comportant cinq phases : 1) la désyncrétisation, qui restructure les savoirs destinés à être enseignés en disciplines d'enseignement, dépendant fortement des politiques éducatives ; 2) la dépersonnalisation, une décontextualisation qui sépare « le savoir de la personne » (Verret, 1975, p. 146) ; 3) « la programmation des apprentissages et des contrôles suivant des séquences raisonnées permettant une acquisition progressive des expertises [...] » (Verret, 1975, p. 146-147), ce que Verret appelle « la programmabilité de l'acquisition des savoirs » ; 4) la publicité, au travers de référentiels, de programmes, qui permet la communication à la société des intentions des décideurs ; 5) et, finalement, le contrôle des acquisitions, défini par Verret comme étant « le contrôle réglé des apprentissages suivant des procédures de vérification autorisant la certification des compétences – c'est-à-dire le contrôle social des apprentissages » (Verret, 1975, p. 146-147).

C'est ce dernier point qui nous intéresse plus particulièrement : la société, par l'intermédiaire de la représentation nationale, de la *bureaucratie* (Verret, 1975) ou de la *Noosphère* (Chevallard, 1985), a déterminé des savoirs à faire acquérir, s'est organisée pour les faire acquérir et doit s'assurer de cette acquisition au travers de procédures de vérification.

Si, dans les propos de Verret, le contrôle des acquisitions repose sur l'évaluation des productions des élèves, on peut également penser au *contrôle social des institutions* délivrant ces savoirs. Ainsi, l'évolution des politiques publiques, issue du New Public Management (Hood, 1991), se traduit en France par l'adoption de la loi organique relative aux lois de finances (LOLF) (Arkwright et al., 2007) qui consacre le passage d'une logique de moyens à celle de résultats. Il faut donc pouvoir évaluer la performance des opérateurs de l'État, conduisant à « la production d'indicateurs dans toutes les organisations sous contrôle direct ou indirect de l'État » (Bessire et Fabre, 2014) et à la refonte des agences et autorités en

charge de l'évaluation. On peut citer, dans le champ qui nous intéresse, la création de l'Agence pour l'évaluation de la recherche et de l'enseignement supérieur (AERES) en France en 2007, transformée en 2013 en Haut Conseil de l'évaluation de la recherche et de l'enseignement supérieur (Hcéres). Si, assez souvent, les évaluations sont rendues publiques (c'est le cas pour le Hcéres), il n'est pas clairement établi que le Parlement s'en saisisse et qu'un contrôle social réel des institutions soit fait.

IV. Contexte territorial et diplômes nationaux : l'exemple du Brevet de technicien supérieur (BTS)

Le premier exemple est celui des examens permettant l'obtention de diplômes nationaux, dont certaines ou la totalité des épreuves sont gérées nationalement, de l'élaboration des sujets à l'évaluation des productions des candidats. Dans une série de travaux, Odacre (2018, 2021) et Odacre et al., (2021a, 2021b) ont exploré le fonctionnement de certains BTS en Guadeloupe. Nous reprenons ici, pour partie, les données de ces travaux, dans le cadre d'une analyse centrée sur les questions d'évaluation. Le BTS est la plus importante formation courte de l'enseignement supérieur en France par ses effectifs. Valorisé par 120 crédits européens (ECTS), le BTS procure à ses lauréats « une qualification professionnelle les rendant aptes à exercer la fonction de technicien supérieur dans les professions industrielles et commerciales, dans les activités de service ou celles relevant des arts appliqués³² ». Ce diplôme, dont les spécialités couvrent la plupart des secteurs économiques, est préparé dans tout le territoire national, dont les outre-mer. Sa première vocation, l'insertion professionnelle immédiate, fait que la carte des formations de chaque territoire doit s'adapter au marché local du travail (Lemistre, 2015). Néanmoins, les différentes instructions relatives au BTS sont « élaborées à l'échelon national, sans qu'il soit prévu de pouvoir les moduler en fonction des contextes locaux d'enseignement et d'insertion professionnelle » (Odacre et al., 2021b, p. 32). Il en est de même pour les épreuves écrites dont l'organisation tient cependant compte des différents calendriers scolaires (ce qui constitue une forme de contextualisation) sans que les sujets soient adaptés aux contextes des différents centres d'examen.

³² Voir <https://www.enseignementsup-recherche.gouv.fr/fr/brevet-de-technicien-superieur-bts-45915>, consulté le 7 février 2022.

L'intérêt des auteurs précités pour des BTS proposés en Guadeloupe à fort contenu en sciences de gestion provient, en particulier, de l'existence de spécificités sur le plan fiscal dans ce territoire comme dans les autres outre-mer. Loin d'être anecdotiques, ces dispositions s'enracinent dans l'histoire coloniale, ce qui peut exercer une influence sur les conceptions des acteurs locaux de l'éducation. Il s'agit, par exemple, de l'octroi de mer, dont l'origine remonte au 17^e siècle (M'Saïdié, 2013) et qui, en taxant certains produits d'importation, vise à protéger la production locale. Si l'efficacité économique de ce dispositif reste à démontrer, son existence possède une influence sur la gestion des entreprises. Ainsi, il semble important que les futurs diplômés de BTS, comme le BTS *Support à l'Action Managériale* (anciennement *Assistant de Manager*) connaissent les dispositifs applicables en Guadeloupe (dont l'octroi de mer), ainsi que les règles générales nationales. Ceci nous conduit à préciser la question de recherche : comment les enseignants se positionnent-ils face au dilemme de favoriser une insertion locale (au travers d'un enseignement contextualisé) ou la réussite d'un examen national (qui ne prend pas en compte le contexte) ?

L'enquête menée par Odacre (2018) porte sur les pratiques de six professeurs intervenant en Guadeloupe dans les enseignements de Sciences de gestion : trois pour le BTS *Assistant de Manager* (la moitié des intervenants du champ disciplinaire concerné dans ce diplôme) et trois pour le BTS *Tourisme* (la totalité des intervenants). Si l'évaluation des étudiants n'était pas l'objectif premier des entretiens semi-directifs menés, la question des interactions délicates entre les contextes nationaux et territoriaux que nous venons d'évoquer a naturellement conduit à ce qu'elle soit abordée. Cette enquête est complétée par quatre entretiens menés en 2018 par l'auteur, selon des modalités analogues, auprès d'enseignants de Sciences de gestion intervenant en Polynésie française dans le Supérieur.

L'exploitation des données issues de ces entretiens montre que l'attitude des professeurs interrogés (micro-contexte) s'inscrit entre deux pôles : 1) une centration sur la tâche, qui doit être adaptée au contexte de l'étudiant, d'ailleurs plus à des fins motivationnelles que pour préparer l'insertion professionnelle ; 2) une centration sur l'objectif, en l'occurrence la réussite au diplôme, plus que l'acquisition des compétences visées par celui-ci. Pour illustrer ces deux pôles, nous prendrons l'exemple de deux professeurs intervenant en Guadeloupe dans le même BTS et qui déclarent des pratiques différentes. Ainsi, pour les notions relatives à la taxe sur la valeur ajoutée (TVA), une professeure exploite des ressources locales mais reste néanmoins consciente que « le taux de taxe sur la valeur ajoutée n'est pas le même ici, il y a un octroi de mer qu'il n'y a pas en Métropole. » Son

but, dans cette utilisation de ressources locales, est essentiellement « d'intéresser ses élèves ». À l'opposé, un de ses collègues va baser son enseignement sur les taux de TVA en vigueur dans l'Hexagone. Il estime, en effet, qu'enseigner les taux locaux pourrait handicaper les étudiants pour l'examen terminal. De ce fait, ce professeur regrette de ne pas parler de l'octroi de mer évoqué ci-dessus : « On ne parle pas d'octroi de mer, [...] c'est un petit peu dommage ! ». Un des enseignants interrogés en Polynésie mentionne son évolution d'un des pôles vers l'autre. Ce professeur a fait plusieurs mutations entre l'Hexagone et la Martinique (au régime similaire à celui de la Guadeloupe), sans d'abord modifier ses pratiques. Sa prise de conscience est intervenue, en Martinique, dans le cadre de formations professionnelles qu'il animait où un stagiaire lui fait remarquer que « Ce n'est pas ça ici ». Il précise alors la solution pragmatique qu'il prend : « J'ai fait faire des "Bristol" [fiches cartonnées] à mes élèves, à gauche : ce que tu dis en France, à droite, ce que tu fais quand tu vas ici dans l'entreprise ». Poursuivant par ses pratiques en Polynésie, il note les limites de son dispositif : les écarts de contextes (il cite le statut d'autonomie du territoire, la situation géographique...) sont tellement plus importants qu'il « faudrait faire autre chose ». Ces trois exemples de pratiques témoignent d'un constat qui peut être étendu, avec des nuances, à sept des dix professeurs. Tous ont conscience de cette tension entre les prescrits et les examens nationaux d'une part et les contextes territoriaux d'autre part.

