

Quelles sont les approches mises en place par les enseignants du secondaire inférieur de la Fédération Wallonie-Bruxelles en matière d'évaluation, leur sentiment d'efficacité à cet égard ainsi que leurs besoins de formations et leurs modalités préférées de formations, et quels liens peut-on établir entre ces dimensions ?

Auteur : Robaye, Clélia

Promoteur(s) : Fagnant, Annick

Faculté : Faculté de Psychologie, Logopédie et Sciences de l'Éducation

Diplôme : Master en sciences de l'éducation, à finalité spécialisée en enseignement

Année académique : 2025-2026

URI/URL : <http://hdl.handle.net/2268.2/25356>

Avertissement à l'attention des usagers :

Tous les documents placés en accès ouvert sur le site le site MatheO sont protégés par le droit d'auteur. Conformément aux principes énoncés par la "Budapest Open Access Initiative"(BOAI, 2002), l'utilisateur du site peut lire, télécharger, copier, transmettre, imprimer, chercher ou faire un lien vers le texte intégral de ces documents, les disséquer pour les indexer, s'en servir de données pour un logiciel, ou s'en servir à toute autre fin légale (ou prévue par la réglementation relative au droit d'auteur). Toute utilisation du document à des fins commerciales est strictement interdite.

Par ailleurs, l'utilisateur s'engage à respecter les droits moraux de l'auteur, principalement le droit à l'intégrité de l'oeuvre et le droit de paternité et ce dans toute utilisation que l'utilisateur entreprend. Ainsi, à titre d'exemple, lorsqu'il reproduira un document par extrait ou dans son intégralité, l'utilisateur citera de manière complète les sources telles que mentionnées ci-dessus. Toute utilisation non explicitement autorisée ci-avant (telle que par exemple, la modification du document ou son résumé) nécessite l'autorisation préalable et expresse des auteurs ou de leurs ayants droit.

Quelles sont les approches mises en place par les enseignants du secondaire inférieur de la Fédération Wallonie-Bruxelles en matière d'évaluation, leur sentiment d'efficacité à cet égard ainsi que leurs besoins de formations et leurs modalités préférées de formations, et quels liens peut-on établir entre ces dimensions ?

Mémoire présenté par Mazzetto Riccardo et Robaye Clélia

En vue de l'obtention du diplôme de
Master en Sciences de l'Éducation,
à finalité spécialisée en enseignement.

Promotrice : Annick Fagnant

Superviseur : Jean Baron

Lecteurs : Christian Monseur et Fabian Pressia

Année académique 2025-2026

Remerciements

Nous remercions tout d'abord Madame Fagnant et Monsieur Baron pour leur accompagnement, leurs conseils avisés, leur patience et leurs encouragements tout au long de ce mémoire de fin d'études. Cela nous a permis plusieurs fois de nous recentrer sur les objectifs principaux de ce mémoire. Ils ont été pour nous nos garde-fous.

Nous remercions ensuite nos lecteurs, Messieurs Monseur et Pressia pour l'intérêt porté à notre recherche et le temps consacré à la lecture et à l'analyse de ce travail.

Nous remercions également les enseignants ayant participé à cette étude, que ce soit lors du laboratoire cognitif ou lors de la complétion du questionnaire en ligne. Le temps consacré à cette recherche, ainsi que la contribution de chacun, le partage d'une partie de leur identité professionnelle, a permis la réalisation de ce travail. Nous leur en sommes extrêmement reconnaissants.

Enfin, nous remercions nos familles, amis et notre chat pour le soutien émotionnel inestimable. Nous savons que notre absence dû au travail engendré par ces études leur a pesé et nous les remercions encore une fois de leur compréhension.

Table des matières

1.	Introduction générale.....	1
2.	Cadre théorique	3
2.1.	Les étapes de conception de l'évaluation	3
2.2.	Teacher assessment literacy.....	5
2.2.1.	Vocabulaire.....	5
2.2.2.	Évolution des définitions.....	5
2.2.3.	Définition finale	7
2.2.4.	Trois différentes approches de l'évaluation.....	9
2.3.	Self-efficacy en évaluation	14
2.3.1.	Définition du self-efficacy.....	14
2.3.2.	Le self-efficacy des enseignants.....	15
2.3.3.	Le self-efficacy des enseignants en lien avec l'acte d'évaluer.....	16
2.4.	Formation et besoins des enseignants	16
2.4.1.	Définition de la formation professionnelle continue.....	17
2.4.2.	Définition du développement professionnel	17
2.4.3.	Définition des besoins de formation.....	18
2.4.4.	Modalités de formations	19
2.4.5.	Perspectives pour les formations futures.....	20
3.	Question de recherche et objectifs.....	21
4.	Méthodologie	23
4.1.	Introduction	23
4.2.	Création du questionnaire et analyses statistiques envisagées	23
4.2.1.	Partie 1 : Identifier, parmi les répondants, des profils d'évaluateurs.	24
4.2.2.	Partie 2 : Constater et questionner dans quelle mesure un lien existe entre le profil d'évaluateur et d'une part, le sentiment d'efficacité lié à l'évaluation et, d'autre part les besoins de formations liées à l'évaluation	27

4.2.3.	Partie 3 : Constater les préférences de modalités de formations des clusters d'enseignants et les liens que nous pourrions établir avec les différentes approches.	30
4.2.4.	Partie 4 : informations démographiques des participants permettant de décrire l'échantillon.....	31
5.	Présentation des résultats : analyse	32
5.1.	Échantillon	32
5.2.	Partie 1 du questionnaire : les approches	32
5.2.1.	Fiabilité, consistance interne et analyse factorielle des approches	32
5.2.2.	Détermination des clusters/profils.....	36
5.2.3.	Description des clusters.....	36
5.2.4.	Pratiques faisant l'unanimité.....	37
5.3.	Partie 2 du questionnaire : Sentiment d'efficacité et besoins de formations.....	38
5.3.1.	Fiabilité, consistance interne et analyse factorielle du sentiment d'efficacité .	38
5.3.2.	Analyse du sentiment d'efficacité de l'échantillon	39
5.3.3.	Fiabilité, consistance interne et analyse factorielle des besoins de formations	41
5.3.4.	Analyse des besoins de formations de l'échantillon	43
5.3.5.	Analyse des liens entre le sentiment d'efficacité et les besoins de formations	45
5.4.	Partie 3 du questionnaire : Modalités de formations.....	47
6.	Interprétation et discussion.....	49
6.1.	Constat des approches d'évaluations.....	49
6.2.	Constat des sentiments d'efficacité	52
6.3.	Constat des besoins de formations	53
6.4.	Constat des modalités de formations.....	54
6.5.	Constat des liens.....	55
6.5.1.	Approches d'évaluation – SE	55
6.5.2.	Approches d'évaluation – BF.....	55
6.5.3.	Approches d'évaluations – modalités de formations	56
6.5.4.	SE – BF	56

6.5.5.	Approches d'évaluations – SE – BF	57
7.	Conclusion, limites et perspectives	58
8.	Bibliographie	60
9.	Annexes	66

Table des figures

Figure 1. Cycle MoSCodEE (Detroz et al., 2020)	3
Figure 2. Schéma de l'AL (Pastore, 2023)	8
Figure 3. Graphique des clusters	37
Figure 4. Graphique des SE	41
Figure 5. Graphique des BF	44
Figure 6. Présentiel vs En ligne	49
Figure 7. Seul vs En groupe	49

Table des tableaux

Tableau 1. Caractérisation de l'AOL, de l'AFL et de l'AAL (Earl, 2003)	11
Tableau 2. Caractéristiques des 3 approches dans ce mémoire	12
Tableau 3. Descriptif de l'échantillon	32
Tableau 4. Ajustement du modèle de la partie 1 du questionnaire	34
Tableau 5. α , ω et ajustement du modèle de la partie 1 du questionnaire après réajustement	35
Tableau 6. Matrice de corrélation entre les scores obtenus dans chaque approche ...	36
Tableau 7. Centres de chaque cluster	37
Tableau 8. Actions faisant l'unanimité	38
Tableau 9. α , ω et ajustement du modèle de la partie SE	39
Tableau 10. α , ω et ajustement du modèle de la partie SE après réajustement	40
Tableau 11. α , ω et ajustement du modèle de la partie BF	42

Table des Annexes

I.	Fiche IA.....	i
II.	Responsabilité partagée.....	iii
III.	Questionnaire (avec les codes utilisés dans le texte).....	v
a.	Partie 1 du questionnaire	vi
b.	Partie 2 du questionnaire	ix
c.	Partie 3 du questionnaire	xviii
d.	Partie 4 du questionnaire	xx
IV.	Laboratoires cognitifs.....	xxiii
a.	Laboratoire cognitif n°1 : enseignante de sciences	xxiv
b.	Laboratoire cognitif n°2 : enseignant de mathématiques	xxvii
c.	Laboratoire cognitif n°3 : enseignante d'anglais.....	xxxi
d.	Laboratoire cognitif n°4 : enseignante de géographie.....	xxxii
e.	Laboratoire cognitif n°5 : enseignante de français	xxxiii
f.	Laboratoire cognitif n°6 : enseignant de français.....	xxxv
g.	Laboratoire cognitif n°7 : enseignante de cours commerciaux	xxxvii
V.	Analyses supplémentaires avant la création du questionnaire	xxxix
a.	Justification des thèmes de la partie 1 du questionnaire	xl
b.	Analyse plus détaillée du lien entre item et approche dans la partie 1 du questionnaire xliii	
c.	Littérature en lien avec les derniers items de la partie 2	l
VI.	Analyse statistique de la construction de la première partie	lv
a.	Fiabilité par approche.....	lvi
b.	Analyse factorielle confirmatoire : contribution des items	lvii
c.	Boîtes à moustache des scores en AOL, AFL et AAL.....	lviii
VII.	Détermination du nombre de clusters.....	lix
a.	2 clusters vs 3 clusters	lx
b.	Nombre optimal de clusters selon Jamovi.....	lxi

VIII.	Les 3 clusters	lxii
a.	Par rapport au sexe	lxiii
b.	Par rapport à l'âge/années d'expériences	lxiv
c.	Par rapport aux matières enseignées	lxv
d.	Par rapport au titre pédagogique	lxvi
e.	Par rapport à la filière	lxvii
f.	Premier choix de carrière	lxviii
g.	Souvenir d'un cours sur l'évaluation pendant la formation initiale	lxix
h.	Souvenir d'un cours sur l'évaluation pendant la formation continue	lxx
IX.	Analyse statistique de la construction de la deuxième partie	lxxi
a.	Analyse factorielle confirmatoire des types de self-efficacy	lxxii
b.	Analyse factorielle confirmatoire des types de besoins de formations	lxxvii
c.	Matrice de corrélation des SE	lxxxii
X.	Le self-efficacy par rapport aux clusters	lxxxiii
a.	Les 4 types de self-efficacy	lxxxiv
b.	Self-efficacy général	lxxxv
c.	Self-efficacy AOL	lxxxvi
d.	Item SE AOL 5	lxxxvii
e.	Self-efficacy AFL	lxxxviii
f.	Self-efficacy AAL	lxxxix
g.	Item SE AAL6	xc
XI.	Les besoins de formations par rapport aux clusters	xc i
a.	Item besoin de formation global	xcii
b.	Les 4 types de besoins de formations	xciv
c.	Besoin de formation général	xcv
d.	Item MOSCO BF2	xcvi
e.	Item MOSCO BF1	xcvii

f.	Besoins de formations AOL	xcix
g.	Item BF AOL1	c
h.	Besoins de formations AFL	ci
i.	Besoins de formations AFL : tableaux de fréquences pour chaque item	cii
j.	Besoins de formations AAL	cvii
k.	Items BF AAL1 et BF AAL2 : tableau de fréquences.....	cviii
XII.	Lien self-efficacy/besoins de formations [SE/BF]	cx
a.	Corrélations faibles binômes SE/BF	cxix
b.	SE/BF global	cxiii
c.	SE/BF AAL2	cxiv
d.	SE/BF changement de pratiques	cxv
e.	2 corrélations SE/BF supplémentaires analysées	cxvii
f.	SE/BF AAL6	cxix
XIII.	Modalités de formations.....	cxx
a.	Modalités de formations sans clusters.....	cxxi
b.	Modalités de formations avec clusters	cxxii

1. Introduction générale

Cette recherche vise à s'intéresser aux dynamiques complexes entre l'approche des enseignants par rapport à l'évaluation, leur confiance en leur capacité à évaluer et leurs besoins de formations ainsi que leurs modalités préférées de formations en évaluation.

Pour y parvenir, il est essentiel de considérer le contexte actuel des politiques éducatives. De nombreux pays de l'OCDE font le choix de s'orienter vers l'« évaluation soutien d'apprentissage » [ESA] comme un outil parmi d'autres pour réduire l'échec scolaire (Pasquini & Morales Villabona, 2022, p. 514). En Fédération Wallonie-Bruxelles [FWB], le Pacte d'Excellence intègre cette logique en promouvant l'« approche évolutive des difficultés d'apprentissage », dans laquelle l'évaluation occupe une place centrale. La FWB a publié en 2025 le guide « L'évaluation au service des apprentissages de tous les élèves dans le cadre du tronc commun », dans lequel l'évaluation décrite est fort proche de l'ESA. Ce tronc commun débutera en première secondaire en septembre 2026. Or, DeLuca et ses collègues (2019a) constatent que les enseignants se sentent souvent insuffisamment préparés à appliquer de nouvelles politiques d'évaluation et Duru-Bellat (2009) met en avant le fait que les réformes venues d'en haut ont une faible portée si les enseignants n'en partagent pas la philosophie : ceux-ci résistent passivement. De ce fait, des propositions de formations existent depuis la rentrée 2025 (FWB, s.d.), et notamment via le catalogue de l'Institut interréseaux de la formation professionnelle continue. Avant de prendre part à des formations, il faudrait, selon nous, identifier les conceptions de l'évaluation des enseignants car, selon Pastore (2023), cela est pertinent pour comprendre les dynamiques impliquées dans la mise en œuvre des politiques éducatives. Ces conceptions de l'évaluation nous mènent, au fil de nos lectures, à comprendre que ce n'est qu'une partie du concept d'assessment literacy [AL] des enseignants (la littératie en évaluation des enseignants). Cette notion d'AL regroupe, entre autres, les connaissances, les capacités, les expériences, les états d'esprit, les croyances, les conceptions, les pratiques à mettre en œuvre pour une évaluation de qualité de manière à soutenir l'apprentissage des élèves (Van Orman, 2024 ; DeLuca et al., 2020).

L'enquête TALIS 2024, comprenant des questions nationales portées, entre autres, sur l'utilisation de l'évaluation, donne ces informations (Quittre et al., 2025) : 93% d'enseignants du premier degré de la FWB déclarent mettre en place « souvent » ou « toujours » des pratiques d'évaluation pour vérifier l'acquisition de la matière vue en classe, 38% font des interrogations non cotées qui aident les élèves à situer leurs apprentissages et 25% demandent aux élèves

d'évaluer eux-mêmes leurs progrès. Ces résultats montrent une faible utilisation des formes d'évaluation qui soutiennent l'apprentissage des élèves, voire qui mettent l'élève au centre de son apprentissage. Cependant, ces résultats, bien qu'obtenus à l'aide d'une étude valable avec un échantillon représentatif, n'ont été obtenus qu'à partir de quelques items précis, bien choisis, vu qu'il y avait d'autres thématiques à traiter. Certains pourraient donc avoir répondu ne pas effectuer certains types d'évaluation alors que, dans la pratique, ils les mettent en place sans s'en rendre compte, par méconnaissance du sujet. De plus, on peut utiliser une conception formative de l'évaluation tout en n'effectuant pas obligatoirement d'évaluation non cotée. Il serait donc intéressant de questionner, sur davantage d'items, les pratiques évaluatives déclarées des enseignants dont des formes d'évaluation mettant l'élève au centre de son apprentissage.

Parmi toutes les notions concernant l'AL des enseignants, le choix s'est porté sur les besoins de formations [BF] en évaluation et le self-efficacy [SE] en évaluation. Ce choix est dû à deux raisons. Premièrement, ce SE est difficilement dissociable de l'AL vu qu'il est considéré comme un élément clé de l'AL qui elle-même est fondée sur les connaissances en évaluation (Levy-Vered & Nasser-Abu Alhija, 2015). D'ailleurs, toujours selon ces auteures, une partie de la variance du SE peut être expliquée par l'AL et la formation en évaluation. Deuxièmement, dans notre pratique, nous avons observé que certains enseignants ne faisant pas partie des plus formés ont une plus grande confiance en leurs capacités. Certains n'effectuant que des évaluations sommatives, ont une grande confiance en leurs capacités à créer et utiliser des évaluations. Les personnes qui auraient donc le plus besoin de formations ne seraient peut-être pas obligatoirement en recherche de formations. Ceci est lié au phénomène de « méta-ignorance » (ne pas savoir qu'on ne sait pas) évoqué par Dunning (2011, cité par Kruse et al., 2020).

Les objectifs de ce mémoire sont donc multiples. Nous voulons questionner les pratiques évaluatives, le SE, le BF et les modalités de formations préférées des enseignants du secondaire inférieur en FWB et les liens entre ceux-ci. Interroger les pratiques évaluatives des enseignants nous semble inévitablement lié à interroger les étapes de conceptions d'une évaluation bien conçue, ceci pour constater le SE et le BF des professeurs à chaque étape.

Pour atteindre ces objectifs, un cadre théorique a été effectué sur les étapes de conceptions de l'évaluation ainsi que sur les trois notions principales de notre mémoire (AL, SE et BF).

Riccardo Mazzetto s'est focalisé sur le SE et Clélia Robaye sur le BF. Notre partenariat dans ce travail nous permet d'avoir une confrontation de points de vue et une confrontation de 2 notions par rapport à l'évaluation qui n'auraient pas été possibles sans ce binôme.

Ce mémoire se poursuit avec la méthodologie de ce travail, envisagée de façon quantitative via la création et la passation d'un questionnaire, dont les données seront traitées statistiquement. Une présentation des résultats et une discussion sur ceux-ci sont réalisées. Une conclusion sur les points saillants de ce travail, des limites et des perspectives clôture ce mémoire.

2. Cadre théorique

Ce cadre théorique débute avec le Modèle Structuré de Conception d'Évaluation des Étudiants, aussi appelé cycle MoSCodEE, ceci permettant de comprendre la réflexion, en étapes, des enseignants quand ils évaluent. Ensuite, l'AL des enseignants est défini. L'AL est le concept général duquel découle cette recherche. Par la suite, des définitions sont données quant aux approches de l'évaluation, au SE en évaluation et aux BF en évaluation. Le cadre théorique s'achève sur les modalités de formations et des perspectives pour les formations futures.

2.1. Les étapes de conception de l'évaluation

L'évaluation en enseignement est un concept qui a évolué, qui a pris de l'ampleur dans les recherches scientifiques et s'est diversifié. Plusieurs auteurs ont essayé de résumer toutes ces activités en différents aspects mais aucun des modèles, rencontrés lors de nos lectures, n'apporte, selon nous, une vision aussi détaillée que celle du cycle MoSCodEE dont nous avons eu connaissance au cours de « Questions d'évaluation » dispensé par M. Detroz (2023) à l'Université de Liège. Ce cycle reprend 9 étapes, présentées dans le schéma ci-dessous (Detroz et al., 2020, Detroz, 2023).

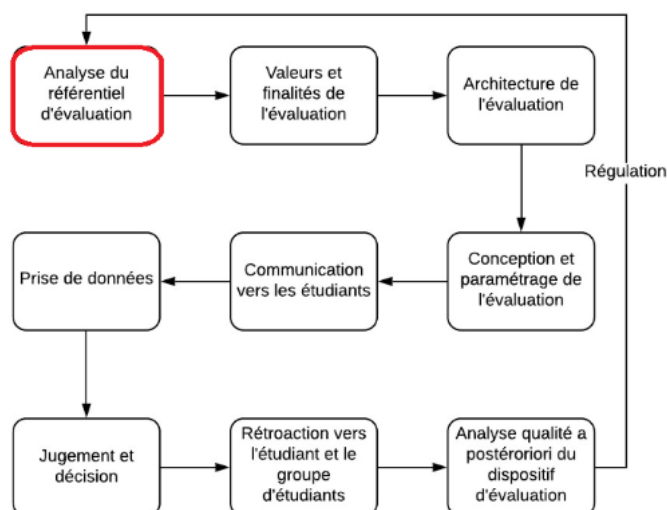


Figure 1. Cycle MoSCodEE (Detroz et al., 2020)

Chaque étape offre une piste de réflexion pour s'approcher de l'évaluation la plus valide possible des compétences des élèves. Ces étapes sont définies ci-après.

L'analyse du référentiel d'évaluation permet de circonscrire la matière à évaluer, donc la matière à enseigner, et donc de s'assurer de l'alignement pédagogique. Une table de spécification peut être réalisée, par exemple, pour guider le processus.

L'étape des valeurs et finalités de l'évaluation permet de définir le modèle d'évaluation que l'on souhaite mettre en place : évaluation certificative, formative, ...

L'architecture de l'évaluation est l'étape dans laquelle les enseignants choisissent les modalités d'évaluation : questions ouvertes ou fermées, QCM ou non, écrit ou oral, papier-crayon ou numérique, ... Ils choisissent également la temporalité de l'évaluation : l'articulation des séquences d'évaluations et les calendriers respectifs.

La conception et le paramétrage de l'évaluation concernent l'opérationnalisation de l'architecture : rédaction des questions, paramétrage de l'évaluation, choix des barèmes de points, choix des pondérations entre questions, ... Il faut également à cette étape s'assurer de la qualité des items.

La communication vers l'étudiant reprend, les enjeux de l'évaluation, la matière de l'évaluation, les modalités d'évaluation utilisées, le mode de correction, ... Il faut être le plus précis possible et laisser le moins de place possible à l'ambiguïté pour diminuer la charge mentale des élèves.

La prise de données est la passation de l'évaluation par les élèves.

L'étape de jugement et décision est celle qui permet d'attribuer une note à l'élève sur base des données recueillies. L'enseignant corrige via une grille d'évaluation (critériée, générale, ...), régule (recalibre sa grille si besoin) et recorrige si nécessaire.

La rétroaction vers l'étudiant et le groupe d'étudiants est l'étape dans laquelle l'enseignant rend des feedbacks aux élèves sur base de l'épreuve passée. L'enseignant soulève les points forts et les faiblesses de chacun. Cela permet aux élèves de s'améliorer et, si nécessaire, de se préparer à une deuxième évaluation en ayant l'aide de l'enseignant pour s'autoréguler. Un feedback comprenant au moins les réponses aux questions ou un corrigé-type, donné le plus rapidement possible après l'évaluation, est conseillé.

L'analyse qualité a posteriori du dispositif d'évaluation est une récolte d'informations sur le processus utilisé afin de faire les modifications nécessaires pour les prochaines évaluations. En effet, les évaluations étant des construits humains, il est nécessaire de porter un regard réflexif

dessus. C'est également à cette étape que l'enseignant prend en compte les retours des élèves sur les évaluations proposées.

Ce processus en 9 étapes est intéressant afin de mettre en lumière différents aspects de l'évaluation. Au départ, pensé en ordre chronologique et successif, il est maintenant pensé en termes de focales pour analyser ses pratiques (Detroz et al., 2020). Cela va justement être un des prochains points d'attention : observer le SE et le BF pour chaque étape de ce cycle et pour le cycle complet. Le SE étant spécifique aux compétences, aux tâches et au domaine d'enseignement (Bong, 2006, citée par Kyttälä et al., 2023), il est important de prendre en compte toutes ces facettes de la conception de l'évaluation pour mieux appréhender le SE en évaluation des enseignants.

2.2. Teacher assessment literacy

Dans cette partie, l'évolution du concept d'AL est analysée en fonction de différentes définitions qu'il a pu avoir. L'évolution du concept fait évoluer les outils permettant de récolter des informations en lien avec l'AL. Un seul outil est envisagé en lien avec les définitions actuelles de l'AL.

2.2.1. Vocabulaire

« Assessment literacy » se traduit littéralement en « littératie en évaluation ». Il existe en revanche d'autres expressions utilisées pour parler du même concept. Les 3 autres manières principales repérées par Pastore (2023, p. 16) dans sa revue systématique sont : « teacher assessment competence » (la compétence des enseignants en évaluation), « teacher assessment capability » (la capacité des enseignants en évaluation) et « teacher assessment approaches » (les approches des enseignants en évaluation). Le fait que des mots avec des différences de sens comme littératie, compétence, capacité et approches, sont utilisés pour un même concept montre déjà que cette notion possède plusieurs facettes. Dans ce travail, le parti pris est d'utiliser l'expression la plus récurrente dans les articles scientifiques lus : « assessment literacy » [AL].

2.2.2. Évolution des définitions

Les tentatives au cours des 30 dernières années de définir l'AL sont aussi nombreuses que différentes sans pour autant parvenir à un consensus. Et cela est logique pour Willis et ses collaborateurs (2013) car l'AL est de nature contextuelle, culturelle et sociale. Elle n'est malheureusement pas très développée dans la communauté de recherche francophone (Pasquini & Morales Villabona, 2022).

Un des premiers à avoir défini l'AL est Stiggins (1991) qui la définit comme une compréhension de base de la signification d'une évaluation de haute et de basse qualité et comme la capacité d'appliquer ces connaissances à diverses mesures de la réussite des élèves de façon critique. Il effectue tout au long de son article une distinction entre les novices et les experts. Cette séparation montre deux des buts principaux des premières études et des politiques : déterminer les connaissances qui font d'un enseignant un bon praticien de l'évaluation et évaluer le niveau de littératie en évaluation des enseignants.

Le même esprit anime les définitions de certaines études des années 2000. Voici une définition de Paterno datant de 2001 :

La possession de connaissances sur les principes de base d'une bonne pratique d'évaluation ; ce qui comprend la connaissance des terminologies, le développement et l'utilisation de méthodologies et de techniques d'évaluation, la familiarité avec les critères de qualité en matière d'évaluation ... et avec des alternatives quant aux façons traditionnelles d'évaluer l'apprentissage [traduction libre] (cité par Mertler, 2004, pp. 50-51).

Cette définition, malgré un plus grand détail sur le type de connaissances, établit toujours un lien positif entre le nombre de connaissances et la qualité d'un évaluateur.

Brookhart (2011) et Popham (2013) définissent l'AL comme les capacités et les savoirs requis par les enseignants pour mesurer et soutenir l'apprentissage des élèves à travers l'évaluation (cités par DeLuca et al., 2016). De façon concordante avec Stiggins et Paterno, il y a toujours la notion de connaissances nécessaires. Cependant, le fait de parler de soutien d'apprentissage montre un grand changement. L'évolution du concept d'évaluation passe d'un acte permettant uniquement de juger la prestation d'un élève à un acte permettant également d'aider l'élève dans son apprentissage. Cette modification est due entre autres à l'intérêt généralisé pour l'évaluation formative (Pastore, 2023). Selon cette auteure, un autre changement substantiel dans les définitions de début 2010 est la reconnaissance de l'importance du contexte, de la nature socioculturelle de la pratique de l'évaluation comme le montre une des définitions les plus utilisées par d'autres chercheurs :

L'AL est une pratique sociale dynamique dépendante du contexte qui implique que les enseignants articulent et négocient les connaissances scolaires et culturelles entre eux et avec les apprenants, dans l'initiation, le développement et la pratique de l'évaluation pour atteindre les objectifs d'apprentissage des élèves [traduction libre] (Willis et al., 2013, p.242).

La notion de l'évaluation au service de l'apprentissage y est toujours présente mais de l'importance est également donnée au contexte. En effet, toute décision prise dépend énormément de la situation dans laquelle on se trouve. Il faut donc que, dans son développement, l'enseignant acquière des compétences conditionnelles adaptables à toutes circonstances. Cela va modifier fortement l'étude de l'AL en analysant la capacité à mettre en

pratique des connaissances à l'aide notamment de scénarios. Dans cette définition de Willis et ses collègues, on relève une autre évolution de l'évaluation grâce au concept de négociation entre eux et avec les apprenants. Ces chercheurs s'accordent sur le fait que l'évaluation est un processus collaboratif entre enseignants d'une part et entre l'enseignant et ses élèves d'autre part. Le fait d'intégrer les élèves dans le processus d'évaluation, augmente leur perception de contrôlabilité. Et, selon Viau (1998), un élève ayant cette perception a une plus grande motivation face à l'apprentissage.

D'autres chercheurs font remarquer que la négociation se fait également en fonction d'autres facteurs. Xu et Brown (2016) précisent que le contexte (l'ensemble des facteurs influençant la conception de l'évaluation) est une combinaison de traits cognitifs, de systèmes affectifs et de croyances, et d'influences socioculturelles et institutionnelles. C'est le dernier mot qui montre une notion encore peu évoquée dans les définitions précédentes. L'État, ses curricula, son discours politique, le climat du pays/région/école et ses normes professionnelles jouent un rôle important dans l'AL. Dans les programmes d'études, il y a ce que l'État indique comme ce qui est à enseigner en priorité mais il y a également les principaux attendus des élèves. Dans ce même article, Xu et Brown mettent en avant un autre facteur : l'identité d'évaluateur. Pour eux, au fur et à mesure que cette identité se construit, les pratiques et conceptions de l'évaluation évoluent. Cependant, il n'y a pas que la partie « évaluateur » de l'identité de l'enseignant qui influence ses pratiques évaluatives mais bien toute son identité. En effet, ses valeurs, son caractère, ses expériences passées, sa formation initiale et continue et ses convictions sur ce qui constitue une preuve valide et utile de l'apprentissage des élèves sont également des éléments qui influencent l'AL d'un enseignant (DeLuca et al., 2018).

2.2.3. Définition finale

Deux dernières définitions regroupant tous les éléments cités depuis le début de la partie concernée à la définition de l'AL sont reprises ci-dessous :

L'AL correspond aux connaissances, capacités, expérience, et états d'esprit nécessaires pour mettre en œuvre une évaluation en classe de qualité de manière à soutenir l'apprentissage ; composantes socio-émotionnelles de l'apprentissage, interaction, et relations avec les étudiants ; et communication/collaboration avec divers acteurs (entre autres : élèves, parents, administration) [traduction libre] (Van Orman, 2024, p. 5).

et

L'AL correspond aux connaissances, croyances et pratiques telles qu'elles se situent dans les contextes d'enseignement et d'apprentissage et qui sont influencées par de multiples facteurs, notamment les exigences politiques, le développement professionnel des enseignants, l'environnement d'apprentissage et les négociations enseignant-élève [traduction libre] (DeLuca et al., 2020, p. 27).

Dans ces deux dernières définitions, on peut noter la difficulté à lister tous les éléments présents dans l'AL. En effet, Van Orman (2024) et DeLuca et ses collaborateurs (2020) utilisent des expressions comme « entre autres », « de multiples facteurs, notamment... ». C'est pourquoi Pastore propose le schéma ci-dessous dans sa revue systématique (2023, p. 22) :

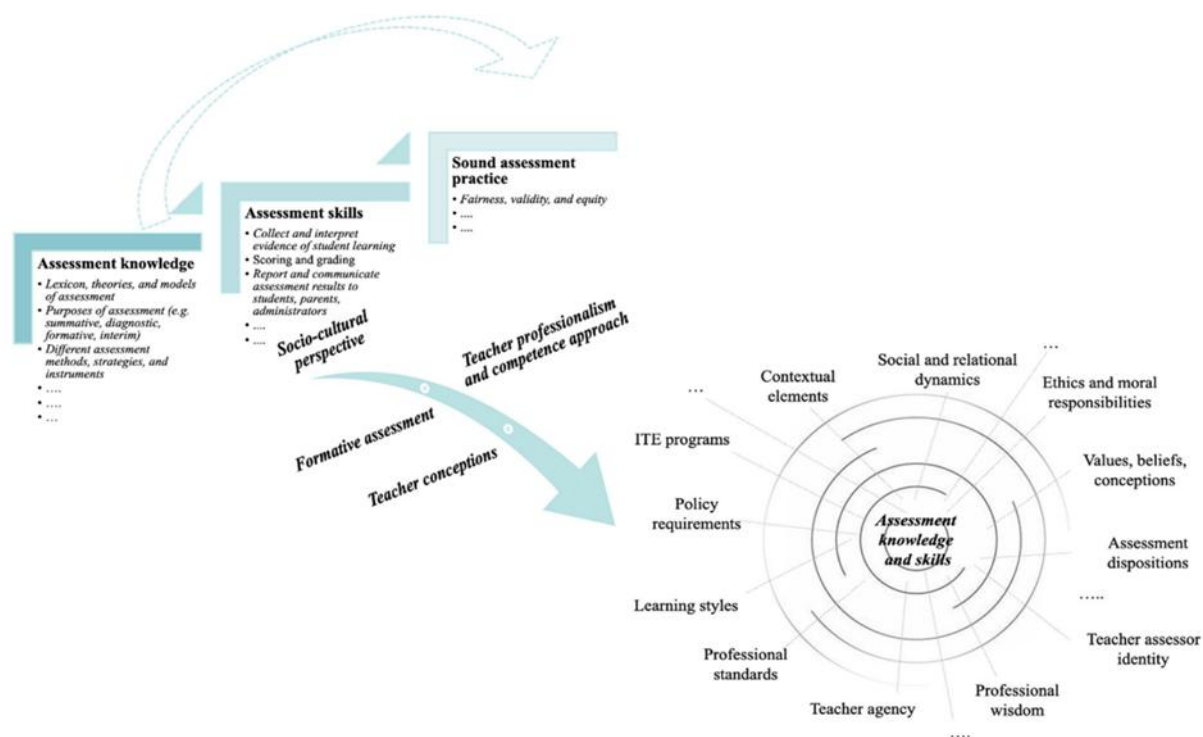


Figure 2. Schéma de l'AL (Pastore, 2023)

Comme DeLuca et ses collègues, des stratagèmes sont utilisés pour montrer l'impossibilité de tout lister comme des points de suspension à plusieurs endroits du schéma. Pourtant, ce schéma nous semble bien plus complet que toute autre définition. Pastore propose ce schéma dans sa revue systématique ce qui signifie que tous les concepts présents viennent d'autres chercheurs dont ceux cités dans ce travail. Parmi les notions de l'AL non encore abordées dans ce travail, nous aimerions parler d'une notion de plus en plus présente ces dernières années dans la littérature en sciences de l'éducation : le concept de fairness (le fait d'être équitable). Même si l'idée qu'il faut lutter contre la reproduction des inégalités sociales par l'institution scolaire existe depuis longtemps, les chercheurs insistent de plus en plus sur la manière de modifier chaque acte pédagogique pour qu'il soit plus équitable. Il en va de même pour l'acte d'évaluer. Cette définition et ce schéma montrent la volonté d'une vision holistique de l'AL. Le mot « développement » et les flèches présentes sur le schéma montrent une dernière idée à mettre en avant par rapport à l'AL : elle ne se résume pas à une liste de composantes cognitives et affectives figées. Elle évolue et s'améliore au cours du temps lors d'un long processus

d'apprentissage complexe et itératif (Pastore, 2023). Ces trajectoires de développement sont influencées par l'évolution de l'identité de l'enseignant. Ceci est en lien avec l'idée de Pasquini et Morales Villabona (2022) d'élargir la définition de la littératie en évaluation à « l'analyse de l'évolution des dimensions identitaires et culturelles de l'enseignant » (p. 547).

Selon Pastore (2023), l'AL doit donc être pensée comme une interaction complexe de composantes (connaissances, compétences et dispositions) dans laquelle il n'y a pas de perspective hiérarchique et ascendante. C'est une des raisons principales pour lesquelles, ce mémoire se focalise sur ce qui est pratiqué en classe. En effet, certains enseignants ont les connaissances et compétences nécessaires à l'utilisation de nouvelles formes d'évaluation mais ne sont pas disposés à les utiliser pour diverses raisons telles que : mauvaises expériences, incompatibilité avec la classe, manque de formation, manque de temps, ... (Graham et al., 2021 ; Jager et al., 2022). D'autres enseignants peuvent essayer de pratiquer les nouvelles formes d'évaluation tout en ayant très peu de connaissances à leur sujet.

2.2.4. Trois différentes approches de l'évaluation

Qu'entend-on par « approche » ?

À la lecture et à l'analyse des nombreux aspects qui composent l'AL d'un enseignant, il nous paraît impossible de définir des catégories distinctes d'AL appelées approches : il y a trop d'éléments (connaissances, compétences, dispositions, valeurs, conceptions, responsabilités éthiques, ...) pour pouvoir obtenir des unités qui définiraient des approches d'AL parmi les enseignants. L'utilisation de l'outil, représentatif des pratiques actuelles d'évaluation, « Approaches to Classroom Assessment Instrument » [ACAI] le confirme (DeLuca et al., 2016). Dans l'outil ACAI, les participants sont immergés dans 5 scénarios. Pour chacun des 4 thèmes (les objectifs, le processus, l'éthique et la théorie de la mesure), 3 actions sont proposées, chacune correspondant à une approche évaluative distincte construite sur des bases théoriques. Différentes tentatives de façons de répondre (choisir une réponse, prioriser ses réponses, se positionner sur une échelle de Likert) ont été testées. Malgré cela, aucune approche n'a été vérifiée empiriquement et aucune validité interne de ce questionnaire n'a pu être dégagée de façon solide (Baron & Fagnant, 2025). Toujours selon ces auteurs, les raisons de cet « échec » pourraient être l'utilisation de différents scénarios qui font varier les profils d'évaluateurs. Selon nous, il est possible que, malgré une base théorique existante, dans la réalité, il n'existe pas de lien entre tous les items A (ou B ou C) des différents thèmes. En effet, un enseignant pourrait mettre l'accent sur les évaluations formatives tout en ne pratiquant pas une réflexion sur le traitement équitable des étudiants.

C'est pourquoi, même si la même expression que celle de DeLuca et ses collègues est utilisée dans ce travail (approches de l'évaluation), ce terme ne renverra qu'à un seul thème d'ACAI : les buts d'évaluation. Ces buts sont : l'évaluation de l'apprentissage (Assessment of Learning [AOL]), l'évaluation pour l'apprentissage (Assessment for Learning [AFL]) et l'évaluation en tant qu'apprentissage (Assessment as Learning [AAL]).

Les trois différentes approches dans la littérature

Dans l'outil ACAI de DeLuca et ses collègues, la multitude de trajectoires existantes sont regroupées en trois approches différentes. La première intitulée approche A et contenant l'AOL se concentre sur les évaluations sommatives montrant une importance portée sur la fiabilité de l'évaluation et sur la standardisation des évaluations. La deuxième appelée approche B et contenant l'AFL met l'accent sur les évaluations formatives, l'aspect validité, l'utilisation et la notation des évaluations et sur le traitement équitable des étudiants. La troisième nommée approche C contenant l'AAL est une approche plus individualisée qui prend davantage en compte la fiabilité, la validité et la différenciation en utilisant la communication et l'évaluation comme soutien d'apprentissage (DeLuca et al., 2016). Comme écrit supra, ces descriptions prennent en compte trop de caractéristiques qui ne forment pas obligatoirement une unité.

Mottier-Lopez (2022) indique qu'il n'y a pas de consensus entre les scientifiques de l'éducation sur les définitions des différentes approches. Dans la suite de ce travail, une analyse de quelques définitions des 3 approches (AOL, AFL et AAL) est réalisée. Ceci a pour but d'en créer une pour chaque approche à utiliser pour l'élaboration du questionnaire.

Parmi les définitions les plus utilisées, il y a celles de Black et William (2018, cités par Fagnant, 2023, p. 15) :

- Assessment of learning fait référence à l'utilisation de l'évaluation pour décider dans quelle mesure les élèves ont atteint les objectifs d'apprentissage ;
- Assessment for learning correspond au processus mis en œuvre par l'enseignant pour utiliser les résultats de l'évaluation en vue d'informer [et de réguler] son enseignement ;
- Assessment as learning est utilisé dans un sens très spécifique pour décrire le rôle de l'élève afin de monitorer et diriger son propre apprentissage.

Dans cette définition, l'utilisation de la préposition « pour » à 3 reprises, ainsi que les expressions « en vue de » et « afin de », indique que ce qu'on va entendre par « approche » est bien le but visé par l'enseignant dans son utilisation de l'évaluation, ce qui rejoint la conception des buts d'évaluation de DeLuca et ses collègues. Or, et c'est ce qui a parfois posé problème dans certains outils d'analyse de l'AL, un enseignant peut, à l'aide de la même évaluation, chercher à atteindre plusieurs objectifs. Il peut même avoir différents buts sans qu'un soit

prioritaire sur l'autre. C'est sans doute ce qui a poussé DeLuca et ses collaborateurs à demander à leurs participants de ne plus choisir une réponse mais de se positionner sur une échelle de Likert pour chaque réponse. L'objectif donc, dans la description des 3 approches, ne va pas être de décrire des buts opposés mais bien des buts distincts qui peuvent être complémentaires. C'est d'ailleurs une raison pour laquelle nous émettons une critique à cette définition. L'expression « dans quelle mesure l'élève a atteint les objectifs d'apprentissage » peut impliquer une analyse des compétences acquises et non acquises et dans quelle mesure elles ne sont pas acquises. Or, cette analyse peut être faite également pour informer l'élève sur son apprentissage. Nous y voyons donc un chevauchement qu'on ne peut pas accepter dans une tentative d'obtenir des définitions distinctes. Une autre critique est que nous ne voyons pas, dans la définition de l'AOL, un but autre qu'évaluer les acquis. Or évaluer les acquis est dans la définition même de l'évaluation. Selon nous, il n'y a donc pas réellement de but qui accompagne cette évaluation.

Le tableau ci-dessous, contenant les caractéristiques de chaque approche de Earl en 2003 (citée et traduite par Fagnant, 2023, p. 16), nous permet d'obtenir ce but qui manque à l'AOL.

Approche	Objectif	Point de référence	Évaluateur principal
Assessment of learning	Jugements concernant le rang (classement), le passage dans un niveau supérieur, les diplômes, etc.	Les autres élèves	L'enseignant
Assessment for learning	Informations utiles pour les prises de décisions pédagogiques par les enseignants	Critères ou attentes externes	L'enseignant
Assessment as learning	Autocontrôle et autocorrection ou ajustement	Objectifs personnels et critères externes	L'étudiant

Tableau 1. Caractérisation de l'AOL, de l'AFL et de l'AAL (Earl, 2003)

Malgré plusieurs critiques émises par certains auteurs, cette caractérisation apporte des éléments non repris dans les définitions de Black et Wiliam (2018, cités par Fagnant, 2023) car il y a davantage de distinctions dans les différents buts attribués aux 3 approches surtout en ce qui concerne l'AOL. Ici, il y a un but précis à l'analyse qui est celui de juger du rang ou du passage dans un niveau supérieur. Nous émettons tout de même une critique : le jugement du passage dans un niveau supérieur ou de la réussite d'une interrogation ne se fait pas obligatoirement consciemment en se référant aux autres élèves. Malgré l'existence du phénomène régulier qui fait qu'à la fin de l'année les résultats d'une classe ressemble à une courbe de Gauss, la comparaison entre élèves est parfois faite de manière inconsciente. De plus, pour juger d'un rang ou d'un classement, une cote, chiffrée ou non, est très souvent utilisée, ce qui n'est nullement mentionné. L'OCDE (2012, cité par Pasquini, 2021, p. 4) confirme que peu

importe les contextes, « l'évaluation scolaire se réduit le plus souvent à passer une épreuve aboutissant à une note ». Parmi les nombreuses critiques émises par Fagnant (2023) sur ce tableau, il y a également l'absence du rôle des pairs dans les 3 approches. Nous éprouvons donc le besoin de créer notre propre tableau comparatif de ces 3 approches.