La première limite à la contextualisation apparaît donc clairement au travers des pratiques professionnelles : que l'enseignant se centre sur la tâche ou sur la réussite au diplôme, la délicate articulation entre le contexte d'apprentissage et l'évaluation nationale est unanimement évoquée. Ceci est résumé par les propos d'une des professeurs de la Guadeloupe : « [...] on n'a pas à trop contextualiser puisque c'est un enseignement, c'est un examen national ». Une autre limite, même si elle est moins directement liée aux questions d'évaluation, mérite d'être mentionnée. Il s'agit du manque de ressources pédagogiques basées sur les contextes territoriaux, ce qui ne permet pas d'aller aussi loin que le souhaiteraient les enseignants. Ainsi, un professeur de Guadeloupe indique que « [...] les manuels nous obligent à travailler sur des sujets qui sont plus généraux, c'est vrai ». Un des professeurs interrogés en Polynésie va dans la même direction et complète en indiquant : « [...] tu pars d'un article [de la presse locale] ou d'une note de l'ISPF [Institut de la statistique de la Polynésie française], mais tu n'as pas toute l'analyse [soupir] ou ça demande du travail pour filtrer ». Cet enseignant développe ensuite deux contraintes complémentaires, le temps (variable contextuelle) à consacrer à la

production d'un enseignement contextualisé (dans le cadre d'un engagement de durée limitée en Polynésie) et le manque de ressources pédagogiques.

V. Une évaluation des établissements d'enseignement supérieur contextualisée ? Le cas de l'université des Antilles

Le deuxième exemple est celui de l'évaluation de l'université des Antilles (UA) par le Hcéres. Avant de poursuivre l'exposé, nous mentionnons une des limites de l'étude, due à la position de l'auteur qui était conseiller du Président de l'université sur la démarche d'évaluation pendant la réalisation de celle-ci. Ce rôle lui confère le statut d'*observateur participant* interne dont les limites sont connues en particulier en matière de distanciation et de déontologie (Pereira, 2015). L'auteur est également tenu, dans le cadre d'une démarche d'éthique et d'intégrité, de ne pas faire état de ce qui pourrait nuire à l'université ou aux personnes.

Chaque établissement d'enseignement supérieur se veut singulier au travers de son identité, sa *marque* qui lui permet de « pouvoir fixer un positionnement performant dans le marché universitaire international » (Aguerreberre, 2014, p. 3). Si l'UA n'échappe pas à cette règle, et se voit reconnaître un positionnement cohérent (Hcéres, 2022), les particularités du statut de cette université sont au centre de notre questionnement. Dans son histoire récente, l'UA a vu son traitement relever d'un parcours législatif particulier : lors de l'adoption de la loi relative aux libertés et responsabilités des universités (LRU) en 2007, le gouvernement traite l'UA, à l'époque université des Antilles et de la Guyane (UAG), à part : le Parlement va ainsi autoriser le gouvernement à prendre, sous forme d'ordonnance, des mesures directement applicables à l'UAG avec force de Loi. Un des facteurs explicatifs de ce choix gouvernemental était l'introduction par la LRU de relations plus étroites entre les universités et les régions d'implantation. Or, l'UAG était la seule université implantée sur plusieurs régions, ce qui demandait des adaptations pouvant retarder le processus législatif. En 2013, ce choix sera réitéré lors de l'élaboration de la Loi relative à l'enseignement supérieur et à la recherche, dans un contexte de territorialisation accrue de l'enseignement supérieur et de la recherche (Morin, 2016). Pour l'UA, la période 2013-2015 sera marquée par plusieurs événements, dont la création d'une université de plein exercice en Guyane au 1^{er} janvier 2015, avec – en miroir – des demandes de création

de deux universités, l'une en Guadeloupe et l'autre en Martinique. En réponse à ces demandes, une organisation particulière sera discutée puis adoptée par voie législative en 2015. Cette voie solennelle peut sembler sceller l'unité de l'établissement, mais rend les adaptations ultérieures plus complexes car elles doivent à nouveau passer par ce même processus. La disposition phare, usuellement appelée l'autonomie des pôles, prévoit que les implantations territoriales (les pôles) de l'UA en Guadeloupe et en Martinique voient leurs prérogatives renforcées. Ainsi, chaque pôle dispose des conseils et comités présents dans toute université : un conseil de pôle, sorte de conseil d'administration territorial, une commission recherche (CR), une commission formation et vie universitaire (CFVU). Cependant, il n'est pas prévu de conseil académique polaire, alors qu'il existe un conseil académique d'établissement formé du rassemblement des deux CR et des deux CFVU, le conseil d'administration (CA) de l'établissement étant la réunion des deux conseils de pôle. Le comité social d'administration (CSA), instance consultée notamment pour les questions relatives à l'organisation et au fonctionnement des services, se décline également en deux CSA spéciaux de pôles. Ainsi, l'UA possède les instances classiques de toute université française, mais dans une structuration plus complexe. Ces différentes mesures aboutissent à un établissement dont l'administration est très dérogatoire, ce qui induit les éléments suivants en relation avec son évaluation.

L'évaluation des universités françaises est réalisée tous les cinq ans par le Hcéres et porte sur trois thèmes clés : 1) le pilotage de l'établissement ; 2) la formation, évaluée ex post (évaluation du contrat antérieur) et ex ante (évaluation des projets pour le nouveau contrat), en deux temps successifs ; 3) la recherche, évaluée en un temps unique, au niveau de chaque unité de recherche. L'établissement peut demander au Hcéres de porter son attention sur des points particuliers et l'UA avait choisi de formuler deux attentes particulières dont nous reproduisons ici la première :

« Une université dont le statut dérogatoire confirmé depuis une douzaine d'années par les textes législatifs successifs est source d'instabilité et de complexité dans l'organisation et le fonctionnement de l'établissement : quel regard du comité vis-à-vis de l'organisation mise en place par l'université (modes de gouvernance, pilotage, organisation interne) dans le cadre des dispositions statutaires et juridiques actuelles ? » (Hcéres, 2022, p. 5)

Nous ne procédons pas ici à une analyse exhaustive du rapport du Hcéres, mais portons notre regard sur des exemples de traitement de cette attente,

dans la perspective d'alimenter la question de la relation entre contexte et évaluation.

Un premier exemple est celui du conseil des directeurs de composantes pédagogiques, c'est-à-dire des Unités de formation et de recherche (UFR), des Instituts nationaux supérieurs du professorat et de l'éducation (INSPE) et des Instituts universitaires de technologie (IUT) dont le rapport indique que :

« Le conseil des directeurs des composantes est convoqué et animé par le président, sans que le VPCA et les VP des pôles y participent. Ainsi, il est difficile de comprendre l'articulation entre ce conseil et les conseils de pôles qui arbitrent les demandes de moyens des composantes de ces pôles » (Hcéres, 2022, p. 12)

Cette remarque paraît en effet fondée, non seulement d'un point de vue fonctionnel, mais presque d'un point de vue réglementaire. En effet, juridiquement parlant, les pôles sont des méta-composantes rassemblant les composantes implantées sur le pôle correspondant. En restant sur l'articulation établissement/pôle, le rapport formule également une remarque qui justifie, selon les experts, l'absence de réponse aboutie à cette attente de l'établissement :

« En l'état actuel, et compte tenu du manque d'informations relatives au fonctionnement administratif au sein des pôles, le comité ne peut pas affirmer que le mode de pilotage choisi aboutisse à un équilibre entre l'unité de l'établissement "université des Antilles" et l'autonomie des deux pôles. » (Hcéres, 2022, p. 14)

On remarque ici la prudence du comité, qui n'indique pas clairement le responsable du manque d'information. La teneur générale du rapport peut laisser penser qu'il s'agit de l'établissement. Cependant, un problème méthodologique se pose ici : est-ce à l'établissement de fournir *a priori* des informations avec le risque d'introduire un biais ? Est-ce au comité de penser une méthodologie spécifique pour traiter cette question, inédite pour une université ? Cette prudence énonciative s'inscrit dans une tradition de lissage que possède de nombreux rapports du Hcéres (Barats et Née, 2020). L'attente d'information de la part de l'établissement s'inscrit dans une autre dimension : le Hcéres produit des *méta-rapports* (Barats et Née, 2020) ou, de manière plus précise, des rapports sur le rapport d'autoévaluation (RAE) que l'établissement doit fournir préalablement au processus d'évaluation proprement dit.

Le rapport se penche également sur l'évaluation du pilotage de l'offre de formation (OF), qui donne notamment lieu au paragraphe suivant :

« Le pilotage politique et opérationnel de l'OF n'est pas réalisé de façon optimale. Le RAE et les entretiens ont montré qu'il n'existe pas d'instance, commune aux deux pôles, de pilotage de l'offre de formation tout au long de la vie. En effet, comme pour la commission de la recherche, il existe deux CFVU, une pour chacun des pôles, mais il n'existe pas d'instance démocratique représentative qui permettrait de définir une vision globale de l'offre de formation et d'opérer une régulation de cette offre, qu'elle soit initiale ou continue, à l'échelle de l'établissement. Le comité recommande à l'établissement d'envisager la fusion des deux CR et CFVU ou, au minimum, de renforcer le pilotage transverse de ces instances. » (Hcéres, 2022, p. 26-27)

Si beaucoup d'acteurs de l'UA considèrent le système à deux CR et deux CFVU lourd, le rapport oublie ici que cette disposition relève la Loi et que la fusion n'en est donc pas possible, sauf intervention du Parlement. Le renforcement du pilotage transversal était déjà acté par l'établissement au moment de l'évaluation (septembre 2021) en lien avec la politique de formation pour la période 2022-2027, votée par les instances de l'établissement en mai 2021³³. Cette méconnaissance des statuts de l'établissement, et surtout l'oubli de la complexité de leur modification, laisse à penser que le comité d'experts ne s'est pas réellement approprié l'organisation dérogatoire de l'université ou plutôt qu'il n'a pas totalement mesuré la part revenant à l'établissement et celle revenant à l'État. Ceci pose donc la question de l'adaptation de l'évaluation menée au contexte particulier de l'établissement.