Les trois différentes approches dans ce mémoire

En essayant de regrouper toutes les informations prises dans la littérature citée supra et d'y apporter des modifications pour qu'elles s'ajustent aux critiques émises, nous avons créé ce tableau permettant de bien différencier les 3 approches. Cette différenciation va être utile lors de la rédaction des items du questionnaire et permettra notamment de clarifier les approches dans le contexte du scénario choisi : la réaction de l'enseignant face à un échec massif des élèves à la suite d'une évaluation sommative.

	But de l'enseignant	Moyen pour atteindre le but	Communication	Principal évaluateur	Sommatif ou formatif	Conséquence sur la suite de l'apprentissage	Objectif pensé pour l'élève
AOL	Juger de la réussite	Les critères et compétences	La cote	Enseignant	Davantage sommatif	Aucune modification	Réussir
AFL	Informé et réguler son enseignement	Les critères et compétences	Feedbacks de l'enseignant	Enseignant	Davantage formative	Modification pédagogique	Apprendre
AAL	Que l'élève, seul ou avec ses pairs, monitore et dirige son apprentissage	Les objectifs personnels et les critères externes	Autocritiques ou critiques par les pairs	Élève(s)	Davantage formatrice ¹	Modification de l'implication/ de la manière de travailler/ de l'acquisition de la matière par l'élève	Développer son autonomie, sa responsabilisation

Tableau 2. Caractéristiques des 3 approches dans ce mémoire

Explication plus détaillée du profil AOL dans ce mémoire :

Dans ce profil, l'enseignant met tout en œuvre pendant l'apprentissage pour que l'élève ait appris tout ce qu'il faut pour réussir l'évaluation. Au moment de l'évaluation sommative, son travail en tant qu'enseignant, dans le sens littéral du terme, est terminé. Il met sa casquette d'évaluateur et reprendra son rôle d'enseignant à partir des prochaines notions à enseigner. Le but principal de l'évaluation et de sa correction, est d'obtenir une cote permettant de juger de la réussite de l'élève. Pour lui, l'élève a également besoin de cette cote qui est l'information la plus claire pour se situer dans sa réussite (Guskey & Link, 2019, cités par Pasquini, 2021). Mettre l'objectif principal sur une décision de passage ou acquisition représentée par une cote,

¹ Formatrice = centrée sur l'activité de l'élève et ayant pour objet principal de « favoriser des processus d'autoévaluation » (Merle, 2018, cité par Fagnant, 2023, p.8)

ajoute, consciemment ou non, une notion normative (Crahay, 2019). Cette notion normative se retrouve également chez les élèves qui vont comparer entre eux les cotes (Crahay, 2009). Cette cote est également considérée par l'enseignant comme une source de motivation pour les élèves (Stiggins, 2017). En effet, en ayant un point de repère si clair, il est plus facile pour les élèves de se fixer des objectifs quantifiables (faire plus, faire autant) ou de créer un esprit de compétition permettant que chacun tire l'autre vers le haut.

Explication plus détaillée du profil AFL dans ce mémoire :

Dans ce profil, l'enseignant met tout en œuvre pendant l'apprentissage pour que l'élève ait acquis le plus de compétences possibles. Au moment de l'évaluation sommative, son travail en tant qu'enseignant, dans le sens littéral du terme, n'est pas terminé. Un enseignant ayant une approche for learning, non seulement utilise des critères pour vérifier si l'élève a atteint les objectifs d'apprentissage mais il va les utiliser également dans ses feedbacks aux élèves pour qu'ils se situent dans leurs apprentissages. En effet, malgré l'utilisation d'une cote, qui peut être présente juste par obligation légale, cette dernière n'est pas suffisante pour l'enseignant dont le but est de faire comprendre à chaque élève ce qui a été acquis de ce qui ne l'est pas. Une évaluation sommative, même de fin de chapitre, reste une évaluation des acquis à un temps t qui peut être utilisée pour aider l'élève à se perfectionner dans ses apprentissages. Cette évaluation peut également servir à l'enseignant pour analyser son propre enseignement et se remettre en question sur la pertinence des pratiques pédagogiques utilisées.

Explication plus détaillée du profil AAL dans ce mémoire :

Dans ce profil, l'enseignant met tout en œuvre pendant l'apprentissage pour que l'élève ait non seulement acquis le plus de compétences possibles mais également pour diminuer sa dépendance à l'enseignant. Avant même l'évaluation de fin de chapitre, il utilisera des pratiques pédagogiques comme l'évaluation par les pairs ou l'auto-évaluation. Lors de la planification de l'évaluation, il veillera à donner à l'élève des prises de décision (exemples : la date de la passation de l'évaluation, une modalité d'évaluation qui respecte les attendus, ...). Au moment de l'évaluation sommative, son travail en tant qu'enseignant de la matière, dans le sens littéral du terme, est presque terminé. Il va maintenant revêtir une casquette de coach ou d'enseignant de compétences transversales. Après avoir corrigé l'évaluation, il va essayer de mettre tout en place pour que l'élève, seul, en groupe, ou en discussion classe-professeur, puisse faire une auto-analyse non seulement des compétences acquises ou non acquises mais également de son investissement, de sa façon d'étudier, des causes des acquis et non acquis. Il aidera les élèves à utiliser toutes ces informations récoltées pour se fixer des objectifs personnels et de classe. Un

enseignant dans le profil AAL va donner plus d'importance à cette auto-analyse dans la suite du cours qu'à l'évaluation elle-même. Par exemple, si un élève n'a pas bien répondu à une question mais qu'il dit qu'il savait, ou sait maintenant, y répondre correctement, ce dernier peut décider ne pas la retravailler. Le contraire existe également : l'élève peut décider de retravailler une compétence, dont l'exercice a été correctement réalisé. De plus, l'enseignant n'hésitera pas à laisser une place importante aux élèves dans l'évaluation en les laissant faire une critique constructive de l'évaluation de fin de chapitre pour vérifier si celle-ci était en adéquation avec l'enseignement et les attentes des élèves avant celle-ci.

2.3. Self-efficacy en évaluation

Un niveau élevé de self-efficacy [SE] des enseignants est considéré comme important pour une carrière d'enseignant réussie et saine (Täschner et al., 2024). Selon Rothinam (2023), pour mettre en œuvre une réforme de l'évaluation en classe, les enseignants doivent avoir une attitude positive en termes de SE et d'AL. Considérant les deux informations précédentes, il est donc important, dans ce contexte de réforme de l'enseignement initié par la FWB via le tronc commun [TC], d'analyser également ce concept de SE. Avant de traiter le SE des enseignants en évaluation, le SE est défini de manière générale, puis spécifiquement aux enseignants.

2.3.1. Définition du self-efficacy

Le SE peut être traduit par « auto-efficacité » mais qui est une traduction, selon nous, apparaissant très peu dans les publications francophones et qui ne résonne pas dans notre culture scientifique au sens que Bandura lui confère. Selon Bandura (1997), le SE est la croyance à produire des effets souhaités à l'aide de ses propres capacités à apprendre, organiser et exécuter des actions. Il continue en précisant que le SE perçu ne concerne pas le nombre de compétences que l'on possède, mais ce que l'on croit pouvoir faire avec ses compétences dans divers contextes. En résumé, analyser le SE d'une personne ne signifie pas déterminer les connaissances et les compétences qu'elle possède mais déterminer sa croyance en ses capacités à les mettre en œuvre.

Ce concept est basé sur le modèle cognitif social, qui postule que le fonctionnement humain résulte des interactions entre des facteurs personnels, comportementaux et environnementaux (Eufemia, 2012). Ce SE est créé et repose sur une ou plusieurs des quatre sources suivantes : expériences de maîtrise, expériences vicariantes, persuasions sociales et états ou réactions physiologiques et émotionnelles (Bandura, 1997).

Plus les individus croient pouvoir influencer positivement l'avenir, plus ils travaillent dur pour atteindre ces objectifs et plus ils ont de chances de les atteindre (Bandura, 1997). Mais le SE n'influence pas que sa propre personne. Selon Luthans et Peterson, les croyances perçues dans le SE d'un leader sont positivement corrélées avec les capacités de performance des employés et la capacité à surmonter les obstacles au changement (2002, cités par Eufemia, 2012). Il est donc naturel d'analyser si, dans la littérature scientifique, le SE des enseignants est également corrélé aux performances et aux comportements des élèves.

2.3.2. Le self-efficacy des enseignants

Skaalvik et Skaalvik (2007, cités par Kyttälä et al., 2023) définissent ce SE comme la croyance individuelle des enseignants en leur capacité à planifier, organiser et mener à bien les activités nécessaires à l'atteinte d'objectifs pédagogiques donnés. Ce SE est donc important vu sa présence avant et pendant chaque acte pédagogique. Tschannen-Moran et Hoy (2001, cités par Rothinam, 2023), eux, ont défini le SE des enseignants comme le jugement d'un enseignant sur sa capacité à obtenir les résultats souhaités en matière d'engagement et d'apprentissage des élèves, même parmi ceux qui peuvent être difficiles ou démotivés. Le modèle cognitif social est bien présent dans cette définition au vu de l'interaction entre des facteurs personnels, ici le jugement de ses propres capacités, et des facteurs comportementaux, ici l'engagement et l'apprentissage des élèves. On peut s'interroger sur l'existence de l'influence exercée par ce jugement ainsi que sur l'ampleur de ses effets.

Le SE des enseignants influence leurs pratiques pédagogiques (Kyttälä et al., 2023) et leur motivation (Levy-Vered & Nasser-Abu Alhija, 2015) qui sont des éléments qui jouent sur l'efficacité en classe. Les enseignants avec un haut SE sont convaincus d'avoir un impact positif sur la réussite des élèves (Levy-Vered & Nasser-Abu Alhija, 2015) et sont plus résilients dans leur pratique et plus enclins à s'investir pour aider tous les élèves à réaliser leur plein potentiel (Pendergast et al., 2011, cités par Levy-Vered & Nasser-Abu Alhija, 2015). Parmi les autres bénéfices d'un haut SE des enseignants, on peut noter une meilleure qualité d'enseignement et un meilleur engagement de la part des élèves (Zee & Koomen, 2016, cités par Täschner et al., 2024), une meilleure relation enseignant-élèves (Holzberger et al., 2014, cités par Täschner et al., 2024), et de meilleurs résultats scolaires (Caprara et al., 2006, cités par Täschner et al., 2024 ; Kim & Seo, 2018, cités par Täschner et al., 2024). Le SE des enseignants est considéré comme un prédicteur important de l'enseignement différencié (Neve et al., 2015, cités par Täschner et al., 2024 ; Suprayogi et al., 2017, cités par Täschner et al., 2024).

Le SE est spécifique aux compétences, aux tâches et au domaine d'enseignement (Bong, 2006, citée par Kyttälä et al., 2023). Le SE dans les pratiques évaluatives peut donc être différent du SE observé dans d'autres domaines de l'enseignement.

2.3.3. Le self-efficacy des enseignants en lien avec l'acte d'évaluer

Dans sa revue systématique sur l'AL, Pastore (2023) conclut que les enseignants ont tendance à se percevoir comme peu confiants dans l'évaluation de l'apprentissage des élèves. Cependant, cette information est à relativiser par le fait, selon Levy-Vered et Nasser-Abu Alhija (2015) et Kyttälä et ses collègues (2023), qu'il existe peu d'études sur le SE en évaluation des enseignants. Dans ces deux articles, ils citent l'expérience de Zhang et Burry-Stock (2003) dans laquelle il a été observé que le SE en évaluation des enseignants est très lié à leurs performances dans leurs pratiques évaluatives en classe, et est lié à leurs connaissances en évaluation, qui sont toutes les deux des parties de l'AL. Les résultats des études respectives de Levy-Vered et Nasser-Abu Alhija et de Kyttälä et ses collègues vont dans le même sens que ceux de Zhang et Burry-Stock. Ces études vont même un peu plus loin dans leurs découvertes. Kyttälä et ses collègues, de leur côté, n'ont pas évalué seulement un SE en évaluation globale mais ont également évalué le SE lié aux conceptions de l'évaluation (sommatrice et formative). Ceci leur a permis d'observer que les conceptions de l'évaluation et le SE en évaluation de leur échantillon sont étroitement liés. Levy-Vered et Nasser-Abu Alhija, quant à eux, ont mis en évidence qu'une partie du SE peut être expliquée par l'AL et la formation à l'évaluation.

Observons le cas des enseignants du premier degré de l'enseignement secondaire de la FWB à l'aide du rapport TALIS 2024 (Quittre et al., 2025). Il y est indiqué que 87% de ces enseignants ont un bon SE en ce qui concerne la rédaction de bonnes questions. Ce pourcentage est en progression par rapport aux résultats de TALIS 2018 (différence significative de 7,7 %) et n'a pas de différence significative avec la moyenne des pays de l'OCDE ayant participé à TALIS 2024. En revanche, seulement 63,4 % de ces enseignants ont un bon SE en ce qui concerne l'utilisation de modalités d'évaluation diverses contre une moyenne de 80% dans l'OCDE. Il serait donc intéressant de sonder les enseignants au sujet des modalités d'évaluation dont celles plus actuelles telles que la différenciation des évaluations, le numérique pour évaluer, ...

2.4. Formation et besoins des enseignants

Les pratiques actuelles tendent au plus grand flou alors que l'évaluation ne s'improvise pas (Jorro, 2009). Elle nécessite une formation. (Jorro, 2009 ; Crahay, 2009). Bien qu'il n'existe pas une approche collective de l'évaluation, les enseignants eux-mêmes aimeraient avoir une harmonisation des pratiques (Jorro, 2009 ; Soussi et al., 2006 ; Braxmeyer et al., 2005) même

s'ils sont peu nombreux à déclarer avoir suivi une formation à l'évaluation (Soussi et al., 2006). D'ailleurs, dans cette dernière étude, ceux ayant suivi ce genre de formations disent qu'ils ont trouvé cela utile, qu'ils utilisent ce qui a été appris et que cela a transformé leurs pratiques (Soussi et al., 2006). L'AL, en ce compris les pratiques évaluatives, est en constante évolution afin de mieux soutenir l'apprentissage des élèves (Willis et al., 2013) et si nous voulons que ce changement de pratiques soit dans la visée du Pacte de lutter contre l'échec scolaire (Crahay, 2019), il est nécessaire de mettre en place des formations spécifiques.

Dans ce mémoire, il ne sera envisagé que les formations professionnelles continues [FPC] étant donné que le public de cette recherche est un public d'enseignants en exercice dans l'enseignement secondaire inférieur de FWB. Nous décidons de nous focaliser sur ce public fort peu sollicité dans les recherches scientifiques actuellement par rapport aux enseignants en formation initiale (Coombs et al., 2018). Dans cette partie, une définition de la FPC est envisagée, suivie d'une définition du développement professionnel. La suite est consacrée aux besoins de formations, aux modalités de formations et aux perspectives de formations futures.

2.4.1. Définition de la formation professionnelle continue

« La formation professionnelle continue des membres de l'équipe éducative des écoles de l'enseignement secondaire s'inscrit dans un processus de développement professionnel de ceux-ci et dans une professionnalisation accrue » (Circulaire 9566, p.593). En FWB, les enseignants ont le droit et le devoir de se former. Il existe 2 types de formations : « celles répondant à des besoins collectifs, lesquelles se caractérisent par une participation obligatoire des bénéficiaires de formation » qui s'organisent pour 1/3 en interréseaux et 2/3 en réseau pour un total de 6 demi-jours par année scolaire et « celles répondant à des besoins personnalisés, lesquelles se caractérisent par une participation facultative et volontaire des bénéficiaires de formation » pour un total de 10 demi-jours par année scolaire (Circulaire 9566, p.596). Ces formations sont le prolongement de la formation initiale qui, elle, vise le développement de la professionnalité, les savoirs et compétences nécessaires à l'entrée dans la profession alors que le long terme oblige à construire des rapports fins à la profession : s'adapter en fonction du contexte tout au long de la carrière (Vanhulle, 2009). Ces FPC ont pour visée un développement professionnel qui est défini ci-après.

2.4.2. Définition du développement professionnel

L'article 6.1.1-2., point 9, du décret portant le livre 6 du Code de l'enseignement relatif à la FPC (FWB, 2021), définit le développement professionnel comme un :

processus graduel d'acquisition et de transformation des compétences et des composantes identitaires conduisant progressivement les individus et les collectivités à améliorer, enrichir et actualiser leur pratique, à agir avec efficacité et efficience dans les différents rôles et responsabilités professionnelles qui leur incombent, à atteindre un nouveau degré de compréhension de leur travail et à s'y sentir à l'aise.

C'est donc un processus qui permet de s'approprier différentes composantes (et invariantes) de son travail en menant des actions réfléchies en autonomie et en investissant ceci de sens (Vanhulle, 2009). Un lien peut être fait avec les définitions de l'AL évoquées supra (DeLuca et al., 2020 ; Pastore, 2023 ; Van Orman, 2024). L'AL (les connaissances, les compétences, les pratiques, ...) évolue constamment pour mieux soutenir les apprentissages des élèves. Un lien peut également être fait avec la figure 2 : les connaissances, compétences et pratiques interagissent et évoluent au regard d'éléments d'influences dont les demandes institutionnelles, l'identité de l'enseignant, les responsabilités morales et éthiques, ... Or, les contextes de travail en FWB sont extrêmement variés et la formation doit répondre selon l'article 6.1.3-10 du décret cité supra (p.12) « au contexte dans lequel le bénéficiaire de formation exerce sa fonction ou sa mission » et comme signalé par Pastore (2023), les parcours de formation devraient inclure des attentes institutionnelles et sociales (normes professionnelles) mais également les besoins de formations des enseignants.

2.4.3. Définition des besoins de formation

Les besoins de formations [BF] sont définis par l'article 6.1.1-2., point 4 du décret cité supra (p.2), comme étant les « écarts entre une situation actuelle, intégrant des apprentissages déjà réalisés, et une situation souhaitée ou souhaitable, qu'une formation est susceptible de diminuer. L'analyse des écarts va déterminer des objectifs de formation ». Donc, une évaluation initiale des compétences des enseignants visant à identifier les priorités d'amélioration est nécessaire pour adapter le contenu de la formation aux besoins professionnels des enseignants et, d'autre part, la formation est plus efficace quand elle correspond aux besoins professionnels des enseignants (Christoforidou & Kyriakides, 2021). En outre, une formation à l'évaluation qui ne partirait pas des préoccupations de ces derniers ne permettrait pas des processus de changement de façon durable (Jorro, 2016). Partir des préoccupations n'est pas suffisant. Selon Kennedy (2016) l'effet du développement professionnel dépend fortement de la motivation et des dispositions des professeurs à modifier leurs pratiques. Les enseignants sont sortis de leurs salles de classe pour se former en espérant qu'ils mettent en pratique ce qu'ils ont appris une fois revenus dans leur école. Cependant, les enseignants en exercice ont déjà développé un système de pratiques qui, selon eux, répond de manière optimale aux différents défis à relever. Il leur reste alors à décider si, quand et comment intégrer les nouvelles idées dans leurs pratiques (Kennedy, 2016).

Les enseignants ayant pris part à la recherche de Braxmeyer et ses collaborateurs (2005) ont même déjà dégagé plusieurs aspects de formations qui les intéressent : être formés à l'évaluation orale et au savoir-être, être formés sur les méthodes d'évaluations, avoir des références nationales pour toutes les matières et disciplines, avoir du temps pour tirer profit des évaluations, se concerter avec les collègues, échanger sur les pratiques. Il est cependant vrai que les enseignants sont en tension entre les demandes et les attentes (en contradiction) des uns et des autres (Vanhulle, 2009 ; Crahay et al., 2019). Par exemple, les enseignants sont tenaillés entre les recommandations pédagogiques, les contraintes du système et les attentes des parents. Donc, un certain nombre d'enseignants en fin d'année mélangent les fonctions des évaluations et tentent de pronostiquer dans quelle mesure un élève peut suivre les enseignements de l'année à venir, ce qui est une dérive importante du système en place (Crahay et al., 2019) surtout quand on sait que les enseignants ne sont pas à l'aise avec la notation des évaluations (Pastore, 2023) qui mènent in fine à ces pronostiques. En effet, les enseignants redoutent la pratique évaluative selon Jorro (2016) car ils sont obligés de la mettre en œuvre alors qu'elle est apprise sur le tas et qu'elle nécessite des savoirs précis, des compétences et gestes professionnels adaptés centrés d'une part sur la mesure et d'autre part sur l'interprétation des informations.

Un référentiel de savoirs et compétences n'existe pas dans le domaine de l'évaluation (car il y a différents contextes en FWB et différentes pressions exercées sur les enseignants) (Vanhulle, 2009). La formation devrait donc, selon DeLuca et ses collègues (2021), mettre le focus sur la dimension culturelle de l'évaluation (c'est-à-dire les différentes façons d'évaluer à travers le monde) et préparer les professeurs à jongler entre compréhension locale et globale de l'évaluation (c'est-à-dire la micro-culture de la classe et la macro-culture du pays). Ceci devrait permettre de donner aux enseignants de nouvelles perspectives en matière d'évaluations.

2.4.4. Modalités de formations

Plusieurs modalités de formations sont mentionnées dans divers textes (décret, littérature scientifique, ...). Par exemple, les professeurs ont besoin d'être accompagnés : l'autonomie et la pensée créative se développent en connaissant les contraintes (Vanhulle, 2009) et ce besoin d'accompagnement doit se penser en dispositif meso (école) et doit s'appliquer entre « amis critiques » (Crahay, 2009, pp. 250-251). En effet, comme dit supra, les formations doivent répondre aux besoins des professionnels (Jorro, 2016 ; Christoforidou & Kyriakides, 2021). Le décret cité supra fait mention d'un portfolio de développement professionnel qui est à compléter tout au long de la carrière et qui peut servir de base à une réflexion sur ses pratiques (FWB, 2021). Ce portfolio favorise la connaissance de soi par l'autoévaluation et renforce l'estime de

soi par la reconnaissance des forces et faiblesses (Mottier Lopez & Van Nieuwenhoven, 2009). Dorénavant, il y a également la possibilité de rejoindre une communauté de pratique ou communauté d'apprentissage professionnelle (FWB, s.d.). Dans cette modalité, tous les participants s'engagent dans un processus de négociation de sens des actions par rapport à un projet commun partagé (Mottier Lopez & Van Nieuwenhoven, 2009). À notre connaissance, ces deux pratiques, que sont le portfolio et les communautés de pratique, n'ont pas été assez déployées pour être connus par la majorité des enseignants de la FWB. Il y a aussi un intérêt à concevoir les formations en modules ou à être en relation avec un expert, un pair expert (ou plusieurs) (Jorro, 2009). Toutes ces modalités permettent un apprentissage actif, reconnu comme pratique efficace en formation par Blank et de las Alas (2009, cités par Kennedy, 2016). En effet, les enseignants doivent être actifs pour apprendre, comme les élèves doivent l'être : les formations à messages prescriptifs ne fonctionnent pas (Kennedy 2016).

D'autre part, il est nécessaire de penser les formations sur la durée pour qu'il y ait un effet durable (Allal & Dauvisis, 2019). De fait, le développement de la profession demande du temps. C'est un processus d'apprentissage complexe et itératif influencé par le parcours de formation initiale et continue et différentes variables (durée, connaissance du contenu, modèles de formation et opportunités) ainsi que l'expérience d'enseignement continue et des facteurs contextuels (collaboration, participation scolaire et normes professionnelles) (Pastore, 2023).

Selon les enseignants ayant participé à la recherche de Soussi et de ses collègues (2006), les deux formations les plus voulues seraient d'une part une formation continue sur mesure c'est-à-dire liée à un projet d'école ou un groupe-discipline et d'autre part, une collaboration entre collègues. DeLuca et ses collaborateurs (2016) ajoutent que les formations en présentiel sont préférées à celles en ligne et que les enseignants moins expérimentés préfèrent l'apprentissage individuel. Jonson (2008, cité par DeLuca et al., 2016) exprime le fait que le mentorat est préconisé pour le début de carrière.

2.4.5. Perspectives pour les formations futures

Même si les formations étaient adaptées aux volontés des enseignants, pour que développement professionnel il y ait, il faut que les apprentissages professionnels soient vus comme accessibles et utiles donc que ce soit un sujet d'intéressement professionnel (Jorro, 2009). Il faut donc créer une problématisation commune et un espace d'intéressement entre praticiens politiques et praticiens enseignants sans passer de la recherche scientifique à l'établissement de réformes directement car sinon, les enseignants ne mettent pas en place des pratiques qui changent profondément (Athias et al., 2020). En effet, par exemple, les réformes des années 90 ont été

prises en place « avec peu d'accompagnement sur le terrain et sans augmentation de ressources ou de moyens » ce qui a creusé « le fossé entre les décideurs politiques et les acteurs de terrain » et qui a, in fine, eu peu de retombées dans les classes (Lafontaine et al., 2009, p.65). D'ailleurs, les FPC obligatoires ne sont pas voulues par les enseignants (Soussi et al., 2006). En bref, modification de pratiques il y a si on entrevoit un gain concret et immédiat (Jorro, 2009) et, pour changer un habitus évaluatif, cela passe par une formation (Crahay, 2009).

Puisqu'une des visées du Pacte pour un enseignement d'excellence est de lutter contre l'échec scolaire et le redoublement (FWB, 2025), et que Crahay (2019) dit que ceci passe obligatoirement par un changement de pratiques d'évaluation des enseignants, il devient primordial de se pencher sur les pratiques des enseignants dans le domaine et sur la formation qui y est associée. Sans formation, les enseignants reprennent leurs propres expériences passées au sein du système éducatif et prennent alors des décisions importantes pour les élèves en ayant soit une compréhension incomplète, soit une mauvaise compréhension de ce que les élèves peuvent faire (Shavelson, 2006, cité par Van Orman et al., 2024). Il faut absolument travailler à une compréhension plus nuancée de ce que font les enseignants, ce qui les motive, la manière dont ils apprennent et se développent (Kennedy, 2016).

Vu les gains potentiels liés aux formations des enseignants, dans le domaine de l'évaluation entre autres, sur le site du Pacte pour un enseignement d'excellence, il est indiqué que la FWB a décidé d'augmenter le budget lié aux formations (de 10 à 27 millions d'euros) et que d'autres modalités de formations sont dorénavant accessibles aux enseignants (Communautés d'apprentissage professionnelles et formation en immersion, entre autres).

3. Question de recherche et objectifs

Un premier constat est qu'il n'existe pas, d'après nos lectures, d'article reprenant les pratiques évaluatives des enseignants de l'enseignement secondaire inférieur en FWB. À l'exception des résultats des enquêtes TALIS, nous n'avons trouvé que très peu d'informations sur les manières d'évaluer de nos enseignants et sur les approches d'évaluation mises en œuvre. Ensuite, un deuxième constat : bien que certains aient pointé des manques de littérature concernant l'AL, le SE et les BF des enseignants en évaluation et les modalités préférées de formations, des recherches ont tout de même été effectuées sur ces 4 variables. Immédiatement, nous pensons à l'outil ACAI (DeLuca et al., 2016) qui aurait été un outil idéal à utiliser en l'état ne serait-ce que pour obtenir des résultats pour ces 4 variables sur un public sur lequel ce questionnaire n'avait pas encore été testé, à savoir les enseignants du secondaire inférieur de la FWB.

Cependant, les chiffres venant des analyses statistiques de fiabilité d'ACAI n'ont jamais atteint les seuils requis (Baron & Fagnant, 2025). Les deux raisons principales évoquées sont les trop nombreuses notions de l'AL visées et la multitude de scénarios utilisés. En ce qui concerne la première raison, dans la partie théorique, une seule notion de l'AL a été ciblée : les approches ou, en d'autres mots, les buts visés par les enseignants.

Dans les recherches sur le SE, nous nous sommes rendu compte que cibler le SE en évaluation était un peu réducteur. En effet, Kyttälä et ses collègues (2023), eux, ont analysé plusieurs types de SE en évaluation car le SE est spécifique aux compétences, aux tâches et au domaine d'enseignement (Bong, 2006, citée par Kyttälä et al., 2023). Quant aux BF en évaluation, aucune typologie sur le sujet n'a été créée à notre connaissance.

De là est née notre idée : au lieu de voir s'il n'y a un lien qu'entre approches, SE en évaluation et BF en évaluation, ne serait-il pas également intéressant d'investiguer un lien possible entre l'approche AOL, le SE dans des actions liées à l'approche AOL et le BF dans des actions liées à l'approche AOL ? La même question se pose pour les deux autres types d'approche.

Donc, pour conclure cette partie, malgré nos lectures de nombreux articles sur les pratiques évaluatives (Soussi et al., 2006, Braxmeyer et al., 2005, ...), sur le SE (Bandura, 1997 ; Levy-Vered & Nasser-Abu Alhija, 2015 ; Kyttälä et al., 2023 ; ...) et sur les FPC (Kennedy, 2016 ; Fagnant & Goffin, 2021 ; Jorro, 2016 ; ...), nous en sommes arrivés à la conclusion qu'aucun article, à notre connaissance, ne fait état des pratiques évaluatives des enseignants de FWB, des approches des enseignants en la matière et des liens qui pourraient lier ces dernières avec le SE des enseignants ; ou leurs BF ou leurs modalités préférées de formations.

Nous en venons donc au besoin d'étudier cela au travers de cette question de recherche :

Quelles sont les approches mises en place par les enseignants du secondaire inférieur de la Fédération Wallonie-Bruxelles en matière d'évaluation, leur sentiment d'efficacité à cet égard ainsi que leurs besoins de formations et leurs modalités préférées de formations, et quels liens peut-on établir entre ces dimensions ?

Cette question nous mène à investiguer les 5 actions suivantes :

1. Identifier, parmi les répondants, des profils d'évaluateurs ;
2. Constater et questionner dans quelle mesure un lien existe entre le profil d'évaluateur et le sentiment d'efficacité lié à l'évaluation ;

3. Constater et questionner dans quelle mesure un lien existe entre le profil d'évaluateur et les besoins de formations liées à l'évaluation ;
4. Constater et questionner dans quelle mesure un lien existe entre le profil d'évaluateur et les modalités préférées de formations ;
5. Identifier les liens existants entre profil, sentiment d'efficacité et besoins de formations.

4. Méthodologie

4.1. Introduction

Nous privilégions une recherche par questionnaire, créé sur base de différentes théories issues de la littérature scientifique. Le but étant d'investiguer des liens possibles entre différentes variables, un questionnaire sondant un certain nombre d'enseignants est le moyen le plus approprié d'y parvenir selon nous. En effet, Demonty et ses collègues (2024) conseillent d'utiliser ce type de recherche pour une récolte d'informations spécifiques, mesurables, dont le sujet n'est pas trop sensible et qui permet de comparer les réponses de différentes catégories de personnes.

Ce questionnaire est mis à l'épreuve (Demonty et al., 2024) lors de 7 laboratoires cognitifs (voir annexe IV). Il est modifié en fonction de chaque laboratoire cognitif pour devenir une version toujours plus compréhensible pour les participants. Enfin, la version finale de ce questionnaire (voir annexe III) est diffusée à large échelle via les réseaux sociaux. Les résultats sont analysés statistiquement et des interprétations sont énoncées en lien avec le contexte/les 5 actions de la présente recherche. Chaque partie de cette recherche est détaillée ci-dessous.

4.2. Création du questionnaire et analyses statistiques envisagées

L'idée de ce travail n'est pas de savoir à quel point les enseignants sont compétents en évaluation mais bien de savoir quelles actions ils disent pratiquer en classe. En effet, les données produites à l'aide de ce questionnaire sont des données auto-rapportées. Ce questionnaire a pour but de discriminer les enseignants du secondaire inférieur sur 4 points : leurs approches en évaluation, leur SE en évaluation, leurs BF en évaluation et leurs modalités préférées de formations à l'évaluation.

La première des 5 actions principales citées au point 3 de ce travail se réalise à l'aide de la partie 1 du questionnaire, les actions 2 et 3 à l'aide la partie 2 du questionnaire, la quatrième action à l'aide de la partie 3 et enfin, la cinquième action se réalise lors de l'analyse des résultats et lors de la discussion finale regroupant toutes les informations récoltées dans les différentes

parties. Enfin, une partie 4 du questionnaire est présente pour obtenir des informations générales (âge, sexe, qualification, ...) qui enrichissent l'analyse des parties 1, 2 et 3.

4.2.1. Partie 1 : Identifier, parmi les répondants, des profils d'évaluateurs.

Dans la partie 2.2.4., trois approches (AOL, AFL et AAL) par rapport à l'évaluation ont été définies en s'appuyant sur la littérature scientifique. Dans cette partie 1, ces 3 approches sont questionnées. Cette partie 1 commence par un scénario qui est un des scénarios utilisés par DeLuca et ses collaborateurs dans son questionnaire de 2016. Le choix de ne se concentrer que sur un seul scénario peut paraître étrange vu l'étendue des activités liées à l'évaluation. Cependant, les actions et les réflexions des professeurs sont fortement dépendantes du contexte dans lequel ils sont. Les derniers résultats de chercheurs voulant utiliser différents scénarios pour déterminer des pratiques évaluatives liées à 3 profils évaluateurs ont montré une impossibilité d'obtenir des indices satisfaisants de consistance interne (Baron & Fagnant, 2025). Nous avons donc choisi, parmi les différents scénarios², le premier scénario dans lequel l'identification par le plus grand nombre de répondants à la situation serait la plus probable : la réaction face à un échec massif lors d'une interrogation sommative. C'est un scénario qui est susceptible d'être rencontré par une majorité d'enseignants car il n'utilise pas des éléments qui ne concernent pas les enseignants de toutes les matières comme les tests standardisés du scénario 5. L'évaluation sommative, elle, est présente partout vu qu'elle est le type d'évaluation que les établissements imposent (Pasquini & Morales Villabona, 2022). En effet, « l'évaluation scolaire se réduit le plus souvent à passer une épreuve aboutissant à une note » (OCDE, 2012, cité par Pasquini, 2021, p. 4) et dans le contexte de la FWB où le redoublement est bien plus présent que dans d'autres pays (Lafontaine et al., 2019), les échecs ne sont pas des événements exceptionnels.

Ce scénario est suivi de 7 thèmes de réaction : Nouvelle évaluation ou pas (Soussi et al., 2006 ; FWB, 2025) ; Remédiation en classe/différenciation (Soussi et al., 2006 ; FWB, 2025) ; Investissement des élèves (Stiggins, 2017 ; DeLuca et al., 2021 ; Quittre et al., 2025) ; Remise en question de la validité de l'évaluation (Raulin & Lebeaume, 2019 ; Detroz et al., 2020 ; Detroz, 2023) ; Discussion avec les pairs (Soussi et al., 2006 ; Crahay, 2009 ; Campanalle et al., 2010 ; Wyatt-Smith et al., 2010 ; Quittre et al., 2025) ; Feedback (Soussi et al., 2006 ; Hattie

² Les autres scénarios sont, avec la numérotation de l'outil ACAI :

2 : Réaction face au plagiat d'un élève ; 3 : Traitement des élèves à besoins spécifiques ; 4 : Priorité dans une planification ; 5 : Réponse à un parent par rapport à un test standardisé (DeLuca et al., 2016).

& Timperley, 2007 ; Detroz et al., 2020 ; FWB, 2025) ; Rôle de l'évaluation (Raulin & Lebeaume, 2019 ; FWB, 2025).

Chaque thème possède de la littérature sous-jacente avec laquelle nous pourrions faire des liens lors de notre analyse. Une justification détaillée du choix de chaque thème se trouve en annexe V.a. Le choix du nombre de thèmes est dû à deux facteurs : le besoin de validité et la dépense énergétique des répondants. Pour obtenir un score valable, il est conseillé d'avoir au minimum 6 items par notion (Demonty et al., 2024). Au départ, 11 thèmes avaient été créés car, plus il existe d'items, plus il est facile d'avoir une validité forte (Laveault & Grégoire, 2014). Cependant, les laboratoires cognitifs nous ont appris que cette partie était très énergivore par rapport aux autres parties (voir remarques générales des annexes IV.a. et IV.b.). Certains thèmes ont donc été supprimés. Pour certains, la cause est qu'ils étaient moins liés, dans la littérature, aux différentes approches ce qui aurait rendu plus difficile l'analyse. D'autres thèmes semblaient être répétitifs. Cette partie a donc finalement été réduite à 7 thèmes.

Chaque thème est composé de 3 actions (items) possibles, une par approche que nous voulons étudier. Chacun de ces items a été créé en se référant aux définitions et aux caractéristiques des approches. L'analyse du lien entre item et approche se trouve dans l'annexe V.b. La validité de ce lien est également vérifiée au point 5.2.1. de ce travail par des analyses statistiques après réception des questionnaires à l'aide d'indices tels que l'alpha de Cronbach, l'oméga de McDonald, une analyse factorielle confirmatoire [AFC] et l'analyse de l'ajustement du modèle (Navarro et al., 2020 ; Demonty et al., 2024). L'alpha de Cronbach permet de savoir si les items mesurent bien la même chose, un seul et même construct, via une corrélation moyenne entre les items. Il permet de valider la cohérence interne de l'échelle. L'oméga de McDonald poursuit les mêmes objectifs mais est plus robuste avec de petits échantillons. Quand un modèle théorique indique la façon dont les variables sont reliées les unes aux autres, l'utilisation d'une AFC est à utiliser pour confirmer cette « structure factorielle latente préétablie » (Navarro et al., 2020, p.453). L'ajustement du modèle qui suit l'AFC « nous indique si notre modèle correspond bien aux données observées » (Navarro et al., 2020, p.458).

Le format choisi pour ces items est celui de l'échelle de Likert. Ce choix s'appuie sur plusieurs avantages qu'apporte ce type de format. La mesure est plus précise grâce à la théorie classique des scores (Laveault & Grégoire, 2014). Par rapport aux questions ouvertes, le temps de passation est bien plus court et permet de couvrir le plus largement possible le domaine visé. Enfin, dans ce format en ligne, il y a une standardisation des modalités de passation qui produisent des données brutes. Tout ceci augmente la fidélité par rapport à des questions

ouvertes grâce au peu de place laissé à l'interprétation des réponses (Laveault & Grégoire, 2014). Ceci ne signifie pas que les questionnaires sont 100% objectifs vu l'existence de biais qui menacent la validité et la comparabilité selon les contextes : acquiescement, choix des extrêmes, tendance centrale, désirabilité sociale, comportements erratiques, ou satisfaisants (Demonty et al., 2024). Ce choix d'ajouter une échelle de Likert pour chaque item au lieu d'imposer des items à choix forcé est en lien avec l'objectif de ce questionnaire. Nous voulons obtenir des profils pour lesquels le score n'est pas influencé par une obligation de sélectionner une approche plutôt qu'une autre. Les items à choix forcé sont très intéressants quand on veut faire une distinction nette entre les approches. Or, les différentes utilisations précédentes de l'outil ACAI montrent qu'il n'existe pas des enseignants 100% AOL ou AFL ou AAL (Baron & Fagnant, 2025). Ici, nous cherchons donc à estimer dans quelle mesure les approches associées à chaque profil sont utilisées.

Chaque item est accompagné d'une échelle de Likert à 6 échelons (1 : Tout à fait improbable ; 2 : Plutôt improbable ; 3 : Légèrement improbable ; 4 : Légèrement probable ; 5 : Plutôt probable ; 6 : Tout à fait probable). Chaque échelle de Likert est ordonnée du plus négatif par rapport à l'item au plus positif (ordre ascendant) pour ne pas perturber les participants sur les modalités de réponses (Demonty et al., 2024). Bien que Likert lui-même conseillait 5 échelons (Laveault & Grégoire, 2014), nous avons choisi de ne pas mettre d'élément neutre, considérant que chaque enseignant possède un avis à donner sur chacune des pratiques évaluatives énoncées (Demonty et al., 2024). Six échelons ont été choisis pour avoir une meilleure possibilité de nuances, pour faciliter la discrimination entre les participants, sans pour autant proposer trop de choix, ce qui empêcherait le participant de se placer judicieusement. Avec 6 échelons, un participant sur 7 aux laboratoires cognitifs a déjà mentionné sa difficulté à se placer sur l'échelle (voir échelle de la partie 1 de l'annexe IV.f.). Cependant, en-dessous de 5 échelons et au-dessus de 9 échelons, la fidélité de la mesure diminue (Monseur, 2023). Le choix de 6 échelons s'est donc imposé à nous. Une dernière raison de ce choix est d'enlever le biais de tendance centrale (Monseur, 2023).

Les 6 échelons sont codés de 1 à 6. Les participants, en se plaçant sur chaque échelle de Likert, obtiennent un score dans chaque approche qui est la moyenne des codes relatifs à l'approche en question. Cette moyenne peut sembler contraire aux règles statistiques vu que l'échelle avec laquelle nous travaillons se rapproche de la définition d'une échelle ordinale (Navarro et al., 2020 ; Monseur, 2022). Nous pouvons considérer que nous sommes dans une échelle quasi-intervalle car les participants se comportent comme si les écarts entre les différents échelons

étaient identiques (Navarro et al., 2020). Le calcul de scores est donc possible. Ceci est confirmé par la réflexion d'un de nos participants au laboratoire cognitif qui, se posant la question sur l'intitulé des échelons, a éliminé sa propre question en disant qu'il comprenait qu'il devait se placer sur un continuum allant de 1 à 6 (voir échelle de la partie 1 de l'annexe IV.f.). Au regard de l'ensemble de ces 3 scores, à l'aide de la méthode k-means (Tchibozo, 2019), des clusters d'enseignants sont créés. En agissant de la sorte, deux, trois ou quatre clusters sont attendus selon le meilleur groupement par rapport aux données obtenues. La méthode k-means va permettre d'obtenir des groupements qui vont se distinguer sur au moins une approche. Par exemple, il est possible d'avoir un groupement de personnes très AOL (avec une moyenne très haute sur les items AOL) et très AFL mais pas du tout AAL. Tous ces groupements proposés par le logiciel sont nos profils.

En plus de créer des profils, cette partie 1 du questionnaire permet d'investiguer les approches en matière d'évaluation mises en place par les enseignants ayant participé à la recherche. Le fait d'avoir créé plusieurs items pour une seule approche permet, dans l'analyse, de mettre en évidence les actions qui sont les plus utilisées au sein d'une même approche. Par exemple, TALIS 2024 montre que 25% des enseignants du premier degré utilise l'auto-évaluation (Quittre et al., 2025). Or, il existe d'autres méthodes que l'auto-évaluation pour mettre l'élève au centre de l'évaluation. Nous voulons donc approfondir la compréhension de ces résultats avec d'autres conceptions qui se rapportent à l'approche AAL.