VI. Discussion et conclusion : vers une évaluation plus contextualisée ?

Plusieurs auteurs convergent vers l'idée que l'évaluation est nécessairement située, comme Mottier Lopez et Dechamboux (2017) qui indiquent que « les processus d'interprétation et de prise de décision qui

³³ Voir, sur le site de l'université, les délibérations du conseil d'administration : <http://www.univ-antilles.fr/decouvrir-luniversite-des-antilles/actes-administratifs-reglementaires>. Le paragraphe réformant le mode de pilotage de la formation figure à la page 11 du document « Présentation du projet d'offre de formation - premier cycle » et à la page 10 du document « Présentation du projet d'offre de formation - deuxième cycle », documents annexés à la délibération n°2021-42. Le comité d'experts du Hcéres était informé de la présence des délibérations des conseils sur le site de l'université.

sous-tendent le jugement en acte de l'enseignant s'inscrivent dans des contextes pluriels qui participent pleinement à la dynamique de co-constitution référentielle [...] » (p. 23). Si les cadres théoriques qui sous-tendent les approches situées de l'évaluation et celles de la contextualisation didactique sont partiellement différents, les objets d'étude et les objectifs sont similaires, liés à la compréhension des situations à travers les différents paramètres en interaction. Sans aboutir à une définition d'évaluations contextualisées, ni *a fortiori* aux possibilités de leur mise en œuvre concrète, nous souhaitons reprendre certains des traits des exemples présentés plus haut pour discuter de quelques conditions, caractéristiques qu'elles pourraient prendre et à quelles limites elles risquent de faire face.

Mouraz et al., (2012) explorent la notion de territorialisation du curriculum qui vise à favoriser les apprentissages en structurant « le socle de départ [...] par les savoirs détenus par le jeune ou par les expériences qu'il aurait vécues intensément » (p. 32), proche du *déjà-su* (Verret, 1975). Ces auteurs insistent sur deux dimensions, dont l'une est relative au macro-contexte qui doit viser « la définition des politiques éducatives et des programmes scolaires faisant la promotion d'une formation pour la diversité » (p. 36). L'autre dimension porte sur les micro-contexte et méso-contexte et les actions locales, y compris concernant la définition du curriculum et sa déclinaison en termes d'action éducative. Ces auteurs soutiennent que la contextualisation doit privilégier « le recours à des activités et à des expériences d'apprentissage proches de la réalité des élèves en augmentant ainsi les chances de succès pour tous » (Mouraz et al., 2012, p. 34). Cette approche permet de comprendre les enjeux et les conditions d'une territorialisation du curriculum, autrement dit d'une dévolution au moins partielle de la programmation des apprentissages aux contextes intermédiaires. Celle-ci doit également s'accompagner de la production de ressources pédagogiques, selon ce que les professeurs de BTS évoquaient, et de la territorialisation de l'évaluation, pour opérer le contrôle social des acquisitions. Ceci interroge également le macro-contexte : jusqu'à quel point, dans un pays comme la France, l'État est-il disposé à confier aux échelons territoriaux la responsabilité de l'évaluation de façon plus étendue qu'actuellement ? Parmi les raisons de cette prudence figure une contrainte sociale, variable selon le type d'examen ou de concours : toute dévolution à un échelon plus local de l'évaluation crée des suspicions sur la qualité des diplômes délivrés. Ainsi, la territorialisation de l'évaluation est sans doute nécessaire, mais son acceptation sociale et politique en France n'est pas acquise.

Dans cette même relation entre le local et le national, l'État a renforcé l'autonomie de certains opérateurs et services décentralisés qui doivent alors rendre compte de leur action au travers d'une évaluation de leur performance. À cette fin, l'existence de référentiels et de protocoles d'évaluation normalisés, basés sur des indicateurs quantitatifs, est jugé nécessaire pour établir des bases de comparaison (Boulet et Harari-Kermadec, 2014). D'un autre côté, les éléments plus qualitatifs demandent une connaissance fine du contexte de chaque institution, ce qui a fait l'objet de l'analyse effectuée *supra* dans le cadre de l'évaluation de l'UA. Deux dimensions peuvent être évoquées ici. Il s'agit en premier d'une composante spatiale, puisque – dans l'évaluation de l'UA par le Hcéres – les échanges avec les différents acteurs des établissements se sont effectués en visio-conférence en raison de la crise sanitaire. En matière scientifique, l'évaluation sur documents est, sinon la règle, du moins très fréquente (évaluations d'articles scientifiques, de projets de recherche...). Les évaluations du Hcéres reposent en premier lieu sur le RAE, complété par des entretiens avec différents acteurs de l'établissement et de son environnement. La réalisation de ces entretiens entièrement à distance pose des questions similaires à celles de l'enseignement à distance quant à la qualité des interactions. Comme le souligne Ferone (2019), des appréciations paradoxales sont portées sur ce point, souvent en fonction de l'opinion préalable des différents acteurs sur le numérique. La seconde dimension, temporelle, interroge les conditions de réalisation des évaluations. En effet, l'analyse et la compréhension des contextes particuliers peut nécessiter un temps non compatible avec celui de l'organisation de l'évaluation. Dans l'exemple retenu ici, il s'agit non seulement de l'étude des statuts particuliers de l'UA et de leurs implications sur le fonctionnement de l'établissement, mais également des contextes régionaux.

En conclusion, les exemples développés ici semblent plaider pour des évaluations plus contextualisées pour satisfaire aux besoins de formation et de développement des territoires. Ceci demande une réflexion sur l'articulation entre les niveaux locaux et globaux : il s'agit d'une voie assez étroite entre les exigences de certification de compétences (pour la délivrance d'un diplôme) ou de capacité à répondre à des normes nationales ou internationales (pour un établissement). À ces éléments, s'ajoutent des résistances sociales (la crainte de diplômes dévalués), institutionnelles (la crainte d'évaluations locales de complaisance) et pédagogiques ou scientifiques : il s'agit de la nécessité de produire des ressources d'enseignement, des protocoles de recherches ou d'évaluation, prenant en compte les contextes, pour reprendre Blanchet (2016) qui

affirmait la nécessité de contextualiser y compris les méthodes. Nous espérons donner envie aux lecteurs d'explorer ce champ des évaluations contextualisées et situées, au profit du développement des compétences nécessaires à chaque territoire.

VII. Bibliographie

- Aguerrebere, P. M. (2014). La construction de la valeur de marque des universités à travers les réseaux sociaux. Le cas de Facebook. *Qua-
derni*, (85). DOI : <https://doi.org/10.4000/quaderni.841>
- Anciaux, F., Forissier, T. et Prudent, L. F. (2013). *Contextualisations didac-
tiques. Approches théoriques*. L'Harmattan.
- Arkwright, E., De Boissieu, C., Lorenzi, J.-H. et Samson, J. (2007). *Économie
politique de la LOLF. Rapport pour le Conseil d'Analyse Économique*.
La documentation française.
- Barats C. et Née, É. (2020). Structurer, encadrer et lisser l'évaluation : les
écrits de l'AERES/HCERES. *Communication & langages*, 203(1),
5-25. URL : [https://www.cairn.info/revue-communication-et-
langages-2020-1-page-5.htm](https://www.cairn.info/revue-communication-et-langages-2020-1-page-5.htm)
- Bazire, M. et Brézillon, P. (2005). Understanding Context Before Using It.
Dans A. Dey, B. Kokinov, D. Leake et R. Turner (dir.), *Modeling
and Using Context*. CONTEXT 2005 (p.29-40). Lecture Notes in
Computer Science, vol 3554. Springer.
https://doi.org/10.1007/11508373_3
- Bessire, D. et Fabre P. (2014). New public management et évaluation de la
recherche publique : L'invention du « produisant ». *Finance Con-
trôle Stratégie*, 17(4) [En ligne]
DOI : <https://doi.org/10.4000/fcs.1554>
- Blanchet, P. (2016). Contextualisations didactiques et didactologiques :
questions en débat. *Contextes et didactiques*, 7.
DOI : <https://doi.org/10.4000/ced.542>
- Boulet, E. et Harari-Kermadec, H. (2014). Le rôle de la quantification dans
le processus de marchandisation de l'université. Étude de cas
d'une université anglaise. *Politiques et management public*,
31(4). URL : <http://journals.openedition.org/pmp/7304>

- Chevallard, Y. (1985). *La transposition didactique du savoir savant au savoir enseigné*. La Pensée Sauvage.
- Delcroix, A., Forissier, T. et Anciaux, F. (2013). Vers un cadre d'analyse opérationnel des phénomènes de contextualisation didactique. Dans F. Anciaux, T. Forissier et L.-F. Prudent (dir.), *Contextualisations didactiques : approches théoriques* (p. 141–185). L'Harmattan.
- Dorville, A. (2005). Diglossie et dyslexies. Rapport de synthèse dans le cadre de l'appel à projets pour les Actions Concertées Incitatives Cognitives pour le ministère de la recherche.
- Duranti, A. et Goodwin, C. (1992). *Rethinking context: language as an interactive phenomenon*. Cambridge University Press.
- Ferone, G. (2019). Numérique et apprentissages : prescriptions, conceptions et normes d'usage. *Recherches en éducation*, (35), 63-75. URL : http://www.recherches-en-education.net/IMG/pdf/REE_35.pdf.
- Forissier, T. (2019). *Contextualisation et effets de contextes dans l'apprentissage des Sciences*. [habilitation à diriger les recherches], Université des Antilles et de la Guyane. URL : <https://hal.archives-ouvertes.fr/tel-02277162/document>
- Giordan, A. et De Vecchi, G. (1987). *Les Origines du savoir : des conceptions des apprenants aux concepts scientifiques*. Delachaux et Nieslté.
- Hcéres - Haut conseil de l'évaluation de la recherche et de l'enseignement supérieur (2022). Rapport d'évaluation de l'université des Antilles. URL : <https://www.hceres.fr/fr/rechercher-une-publication/universite-des-antilles-1>
- Hood, C. (1991). A Public Management for all Seasons? *Public administration*, (69), 3-19. DOI : <https://doi.org/10.1111/j.1467-9299.1991.tb00779.x>
- Leite, C. (2005). A territorialização das políticas e das práticas educativas. Dans C. Leite (dir.), *Mudanças curriculares em Portugal. Transição para o século XXI* (p. 15-32). Porto Editora.
- Lemistre, P. (2015). De bac moins trois à bac plus trois : le sens de la professionnalisation des filières, au vu des parcours et de l'insertion. Relief, 50. Récupéré le 21 septembre 2017 de : http://jdl2015.univ-lille1.fr/wp-content/uploads/R50_JDL22_051_Lemistre.pdf.