4.2.2. Partie 2 : Constater et questionner dans quelle mesure un lien existe entre le profil d'évaluateur et d'une part, le sentiment d'efficacité lié à l'évaluation et, d'autre part les besoins de formations liées à l'évaluation

Dans cette partie, il n'y a plus de scénario car l'objectif est de déterminer le SE et les BF de manière plus générale. Il est donc demandé de répondre par rapport à la pratique quotidienne. Certains raisonnements effectués pour créer la partie 1 sont également de mise ici (se placer sur une échelle de Likert à 6 échelons, sans choix forcé, sans valeur centrale, ...). Nous avons créé également les items de la partie 2. Chaque item est accompagné de deux échelles de Likert : une allant de « pas du tout capable » à « entièrement capable » pour répondre à la question « à quel point vous vous sentez capable d'effectuer chaque action listée ? », l'autre allant de « pas du tout » à « tout à fait » pour répondre à la question « ressentez-vous le besoin de vous former sur celles-ci ? ».

L'idée de mettre 2 échelles pour une même action est liée à 3 raisons. Une première est que les actions abordées pour le SE sont les mêmes que pour les BF. En effet, dans la partie théorique,

il a été énoncé qu'une partie de la variance du SE peut être expliquée par l'AL et la formation en évaluation (Levy-Vered & Nasser-Abu Alhija, 2015). Nous voulons donc absolument lier les 2 notions dans notre recherche. Une deuxième raison est que cela facilite les réponses et économise les réflexions du participant par rapport à un questionnaire où les questions auraient été séparées en 2 parties. Cette deuxième raison a été confirmée par la majorité de nos laboratoires cognitifs. Enfin, placer les 2 échelles l'une à côté de l'autre induit un lien entre les 2 réponses, ce qui est souhaité dans notre questionnaire.

Pour créer ces items nous nous sommes inspirés en partie du questionnaire de Kyttälä et ses collaborateurs (2023), ayant pour but de déterminer 4 types de SE : pour l'évaluation sommative, pour l'évaluation formative, l'auto-efficacité globale en évaluation, et l'évitement de l'évaluation. Notre questionnaire investiguant les différentes approches, il serait aussi utile, selon nous, de distinguer également les divers SE et BF reliés à chacune de ces approches. C'est pourquoi nous avons décidé de transformer ce questionnaire pour avoir 4 dimensions : 3 liées aux actions relevant des approches AOL, AFL et AAL et une quatrième dimension générale. Nous nous sommes donc inspirés du SE lié à l'évaluation sommative pour notre SE AOL, du SE lié à l'évaluation formative pour notre SE AFL et de l'auto-efficacité globale pour notre SE général. Vu la présence de la dimension AAL dans les approches, il nous a semblé logique de mesurer un SE AAL. La décision d'enlever le SE lié à l'évitement de l'évaluation est dû au fait qu'en FWB, comme dans d'autres systèmes éducatifs, il est demandé aux enseignants de justifier leurs décisions quant à l'apprentissage des élèves. En effet, depuis longtemps la note représente l'information la plus concrète que les élèves puissent obtenir sur leurs apprentissages (Guskey & Link, 2019, cités par Pasquini, 2021). Il y a d'ailleurs une augmentation des contraintes légales encadrant la production et l'usage des notes (Pasquini, 2021).

La validité des dimensions SE et BF est vérifiée aux points 5.3.1. et 5.3.3 de ce travail par des analyses statistiques après réception des questionnaires à l'aide d'indices tels que l'alpha de Cronbach, l'oméga de McDonald, une analyse factorielle confirmatoire et l'analyse de l'ajustement du modèle (Navarro et al., 2020 ; Demonty et al., 2024). Ces analyses statistiques sont effectuées sur les items ne traitant que du SE, et aussi sur les items ne traitant que du BF. Ces analyses sont ensuite effectuées sur toutes les sous-parties de la partie 2 (SE général, BF général, SE AOL, BF AOL, ...). Après ces analyses, un score peut être calculé pour chacune d'elles. Ce score est une moyenne des codes obtenus dans les items correspondants.

La première paire d'items « J'ai confiance en mes compétences d'évaluateur » et « J'ai besoin de formations pour développer mes compétences d'évaluateur » a été placée avant toutes les

autres paires d'items pour que la lecture de toutes les étapes du cycle MoSCodEE n'ait pas d'influence sur une sensation globale que nous voulions la plus intuitive possible. Deux moyennes par cluster, appelées « SE global » et « BF global », sont calculées puis comparées.

Les 12 paires d'items suivantes ont pour but de déterminer le SE général en évaluation et les BF généraux. L'évaluation est un vaste concept. Le cycle MoSCodEE, explicité au point 2.1. de ce travail, reprenant les 9 étapes pour mettre en place une évaluation valide et fiable, nous a semblé balayer tous les stades utiles pour l'évaluation. Ceci permet d'investiguer de façon assez complète le SE et le BF à chaque étape du cycle. Par rapport aux clusters créés à la partie 1, peu importe le cluster auquel le participant appartient finalement, ces étapes font partie intégrante d'une évaluation pensée et donc nous ne considérons pas qu'il faille les relier à une approche plutôt qu'une autre. La séparation des paires d'items 2 à 13 par rapport aux différentes étapes du cycle est reprise à l'annexe III.b. Pour structurer les items, un codage a été utilisé. Par exemple : MOSCO SE1P1 se rapporte à la première partie de la première étape du cycle MoSCodEE pour laquelle on investigate le SE des enseignants. Les scores « SE général » et « BF général » sont calculés à l'aide de ces 12 paires d'items. Ces scores sont ensuite comparés entre les clusters. Par la suite, il est possible d'investiguer les étapes du cycle dans lesquelles les clusters sont les plus confiants en évaluation et celles dans lesquelles ils ressentent le plus de BF. Pour ce faire, comme pour la première paire d'items, les moyennes des SE et des BF sont calculées pour chaque étape du cycle. Nous comparons ces moyennes entre clusters. Toutes ces analyses ont pour but de constater si un lien existe entre approche et SE, et si un lien existe entre approche et BF.

Ensuite, 18 items ont été créés reprenant les pratiques liées aux approches AOL, AFL et AAL. Six items ont été créés par approche car c'est le minimum conseillé pour pouvoir faire des traitements statistiques valides et fiables (Demonty et al., 2024). La séparation des paires d'items 14 à 31 par rapport aux différentes pratiques et les approches associées est indiquée dans le codage de l'annexe III.b. C'est à partir de cette séparation d'items que sont calculés les scores (moyennes) : un score par approche pour le SE et un score par approche pour les BF. À l'aide de ces scores, 2 types de comparaisons sont effectuées. Premièrement, à l'intérieur d'un même type de SE ou BF (par exemple SE AAL), les scores entre les clusters sont comparés. Le but est de constater par exemple si le cluster dans lequel on déclare pratiquer le plus l'AAL est également le cluster qui possède un plus grand SE dans les actions liées à l'AAL. Deuxièmement, à l'intérieur d'un même cluster, les différents scores obtenus sont comparés pour chaque type de SE et BF. Le but est de constater par exemple si le cluster dans lequel on

déclare pratiquer le plus l'AAL a un score plus haut en SE AAL qu'en SE AOL. À l'intérieur de ces clusters, nous pouvons également constater les actions (items) dans lesquelles les clusters sont les plus confiants et ceux dans lesquelles les clusters ressentent le plus de BF. Nous pourrions alors tirer une conclusion par rapport au versant AOL, AFL ou AAL avec lequel est reliée l'action investiguée.

Les 8 paires d'items suivants (32 à 39) sont issues de nos lectures. En effet, nous avons lu un certain nombre d'articles, d'ouvrages et de rapports dans lesquels les mêmes thèmes revenaient : la collaboration avec les collègues, la différenciation, l'utilisation du numérique, l'hétérogénéité des classes, questionner ses choix de pratiques et utiliser les résultats des épreuves externes. Le détail des liens entre chaque item et la littérature scientifique est à l'annexe V.c. Un de nos buts secondaire étant d'enrichir la compréhension des pratiques évaluatives des enseignants, ces 8 dernières paires d'items nous aident dans ce processus. Le dernier item « J'ai confiance en mes capacités à utiliser ce que j'ai appris en formation (initiale ou continue) pour améliorer la qualité de mes pratiques évaluatives » permet d'interroger le lien entre les clusters et le SE en application de ce qui a été appris en formation. Pour ces derniers items, une moyenne est calculée par item par cluster. Les clusters sont comparés par rapport à leurs moyennes en SE sur un même item. Il en va de même pour les moyennes en BF.

Pour constater si un lien existe entre SE et BF, une corrélation est calculée pour chaque paire d'items. Une corrélation négative est attendue : si un participant se sent très capable dans une action, on peut s'attendre à ce que celui-ci ne ressente pas le besoin de s'y former. Nous nous attardons ensuite sur les corrélations « surprenantes » : celles qui seraient proches de 0 ou de 1. Une corrélation proche de 0 indique qu'il n'y a pas de lien entre les deux variables. Une corrélation proche de 1 pourrait indiquer des paires d'items avec des SE et BF tous les deux élevés, ce qui impliquerait qu'on y apporte beaucoup d'importance ou pourrait indiquer des paires d'items avec des SE et BF tous les deux bas, ce qui impliquerait qu'on y apporte très peu d'importance. Ces paires d'items dont les corrélations sont surprenantes sont analysées en observant les tableaux de fréquences selon les clusters pour constater s'il y a une différence.

4.2.3. Partie 3 : Constater les préférences de modalités de formations des clusters d'enseignants et les liens que nous pourrions établir avec les différentes approches.

Cette partie du questionnaire est composée de 14 items qui sont repris du questionnaire de DeLuca et ses collaborateurs (2016) et traduite par nos soins. Elle a été modifiée selon les laboratoires cognitifs effectués (besoins de clarification de certaines notions). Elle permet de

questionner le lien entre les clusters créés à la partie 1 et les préférences de formations : individuelles ou en groupe ; en distanciel ou en présentiel ; avec ses propres collègues/ son propre contexte ou avec des collègues de tous horizons. Pour chaque item et chaque cluster une moyenne est calculée. Nous comparons les clusters par rapport à ces moyennes. Cette comparaison permet de répondre à des questions comme : Quelle modalité de formation est la plus voulue et la moins voulue par chaque cluster ? Est-ce une modalité en ligne ou en présentiel ? En groupe ou en individuel ? ...

4.2.4. Partie 4 : informations démographiques des participants permettant de décrire l'échantillon

La dernière partie du questionnaire est composée de 10 items permettant de décrire l'échantillon mais également de décrire les clusters. Elle sert également à mettre en lien les informations données dans cette partie (âge, sexe, années d'expérience, filière d'enseignement majoritaire, ...) avec les résultats issus de la littérature constatés dans nos lectures. Par exemple, dans l'objectif de faire un état des lieux des différents SE, nous pouvons vérifier si le cluster le plus AFL est plus féminin que les autres clusters. Cette constatation peut être mise en parallèle avec l'étude de Coombs et ses collaborateurs (2022), selon laquelle les femmes auraient un meilleur SE en évaluation formative que les hommes.

Cette partie, parfois présente dès le début des questionnaires, a été placée à la fin pour diminuer au maximum les menaces de stéréotypes liés à l'âge, ancienneté, sexe, ... sur l'ensemble des questions précédentes. En effet, la menace du stéréotype est due aux croyances que le participant détient quant au regard que les autres portent sur son groupe et pousse ce dernier à agir différemment pour ne pas confirmer cette croyance (Fresson & Dardenne, 2019). Nous ne voulions pas que des stéréotypes comme « les plus vieux sont moins bons avec le numérique », ou « les femmes sont plus maternantes que les hommes » ou encore « les jeunes enseignants sont moins légitimes que les enseignants avec plus d'expérience au sujet de l'évaluation des étudiants » influencent les réponses de nos répondants. Qui plus est, Monseur (2023) nous dit qu'il faut commencer un questionnaire avec des questions simples mais qui présentent un intérêt pour encourager la participation (ce que nous faisons avec une situation connue des enseignants) et qu'il faut aussi éviter que le participant ne développe une attitude négative, donc qu'il faut placer les questions sensibles ultérieurement (ce que nous faisons en plaçant les questions à caractère personnel à la fin).

5. Présentation des résultats : analyse

5.1. Échantillon

Genre	Femme		Homme		Autre		
	77		29		0		
Âge	22 – 29		30 – 39	40 – 49	50 – 59	60 – 69	
	12		59	17	14	4	
Années d'expérience	0 – 4	5 – 9	10 – 19	20 – 29	30 – 39	40 – 44	
	10	22	57	12	4	1	
Qualification	AESI		Master à finalité didactique	Autre qualification pédagogique		Sans qualification pédagogique	
	78		12	13		3	
Matière enseignée	Français	Mathématiques	Sciences	Langues	Géographie-Histoire	Autre	
	22	21	20	10	9	33	
Réseau	Privé	Libre confessionnel	Libre indépendant	Officiel WBE	Communal	Provincial	
	2	10	2	69	20	3	
Filière	Différencié	1 ^{er} degré commun ou général	Technique ou artistique de transition	Technique ou artistique de qualification	Professionnel	Spécialisé	DASPA
	1	72	5	12	10	2	4
Questions supplémentaires	1 ^{er} choix de carrière		Cours sur l'évaluation en FI			Cours sur l'évaluation en FC	
	Oui	Non	Oui	Non	Pas d'AESI	Oui	Non
	74	32	64	31	11	51	55

Tableau 3. Descriptif de l'échantillon

Cent six personnes ont participé à cette recherche dont près de 73% sont des femmes. L'âge moyen est de 39 ans et les participants ont en moyenne 13 années d'expérience professionnelle. Ces enseignants proviennent de tous les réseaux et de toutes les filières présents en FWB. Les enseignants participants à cette recherche sont 84,9% à avoir soit une agrégation de l'enseignement secondaire inférieur [AESI] soit un master à finalité didactique et pour 69,8% de cet échantillon, l'enseignement est le premier choix de carrière. Une formation à l'évaluation aurait été donnée à 60,4% des enseignants de cette recherche lors de leur formation initiale et à 48,1% lors d'une FPC.

5.2. Partie 1 du questionnaire : les approches

5.2.1. Fiabilité, consistance interne et analyse factorielle des approches

Les données récoltées ont été analysées avec le logiciel JAMOVI. L'alpha de Cronbach est un test largement utilisé en sciences sociales et peut suffire à établir la fiabilité de la cohérence interne mais l'oméga de McDonald est plus robuste avec de petits échantillons, comme c'est le cas dans ce travail (Navarro et al., 2020) et il accepte que les données soient multifactorielles ce qui est plus en lien avec des données recueillies dans des études en sciences de l'éducation

(Béland et al., 2017). Si les deux indices de fiabilité ne sont pas identiques (ce qui est le cas ici : $\alpha=0,668$ et $\omega=0,698$), il faut privilégier l'oméga de McDonald et les deux sont satisfaisants à partir de 0,7 (Navarro et al., 2020). Nous considérons que notre oméga de McDonald est suffisamment proche de cette valeur pour être considéré comme satisfaisant.

Avant de créer des clusters à partir des scores en AOL, AFL et AAL dans la partie 1, les scores, qui sont obtenus en effectuant la moyenne des réponses des participants sur les items spécifiques à chaque approche, sont calculés et une vérification de la fiabilité de ces 3 dimensions doit être effectuée. Pour l'AOL et l'AFL, ni l'alpha ni l'oméga ne dépasse 0,58 (voir annexe VI.a.). En ce qui concerne l'AAL ($\alpha = 0,624$ et $\omega = 0,642$), malgré des résultats plus proches de 0,7 que les deux autres approches de l'évaluation, ces indices n'atteignent pas les seuils recommandés. Les résultats de l'analyse factorielle confirmatoire vont dans le même sens.³ En effet, malgré un khi carré relatif respectable ($\chi^2/DDL = 451/186 \simeq 2,424$) puisqu'il est fort proche du seuil de 2 (tout en lui étant malheureusement supérieur alors qu'il devrait lui être inférieur) les autres indices (CFI, TLI, RMSEA) sont très loin des seuils recommandés pour un ajustement satisfaisant : CFI>0,9 ; TLI>0,9 ; $0,05 < RMSEA < 0,08$ (Navarro et al., 2020 ; Béland & Michelot, 2020) ; SRMR<0,08 (Hu & Bentler, 1999).

Mesures ajustées					
CFI (indice d'ajustement comparatif)	TLI	SRMR	Approximation de RMSE	IC90% de la RMSEA	
				Borne inf	Supérieur
0.391	0.313	0.130	0.116	0.102	0.130

Tableau 4. Ajustement du modèle de la partie 1 du questionnaire

Dans le tableau (annexe VI.b.) indiquant la contribution de chaque item à sa variable (AOL, AFL ou AAL) deux items ont un $p > 0.005$ ce qui signifie que ces items ne contribuent pas de manière fiable au facteur. Il s'avère que ce sont 2 items sur lesquels les participants au laboratoire cognitif avaient beaucoup réagi. Le premier item (« vérifier si la cote finale obtenue par chaque élève représente bien le niveau d'acquisition global de la matière, selon ce que chacun a produit dans l'évaluation ») a été difficilement compris (voir item 16 de la partie 1 de l'annexe IV.b ; voir item 10 de la partie 1 de l'annexe IV.e. ; voir item 10 de la partie 1 de l'annexe IV.f.). Malgré plusieurs reformulations, il est probable que cet item n'ait toujours pas une formulation claire et compréhensible de tous. L'autre item problématique est « écrire sur l'évaluation de chaque élève des retours personnalisés sur les types d'erreurs effectués et des

³ Une tentative d'analyse factorielle exploratoire [AFE] a donc été effectuée mais il n'y avait pas de lien, selon nous, aux groupements d'items proposés par l'AFE.

conseils sur la manière de les travailler ». C'est l'item qui avait le plus fait réagir les participants du laboratoire cognitif, même ceux qui répondaient positivement aux autres items d'approche AFL : « on n'est pas payé pour ça » (voir item 23 de la partie 1 de l'annexe IV.b.), « on sait qu'on doit le faire mais on n'a pas le temps » (voir remarques générales de l'annexe IV.d.). Malgré le fait que ce genre de feedback est reconnu par la littérature scientifique comme parmi les plus efficaces et les moins coûteux, dans la pratique, cela reste une démarche chronophage. Certains enseignants n'ayant que des cours à 1h par semaine et un nombre de classes élevé, nous pouvons imaginer la difficulté à mettre cela en place. Nous décidons donc d'éliminer ces 2 items.

Dans 2 des 3 boîtes à moustache représentant les scores sur chaque approche, 2 participants sont marqués comme données aberrantes par le logiciel (voir annexe VI.c.). Le premier (n°59) a un score de 1,7 en AFL alors que tous les autres participants ont des scores supérieurs à 2,5. Le deuxième (n°39) a un score de 5,286 en AAL mais il y a 6 autres participants dont le score est compris entre 5 et 5,3. À la suite de ces observations, nous décidons donc de ne supprimer que le participant n°59.

Test d'ajustement exact

χ^2	ddl	p
326	149	< .001

$$\frac{\chi^2}{ddl} = \frac{326}{149} \approx 2,21$$

Mesures ajustées

	AOL	AFL	AAL
α	0,606	0,534	0,622
ω	0,612	0,552	0,640

CFI (indice d'ajustement comparatif)	TLI	SRMR	Approximation de RMSE	IC90% de la RMSEA	
				Borne inf	Supérieur
0.468	0.390	0.117	0.106	0.0906	0.122

Tableau 5. α , ω et ajustement du modèle de la partie 1 du questionnaire après réajustement

À l'exception des α et ω de la dimension AFL, tous les indices présents dans le tableau ci-dessus se sont rapprochés des seuils recommandés, mais n'apportent pas une validité statistique satisfaisante à notre modèle. Cela semblerait signifier que les items ne mesurent pas les dimensions souhaitées. Mais les raisons de ces résultats peuvent être multiples. Ils pourraient être en lien avec une mauvaise formulation des items. Les participants pourraient avoir mal compris des items ou n'ont peut-être pas décelé le but sous-jacent à chaque action. Une dernière cause possible est la taille de l'échantillon (N=106) : c'est peu pour obtenir une fidélité à l'intérieur d'un modèle (Hu & Bentler, 1999).

Ces résultats n’empêchent pas forcément l’utilisation des scores. Examinons ce que signifient ces résultats. Tous ces résultats sont obtenus à partir d’une covariance entre les facteurs à l’intérieur d’une même approche mais également à partir de corrélations entre les approches. En ce qui concerne la covariance entre les facteurs à l’intérieur d’une même approche, cela signifierait que les participants, même ceux qui auraient le même score dans cette approche, ne répondraient pas de manière semblable dans les items relatifs à cette approche. Cela peut être expliqué par le fait que certains participants, en raison de leur environnement d’apprentissage, vont être plus susceptibles de mener ou non certaines actions. Ceci est en lien avec la littérature car l’AL, dont les pratiques évaluatives, est influencé par l’environnement d’apprentissage (DeLuca et al., 2020) mais ceci est également en lien avec un de nos laboratoire cognitif dans lequel la participante déclare être d’accord avec le principe de les laisser se placer dans les groupes de besoins pour les remédiations mais que dans la réalité, les élèves se placent avec leurs copains et le but de remédiations de leurs propres lacunes n’est pas rencontré donc in fine, elle ne ferait pas cette action. En ce qui concerne la corrélation entre les approches, ce modèle montrerait qu’il n’y a pas de lien entre les approches. Cette matrice ci-dessous le confirme, surtout en ce qui concerne les liens entre l’approche AOL et les autres.

Matrice de corrélation		Moyenne AOL	Moyenne AFL	Moyenne AAL
Moyenne AOL	Rho de Spearman	—		
	ddl	—		
	valeur p	—		
Moyenne AFL	Rho de Spearman	0.048	—	
	ddl	103	—	
	valeur p	0.624	—	
Moyenne AAL	Rho de Spearman	-0.161	0.414***	—
	ddl	103	103	—
	valeur p	0.100	< .001	—

Tableau 6. Matrice de corrélation entre les scores obtenus dans chaque approche

Or, et c’est un des buts de notre recherche (celui de constater des profils selon leurs positionnements sur les différentes approches), 2 participants ayant le même rapport à l’AOL pourraient avoir des rapports diamétralement opposés sur l’AFL ou l’AAL. De plus, un alpha et un oméga proche de 0,6 indiquent une variance d’erreur – le fait que les items ne mesurent pas la même chose (Monseur, 2023) – dans le score de l’échelle proche de 40%, ce qui est élevé et indique par exemple que l’approche AFL est moins cohérente comme mesure fiable que les autres approches (Navarro et al., 2020). Cela indique donc une variance d’erreur mais n’empêche pas l’utilisation de ces approches. Nous décidons donc d’utiliser tout de même ces scores pour la réalisation de nos clusters mais à la vue des analyses de fiabilité statistique, nous nous contentons de rester dans une approche exploratoire.

5.2.2. Détermination des clusters/profils

Sur base des 3 scores AOL, AFL et AAL, plusieurs tentatives de création de 2, 3, 4, ... groupements de participants (clusters) sont effectuées à l'aide de la méthode k-means. À partir de 4, le logiciel Jamovi refuse de créer des clusters. Le choix se porte donc sur 2 ou 3 clusters. Dix puis 20 puis 30, ... valeurs de départ aléatoires sont choisies par le logiciel jusqu'à ce que les clusters soient stables, c'est-à-dire jusqu'à ce que les centres de chaque cluster ne se modifient plus. Que ce soit pour 2 ou 3 clusters, une stabilité est rencontrée (annexe VII.a.).

Le logiciel Jamovi, au travers du graphique détaillant le nombre optimal de clusters (annexe VII.b.), indique qu'1 cluster est le nombre optimal ce qui signifie probablement qu'il n'y a pas de groupements naturels parmi les répondants. Cependant, vu les objectifs de notre recherche, nous éliminons l'option « 1 cluster ». Dans ce même graphique, il est indiqué que 3 clusters sont préférables à 2. Pourtant, quel que soit le nombre de clusters, quand on compare les boîtes à moustaches dans les approches dans lesquelles ils sont censés se différencier, il n'y pas vraiment de superposition. Il y a une autre raison, en plus du graphique de l'annexe VII.b. cité supra, incitant au choix de 3 clusters. Dans l'option « 2 clusters », il y a davantage de variance intra-cluster qu'inter-clusters (annexe VII.a.) ce qui n'est pas le cas dans l'option « 3 clusters ». Nous choisissons donc l'option « 3 clusters ».

5.2.3. Description des clusters

	N	AOL	AFL	AAL
Cluster 1	40	4,371	5,008	3,704
Cluster 2	39	2,944	4,491	3,773
Cluster 3	26	4,199	3,647	2,478

Tableau 7. Centres de chaque cluster

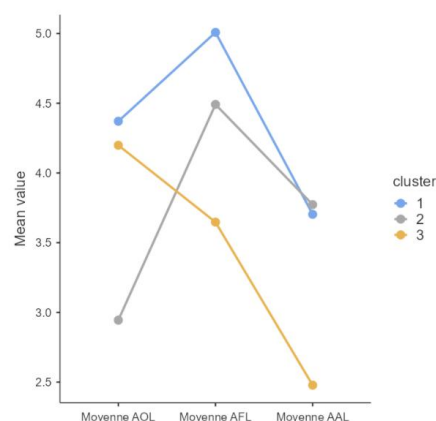


Figure 3. Graphique des clusters

Ces résultats signifient que le cluster 1 déclare pratiquer assez bien des actions liées à l'AOL, fortement des actions liées à l'AFL et moyennement des actions liées à l'AAL ; le cluster 2 déclare pratiquer rarement des actions liées à l'AOL, assez bien des actions liées à l'AFL et moyennement des actions liées à l'AAL ; le cluster 3 déclare pratiquer assez bien des actions liées à l'AOL, moyennement des actions liées à l'AFL et rarement des actions liées à l'AAL.

La part d'hommes et de femmes dans les clusters ne sont pas équivalentes proportionnellement parlant. C'est à dire que les femmes semblent avoir évité le cluster 3 et que les hommes semblent avoir évité le cluster 1. En effet, seulement 19,5% des femmes interrogées se retrouvent dans le cluster 3 et seulement 17,9% des hommes se retrouvent dans le cluster 1 (voir annexe VIII.a.)

Les clusters ne se différencient pas sur l'âge (voir annexe VIII.b.) puisque, même si les clusters possèdent des moyennes d'âge différentes (respectivement 41,3 ; 37,6 et 37,8), l'écart-type dépasse 8 pour les 3 clusters. En ce qui concerne les années d'expérience (voir annexe VIII.b.), le même constat est observé. De plus, les novices (< 5 ans d'expérience) ne se retrouvent pas tous dans un même cluster. Sur les 10 participants novices, la répartition est respectivement de 3 ; 5 et 2.

Quelques données sont à mettre en évidence en ce qui concerne les matières enseignées à l'intérieur des clusters (voir annexe VIII.c.). Premièrement, alors qu'ils sont 20 enseignants de sciences à avoir participé à cette étude, seulement 2 sont dans le cluster 3. Les professeurs d'histoire/géographie sont, eux, davantage présents dans le cluster 3 (6 professeurs sur 9).

En ce qui concerne le titre pédagogique, que ce soit un AESI, un master à finalité didactique ou pas de titre pédagogique, la répartition est assez équilibrée entre les différents clusters, c'est à dire qu'on observe une répartition proportionnelle à la taille des clusters (voir annexe VIII.d.). Il en va de même en ce qui concerne les filières d'enseignement, si l'enseignement est le premier choix de carrière ou non, si l'enseignant se souvient avoir reçu une formation à l'évaluation que ce soit lors de sa formation initiale ou en FPC (voir annexes VIII.e., VIII.f., VIII.g, VIII.h.).

5.2.4. Pratiques faisant l'unanimité

	Items récoltant le plus d'adhésion			Items récoltant le moins d'adhésion	
	AFL7 ⁴	AOL5 ⁵	AFL5 ⁶	AOL6 ⁷	AAL2 ⁸
Moyenne	5,12	4,97	4,95	2,20	2,44

Tableau 8. Actions faisant l'unanimité

⁴ AFL7 = « Rappeler que les évaluations servent à savoir où ils en sont dans l'acquisition de leurs compétences mais qu'ils peuvent utiliser ces informations pour progresser ».

⁵ AOL5 = « Discuter des cotes avec l'équipe pédagogique pour vérifier si le phénomène se répète également chez eux ».

⁶ AFL5 = « Discuter avec des collègues de la même discipline des compétences non acquises en vue d'échanger sur les pratiques pédagogiques ».

⁷ AOL6 = « Renseigner l'élève sur le fait qu'il a réussi ou échoué à l'évaluation en écrivant uniquement la cote globale, sans faire de commentaire ».

⁸ AAL2 = « Créer des zones en classe liées aux compétences non acquises et laisser les élèves se grouper selon ce qu'ils estiment devoir retravailler ».

Le tableau ci-dessus reprend les actions pour lesquelles l'échantillon, quel que soit le cluster, a montré le plus ou le moins d'adhésion.

Les écarts-types de tous les items de la partie 1, que ce soit globalement ou au sein de chaque cluster, sont tous autour de 1-1,5 ce qui indique que toutes les réponses sont moyennement dispersées autour de leurs moyennes. Aucun écart-type n'est assez proche de 0 ou de 2,5 pour être relevé comme notable.

5.3. Partie 2 du questionnaire: Sentiment d'efficacité et besoins de formations

Cette deuxième partie du questionnaire est une combinaison des questionnements sur le SE et les BF. Ces 2 variables sont analysées de manière distincte puis une analyse du lien entre ces 2 variables est effectuée.

5.3.1. Fiabilité, consistance interne et analyse factorielle du sentiment d'efficacité

L'alpha de Cronbach et l'oméga de McDonald sur tous les items traitant le SE sont respectivement de 0,944 et 0,954.

Comme dit dans la partie théorique, 4 types de SE sont à vérifier statistiquement : en général (obtenu à l'aide des items sur le cycle MoSCodEE), lié à l'AOL, lié à l'AFL et lié à l'AAL.

Les consistances internes à chaque type de SE sont très bonnes statistiquement puisque le plus petit indice (alpha et oméga confondus) est de 0,848 (voir tableau ci-dessous).

	SE général	SE AOL	SE AFL	SE AAL				
α	0,893	0,862	0,848	0,854				
ω	0,896	0,869	0,850	0,869				

Test d'ajustement exact			Mesures ajustées					
χ^2	ddl	P					IC90% de la RMSEA	
			CFI (indice d'ajustement comparatif)	TLI	SRMR	Approximation de RMSE	Borne inf.	Supérieur
846	399	<.001	0.763	0.742	0.0795	0.103	0.0936	0.113

$$\frac{\chi^2}{ddl} = \frac{846}{399} \approx 2,12$$

Tableau 9. α , ω et ajustement du modèle de la partie SE

Les résultats de l'ajustement du modèle sont proches des seuils recommandés mais, à l'exception du SRMR qui l'atteint, les autres n'atteignent pas ces seuils et donc ne correspondent pas à un modèle dit satisfaisant. Nous concluons donc que le modèle ne

correspond pas bien statistiquement aux données observées mais vu la proximité avec ces seuils arbitraires nous acceptons ce modèle tout en continuant cette recherche de manière exploratoire.

Dans le tableau relevant la contribution des items aux facteurs (voir annexe IX.a.), chaque item possède un p inférieur à 0,001 ce qui signifie que chaque item apporte « une contribution raisonnable au modèle » (Navarro et al., 2020, p.459). Dans le tableau des indices de modification utile pour repérer un éventuel item problématique, l'indice le plus grand est obtenu dans la colonne SE AOL par le SE AFL4 : « J'ai confiance en mes capacités à communiquer les objectifs et les buts de l'évaluation aux parents/tuteurs lorsque cela est approprié ». En relisant cet item, nous nous rendons compte que notre formulation est à revoir. En effet, cet item pourrait être interprété par des buts de réussite (exemple : obtenir au moins 60%) au lieu de l'interpréter par des buts d'apprentissage comme nous l'avions conçu au départ. Nous réfléchissons donc à un déplacement ou une suppression. Cependant, chaque item SE étant relié avec un item BF, si changement il y a, il doit s'opérer sur les 2 items. En observant les indices de modification de l'analyse factorielle des BF, nous remarquons que le BF AFL4 serait à déplacer dans les BF AAL. Cette différence dans les déplacements ajoutée à la mauvaise formulation, nous poussent à supprimer les items SE AFL4 et BF AFL4.

			SE général	SE AOL	SE AFL	SE AAL		
α			0,893	0,862	0,840	0,854		
ω			0,896	0,869	0,842	0,869		

Test d'ajustement exact			Mesures ajustées					IC90% de la RMSEA	
χ^2	ddl	p	CFI (indice d'ajustement comparatif)	TLI	SRMR	Approximation de RMSE	Borne inf	Supérieur	
750	371	<.001	0.785	0.765	0.0778	0.0986	0.0884	0.109	
$\frac{\chi^2}{ddl} = \frac{750}{371} \approx 2,02$									

Tableau 10. α , ω et ajustement du modèle de la partie SE après réajustement

Après avoir supprimé ces items, la plupart des indices se sont rapprochés des seuils recommandés, sans les atteindre.

5.3.2. Analyse du sentiment d'efficacité de l'échantillon

Sentiment d'efficacité global

En ce qui concerne le SE global (1 item seulement reprenant le sentiment d'efficacité global des participants), les moyennes des réponses des participants des 3 clusters sont : 4,78 pour le premier cluster, 4,49 pour le deuxième cluster et 4,96 pour le troisième cluster. Les écarts-types

vont de 0,662 à 0,756. Les clusters ne se différencient donc pas sur cet item. Ces moyennes hautes sont dues notamment au fait que seulement 3 personnes sur les 105 personnes ont répondu se sentir peu capable en évaluation.

Clusters et scores des types de sentiment d'efficacité

Un score (moyenne) est calculé par variable (SE général, SE AOL, SE AFL et SE AAL) pour chaque cluster. Vu les résultats statistiques de fiabilité que ce soit pour les approches ou pour les SE, nous nous contentons de rester dans une visée exploratoire.

Premièrement, si on compare les scores de chaque cluster dans les 4 types de SE, on peut voir que les scores pour un même type entre les clusters sont forts proches puisqu'ils diffèrent de maximum 0,3 (voir annexe X.a.).

Deuxièmement, si on compare les différents types de SE entre eux, on peut remarquer que le général, l'AOL et l'AFL sont également très proches et se trouvent aux environs de 4,86

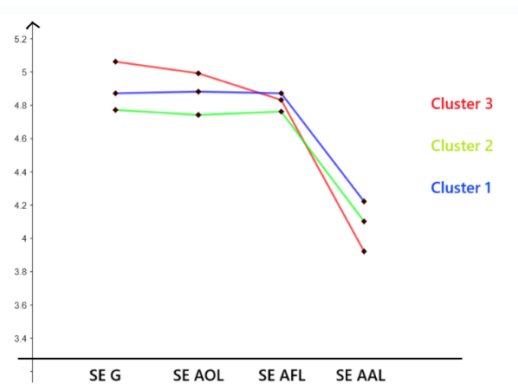


Figure 4. Graphique des SE

avec des écarts-types aux environs de 0,671 (voir annexe X.a.). La très grande partie des participants se déclareraient donc au moins assez capable dans la plupart des items de ces 3 types. Nous ne nous attardons donc pas trop dessus dans la suite pour ces 3 types de SE. Le SE en AAL, lui, diffère des 3 autres avec des scores se situant aux alentours de 4,08 et des écarts-types aux alentours de 0,925.

Analyse des types de sentiment d'efficacité pris séparément

Les moyennes de l'ensemble des participants sont assez hautes (entre 4,64 et 5,20) pour le SE général (voir annexe X.b.). Le score le plus bas a été obtenu à l'étape 1 du cycle MoSCodEE, même si 4,64 signifie être entre assez capable et très capable. L'étape 1 consiste à analyser un programme ou un référentiel pour identifier la matière mais également pour créer une triple concordance objectifs-méthodes-évaluations. Le score le plus haut est obtenu à l'étape 5 dans laquelle il est question de communiquer clairement aux élèves les informations utiles pour la préparation à l'évaluation. L'ensemble des participants y a répondu de façon positive (échelons 4 à 6). Quand on compare les moyennes pour chaque étape du cycle MoSCodEE entre les clusters (voir annexe X.b.), nous notons que les différences entre celles-ci sont toutes inférieures

aux écarts-types de chaque cluster. Ce qui veut dire que les clusters ne se différencient pas réellement sur ces étapes du cycle MoSCodEE au regard de la variable SE.

Si on compare les moyennes selon les items SE en AOL (voir annexe X.c.), l'item SE AOL5 obtient la moins bonne moyenne même si la différence avec les autres est minime mais avec un bien plus grand écart-type que les autres. Cet item est celui qui traite de la capacité à proposer un redoublement ou une réorientation sur base des résultats finaux. En regardant attentivement l'annexe X.d., 18 enseignants sur les 105 participants ont répondu de « pas du tout capable » à « peu capable », alors que pour chacun des autres items AOL ils étaient maximum 7 sur 105 dans ces catégories. À travers les clusters, il y a 13 enseignants sur les 18 précédemment évoqués qui ressortent du cluster 2, le cluster le moins porté sur l'AOL.

Concernant les items SE en AFL (voir annexe X.e.), les moyennes sont à nouveau des nombres élevés (de 4,65 à 4,91) témoignant d'un bon SE mais avec des écarts-types plus grands allant de 0,876 à 1,02. En allant regarder la répartition, cela s'explique par un peu plus de participants donnant des réponses codées 3 et 6 (« peu capable » et « entièrement capable ») mais rien de très particulier n'est à constater pour cette variable même lors d'une observation par cluster.

Nous nous attendions à des résultats plus mitigés pour le SE en AAL par rapport aux autres types de SE (voir annexe X.f.). Les résultats sont, comme prévu, plus bas mais témoignent d'un SE plutôt bon vu que les moyennes pour chaque item varient de 3,8 à 4,61. Les écarts-types eux aussi divergent en étant plus grand que dans les autres types de SE (de 1,02 à 1,36). Cela montre que malgré des moyennes plus grandes que 3,5 (considéré comme le centre de l'échelle), les participants ont répondu de façon très variée à ces items. Si on observe au niveau des clusters, les moyennes pour un même item ne varient pas beaucoup entre les clusters sauf pour l'AAL 6 avec le cluster 3 et le cluster 1 ayant respectivement des moyennes de 3,65 et de 4,6. Les grands écarts-types nous indiquent qu'il n'y a peut-être finalement pas de grande différence entre ces 2 clusters. Quand on regarde la répartition des réponses dans le tableau de fréquences (voir annexe X.g.), en effet elle est très similaire, à l'exception de 5 personnes ayant répondu ne pas être du tout capable d'animer une discussion collective avec les élèves autour du sens de l'évaluation.

5.3.3. Fiabilité, consistance interne et analyse factorielle des besoins de formations

L'alpha de Cronbach et l'oméga de McDonald, sur tous les items traitant les BF, sont respectivement de 0,970 et 0,973.

Les consistances internes à chaque type de BF (en général, lié à l'AOL, lié à l'AFL et lié à l'AAL) sont très bonnes statistiquement puisque le plus petit indice (alpha et oméga confondus) est de 0,893 (voir tableau ci-dessous).

			BF général	BF AOL	BF AFL	BF AAL	
	α		0,939	0,900	0,908	0,893	
	ω		0,939	0,901	0,909	0,895	
Test d'ajustement exact			Mesures ajustées				
χ^2	<u>ddl</u>	<u>p</u>	IC90% de la RMSEA				
1132	399	<.001	CFI (indice d'ajustement comparatif)	TLI	SRMR	Approximation de RMSE	Borne inf. Supérieur
$\frac{\chi^2}{ddl} = \frac{1132}{399} \approx 2,84$			0.752	0.730	0.0733	0.132	0.123 0.141

Tableau 11. α , ω et ajustement du modèle de la partie BF

Les résultats de l'ajustement du modèle sont proches des seuils recommandés mais, à l'exception du SRMR qui l'atteint, les autres n'atteignent pas ces seuils et donc ne correspondent pas à un modèle dit satisfaisant. Nous concluons donc que le modèle ne correspond pas bien statistiquement aux données observées mais vu la proximité avec ces seuils arbitraires nous acceptons ce modèle tout en continuant cette recherche de manière exploratoire.

Dans le tableau relevant la contribution des items aux facteurs (voir annexe IX.b.), chaque item possède un p inférieur à 0,001 ce qui signifie que chaque item apporte « une contribution raisonnable au modèle » (Navarro et al., 2020, p.459). Dans le tableau des indices de modification (voir annexe IX.b.) utile pour repérer un éventuel item problématique, deux items possèdent un plus grand indice que les autres. L'item MOSCO BF8 concerne le retour constructif à donner aux élèves à la suite d'une évaluation. Ce retour pourrait servir, selon l'enseignant, à des buts de réussite, d'apprentissage et/ou d'autonomie de l'élève. Cependant, dans les items MoSCodEE, c'est le cas d'autres items également. De plus, une suppression ou un déplacement enlèverait une des étapes du cycle MoSCodEE. Nous décidons donc de ne pas déplacer ou supprimer l'item car il ne nous semble pas y avoir de sens à le faire. Un autre item est potentiellement problématique : le BF AAL2 placé dans la colonne BF AFL. Il concerne l'utilisation des retours des élèves pour modifier ses pratiques. Nous comprenons l'idée sous-jacente des enseignants de voir cela comme une réflexion sur ses pratiques qui pourrait s'apparenter à de l'AFL mais c'est bien sur base des propos des élèves que l'action est envisagée. L'élève est donc bien au centre de l'évaluation et c'est lui qui est sollicité pour faire évoluer ces pratiques. Nous décidons également de ne pas le déplacer ou le supprimer.

5.3.4. Analyse des besoins de formations de l'échantillon

Besoin de formation global

En ce qui concerne le BF global (1 item seulement reprenant le besoin de formation global des participants), tous les échelons ont été choisis par au moins un participant (voir annexe XI.a.) et les moyennes des réponses des participants des 3 clusters sont : 3,75 pour le premier cluster, 3,69 pour le deuxième cluster et 3,08 pour le troisième cluster. Les écarts-types vont de 1,28 à 1,47. Les clusters ne se différencient donc pas réellement sur cet item. Mais vu la différence assez importante entre les deux premiers clusters et le 3^e, une observation du tableau des fréquences a été effectuée. Dans le cluster 1, il y a deux fois plus de personnes qui ressentent un besoin de formation que de personnes qui en ressentent moins alors que dans les autres clusters, la moitié des participants sont de part et d'autre de 3,5.

Clusters et scores des types de besoins de formations

Un score (moyenne) est calculé par variable (BF général, BF AOL, BF AFL et BF AAL) pour chaque cluster (voir annexe XI.b.). Vu les résultats statistiques de fiabilité que ce soit pour les approches ou pour les BF, nous nous contentons de rester dans une visée exploratoire.

Premièrement, si on compare à l'intérieur d'un même type de BF, encore une fois, les différences de moyennes sont inférieures aux écarts-types ce qui montre que les clusters ne se différencient pas réellement. Les différences de moyennes sont plus grandes en BF AAL et en BF général (respectivement 0,64 et 0,54 contre 0,34 et 0,3 pour les autres types de BF). De ce fait, nous décidons de regarder plus attentivement les items liés à

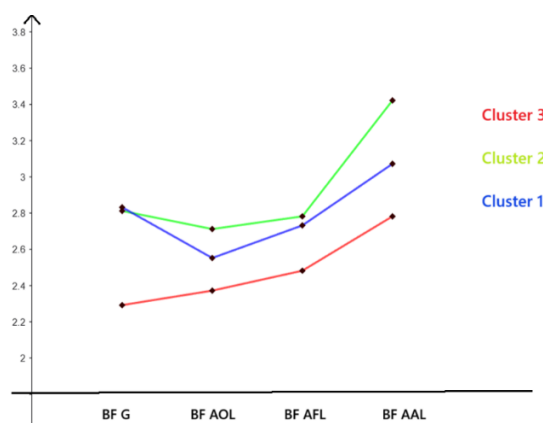


Figure 5. Graphique des BF

l'AAL et aux étapes du cycle MoSCodEE. Deuxièmement, si on compare les 4 types de BF, on constate que tous les clusters ont des scores aux alentours de 2,6 à l'exception du BF AAL pour lequel les scores de chaque cluster sont les plus élevés puisqu'ils sont aux alentours de 3. Ceci pourrait signifier une légèrement plus grande envie de se former sur cette approche.

Analyse des types de besoins de formations pris séparément

Les moyennes de l'ensemble des participants sont moyennement basses (entre 2,20 et 3,22) pour le BF général (voir annexe XI.c.). Le score le plus bas a été obtenu à l'étape 3 du cycle MoSCodEE, ce qui signifie que les participants ont fait des déclarations autour de l'échelon

« ne presque pas ressentir le besoin de se former » pour planifier des évaluations. Le score le plus haut, tout en étant sous 3,5, est obtenu à l'étape 1 dans laquelle il est question d'analyser un programme, un référentiel et de créer une triple concordance entre les objectifs, les méthodes et les évaluations. Quand on compare les moyennes pour chaque étape du cycle MoSCodEE entre les clusters, on note, comme pour le SE, que les différences entre celles-ci sont toutes inférieures aux écarts-types de chaque cluster. Ce qui veut dire que les clusters ne se différencient pas réellement sur ces étapes du cycle MoSCodEE au regard de la variable BF. Cependant, 2 étapes du cycle, l'étape 2 (choix de la bonne modalité d'évaluation) et l'étape 1 (analyse d'un programme et créer une triple concordance), possèdent une très grande différence entre les moyennes des clusters. Quand on analyse le tableau de fréquence de l'étape 2 (voir annexe XI.d.), dont la différence de moyennes entre les clusters 2 et 3 est de 1,13, nous remarquons que, sur les 42 personnes qui ont exprimé leur besoin de se former, seules 5 appartiennent au cluster 3 (celui qui est le moins AAL). Sur les items MOSCO BF1P1 (traitant de l'analyse des programmes/référentiels) et MOSCO BF1P2 (traitant de la triple concordance) qui forment le BF de l'étape 1, seulement une personne du cluster 3 sur les 26 se retrouve dans les échelons 5 et 6 pour le BF1P1 et seulement 2 personnes du cluster 3 sur les 26 se retrouvent dans les échelons 5 et 6 pour le BF1P2 (voir annexe XI.e.). Même si le cluster 3 est le cluster qui compte le moins de participants, cela ne justifie pas une telle répartition (moins de 10% du cluster 3 choisissent ces échelons contre environ 20% pour les 2 autres clusters).

Si on compare les moyennes des items BF en AOL (voir annexe XI.f.), l'item BF AOL1 obtient la meilleure moyenne (2,96) même si la différence avec les autres est minime. Il s'agit de l'item qui traite de l'élaboration et l'utilisation d'épreuves bien construites pour obtenir des cotes qui reflètent les acquis des élèves. L'item BF AOL4, qui traite de la communication des résultats finaux aux parents/tuteurs, est celui qui récolte la moins bonne moyenne (2,14). À travers les clusters, les items BF AOL1 et BF AOL4 conservent leur position, pour chacun des clusters. En regardant le tableau des fréquences (voir annexe XI.g.), il n'y a que dans celui de l'AOL1 qu'on constate un élément surprenant : il y a 15% des enseignants de chacun des clusters 1 et 2 qui souhaitent fortement être formés (échelons 5 et 6) pour l'AOL1 et 0% des enseignants du cluster 3 qui le souhaitent alors que le cluster 3 est un des clusters les plus AOL.

Concernant les items BF en AFL (voir annexe XI.h.), les moyennes sont très similaires et basses (toutes aux alentours de 2,7) que ce soit entre les items ou entre les clusters. Les tableaux de fréquences montrent que, par item, il y a maximum 2 participants du cluster 3 qui se sont placés dans les échelons 5 et 6, ce qui est loin d'être le cas dans les autres clusters (voir annexe XI.i.).

Le cluster 3, dans sa grande majorité, indique donc ne pas ressentir fortement des besoins de formations liées à l'AFL et ce cluster est le moins AFL des 3 clusters.

Concernant le BF en AAL, les moyennes sont, comme prévu, plus hautes mais elles varient seulement de 2,65 à 3,58 (voir annexe XI.j.). Quel que soit le cluster observé, le BF AAL1 est l'item ayant la plus haute moyenne, et le BF AAL2 la plus petite moyenne. C'est-à-dire que d'une part, 54% des participants ressentent le besoin d'être formés à engager les élèves dans des dispositifs d'autoévaluation et d'évaluation par les pairs et, d'autre part 71% des enseignants interrogés ne ressentent pas le besoin de se former sur l'utilisation des retours constructifs de la part des élèves sur leurs pratiques évaluatives pour adapter celles-ci à leurs besoins (voir annexe XI.k.).

5.3.5. Analyse des liens entre le sentiment d'efficacité et les besoins de formations

La deuxième partie du questionnaire a été conçue de manière à proposer pour une même action, 2 items : un concernant le SE et un concernant le BF. Ces paires d'items sont appelés dans la suite de ce travail les binômes SE-BF. Pour identifier le lien entre SE et BF, des corrélations sont réalisées pour chaque binôme SE-BF. Toutes les corrélations obtenues sont négatives. La grande majorité des corrélations, 35 binômes sur les 39 (environ 90%), sont entre $-0,4$ et $-0,9$ de façon statistiquement très significative et donc représentent des relations allant de modérées à fortes (Navarro et al., 2020). Parmi ces 35 corrélations, 8 se trouvent même entre $-0,7$ et $-0,9$ ce qui signifie un lien fort. Parmi ces 8 corrélations, 4 lient des binômes SE/BF AOL et aucune ne lie des binômes AAL. Mais ce qui va nous intéresser le plus sont les 4 binômes sur les 39 qui ont une corrélation comprise entre 0 et $-0,4$ (voir annexe XII.a.). Ce choix peut sembler étrange, mais plus les corrélations sont proches de -1 , plus ce sont des corrélations pour lesquelles une certaine logique est de mise : plus on se sent capable, moins on devrait ressentir le besoin de se former. Comme écrit dans la méthodologie, nous voulons donc investiguer les items sur lesquels les participants ont développé une autre logique (par exemple : se sentir capable mais ressentir quand même le besoin de se former ou ne pas se sentir capable et ne pas ressentir le besoin de se former).

Parmi ces 4 binômes, il y a le binôme SE global/BF global ayant un coefficient de corrélation égal à $-0,389$, statistiquement très significatif. L'analyse révèle que les enseignants se déclarent très confiants en leurs compétences en évaluation (échelons 4-5-6 utilisés pour 102 participants sur les 105 interrogés) et qu'aucun enseignant ayant participé à cette recherche n'a déclaré être « pas du tout » ou « très peu » confiant en ses compétences d'évaluateur. Ils déclarent qu'ils veulent des formations pour développer leurs compétences d'évaluateur pour 61 participants

sur 105 (voir annexe XII.b.). Parmi les 44 participants restants, c'est-à-dire ceux déclarant ne pas ressentir le besoin de se former en évaluation, il n'y a que 7 participants qui ont déclaré « ne pas du tout » avoir besoin de formations pour développer leurs compétences d'évaluateur mais ceux-ci déclarent, pour 4 d'entre eux, dans les questions suivantes, « presque pas » ou « très peu » avoir besoin de se former, ce qui n'est plus un non ferme et définitif. Si on résume, il y a donc 2,9% (3 participants sur 105) se déclarant peu capables alors que 58,1% déclarent ressentir le besoin de se former. Il y a donc au minimum 55,2% des personnes interrogées qui se disent capables mais qui ressentent le besoin de se former.

Parmi les 4 binômes cités supra, un deuxième est le binôme SE/BF AAL2 (traitant les retours constructifs des élèves) ayant un coefficient de corrélation égal à -0,383, statistiquement très significatif. Ce nombre est assez étonnant au regard des tableaux de fréquences (voir annexe XII.c.) puisqu'on y note que 92,4% se sentent capables dans cette action et 71,4% ne ressentent pas réellement (échelons 1 à 3) le besoin de s'y former. Cela signifie tout de même qu'il y a, pour cette action, au minimum 21% de participants qui déclarent être capables mais déclarent ressentir le besoin de s'y former. Cette même analyse à travers les clusters nous montre que ces personnes « capables mais ayant besoin de formation » ne sont pas réparties de façon homogène dans les clusters. Ces personnes représentent au minimum 30 % du cluster 1, au minimum 20,5% du cluster 2 et au minimum 7,7% du cluster 3 (le cluster le moins AAL).

Parmi les 4 binômes cités supra, un troisième est le binôme SE/BF, traitant le changement des pratiques évaluatives à la suite de nouvelles exigences ou demandes, avec un coefficient de corrélation égal à -0,324, statistiquement très significatif. Ici, les tableaux de fréquences et les graphiques (voir annexe XII.d.) montrent clairement d'où vient ce nombre. Alors que la grande majorité des participants (93,3%) ne se trouve que dans les échelons 4-5-6 du SE, 96,2% des participants se sont placés plus ou moins équitablement entre les échelons de 1 à 5 par rapport au BF. En effectuant le même calcul que pour les binômes précédents, il y a au minimum 28,5% des participants se déclarant capable mais en ressentant tout de même le besoin de s'y former. Ces personnes représentent au minimum 35% du cluster 1, au minimum 30,8% du cluster 2 et au minimum 15,4% du cluster 3.

Le dernier des 4 binômes est le binôme SE/BF traitant des épreuves externes avec un coefficient de corrélation égal à -0,189, non statistiquement significatif. C'est un binôme que nous avons recalibré pour ne tenir compte que des personnes ayant déclaré enseigner une branche liée à une épreuve externe. Cependant, vu la non-significativité de la relation ainsi que, peut-être, la

difficulté de réponse puisque c'est le seul binôme ayant 7 échelons sur une partie dans laquelle tous les items sont sur 6 échelons, nous décidons de ne pas traiter davantage ce binôme.

En plus de ces 4 binômes, nous analysons deux autres binômes SE/BF (voir annexe XII.e.). Le premier est le binôme SE/BF AAL6 traitant la discussion collective sur le sens de l'évaluation avec un coefficient de corrélation égal à -0,402, statistiquement très significatif. Cette valeur aussi proche du -0,4 est la raison principale qui nous convainc de l'analyser. Le tableau des fréquences (voir annexe XII.f.) indique qu'environ 22% ne se sentent pas capable d'animer une telle discussion mais environ 35% désirent s'y former. Ceci signifie qu'au moins 13% des participants sont des enseignants se sentant capable dans cette action mais ressentant tout de même le besoin de s'y former. Ces personnes représentent au minimum 15% du cluster 1, au minimum 20,5% du cluster 2 et au minimum 0% du cluster 3. Il est donc possible que pour l'animation d'une discussion collective au sujet du sens de l'évaluation, il n'y ait aucun participant du cluster 3 se sentant capable mais ressentant tout de même le besoin de s'y former.

Le second binôme supplémentaire analysé est le binôme SE/BF MoSCodEE étape 1 partie 1, celui traitant l'analyse du programme/référentiel pour identifier les éléments à évaluer, car c'est un item qui a été traité dans les analyses de SE et dans les analyses de BF car les résultats ressortaient par rapport aux autres items. De plus, il pourra être mis en relation avec le binôme SE/BF traitant des changements de pratiques évaluatives (selon des exigences sociétales ou politiques par exemple). Le lien entre le SE et le BF possède un coefficient de corrélation égal à -0,598, statistiquement très significatif. On peut donc s'attendre à ne pas avoir beaucoup de personnes se disant capable mais ressentant tout de même le besoin de se former. Les calculs montrent une réalité totalement différente. Au minimum 33,4% des participants sont des personnes se disant capable d'analyser un programme/référentiel pour identifier les éléments à évaluer, tout en ressentant le besoin de s'y former. Ces personnes représentent au minimum 52,5% du cluster 1, au minimum 25,7% du cluster 2 et au minimum 15,4% du cluster 3. Ce contraste entre le coefficient de corrélation et ces pourcentages peut être expliqué en partie par le fait qu'elle fait partie de l'étape du cycle MoSCodEE possédant la moyenne en SE la plus basse (4,64) et la moyenne en BF la plus haute (3,22) par rapport à toutes les autres étapes du cycle. Leur plus grande proximité avec le centre de l'échelle (3,5) par rapport aux autres items peut être une explication.

5.4. Partie 3 du questionnaire : Modalités de formations

Avant de donner les résultats des analyses, rappelons que dans cette partie, il y a 5 échelons allant de 1 à 5. Le centre de l'échelle ne se trouve donc plus à 3,5 mais à 3 et l'écart-type

maximal est à 2 et non plus à 2,5. Puisqu'ici, une modalité centrale était possible, une partie des participants ne s'est pas prononcée en faveur ou en défaveur pour certains items.

Par rapport à la population générale (voir annexe XIII.a.), la modalité qui obtient la plus haute moyenne est MF7, le travail collaboratif en classe avec des collègues et celle qui obtient la plus petite moyenne est MF12, le mentorat individuel avec un enseignant d'une autre discipline.

L'analyse au niveau des clusters (voir annexe XIII.b.), révèle que ces 2 items conservent leur place de première et dernière modalité de formation voulue.

Ci-contre, les types de formation en présentiel (MF1, 2, 7 et 8) sont mis en opposition avec un groupement rassemblant les types de formation en ligne (MF3, 4, 5, 6 et 10). Les modes de formation en présentiel sont majoritairement préférés alors que les modes en ligne sont plus clivants. Le cluster 3, en particulier, est majoritairement défavorable aux modes en ligne alors que la part de favorables et de défavorables ne dépasse pas les 50% dans les 2 autres clusters.

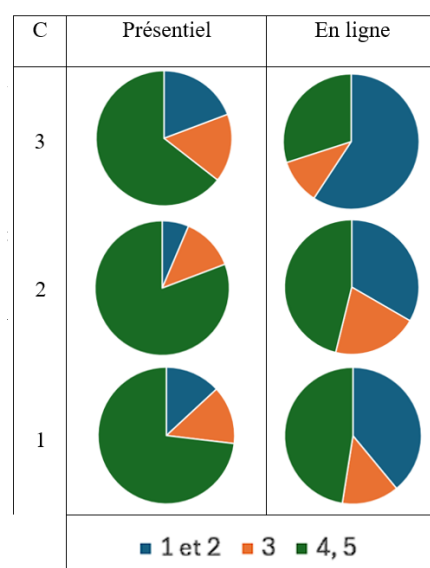


Figure 6. Présentiel vs En ligne

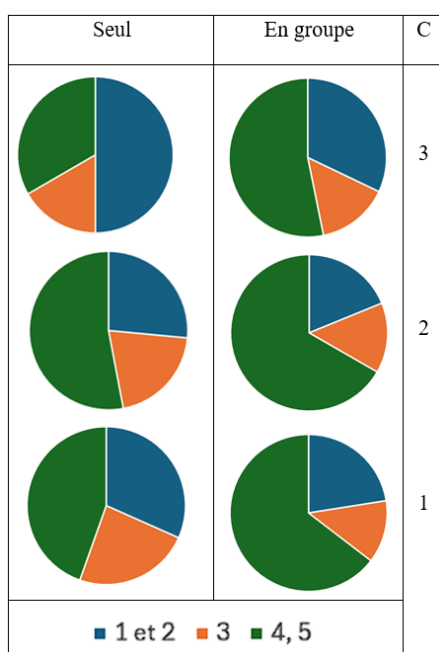


Figure 7. Seul vs En groupe

Ci-contre, les types de formation seul (MF3, 5, 10, 11, 12 et 13) sont mis en opposition avec un groupement rassemblant les types de formation en groupe (MF1, 2, 4, 6, 7 et 8). Les formations en groupe sont majoritairement préférées. Au niveau des clusters, le cluster 2, bien que majoritairement favorable aux formations en groupe, veut également, pour plus de la moitié de ses membres, se former seul. Ce cas de figure ne se retrouve pas dans les clusters 1 et 3. Le cluster 3, quant à lui, se démarque par le fait que la moitié de ses membres ne voudraient pas se former seul.

6. Interprétation et discussion

6.1. Constat des approches d'évaluation

Comme mentionné par Pastore (2023), identifier les conceptions de l'évaluation des professeurs est pertinent pour comprendre les dynamiques impliquées dans la mise en œuvre des politiques éducatives. Il semblerait que dans l'échantillon, il y ait des profils cadrant avec les politiques actuelles du Pacte pour un enseignement d'excellence puisqu'on constate clairement qu'au niveau des approches de l'évaluation, les participants déclarent se diriger vers des actions à buts AFL. En effet, parmi les trois approches, l'AFL est la seule dans laquelle aucun de nos profils n'a une moyenne inférieure à 3,5.

En revanche, l'AAL est l'approche pour laquelle les répondants ont émis le plus de réserve puisque c'est la seule qui n'atteint pas le 4 de moyenne quel que soit le cluster. Il est à noter tout de même que nos deux plus gros clusters ont une moyenne comprise entre 3,5 et 4 ce qui montre un certain intérêt de la majorité de nos participants. Le fait qu'aucun des clusters n'ait au moins un bon score en AAL confirme ce que nous avait appris l'enquête TALIS 2024 (Quittre et al., 2025) : peu (25%) d'enseignants du secondaire inférieur en FWB mettent en place des pratiques d'évaluation permettant à l'élève d'être activement impliqué dans le processus évaluatif (avec des pratiques telles que, dans TALIS 2024 : l'auto-évaluation de ses progrès). Même si certaines actions AAL du questionnaire du présent travail donnent des résultats aussi proches, pour d'autres actions (comme la discussion avec les élèves de la date de passation de la nouvelle évaluation ou comme laisser du temps aux élèves pour qu'ils échangent sur la façon de se préparer à une évaluation), ce sont presque 50% des participants qui les ont jugées plutôt probables ou tout à fait probables.

Les analyses ont montré qu'au sein même d'une approche, les items ne sont pas corrélés. En effet, même si une personne a comme but la réussite scolaire des élèves, son expérience la pousse à effectuer davantage certaines actions que d'autres visant ce but. De fait, l'AL est sujet aux systèmes affectifs et de croyances mais a également des influences socioculturelles et institutionnelles (Xu & Brown, 2016). Elle est donc bien propre à chaque personne.

Les analyses ont également montré qu'il n'y a pas de corrélation entre les approches AOL et AFL ou entre les approches AOL et AAL. En effet, un enseignant peut viser des buts de réussite et d'apprentissage. Un autre peut ne viser que des buts de réussite. Un autre encore peut ne viser que des buts d'apprentissage. Il existe en revanche un lien modéré (corrélation positive et statistiquement très significative) entre les scores AFL et AAL. Ceci signifie que si une

personne est plus portée qu'une autre sur des évaluations visant des buts d'apprentissage, elle sera également plus portée à impliquer activement l'élève dans le processus évaluatif. La création de clusters montre différents profils combinant différemment l'ensemble des 3 approches.

Le cluster 1 est caractérisé par un score haut en AOL, un score très haut en AFL et un score moyen en AAL. Très peu d'hommes y sont alors que c'est le cluster le plus AFL des 3. L'étude de Coombs et ses collaborateurs (2022) expose le fait que les femmes sont davantage dans une approche AFL que les hommes. Cette tendance apparaît ici également.

Le cluster 2 est caractérisé par un score bas en AOL, un score haut en AFL et un score moyen en AAL. C'est le seul cluster avec un score bas en AOL. Il y a donc une partie des enseignants ne visant pas des buts de réussite chez leurs élèves privilégiant des buts d'apprentissage. Ce cluster n'est pas genré et ne se différencie pas des autres clusters par toute autre caractéristique (qualification, matière, âge, ...).

Le cluster 3 est caractérisé par un score haut en AOL, un score moyen en AFL et un score très bas en AAL. C'est un cluster évité par les professeurs de sciences mais fortement choisi par les enseignants de géographie/histoire. Une première raison possible selon nous est que les enseignants des cours d'une ou deux heures par semaine ont moins de possibilité de mettre l'élève au centre de son évaluation en raison de la trop grande multitude d'élèves. Une seconde cause possible est que ces enseignants privilégient l'apprentissage de la matière durant le peu d'heures de cours par semaine car ils n'ont pas la possibilité temporelle de faire autre chose que cela. D'autre part, une constatation surprenante doit être mentionnée ici : alors que Soussi et ses collègues (2006) font part du fait que les évaluations formatives interviennent autant pour les enseignants de français, sciences, mathématiques et histoire-géographie-sciences sociales, nous constatons dans notre étude que ce n'est pas entièrement le cas pour les professeurs d'histoire-géographie qui se retrouvent dans le cluster le moins porté sur les approches AFL/AAL.

Il est à remarquer que le cluster 3 n'est composé que de 26 personnes contrairement aux clusters 1 et 2 respectivement composés de 40 et 39 personnes. Il y a donc une majorité (75 %) des participants qui sont plutôt favorables aux approches AFL et légèrement plus favorables que défavorables pour l'approche AAL. Ceci permet de confirmer ce qui a été dit supra : la majorité des enseignants participant à cette recherche sont déjà en partie engagés dans les approches voulues par le Pacte pour un enseignement d'excellence en ce qui concerne l'évaluation, c'est-à-dire soutenir les apprentissages des élèves au travers de l'évaluation.

Tous les résultats trouvés dans la littérature scientifique concernant les novices trouvent très peu d'écho dans la présente recherche dû au petit nombre de novices (10) et à sa répartition plus ou moins équilibrée au niveau des clusters.

Au niveau des actions, dans cette partie 1, certaines ont été jugées très probables et d'autres très improbables, quel que soit le cluster. Une des informations les plus notables est que, parmi les 3 items les plus choisis, 2 concernent les discussions entre collègues, que ce soit au sujet des cotes ou des compétences. Braxmeyer et ses collègues (2005) mentionnent d'ailleurs le fait que 75% des enseignants s'informent des résultats obtenus par les élèves dans les autres matières. L'action dans laquelle les enseignants se sont placés le plus positivement, même ceux des clusters avec un score haut en AOL, est celui dans lequel on rappelle aux élèves que l'évaluation montre l'apprentissage à un temps t mais qu'il est possible d'évoluer après. Ceci signifie que, même pour les enseignants qui considèrent que l'enseignement s'arrête à l'interrogation de fin de chapitre, l'apprentissage peut continuer pour l'élève. D'ailleurs, DeLuca et ses collègues (2019b) établissent le fait qu'il n'existe pas d'enseignant avec un état d'esprit fixe. Les enseignants sont soit avec un état d'esprit mixte (davantage tendance AOL-AFL) soit avec un état d'esprit de croissance (davantage tendance AAL).

Parmi les items dans lesquels les enseignants sont le plus défavorables quel que soit le cluster, il y a celui dans lequel l'enseignant crée des zones de remédiations liées aux compétences mais pour lesquelles il laisse l'élève se placer en fonction de ce qu'il estime devoir retravailler. Un des participants aux laboratoires cognitifs donne une explication à ce rejet : il évoque le fait que son expérience lui a appris que les élèves vont se placer selon les affinités sociales et non par rapport aux compétences et donc ne le ferait pas car cela créerait une dynamique qui entraverait l'apprentissage (voir item 6 de la partie 1 de l'annexe IV.b.).

Ce qu'il faut en retenir :

- L'approche AFL est la plus présente ;
- L'approche AAL est la moins présente ;
- Une importance est donnée à la discussion entre collègues ;
- L'expérience influence fortement les choix d'actions
- Les femmes sont plus AFL que les hommes
- Il existe différentes combinaisons des approches

6.2. Constat des sentiments d'efficacité

Les analyses statistiques révèlent qu'il est possible que, non seulement il y ait bien différents types de SE mais qu'en plus, ils soient assez corrélés. D'ailleurs dans la matrice de corrélation entre les différents types de SE, tous les liens sont statistiquement très significatifs. Le plus petit coefficient est de 0,528 (voir annexe IX.c.). Ceci signifie que plus on se sent capable dans un type de SE, plus on se sent capable dans les autres types de SE.

Quel que soit l'item, la moyenne du SE est toujours supérieure à 3,5, le milieu de notre échelle. Si on regarde les scores selon les types de SE, ceux-ci sont tous supérieurs à 4. On peut donc en conclure que les participants se sentent globalement capables en évaluation. Ceci est en contraste avec le constat de Pastore, qui, dans sa revue systématique (2023), arrivait à la conclusion que les enseignants se perçoivent comme peu confiants en évaluation.

Si on compare les types de SE, à l'exception du SE AAL, les scores se trouvent tous aux environs de 4,86. De plus, si on compare les écarts-types dans chaque type de SE, à l'exception du SE AAL, ceux-ci sont entre 0,5 et 0,75. Une très grande majorité de l'échantillon est donc au-dessus du 3,5. Cela signifie que les enseignants de l'échantillon se sentent très confiants que ce soit pour l'évaluation de manière générale ou pour les actions à buts de réussite ou à buts d'apprentissage.

Dans le SE général, les étapes du cycle MoSCodEE dans lesquelles les participants se sentent le plus et le moins capables ont été déterminées. En ce qui concerne la moyenne la plus haute, aucune constatation particulière n'a été relevée. En ce qui concerne la moyenne la plus basse, il s'agit de la première étape comprenant notamment l'analyse du programme/référentiel, même si la moyenne est proche des autres. Nous nous demandons si c'est un effet du TC qui arrive en 1^{ère} secondaire en septembre 2026, une appréhension des enseignants.

En ce qui concerne le SE AAL, sa moyenne est plus basse (4,1). Son écart-type est plus grand (0,89). Cela signifie que pour ce type de SE, les participants se différencient plus entre ceux qui se sentent capables et ceux qui ne se sentent pas capable.

Ce qu'il faut en retenir :

- Les répondants se sentent capables en évaluation.
- Les participants se sentent moins capables dans les actions qui visent à mettre l'élève au centre de son apprentissage.
- Il y a davantage de variabilité dans le sentiment d'efficacité en AAL.

6.3. Constat des besoins de formations

Le sentiment du besoin de formation global est mitigé. La dispersion sur les différents échelons de l'échelle de Likert est assez grande (écart-type entre 1,28 et 1,47) ce qui semblerait vouloir dire que les enseignants n'ont pas une vision homogène sur le sujet. Cela est intéressant puisque l'effet du développement professionnel dépend de la motivation et de la disposition des professeurs à modifier leurs pratiques (Kennedy, 2016).

Le fait que les enseignants aient déclaré un léger plus grand besoin de formation sur les items AAL voudrait potentiellement dire qu'ils s'intéressent aux approches de l'évaluation mettant l'élève au centre de son apprentissage telles que l'auto-évaluation et l'évaluation par les pairs. DeLuca et ses collègues (2016) disent à ce sujet que les enseignants valorisent les opportunités d'apprentissage intégrées au travail avec mise en œuvre soutenue de nouvelles pratiques pédagogiques.

Le fait que l'item BF AOL1 (élaborer et utiliser des épreuves bien construites pour obtenir des cotes qui reflètent les acquis des élèves) possède la moyenne la plus haute parmi les items BF AOL semble mettre en avant le souhait principal des enseignants de se former pour déterminer une cote la plus représentative possible des acquis des élèves. Cette volonté de justesse montre une recherche continue d'une certaine professionnalité et équité envers les élèves.

En outre, le fait que 71% des professeurs ne ressentent pas le besoin d'être formé à l'utilisation des retours constructifs des élèves montre selon nous, une grande maîtrise de cela, ou un manque de confiance dans les capacités des élèves à faire des retours constructifs (propos tenu par un participant aux laboratoires cognitifs – voir item 33 de la partie 2 de l'annexe IV.b.), ou encore le fait que les enseignants considèrent que cela remettrait en cause la crédibilité de l'enseignant (Jorro, 2009). Or pour l'item AAL4 de la partie 1 traitant également de l'utilisation des retours constructifs des élèves, 59% des répondants déclarent des possibilités de mener l'action de « tout à fait improbable » à « légèrement improbable ». Il y a donc au minimum 30% des enseignants qui ne pratiquent pas et ne ressentent pas le besoin de s'y former.

Ce qu'il faut en retenir :

- Les enseignants voudraient être davantage formés aux pratiques liées à l'AAL.
- Une partie des enseignants ne semblent pas considérer une formation sur l'utilisation des retours des élèves sur leurs évaluations comme un levier de développement professionnel.

6.4. Constat des modalités de formations

Kennedy (2016) signale qu'actuellement, il n'y a pas d'idée bien développée de l'apprentissage des professeurs ou de la façon d'aider les enseignants à intégrer de nouvelles idées dans leur système de pratiques. L'analyse des modalités de formation est une tentative d'apporter des éléments de réponses à ceci.

La modalité de formation la plus souhaitée par les enseignants est MF7 « travail collaboratif en classe avec des collègues » (moyenne de 4,15) ce qui confirme que l'évaluation est une pratique sociale située (Pastore, 2023) et que les enseignants en sont bien conscients et en sont demandeurs. Ceci est tout de même en contraste avec le fait que, dans la partie 2 du questionnaire, 79% des enseignants ont répondu ne pas avoir besoin de se former à la collaboration pour l'évaluation. Selon nous, cela peut être interprété comme « je ne veux pas me former à la collaboration, je veux la pratiquer, en contexte ». Soussi et ses collègues (2006) affirment qu'il y a une préférence des enseignants du secondaire inférieur pour les travaux faits en collaboration entre collègues plutôt que l'utilisation d'épreuves externes. Si notre interprétation est la bonne, il s'agit d'un point important soulevé étant donné que concevoir des modalités d'évaluation en équipe est plus efficace et plus équitable (FWB, 2025). D'après Braxmeyer et ses collaborateurs (2005), 90% des enseignants sont souvent seuls pour construire et corriger des évaluations (pas d'approche collective de l'évaluation dans le projet d'établissement) et si construction ensemble il y a, c'est entre enseignants de la même discipline. Ce tout dernier point est confirmé par les réponses obtenues au MF11 et MF12. En effet, MF12 « mentorat individuel avec un enseignant d'une autre discipline » est la modalité la moins souhaitée par les enseignants (moyenne de 2,50). Il est certain que ce n'est pas le mentorat individuel qui est rejeté car le mentorat individuel avec un professeur d'une même discipline (MF11), est une formation pour laquelle 65,7% des enseignants seraient « plutôt » à « très favorables ».

Les enseignants préfèrent les modalités en présentiel par rapport aux modalités en ligne. Cela confirme les résultats de DeLuca et ses collaborateurs (2016) : les formations en présentiel sont préférées à celles en ligne.

Les enseignants préfèrent les modalités en groupe par rapport aux modalités seul. Cela confirme les résultats de Soussi et ses collègues (2006) qui informent que les 2 modalités de formations les plus voulues seraient d'une part une formation continue sur mesure c'est-à-dire liée à un projet d'école ou un groupe-discipline et d'autre part, une collaboration entre collègues.

Ce qu'il faut en retenir :

- La collaboration entre collègues en classe est la modalité préférée par les enseignants de cette recherche et, cela est justement encouragé par la FWB.
- Les enseignants préfèrent les modalités en présentiel et en groupe.
- Les enseignants ne reconnaissent pas leurs pairs d'autres disciplines comme pouvant être un levier de développement professionnel.

6.5. Constat des liens

6.5.1. Approches d'évaluation – SE

Dans nos résultats, un élément est récurrent : les clusters ne se différencient pas réellement sur le SE. Cette conclusion est due au fait que les SE, que ce soit dans les scores des types de SE ou dans les items, sont assez proches et avec des écart-types plus grands que la différence des moyennes. C'est un résultat assez étonnant au vu des éléments mis en avant dans la littérature. En effet, plusieurs chercheurs avaient conclu que le SE en évaluation et les conceptions de l'évaluation étaient étroitement liés (Kyttälä et al., 2023 ; Levy-Vered & Nasser-Abu Alhija, 2015).

Certains items ont attiré tout de même notre attention. Parmi ces items, il y a le SE AOL5 traitant la capacité à proposer un redoublement ou une réorientation. C'est l'item le plus bas en SE dans les items du type SE AOL. Dans cet item, 18 personnes ont répondu ne pas être capables (échelons 1 à 3) d'effectuer cette action. Parmi ces 18 personnes, 13 enseignants, donc plus de 70%, ressortent du cluster 2 qui est le seul cluster ayant une moyenne inférieure à 3,5 en AOL dans la première partie. De plus malgré le fait que les moyennes sont trop proches pour déclarer que les clusters se différencient réellement, ce même cluster qui se dit globalement moins capable que les autres clusters, a des moyennes plus basses que les autres clusters dans 5 des 6 items SE AOL, ce qui est loin d'être le cas dans les items SE AFL et SE AAL.

Ce dernier paragraphe montre donc un certain lien entre les conceptions et le SE mais pas de manière aussi catégorique que ce que faisait présager le cadre théorique.

6.5.2. Approches d'évaluation – BF

Nous constatons que le cluster 3 est celui qui, sur les 4 types de BF, comporte le plus de participants ressentant le moins le besoin de se former en évaluation. En effet, celui-ci renferme le plus de participants déclarant ne pas ressentir le besoin de se former à certaines étapes du cycle MoSCodEE (MOSCO1 et MOSCO2), déclarant pour 0% ressentir fortement le besoin de

se former (échelons 5 et 6) au BF AOL1 (traitant l'élaboration et l'utilisation d'épreuves bien construites pour obtenir des cotes qui reflètent les acquis des élèves) et déclarant ressentir le moins le besoin de se former en AFL et en AAL.

Or, ce cluster 3 est composé d'enseignants déclarant pratiquer assez bien des actions liées à l'AOL, moyennement des actions liées à l'AFL et rarement des actions liées à l'AAL. Apparemment, ces enseignants ne souhaitent pas développer des actions liées à d'autres approches que celles qu'ils pratiquent déjà. Or, selon DeLuca et ses collègues (2019), l'état d'esprit de l'enseignant influence les raisons d'évaluer et donc influence la culture d'apprentissage. Un enseignant qui serait fortement porté sur l'évaluation des apprentissages dont le but est de juger d'une réussite, peu porté sur l'évaluation pour l'apprentissage dont le but est de réguler son enseignement et d'informer l'élève pour qu'il développe son apprentissage et presque pas porté sur l'évaluation en tant qu'apprentissage dont le but est de faire en sorte que l'élève monitore et dirige son apprentissage est un enseignant qui semble aller à l'encontre des volontés en matière d'évaluation du Pacte pour un enseignement d'excellence et du guide sur l'évaluation paru récemment (FWB, 2025).

6.5.3. Approches d'évaluation – modalités de formations

Les clusters 1 et 2 ne se différencient pas vraiment sur les modalités de formations. En revanche, le cluster 3 a la moins bonne moyenne sur toutes les modalités de formation sauf pour MF7 pour laquelle il a la meilleure moyenne des 3 clusters (moyenne de 4,27) et pour MF11 pour laquelle il a une meilleure moyenne que le cluster 1 (respectivement 3,69 et 3,58). Il semblerait donc que le cluster 3 préfère se former en faisant du travail collaboratif en classe avec des collègues et en participant à un mentorat individuel avec un enseignant de la même discipline.

L'idée de cette comparaison modalités-clusters a pour but de vérifier si la façon d'enseigner a un lien avec la façon d'apprendre (Coombs et al., 2020). Malgré la diversité montrée au paragraphe précédent, aucun lien n'a été trouvé vu que les 2 items ont des caractéristiques différentes (ensemble vs individuel).

6.5.4. SE – BF

Trente-cinq des 39 binômes SE/BF sont liés, au minimum modérément, par des corrélations négatives statistiquement très significatives. Nous pouvons donc en conclure que SE et BF sont des notions très étroitement liées l'une à l'autre. Mais nous nous en doutions et pas seulement parce que des études l'avaient montré (Levy-Vered & Nasser-Abu Alhija, 2015). Ceci a une certaine logique vu que, plus une personne se sent capable de faire une action, moins elle devrait

ressentir le besoin de se former. Et dans l'autre sens, la logique est toujours présente : moins une personne est capable de faire une action, plus elle devrait ressentir le besoin de se former.

Mais dans les 4 autres binômes qui n'ont pas montré des liens modérés, que signifierait le fait de se sentir capable et de ressentir tout de même le besoin de se former ? Et que signifierait le fait de se sentir incapable mais de ne pas ressentir le besoin de se former ? Notre hypothèse est que cela communiquera une information sur l'importance que le participant donne à cet item. C'est pourquoi nous analyserons au point 6.5.5. les binômes SE/BF AOL, SE/BF AFL et SE/BF AAL (dont 2 des 4 binômes mentionnés).

Parmi les binômes qui nous ont interpellés, le plus marquant est le binôme SE/BF global ($r = -0,389$, statistiquement très significatif). Comme écrit dans la partie des résultats, au moins 55,2 % des participants sont des enseignants qui se disent capables de manière globale en évaluation mais qui ressentent le besoin de s'y former. Ceci signifie que plus de la moitié de l'échantillon considère l'évaluation comme importante dans leur rôle d'enseignant. Ceci est une bonne nouvelle vu que, comme dit dans l'introduction, l'évaluation qui soutient les apprentissages occupe une place centrale dans le TC qui arrive en secondaire à la rentrée 2026.

En lien avec cette arrivée, nous avons déjà traité plusieurs fois l'étape 1 du cycle MoSCodEE traitant notamment de l'analyse du référentiel/programme. Lors du lien SE/BF, nous avons calculé que plus de 33,4% des participants sont « capables mais en ressentant tout de même le besoin de s'y former ». En plus de l'importance donnée à l'analyse d'un référentiel/programme, le stress de l'arrivée du TC a pu aussi influencer ce pourcentage.

6.5.5. Approches d'évaluations – SE – BF

Les autres binômes SE/BF analysés dans les résultats apportent des informations inattendues et remarquables. Pour le binôme SE/BF AAL2 traitant les retours constructifs des élèves, il y a au minimum 21% des participants à notre enquête qui sont « capables mais en ressentant tout de même le besoin de s'y former ». Comme montré dans les résultats, ces 21% de l'échantillon représentent moins de 8% du cluster 3. Or, ce cluster est le seul ayant une moyenne très basse en AAL. Un lien est donc bien présent entre le cluster, la probabilité d'effectuer une action, et le lien SE/BF. Mais ce lien n'est pas un cas isolé.

Au minimum 13% des participants se sont déclarés « capables mais en ressentant tout de même le besoin de s'y former » pour le SE/BF AAL 6 traitant la discussion collective sur le sens de l'évaluation. Ces 13% représentent 0% du cluster 3. Le lien est à nouveau présent.

Pour 64,8% des enseignants, le besoin de se former sur les changements de pratiques évaluatives ne se fait pas ressentir. Ils s'en sentent capables pour 93,3%. Cela aussi est une bonne nouvelle pour l'arrivée du TC dans l'enseignement secondaire inférieur et les demandes formulées dans le Pacte pour un enseignement d'excellence et le guide FWB (2025). Au minimum 28,5% des participants se sont déclarés « capables mais en ressentant tout de même le besoin de s'y former » pour changer leurs pratiques évaluatives. Ces 28,5% ne représentent que 15,4% du cluster 3. Or, même si ce binôme n'est placé dans aucun des 4 types de SE/BF, la notion de changer ses pratiques évaluatives a une certaine résonance avec l'AFL. Or le cluster 3 est celui qui a le score le plus bas en AFL parmi les clusters. Un lien est donc à nouveau présent.

7. Conclusion, limites et perspectives

S'il faut retenir quelque chose de ce travail, c'est que les buts poursuivis dans les évaluations influenceraient le lien entre SE et BF. En d'autres termes, si un enseignant évalue avec des buts d'apprentissage (comme c'est le cas dans l'AFL par exemple), il donnera de l'importance aux actions permettant d'atteindre ces buts et donc voudra davantage se former à cela (ou rejettera moins la proposition), même s'il se sent compétent en la matière.

Tout cela implique de faire évoluer les conceptions des enseignants pour qu'ils considèrent comme très porteuses, très importantes, des actions comme l'auto-évaluation ou l'évaluation par les pairs. Sans cela, même en se sentant incapable d'effectuer ces actions, ils ne ressentiront pas le besoin de s'y former. Donc une formation n'aurait pas beaucoup d'impact vu qu'il faut partir des préoccupations des enseignants (Jorro, 2016). En faisant évoluer ces conceptions, même un enseignant qui se croirait capable sans l'être réellement, dû à l'effet Dunning de « méta-ignorance » (2011, cité par Kruse et al., 2020), ressentirait le besoin de s'y former et donc recevra un effet bénéfique de cette formation.

De plus, il faudrait encourager la collaboration entre les collègues de la même discipline, comme voulu par des enseignants du secondaire inférieur de la FWB, pour développer des systèmes d'évaluations pertinents et efficaces. Comme déjà mentionné, les formations pensées en dispositif meso (école) et s'appliquant entre « amis critiques » (Crahay, 2009) sont de l'apprentissage actif, reconnu comme pratique efficace en formation par Blank et de las Alas (2009, cités par Kennedy, 2016). Cela est appuyé par le guide sur l'évaluation publié récemment (FWB, 2025) qui nous confirme que concevoir des modalités d'évaluation en équipe est plus efficace et plus équitable.

Ce travail comporte des limites telles que l'échantillon qui n'est pas représentatif et qui est trop petit pour que certains traitements statistiques soient fiables. Pour la recherche future, il serait donc bon de déterminer si les résultats de la présente étude ressortent dans un échantillon plus large et représentatif. Cela permettrait de vérifier si les traitements statistiques montrent une plus grande fiabilité. Dans le cas contraire, cela appuierait la présence d'influences, multiples et diverses selon la personne, dans les choix d'action à l'intérieur même d'une approche.

Les formulations des items de cette recherche ont été relues plusieurs fois, que ce soit par nous-mêmes ou par les participants du laboratoire cognitif. Cependant, il reste sûrement des formulations perfectibles. Dans le cas d'une reprise de ce questionnaire, il faudrait à nouveau questionner les formulations, de préférence par des experts, pour que les buts visés soient bien identifiés à la lecture. Il serait également bon de requestionner ces liens dans quelques années pour voir si le TC et les formations à l'évaluation ont fait évoluer les conceptions des enseignants.