- Morin, Y. (2016). *L'Université en ses territoires : l'Enseignement Supérieur et la Recherche comme opérateur territorial*. [thèse de doctorat], Université Grenoble Alpes. URL : <https://tel.archives-ouvertes.fr/tel-01433816/document>.
- Mottier Lopez, L. et Dechamboux, L. (2017). D'un référentiel d'évaluation fixe à une co-constitution référentielle dynamique, ce que nous apprend le jugement situé de l'enseignant. *Contextes et didactiques*, (9), 12-29. DOI : <https://doi.org/10.4000/ced.706>
- Mouraz, A., Fernandes, P. et Morgado, J.-C. (2012). Contextualisation curriculaire : des discours aux pratiques. *La Recherche en Éducation*, (7), 31-44.
- M'Saïdié, T. (2013). L'octroi de mer face à la Constitution : réflexions à partir de la décision du Conseil constitutionnel du 26 juillet 2013. *Revue française de droit constitutionnel*, (96), 893-915. DOI : <https://doi.org/10.3917/rfdc.096.0893>
- Odacre, E. (2018). *L'enseignement des sciences économiques et gestion en BTS en Guadeloupe : entre prescrits nationaux et contexte socio-économique régional*. [thèse de doctorat], Université des Antilles. URL : <http://theses.fr/2018ANTI0252>
- Odacre, E. (20201). Le poids du contexte socioéconomique local dans l'abandon des bacheliers professionnels en brevet de technicien supérieur en Guadeloupe. *Éducation et socialisation - Les cahiers du CERFEE*, 58. DOI : <https://doi.org/10.4000/edso.13203>
- Odacre, E., Issaeva, I. et Delcroix A. (2021a). Les représentations des étudiants de la formation BTS en Guadeloupe à l'épreuve des stages professionnels. *Revue internationale de pédagogie de l'enseignement supérieur*, 37(2). URL : <http://journals.openedition.org/ripes/3219>.
- Odacre, E., Issaeva, I. et Delcroix, A. (2021b). Pratiques de contextualisation en brevet de technicien supérieur en Guadeloupe : quelles articulations entre contexte local et programmes nationaux ? *Revue des sciences de l'éducation*, 47(1), 31-52. DOI : <https://doi.org/10.7202/1078161ar>
- Pereira, I. (2015). L'expérimentation en immersion. *Revue des sciences sociales*, (54), 148-155. DOI : <https://doi.org/10.4000/revss.2418>

Robo, E. (2021). *Etude comparée des interactions à visée d'apprentissage lors de séances de géométrie au cycle 3 de l'école primaire : le cas de la Polynésie française et de la Guyane française*. [thèse de doctorat], Université des Antilles. URL : <https://www.theses.fr/2021ANTI0615>.

Verret, M. (1975). *Le temps des études*. Honoré Champion.

Conclusion : 'De quels contextes parle-t-on pour problématiser l'évaluation ?'

Quelques apports du 33ème colloque de l'ADMEE-Europe pour conclure

Lucie Mottier Lopez

I. Introduction

« Pluralité de contextes, pluralité d'évaluations en éducation : quelles interactions et quels enjeux ? », telle était la thématique du 33ème colloque de l'ADMEE-Europe qui a eu lieu à l'Université des Antilles en Guadeloupe. Dans la continuité des travaux scientifiques du laboratoire CRREF « Centre de Recherches et de Ressources en Éducation et Formation », cette thématique invitait à interroger les problématiques d'évaluation en éducation au regard des contextes et de leur pluralité. Comme énoncé par l'introduction de cet ouvrage, comment penser les transitions entre des contextes pluriels ? Comment former à des cultures plurielles ? Comment conjuguer des sources multiples ?

La notion de contexte apparaît souvent comme un allant de soi, avec l'illusion d'un mot partagé. C'est un lieu faussement commun, comme l'ont d'emblée souligné Malet et Prudent dans leurs conférences respectives à l'occasion du colloque.³⁴ Sans prétention à l'exhaustivité, ce

³⁴ A chaque fois que je citerai dans ce texte un auteur sans indiquer une année, c'est parce qu'il s'agit d'une communication ou conférence effectuée au 33ème colloque de l'ADMEE-Europe, dont la référence se trouve au lien suivant : <https://ad->

texte conclusif présente quelques éléments de synthèse à la suite d'une lecture attentive du [livret des résumés](#) des communications et l'écoute des présentations que j'ai eu l'occasion de suivre.³⁵ Une première catégorisation est proposée dans la deuxième partie de ce texte, visant à identifier quelques dimensions que les auteurs et les autrices ont eu tendance à privilégier quand il s'agit de problématiques évaluatives interrogées à la lumière des contextes. A partir de ce premier tour d'horizon, non exhaustif, je poursuis en explicitant une position qui est certainement aujourd'hui admise : aucune activité humaine instituée (dont l'activité évaluative en éducation) n'est marquée que de son seul contexte immédiat de développement. Les contextes sont nécessairement pluriels, de façon plus ou moins reconnue et explicitement problématisée, de façon plus ou moins hiérarchisée, emboîtée et/ou en interaction. Cela me donne l'occasion d'explicitier, dans les parties suivantes du texte, un cadre épistémologique et théorique, parmi d'autres possibles, pour comprendre et interpréter les évaluations dites situées, conçues comme étant foncièrement liées aux contextes pluriels qui les constituent et réciproquement. Quelques questionnements terminent ce texte, ouverts à de nouvelles investigations et développements.

II. De quel(s) contexte(s) parle-t-on ?

Dans le [livret des résumés](#), les auteurs ont utilisé une large palette de termes pour approcher la thématique du colloque et parler de contexte(s) : culture, situation, environnement, lieu, milieu, terrain, espace, système, sphère, champ, pratique, climat, etc. Le « contexte » se caractérise notamment à partir d'*attributs* géographiques, historiques, politiques, culturels, sociaux, linguistiques, éducatifs, professionnalisants, etc. Il l'est également en termes de *niveaux* : macro-méso-micro ; local-global ; interne-externe ; collectif-individuel ; social-psychologique ; préexistant-émergent ; donné-construit, notamment. Y compris pour convoquer des dimensions complexes, tels les cinq niveaux socio-écologiques développés par Nizet dans cet ouvrage, et les sous-systèmes et leurs facteurs d'influence. Les *relations* entre contextes sont diversement qualifiées :

[meeqp2022.sciencesconf.org/data/pages/Livret des résumés V1 Site du Colloque 1.pdf](https://meeqp2022.sciencesconf.org/data/pages/Livret_des_resumes_V1_Site_du_Colloque_1.pdf)

³⁵ En tout, 95 communications, 11 symposia réunissant 63 communications, 7 communications affichées, 2 tables rondes, 8 conférences plénières ont composé le programme du colloque.

emboitements hiérarchiques *versus* constitutions dynamiques pour penser les interactions entre les contextes, leurs imbrications, leurs interdépendances, leurs co-transformations ou co-régulations. Quelques travaux illustratifs présentés au colloque sont cités ci-dessous.

D'une façon relativement convenue quand il s'agit de présenter une étude empirique, les auteurs convoquent la notion de contexte pour désigner les conditions spécifiques dans lesquelles l'étude a été menée (lieu(x), participants, temporalités). Dans les communications, la notion de contexte permet aussi de spécifier l'ordre d'enseignement selon des *structures institutionnelles* : contexte scolaire, contexte d'enseignement spécialisé, contexte d'enseignement secondaire, contexte d'enseignement supérieur, contexte universitaire, contexte de formation (formelle, informelle), contexte de stages, contexte de VAE, etc. D'autres auteurs parlent, quant à eux, de contexte pour décrire une situation particulière, délimitée dans le temps, marquée par des *circonstances* qui le caractérisent : contexte sanitaire, contexte de crise, contexte du Covid, contexte précaire, contexte incertain, etc. D'autres associent le contexte à une *fonction* spécifique : contexte d'inclusion, contexte de professionnalisation, contexte d'accompagnement, contexte d'alternance, etc. Ou encore par rapport à l'objet étudié : contexte d'implémentation de l'innovation (étudiée), contexte de mise en œuvre de la réforme (étudiée), etc.