Une autre limite de notre questionnaire est que les données traitées dans ce travail sont des données auto-rapportées ce qui ne traduit pas réellement ce qui est fait effectivement dans les salles de classe. Effectuer une recherche sur les pratiques évaluatives effectives des enseignants dans leurs classes en FWB (entretiens, observations, ...) permettrait de lister ce qui est réellement réalisé en classe.

Selon les traitements statistiques choisis et utilisés, des résultats différents peuvent être obtenus. Qui plus est, il existe une part de subjectivité quant à l'interprétation des chercheurs et de ce qu'ils mettent en avant ou pas.

Enfin, ce mémoire s'est conclu sur l'importance de faire évoluer les conceptions mais ce mémoire ne donne pas d'explication sur la manière d'y arriver. Une recherche supplémentaire est donc nécessaire pour comprendre comment faire évoluer les conceptions des enseignants.

8. Bibliographie

- Allal, L., & Dauvisis, M.-C. (2019). Lectures croisées : l'activité évaluative, ses acteurs et ses contextes. Dans A. Jorro & N. Droyer (dirs.), *L'évaluation, levier pour l'enseignement et la formation* (pp. 159-169). De Boeck Supérieur.
- Athias, F., Morellato, M., & Gruson, B. (2020). Pour mieux enseigner, il faut s'appuyer sur les résultats de la recherche. Dans Collectif « Didactique pour Enseigner ». *Enseigner, ça s'apprend* (pp. 155-171). Retz
- Bandura, A. (1997). Self-efficacy: The exercise of control. W.H. *Freeman*, 1-610. https://www.academia.edu/28274869/Albert_Bandura_Self_Efficacy_The_Exercise_of_Control_W_H_Freeman_and_Co_1997_pdf
- Baron, J., & Fagnant, A. (2025). Comment appréhender les compétences évaluatives des enseignants ? Adaptation d'un outil au contexte de l'enseignement secondaire supérieur belge francophone. *Évaluer. Journal international de recherche en éducation et formation*, 11(1), 29-48. <https://doi.org/10.48782/e-jiref-11-1-29>
- Béland, S., Cousineau, D., & Loye, N. (2017). Utiliser le coefficient omega de McDonald à la place de l'alpha de Cronbach. *McGill Journal of Education / Revue des sciences de l'éducation de McGill*, 52(3), 791–804. <https://doi.org/10.7202/1050915ar>
- Béland, S., & Michelot, F. (2020). Une note sur le coefficient oméga (ω) et ses déclinaisons pour estimer la fidélité des scores. *Mesure et évaluation en éducation*, 43(3), 103–122. <https://doi.org/10.7202/1084526ar>
- Braxmeyer, N., Guillaume, J.-C., Lairez, Th., & Lévy, J.-F. (2005). Les pratiques d'évaluation des enseignants en collège. *Ministère de l'Éducation nationale, de l'Enseignement supérieur et de la Recherche, Direction de l'évaluation et de la prospective [DEP]*. [https://archives-statistiques-depp.education.gouv.fr/Default/doc/SYRACUSE/10844/les-pratiques-d-evaluation-des-enseignants-au-college-nicole-braxmeyer-jean-claude-guillaume-et-jean-FRes enseignants au collège / Nicole Braxmeyer, Jean-Claude Guillaume et Jean-François Lévy - Portail des publications de la DEPP](https://archives-statistiques-depp.education.gouv.fr/Default/doc/SYRACUSE/10844/les-pratiques-d-evaluation-des-enseignants-au-college-nicole-braxmeyer-jean-claude-guillaume-et-jean-FRes%20enseignants%20au%20coll%C3%A8ge%20-%20Nicole%20Braxmeyer,%20Jean-Claude%20Guillaume%20et%20Jean-Fran%C3%A7ois%20L%C3%A9vy%20-%20Portail%20des%20publications%20de%20la%20DEPP)
- Campanale, F., Dejemeppe, X., Vanhulle, S., & Saussez, F. (2010). Dispositifs d'autoévaluation socialisée en formation : une contrainte pervertissante ou une opportunité transformatrice ? Dans L. Paquay, C. Van Nieuwenhoven & P. Wouters (dirs). *L'évaluation, levier du développement professionnel ?* (pp. 193-205). De Boeck.
- Christoforidou, M., & Kyriakides, L. (2021). Developing teacher assessment skills: the impact of the dynamic approach to teacher professional development. *Stud. Educ. Eval.* 70, 1-12. <https://doi.org/10.1016/j.stueduc.2021.101051>
- Coombs, A. DeLuca, C., LaPointe-McEwan, D., & Chalas, A. (2018). Changing approaches to classroom assessment: An empirical study across teacher career stages. *Teaching and Teacher Education*, 71, 134-144. <https://doi.org/10.1016/j.tate.2017.12.010>
- Coombs, A., Rickey, N., DeLuca, C., & Liu, S. (2022). Chinese teachers' approaches to classroom assessment. *Educational Research for Policy and Practice*, 21, 1-18. <https://doi.org/10.1007/s10671-020-09289-z>
- Crahay, M. (2009). Articuler l'évaluation en classe et le pilotage des systèmes, est-ce possible ? Dans L. Mottier Lopez & M. Crahay (dirs), *Évaluations en tension* (pp. 233-251). De Boeck.

- Crahay, M. (2019). Introduction. Dans M. Crahay (dir), *Peut-on lutter contre l'échec scolaire ?* (pp. 21-51). De Boeck Supérieur.
- Crahay, M., Mottier Lopez, L., & Marcoux, G. (2019). L'évaluation des élèves : Docteur Jekyll and Mister Hyde de l'enseignement. Dans M. Crahay (dir), *Peut-on lutter contre l'échec scolaire ?* (pp. 357-425). De Boeck Supérieur.
- Crahay, M. (2019). Du redoublement à la régulation des apprentissages. Dans M. Crahay (dir), *Peut-on lutter contre l'échec scolaire ?* (pp. 427-446). De Boeck Supérieur.
- DeLuca, C., LaPointe-McEwan, D., & Luhanga, U. (2016). Approaches to Classroom Assessment Inventory: A New Instrument to Support Teacher Assessment Literacy. *Educational Assessment*, 21(4), 248-266. <https://doi.org/10.1080/10627197.2016.1236677>
- DeLuca, C., Valiquette, A., Coombs, A., LaPointe-McEwan, D., & Luhanga, U. (2018). Teachers' approaches to classroom assessment: a large-scale survey. *Assessment in Education: Principles, Policy & Practice*, 25(4), 355-375. <https://doi.org/10.1080/0969594X.2016.1244514>
- DeLuca, C., Coombs, A., MacGregor, S., & Rasooli, A. (2019a). Toward a Differential and Situated View of Assessment Literacy: Studying Teachers' Responses to Classroom Assessment Scenarios. *Front. Educ.*, 4(94), 1-10. <https://doi.org/10.3389/educ.2019.00094>
- DeLuca, C., Coombs, A., & LaPointe-McEwan, D. (2019b). Assessment mindset: Exploring the relationship between teacher mindset and approaches to classroom assessment. *Studies in Educational Evaluation*, 61, 159-169. <https://doi.org/10.1016/j.stueduc.2019.03.012>
- DeLuca, C., Schneider, C., Coombs, A., Pozas, M., & Rasooli, A. (2020). A cross-cultural comparison of German and Canadian student teachers' assessment competence. *Assessment in Education: Principles, Policy & Practice*, 27(1), 26-45. <https://doi.org/10.1080/0969594X.2019.1703171>
- DeLuca, C., Rickey, N., & Coombs, A. (2021). Exploring assessment across cultures: Teachers' approaches to assessment in the U.S., China, and Canada. *Cogent Education*, 8(1), 1-26. <https://doi.org/10.1080/2331186X.2021.1921903>
- Demonty, I., Jaegers, D., & Dupont, V. (2024). *Construction et analyse de tests et de questionnaires*. Notes de cours non publiées, Université de Liège, Liège.
- Detroz, P., Malay, L., & Crahay, V. (2020). Une démarche structurée pour définir quelques conseils en vue de limiter l'impact de la pandémie sur l'évaluation de nos étudiants. *Évaluer. Journal international de recherche en éducation et formation, Numéro Hors-série 1*, 97-100.
- Detroz, P. (2023). *Questions d'évaluation*. Notes de cours non publiées, Université de Liège, Liège.
- Duru-Bellat, M. (2009). Évaluation externe du système éducatif et ressenti des acteurs. Dans L. Mottier Lopez & M. Crahay (dirs), *Évaluations en tension* (pp. 47-59). De Boeck.
- Eufemia, F. (2012). *the relationship between formative assessment and teachers' self-efficacy*. [Thèse de doctorat, université de Floride]. [https://ufdcimages.uflib.ufl.edu/UF/E0/04/39/38/00001/EUFEMIA .pdf](https://ufdcimages.uflib.ufl.edu/UF/E0/04/39/38/00001/EUFEMIA.pdf) UNIVERSITY OF FLORIDA THESIS OR DISSERTATION FORMATTING TEMPLATE

- Fagnant, A., & Goffin, C. (2021). Note de synthèse – Comment se conjuguent « évaluation » et « développement professionnel des enseignants » dans les écrits scientifiques de l'Admee. *Évaluer. Journal international de recherche en éducation et formation*, 7(3), 5-26.
- Fagnant, A. (2023). Les pratiques d'évaluation en classe : des compétences professionnelles pour soutenir l'apprentissage des élèves. *Cnesco-Cnam*, 1-53. https://www.cnesco.fr/wp-content/uploads/2023/08/Cnesco-CC-Eval_FAGNANT.pdf
- Fédération Wallonie-Bruxelles [FWB]. (s.d.). Pacte pour un enseignement d'excellence. Enseignement.be consulté le 5/04/2026 sur <https://pactepourunenseignementdexcellence.cfwb.be/mesures/le-renforcement-de-la-formation-professionnelle-continue/>
- Fédération Wallonie-Bruxelles [FWB]. (2025). Document : Circulaire 9566 du 21/08/2025 Circulaire générale relative à l'organisation de l'enseignement secondaire ordinaire et à la sanction des études Année scolaire 2025-2026. Consulté sur https://gallilex.cfwb.be/sites/default/files/circulaires/2025-09/53066_0001.pdf
- Fédération Wallonie-Bruxelles [FWB]. (2021). Décret portant le Livre 6 du Code de l'enseignement fondamental et de l'enseignement secondaire et portant le titre relatif à la formation professionnelle continue des membres de l'équipe éducative des écoles et des membres du personnel de l'équipe pluridisciplinaire des Centres PMS. Centre de documentation administrative. https://www.gallilex.cfwb.be/document/pdf/49421_000.pdf
- Fédération Wallonie-Bruxelles. (2025). Guide « L'évaluation au service des apprentissages de tous les élèves dans le cadre du tronc commun ». Auteur
- Fresson, M., & Dardenne, B. (2019). Les redoublants, victimes de stéréotypes dévalorisants. Dans M. Crahay (dir), *Peut-on lutter contre l'échec scolaire ?* (pp. 231-260). De Boeck Supérieur.
- Graham, L. J., de Bruin, K., Lassig, C., & Spandagou, I. (2021). A scoping review of 20 years of research on differentiation : investigating conceptualisation, characteristics, and methods used. *Review of education*, 9(1), 161-198. https://www.researchgate.net/publication/347774299_A_scoping_review_of_20_years_of_research_on_differentiation_investigating_conceptualisation_characteristics_and_methods_used
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77 (1), 81-112. <https://www.jstor.org/stable/4624888>
- Hu, L.-T., & Bentler, P.M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, 6(1), 1-55. <https://doi.org/10.1080/10705519909540118>
- Jager, L., Denessen, E., Cillessen, A., & Meijer, P.C. (2022). Capturing instructional differentiation in educational research : investigating opportunities and challenges. *Educational Research*, 64(2), 224-241. <https://doi.org/10.1080/00131881.2022.2063751>
- Jorro, A. (2009). L'évaluation comme savoir professionnel. Dans L. Mottier Lopez & M. Crahay (dirs), *Évaluations en tension* (pp. 219-231). De Boeck.

- Jorro, A. (2016). Se former à l'activité évaluative. *Éducation permanente*, 208, 53-64. <https://cnam.hal.science/hal-02290580>
- Kennedy, M. (2016). How does professional development improve teaching? *Review of Educational Research*, 86(4), 945-980. <https://doi.org/10.3102/0034654315626800>
- Kruse, L. M., Impellizeri, W., Witherel, C. E., & Sondergeld, T. A. (2020). Evaluating the impact of an assessment course on preservice teachers' classroom assessment literacy and self-efficacy. *Mid-Western educational*, 32(2), 106-132. https://scholarworks.bgsu.edu/mwer/vol32/iss2/3/valuating_the_impact_of_an_Assessment_Course_on_Preservice_Teachers_'_by_Lance_Kruse,_Whitney_Impellizeri_et_al.
- Kyttälä, M., Björn, P. M., Rantamäki, M., Närhi, V., & Aro, M. (2023). Exploring pre-service special needs teachers' assessment conceptions and assessment self-efficacy. *European Journal of Special Needs Education*, 38(1), 63-78, <https://doi.org/10.1080/08856257.2021.2021871>
- Lafontaine, D., Soussi, A., & Nidegger, C. (2009). Évaluations internationales et/ou épreuves nationales : tensions et changement de pratiques. Dans L. Mottier Lopez & M. Crahay (dirs), *Évaluations en tension* (pp. 61-80). De Boeck.
- Lafontaine, D., Baye, A., & Monseur, C. (2019). Redoublement, efficacité et équité : l'éclairage des enquêtes internationales. Dans M. Crahay (dir), *Peut-on lutter contre l'échec scolaire ?* (pp. 111-153). De Boeck Supérieur.
- Laveault, D., & Grégoire, J. (2014). *Introduction aux théories des tests en psychologie et en sciences de l'éducation*. 3e éd. Méthodes en sciences humaines. De Boeck supérieur.
- Levy-Vered, A., & Nasser-Abu Alhija, F. (2015). Modelling beginning teachers' assessment literacy: the contribution of training, self-efficacy, and conceptions of assessment. *Educational Research and Evaluation*, 21(5-6), 378-406, <https://doi.org/10.1080/13803611.2015.1117980>
- Mertler, C. (2004). Secondary Teachers' Assessment Literacy: Does Classroom Experience Make A Difference ? *American Secondary Education*, 33(1), 49-64. <https://www.jstor.org/stable/41064623>
- Monseur, C. (2022). *Notions de statistiques appliquées à l'éducation*. Notes de cours non publiées, Université de Liège, Liège.
- Monseur, C. (2023). Syllabus d'Introduction à la psychométrie. Presses Universitaires de Liège.
- Mottier Lopez, L., & Van Nieuwenhoven, C. (2009). Co-élaboration d'un dispositif d'évaluation entre formateurs : quelles relations de régulation au sein de la communauté de pratique ? Dans L. Mottier Lopez & M. Crahay (dirs), *Évaluations en tension* (pp. 181-199). De Boeck.
- Mottier Lopez, L. (2022, 23 novembre). *ENJEUX DE L'ÉVALUATION EN CLASSE – Lucie Mottier Lopez – CC évaluation en classe – 23 novembre* [vidéo]. YouTube. <https://www.youtube.com/watch?v=795YCJovKOE>
- Navarro, D., Foxcroft, D., & Meunier, J.-M. (2020). Apprentissage des statistiques avec Jamovi. *HAL open science*, 1-518. <https://hal.science/hal-02335912v2>
- Pastore, S. (2023). Teacher assessment literacy: a systematic review. *Frontiers in Education*, 8. <https://doi.org/10.3389/feduc.2023.1217167>

- Pasquini, R. (2021). *Quand la note devient constructive. Évaluer pour certifier et soutenir les apprentissages*. Hermann.
- Pasquini, R., & Morales Villabona, F. (2022). Comment appréhender une littératie en évaluation pour décrire et comprendre le développement des compétences des enseignants ? *Canadian Journal of Education / Revue canadienne de l'éducation*, 45(2), 512–554. <https://doi.org/10.53967/cje-rce.v45i2.51710>
- Quittre, V., Collinge, A., Crepin, F., Schillings, P., & Fagnant, A. (2025). *TALIS 2024—Enseigner aujourd'hui*. <https://orbi.uliege.be/handle/2268/336950>
- Raulin, D., & Lebeaume, J. (2019). Pour un élève, acteur de l'évaluation. Dans A. Jorro & N. Droyer (dirs.), *L'évaluation, levier pour l'enseignement et la formation* (pp. 85-101). De Boeck Supérieur.
- Rothinam, N. (2023). The relationship between self-efficacy and attitude towards assessment literacy among teachers. *European Proceedings of Educational Sciences*, 499-511. https://www.researchgate.net/publication/374133077_The_Relationship_Between_Self-Efficacy_and_Attitude_Towards_Assessment_Literacy_Among_Teachers_Relationship_Between_Self-Efficacy_and_Attitude_Towards_Assessment_Literacy_Among_Teachers
- Soussi, A., Ducrey, F., Ferrez, E., Nidegger, C., & Viry, G. (2006). Pratiques d'évaluation : ce qu'en disent les enseignants (à l'école obligatoire et dans l'enseignement postobligatoire général). *Evalepcopo. Service de la recherche en éducation*, 1-123. https://www.syneval.ch/database/pdf/Soussi_et_al_2006_Evalepcopo.pdf
- Stiggins, R. J. (1991). Assessment literacy. *The Phi Delta Kappan*, 72(7), 534-539. <https://www.jstor.org/stable/20404455>
- Stiggins, R. (2017). *The perfect assessment system*. ASCD.
- Täschner, J., Dicke, T., Reinhold, S., & Holzberger, D. (2024). “Yes, I Can!” A Systematic Review and Meta-Analysis of Intervention Studies Promoting Teacher Self-Efficacy. *Review of Educational Research*, XX(X), 1-50. <https://doi.org/10.3102/00346543231221499>
- Tchibozo, G. (2019). Introduction pratique aux méthodes quantitatives en Sciences de l'Éducation et de la Formation. *Atramenta*, 1-768. https://www.researchgate.net/publication/335587663_Introduction_pratique_aux_methodes_quantitatives_en_Sciences_de_l%27education_et_de_la_formationuction_pratique_aux_méthodes_quantitatives_en_Sciences_de_l%27éducation_et_de_la_formation
- Vanhulle, S. (2009). Évaluer la professionnalité émergente des futurs enseignants : un pari entre cadres contraignants et tensions formatives. Dans L. Mottier Lopez & M. Crahay (dirs.), *Évaluations en tension* (pp. 165-180). De Boeck.
- Van Orman, D., Gotch, C., & Carbonneau, K. (2024). Preparing teacher candidates to assess for learning: A systematic review. *Educational Research*, 20(10). 1 –37. <https://doi.org/10.3102/00346543241233015>
- Viau, R. (1998). Les perceptions de l'élève : Sources de sa motivation dans les cours de français. *Québec français*, 110, 45-47. <https://www.erudit.org/fr/revues/qf/1998-n110-qf1203783/56310ac/>
- Willis, J., Adie, L., & Klenowski, V. (2013). Conceptualising teachers' assessment literacies in an era of curriculum and assessment reform. *The Australian Educational Researcher*, 40(2), 241-256. <https://doi.org/10.1007/s13384-013-0089-9>

- Wyatt-Smith, C., Klenowski, V., & Gunn, S. (2010). The centrality of teachers' judgement practice in assessment: a study of standards in moderation. *Assessment in Education: Principles, Policy & Practice*, 17(1), 59-75.
<https://doi.org/10.1080/09695940903565610>
- Xu, Y., & Brown, G.T.L. (2016). Teacher assessment literacy in practice: A reconceptualization. *Teaching and Teacher Education*, 58, 149-162.
<https://www.sciencedirect.com/science/article/abs/pii/S0742051X16300907?via%3Dihub>

9. Annexes

I. Fiche IA

Dans ce travail de recherche, l'utilisation des IA Génératives pour ...	Utilisation autorisée	Utilisation de l'étudiant	Nom de/des IA générative(s) utilisée(s)
Aider à la rédaction du plan et de la structure du travail	✓	/	/
Vérifier l'orthographe et la grammaire	✓	/	/
Reformuler le contenu de votre travail ou améliorer le style	✓	/	/
Traduire une autre langue	✓	Clélia : oui Riccardo : oui	Deepl / google lens Traducteur d'adobe
Aider à la planification et gestion de projet	✓	/	/
Obtenir des informations générales sur un sujet, expliquer et approfondir la compréhension de concepts	✓	Clélia : oui Riccardo : oui	Perplexity Copilot
Faire un brainstorming et évaluer ses propres idées, pour trouver d'autres perspectives ou des contre-arguments	✓	/	/
Aider à la recherche documentaire	✓	Clélia : oui Riccardo : oui	Perplexity ChatGPT, Copilot
Aider à la synthèse de la littérature	✓	/	/
Aider à la formulation de questions de recherche et d'hypothèses	✓	/	/
Aider à l'analyse de données (identification d'analyses à réaliser)	✓	Clélia : oui Riccardo : oui	Perplexity ChatGPT, Copilot
Aider à l'interprétation des résultats	✓	/	/
Aider au respect des normes APA de la liste des références	✓	/	/
Aider à la programmation de code et au débogage	✓	/	/
Générer un feedback critique et révision	✓	/	/
Créer du contenu multimédia, par exemple des images, des vidéos, des animations ou du son (mais toujours expliquer l'utilisation des IA)	✓	/	/

Générer le contenu du travail à partir de mots clés, d'un thème ou d'une question de départ sans vérification et appropriation personnelle	×		
Reproduire ou reformuler du contenu existant sans citer la source (plagiat)	×		
Falsifier, créer des données	×		
Générer des références bibliographiques sans en contrôler l'existence ou l'exactitude	×		
Utiliser dans une IA des données non anonymisées issues de l'étude (participants, résultats, verbatim, etc.)	×		

Je, soussignée Robaye Clélia, déclare sur l'honneur avoir complété la grille avec exactitude et honnêteté, en toute bonne foi.

Je, soussigné Mazzetto Riccardo, déclare sur l'honneur avoir complété la grille avec exactitude et honnêteté, en toute bonne foi.

Légende :

= L'utilisation des IA Génératives est autorisée (ce qui ne veut pas dire qu'elle est encouragée. Voir ci-dessous)

+ = L'utilisation des IA Génératives n'est pas autorisée, les infractions entraîneront des sanctions.

n.a = Non-applicable pour ce cours

L'étudiant rempli sa colonne par « Oui » ou par « Non » en fonction de s'il a utilisé ou non une IA générative.

***Attention :** le fait que certaines pratiques soient autorisées ne signifie pas que l'on attende ou que l'on encourage l'utilisation d'IA Génératives pour cette évaluation. Dans de nombreuses situations, vous obtiendrez probablement de meilleurs résultats sans utiliser d'IA Génératives.

Cette grille est une adaptation du « Tableau d'utilisation des IA Génératives » de la Faculté des Arts et des Sciences Sociales (FASoS) de l'Université de Maastricht.

II. Responsabilité partagée

L'ensemble du travail a été relu et critiqué par les deux étudiants.

Cette version du travail est un consensus entre les 2 étudiants.

	Riccardo	Clélia
Organisation du travail selon les normes demandées par le règlement facultaire		X
Fil rouge du travail et conception de l'organisation interne de chaque partie du travail	X	
Remerciements	X	X
Table des matières	X	
Table des figures	X	
Table des annexes	X	
1 Introduction générale	X	
2.1 Les étapes de conception de l'évaluation		X
2.2 Teacher assessment literacy	X	
2.3 Self-efficacy en évaluation	X	
2.4 Formation et besoins des enseignants		X
3 Question de recherche et objectifs		X
4.1 Méthodologie : introduction		X
4.2 Méthodologie : création du questionnaire et analyses statistiques envisagées	X	X
5. Présentation des résultats : analyse	X	X
6. Interprétation et discussion	X	X
7. Conclusion		X
8. Bibliographie		X
9. Annexes	X	
Lectures pour ce travail	30%	70%
Informations issues de nos pré-mémoires	50%	50%
Création du questionnaire en tant que tel	50%	50%

Implémentation du questionnaire dans la plateforme universitaire	90%	10%
Laboratoires cognitifs effectués	43%	57%
Dossier pour la CVE	60%	40%
Participations aux séances de rédaction du mémoire et de défense orale du mémoire	50%	50%

III. Questionnaire (avec les codes utilisés dans le texte)

a. Partie 1 du questionnaire

Voici un scénario :

"Vous donnez une évaluation sommative de fin de chapitre sous format papier/crayon, incluant des aménagements raisonnables pour les élèves à besoins spécifiques. 16 des 24 élèves échouent."

En réaction à cette situation, en tant qu'enseignant(e) et selon votre contexte, indiquez la probabilité d'effectuer les actions suivantes dans leur entièreté.

1 : Tout à fait improbable ;

2 : Plutôt improbable ;

3 : Légèrement improbable ;

4 : Légèrement probable ;

5 : Plutôt probable ;

6 : Tout à fait probable

AOL1	Inscrire le résultat de chaque élève dans le cahier de cotes et passer à la matière suivante.	1, 2, 3, 4, 5 ou 6
AFL1	Faire une séance de correction pour ensuite faire repasser une épreuve similaire.	1, 2, 3, 4, 5 ou 6
AAL1	Après avoir laissé du temps aux élèves pour prendre connaissances de leurs erreurs et pour poser des questions, demander aux élèves s'ils veulent repasser l'interrogation. Si oui, discuter de la date de passation.	1, 2, 3, 4, 5 ou 6
AFL2	Faire une activité de remédiation en classe en groupant les élèves sur base des compétences acquises ou non lors de l'évaluation.	1, 2, 3, 4, 5 ou 6
AOL2	Inviter les élèves, qui ont échoué à l'évaluation, à revoir le chapitre chez eux.	1, 2, 3, 4, 5 ou 6
AAL2	Créer des zones en classe liées aux compétences non acquises et laisser les élèves se grouper selon ce qu'ils estiment devoir retravailler.	1, 2, 3, 4, 5 ou 6
AOL3	Lors du prochain chapitre, utiliser les cotes obtenues dans cette évaluation pour motiver les élèves à augmenter leur investissement.	1, 2, 3, 4, 5 ou 6

AFL3	Lors du prochain chapitre, modifier sa propre façon d'évaluer pour augmenter l'investissement des élèves.	1, 2, 3, 4, 5 ou 6
AAL3	Demander aux élèves de réfléchir individuellement à leur investissement à l'aide de questions guides sur le temps consacré à travailler le chapitre, l'étalement des séances d'étude et la participation en classe. Leur proposer ensuite de se fixer des objectifs personnels pour la prochaine évaluation. Être disponible pour les élèves qui voudraient être assistés dans cette tâche.	1, 2, 3, 4, 5 ou 6
AOL4	Vérifier si la cote finale obtenue par chaque élève représente bien le niveau d'acquisition global de la matière, selon ce que chacun a produit dans l'évaluation.	1, 2, 3, 4, 5 ou 6
AFL4	Vérifier si les questions du contrôle évaluent les compétences visées.	1, 2, 3, 4, 5 ou 6
AAL4	Demander aux élèves d'énoncer une critique de l'évaluation (cela peut même se faire de façon écrite et anonyme).	1, 2, 3, 4, 5 ou 6
AOL5	Discuter des cotes avec l'équipe pédagogique pour vérifier si le phénomène se répète également chez eux.	1, 2, 3, 4, 5 ou 6
AFL5	Discuter avec des collègues de la même discipline des compétences non acquises en vue d'échanger sur les pratiques pédagogiques.	1, 2, 3, 4, 5 ou 6
AAL5	Permettre aux élèves de discuter et échanger sur leur façon de se préparer aux évaluations.	1, 2, 3, 4, 5 ou 6
AOL6	Renseigner l'élève sur le fait qu'il a réussi ou échoué à l'évaluation en écrivant uniquement la cote globale, sans faire de commentaire.	1, 2, 3, 4, 5 ou 6
AFL6	Écrire sur l'évaluation de chaque élève des retours personnalisés sur les types d'erreurs effectués et des conseils sur la manière de les travailler.	1, 2, 3, 4, 5 ou 6
AAL6	À l'aide de la copie corrigée de son évaluation et d'un correctif, faire réaliser en classe, à chaque élève, une auto-analyse sur la compréhension, l'application et le transfert des concepts dans divers types de tâches. Être disponible pour les élèves qui demandent de l'aide.	1, 2, 3, 4, 5 ou 6

AOL7	Rappeler l'importance de la réussite scolaire pour leur avenir.	1, 2, 3, 4, 5 ou 6
AFL7	Rappeler que les évaluations servent à savoir où ils en sont dans l'acquisition de leurs compétences mais qu'ils peuvent utiliser ces informations pour progresser.	1, 2, 3, 4, 5 ou 6
AAL7	Demander aux élèves de réfléchir seuls à l'utilisation de l'évaluation pour améliorer leurs apprentissages. Se servir des réflexions personnelles des élèves pour lancer une discussion de réflexion commune.	1, 2, 3, 4, 5 ou 6

b. Partie 2 du questionnaire

Voici des actions liées à votre pratique quotidienne de l'évaluation. Pour chaque action, il vous est demandé de vous positionner sur 2 aspects :

- **Dans votre pratique quotidienne, à quel point vous vous sentez capable d'effectuer chaque action listée ?**

1 : Pas du tout capable ;

2 : Très peu capable ;

3 : Peu capable ;

4 : Assez capable ;

5 : Très capable ;

6 : Entièrement capable.

- **À quel point ressentez-vous le besoin de vous former sur celles-ci ?**

1 : Pas du tout ;

2 : Presque pas ;

3 : Très peu ;

4 : Un peu ;

5 : Beaucoup ;

6 : Tout à fait.

SE GLOBAL	J'ai confiance en mes compétences d'évaluateur.	1, 2, 3, 4, 5 ou 6
BF GLOBAL	J'ai besoin de formations pour développer mes compétences d'évaluateur.	1, 2, 3, 4, 5 ou 6

MOSCO SE1P1	J'ai confiance en mes capacités à analyser un programme ou un référentiel pour identifier les éléments à évaluer.	1, 2, 3, 4, 5 ou 6
----------------	---	-----------------------

MOSCO BF1P1	J'ai besoin de formations pour m'aider à analyser un programme ou un référentiel pour identifier les éléments à évaluer.	1, 2, 3, 4, 5 ou 6
----------------	--	-----------------------

MOSCO SE1P2	J'ai confiance en mes capacités à créer une triple concordance entre : objectifs visés, méthodes d'enseignement et évaluations.	1, 2, 3, 4, 5 ou 6
----------------	---	-----------------------

MOSCO BF1P2	J'ai besoin de formations pour m'aider à créer une triple concordance entre : objectifs visés, méthodes d'enseignement et évaluations.	1, 2, 3, 4, 5 ou 6
----------------	--	-----------------------

MOSCO SE2	J'ai confiance en mes capacités à choisir la bonne modalité d'évaluation (travaux, QCM, questions ouvertes, en groupe, seul, ...) suivant les avantages et les inconvénients de chaque modalité.	1, 2, 3, 4, 5 ou 6
--------------	--	-----------------------

MOSCO BF2	J'ai besoin de formations pour m'aider à choisir la bonne modalité d'évaluation (travaux, QCM, questions ouvertes, en groupe, seul, ...) suivant les avantages et les inconvénients de chaque modalité.	1, 2, 3, 4, 5 ou 6
--------------	---	-----------------------

MOSCO SE3	J'ai confiance en mes capacités à planifier mes évaluations.	1, 2, 3, 4, 5 ou 6
--------------	--	-----------------------

MOSCO BF3	J'ai besoin de formations pour m'aider à planifier mes évaluations.	1, 2, 3, 4, 5 ou 6
--------------	---	-----------------------

MOSCO SE4P1	J'ai confiance en mes capacités, dans les évaluations, à donner plus de poids aux questions portant sur les objectifs principaux.	1, 2, 3, 4, 5 ou 6
----------------	---	-----------------------

MOSCO BF4P1	J'ai besoin de formations pour m'aider, dans les évaluations, à donner plus de poids aux questions portant sur les objectifs principaux.	1, 2, 3, 4, 5 ou 6
----------------	--	-----------------------

MOSCO SE4P2	J'ai confiance en mes capacités à rédiger des questions de qualité.	1, 2, 3, 4, 5 ou 6
----------------	---	-----------------------

MOSCO BF4P2	J'ai besoin de formations pour m'aider à rédiger des questions de qualité.	1, 2, 3, 4, 5 ou 6
----------------	--	-----------------------

MOSCO SE5	J'ai confiance en mes capacités à communiquer clairement aux élèves les informations utiles pour leur préparation à l'évaluation.	1, 2, 3, 4, 5 ou 6
MOSCO BF5	J'ai besoin de formations pour m'aider à communiquer clairement aux élèves les informations utiles pour leur préparation à l'évaluation.	1, 2, 3, 4, 5 ou 6

MOSCO SE6	J'ai confiance en mes capacités à m'informer, au travers de différents types d'évaluations, des apprentissages de mes élèves.	1, 2, 3, 4, 5 ou 6
MOSCO BF6	J'ai besoin de formations pour m'aider à m'informer, au travers de différents types d'évaluations, des apprentissages de mes élèves.	1, 2, 3, 4, 5 ou 6

MOSCO SE7P1	J'ai confiance en mes capacités à corriger une évaluation sur base d'une grille de correction.	1, 2, 3, 4, 5 ou 6
MOSCO BF7P1	J'ai besoin de formations pour m'aider à corriger une évaluation sur base d'une grille de correction.	1, 2, 3, 4, 5 ou 6

MOSCO SE7P2	J'ai confiance en mes capacités à créer une grille de correction.	1, 2, 3, 4, 5 ou 6
MOSCO BF7P2	J'ai besoin de formations pour m'aider à créer une grille de correction.	1, 2, 3, 4, 5 ou 6

MOSCO SE8	J'ai confiance en mes capacités à faire un retour constructif aux élèves à la suite d'une évaluation.	1, 2, 3, 4, 5 ou 6
MOSCO BF8	J'ai besoin de formations pour m'aider à faire un retour constructif aux élèves à la suite d'une évaluation.	1, 2, 3, 4, 5 ou 6

MOSCO SE9	J'ai confiance en mes capacités à porter un regard réflexif sur mes évaluations.	1, 2, 3, 4, 5 ou 6
MOSCO BF9	J'ai besoin de formations pour m'aider à porter un regard réflexif sur mes évaluations.	1, 2, 3, 4, 5 ou 6

SE AAL1	J'ai confiance en mes capacités à engager les élèves dans des dispositifs d'auto-évaluation et d'évaluation entre pairs.	1, 2, 3, 4, 5 ou 6
BF AAL1	J'ai besoin de formations pour m'aider à engager les élèves dans des dispositifs d'auto-évaluation et d'évaluation entre pairs.	1, 2, 3, 4, 5 ou 6

SE AOL1	J'ai confiance en mes capacités à élaborer et utiliser des épreuves bien construites de manière à obtenir des cotes représentant les acquis des élèves.	1, 2, 3, 4, 5 ou 6
BF AOL1	J'ai besoin de formations pour m'aider à élaborer et utiliser des épreuves bien construites de manière à obtenir des cotes représentant les acquis des élèves.	1, 2, 3, 4, 5 ou 6

SE AFL1	J'ai confiance en mes capacités à critiquer mes évaluations pour améliorer la qualité de mon enseignement.	1, 2, 3, 4, 5 ou 6
BF AFL1	J'ai besoin de formations pour m'aider à critiquer mes évaluations pour améliorer la qualité de mon enseignement.	1, 2, 3, 4, 5 ou 6

SE AAL2	J'ai confiance en mes capacités à utiliser les retours constructifs de la part des élèves sur mes pratiques évaluatives pour adapter ces pratiques à leurs besoins.	1, 2, 3, 4, 5 ou 6
BF AAL2	J'ai besoin de formations pour m'aider à utiliser les retours constructifs de la part des élèves sur mes pratiques évaluatives pour adapter ces pratiques à leurs besoins.	1, 2, 3, 4, 5 ou 6

SE AOL2	J'ai confiance en mes capacités à commenter les cotes obtenues par les élèves pour améliorer les résultats de leur prochaine évaluation.	1, 2, 3, 4, 5 ou 6
BF AOL2	J'ai besoin de formations pour m'aider à commenter les cotes obtenues par les élèves pour améliorer les résultats de leur prochaine évaluation.	1, 2, 3, 4, 5 ou 6

SE AFL2	J'ai confiance en mes capacités à analyser l'évaluation dans le but d'identifier les besoins des élèves.	1, 2, 3, 4, 5 ou 6
BF AFL2	J'ai besoin de formations pour m'aider à analyser l'évaluation dans le but d'identifier les besoins des élèves.	1, 2, 3, 4, 5 ou 6

SE AFL3	J'ai confiance en mes capacités à commenter l'évaluation en fonction des erreurs effectuées dans le but de donner des pistes pour les travailler.	1, 2, 3, 4, 5 ou 6
BF AFL3	J'ai besoin de formations pour m'aider à commenter l'évaluation en fonction des erreurs effectuées dans le but de donner des pistes pour les travailler.	1, 2, 3, 4, 5 ou 6

SE AAL3	J'ai confiance en mes capacités à aider les élèves à commenter leur propre évaluation afin d'effectuer une auto-critique (apprentissage réalisé, gestion du temps, préparation, implication pendant l'apprentissage, ...).	1, 2, 3, 4, 5 ou 6
BF AAL3	J'ai besoin de formations pour m'aider à aider les élèves à commenter leur propre évaluation afin d'effectuer une auto-critique (gestion du temps, préparation, implication pendant l'apprentissage, ...).	1, 2, 3, 4, 5 ou 6

SE AOL3	J'ai confiance en mes capacités à créer une cote globale qui fournit un résumé approprié de l'apprentissage des élèves.	1, 2, 3, 4, 5 ou 6
BF AOL3	J'ai besoin de formations pour m'aider à créer une cote globale qui fournit un résumé approprié de l'apprentissage des élèves.	1, 2, 3, 4, 5 ou 6

SE AFL4	J'ai confiance en mes capacités à communiquer les objectifs et les buts de l'évaluation aux parents/tuteurs lorsque cela est approprié.	1, 2, 3, 4, 5 ou 6
BF AFL4	J'ai besoin de formations pour m'aider à communiquer les objectifs et les buts de l'évaluation aux parents/tuteurs lorsque cela est approprié.	1, 2, 3, 4, 5 ou 6

SE AOL4	J'ai confiance en mes capacités à communiquer les résultats finaux aux parents/tuteurs.	1, 2, 3, 4, 5 ou 6
BF AOL4	J'ai besoin de formations pour m'aider à communiquer les résultats finaux aux parents/tuteurs.	1, 2, 3, 4, 5 ou 6

SE AOL5	J'ai confiance en mes capacités à proposer un redoublement ou une réorientation sur base des résultats finaux.	1, 2, 3, 4, 5 ou 6
BF AOL5	J'ai besoin de formations pour m'aider à proposer un redoublement ou une réorientation sur base des résultats finaux.	1, 2, 3, 4, 5 ou 6

SE AOL6	J'ai confiance en mes capacités à rédiger un rapport synthétique des résultats d'évaluation pour le conseil de classe.	1, 2, 3, 4, 5 ou 6
BF AOL6	J'ai besoin de formations pour m'aider à rédiger un rapport synthétique des résultats d'évaluation pour le conseil de classe.	1, 2, 3, 4, 5 ou 6

SE AFL5	J'ai confiance en mes capacités à concevoir des activités de remédiation basées sur les difficultés observées lors d'une évaluation.	1, 2, 3, 4, 5 ou 6
BF AFL5	J'ai besoin de formations pour m'aider à concevoir des activités de remédiation basées sur les difficultés observées lors d'une évaluation.	1, 2, 3, 4, 5 ou 6

SE AFL6	J'ai confiance en mes capacités à adapter la progression pédagogique en fonction des résultats d'évaluations.	1, 2, 3, 4, 5 ou 6
------------	---	-----------------------

BF AFL6	J'ai besoin de formations pour m'aider à adapter la progression pédagogique en fonction des résultats d'évaluations.	1, 2, 3, 4, 5 ou 6
------------	--	-----------------------

SE AAL4	J'ai confiance en mes capacités à faire participer activement les élèves à déterminer critères de réussite.	1, 2, 3, 4, 5 ou 6
BF AAL4	J'ai besoin de formations pour m'aider à faire participer activement les élèves à déterminer des critères de réussite.	1, 2, 3, 4, 5 ou 6

SE AAL5	J'ai confiance en mes capacités à élaborer des dispositifs d'auto-évaluation et d'évaluation entre pairs.	1, 2, 3, 4, 5 ou 6
BF AAL5	J'ai besoin de formations pour m'aider à élaborer des dispositifs d'auto-évaluation et d'évaluation entre pairs.	1, 2, 3, 4, 5 ou 6

SE AAL6	J'ai confiance en mes capacités à animer une discussion collective avec les élèves autour du sens des évaluations.	1, 2, 3, 4, 5 ou 6
BF AAL6	J'ai besoin de formations pour m'aider à animer une discussion collective avec les élèves autour du sens des évaluations.	1, 2, 3, 4, 5 ou 6

SE NUM 1	J'ai confiance en mes capacités à utiliser des outils numériques pour évaluer.	1, 2, 3, 4, 5 ou 6
BF NUM 1	J'ai besoin de formations pour m'aider à utiliser des outils numériques pour évaluer.	1, 2, 3, 4, 5 ou 6

SE NUM 2	J'ai confiance en mes capacités à utiliser des outils numériques pour aider les élèves à s'auto-évaluer ou évaluer leurs pairs.	1, 2, 3, 4, 5 ou 6
BF NUM 2	J'ai besoin de formations pour m'aider à utiliser des outils numériques pour aider les élèves à s'auto-évaluer ou évaluer leurs pairs.	1, 2, 3, 4, 5 ou 6

SE NUM 3	J'ai confiance en mes capacités à utiliser des outils numériques pour m'aider à différencier les évaluations.	1, 2, 3, 4, 5 ou 6
BF NUM 3	J'ai besoin de formations pour m'aider à utiliser des outils numériques pour m'aider à différencier les évaluations.	1, 2, 3, 4, 5 ou 6

SE hétérogénéité	J'ai confiance en mes capacités à utiliser l'évaluation pour gérer l'hétérogénéité de mes classes.	1, 2, 3, 4, 5 ou 6
BF hétérogénéité	J'ai besoin de formations pour m'aider à utiliser l'évaluation pour gérer l'hétérogénéité de mes classes.	1, 2, 3, 4, 5 ou 6

SE collaboration	J'ai confiance en mes capacités à collaborer avec mes collègues pour évaluer.	1, 2, 3, 4, 5 ou 6
BF collaboration	J'ai besoin de formations pour m'aider à collaborer avec mes collègues pour évaluer.	1, 2, 3, 4, 5 ou 6

SE changement de pratiques	J'ai confiance en mes capacités à changer mes pratiques d'évaluation si besoin (nouvelles exigences sociétales ou politiques, besoins des élèves, ...).	1, 2, 3, 4, 5 ou 6
BF changement de pratiques	J'ai besoin de formations pour m'aider à changer mes pratiques d'évaluation si besoin (nouvelles exigences sociétales ou politiques, besoins des élèves, ...).	1, 2, 3, 4, 5 ou 6

SE Défense de ses choix	J'ai confiance en mes capacités à défendre mes choix concernant mes pratiques évaluatives.	1, 2, 3, 4, 5 ou 6
BF Défense de ses choix	J'ai besoin de formations pour m'aider à défendre mes choix concernant mes pratiques évaluatives.	1, 2, 3, 4, 5 ou 6

SE Epreuves externes	J'ai confiance en mes capacités à utiliser les résultats d'épreuves externes pour améliorer la qualité de mon enseignement.	1, 2, 3, 4, 5 ou 6
BF Epreuves externes	J'ai besoin de formations pour m'aider à utiliser les résultats d'épreuves externes pour améliorer la qualité de mon enseignement.	1, 2, 3, 4, 5 ou 6

SE Formations	J'ai confiance en mes capacités à utiliser ce que j'ai appris en formation (initiale ou continue) pour améliorer la qualité de mes pratiques évaluatives.	1, 2, 3, 4, 5 ou 6
------------------	---	-----------------------

c. Partie 3 du questionnaire

Veillez indiquer vos modes préférés de formation pour l'apprentissage de l'évaluation :

1 = très défavorable,

2 = plutôt défavorable,

3 = ni défavorable, ni favorable,

4 = plutôt favorable,

5 = très favorable.

MF1	Une formation traitant l'évaluation sous tous ses aspects, en présentiel, avec d'autres enseignants.	1, 2, 3, 4 ou 5
MF2	Plusieurs modules de formation traitant chacun un aspect spécifique de l'évaluation, en présentiel, avec d'autres enseignants.	1, 2, 3, 4 ou 5
MF3	Une formation traitant l'évaluation sous tous ses aspects, en ligne, en mode autonome.	1, 2, 3, 4 ou 5
MF4	Une formation traitant l'évaluation sous tous ses aspects, en ligne, en collaboration avec d'autres enseignants.	1, 2, 3, 4 ou 5
MF5	Plusieurs modules de formation traitant chacun un aspect spécifique de l'évaluation, en ligne, en mode autonome.	1, 2, 3, 4 ou 5
MF6	Plusieurs modules de formation traitant chacun un aspect spécifique de l'évaluation, en ligne, en collaboration avec d'autres enseignants.	1, 2, 3, 4 ou 5
MF7	Travail collaboratif en classe avec des collègues.	1, 2, 3, 4 ou 5
MF8	Travail collaboratif en classe avec un expert.	1, 2, 3, 4 ou 5
MF9	Formation comprenant une partie en ligne et une autre en présentiel.	1, 2, 3, 4 ou 5
MF10	Webinaires. (Webinaire = séminaire en ligne)	1, 2, 3, 4 ou 5

MF11	Mentorat individuel avec un enseignant de la même discipline. (Mentorat = aide apportée à un débutant par son conseiller expérimenté)	1, 2, 3, 4 ou 5
MF12	Mentorat individuel avec un enseignant d'une autre discipline.	1, 2, 3, 4 ou 5
MF13	Auto-apprentissage à l'aide de ressources (par exemple : littérature, vidéo, ...).	1, 2, 3, 4 ou 5
MF14	Proposition d'un autre mode de formation souhaité	Oui-non <i>Autre</i> <i>(veuillez</i> <i>décrire)</i>

d. Partie 4 du questionnaire

Veillez indiquer quelques informations vous concernant. Les informations que vous allez communiquer maintenant ne sont pas de nature à compromettre votre anonymat.

- Sexe :

1	<i>Femme</i>
2	<i>Homme</i>
3	<i>X</i>

- Âge : *[case à remplir]*

- Matière principale enseignée :

1	Français
2	Mathématiques
3	Sciences
4	Technologie
5	Langues
6	Géographie/Histoire
7	Latin
8	Éducation artistique
9	Éducation physique
10	Cours philosophiques (religion, morale, CPC)
11	<i>Autre</i>

- Années d'expérience professionnelle en tant qu'enseignant(e) : ...

Ne tenez pas compte du fait que vous avez travaillé à temps partiel ou à temps plein. Ne tenez pas compte des périodes de congé prolongées telles qu'un congé de maternité ou paternité.

Arrondissez les nombres d'années à l'unité la plus proche.

- Plus haut type de qualification pédagogique obtenue :

Sélectionnez une seule réponse.

1	Une formation initiale d'enseignant : AESI
2	Un master à finalité didactique.
3	Une formation d'enseignant après une formation universitaire : AESS ou CAPAES
4	Une formation d'enseignant après une formation non universitaire : CAPAES
5	Un certificat d'aptitudes pédagogiques : CAP
6	Une formation dans un autre domaine à caractère pédagogique (par exemple, éducateur spécialisé).
7	Une formation relative à des contenus disciplinaires uniquement (formation professionnelle sans titre pédagogique).
8	Je n'ai aucune qualification formelle que ce soit pédagogique ou en relation avec la matière que j'enseigne.
9	Autre

- Dans quel réseau travaillez-vous majoritairement ?

1	<i>Privé</i>
2	<i>Libre subventionné confessionnel</i>
3	<i>Libre subventionné indépendant</i>
4	<i>Officiel WBE</i>
5	<i>Communal</i>
6	<i>Provincial</i>

Filière des classes auxquelles vous donnez cours majoritairement ?

1	<i>Premier degré différencié</i>
2	<i>Premier degré commun ou enseignement général</i>
3	<i>Enseignement technique ou artistique de transition</i>
4	<i>Enseignement technique ou artistique de qualification</i>
5	<i>Enseignement professionnel</i>
6	<i>Enseignement spécialisé</i>
7	DASPA

- L'enseignement était-il votre premier choix de carrière ? *Par « carrière », on entend avoir un travail rémunéré que vous avez souhaité exercer tout au long de votre vie.*

1	<i>Oui</i>
2	<i>Non</i>

- Vous souvenez-vous avoir suivi un cours sur l'évaluation lors de votre formation initiale (AESI) ?

1	<i>Oui</i>
2	<i>Non</i>
3	Je n'ai pas l'AESI

- Vous souvenez-vous avoir suivi une formation sur l'évaluation dans le cadre d'une formation en cours de carrière ?

1	<i>Oui</i>
2	<i>Non</i>

FIN DU QUESTIONNAIRE

Nous vous remercions d'avoir pris du temps pour remplir ce questionnaire.

IV. Laboratoires cognitifs

a. Laboratoire cognitif n°1 : enseignante de sciences

Plus d'une heure de discussion, de demande de clarifications/questionnements

Très/trop long

- Partie 1 :
 - Mettre des numéros aux questions.
 - Moins compréhensible que la partie 2...
 - Item 4 : « Faire une activité de remédiation en groupant les élèves en classe sur base des compétences acquises ou non lors de l'évaluation » → à changer pour faire émerger l'idée de petits groupes au sein de la classe.
 - Item 6 : « Créer des zones liées aux compétences non acquises et laisser les élèves se grouper selon ce qu'ils estiment devoir retravailler. » → ajouter « des zones de compétences dans la classe »
 - Item 14 : « réfléchir sur ses pratiques évaluatives pour augmenter l'investissement des élèves » → item non compris car ne elle voit pas comment le fait de réfléchir à ses propres pratiques va avoir une incidence sur l'investissement des élèves.
 - Item 17 : « Vérifier si les questions du contrôle évaluent les compétences visées » → enlever le « re » de vérifier.
 - Item 21 : « Permettre aux élèves de discuter et échanger sur leurs méthodes de travail liées à la préparation de cette évaluation » → enlever le « liées à »
 - Item 22 : « écrire uniquement une cote sur l'évaluation pour renseigner l'élève sur le fait qu'il a échoué ou réussi l'évaluation » → remplacer « une cote » par « la cote globale »
 - Item 24 : trop long, pas assez concis, on se perd, idée pas claire.
 - Item 25 : “Attribuer des exercices complémentaires aux élèves ayant échoué, à effectuer à domicile, sur base de la cote de l'évaluation” --> “sur base d'exercices ciblés”. Propose de modifier l'item mais ça ne colle pas avec notre idée de départ.
 - Item 26 : “attribuer des exercices complémentaires différenciés à chaque élève selon les erreurs commises, à effectuer à domicile” --> “pour chaque élève ayant échoué” Propose de modifier l'item mais ça ne colle pas avec notre idée de départ. Semble ne pas envisager qu'on puisse donner des exercices supplémentaires à des élèves ayant réussi l'évaluation.
 - Item 28 : “Rappeler l'importance de la réussite scolaire pour leur futur” --> “pour leur avenir” --> se pose la question “future évaluation ? On dirait qu'il manque quelque chose à cette phrase...”
 - Item 29 : “Rappeler que les évaluations servent à prendre une photo à un temps t de leurs compétences mais qu'ils peuvent utiliser ces informations pour faire évoluer ces compétences” --> “leurs compétences”
 - Item 30 : faire 2 items “Demander aux élèves de réfléchir à l'utilisation de l'évaluation pour améliorer leurs apprentissages. Discuter ensuite avec la classe de leurs réflexions.” --> “demander aux élèves de réfléchir en quoi cette évaluation leur permet d'améliorer leurs apprentissages” et “se servir

de la réflexion personnelle des élèves pour lancer une discussion de réflexion commune”.

- Item 32 : “fournir aux élèves les critères de l’évaluation afin de permettre aux élèves de comprendre leurs erreurs”. --> se pose les questions suivantes : “critères = grille d’évaluation ? Avant ou après l’interrogation ?”
- Item 33 : “fournir aux élèves les critères de l’évaluation et discuter avec eux de ce qui manquait dans leurs réponses pour l’acquisition des critères” --> se pose la question et en vient à voir une incohérence vis-à-vis de l’item 32 : “critères = acquis ici ? Parce que dans l’item précédent critères = grille d’évaluation, ce n’est pas la même chose mais vous utilisez le même mot...”
- Partie 2 :
 - Item 3 : “j’ai confiance en mes capacités d’analyse d’un programme ou un référentiel pour identifier les éléments à évaluer” --> l’item pose problème sans savoir pourquoi... Écrirait davantage “je sais” que “je suis capable de”.
 - Item 9 (et 10) : “j’ai confiance en mes capacités à articuler des séquences d’évaluation dans une temporalité réfléchie” --> “qu’est-ce que vous mettez derrière temporalité ? Temps pour passer l’interrogation ou temps pour voir la matière avant de faire une évaluation ?” --> le participant réfléchit aux options mais nous laissons volontairement l’idée large pour impliquer plusieurs interprétations.
 - Item 13 et item 14 : enlever “bonne” dans “j’ai confiance en mes capacités à rédiger des questions de bonne qualité” et “j’ai besoin de formations pour m’aider à rédiger des questions de bonne qualité”
 - Item 17 (et item 18) : “j’ai confiance en mes capacités à collecter des données utiles sur les apprentissages de mes élèves” --> ne comprend pas l’item (“données utiles ?”)
 - Item 31 (et item 32) : “j’ai confiance en mes capacités à analyser mes pratiques d’évaluation pour améliorer la qualité de mon enseignement” --> “mes capacités d’auto-critique”
 - Item 35 : “j’ai confiance en mes capacités à commenter l’interrogation en fonction de la note obtenue dans le but d’améliorer la note de la prochaine évaluation” --> propose de retirer “en fonction de la note obtenue” mais ça ne colle pas à notre idée, on veut mettre en avant le fait de se baser sur la cote pour commenter.
 - Item 37 (et 38) : “j’ai confiance en mes capacités à analyser l’interrogation dans le but d’identifier les besoins des élèves” --> propose d’ajouter “spécifiques” mais non, pas forcément référence aux besoins spécifiques ici.
 - Item 39 (et 40) : “j’ai confiance en mes capacités à commenter l’interrogation en fonction des erreurs effectuées dans le but de donner des pistes pour les retravailler” --> “des pistes réflexives pour les travailler”
 - Item 41 (et 42) : “j’ai confiance en mes capacités à aider les élèves à commenter leur propre évaluation pour effectuer une auto-critique

(apprentissage réalisé, gestion du temps, préparation, implication pendant l'apprentissage, ...) --> "afin d'effectuer"

- Item 43 (et 44) : "j'ai confiance en mes capacités à créer une note finale qui fournit un résumé approprié de l'apprentissage des élèves" --> "note globale"
- Item 55 (et 56) : "j'ai confiance en mes capacités à adapter la progression pédagogique en fonction des résultats d'évaluation" --> ne connaît pas le concept de "progression pédagogique" --> à modifier ?
- Item 67 (et 68) : "j'ai confiance en mes capacités à utiliser les outils numériques pour m'aider à différencier les évaluations" --> "... à adapter les évaluations en fonction des difficultés de l'élève".
- Item 71 (et 72) : "j'ai confiance en mes capacités à collaborer avec mes collègues pour évaluer" --> "... avec mes collègues de la même discipline ..."
- Item 73 (et 75) : "J'ai confiance en mes capacités à changer mes pratiques d'évaluation (en fonction de nouvelles exigences sociétales ou politiques, en fonction des besoins des élèves, ...)" --> propose de faire 2 items pour séparer "en fonction de nouvelles exigences sociétales ou politiques" et "en fonction des besoins des élèves" mais nous voulons juste savoir s'ils se sentent capable de changer leurs pratiques, peu importe la raison finalement.
- Partie 3 :
 - 5 : "une série de courts modules sur l'évaluation, en ligne, proposés sous forme de modules indépendants" --> redondant ! Enlever "proposés sous forme de modules indépendants".
 - 11 : "mentorat individuel avec un enseignant de la même discipline" --> ne connaît pas le terme mentorat donc propose de l'expliquer une fois.
 - 13 : "auto-apprentissage à l'aide de ressources sélectionnées (par exemple : littérature, vidéos, ...)" --> enlever "sélectionnées"
- Partie 4 :
 - En général, propose de faire des questions utilisant un ton moins familier : "de quel sexe êtes-vous ?" plutôt que "quel est votre sexe ?"
 - Point 5 (qualification pédagogique) : propose de dire "formation universitaire ou non universitaire"

Remarque générale : trop trop trop long : sensation beaucoup plus présente pour la première partie

b. Laboratoire cognitif n°2 : enseignant de mathématiques

• Partie 1 :

- Scénario : “Vous donner à votre classe une évaluation sommative de fin de chapitre sous format papier/crayon, incluant des aménagements et des modifications pour les apprenants identifiés. 16 des 24 élèves échouent. --> **difficulté de compréhension sur les modifications pour les apprenants identifiés.**
- Item 1 : “Inscrire la cote de chaque élève dans le cahier de cotes et passer à la matière suivante.” --> **conseil : changer en inscrire le résultat de chaque élève dans le cahier de cote du professeur...**
- Item 6 : “Créer des zones liées aux compétences non acquises et laisser les élèves se grouper selon ce qu’ils estiment devoir retravailler.” --> **réflexion sur le fait que personne ne fera cette action car les élèves ne se regrouperont jamais par rapport à ce qu’ils veulent travailler mais ils iront là où il y aura ceux avec qui ils s’entendent.**
- Item 9 : “Transmettre l’interrogation aux élèves. Leur laisser quelques minutes pour prendre connaissance des corrections...” --> **demande de si dans la transmission, l’idée était de transmettre toutes les questions ou juste une partie.**
- Item 10 à 12 : “Transmettre les cotes aux parents.” “Transmettre les compétences acquises et non acquises aux parents.” “Laisser quelques minutes aux élèves pour qu’ils puissent écrire ce qui a été et ce qui n’a pas été lors de cette évaluation à leurs parents.” --> **pour les trois items la question était de savoir comment transmettre toutes ces informations (voie journal de classe, informatisée, ...). Le dernier item : Comment peuvent-ils expliquer à leurs parents des erreurs qu’ils ne savent même pas reconnaître.**
- Item 14 : “Réfléchir sur ses pratiques évaluatives pour augmenter l’investissement des élèves. --> **question sur quand avait lieu cette réflexion ? Avant l’interrogation ?**
- Item 16 : “Vérifier si la cote finale représente bien le niveau d’acquisition global de la matière du chapitre selon ce que l’élève a produit dans l’évaluation” --> **Difficulté de compréhension : nécessité de plusieurs relectures**
- Item 19 : “Discuter des cotes avec l’équipe pédagogique pour vérifier si le phénomène se répète également chez eux.” --> **Question sur le fait si on parle bien du fait que les cotes sont aussi basses chez eux. La personne ajoute ensuite que cet item est compliqué vue la différence dans la manière de coter entre les enseignants.**
- Item 23 : “Écrire sur l’évaluation de chaque élève des retours personnalisés sur les types d’erreurs effectués et des conseils sur la manière de les travailler.” --> **On n’est pas payé pour ça.**
- Item 24 : “Faire réaliser en classe, à chaque élève, pour l’évaluation corrigée qu’ils reçoivent, une auto-analyse de sa compréhension des concepts, sa gestion du temps, sa préparation, son implication pendant l’apprentissage, ... avec ou sans discussion avec l’enseignant, selon les besoins des élèves.” --> **trop d’informations ce qui entraîne d’être d’accord**

avec une bonne partie mais pas avec d'autres...ce qui entraîne une difficulté de placement sur l'échelle de Likert. --> Ajout de l'idée d'action dans son entièreté dès l'énoncé du scénario et réflexion à mener sur la réduction de cet item.

- Item 25 : "Attribuer des exercices complémentaires aux élèves ayant échoué, à effectuer à domicile, sur base de la cote de l'évaluation." --> Incompréhension sur le fait que vu qu'on parlait de cote, est-ce que cela implique que la réalisation de ceux-ci peut modifier la cote.
- Item 29 : "Rappeler que les évaluations servent à prendre une photo à un temps t de leurs compétences mais qu'ils peuvent utiliser ces informations pour faire évoluer ces compétences." --> relecture nécessaire pour une bonne compréhension. Suggestion de remplacer "ces" par "leurs".
- Item 33 : "Fournir aux élèves les critères de l'évaluation et discuter avec eux de ce qui manquait dans leurs réponses pour l'acquisition des critères." --> Comment faire pour les élèves qui ont rendu leurs feuilles blanches ?
- Partie 2 :
 - Observation : assez rapidement il lit les items deux à deux (lentement le premier et presque pas le deuxième).
 - Item 9 : "J'ai confiance en mes capacités à articuler des séquences d'évaluation dans une temporalité réfléchie." --> Incompréhension de la phrase. Après explication une remarque a été faite dans le fait qu'on peut réfléchir au départ à une temporalité mais qu'elle n'est jamais respectée du aux nombreux aléas (activités, grève, absences, ...).
 - Item 11 : "J'ai confiance en mes capacités à donner plus de poids aux questions portant sur les objectifs principaux." --> Demande si on parle de poids dans les questions de l'interrogation ou dans les questions même en cours.
 - Item 23 : "J'ai confiance en mes capacités à communiquer des informations utiles aux élèves à la suite de l'évaluation" --> Demande si on parle de communication à toute la classe ou à chaque élève en individuel.
 - Item 29 : "J'ai confiance en mes capacités à élaborer et utiliser des épreuves bien construites de manière à obtenir des notes représentant les acquis des élèves." --> plusieurs relectures nécessaires mais qui ont mené à une compréhension sans poser de question.
 - Item 33 : "J'ai confiance en mes capacités à utiliser les retours constructifs de la part des élèves sur mes pratiques évaluatives pour m'adapter à leurs besoins." --> Demande si, pour nous, les élèves sont capables de faire des retours constructifs. (Sous-entendu que pour lui, ils ne le sont pas.)
 - Item 59 : "J'ai confiance en mes capacités à élaborer un dispositif de co-évaluations entre pairs." --> Demande ce qu'est la co-évaluation.
 - Item 61 : "J'ai confiance en mes capacités à animer une discussion collective avec les élèves autour du sens d'une évaluation." --> Demande ce qu'on entend par "le sens d'une évaluation". Demande également si cette discussion se passe avant ou après car pour lui on suivait une chronologie et on est arrivé à des actions qui se passent après l'interrogation.

- Réflexion sans item : On cherche plus à voir comment aider les élèves sans voir comment aider les profs.
- Item 63 : “J’ai confiance en mes capacités à utiliser les outils numériques pour évaluer” --> “les” outils lui font demander de quels outils on parle.
- Item 67 : “J’ai confiance en mes capacités à utiliser les outils numériques pour m’aider à différencier les évaluations”. --> Demande ce qu’on entend par différencier les évaluations.
- Partie 3 :
 - Énoncé : “Veuillez indiquer vos modes préférés de soutien pédagogique pour l’apprentissage de l’évaluation.” --> Incompréhension de ce qu’on entend par soutien pédagogique. Quand on lui dit que ce sont les formations, il avoue qu’il pensait que c’était du soutien pour les élèves et demande qu’on reformule l’énoncé.
 - Item 1 : “Cours complet sur l’évaluation, en face à face, avec d’autres enseignants en exercice.” --> Demande pourquoi on a mis “en exercice”. Il trouve ça étrange qu’on irait rechercher des enseignants plus en exercice.
 - Item 2 : “Modules courts sur l’évaluation, en face à face, avec d’autres enseignants en exercice.” --> dit de reformuler “courts” en “petits” car fait penser à un cours surtout à cause des items 3 et 4 qui démarrent par le mot “cours”.
 - Item 4 : “Court complet sur l’évaluation, en ligne, en collaboration avec d’autres enseignants en exercices.” --> “s” de trop dans le mot “exercices”.
 - Item 7-8-11 : “Travail collaboratif en classe avec *des collègues*.” “Travail collaboratif en classe avec *un expert*.” “Mentorat individuel avec un enseignant de la *même discipline*.” --> comparaison des trois expressions mises en italique pour savoir si l’expert et ceux de la même discipline sont obligatoirement des non-collègues et si les collègues sont obligatoirement d’une autre discipline.
 - Réflexion sans item : Ne comprend pas si tous ces items sont des formations purement théoriques et si elles sont suivies d’un questionnaire ou suivies d’exercices à faire après en classe pour aider à mettre en place.
- Partie 4 :
 - Pour âge et années d’expérience --> conseille des tranches.
 - Trouve la fin abrupte et conseille de terminer par un mot de remerciement vu le temps passé pour pouvoir réaliser

Remarques générales :

- Encore un questionnaire où on cherche à savoir ce que les professeurs peuvent faire en plus sans se questionner sur ce que les élèves peuvent faire en plus.
- Questionnaire à raccourcir surtout la première partie qui demande plus de concentration

- À la question s'il a apprécié quelque chose ou si quelque chose lui a déplu, il a répondu que rien ne l'avait choqué et que c'étaient des questions qui étaient logiques par rapport à la pratique de l'évaluation en général.

c. Laboratoire cognitif n°3 : enseignante d'anglais

- Partie 1 :
 - Item 7 : “utiliser les cotes obtenues pour motiver les élèves à augmenter leur investissement” --> elle tique sur le fait de motiver les élèves de cette façon mais comprend l’item.
 - Item 18 : “faire réaliser en classe, à chaque élève, une auto-analyse (compréhension des concepts, gestion du temps, préparation, implication pendant l’apprentissage, ...) avec ou sans discussion avec l’enseignant, selon les besoins des élèves” --> enlever “selon les besoins des élèves” et faire 2 items, un avec discussion et un sans discussion --> non car on veut laisser l’élève décider.
 - Item 21 : “demander aux élèves de réfléchir seul à l’utilisation de l’évaluation pour améliorer leurs apprentissages. Se servir des réflexions personnelles des élèves pour lancer une discussion de réflexion commune.” --> elle se demande si c’est de cette évaluation qu’on parle ou dans l’absolu.
- Partie 2 :
 - Item 9 : “J’ai confiance en mes capacités à articuler des séquences d’évaluation dans une temporalité réfléchie” --> se demande si c’est le temps qu’on laisse aux élèves pour une évaluation ou programmer des évaluations à différents moments de l’année.
- Partie 3 :
- Item 10 : “webinaires” --> demande si c’est bien les conférences via internet.
- Partie 4 :
 - Item 5 : “type de qualification pédagogique obtenue” --> propose d’ajouter “une expérience professionnelle et pas de titre pédagogique” (mais c’est déjà ce que recouvre pour nous le point 7 “une formation relative à des contenus disciplinaires uniquement”)
 - Item 7 : “niveau des classes auxquelles vous donnez cours majoritairement” --> le fait de ne permettre qu’un choix possible lui semble embêtant --> pour permettre un traitement statistique plus aisé

Remarques générales : le questionnaire prend un certain temps à remplir... et ça laisse un sentiment culpabilisant.

d. Laboratoire cognitif n°4 : enseignante de géographie

- Partie 1 :
 - Item 8 : “modifier sa propre façon d’évaluer pour augmenter l’investissement des élèves” --> bloque sur le fait que ça puisse influencer l’investissement des élèves mais comprend l’item.
 - Item 18 : “faire réaliser en classe, à chaque élève, une auto-analyse (compréhension des concepts, gestion du temps, préparation, implication pendant l’apprentissage, ...) avec ou sans discussion avec l’enseignant, selon les besoins des élèves” --> phrase pas claire... à scinder ?
 - Item 20 : “Rappeler que les évaluations servent à prendre une photo à un temps t de leurs compétences mais qu’ils peuvent utiliser ces informations pour faire évoluer leurs compétences” --> prendre une photo est mal compris, pense que les élèves prennent une photo du test avec leur téléphone --> propose de changer de métaphore sans en proposer une.
- Partie 2 :
 - Item 5 : “J’ai confiance en mes capacités à créer une concordance entre les objectifs visés, mes méthodes d’enseignement et mes évaluations” --> ne comprend pas que c’est une concordance entre les 3 en même temps donc propose de rendre ça un peu plus explicite.
 - Item 9 : “J’ai confiance en mes capacités à articuler des séquences d’évaluation dans une temporalité réfléchie” --> ne comprend pas “articuler des séquences d’évaluation” donc à modifier pour expliciter ce qu’on veut dire.
 - Item 57 : “j’ai confiance en mes capacités à faire participer activement les élèves à l’établissement de critères de réussite” --> “établissement” mot à changer car pense que nous parlons d’école
- Partie 3 :
 - Item 5 : “une série de modules indépendants sur l’évaluation, en ligne” --> ne comprend pas l’idée de modules indépendants
 - Item 10 : “webinaires” --> ne comprend pas ce que c’est.
 - Item 11 : “mentorat individuel avec un enseignant de la même discipline”
 - Item 12 : “mentorat individuel avec un enseignant” --> ne comprend pas ce que veut dire “mentorat” --> propose de l’expliquer une fois.
- Partie 4 :
 - Item 9 : “vous souvenez-vous avoir suivi un cours sur l’évaluation en formation initiale ?” --> ne comprend pas l’expression “formation initiale” Mais cette formulation a été utilisée dans l’item 5 et ça ne lui avait pas posé question.

Remarques générales : ça prend du temps de répondre à ce questionnaire... et c’est assez culpabilisant. Et puis, selon le type d’élèves qu’on a, on n’a pas les mêmes ressentis et possibilités de travail. Elle, par exemple, sait qu’elle devrait écrire des commentaires sur chaque interrogation mais elle n’a pas le temps de le faire vu le nombre d’élèves.

e. Laboratoire cognitif n°5 : enseignante de français

• Partie 1 :

- Item 4 à 6 traitant les différentes façons de faire une remédiation en classe -
--> Indique que ce pour ce genre d'actions, elle les envoie en remédiations car elle n'a pas le temps d'effectuer une remédiation en classe.
- Item 7 : "Utiliser les cotes obtenues pour motiver les élèves à augmenter leur investissement." --> s'arrête à la lecture, sourit, me regarde, ne me dit rien.
- Item 9 : "Demander aux élèves de réfléchir individuellement sur leur investissement dans le but de se fixer des objectifs personnels pour la prochaine évaluation." --> doit relire puis réindique que c'est le genre d'actions qu'elle pourrait faire mais aussi qu'elle pourrait demander au professeur qui effectue la remédiation de le faire avec les élèves.
- Item 10 : "Vérifier si la cote finale obtenue par chaque élève représente bien le niveau d'acquisition global de la matière du chapitre selon ce que chacun a produit dans l'évaluation" --> doit relire plusieurs fois, dit que l'item est trop long et qu'il faudrait au minimum rajouter une virgule après le mot "chapitre".
- Item 12 : "Demander aux élèves d'énoncer une critique de l'évaluation (cela peut même se faire de façon écrite et anonyme)." --> trouve l'idée super intéressante. Elle nous a recontacté indiquant que, suite à une interrogation complètement ratée, elle a demandé une critique aux élèves et qu'ils ont répondu que cette interrogation respectait les attendus énoncés et que c'était juste dû à un manque de travail de leur part.
- Item 15 : "Permettre aux élèves de discuter et échanger sur leur façon de se préparer aux évaluations." --> trouve que c'est quelque chose à faire dans le cours de soutien appelé "méthodo".
- Item 18 : "Faire réaliser en classe, à chaque élève, une auto-analyse (compréhension des concepts, gestion du temps, préparation, implication pendant l'apprentissage, ...). Être disponible pour les élèves qui demandent de l'aide." --> Demande si cette action, on lui demande de la faire avant ou après l'interrogation (cela montre que le scénario n'est plus trop présent pour elle).
- Item 21 : "Demander aux élèves de réfléchir seuls à l'utilisation de l'évaluation pour améliorer leurs apprentissages. Se servir des réflexions personnelles des élèves pour lancer une discussion de réflexion commune." --> Pour elle, les élèves, surtout ceux de première année, sont incapables seuls d'une telle réflexion et qu'il faut plus les accompagner, les aider dans un premier temps.

• Partie 2 :

- Observation : il a fallu 10 items avant qu'elle lise les items deux à deux (lentement le premier et presque pas le deuxième).
- Item 7 : "J'ai confiance en mes capacités à choisir la bonne modalité d'évaluation (travaux, QCM, questions ouvertes, en groupe, seul, numérique, ...) suivant les avantages et les inconvénients de chaque modalité." --> s'arrête sur le mot "numérique" car se dit capable d'utiliser

toutes les différentes modalités mais ne se dit pas capable d'utiliser une modalité numérique de l'évaluation et ajoute qu'elle n'a pas les moyens technologiques pour le faire avec chaque classe.

- Item 23 : "J'ai confiance en mes capacités à communiquer des informations utiles aux élèves à la suite de l'évaluation". --> trouve l'item pas très clair et se demande ce qu'on entend par la communication d'informations utiles.
- Item 33 : "J'ai confiance en mes capacités à utiliser les retours constructifs de la part des élèves sur mes pratiques évaluatives pour m'adapter à leurs besoins" --> s'arrête sur l'item, le relit mais ne fait aucun commentaire.
- Item 41 : "J'ai confiance en mes capacités à aider les élèves à commenter leur propre évaluation afin d'effectuer une auto-critique (apprentissage réalisé, gestion du temps, préparation, implication pendant l'apprentissage, ...)." --> s'arrête sur l'item, le relit mais ne fait aucun commentaire.
- Item 50 : "J'ai besoin de formations pour m'aider à proposer un redoublement ou une réorientation sur base des résultats finaux." --> trouve très bizarre qu'une formation soit proposée pour ça et le répète à deux reprises.
- Partie 3 :
 - Demande s'il n'y a pas besoin d'un titre pour indiquer une nouvelle partie.
 - Item 14 : "Atelier de développement professionnel." --> trouve qu'il faudrait une explication de ce qu'on entend par atelier de développement professionnel.
 - Item 15 : "Autre" (veuillez décrire) (réponse ouverte)" --> problème technique : il fallait obligatoirement écrire quelque chose pour passer à la page suivante.
- Partie 4 :
 - Aucun commentaire et aucune observation notable

Remarques générales :

- Elle a aimé le site qui permet d'avoir un rappel quand on clique sur une bulle le nom de l'échelon (ex : tout à fait probable).
- Elle a apprécié faire le questionnaire car ça lui a fait réfléchir à sa pratique. Elle nous a recontacté pour nous dire que depuis qu'elle l'avait passé, elle réfléchissait à sa pratique et qu'elle s'est sentie légèrement culpabilisée par tous les items qu'elle ne faisait pas.

f. Laboratoire cognitif n°6 : enseignant de français

Problème avec le site qui ne permet sur aucun smartphone de se connecter au lien en affichant le message : “Erreur de sécurité. Le certificat ...”

- Partie 1 :
 - L'échelle : “En tant qu'enseignant dans cette situation et selon votre contexte, indiquez la probabilité d'effectuer les actions suivantes dans leur entièreté. 1. Tout à fait improbable ; 2 : Plutôt improbable ; 3 : Légèrement improbable ; 4 : Légèrement probable ; 5 : Plutôt probable ; 6 : Tout à fait probable.” --> **Difficulté de voir la différence entre 3 niveaux positifs. Il nous conseille de passer à une échelle à 5 ou 4 échelons car c'est beaucoup plus facile quand il n'y a que 2 échelons positifs.**
 - Item 10 : “Vérifier si la cote finale obtenue par chaque élève représente bien le niveau d'acquisition global” --> **Il demande si on parle d'une cote représentative par rapport aux autres élèves de la classe ou par rapport aux compétences.**
 - Item 16 : “Ne pas écrire de commentaire. La cote globale suffit pour renseigner l'élève sur le fait qu'il a échoué ou réussi l'évaluation.” --> **Difficulté avec la négation car s'il voulait dire qu'il voulait écrire un commentaire, il devrait aller dans les échelons négatifs.**
 - Item 20 : “Rappeler que les évaluations servent à savoir où ils en sont dans l'acquisition de leurs compétences mais qu'ils peuvent utiliser ces informations pour progresser.” --> **Il dit qu'il ne le ferait pas avec ces mots mais faire un speech dans cette idée générale oui.**
- Partie 2 :
 - L'échelle : “À quel point ressentez-vous le besoin de vous former sur celles-ci ? 1 : Pas du tout ; 2 : Presque pas ; 3 : Très peu ; 4 : Un peu ; 5 : Beaucoup ; 6 : Tout à fait.” --> **Demande s'il ne faut pas échanger le 2e et le 3e en termes de sens. + Pose la question de ce qu'il se passe si on en a besoin mais pas envie.**
 - **Observation : comprend immédiatement le doublon et les lit deux à deux.**
 - 7-8 items du type for ou as comme 39-40 : “Commenter en fonction des erreurs dans le but de donner des pistes pour les travailler.” ou items 53-54 “Concevoir des activités de remédiations” --> **Il se dit très capable mais il ne veut pas le faire et ne le fera pas.**
 - Item 57 : “J'ai confiance en mes capacités à faire participer activement les élèves à déterminer des critères de réussite.” --> **Il a répété à plusieurs reprises qu'on parlait bien de ses capacités à lui, donc de sa capacité à faire en sorte que les élèves soient capables de déterminer les critères. “Moi”, “pour l'élève”, “moi”, “pour l'élève”. Conseille de reformuler l'item.**
 - Item 69 : “J'ai confiance en mes capacités à utiliser l'évaluation pour gérer l'hétérogénéité de mes classes.” --> **Il aurait aimé une précision sur ce qu'on entend par “gérer” et par “hétérogénéité”. Il demande ensuite si cela signifie qu'il doit adapter l'interrogation au niveau de chaque élève.**
- Partie 3 :

- Rien
- Partie 4 :
 - Remarque de moi-même pas du participant : “Âge” à changer car commence avant 18 ans alors que dans le consentement il est mis 18 ans.

Remarques générales :

- Rien ne l’a choqué
- Il conseille de terminer avec quelque chose de plus joyeux qu’un simple remerciement.
- Il a aimé le site qui permet d’avoir un rappel quand on clique sur une bulle le nom de l’échelon (ex : tout à fait probable).

g. Laboratoire cognitif n°7 : enseignante de cours commerciaux

23 minutes prises pour répondre aux questions et me dire ce qui lui pose problème

Prof de pratique professionnelle en 3e

- Partie 1 :
 - Item 14 : “Discuter avec des collègues de la même discipline des compétences non acquises en vue d’échanger sur les pratiques pédagogiques” --> remarque générale sur le fait qu’elle n’a pas de collègue directe de la même discipline (option) pour échanger sur les pratiques... et c’est difficile quand le personnel n’est pas fixe (gros turnover) --> réponse plutôt négative mais pas par manque d’envie, par manque de possibilités. --> je pense aux nouvelles possibilités de formation telles que les communautés d’apprentissages professionnelles qui pourraient lui convenir dans le futur.
- Partie 2 :
 - Item 9 : “j’ai confiance en mes capacités à planifier mes évaluations” --> soulève le fait que ce n’est pas le problème de planifier des évaluations mais de gérer les imprévus qui est compliqué car il y en a beaucoup dans son école, ce qui entrave sa planification.
 - Item 11 : “j’ai confiance en mes capacités, dans les évaluations, à donner plus de poids aux questions portant sur les objectifs principaux” --> tique sur la pertinence de la question car, selon elle, comme il n’y a quand même qu’une seule cote au bulletin, qu’il n’y a pas de détail, on ne sait quand même pas ce qu’il y a derrière... (pour elle, il y a un manque de transparence dans nos objectifs et leur “poids” et une communication opaque aux élèves et à leurs parents)
 - Item 25 : “j’ai confiance en mes capacités à porter un regard réflexif sur mes évaluations” --> seul c’est un peu compliqué... On est forcément d’accord avec soi-même... c’est mieux à plusieurs.
 - Item 69 : “j’ai confiance en mes capacités à utiliser l’évaluation pour gérer l’hétérogénéité de mes classes” --> se demande si c’est lié aux différents niveaux d’élèves (différencier ses évaluations) ou si c’est l’idée de faire des évaluations diagnostiques.
 - Item 77 : “j’ai confiance en mes capacités à utiliser les résultats d’épreuves externes pour améliorer la qualité de mon enseignement” --> ajouter “ça ne me concerne pas” dans les réponses possibles.
- Partie 3 : /
- Partie 4 :
 - Item 3 : “Matière principale enseignée” --> frustration de ne pas pouvoir renseigner son intitulé de cours dans “autre”... laisser la possibilité de l’écrire pour se rendre compte de toutes les options des profs qui ont répondu à l’enquête et... pour la reconnaissance de ces enseignants ?

Remarques générales :

- Pas trop long.
- Accessible à tous les enseignants (facile à comprendre).
- Intéressant pour se remettre en question.

V. Analyses supplémentaires avant la création du questionnaire

a. Justification des thèmes de la partie 1 du questionnaire

Nombre et ordre

Les thèmes de pratiques évaluatives investigués dans la partie 1 sont des thèmes qui reviennent couramment dans la littérature. Nous en avons d'abord créé 11 et pour chaque thème, 3 sous-questions ont été pensées pour déterminer des profils d'enseignants. Nos laboratoires cognitifs nous ont convaincus de la nécessité d'en enlever au vu de la fatigue engendrée par le positionnement sur 33 items liés à un seul scénario. Nous en avons finalement gardé 7, en discutant avec les deux premiers participants aux laboratoires cognitifs, ce qui fait tout de même un positionnement à avoir sur 21 items.

Ces items ne sont ni des pratiques « efficaces » ni des pratiques « médiocres » (Casalfiore, 2000) car aucun article scientifique à ce jour (selon nos recherches) ne semble avoir mis à jour des pratiques d'un ou l'autre type. En revanche, des pratiques sont recommandées/encouragées par certains organismes (par l'IES, par la FWB, ...) car il y a de nombreuses recherches qui ont trouvé des résultats positifs à leur utilisation.

L'ordre des thèmes n'est pas pensé en progression ou en hiérarchie parce qu'aucun thème n'est plus légitime qu'un autre, plus important qu'un autre. Ce sont diverses facettes d'un même concept : l'évaluation.

Littérature derrière chaque thème

Le thème de la nouvelle évaluation à faire ou pas (premier thème) est documentée dans la recherche de Soussi et ses collègues en 2006 : les enseignants interrogés sont favorables à cette façon de procéder. Ensuite, l'idée est sous-entendue dans le guide pour l'évaluation édicté par la FWB en 2025. En effet, une idée forte voulue d'une évaluation sommative est la prise d'informations utiles, de qualité. Soussi et ses collègues (2006) nous informent que les enseignants de l'enseignement secondaire préfèrent créer des évaluations eux-mêmes que se baser sur des modèles existants (épreuves externes par exemple). Or, par exemple, Allal et Dauvisis (2019) nous disent que peu d'enseignants sont formés à la délicate tâche de rédaction de questions d'évaluations, questions qui dans leurs formulations sont parfois à l'origine des erreurs des élèves. Les enseignants, conscients ou pas de ce biais, font parfois repasser une évaluation si cela est jugé nécessaire (Soussi et al., 2006). Raulin et Lebeaume (2019) nous disent d'ailleurs que l'élève, de par sa conscience évaluative, et de ses actions, peut impacter les décisions ultérieures de son enseignant.

Le deuxième thème qui est celui de la remédiation en classe est aussi un thème émanant du guide pour l'évaluation de la FWB de 2025 : il faut rejoindre les besoins des élèves, que ce soit

en remédiations ou en consolidation. Soussi et ses collègues mentionnent qu'une des finalités de l'évaluation perçues par les enseignants (déjà en 2006) est de pratiquer des activités de remédiations.

Le troisième thème qui est celui de l'investissement des élèves est un thème controversé mais qui est inhérent à la situation d'enseignement-apprentissage. En effet, pour qu'apprentissage il y ait, il faut que les deux parties fassent une part du travail : l'enseignant aussi bien que l'élève (Stiggins, 2017). D'ailleurs, dans le code de l'enseignement, il est rappelé qu'un travail scolaire de qualité implique, entre autres, de « participer activement aux activités scolaires en montrant de l'écoute, de l'implication, de la prise d'initiative, de l'engagement et du sens des responsabilités » (2019, Titre V, chapitre 1^{er}, section IV, article 1.5.1-8., § 3, 4^o, p.18). Il est essentiel d'impliquer l'élève dans le processus évaluatif pour qu'il puisse, in fine, tirer réellement parti des interventions de l'enseignant pour améliorer ses apprentissages (Fagnant, 2022). En Chine, par exemple, dans la cote est comprise une partie sur l'investissement de l'élève car l'engagement envers l'apprentissage fait partie selon eux d'une cote globale liée aux cours (DeLuca et al., 2021). D'ailleurs, nous savons qu'à l'adolescence, il y a une baisse de la motivation (Berger, 2021 cité par Quittre et al., 2025) qui rend difficile le travail de l'enseignant. Celui-ci doit susciter l'intérêt des élèves et les faire travailler, malgré cette baisse de la motivation (Quittre et al., 2025 ; Soussi et al., 2006). D'autre part, à travers le guide pour l'évaluation (2025), la FWB nous renforce dans l'idée de travailler sur le statut de l'erreur pour ne pas que la motivation des élèves à apprendre se détériore. Zilberberg l'énonçait déjà en 2019 car l'erreur est vue très négativement depuis le siècle passé et doit être évitée selon l'opinion générale. Il faut donc échanger avec les élèves pour permettre à ceux-ci de dégager des pistes pour ajuster leurs stratégies d'apprentissage (FWB, 2025).