Au regard des caractéristiques (étant seulement énoncées³⁶ *versus* faisant l'objet d'une conceptualisation) des contextes concernés, les questions ont été multiples, notamment dans des approches à visée descriptive :

- Quelles pratiques évaluatives ? Quelles normes ? Quelles cultures ? Quelles représentations ? Quelles valeurs ? Quelles croyances ? ...
- Quels dispositifs d'évaluation ? Quels outils ? ...
- Quels acteurs ? Quel(s) rôle(s) ? Quelles activités ? Quels ressentis ? Quelles émotions ? ...
- Quelles actions ? Quels gestes ? Quels langages ? ...

³⁶ Un ensemble de communications se sont arrêtées à ces éléments descriptifs sans problématiques particulières à leur propos.

- Quels repères ? Quelle référentialisation ? Quelle contextualisation ? ...

Dans certains travaux, ces caractéristiques (nommées sous différents vocables, dans des approches méthodologiques variées (quantitatives et/ou qualitatives) sont *mises en relation* à l'objet « évaluation » en termes d'influences, d'impacts, d'effets des contextes sur..., d'obstacles, de limites, de freins, de pressions des contextes sur..., de ressources, de facteurs facilitants, de conditions contextuelles facilitantes à... Par exemple, la conférence d'El Hage s'est intéressée à *l'influence* du contexte social sur le développement des compétences des élèves, des idéologies et des croyances en évaluation ; celle de Cerda et Heritage aux *influences* des normes et des pratiques contextuelles sur l'autoévaluation et la co-régulation dans l'évaluation formative. Le projet du symposium de Lafont était, quant à lui, d'identifier *l'impact* des différents contextes pluriels aux niveaux macro, méso et micro sur l'évaluation, ainsi que les relations et interactions entre les contextes et les différentes formes d'évaluation. Les *effets des contexte* produits dans les situations d'évaluation étaient aussi au cœur du symposium coordonné par Poggi et Verscheure.

En lien avec les *enjeux didactiques* et les travaux du laboratoire CRREF, des problématiques d'effets de contexte, de contextualisation didactique ou encore de didactique contextualisée ont fait l'objet de plusieurs communications. Celle de Bonbonne par exemple visait à étudier les tensions entre les usages sociaux et les usages didactiques, dans une approche dite socio-didactique, notamment pour appréhender différents effets de contexte « en actes » générés par les pratiques enseignantes, susceptibles d'être productrices d'inégalités. Delcroix s'est interrogé sur une contextualisation didactique de l'évaluation (voir également sa contribution dans cet ouvrage) et Psyche sur les enjeux, en termes d'outil d'aide, à une scénarisation pédagogique contextualisée.

Inviter les chercheurs à interroger la pluralité des contextes en évaluation a conduit certains auteurs à effectuer des *comparaisons* entre contextes, mettant en tension les contextes micro, méso, macro ou globaux et locaux, par exemple. C'est le cas de la contribution de Delcroix qui a notamment interrogé la tension entre des contextes géographiques et sociétaux très variés et un système éducatif centralisé, invitant à problématiser les processus de normalisation et d'adaptation. Ainsi, par exemple, dans

certaines communications, le terme d'*évaluation contextualisée* permet d'établir une distinction entre une évaluation standardisée (commune, externe) qui prescrit une procédure associée à un ou plusieurs outils communs à appliquer, à mettre en œuvre dans des contextes spécifiés (qui portent leurs propres pratiques, normes, langages, significations, valeurs, etc.) engageant un processus de contextualisation.

La *comparaison* entre contextes a également conduit certains auteurs à observer des variations entre eux, des transformations, des adaptations. Le projet de Malet dans sa conférence a consisté, par exemple, à comparer des variations sémiotiques dans la construction de l'action éducative en contexte. Quant à Morales-Perlaza et Fonseca Grandón, ils ont procédé à une analyse comparative de programmes de formation à l'enseignement entre différents pays (Québec, Finlande, Chili) pour interroger les potentialités des dispositifs, des pratiques pour mieux former.

Un grand nombre de communications ont thématiqué la multiplicité des contextes, en termes de diversité, d'hétérogénéité, de pluri-contextualité pour penser leur prise en compte dans l'évaluation et les négociations de significations et de référentiels en jeu par exemple. Ainsi, le symposium de Bieri et al. s'est notamment penché sur la prise en compte de la diversité des contextes (conception, accompagnement, exploitation des résultats de l'évaluation, implication des divers acteurs de la formation, entre autres) pour un pilotage optimal de la formation au sein d'un hôpital. Roblez s'est interrogé sur la médiation de l'expert-évaluateur entre les différentes significations évaluatives contextualisées. Chez Salvisberg, la multiplicité des contextes a été appréhendée du point de vue de la multiréférentialisation qui traverse l'ensemble des activités d'un enseignant-évaluateur. Dans ces communications, la pluralité des contextes est reconnue comme pleinement constitutive des activités de pilotage, des activités des experts, des activités des enseignants.

Cette pluralité des contextes a également été thématisée en termes de changements, de *transitions*, de continuités et de discontinuités, de consonnances et de dissonances. Ainsi, la conférence de Nizet a explicitement posé les relations entre contextes et cultures d'évaluation notamment en termes de ruptures. Dans le cadre de la création de centres d'appui à l'évaluation dans des universités marocaines, Rigo et al. ont étudié la contextualisation des dispositifs et des outils pour interroger le

transfert de compétences entre Nord-Sud. Quant à Issaieva, ce sont des enjeux de *développements professionnels* qui ont été interrogés au regard des contextes guadeloupéens marqués par des changements spécifiques d'ordre politique, socioéconomique et culturel. Dans le symposium coordonné par Younès et Berger, ce sont les *processus de référentialisation*, caractéristiques de l'évaluation, qui ont servi à se demander dans quelle mesure ils étaient susceptibles de tirer parti des consonances et des dissonances entre contextes pour tenter de mieux accorder les groupes, les organisations et les institutions. Toujours dans des contextes pluriels en interaction, la communication de De Paor a introduit les notions de *boundary crossing* et de *boundary object* (un portfolio) pour penser la complexité des évaluations invitant à négocier les transitions dans des validations des acquis d'expérience (VAE).

Pour clore cette brève évocation, c'est aussi un niveau des recherches elles-mêmes que la thématique du colloque a été traitée. Challah et al. ont posé la question de savoir comment interroger la complexité des contextes pluriels convoqués dans une recherche-action. Quant à Marcel, il a invité dans sa conférence à concevoir un tiers-espace de l'évaluation, dans une perspective dite écologique pour prendre en compte l'hétérogénéité des contextes qui caractérisent les recherches intervention.

Je terminerai en reformulant certains questionnements qui ont été énoncés par un ensemble d'acteurs locaux, de la Guadeloupe donc, quand il leur a été demandé quels étaient leurs points de vue sur l'évaluation en éducation. Ces témoignages, je dois le dire, m'ont particulièrement touchée par leur sincérité face aux défis rencontrés et relevés, exprimés dans le cadre de la table ronde organisée par Elisabeth Issaieva et discutée par la regrettée Muriel Joseph-Théodore :

- Comment construire une culture d'évaluation commune ? Comment assurer une cohérence nécessaire entre les contextes tout en reconnaissant les différences légitimes (d'engagements, de ressources, de compétences, de motivation, etc.) ?
- Que faire face au constat de morcellement de l'évaluation, de ruptures situationnelles entre les contextes, de pertes de sens et de confiance ? Au nom de quoi parle-t-on de l'évaluation, au nom de quelles valeurs, visés, intentions, buts, projets ? Au nom de quel sens ?

- Comment construire *ensemble* des diagnostics réflexifs, des autoévaluations constructives, de la co-régulation, du jugement professionnel partagé en puisant dans les ressources, parfois minces, des contextes pluriels qui se rencontrent, se confrontent, ne se comprennent pas ?

III. Perspective située de l'activité sociale et humaine

A partir de ce bref tour d'horizon, je vous livre ci-dessous certaines des interrogations qui m'ont habitée en explorant la richesse des travaux des collègues. Ces interrogations trouvent leurs sources dans mes premières « amours scientifiques » et qui, depuis, traversent constamment mes réflexions.³⁷ A l'époque de ma thèse, après avoir été plusieurs années enseignante et formatrice d'enseignants dans le canton de Vaud en Suisse Romande, je m'interrogeais sur les significations locales, construites entre un enseignant et ses élèves, propres à la classe, à leur histoire, à leurs relations, à leurs valeurs. Ceci m'a amené à conceptualiser les normes sociales négociées dans chaque microculture de classe (au sens de Cobb et al., 1997, 2001) en tant que « référentiel » contribuant à « marquer » foncièrement les apprentissages des élèves et leurs processus de régulation (Mottier Lopez, 2008, 2016). Mon expérience d'enseignante me soufflait que beaucoup de choses s'y jouaient, entre ce qui est annoncé comme devant être commun (le plan d'études et ses objectifs d'apprentissage, les manuels utilisés³⁸, les prescriptions, etc.) et ce qui restait indéniablement singulier, émergent, coconstruit dans la classe – y compris dans la renégociation de la compréhension du « commun prescrit ».

Les travaux situationnistes de Lave (1988), entre autres, ont représenté pour moi une inspiration fondatrice pour penser une relation dite « dialectique » (de co-constitution³⁹) entre contextes et activités

³⁷ J'y reconnais une certaine obsessionnalité de ma part.

³⁸ En Suisse romande, nous parlons de « moyens d'enseignement » ; ils sont communs aux enseignants.