Le thème suivant est celui de la remise en question de l'évaluation. En effet, Detroz (2023) nous rappelle que l'évaluation est un construit humain, donc faillible. Dans le cycle MoSCodEE, la dernière étape est l'analyse qualité a posteriori du dispositif d'évaluation en vue de réguler pour une prochaine évaluation. Les biais, dont celui cité supra, sont nombreux dans l'évaluation des étudiants et des erreurs de jugement de l'enseignant peuvent toujours être faites (Felouzis, 1993, cité par Detroz & al., 2020). Les élèves, via leur conscience évaluative, ont leur part de réflexion également, leur point de vue à apporter, puisqu'ils sont acteurs de ce processus d'évaluation (Raulin & Lebeaume, 2019).

Le thème 5, les discussions entre pairs, vient de la nécessité de collaborer pour comprendre, pour réguler, que ce soit entre collègues enseignants ou entre élèves. Soussi et ses collègues

(2006) mettent en évidence la volonté des enseignants de secondaire de collaborer autour de l'évaluation (discussions, échanges de travaux, construction d'épreuves mais également, harmonisation des pratiques). D'ailleurs, Crahay nous confirme que les enseignants échangent autour des compétences des élèves (2009). Talis (exigences, ressources professionnelles et bien-être, 2024) nous informe que la circulaire 7167 reconnaît dorénavant plusieurs formes de travail collaboratif dont la collaboration autour des évaluations, au sein des 60 périodes de travail collaboratif obligatoires depuis 2019. D'autre part, selon Campanalle et ses collègues (2010) : « les dispositifs les plus efficaces pour provoquer l'évaluation régulation sont ceux qui s'accompagnent d'interactions langagières entre pairs et entre formés et formateurs (...) autrement dit des dispositifs d'autoévaluation socialisée » et pour Wyatt-Smith et ses collaborateurs (2010), le développement d'une identité d'évaluateur se fait via le groupe de travail et des négociations avec les pairs.

Le thème 6 quant à lui est une notion bien documentée dans les sciences de l'éducation, un incontournable : le feedback. Celui-ci est présent dans le cycle MoSCodEE, dans le point « rétroaction vers l'étudiant et le groupe d'étudiants » (Detroz et al., 2020). Soussi et ses collaborateurs (2006) mentionnent des enseignants disant utiliser des retours sous forme de commentaires vers les élèves à la suite des évaluations ainsi qu'un partage de pistes pour s'améliorer. Hattie et Timperley (2007) décrivent, quant à eux, le pouvoir d'un bon feedback, construit et porteur d'informations importantes. Depuis, l'idée ayant fait son chemin, le guide sur l'évaluation (FWB, 2025) enjoint les enseignants à s'en emparer et à l'utiliser à bon escient.

Le dernier thème, celui de du rôle de l'évaluation est envisagé sous la forme d'une implication de l'élève dans la réflexion, un pas vers la construction de la conscience évaluative des élèves (Raulin & Lebeaume, 2019), le pourquoi évaluer, ce que cela permet pour eux (informations sur les acquis et les matières à retravailler) et pour l'enseignant (émergence de difficultés et des besoins des élèves) (FWB, 2025). C'est une discussion qui permet de lever les implicites, de rendre transparent ce qui ne l'est pas, afin de renforcer le lien enseignement-apprentissage : la finalité des évaluations, qu'elles soient formatives ou sommatives, est bien de soutenir l'apprentissage de chacun (FWB, 2025). La capacité d'« apprendre à apprendre », une des visées transversales du Tronc Commun, s'en trouvera renforcée (FWB, 2025).

b. Analyse plus détaillée du lien entre item et approche dans la partie 1 du questionnaire

Thème : nouvelle évaluation (ou pas)	
Comment gère-t-on l'idée de faire refaire une évaluation ?	
Inscrire le résultat de chaque élève dans le cahier de cotes et passer à la matière suivante. AOL	1, 2, 3, 4, 5 ou 6
Faire une séance de correction pour ensuite faire repasser une épreuve similaire. AFL	1, 2, 3, 4, 5 ou 6
Après avoir laissé du temps aux élèves, pour prendre connaissance de leurs erreurs et pour poser des questions, demander aux élèves s'ils veulent repasser l'interrogation. Si oui, discuter de la date de passation. AAL	1, 2, 3, 4, 5 ou 6

Le premier item rentre dans notre catégorie AOL. L'enseignant considère que l'interrogation sommative de fin de chapitre clôture son travail en tant qu'enseignant. Le fait d'inscrire la cote est également en lien avec sa logique que la cote et la réussite sont sources de motivation. Nous sommes ici dans une logique d'évaluation sommative, sanctionnant l'apprentissage avec une cote globale représentant la matière acquise et non acquise.

Le deuxième item rentre dans notre catégorie AFL. L'enseignant montre que son but principal est l'apprentissage et non la cote. En corrigeant et en faisant repasser l'épreuve, le but est de montrer que le processus d'apprentissage ne s'arrête pas à l'interrogation sommative : cette dernière est aussi une forme de renseignement sur les apprentissages des élèves dans le but de réguler son enseignement. Cela vise un meilleur apprentissage des élèves via l'évaluation.

Le troisième item (AAL) montre que l'enseignant place les élèves au centre de l'évaluation puisque c'est à eux de décider si une nouvelle interrogation est à planifier et également du moment de la date de passation. L'élève doit donc monitorer son apprentissage : réfléchir à ce qui est le mieux pour lui, les actions à mettre en place (planifier son étude, la façon de réviser, ...) et prendre une décision à communiquer à l'enseignant.

Thème : remédiation en classe	
Comment s'y prendre si une décision de remédiation en classe a été prise ?	

Faire une activité de remédiation en groupant les élèves en classe sur base des compétences acquises ou non lors de l'évaluation. AFL	1, 2, 3, 4, 5 ou 6
Inviter les élèves, qui ont échoué à l'évaluation, à revoir le chapitre chez eux. AOL	1, 2, 3, 4, 5 ou 6
Créer des zones en classe liées aux compétences non acquises et laisser les élèves se grouper selon ce qu'ils estiment devoir retravailler. AAL	1, 2, 3, 4, 5 ou 6

Le 1^{er} item : AFL. Le focus est mis sur l'apprentissage continu grâce à une analyse faite par le professeur. La cote n'influence pas mais ce sont bien les compétences acquises et non acquises remarquée par la correction du professeur dans l'interrogation qui guident les groupements.

Le 2^e item : AOL. Dans cette approche, l'interrogation sommative est censée représenter la fin du processus d'apprentissage. Le professeur n'a plus de rôle à jouer dans l'apprentissage des élèves. Ceux-ci doivent donc combler les lacunes subsistantes de leur côté. De plus, le professeur qui met la réussite comme objectif principal, ne demandera qu'aux élèves ayant échoué de retravailler la matière.

Le 3^e item : AAL. Comme dans le premier item, les groupes sont effectués par compétences. Mais ici ce ne sont pas les compétences dont la non-acquisition a été vue dans l'interrogation qui guident le placement, mais ce sont l'auto-analyse et les objectifs personnels de l'élève. Ceci montre la volonté de l'enseignant de placer l'élève au centre de l'évaluation, de lui laisser monitorer son apprentissage via une réflexion personnelle sur ses besoins.

Thème : investissement des élèves	
Comment, à partir de cette situation, réussir à augmenter l'investissement des élèves ?	
Lors du prochain chapitre, utiliser les cotes obtenues dans cette évaluation pour motiver les élèves à augmenter leur investissement. AOL	1, 2, 3, 4, 5 ou 6
Lors du prochain chapitre, modifier sa propre façon d'évaluer pour augmenter l'investissement des élèves. AFL	1, 2, 3, 4, 5 ou 6
Demander aux élèves de réfléchir individuellement à leur investissement à l'aide de questions guides sur le temps consacré à travailler le chapitre, l'étalement des séances d'études et la participation en classe. Leur	1, 2, 3, 4, 5 ou 6

proposer ensuite de se fixer des objectifs personnels pour la prochaine évaluation. Être disponible pour les élèves qui voudraient être assistés dans cette tâche. AAL	
--	--

Le 1^{er} item : AOL. Les sources de motivation, pour ce profil, étant la réussite et la cote, il est logique d'utiliser également ces sources pour augmenter l'investissement des élèves. On ne parle pas ici uniquement de la cote en tant que feedback mais de l'idée d'"utiliser les cotes" qui signifie également motiver les 16 élèves en échec à davantage agir pour réussir et féliciter ceux qui ont réussi. Les laboratoires cognitifs ont montré que cet implicite était bien compris.

Le 2^e item : AFL. Cela pourrait paraître étrange et non adapté à l'AFL car il n'y a aucune référence aux compétences acquises et non acquises. Ici, nous traitons un domaine hors de la portée de contrôle d'un enseignant qui se place comme principal évaluateur et qui vise principalement l'acquisition des compétences. Cependant, il considère l'interrogation comme des informations qu'il doit traiter pour l'élève mais également pour son enseignement. C'est dans ce deuxième but, qu'il est logique de se remettre en question pour améliorer ses méthodes d'enseignement dans ce profil AFL.

Le 3^e item : AAL. Pour montrer que l'évaluateur et l'agent principal de son apprentissage est l'élève, l'auto-analyse doit également porter sur des actions non disciplinaires comme l'investissement. L'élève peut sembler un peu seul dans cette action et peu aidé par le professeur. Cependant nous sommes ici avec une notion où le professeur ne peut pas avoir un grand contrôle excepté l'aiguiller à l'aide de questions guides. Un élève peut sembler s'investir alors que non et le contraire est également possible. Or, même si l'élève n'est pas le seul juge, et que le professeur analyse également ce qu'il se passe pendant l'enseignement, c'est l'élève qui est considéré comme celui qui sait le mieux juger son investissement. Laisser du temps à l'élève pour la réflexion personnelle nous paraît dans ce cas-ci une action possible pour soutenir son auto-évaluation.

Thème : remise en question de l'évaluation	
Comment s'y prendre si on veut vérifier qu'il n'y a pas eu un problème de conception de l'évaluation ?	

Vérifier si la cote finale obtenue par chaque élève représente bien le niveau d'acquisition global de la matière, selon ce que chacun a produit dans l'évaluation. AOL	1, 2, 3, 4, 5 ou 6
Vérifier si les questions du contrôle évaluent les compétences visées. AFL	1, 2, 3, 4, 5 ou 6
Demander aux élèves d'énoncer une critique de l'évaluation (cela peut même se faire de façon écrite et anonyme). AAL	1, 2, 3, 4, 5 ou 6

Le 1^{er} item : AOL. Le principal objectif dans la conception et la correction d'une interrogation, c'est qu'elles engendrent une cote représentative de la réussite de l'élève. Si la conception est à revoir, c'est l'élément le plus important à vérifier pour l'AOL.

Le 2^e item : AFL. Le principal objectif étant l'acquisition des compétences par les élèves, il est logique de vérifier si les questions de l'évaluation ne font pas intervenir d'autres concepts qui pourraient empêcher la vérification des compétences acquises.

Le 3^e item : AAL. Impliquer l'élève dans la critique de l'évaluation montre que le professeur lui laisse une place centrale dans l'évaluation. Cette action permettra à l'élève d'énoncer en quoi l'interrogation reçue a été différente des attentes de l'élève.

Thème : discussion avec les pairs	
Comment gérer la discussion post évaluation entre pairs ?	
Discuter des cotes avec l'équipe pédagogique pour vérifier si le phénomène se répète également chez eux. AOL	1, 2, 3, 4, 5 ou 6
Discuter avec des collègues de la même discipline des compétences non acquises en vue d'échanger sur les pratiques pédagogiques. AFL	1, 2, 3, 4, 5 ou 6
Permettre aux élèves de discuter et échanger sur leur façon de se préparer aux évaluations. AAL	1, 2, 3, 4, 5 ou 6

Le 1^{er} item : AOL. Les enseignants étant les principaux évaluateurs, il est logique de se retourner vers d'autres enseignants. La cote étant le meilleur représentant du niveau de réussite, la discussion avec les collègues va porter sur la réussite des dernières interrogations chez eux. Même si cette discussion n'énoncera pas obligatoirement les cotes des élèves, elle aura pour but de vérifier si cette tendance basse dans les cotes est uniquement chez un enseignant ou plusieurs. Les enseignants, dans ce cas, chercheraient une homogénéité dans les

attentes non atteintes par les élèves dans d'autres matières par rapport à la norme édictée par l'institution.

Le 2^e item : AFL. Les enseignants étant les principaux évaluateurs, il est logique de se retourner vers d'autres enseignants. Une action possible aurait été de réfléchir avec d'autres collègues si les prérequis présents dans d'autres matières avaient été acquis. Cela aurait montré à nouveau comme objectif principal l'acquisition des compétences par les élèves. Nous n'avons pas gardé cette idée car cela n'aurait pas fonctionné dans toute évaluation et pour toute matière enseignée. Nous sommes revenus sur le deuxième objectif qui est d'améliorer son enseignement. Dans cette idée, il est dans la conception AFL de remettre en question ses propres pratiques pédagogiques en les comparant avec celles de collègues de la même discipline. L'expression "même discipline" a été conservée malgré le fait que cela signifie pour certains enseignants de parler avec des collègues d'une autre école mais il nous semblait important de lier la pratique pédagogique à la matière enseignée.

Le 3^e item : AAL. Les élèves étant les principaux évaluateurs dans ce profil, il est logique, dans la conception AAL, de se retourner vers d'autres élèves. Avec cette formulation, la discussion peut être faite dans différentes modalités (en groupe, en classe, avec ou sans le professeur) selon les précédentes expériences faites avec les élèves. Le professeur est donc à disposition pour donner des aides mais ce sont les élèves les principaux artisans de leur préparation à l'évaluation.

Thème : Feedback	
Qu'est-ce qu'il faut écrire comme retour aux élèves sur leur interrogation et qui doit le faire ?	
Renseigner l'élève sur le fait qu'il a réussi ou échoué à l'évaluation en écrivant uniquement la cote globale, sans faire de commentaire. AOL	1, 2, 3, 4, 5 ou 6
Écrire sur l'évaluation de chaque élève des retours personnalisés sur les types d'erreurs effectués et des conseils sur la manière de les travailler. AFL	1, 2, 3, 4, 5 ou 6
À l'aide de la copie corrigée de son évaluation et d'un correctif, faire réaliser en classe, à chaque élève, une auto-analyse sur la compréhension, l'application et le transfert des concepts dans divers types de tâches. Être disponible pour les élèves qui demandent de l'aide. AAL	1, 2, 3, 4, 5 ou 6

1^{er} item : AOL. On considère l'interrogation sommative comme la fin du processus d'apprentissage. Il ne sert donc à rien d'y mettre un commentaire dans la conception AOL. De plus, la cote est considérée comme l'information la plus facile à comprendre d'un coup d'œil. Enfin, l'expression "a réussi ou échoué" montre également comme objectif la réussite, sans référence à la matière qui devait être apprise.

2^e item : AFL. Malgré la notion d'interrogation sommative, la volonté est de montrer que l'apprentissage ne s'arrête pas là. L'enseignant comme évaluateur principal va donc effectuer une analyse pour indiquer à l'élève non seulement ce qui n'a pas été acquis mais également comment le travailler.

3^e item : AAL. L'élève étant le principal évaluateur, c'est à lui d'effectuer l'analyse fine de sa compréhension. Cela permettra d'identifier les concepts réellement acquis ou non. Cependant, l'enseignant n'est pas absent de cette réflexion car il aide à cette identification grâce au correctif, il est présent pour l'élève pour savoir si c'est plus un souci de compréhension, d'application, ou de transfert.

Thème : rôle de l'évaluation	
Comment gérer une prise de parole sur le rôle de l'évaluation ?	
Rappeler l'importance de la réussite scolaire pour leur avenir. AOL	1, 2, 3, 4, 5 ou 6
Rappeler que les évaluations servent à savoir où ils en sont dans l'acquisition de leurs compétences mais qu'ils peuvent utiliser ces informations pour progresser. AFL	1, 2, 3, 4, 5 ou 6
Demander aux élèves de réfléchir seuls à l'utilisation de l'évaluation pour améliorer leurs apprentissages. Se servir des réflexions personnelles des élèves pour lancer une discussion de réflexion commune. AAL	1, 2, 3, 4, 5 ou 6

1^{er} item : AOL. Dans l'AOL, nous sommes dans la recherche de la réussite dans toutes ses formes à travers la cote. Le but donc à énoncer pour une personne qui veut montrer l'importance de l'évaluation est que la réussite scolaire passe par la réussite des évaluations. La réussite scolaire leur permettra d'avoir plus de facilité de réussir professionnellement (diplôme, réputation, ...).

2^e item : AFL. L'interrogation sommative est également considérée comme celle indiquant des points à noter dans le cahier de cote mais ce n'est pas le plus important. Dans l'AFL, le plus important est l'apprentissage des compétences. Il faut donc voir cette interrogation, même si

c'est peut-être la dernière du chapitre, comme une nouvelle prise d'information servant à continuer l'apprentissage par la suite.

3^e item : AAL. Ici, le plus important dans l'analyse d'une interrogation est de tirer des leçons pour le développement personnel. Pour que l'élève ne soit pas à nouveau livré à lui-même, l'enseignant va diriger la discussion commune sur le rôle de l'évaluation dans les apprentissages. Cependant, pour que les idées viennent principalement des élèves, nous l'avons précédé d'un moment de réflexion individuelle.

c. Littérature en lien avec les derniers items de la partie 2

→ DeLuca ; Valiquette ; Coombs ; LaPointe McEwan & Luhanga 2016 : Les enseignants ont davantage l'envie d'apprendre des méthodes contemporaines d'évaluation que des méthodes traditionnelles. --> *évaluation avec le numérique ?*

→ Mélange “numérique” et “nouveau paradigme AAL” → par rapport à la question TALIS sur l'autoévaluation (25% des enseignants FWB qui le font contre 50% OCDE27, sans forcément d'utilisation du numérique) p.16/33 du livret “enseigner aujourd'hui”.

J'ai confiance en mes capacités à utiliser des outils numériques pour évaluer.	1, 2, 3, 4, 5 ou 6
J'ai besoin de formations pour m'aider à utiliser des outils numériques pour évaluer.	1, 2, 3, 4, 5 ou 6
J'ai confiance en mes capacités à utiliser des outils numériques pour aider les élèves à s'auto-évaluer ou évaluer leurs pairs.	1, 2, 3, 4, 5 ou 6
J'ai besoin de formations pour m'aider à utiliser des outils numériques pour aider les élèves à s'auto-évaluer ou évaluer leurs pairs.	1, 2, 3, 4, 5 ou 6
J'ai confiance en mes capacités à utiliser les outils numériques pour m'aider à différencier les évaluations.	1, 2, 3, 4, 5 ou 6
J'ai besoin de formations pour m'aider à utiliser les outils numériques pour m'aider à différencier les évaluations.	1, 2, 3, 4, 5 ou 6

→ Graham, De Bruin, Lassig, Spandagou 2021 : Nécessité de méthodes pédagogiques capables de rencontrer les besoins d'une diversité d'apprenants (Rose, 2015). La philosophie de la différenciation : adapter l'enseignement pour rencontrer les besoins d'une diversité d'élèves dans des classes hétérogènes (Tomlinson, 2001, 2017). Compréhension, adoption et mise en œuvre de la différenciation par les enseignants est inégale.

→ Dispositifs d'apprentissages personnalisés (12% en FWB contre 35% en OCDE27)

→ Ces 3 items nous permettront d'enrichir les résultats TALIS 2024 disant que les enseignants n'utilisent que peu le numérique pour évaluer. (12% des enseignants du premier degré en FWB disent utiliser les ressources et outils numériques pour évaluer les apprentissages des élèves, p.19/33 du livret “enseigner aujourd'hui, contre 45% dans l'OCDE27) et nous permettront de questionner les pratiques actuelles des enseignants.

J'ai confiance en mes capacités à utiliser l'évaluation pour gérer l'hétérogénéité de mes classes.	1, 2, 3, 4, 5 ou 6
J'ai besoin de formations pour m'aider à utiliser l'évaluation pour gérer l'hétérogénéité de mes classes.	1, 2, 3, 4, 5 ou 6

→ Christoforidou et Kyriakides 2021 : Différenciation seulement présente en stade 4 (dernier stade) de développement professionnel --> compétence + dure à maîtriser/développer

→ DeLuca ; Valiquette ; Coombs ; LaPointe McEwan & Luhanga 2016 : Il existe des différences significatives en fonction des stades de carrière : Les enseignants novices seraient davantage portés sur l'évaluation standardisée. Les enseignants « mid-career » seraient davantage portés sur des évaluations plus équitables. Les enseignants « seniors » seraient davantage portés vers les évaluations plus différenciées. Les enseignants adoptent un style plus complexe [« student-focused approach to assessment fairness » → évaluations différenciées] à mesure qu'ils gagnent de l'expérience sur terrain. P.367

→ Graham, De Bruin, Lassig, Spandagou 2021 : Nécessité de méthodes pédagogiques capables de rencontrer les besoins d'une diversité d'apprenants (Rose, 2015). La philosophie de la différenciation : adapter l'enseignement pour rencontrer les besoins d'une diversité d'élèves dans des classes hétérogènes (Tomlinson, 2001, 2017). Compréhension, adoption et mise en œuvre de la différenciation par les enseignants est inégale.

→ Graham, De Bruin, Lassig, Spandagou 2021 : Les enseignants différencient moins le produit que le contenu et le processus (Adami, 2004) et ce même dans le cadre de recherches sur l'évaluation différenciée (Kaur et al., 2019)

→ Maroy et Cattonar 2002 : Différencier pour répondre aux besoins de chaque élève (particulariser ses méthodes d'apprentissage mais pas ses objectifs) → Pouvoir remettre en question les contenus et maîtriser les compétences de l'acte d'enseigner, le « savoir-enseigner »

J'ai confiance en mes capacités à collaborer avec mes collègues pour évaluer.	1, 2, 3, 4, 5 ou 6
J'ai besoin de formations pour m'aider à collaborer avec mes collègues pour évaluer.	1, 2, 3, 4, 5 ou 6

→ Wyatt-Smith, Klenowski & Gunn 2010 : Le développement d'une identité d'évaluateur se fait via un groupe de travail et des négociations avec les pairs.

→ Wyatt-Smith, Klenowski & Gunn 2010 : Au-delà de servir les meilleurs intérêts des initiatives d'évaluation à l'échelle du système, ce partage [via modération] permet aux enseignants de revoir leurs pratiques d'enseignement et d'évaluation, se concentrer sur la question de la qualité, et, le cas échéant, modifier la pratique.

→ DeLuca ; Rickey ; Coombs 2021 : Les dynamiques de l'évaluation sont façonnées par des compréhensions individuelles et locales (micro-cultures), les croyances et pratiques qui viennent du système comme les études à large échelles (macro-cultures). Les profs naviguent entre ces 2 pôles pour soutenir l'apprentissage et les réalisations des élèves.

→ Graham, De Bruin, Lassig, Spandagou 2021 : Une forte culture collaborative à travailler ensemble pour planifier, pour l'enseignement, pour améliorer l'établissement et la formation influence statistiquement positivement les profs à utiliser la différenciation et celle-ci influence leur sentiment d'efficacité personnelle à enseigner à tous les Es (Goddard & Kim, 2018)

→ Maroy et Cattonar 2002 : Travailler en équipe et développer des pratiques institutionnelles au sein de la gestion collective de la vie scolaire de son établissement

- Développer la collaboration avec collègues **et** avec le Pouvoir Organisateur, inspection, direction, parents, chercheurs, ...
- Isolement vu comme source d'inefficacité
- Nouvelles compétences : communiquer, gérer un groupe, écouter des avis divergents, négocier, ... et élaborer des projets en collégialité, gérer des projets en collégialité, ajuster, évaluer, ...

J'ai confiance en mes capacités à changer mes pratiques d'évaluation si j'en ai besoin (nouvelles exigences sociétales ou politiques, besoins des élèves, ...).	1, 2, 3, 4, 5 ou 6
J'ai besoin de formations pour m'aider à changer mes pratiques d'évaluation si j'en ai besoin (nouvelles exigences sociétales ou politiques, besoins des élèves, ...).	1, 2, 3, 4, 5 ou 6

→ Pasquini et Morales Villabona 2022 : L'évaluation sommative reste, dans la majorité des systèmes éducatifs, la *démarche d'évaluation que les profs privilégient et que les établissements imposent* (Klenowski et Wyatt-Smith, 2014; OCDE, 2012)

→ Pasquini et Morales Villabona 2022 : la Hiérarchie ou les parents sont très influents sur possible développements de pratiques évaluatives innovantes (Fulmer et al., 2015)

→ Shepard 2000 : Pour que les professeurs puissent apporter des changements significatifs aux convictions pédagogiques et aux pratiques qui les accompagnent, ils devront eux-mêmes essayer et réfléchir à de nouvelles approches dans le contexte de leurs propres classes (Putnam & Borko, 1997)

J'ai confiance en mes capacités à défendre mes choix concernant mes pratiques évaluatives.	1, 2, 3, 4, 5 ou 6
J'ai besoin de formations pour m'aider à défendre mes choix concernant mes pratiques évaluatives.	1, 2, 3, 4, 5 ou 6

→ Pasquini et Morales Villabona 2022 : La Hiérarchie ou les parents sont très influents sur possible développement de pratiques évaluatives innovantes (Fulmer et al., 2015)

J'ai confiance en mes capacités à utiliser les résultats d'épreuves externes pour améliorer la qualité de mon enseignement.	1, 2, 3, 4, 5 ou 6
J'ai besoin de formations pour m'aider à utiliser les résultats d'épreuves externes pour améliorer la qualité de mon enseignement.	1, 2, 3, 4, 5 ou 6

→ Wyatt-Smith, Klenowski & Gunn 2010 : Il y a des débats entre les tests locaux et les tests à grande échelle --> les résultats ne sont pas pris en compte de la même manière ...

→ DeLuca ; Rickey ; Coombs 2021 : Les dynamiques de l'évaluation sont façonnées par des compréhensions individuelles et locales (micro-cultures), les croyances et pratiques qui viennent du système comme les études à large échelles (macro-cultures). Les professeurs naviguent entre ces 2 pôles pour soutenir l'apprentissage et les réalisations des élèves.

J'ai confiance en mes capacités à utiliser ce que j'ai appris en formation (initiale ou continue) pour améliorer la qualité de mes pratiques évaluatives.	1, 2, 3, 4, 5 ou 6
---	--------------------

→ Christoforidou et Kyriakides 2021: Nombreuses interventions de développement professionnel des professeurs se concentrent sur l'amélioration des bases de connaissances et compétences des professeurs et pas sur les opportunités de mise en pratique dans leurs pratiques évaluatives (Darling-Hammond et al., 2019) --> *quid suivi des apprentissages de formations ?*

→ Christoforidou et Kyriakides 2021 : Ce que les enseignants savent ne définit pas nécessairement ce qu'ils pratiquent (Darling Hammond et al., 2019 ; Greenwood et Abbott, 2001 ; Kennedy, 2016)

→ Kennedy 2016 : Il y a un problème dans la mise en œuvre : les professeurs sont sortis de leurs salles de classe pour se former et on espère qu'ils mettront en pratique. Or ils ont de nouvelles idées et connaissances mais continuent à mettre en pratique des idées différentes, par habitude, sans remarquer la contradiction (Kennedy 1999) --> toute nouvelle idée apprise en développement professionnel nécessite l'adoption mais aussi l'abandon d'une autre.

VI. Analyse statistique de la construction de la première partie

a. Fiabilité par approche

Partie AOL

Analyse de fiabilité

Statistiques de fidélité de l'objet

	α de Cronbach	ω de McDonald
échelle	0.571	0.578

[3]

Statistiques de fidélité de l'objet

	Si l'item est retiré	
	α de Cronbach	ω de McDonald
2	0.535	0.553
6	0.547	0.557
8	0.440	0.461
11	0.592	0.597
14	0.565	0.571
17	0.528	0.539
20	0.499	0.511

Partie AFL

Analyse de fiabilité

Statistiques de fidélité de l'objet

	α de Cronbach	ω de McDonald
échelle	0.560	0.573

[3]

Statistiques de fidélité de l'objet

	Si l'item est retiré	
	α de Cronbach	ω de McDonald
3	0.518	0.533
5	0.547	0.561
9	0.483	0.500
12	0.489	0.516
15	0.481	0.486
18	0.573	0.585
21	0.549	0.567

Partie AAL

Analyse de fiabilité

Statistiques de fidélité de l'objet

	α de Cronbach	ω de McDonald
échelle	0.624	0.642

[3]

Statistiques de fidélité de l'objet

	Si l'item est retiré	
	α de Cronbach	ω de McDonald
4	0.634	0.649
7	0.595	0.626
10	0.534	0.552
13	0.546	0.579
16	0.631	0.640
19	0.600	0.628
22	0.555	0.570

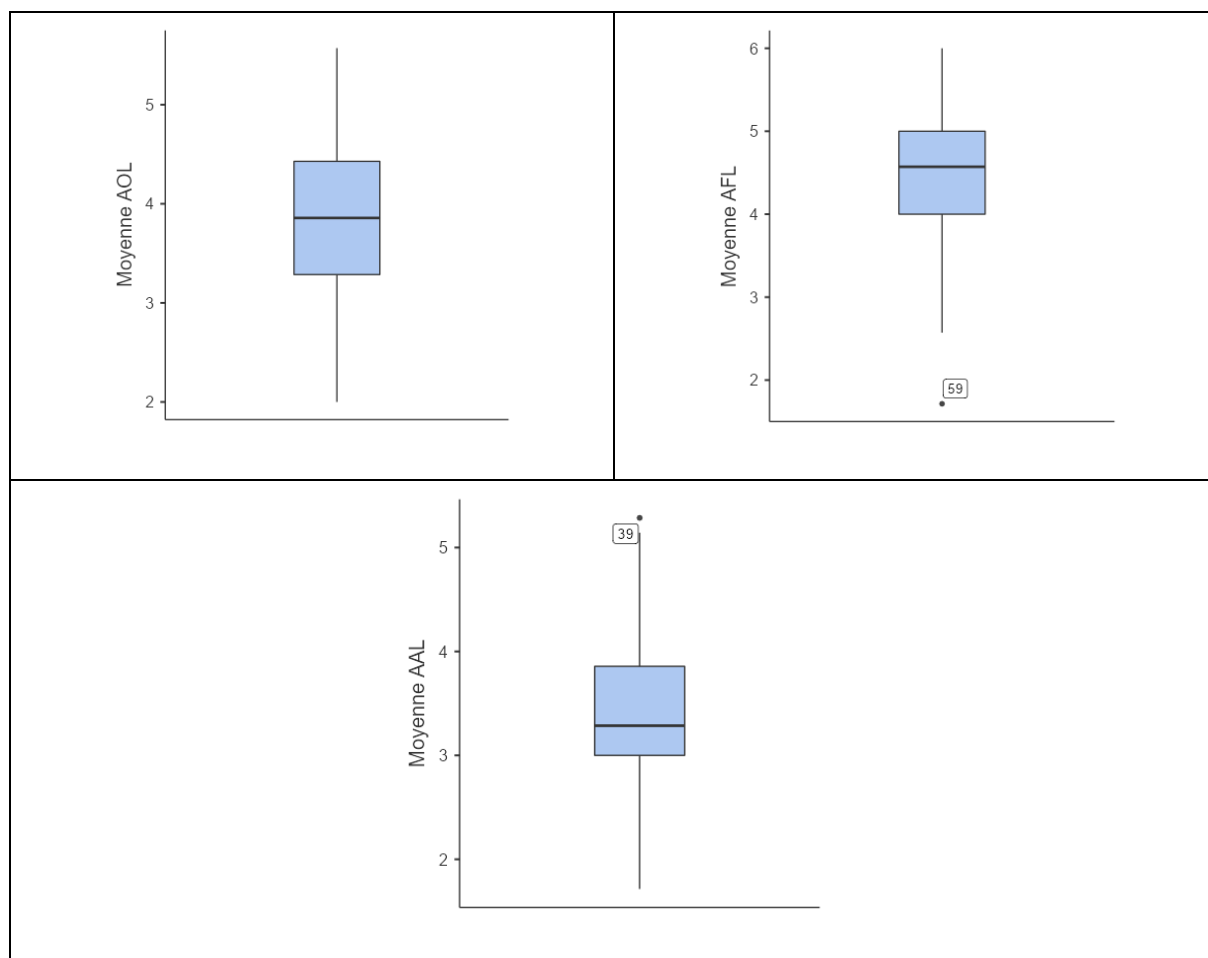
b. Analyse factorielle confirmatoire : contribution des items

Contributions des facteurs

Facteur	Indicateur	Estimation	Erreur standard	Z	p	Estimation standard
AOL	AOL1	0.525	0.247	2.12	0.034	0.317
	AOL2	0.471	0.187	2.52	0.012	0.312
	AOL3	1.042	0.278	3.75	< .001	0.620
	AOL4	0.390	0.214	1.83	0.068	0.280
	AOL5	0.478	0.180	2.65	0.008	0.388
	AOL6	0.568	0.249	2.28	0.022	0.394
	AOL7	0.739	0.220	3.36	< .001	0.538
AFL	AFL1	0.615	0.183	3.37	< .001	0.425
	AFL2	0.498	0.195	2.55	0.011	0.304
	AFL3	0.708	0.175	4.05	< .001	0.478
	AFL4	0.799	0.193	4.13	< .001	0.500
	AFL5	0.685	0.147	4.66	< .001	0.547
	AFL6	0.306	0.164	1.87	0.062	0.222
	AFL7	0.351	0.130	2.71	0.007	0.347
AAL	AAL1	0.494	0.221	2.23	0.026	0.295
	AAL2	0.510	0.174	2.93	0.003	0.354
	AAL3	1.036	0.193	5.37	< .001	0.653
	AAL4	0.831	0.178	4.67	< .001	0.568
	AAL5	0.510	0.190	2.69	0.007	0.342
	AAL6	0.498	0.203	2.45	0.014	0.309
	AAL7	0.887	0.175	5.06	< .001	0.620

[4]

c. Boîtes à moustache des scores en AOL, AFL et AAL



VII. Détermination du nombre de clusters

a. 2 clusters vs 3 clusters

2 clusters

Sum of squares Table

	Value
Cluster 1	101.5
Cluster 2	47.7
Between clusters	63.6
Total	212.8

Clustering Table

Cluster No	Count
1	71
2	34

Centroids of clusters Table

Cluster No	Moyenne AOL	Moyenne AFL	Moyenne AAL
1	1.00	3.772	4.871
2	2.00	3.853	3.662

3 clusters

Sum of squares Table

	Value
Cluster 1	37.7
Cluster 2	41.3
Cluster 3	27.8
Between clusters	106.0
Total	212.8

[3]

Clustering Table

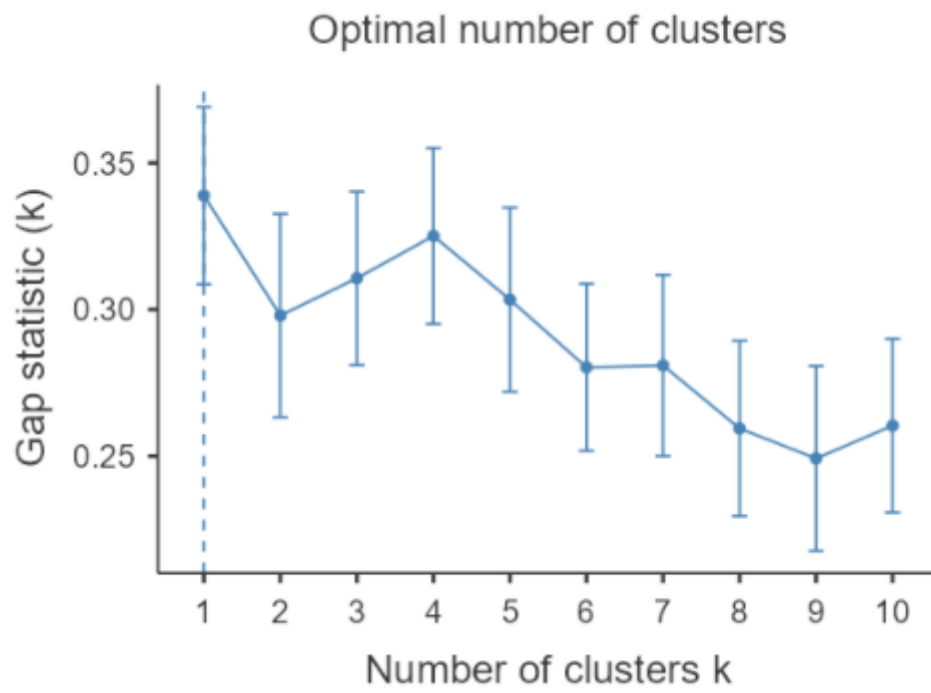
Cluster No	Count
1	40
2	39
3	26

[3]

Centroids of clusters Table

	Moyenne AOL	Moyenne AFL	Moyenne AAL
1	4.371	5.008	3.704
2	2.944	4.491	3.773
3	4.199	3.647	2.478

b. Nombre optimal de clusters selon Jamovi



VIII. Les 3 clusters

a. Par rapport au sexe

	Cluster 1	Cluster 2	Cluster 3
Femmes	45,5%	35,1%	19,5%
	Cluster 1	Cluster 2	Cluster 3
Hommes	17,9%	42,9%	39,3%

b. Par rapport à l'âge/années d'expériences

Statistiques descriptives

	Clustering	Âge	Années d'expérience
Moyenne	3	37.8	13.0
	2	37.6	13.2
	1	41.3	13.8
Ecart-type	3	8.53	7.01
	2	8.72	8.92
	1	10.8	7.13

c. Par rapport aux matières enseignées

	Cluster 1	Cluster 2	Cluster 3	
Français	10	8	4	22
Mathématiques	5	9	7	21
Sciences	10	8	2	20
Technologie	0	0	1	1
Langues	3	3	3	9
Géo/Histoire	3	0	6	9
Latin	1	0	0	1
Ed. art	2	1	0	3
Ed. phys.	1	3	0	4
Cours philo	0	2	0	2
(Secrétariat-)Bureautique	2	1	3	6
Remédiations	0	1	0	1
Communication dans les spécialisé	0	1	0	1
Je n'enseigne pas	0	1	0	1
Sciences-éco/éco	1	1	0	2
Cours commerciaux	1	0	0	1
Informatique	1	0	0	1
	40	39	26	105

d. Par rapport au titre pédagogique

Qualification	Clustering	Quantités
1	3	20
	2	27
	1	28
2	3	3
	2	3
	1	4
3	3	0
	2	1
	1	0
4	3	0
	2	0
	1	1
5	3	2
	2	4
	1	4
7	3	0
	2	0
	1	1
8	3	0
	2	0
	1	1
9	3	1
	2	4
	1	1

1 = AESI 2 = Master 3 = AEES 4 = CAPAES 5 = CAP

7 = discipline sans pédagogique 8 = ni discipline ni pédagogique

9 = autre

e. Par rapport à la filière

Fréquences de Filière			
Filière	Clustering	Quantités	
1	3	0	1 = Premier degré différencié
	2	1	
	1	0	
2	3	17	2 = Premier degré commun ou général
	2	25	
	1	30	
3	3	1	3 = technique ou artistique de transition
	2	1	
	1	2	
4	3	4	4 = technique ou artistique de qualification
	2	6	
	1	2	
5	3	4	5 = professionnel
	2	3	
	1	3	
6	3	0	6 = spécialisé
	2	2	
	1	0	
7	3	0	7 = DASPA
	2	1	
	1	3	

f. Premier choix de carrière

Fréquences de 1er choix de carrière

1er choix de carrière	Clustering	Quantités
1	3	17
	2	27
	1	30
2	3	9
	2	12
	1	10

1 = Oui

2 = Non

g. Souvenir d'un cours sur l'évaluation pendant la formation initiale

Fréquences de Cours éval FI

Cours éval FI	Clustering	Quantités
1	3	17
	2	24
	1	22
2	3	7
	2	11
	1	13
3	3	2
	2	4
	1	5

1 = Oui

2 = Non

3 = Je n'ai pas l'AESI

h. Souvenir d'un cours sur l'évaluation pendant la formation continue

Fréquences de Cours éval FC

Cours éval FC	Clustering	Quantités
1	3	14
	2	14
	1	23
2	3	12
	2	25
	1	17

1 = Oui 2 = Non

IX. Analyse statistique de la construction de la deuxième partie

a. Analyse factorielle confirmatoire des types de self-efficacy

Ajustement du modèle

Test d'ajustement exact

χ^2	ddl	P
846	399	<.001

Mesures ajustées

CFI (indice d'ajustement comparatif)	TLI	SRMR	Approximation de RMSE	IC90% de la RMSEA	
				Borne inf	Supérieur
0.763	0.742	0.0795	0.103	0.0936	0.113

Contributions des facteurs

Facteur	Indicateur	Estimation	Erreur standard	Z	P
SE MOSCODEE	MOSCO SE1P1	0.617	0.0875	7.05	<.001
	MOSCO SE1P2	0.588	0.0716	8.21	<.001
	MOSCO SE2	0.583	0.0870	6.70	<.001
	MOSCO SE3	0.410	0.0715	5.73	<.001
	MOSCO SE4P1	0.491	0.0631	7.79	<.001
	MOSCO SE4P2	0.687	0.0736	9.33	<.001
	MOSCO SE5	0.327	0.0633	5.17	<.001
	MOSCO SE6	0.560	0.0647	8.65	<.001
	MOSCO SE7P1	0.544	0.0744	7.31	<.001
	MOSCO SE7P2	0.577	0.0848	6.80	<.001
	MOSCO SE8	0.449	0.0701	6.41	<.001
	MOSCO SE9	0.607	0.0794	7.65	<.001

Contributions des facteurs

Facteur	Indicateur	Estimation	Erreur standard	Z	P
SE AOL	SE AOL1	0.605	0.0650	9.31	<.001
	SE AOL2	0.710	0.0822	8.63	<.001
	SE AOL3	0.617	0.0763	8.09	<.001
	SE AOL4	0.592	0.0666	8.89	<.001
	SE AOL5	0.732	0.1021	7.18	<.001
	SE AOL6	0.615	0.0775	7.93	<.001
SE AFL	SE AFL1	0.653	0.0805	8.11	<.001
	SE AFL2	0.555	0.0837	6.63	<.001
	SE AFL3	0.659	0.0744	8.86	<.001
	SE AFL4	0.635	0.0839	7.57	<.001
	SE AFL5	0.671	0.0880	7.63	<.001
	SE AFL6	0.726	0.0889	8.16	<.001
SE AAL	SE AAL1	0.935	0.1077	8.68	<.001
	SE AAL2	0.606	0.0963	6.29	<.001
	SE AAL3	0.835	0.0984	8.49	<.001
	SE AAL4	0.911	0.0986	9.24	<.001
	SE AAL5	0.904	0.1081	8.37	<.001
	SE AAL6	0.841	0.1268	6.63	<.001