³⁹ Une relation de constitution et de structuration réciproques entre différents plans, dans le prolongement de la thèse que « learning, thinking, and knowing are relations among people engaged in activity in, with, and arising from the socially and culturally structured world (souligné par l'autrice). This world is itself socially

(d'enseignement, d'apprentissage, d'évaluation), entre contextes et apprentissages, entre contextes et régulations. L'hypothèse situationniste est que la connaissance acquise (ses unités de sens et sa finalité en contextes scolaire et extra-scolaire) ne peut être séparée de l'activité qui a contribué à son développement. Non neutre, cette activité amène l'apprenant à « apprendre » les conditions contextuelles de la connaissance acquise, en tant qu'apprentissage à part entière (Brown et al., 1989). Cette hypothèse a débouché sur des débats fameux, dont celui entre Greeno (1997 ; Greeno et al., 1998), tenant de la perspective située, et Anderson et al. (1996, 1997, 2000), détracteurs. De mon point de vue, une des critiques majeures qui a conduit à des malentendus persistants concernant les travaux situationnistes anglosaxons, a été de considérer la relation de l'activité à une situation strictement concrète et matérielle (Anderson et al., 1996), faisant fi des dimensions institutionnelles, culturelles, historiques préexistantes (Filliettaz et Schubauer Leoni, 2008).

Pourtant, comme analysé par Kirshner et Whitson (1998) par exemple, le projet de Lave est bien de penser les environnements locaux avec, en même temps, ce que l'autrice appelle « les arènes sociales et politiques plus larges ». Non directement négociables par l'individu, ces arènes fournissent un cadre institutionnel (incluant donc des enjeux politiques, économiques, etc.) à l'activité située, avec la conception cependant que ce cadre préexistant n'est pas pleinement déterministe sur l'activité (au sens du déterminisme social de Bourdieu). Pour soutenir cette thèse, Lave convoque la notion de « relation dialectique » entre les personnes qui agissent et les ressources structurantes des environnements expérimentés. Ces ressources structurantes peuvent être des outils, des artefacts, des interactions sociales, mais également des attentes perçues, des normes, des valeurs, etc. Autrement dit, toute affordance, y compris évaluative, qui contraint et tout à la fois structure les activités, leurs processus, leurs produits.

Les travaux empiriques de Grossen (2000), par exemple, s'attachent à montrer la part constitutive des *institutions* dans la construction située de l'interaction sociale en éducation :

constituted » (Lave, 1991, p. 67).

Institutions play an integral part in thinking, learning and teaching, and solving a problem or learning a new body of knowledge are only parts of a broader activity: constructing the interaction itself, and more specifically constructing the institutional frame in which some roles are enacted, others rejected, identities are negotiated, questions, bodies of knowledge or answers, are considered as legitimate or not. (p. 33)

De ce point de vue, les interactions sont symboliquement médiatisées par les institutions sociales, dont l'institution scolaire qui a la légitimité de déterminer, par exemple, quelles connaissances doivent être enseignées, apprises, évaluées, en fonction de quel curriculum, avec quelle méthode, pour quels objectifs, etc. Grossen parle de « *invisible audience* » apportant plus ou moins directement ses règles, ses normes, ses valeurs dans la situation. Au-delà de ces quelques développements, l'enjeu est toujours actuel de parvenir à mieux appréhender et mieux comprendre comment opère la « relation dialectique » entre l'activité humaine et ses contextes pluriels (immédiats et plus larges), incluant leurs dimensions multiples, individuelles et sociales, concrètes / matérielles et symboliques, singulières et culturelles, immédiates et historiques, pour ce qui concerne des problématiques d'évaluation en éducation.

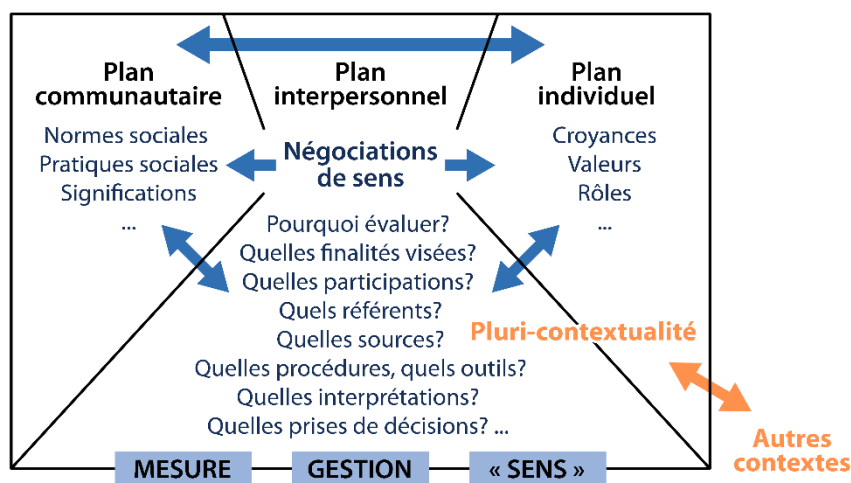
IV. L'évaluation située, le choix d'un ancrage épistémologique

A la suite de ces travaux, le projet de conceptualiser une « évaluation située » renvoie ainsi à un ancrage épistémologique et théorique défini : toute activité évaluative (en tant qu'activité humaine et sociale) renvoie à des pratiques socio-institutionnelles, culturelles, communautaires et individuelles, avec l'hypothèse d'une indissociabilité entre activité et contextes. De ce point de vue, l'évaluation située n'est pas « un modèle » d'évaluation supplémentaire comme proposé par Vial (2012) et par d'autres auteurs qui l'ont suivi, visant notamment à différencier les caractéristiques d'une « évaluation-mesure », d'une « évaluation-gestion » ou d'une « évaluation-négociation de sens ». De fait, le choix de l'ancrage situationniste invite à penser chaque pratique d'évaluation (et les modèles de mesure, de gestion, de sens qui les fondent) comme étant foncièrement située, marquée par des normes, des valeurs, des significations, des langages, des outils, etc., autrement dit, une culture (ou plusieurs cultures en tension) qui la « marque(nt) ». Réciproquement, cette pratique évaluative contribue elle-même à transmettre ou à soutenir la négociation,

la construction de normes, de valeurs, de significations, de langages partagés, etc.

La figure 1 ci-dessous explicite un cadre possible d'analyse et d'interprétation, largement adapté de Cobb et al. (1997, 2001)⁴⁰, pour étudier l'évaluation située (par ex., Mottier Lopez, 2008, 2016). J'y ajoute ici la référence aux modèles d'évaluation « mesure », « gestion », « sens » (Vial, 2012) et quelques exemples de questionnements qui « aident » à différencier les pratiques associées à chacun des modèles.

Figure 1 : Cadre d'analyse et d'interprétation des pratiques d'évaluation située



Le plan communautaire désigne les pratiques et les normes sociales « vues comme reconnues et partagées (*taken-as-shared*)⁴¹ » entre les membres d'un même groupe social étant amené à se coordonner régulièrement pour

⁴⁰ Ces auteurs analysent les pratiques d'enseignement et d'apprentissage du point de vue des microcultures de classe qui les « situent », mais sans introduire dans leur modélisation le plan interpersonnel ni l'objet « évaluation » comme je le fais dans mes travaux.

⁴¹ Cette expression est une traduction du concept « *taken-as-shared* » utilisé par les auteurs, afin de souligner que la compréhension des significations rattachées aux pratiques et normes communautaires gardent toujours une part de singularité sur le plan interprétatif individuel (Cobb et al., 1997, 2001).

mener des activités conjointes. Elles se négocient⁴² de façon continue entre les personnes, souvent de façon implicite. Quant au plan individuel, il désigne les interprétations et raisonnements des acteurs quand ils participent aux pratiques sociales, ainsi que leurs valeurs et croyances individuelles à propos de leur rôle, du rôle des pairs, d'autrui, de leur rapport aux objets (enseignés, appris, évalués). Le plan interpersonnel cible sur les processus de négociation de sens et de construction interactive des significations vues comme partagées entre les membres de la communauté sociale concernée. L'hypothèse situationniste amène à postuler une relation de constitution et de structuration réciproques entre les plans, observée par des phénomènes de régulation dite réflexive ou de co-régulation (Mottier Lopez, 2008).

La thématique du 33^{ème} colloque de l'ADMEE-Europe, alimentée par ses communications inspirantes, encourage à problématiser la façon dont les contextes multiples, sans être directement présents (par ex., Grossen, 2000, citée plus haut) ou sans être explicitement vus comme tels (par exemple, l'usage de référentiels externes qui véhiculent leurs propres normes) « pluri-situent » l'évaluation (Mottier Lopez, 2019). De ce point de vue, les situations évaluatives se complexifient nécessairement. L'évaluation ne peut plus être réduite à des procédés et à des outils techniques qu'il suffirait de savoir appliquer, maîtriser, reproduire. L'essence même de l'évaluation inclut des dilemmes professionnels, bien documentés quand il s'agit par exemple de concilier les fonctions formative (interne à la classe), certificative (pour attester les acquis aux partenaires sociaux) et pronostique (qui projette vers de nouvelles étapes de formation et de vie), ou encore quand il s'agit de concilier des normes institutionnelles exogènes avec des normes professionnelles incarnées (par ex., Mercier-Brunel, dans cet ouvrage). L'évaluation est source de demandes contradictoires, d'exigences souvent difficiles à réunir, de situations ambiguës, indéterminées pour lesquelles il n'y a pas de solutions toutes faites. Autrement dit, dans les termes de Nizet (dans cet ouvrage), l'évaluation est constamment aux prises avec les « niveaux socio-écossystémiques » et leur interdépendance, faites de continuités et de ruptures, entre systèmes et sous-systèmes.

De façon plus pragmatique, les résultats d'une évaluation sont souvent destinés à d'autres personnes que les évalués et évaluateurs concernés, personnes qui ont leurs propres « contextes », systèmes d'activité, normes professionnelles et/ou normes sociales et culturelles, personnes qui ont

⁴² Au sens interactionniste du terme.

leurs propres valeurs et croyances individuelles. Dans le cas de certaines évaluations externes, comme dans le canton de Genève en Suisse par exemple, elles sont décidées par l'institution scolaire (productrice de prescriptions), conçues par des acteurs qui ont un statut d'expert (de la noosphère), puis elles sont utilisées par les enseignants et les élèves dans leurs classes. Les résultats de ces évaluations externes sont ensuite communiqués à l'institution scolaire (à des fins de pilotage du système éducatif) mais également aux parents des élèves dans le cadre d'un « livret scolaire » dans une visée de certification (combinée à d'autres résultats produits par des évaluations internes). Cet exemple montre que l'évaluation est un objet qui *circule* entre des contextes et des cultures pluriels et qui, potentiellement, s'épaissit de sens (pas toujours cohérent) à leur contact. Il paraît alors intéressant de se demander dans quelle mesure elle, et les documents à visée instituante (Girardet, 2023 ; Mottier Lopez, 2023) et autres symboles abstraits qui l'accompagnent et la médiatisent, sont susceptibles d'être des *objets-frontières* (Trompette et Vinck, 2009) « dont la structure est suffisamment commune à plusieurs mondes sociaux pour (qu'ils) assure(nt) un minimum d'identité au niveau de l'intersection tout en étant suffisamment souple pour s'adapter aux besoins et contraintes spécifiques de chacun de ces mondes » (p. 8).

Concluons ce texte avec quelques questions d'ouverture : quels sont les espaces, les outils, les langages de médiation entre les mondes par rapport aux enjeux évaluatifs et leur spécificité, et quels sont leurs affordances ou conditions capacitanes ? Quels sont les « passeurs de frontière » (*boundary spanners*) ou « courtiers de connaissance » (*broker*, Nizet et al., 2019) dont le rôle est de créer des liens entre les groupes sociaux et culturels à propos de questions évaluatives sensibles, incomprises, sources de stress et de pression négative ? Quelles connaissances et compétences professionnelles, y compris de nature psycho-sociale, cela leur demandent-il ? Quels phénomènes propres à l'évaluation observe-t-on dans sa pluricontextualité, en termes par exemple de normalisation, de multiréférentialisation, de co-régulation des contextes et à la fois des acteurs individuels et collectifs impliqués ? etc. Les questions sont nombreuses et invitent, je l'espère, à poursuivre collectivement la problématisation de la pluralité des évaluations et de leurs contextes à des fins éducatives et de formation.

V. Bibliographie

- Anderson, J. R., Reder, L. A. et Simon, H. A. (1996). Situated learning and education. *Educational Researcher*, 25(4), 5-11.
- Anderson, J. R., Reder, L. A. et Simon, H. A. (1997). Situative versus cognitive perspectives: Form versus substance. *Educational Researcher*, 26(1), 18-21.
- Anderson, J. R., Greeno, J. G., Reder, L. A. et Simon, H. A. (2000). Perspectives on learning, thinking and activity. *Educational Researcher*, 29(4), 11-13.
- Brown, J. S., Collins, A. et Duguid, P. (1989). Situated cognition and the culture of learning. *Educational Researcher*, 18(1), 32-42.
- Cobb, P., Stephan, M., McClain, K. et Gravemeijer, K. (2001). Participating in classroom mathematical practices. *Journal of the Learning Sciences*, 10, 113-164.
- Cobb, P., Gravemeijer, K., Yackel, E., McClain, K. et Whitenack, J. (1997). Mathematizing and symbolizing: The emergence of chains of signification in one first-grade classroom. Dans D. Kirshner et J. A. Whitson (dir.), *Situated cognition, social, semiotic, and psychological perspectives* (p. 151-233). Lawrence Erlbaum.
- Filliettaz, L. et Schubauer-Leoni, M. L. (2008). Les processus interactionnels dans leurs dimensions interpersonnelles, socio-historiques et sémiotiques. Dans L. Filliettaz et M.-L. Schubauer-Leoni (dir.), *Processus interactionnels et situations éducatives* (p. 7-39). De Boeck.
- Girardet, C. (2023, prépublication). Co-élaborer des documents à visée instituyente : quelles démarches, quels défis ? Synthèse des textes cantonaux et intercantonal du numéro. *La Revue LEeE*, 8. <https://revue.leee.online/index.php/info/article/view/203>
- Greeno, J. G. (1997). On claims that answer the wrong questions. *Educational Researcher*, 26(1), 5-17.
- Greeno, J. G. et Moore, J. L. (1993). Situativity and symbols: Response to Vera and Simon. *Cognitive Science*, (17), 49-61.
- Greeno, J. G. et the Middle School Mathematics through Applications Project Group. (1998). The situativity of knowing, learning, and research. *American Psychologists*, (53), 5-26.

- Grossen, M. (2000). Institutional framing: In thinking, learning, and teaching. Dans H. Cowie et G. van der Aalsvoort (dir.), *Social interaction in learning and instruction, the meaning of discourse for the construction of knowledge* (p. 21-34). Pergamon Elsevier Science Ltd.
- Kirshner, D. et Whitson, J. A. (1998). Obstacles to Understanding Cognition As Situated. *Educational Researcher*, 27(8), 22-28. <https://doi.org/10.3102/0013189X027008022>
- Lave, J. (1988). *Cognition in practice: Mind, mathematics and culture in everyday life*. Cambridge University Press.
- Lave, J. (1991). Situated learning in communities of practice. Dans L. B. Resnick, J. M. Levine et S. D. Tealey (dir.), *Perspectives on socially shared cognition* (p. 63-82). American Psychological Association.
- Mottier Lopez, L. (2008). *Apprentissage situé : la microculture de classe en mathématiques*. Peter Lang.
- Mottier Lopez, L. (2016). La microculture de classe : un cadre d'analyse et d'interprétation de la régulation située des apprentissages des élèves. Dans B. Noël et S.C. Cartier (dir.), *De la métacognition à l'apprentissage autorégulé* (p. 67-78). De Boeck.
- Mottier Lopez, L. (2019). Penser l'évaluation des apprentissages comme une activité polysituée. *Education permanente*, 220-221, 195-202.
- Mottier Lopez, L. (2023, prépublication). Construire des cultures en faveur d'évaluations positives pour l'avenir : enjeux de la collaboration autour de la production de textes à visée instituante. *La Revue LEEe*, 8. <https://revue.leeonline/index.php/info/article/view/190>
- Nizet, I., Monod-Ansaldi, R., Aldon, G., Prieur, M. et Criquet, A. (2019). L'analyse de valuations dans une démarche collaborative de recherche. *La Revue LEEe*, 1. <https://doi.org/10.48325/rlee.001.04>
- Trompette, P. et Vinck, D. (2009). Retour sur la notion d'objet-frontière. *Revue d'anthropologie des connaissances*, 3(1), 5-27. <https://doi.org/10.3917/rac.006.0005>
- Vial, M. (2012). *Se repérer dans les modèles de l'évaluation. Méthodes, dispositifs, outils*. De Boeck.

Les contributeurs

Marine André est chargée de recherche au service Analyse et Accompagnement du Développement Professionnel des Enseignants de l'Université de Liège. Ses recherches en sciences de l'éducation se concentrent principalement sur le développement professionnel des enseignants et son accompagnement dans le domaine de la littératie, notamment à travers la conception d'outils de formation.

Ashley Byrne est étudiante au doctorat et coordonne les activités de développement durable à l'Institut en Environnement, Développement et Société (EDS) à l'Université Laval.

Janet Cerda termine actuellement son doctorat en développement humain et en psychologie de l'Université de Californie à Los Angeles. Ses recherches se centrent sur les pratiques d'évaluation en classe et l'enseignement et l'apprentissage en contextes multilingues et multiculturels. Ses travaux s'inspirent de ses expériences en tant qu'enseignante et des expériences de ses élèves, de leurs familles et de sa famille en matière de soutien de la langue parlée à la maison dans différents contextes.

Lionel Dechamboux est chargé d'enseignement à l'Université de Genève (Section des sciences de l'éducation) dans l'équipe EReD (évaluer, réguler et différencier pour apprendre) sous la direction de la professeure Lucie Mottier Lopez qui a également co-dirigé sa thèse de doctorat avec le professeur Germain Poizat (Université de Genève). Ses intérêts scientifiques actuels sont liés à l'évaluation formative et certificative, la référentialisation, l'activité des évaluateurs en situation.

Charlotte Dejaegher est chargée de cours à l'Université de Namur à la Faculté des Sciences de l'Éducation et de la Formation. Son parcours doctoral, à la croisée de l'ingénierie en didactique du français et de l'ingénierie de formation, l'a conduite à explorer les questions de l'analyse et l'accompagnement du développement professionnel des enseignants, principalement dans le domaine de la littératie. Ses champs d'intérêt professionnels concernent plus largement l'ingénierie de formation, l'analyse de l'activité, la formation et l'accompagnement des (futurs) enseignants.

Antoine Delcroix est professeur des universités en sciences de l'éducation et de la formation à l'université des Antilles. Il s'intéresse aux phénomènes de contextualisation dans l'enseignement et la formation et a contribué à la création du CRREF qu'il a dirigé pendant dix ans. Il est actuellement directeur de l'Institut national du professorat et de l'éducation de l'académie de la Guadeloupe et responsable d'une revue et d'une collection *aux Presses universitaires des Antilles*.