Covariances des facteurs

		Estimation	Erreur standard	Z	P
SE MOSCODEE	SE MOSCODEE	1.000 ^a			
	SE AOL	0.919	0.0310	29.63	<.001
	SE AFL	0.889	0.0379	23.47	<.001
	SE AAL	0.581	0.0787	7.39	<.001
SE AOL	SE AOL	1.000 ^a			
	SE AFL	0.936	0.0355	26.34	<.001
	SE AAL	0.613	0.0789	7.76	<.001
SE AFL	SE AFL	1.000 ^a			
	SE AAL	0.718	0.0661	10.87	<.001
SE AAL	SE AAL	1.000 ^a			

^a paramètres fixes

Contributions des facteurs - indices de modification

	SE MOSCODEE	SE AOL	SE AFL	SE AAL
MOSCO SE1P1		0.8579	0.0154	2.45059
MOSCO SE1P2		3.3984	4.5294	0.28919
MOSCO SE2		0.0245	0.6115	1.73520
MOSCO SE3		0.3293	0.7560	0.16485
MOSCO SE4P1		1.7234	1.9808	0.11012
MOSCO SE4P2		0.9082	0.5134	0.81942
MOSCO SE5		0.8662	0.1412	0.23664
MOSCO SE6		0.0326	0.0351	0.30272
MOSCO SE7P1		4.1410	2.6374	0.87791
MOSCO SE7P2		0.4490	1.4665	0.75984
MOSCO SE8		11.5688	11.0031	8.20703
MOSCO SE9		0.5937	1.5342	0.73708
SE AOL1	7.01153		0.4276	1.82154
SE AOL2	0.23383		0.9758	0.17417
SE AOL3	0.44696		0.1845	0.69504
SE AOL4	0.87152		0.4783	3.53216
SE AOL5	0.78454		1.5911	1.96620

Contributions des facteurs - indices de modification

	SE MOSCODEE	SE AOL	SE AFL	SE AAL
SE AOL6	1.28160		3.2028	1.93319
SE AFL1	0.02960	7.7257		0.00982
SE AFL2	0.00782	1.7555		0.12362
SE AFL3	1.30105	2.7764		0.05383
SE AFL4	2.21658	20.9973		2.63379
SE AFL5	1.82939	0.4878		0.00300
SE AFL6	1.56590	2.3262		2.46349
SE AAL1	6.56261	9.2459	7.9736	
SE AAL2	5.14802	8.3824	9.8612	
SE AAL3	1.52715	2.5413	1.9764	
SE AAL4	0.52599	0.3010	0.4169	
SE AAL5	4.47248	7.5272	7.6667	
SE AAL6	2.09439	4.4495	2.9198	

b. Analyse factorielle confirmatoire des types de besoins de formations

Ajustement du modèle

Test d'ajustement exact

χ^2	ddl	p
1132	399	<.001

Mesures ajustées

CFI (indice d'ajustement comparatif)	TLI	SRMR	Approximation de RMSE	IC90% de la RMSEA	
				Borne inf	Supérieur
0.752	0.730	0.0733	0.132	0.123	0.141

Contributions des facteurs

Facteur	Indicateur	Estimation	Erreur standard	Z	p
BF MOSCODEE	MOSCO BF1P1	0.906	0.131	6.90	<.001
	MOSCO BF1P2	0.935	0.127	7.36	<.001
	MOSCO BF2	1.085	0.123	8.83	<.001
	MOSCO BF3	0.888	0.115	7.73	<.001
	MOSCO BF4P1	1.035	0.107	9.67	<.001
	MOSCO BF4P2	1.086	0.119	9.09	<.001
	MOSCO BF5	0.877	0.108	8.10	<.001
	MOSCO BF6	1.104	0.113	9.76	<.001
	MOSCO BF7P1	1.080	0.111	9.75	<.001
	MOSCO BF7P2	1.138	0.117	9.74	<.001
	MOSCO BF8	0.972	0.104	9.34	<.001
	MOSCO BF9	1.152	0.111	10.37	<.001

Contributions des facteurs

Facteur	Indicateur	Estimation	Erreur standard	Z	p
BF AOL	BF AOL1	1.002	0.108	9.32	<.001
	BF AOL2	1.084	0.111	9.80	<.001
	BF AOL3	1.057	0.113	9.33	<.001
	BF AOL4	0.974	0.104	9.34	<.001
	BF AOL5	1.115	0.121	9.21	<.001
	BF AOL6	0.969	0.110	8.85	<.001
BF AFL	BF AFL1	1.138	0.115	9.91	<.001
	BF AFL2	1.033	0.109	9.44	<.001
	BF AFL3	1.102	0.102	10.79	<.001
	BF AFL4	1.033	0.116	8.88	<.001
	BF AFL5	1.062	0.115	9.25	<.001
	BF AFL6	1.062	0.112	9.51	<.001
BF AAL	BF AAL1	1.130	0.125	9.08	<.001
	BF AAL2	1.046	0.110	9.55	<.001
	BF AAL3	1.118	0.120	9.35	<.001
	BF AAL4	1.158	0.123	9.39	<.001
	BF AAL5	1.190	0.124	9.62	<.001
	BF AAL6	0.947	0.132	7.16	<.001

Covariances des facteurs

		Estimation	Erreur standard	Z	p
BF MOSCODEE	BF MOSCODEE	1.000 ^a			
	BF AOL	0.928	0.0248	37.5	<.001
	BF AFL	0.929	0.0241	38.5	<.001
	BF AAL	0.779	0.0491	15.9	<.001
BF AOL	BF AOL	1.000 ^a			
	BF AFL	1.000	0.0166	60.2	<.001
	BF AAL	0.873	0.0383	22.8	<.001
BF AFL	BF AFL	1.000 ^a			
	BF AAL	0.922	0.0299	30.9	<.001
BF AAL	BF AAL	1.000 ^a			

^a paramètres fixes

Contributions des facteurs - indices de modification

	BF MOSCODEE	BF AOL	BF AFL	BF AAL
MOSCO BF1P1		8.0715	5.80671	2.33852
MOSCO BF1P2		9.0140	6.24057	1.37410
MOSCO BF2		5.6112	4.69548	4.40303
MOSCO BF3		0.4760	0.01455	1.02172
MOSCO BF4P1		0.1213	0.00116	0.09718
MOSCO BF4P2		3.0680	2.49120	0.82777
MOSCO BF5		0.1127	8.21e-4	0.15424
MOSCO BF6		1.4110	0.53341	0.01512
MOSCO BF7P1		0.0764	0.00116	0.25117
MOSCO BF7P2		0.3204	0.19047	0.18075
MOSCO BF8		36.2292	32.88239	20.23272
MOSCO BF9		3.8751	3.74785	5.16954
BF AOL1	5.3990		7.09911	4.60146
BF AOL2	0.0278		0.57948	0.20065
BF AOL3	1.8047		0.32597	0.01442
BF AOL4	0.9069		1.39735	3.45352
BF AOL5	0.0402		0.16580	0.22156
BF AOL6	4.1380		0.26986	0.16000
BF AFL1	4.2873	1.3174		0.00102
BF AFL2	2.1504	0.4176		0.63147

Contributions des facteurs - indices de modification

	BF MOSCODEE	BF AOL	BF AFL	BF AAL
BF AFL3	1.8195	2.0372		0.86710
BF AFL4	5.4569	4.1782		15.97377
BF AFL5	0.4682	0.1071		0.75903
BF AFL6	0.3172	11.0757		9.61074
BF AAL1	0.5711	3.3236	2.79048	
BF AAL2	7.3525	18.0123	14.68888	
BF AAL3	1.4832	2.1553	2.42619	
BF AAL4	4.2192	3.8779	3.90196	
BF AAL5	8.6904	12.2843	11.58565	
BF AAL6	6.8595	4.6544	5.23980	

c. Matrice de corrélation des SE

Matrice de corrélation

				moyenne Général	SE	moyenne AOL	SE	moyenne AFL	SE	moyenne SE AAL
moyenne Général	SE	Rho Spearman	de	—						
		ddl		—						
		valeur p		—						
moyenne AOL	SE	Rho Spearman	de	0.789***		—				
		ddl		103		—				
		valeur p		<.001		—				
moyenne AFL	SE	Rho Spearman	de	0.711***		0.767***		—		
		ddl		103		103		—		
		valeur p		<.001		<.001		—		
moyenne AAL	SE	Rho Spearman	de	0.528***		0.576***		0.664***		—
		ddl		103		103		103		—
		valeur p		<.001		<.001		<.001		—

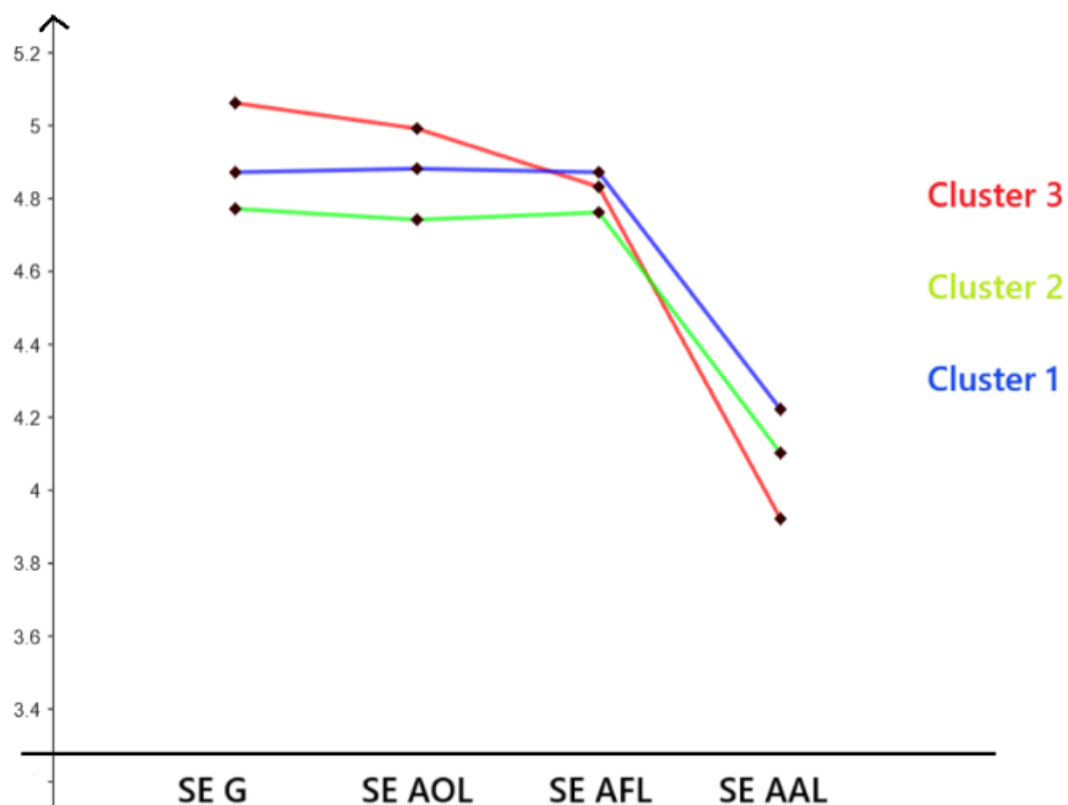
Note. * p < .05, ** p < .01, *** p < .001

X. Le self-efficacy par rapport aux clusters

a. Les 4 types de self-efficacy

Statistiques descriptives

	Clustering	moyenne MOSCODEE	SE	moyenne SE AOL	moyenne SE AFL	moyenne SE AAL
Moyenne	3	5.06		4.99	4.83	3.92
	2	4.77		4.74	4.76	4.10
	1	4.87		4.88	4.87	4.22
Ecart- type	3	0.585		0.723	0.775	1.22
	2	0.588		0.745	0.774	0.788
	1	0.522		0.642	0.687	0.768



b. Self-efficacy général

Statistiques descriptives

	MOSCO SE1 p1+p2	MOSCO SE2	MOSCO SE3	MOSCO SE4 P1+P2	MOSCO SE5	MOSCO SE6	MOSCO SE7 P1+P2	MOSCO SE8	MOSCO SE9
Moyenne	4.64	4.74	5.12	4.81	5.20	4.74	5.02	5.03	4.77
Ecart- type	0.799	0.961	0.768	0.708	0.671	0.760	0.795	0.765	0.901
Minimum	2.50	2	3	3.00	4	3	2.50	3	2
Maximum	6.00	6	6	6.00	6	6	6.00	6	6

	Clustering	MOSCO SE1 p1+p2	MOSCO SE2	MOSCO SE3	MOSCO SE4 P1+P2	MOSCO SE5	MOSCO SE6	MOSCO SE7 P1+P2	MOSCO SE8	MOSCO SE9
Moyenne	3	4.79	5.08	5.12	5.04	5.23	4.85	5.42	5.08	4.88
	2	4.58	4.44	5.10	4.68	5.10	4.69	4.85	5.00	4.74
	1	4.61	4.83	5.15	4.80	5.28	4.72	4.94	5.03	4.72
Ecart- type	3	0.839	0.744	0.864	0.615	0.587	0.925	0.578	0.845	0.864
	2	0.823	0.968	0.718	0.721	0.718	0.694	0.836	0.761	1.02
	1	0.755	1.01	0.770	0.732	0.679	0.716	0.802	0.733	0.816

c. Self-efficacy AOL

Statistiques descriptives

	SE AOL1	SE AOL2	SE AOL3	SE AOL4	SE AOL5	SE AOL6
Moyenne	4.68	4.84	4.81	5.22	4.62	4.96
Ecart-type	0.778	0.962	0.878	0.784	1.14	0.887
Minimum	3	1	2	3	1	2
Maximum	6	6	6	6	6	6

Statistiques descriptives

	Clustering	SE AOL1	SE AOL2	SE AOL3	SE AOL4	SE AOL5	SE AOL6
Moyenne	3	4.88	4.88	5.08	5.31	4.92	4.85
	2	4.56	4.82	4.72	5.13	4.26	4.92
	1	4.65	4.83	4.72	5.25	4.78	5.08
Ecart-type	3	0.766	0.952	0.796	0.618	0.977	1.05
	2	0.852	1.07	0.826	0.923	1.29	0.929
	1	0.700	0.874	0.960	0.742	1.000	0.730

d. Item SE AOL 5

Fréquences de SE AOL5

SE AOL5	Clustering	Fréquences	% du Total	% cumulés
1	3	0	0.0%	0.0%
	2	1	1.0%	1.0%
	1	0	0.0%	1.0%
2	3	0	0.0%	1.0%
	2	2	1.9%	2.9%
	1	0	0.0%	2.9%
3	3	2	1.9%	4.8%
	2	9	8.6%	13.3%
	1	4	3.8%	17.1%
4	3	7	6.7%	23.8%
	2	8	7.6%	31.4%
	1	13	12.4%	43.8%
5	3	8	7.6%	51.4%
	2	12	11.4%	62.9%
	1	11	10.5%	73.3%
6	3	9	8.6%	81.9%
	2	7	6.7%	88.6%
	1	12	11.4%	100.0%

e. Self-efficacy AFL

Statistiques descriptives

	SE AFL1	SE AFL2	SE AFL3	SE AFL5	SE AFL6
Moyenne	4.91	4.77	4.90	4.87	4.65
Ecart-type	0.911	0.912	0.876	0.991	1.02

f. Self-efficacy AAL

Statistiques descriptives

	SE AAL1	SE AAL2	SE AAL3	SE AAL4	SE AAL5	SE AAL6
Moyenne	3.80	4.61	3.96	4.12	3.84	4.28
Ecart-type	1.24	1.02	1.13	1.16	1.23	1.36

Statistiques descriptives

	Clustering	SE AAL1	SE AAL2	SE AAL3	SE AAL4	SE AAL5	SE AAL6
Moyenne	3	3.69	4.35	4.00	3.96	3.85	3.65
	2	3.87	4.69	3.82	4.08	3.79	4.36
	1	3.80	4.70	4.08	4.28	3.88	4.60
Ecart- type	3	1.52	1.44	1.33	1.37	1.41	1.72
	2	1.10	0.922	1.07	1.01	1.15	1.29
	1	1.18	0.758	1.05	1.15	1.20	1.03

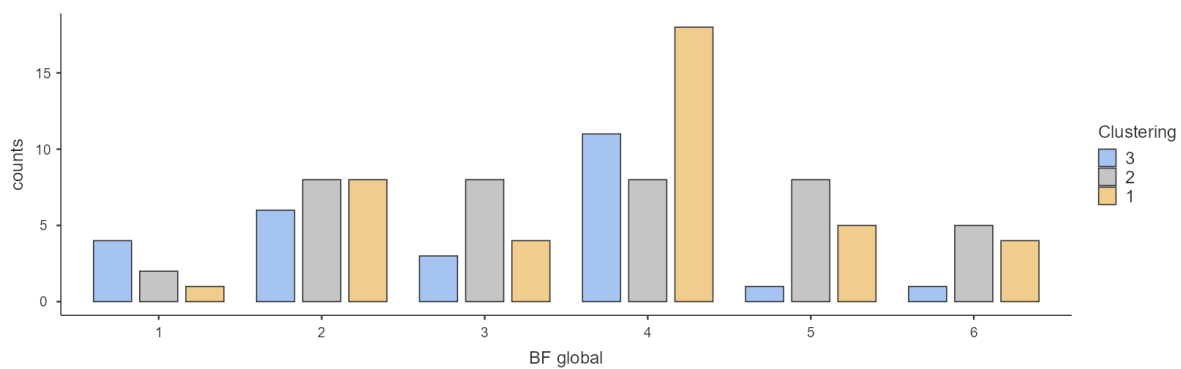
g. Item SE AAL6

Fréquences de SE AAL6

SE AAL6	Clustering	Fréquences	% du Total	% cumulés
1	3	5	4.8%	4.8%
	2	1	1.0%	5.7%
	1	0	0.0%	5.7%
2	3	2	1.9%	7.6%
	2	4	3.8%	11.4%
	1	1	1.0%	12.4%
3	3	3	2.9%	15.2%
	2	3	2.9%	18.1%
	1	4	3.8%	21.9%
4	3	7	6.7%	28.6%
	2	9	8.6%	37.1%
	1	14	13.3%	50.5%
5	3	5	4.8%	55.2%
	2	16	15.2%	70.5%
	1	12	11.4%	81.9%
6	3	4	3.8%	85.7%
	2	6	5.7%	91.4%
	1	9	8.6%	100.0%

XI. Les besoins de formations par rapport aux clusters

a. Item besoin de formation global



Statistiques descriptives

	Clustering	BF global
Moyenne	3	3.08
	2	3.69
	1	3.75
Ecart-type	3	1.35
	2	1.47
	1	1.28

Fréquences de BF global

BF global	Clustering	Fréquences	% du Total	% cumulés
1	3	4	3.8%	3.8%
	2	2	1.9%	5.7%
	1	1	1.0%	6.7%
2	3	6	5.7%	12.4%
	2	8	7.6%	20.0%
	1	8	7.6%	27.6%
3	3	3	2.9%	30.5%
	2	8	7.6%	38.1%
	1	4	3.8%	41.9%
4	3	11	10.5%	52.4%
	2	8	7.6%	60.0%
	1	18	17.1%	77.1%
5	3	1	1.0%	78.1%
	2	8	7.6%	85.7%
	1	5	4.8%	90.5%
6	3	1	1.0%	91.4%
	2	5	4.8%	96.2%
	1	4	3.8%	100.0%

b. Les 4 types de besoins de formations

Statistiques descriptives

	Clustering	moyenne BF MOSCODEE	moyenne BF AOL	moyenne BF AFL	moyenne BF AAL
Moyenne	3	2.29	2.37	2.48	2.78
	2	2.81	2.71	2.78	3.42
	1	2.83	2.55	2.73	3.07
Ecart- type	3	0.918	1.01	1.04	1.12
	2	1.16	1.16	1.27	1.18
	1	1.02	1.09	1.13	1.14

c. Besoin de formation général

Statistiques descriptives

	MOSCO BF1 p1+p2	MOSCO BF2	MOSCO BF3	MOSCO BF4 P1+P2 (2)	MOSCO BF5	MOSCO BF6	MOSCO BF7 P1+P2	MOSCO BF8	MOSCO BF9
Moyenne	3.22	2.90	2.20	2.75	2.28	2.87	2.44	2.45	2.71
Ecart- type	1.34	1.46	1.32	1.21	1.26	1.39	1.31	1.26	1.39

Statistiques descriptives

	Clustering	MOSCO BF1 p1+p2	MOSCO BF2	MOSCO BF3	MOSCO BF4 P1+P2 (2)	MOSCO BF5	MOSCO BF6	MOSCO BF7 P1+P2	MOSCO BF8	MOSCO BF9
Moyenne	3	2.77	2.08	1.92	2.29	2.08	2.62	1.94	2.12	2.62
	2	3.09	3.21	2.28	2.96	2.44	2.92	2.63	2.69	2.82
	1	3.64	3.13	2.30	2.85	2.25	2.98	2.59	2.42	2.67
Ecart- type	3	1.05	1.09	1.29	1.00	1.26	1.39	0.887	1.18	1.30
	2	1.38	1.56	1.28	1.29	1.33	1.42	1.44	1.40	1.48
	1	1.37	1.40	1.38	1.20	1.19	1.37	1.36	1.13	1.38

d. Item MOSCO BF2

Fréquences de MOSCO BF2

MOSCO BF2	Clustering	Fréquences	% du Total	% cumulés
1	3	9	8.6%	8.6%
	2	6	5.7%	14.3%
	1	6	5.7%	20.0%
2	3	11	10.5%	30.5%
	2	11	10.5%	41.0%
	1	8	7.6%	48.6%
3	3	1	1.0%	49.5%
	2	2	1.9%	51.4%
	1	9	8.6%	60.0%
4	3	5	4.8%	64.8%
	2	12	11.4%	76.2%
	1	11	10.5%	86.7%
5	3	0	0.0%	86.7%
	2	5	4.8%	91.4%
	1	4	3.8%	95.2%
6	3	0	0.0%	95.2%
	2	3	2.9%	98.1%
	1	2	1.9%	100.0%

e. Item MOSCO BF1

Fréquences de MOSCO BF1P1

MOSCO BF1P1	Clustering	Fréquences	% du Total	% cumulés
1	3	6	5.7%	5.7%
	2	5	4.8%	10.5%
	1	5	4.8%	15.2%
2	3	5	4.8%	20.0%
	2	14	13.3%	33.3%
	1	5	4.8%	38.1%
3	3	8	7.6%	45.7%
	2	5	4.8%	50.5%
	1	7	6.7%	57.1%
4	3	6	5.7%	62.9%
	2	8	7.6%	70.5%
	1	13	12.4%	82.9%
5	3	1	1.0%	83.8%
	2	4	3.8%	87.6%
	1	5	4.8%	92.4%
6	3	0	0.0%	92.4%
	2	3	2.9%	95.2%
	1	5	4.8%	100.0%

Fréquences de MOSCO BF1P2

MOSCO BF1P2	Clustering	Fréquences	% du Total	% cumulés
1	3	4	3.8%	3.8%
	2	5	4.8%	8.6%
	1	3	2.9%	11.4%
2	3	7	6.7%	18.1%
	2	11	10.5%	28.6%
	1	7	6.7%	35.2%
3	3	6	5.7%	41.0%
	2	6	5.7%	46.7%
	1	4	3.8%	50.5%
4	3	7	6.7%	57.1%
	2	10	9.5%	66.7%
	1	15	14.3%	81.0%
5	3	1	1.0%	81.9%
	2	4	3.8%	85.7%
	1	7	6.7%	92.4%
6	3	1	1.0%	93.3%
	2	3	2.9%	96.2%
	1	4	3.8%	100.0%

f. Besoins de formations AOL

Statistiques descriptives

	BF AOL1	BF AOL2	BF AOL3	BF AOL4	BF AOL5	BF AOL6
Moyenne	2.96	2.61	2.66	2.14	2.65	2.38
Ecart-type	1.30	1.36	1.37	1.26	1.46	1.30

Statistiques descriptives

	Clustering	BF AOL1	BF AOL2	BF AOL3	BF AOL4	BF AOL5	BF AOL6
Moyenne	3	2.58	2.38	2.50	1.92	2.27	2.58
	2	3.05	2.67	2.74	2.38	3.03	2.41
	1	3.13	2.70	2.67	2.05	2.52	2.23
Ecart-type	3	1.10	1.27	1.36	1.02	1.34	1.30
	2	1.32	1.46	1.39	1.46	1.51	1.45
	1	1.38	1.34	1.38	1.18	1.43	1.14

g. Item BF AOL1

Fréquences de BF AOL1

BF AOL1	Clustering	Fréquences	% du Total	% cumulés
1	3	6	5.7%	5.7%
	2	3	2.9%	8.6%
	1	4	3.8%	12.4%
2	3	5	4.8%	17.1%
	2	15	14.3%	31.4%
	1	11	10.5%	41.9%
3	3	9	8.6%	50.5%
	2	5	4.8%	55.2%
	1	10	9.5%	64.8%
4	3	6	5.7%	70.5%
	2	10	9.5%	80.0%
	1	9	8.6%	88.6%
5	3	0	0.0%	88.6%
	2	5	4.8%	93.3%
	1	3	2.9%	96.2%
6	3	0	0.0%	96.2%
	2	1	1.0%	97.1%
	1	3	2.9%	100.0%

h. Besoins de formations AFL

Statistiques descriptives

	Clustering	BF AFL1	BF AFL2	BF AFL3	BF AFL5	BF AFL6
Moyenne	3	2.50	2.38	2.23	2.69	2.58
	2	2.97	2.74	2.59	2.62	3.00
	1	2.75	2.92	2.75	2.50	2.73
Ecart-type	3	1.17	1.27	1.11	1.41	1.39
	2	1.55	1.33	1.33	1.48	1.54
	1	1.45	1.35	1.37	1.30	1.15

i. Besoins de formations AFL : tableaux de fréquences pour chaque item

Fréquences de BF AFL1

BF AFL1	Clustering	Fréquences	% du Total	% cumulés
1	3	6	5.7%	5.7%
	2	6	5.7%	11.4%
	1	10	9.5%	21.0%
2	3	9	8.6%	29.5%
	2	15	14.3%	43.8%
	1	9	8.6%	52.4%
3	3	3	2.9%	55.2%
	2	3	2.9%	58.1%
	1	9	8.6%	66.7%
4	3	8	7.6%	74.3%
	2	6	5.7%	80.0%
	1	6	5.7%	85.7%
5	3	0	0.0%	85.7%
	2	7	6.7%	92.4%
	1	5	4.8%	97.1%
6	3	0	0.0%	97.1%
	2	2	1.9%	99.0%
	1	1	1.0%	100.0%

Fréquences de BF AFL2

BF AFL2	Clustering	Fréquences	% du Total	% cumulés
1	3	9	8.6%	8.6%
	2	6	5.7%	14.3%
	1	7	6.7%	21.0%
2	3	5	4.8%	25.7%
	2	15	14.3%	40.0%
	1	9	8.6%	48.6%
3	3	6	5.7%	54.3%
	2	7	6.7%	61.0%
	1	10	9.5%	70.5%
4	3	5	4.8%	75.2%
	2	6	5.7%	81.0%
	1	9	8.6%	89.5%
5	3	1	1.0%	90.5%
	2	4	3.8%	94.3%
	1	4	3.8%	98.1%
6	3	0	0.0%	98.1%
	2	1	1.0%	99.0%
	1	1	1.0%	100.0%

Fréquences de BF AFL3

BF AFL3	Clustering	Fréquences	% du Total	% cumulés
1	3	9	8.6%	8.6%
	2	7	6.7%	15.2%
	1	10	9.5%	24.8%
2	3	6	5.7%	30.5%
	2	17	16.2%	46.7%
	1	7	6.7%	53.3%
3	3	7	6.7%	60.0%
	2	6	5.7%	65.7%
	1	11	10.5%	76.2%
4	3	4	3.8%	80.0%
	2	4	3.8%	83.8%
	1	8	7.6%	91.4%
5	3	0	0.0%	91.4%
	2	4	3.8%	95.2%
	1	3	2.9%	98.1%
6	3	0	0.0%	98.1%
	2	1	1.0%	99.0%
	1	1	1.0%	100.0%

Fréquences de BF AFL5

BF AFL5	Clustering	Fréquences	% du Total	% cumulés
1	3	6	5.7%	5.7%
	2	10	9.5%	15.2%
	1	11	10.5%	25.7%
2	3	8	7.6%	33.3%
	2	14	13.3%	46.7%
	1	11	10.5%	57.1%
3	3	3	2.9%	60.0%
	2	3	2.9%	62.9%
	1	8	7.6%	70.5%
4	3	7	6.7%	77.1%
	2	6	5.7%	82.9%
	1	8	7.6%	90.5%
5	3	1	1.0%	91.4%
	2	5	4.8%	96.2%
	1	1	1.0%	97.1%
6	3	1	1.0%	98.1%
	2	1	1.0%	99.0%
	1	1	1.0%	100.0%

Fréquences de BF AFL6

BF AFL6	Clustering	Fréquences	% du Total	% cumulés
1	3	7	6.7%	6.7%
	2	7	6.7%	13.3%
	1	7	6.7%	20.0%
2	3	7	6.7%	26.7%
	2	11	10.5%	37.1%
	1	9	8.6%	45.7%
3	3	5	4.8%	50.5%
	2	6	5.7%	56.2%
	1	14	13.3%	69.5%
4	3	5	4.8%	74.3%
	2	8	7.6%	81.9%
	1	9	8.6%	90.5%
5	3	1	1.0%	91.4%
	2	4	3.8%	95.2%
	1	0	0.0%	95.2%
6	3	1	1.0%	96.2%
	2	3	2.9%	99.0%
	1	1	1.0%	100.0%

j. Besoins de formations AAL

Statistiques descriptives

	BF AAL1	BF AAL2	BF AAL3	BF AAL4	BF AAL5	BF AAL6
Moyenne	3.58	2.65	3.23	3.13	3.36	2.82
Ecart-type	1.47	1.32	1.44	1.47	1.49	1.48

Statistiques descriptives

	Clustering	BF AAL1	BF AAL2	BF AAL3	BF AAL4	BF AAL5	BF AAL6
Moyenne	3	3.42	2.23	2.69	2.92	2.77	2.65
	2	3.87	2.87	3.54	3.41	3.74	3.10
	1	3.40	2.70	3.27	3.00	3.38	2.65
Ecart-type	3	1.42	1.27	1.52	1.47	1.34	1.67
	2	1.45	1.36	1.39	1.53	1.53	1.52
	1	1.52	1.26	1.36	1.41	1.44	1.29

k. Items BF AAL1 et BF AAL2 : tableau de fréquences

Fréquences de BF AAL1

BF AAL1	Clustering	Fréquences	% du Total	% cumulés
1	3	4	3.8%	3.8%
	2	2	1.9%	5.7%
	1	6	5.7%	11.4%
2	3	1	1.0%	12.4%
	2	5	4.8%	17.1%
	1	7	6.7%	23.8%
3	3	8	7.6%	31.4%
	2	10	9.5%	41.0%
	1	5	4.8%	45.7%
4	3	8	7.6%	53.3%
	2	7	6.7%	60.0%
	1	11	10.5%	70.5%
5	3	3	2.9%	73.3%
	2	9	8.6%	81.9%
	1	9	8.6%	90.5%
6	3	2	1.9%	92.4%
	2	6	5.7%	98.1%
	1	2	1.9%	100.0%

Fréquences de BF AAL2

BF AAL2	Clustering	Fréquences	% du Total	% cumulés
1	3	11	10.5%	10.5%
	2	3	2.9%	13.3%
	1	8	7.6%	21.0%
2	3	5	4.8%	25.7%
	2	18	17.1%	42.9%
	1	11	10.5%	53.3%
3	3	3	2.9%	56.2%
	2	7	6.7%	62.9%
	1	9	8.6%	71.4%
4	3	7	6.7%	78.1%
	2	6	5.7%	83.8%
	1	10	9.5%	93.3%
5	3	0	0.0%	93.3%
	2	2	1.9%	95.2%
	1	1	1.0%	96.2%
6	3	0	0.0%	96.2%
	2	3	2.9%	99.0%
	1	1	1.0%	100.0%

XII.Lien self-efficacy/besoins de formations [SE/BF]

a. Corrélations faibles binômes SE/BF

Matrice de corrélation

		SE global	BF global
SE global	Rho de Spearman	—	
	ddl	—	
	valeur p	—	
BF global	Rho de Spearman	-0.389***	—
	ddl	103	—
	valeur p	<.001	—

Note. H_a est une corrélation négative

Note. * $p < 0,05$; ** $p < 0,01$; *** $p < 0,001$, unilatéral

Matrice de corrélation

		SE AAL2	BF AAL2
SE AAL2	Rho de Spearman	—	
	ddl	—	
	valeur p	—	
BF AAL2	Rho de Spearman	-0.383***	—
	ddl	103	—
	valeur p	<.001	—

Note. H_a est une corrélation négative

Note. * $p < 0,05$; ** $p < 0,01$; *** $p < 0,001$, unilatéral

Matrice de corrélation

		SE CHANGEMENT PRATIQUES		BF CHANGEMENT PRATIQUES
SE CHANGEMENT PRATIQUES	Rho de Spearman	—		
	ddl	—		
	valeur p	—		
BF CHANGEMENT PRATIQUES	Rho de Spearman	-0.324***		—
	ddl	103		—
	valeur p	<.001		—

Note. H_a est une corrélation négative

Note. * p < 0,05 ; ** p < 0,01 ; *** p < 0,001 , unilatéral

Matrice de corrélation

		P2 SE EXTERNE	P2 BF EXTERNE
P2 SE EXTERNE	Rho de Spearman	—	
	ddl	—	
	valeur p	—	
P2 BF EXTERNE	Rho de Spearman	-0.189	—
	ddl	60	—
	valeur p	0.071	—

b. SE/BF global

Fréquences de SE global

SE global	Fréquences	% du Total	% cumulés
3	3	2.9%	2.9%
4	38	36.2%	39.0%
5	50	47.6%	86.7%
6	14	13.3%	100.0%

Fréquences de BF global

BF global	Fréquences	% du Total	% cumulés
1	7	6.7%	6.7%
2	22	21.0%	27.6%
3	15	14.3%	41.9%
4	37	35.2%	77.1%
5	14	13.3%	90.5%
6	10	9.5%	100.0%

c. SE/BF AAL2

Fréquences de SE AAL2

SE AAL2	Fréquences	% du Total	% cumulés
1	2	1.9%	1.9%
2	3	2.9%	4.8%
3	3	2.9%	7.6%
4	36	34.3%	41.9%
5	43	41.0%	82.9%
6	18	17.1%	100.0%

Fréquences de BF AAL2

BF AAL2	Fréquences	% du Total	% cumulés
1	22	21.0%	21.0%
2	34	32.4%	53.3%
3	19	18.1%	71.4%
4	23	21.9%	93.3%
5	3	2.9%	96.2%
6	4	3.8%	100.0%

d. SE/BF changement de pratiques

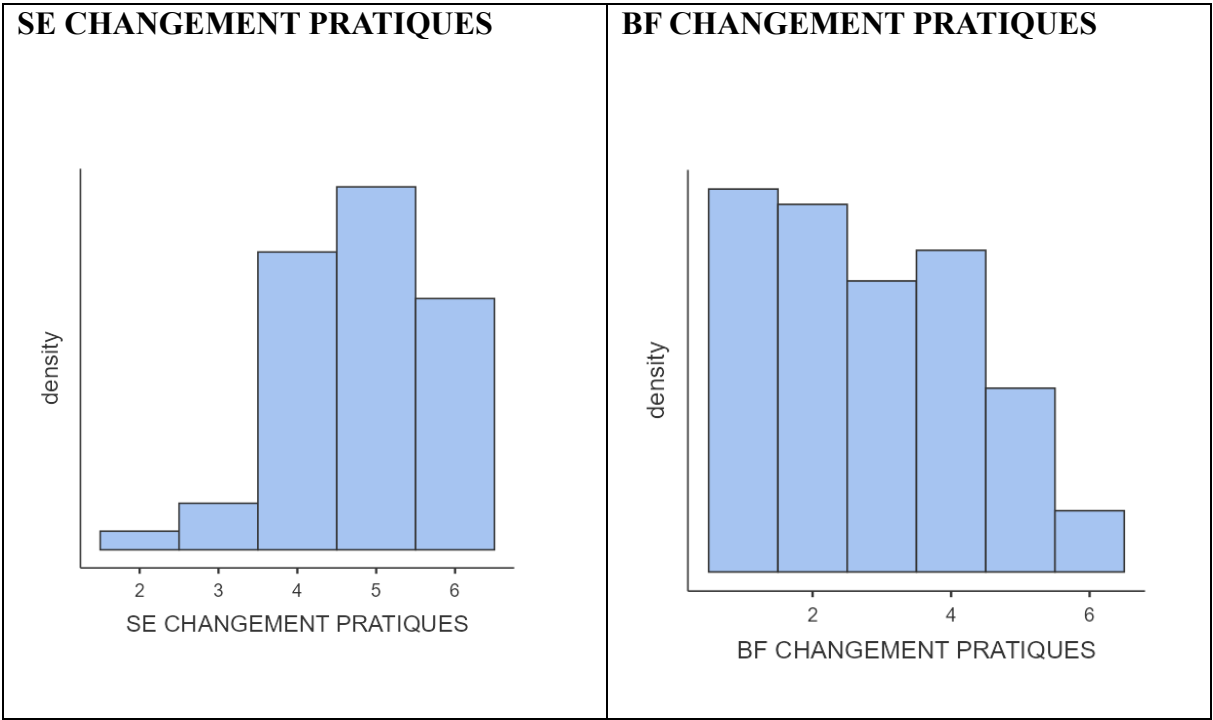
Fréquences de SE CHANGEMENT PRATIQUES

SE CHANGEMENT PRATIQUES	Fréquences	% du Total	% cumulés
2	2	1.9%	1.9%
3	5	4.8%	6.7%
4	32	30.5%	37.1%
5	39	37.1%	74.3%
6	27	25.7%	100.0%

Fréquences de BF CHANGEMENT PRATIQUES

BF CHANGEMENT PRATIQUES	Fréquences	% du Total	% cumulés
1	25	23.8%	23.8%
2	24	22.9%	46.7%
3	19	18.1%	64.8%
4	21	20.0%	84.8%
5	12	11.4%	96.2%
6	4	3.8%	100.0%

Graphes



e. 2 corrélations SE/BF supplémentaires analysées

Matrice de corrélation

		SE AAL6	BF AAL6
SE AAL6	Rho de Spearman	—	
	Ddl	—	
	valeur p	—	
BF AAL6	Rho de Spearman	-0.402***	—
	Ddl	103	—
	valeur p	<.001	—

Note. H_a est une corrélation négative

Note. * $p < 0,05$; ** $p < 0,01$; *** $p < 0,001$, unilatéral

Matrice de corrélation

		MOSCO SE1P1	MOSCO BF1P1	MOSCO SE1P2	MOSCO BF1P2
MOSCO SE1P1	Rho de Spearman	—			
	Ddl	—			
	valeur p	—			
MOSCO BF1P1	Rho de Spearman	-0.598***	—		

Matrice de corrélation

			MOSCO SE1P1	MOSCO BF1P1	MOSCO SE1P2	MOSCO BF1P2
Ddl			103	—		
valeur p			<.001	—		
MOSCO SE1P2	Rho de Spearman		0.598	-0.341***	—	
Ddl			103	103	—	
valeur p			1.000	<.001	—	
MOSCO BF1P2	Rho de Spearman		-0.382***	0.667	-0.552***	—
Ddl			103	103	103	—
valeur p			<.001	1.000	<.001	—

Note. H_a est une corrélation négative

Note. * p < 0,05 ; ** p < 0,01 ; *** p < 0,001 , unilatéral

f. SE/BF AAL6

Fréquences de SE AAL6

SE AAL6	Fréquences	% du Total	% cumulés
1	6	5.7%	5.7%
2	7	6.7%	12.4%
3	10	9.5%	21.9%
4	30	28.6%	50.5%
5	33	31.4%	81.9%
6	19	18.1%	100.0%

Fréquences de BF AAL6

BF AAL6	Fréquences	% du Total	% cumulés
1	25	23.8%	23.8%
2	26	24.8%	48.6%
3	17	16.2%	64.8%
4	21	20.0%	84.8%
5	12	11.4%	96.2%
6	4	3.8%	100.0%

XIII. Modalités de formations

a. Modalités de formations sans clusters

Statistiques descriptives

	MF1 COMP LET PRESE NTIEL ENSE MBLE	MF2 MODU LES PRESE NTIEL ENSE MBLE	MF3 COM PLET EN LIGN E ENSE SEUL	MF4 COM PLET EN LIGN E ENSE MBLE	MF5 MOD ULES EN LIGN E ENSE SEUL	MF6 MOD ULES EN LIGN E ENSE MBLE	MF7 EN CLAS SE ENSE MBLE	MF8 EN CLA SSE EXP ERT	MF9 EN LIGNE ET EN PRESE NTIEL	MF10 WEBI NAIRE S	MF11 MENT ORAT MEM E DISCI PLINE	MF12 MENT ORAT AUTR E DISCI PLINE	MF13 AUTOAPP RENTISSA GE
Mo yen ne	3.69	3.64	3.1 4	2.8 5	3.0 6	2.7 6	4.1 5	3. 8 9	2.98	2.84	3.6 7	2.5 0	3.23
Ec art - typ e	1.06	1.12	1.4 0	1.4 5	1.4 4	1.4 1	0.8 30	1. 0 7	1.42	1.42	1.1 6	1.1 9	1.15

b. Modalités de formations avec clusters

Statistiques descriptives

		MF1 COMP LET PRESE NTIEL ENSEM BLE	MF2 MODU LES PRESE NTIEL ENSEM BLE	MF3 COMP LET EN LIGN E SEUL	MF4 COMP LET EN LIGN E MBLE	MF5 MOD ULES EN LIGN E SEUL	MF6 MODU LES EN LIGN E ENSE MBLE	MF7 EN CLASS E ENSE MBLE	MF8 EN CLAS SE EXPE RT	MF9 EN LIGNE ET PRESE NTIEL	MF10 WEBIN AIRES	MF11 MENT ORAT MEME DISCIP LINE	MF12 MENT ORAT AUTR E DISCIP LINE	MF13 AUTOAPPR ENTISSAG E
Mo yen ne	3	3.42	3.04	2.5 4	2.3 5	2.2 7	2.2 7	4.2 7	3. 5 0	2.19	2.31	3.6 9	2.0 0	2.77
	2	3.82	4.10	3.4 4	2.9 2	3.4 1	2.9 2	4.2 6	4. 0 5	3.31	3.13	3.7 4	2.8 2	3.41
	1	3.73	3.58	3.2 5	3.1 0	3.2 3	2.9 2	3.9 8	3. 9 8	3.17	2.90	3.5 8	2.5 2	3.35
Eca rt- typ e	3	1.17	1.22	1.4 2	1.4 4	1.4 3	1.3 7	0.6 67	1. 2 4	1.36	1.38	1.2 3	1.0 2	1.37
	2	0.94 2	0.82 1	1.3 5	1.3 8	1.2 9	1.3 5	0.9 10	0. 8 2 6	1.32	1.32	1.1 6	1.2 3	1.07
	1	1.09	1.13	1.3 3	1.4 6	1.4 2	1.4 6	0.8 32	1. 1 2	1.39	1.46	1.1 3	1.1 5	1.00