Fadi El Hage est professeur et chercheur à la Faculté des Sciences de l'Éducation de l'Université Saint-Joseph au Liban. Il est Délégué du Recteur pour le développement professionnel et la formation continue, Directeur de la Chaire en Éducation pour le Développement Durable. Ses recherches portent sur les méthodes d'enseignement actives, gestion de classe et éducation à la santé dans les écoles et en enseignement supérieur.

Céline Girardet est maîtresse d'enseignement et de recherche dans l'équipe EReD (évaluer, réguler et différencier pour apprendre dans les systèmes d'enseignement et de formation) à l'Université de Genève. Ses intérêts portent notamment sur les dispositifs d'évaluation continue pour apprendre dans divers contextes éducatifs, l'évaluation en tant que compétence transversale dans la perspective d'un apprentissage tout au long de la vie, et les enjeux des recherches participatives et collaboratives.

Elisabeth Issaieva est actuellement enseignante-chercheuse à l'INSEL, CY Cergy Paris Université. Elle a auparavant enseigné, pendant plusieurs années, à l'Université des Antilles et est régulièrement intervenue dans la formation des enseignant.es du premier et second degré ainsi que dans la formation des enseignant.es spécialise.es, des formateur.trices d'adultes et des conseiller.es d'éducation. Elle a aussi été collaboratrice scientifique au SRED du canton de Genève et à l'Université de Genève en Suisse. Ses travaux menés dans différents contextes visent à contribuer à une meilleure compréhension des mécanismes des processus d'enseignement, d'évaluation et d'apprentissage en vue de combattre l'échec scolaire et de répondre à des enjeux inclusifs.

Lucie Héon est professeure titulaire au Département des fondements et pratiques en éducation de l'Université Laval et directrice des programmes d'études supérieures en administration et politiques de l'éducation. Ses travaux portent sur l'étude des systèmes et des organisations d'enseignement supérieur, la planification et la gestion de l'éducation, et l'analyse des compétences et des pratiques des directeurs d'établissement.

Catherine Larouche est professeure agrégée au département d'anthropologie de la faculté des sciences sociales de l'Université Laval. Ses travaux portent sur les acteurs philanthropiques et non-étatiques du développement, l'organisation des services sociaux, les enjeux liés à la privatisation ainsi que les rapports entre identité et redistribution.

Camille Larouche a obtenu son M.A. en sociologie (Université Laval) en 2023. Elle travaille comme professionnelle de recherche depuis, ayant pour principal intérêt la co-construction de connaissances.

Olivier Leyh est enseignant et chargé de recherche au service Analyse et Accompagnement du Développement Professionnel des Enseignants de l'Université de Liège. Ses activités de recherche portent sur les dispositifs qui favorisent l'entrée dans le monde de l'écrit, l'enseignement de la compréhension en lecture et l'exploitation de la littérature de jeunesse dans les classes. Dans ce cadre, il travaille à la conception d'outils didactiques et de dispositifs de formation destinés aux enseignants et s'intéresse tout particulièrement aux processus qui permettent l'évaluation de leur qualité.

Jean-François Marcel est professeur émérite à l'Université Toulouse Jean-Jaurès où, pendant deux mandats, il a dirigé l'UMR EFTS. Préalablement, il a été Directeur de la Recherche à l'ENSFEA puis Directeur de l'École Doctorale CLESCO. Par ailleurs, il a dirigé une revue internationale et plusieurs collections scientifiques. Ses recherches portent sur l'accompagnement du changement et l'émancipation. À cet effet, il mobilise, de manière privilégiée, la démarche de recherche-intervention.

Natalia Constenla Martinez est chargée d'enseignement, formatrice Universitaire IUFE (CFO-MESP) & Collaboratrice scientifique (EReD) en Sciences de l'éducation à l'Université de Genève.

Yann Mercier Brunel est professeur des universités au Laboratoire Apprentissage Didactique Évaluation Formation, Aix-Marseille Université. Il s'intéresse à la façon dont les feedbacks des enseignants sont susceptibles de soutenir ou de faire obstacle aux processus d'autorégulation de leur apprentissage par les élèves. Il cherche à mettre en évidence que ces feedbacks reposent sur des normes sous-tendant leurs pratiques évaluatives, développées en contexte social et scolaire et relayées par l'ensemble des acteurs de l'éducation : enseignants et élèves, mais également inspecteurs, chefs d'établissement et parents.

Lucie Mottier Lopez est professeure ordinaire à la Faculté de Psychologie et des Sciences de l'Éducation de l'Université de Genève. Elle dirige le groupe de recherche « Evaluer, réguler et différencier pour apprendre dans les systèmes d'enseignement et de formation » (EReD). Ses études actuelles portent sur la voix des élèves dans des évaluations à visée émancipatrice et sur les processus de co-constitution référentielle et de régulation dans une perspective d'évaluation située. Elle a présidé l'ADMEE-Europe de 2009 à 2013.

Isabelle Nizet est professeure titulaire retraitée, associée au département de pédagogie de la faculté d'éducation de l'Université de Sherbrooke. Ses travaux portent sur le développement professionnel en évaluation dans des contextes innovants tels les parcours de professionnalisation, l'IAG en évaluation des apprentissages et les perspectives de décolonisation. Elle œuvre également comme experte internationale auprès de l'Unesco et de l'OCDE sur les thématiques de l'évaluation diagnostique et différenciée en contextes de ruptures pédagogiques en Afrique et en Asie.

Lambert Félix Prudent est professeur émérite de sciences du langage à l'Université des Antilles. Il a été directeur de Centre IUFM de Martinique, directeur de l'ESPE Guadeloupe, responsable de l'Unité Mixte de Recherches LCF (Créoles et Francophonie) basée à l'Université de La Réunion et rédacteur en chef de la revue *Etudes Créoles*. Son champ de recherches embrasse la sociogenèse des langues créoles, l'analyse des pratiques langagières interlectales, et l'aménagement glottopolitique des collectivités en Outremer.

Denis Savard est professeur en évaluation institutionnelle à l'Université Laval. Il dirige le Centre de recherche et d'intervention sur la réussite scolaire (CRIRES), il est chercheur régulier au Laboratoire interdisciplinaire de recherche sur l'enseignement supérieur (LIREs), et chercheur associé au Centre de recherche et d'intervention sur l'éducation et la vie au travail (CRIEVAT). Ses recherches portent sur la réussite scolaire et la performance des établissements et des systèmes éducatifs.

Patricia Schillings est chargée de cours, responsable du service Analyse et Accompagnement du Développement Professionnel des Enseignants de l'Université de Liège. Ses recherches qui conjuguent visée explicative et compréhensive ciblent des thématiques telles que le développement professionnel des enseignants et de leurs formateurs dans le domaine de la littératie. L'accompagnement d'équipes éducatives dans la mise en place de pratiques didactiques ancrées dans des données probantes ainsi que

l'analyse de leur activité en situation constituent également un des axes majeurs de son travail.

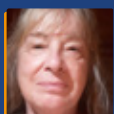
Gonzague Yerly est Professeur ordinaire à la Haute École pédagogique Fribourg (Suisse), où il est responsable de l'unité de recherche Évaluation & Formation (EVAFOR). Ses enseignements portent sur la didactique générale, l'évaluation des apprentissages, les difficultés d'apprentissage et le soutien à l'apprentissage. Ses recherches portent sur les pratiques enseignantes (évaluation, planification, collaboration, différenciation, classes flexibles) et sur les politiques éducatives (évaluation externe, politiques de pilotage).

Cet ouvrage regroupe un ensemble de textes dont le contenu a été présenté au 33^{ème} colloque de l'ADMEE-Europe en 2022. Tenu pour la première fois en Guadeloupe, il a été organisé par l'unité de recherche "Contextes, Recherches et Ressources en Éducation et Formation" de l'université des Antilles, avec l'appui du pôle Guadeloupe de l'université, et en partenariat avec la Région académique de Guadeloupe.

Ce colloque et l'ouvrage qui en résulte traduisent de manière unique la collaboration entre des communautés scientifiques et éducatives issues de la Caraïbe, d'Europe et du Québec, renforçant ainsi la portée interculturelle de sa thématique qui questionne la pluralité des contextes et des évaluations en éducation ainsi que leurs relations et enjeux. Les textes présentés abordent des problématiques évaluatives actuelles tant sur les plans conceptuels que pratiques.



Elisabeth Issaieva est maîtresse de conférences en sciences de l'éducation et de la formation à l'INSEI, Université de Cergy Paris. Ses travaux ciblent l'étude des interactions entre les pratiques d'enseignement, d'évaluation et les apprentissages des élèves dans différents contextes et notamment en contextes de grande diversité.



Isabelle Nizet est professeure titulaire retraitée, associée au département de pédagogie de la faculté d'éducation de l'Université de Sherbrooke. Elle a présidé l'ADMEE-Canada de 2016 à 2022. Elle agit à titre d'expert international sur les enjeux contemporains de l'évaluation scolaire.



Lucie Mottier Lopez est professeure ordinaire à la Faculté de psychologie et des sciences de l'éducation de l'Université de Genève. Elle a présidé l'ADMEE-Europe entre 2009 et 2013. Elle dirige la Revue LEeE sur l'évaluation en éducation et en formation.



Presses de l'ADMEE, Faubourg de l'Hôpital 43,
Case postale 556 CH-2002 Neuchâtel SUISSE

www.presses-admee.com