



Mémoire de fin d'études à la HEP-VS

Titre

***Représentations et modes de gestion de quatre enseignants
autour des biais liés à l'évaluation sommative***

Auteure

Alizée Buthey

Directrice de mémoire

Renée Sandmeier

Lieu et date de dépôt du mémoire

St-Maurice, le 21 février 2011

Remerciements

Tout d'abord, nous souhaitons remercier notre directrice de mémoire, Mme Sandmeier, qui nous a accompagnée, aidée et soutenue tout au long de cette importante tâche. Nous sommes reconnaissante du temps qu'elle nous a accordé, de sa disponibilité pour répondre à nos questions et pour nous guider dans la construction du mémoire de fin d'études. Merci pour tout le travail accompli.

Ensuite, nous aimerions préciser notre reconnaissance à notre mère, enseignante au cycle d'orientation, qui a contribué à notre questionnement autour de l'objet de la présente recherche. Elle a participé à nos discussions concernant le thème choisi, en donnant son opinion, en nous encourageant à continuer sur cette voie, et en récoltant certains documents que nous avons pu utiliser pour élaborer le cadre conceptuel du travail. Elle a également répondu positivement à notre demande pour tester notre guide d'entretien, avant que nous commencions notre recherche auprès des enseignants sélectionnés. Grâce à elle, nous avons donc eu l'occasion d'améliorer nos questions avant de réaliser les quatre rencontres prévues dans le cadre de ce mémoire.

Puis, nous voudrions ajouter combien nous sommes reconnaissante de la confiance que nous ont accordée les différentes directions des écoles qui nous ont autorisée à mener nos entretiens auprès de leur personnel enseignant. Nous remercions également les quatre personnes qui ont participé à notre recherche, et qui se sont montrées très disponibles et flexibles pour ces rencontres. Merci pour leurs réponses authentiques et honnêtes.

Enfin, les derniers remerciements que nous désirons accorder vont à notre ami, qui nous supporte, nous soutient et nous encourage au quotidien, ainsi qu'à nos amies et collègues de la formation initiale à la HEPvs, Véréne Gex-Collet, Roxane Bürki et Pascale Vrugink, qui ont su nous écouter, être présentes pour nous et nous « coacher » dans les moments plus difficiles et durant les périodes de grande fatigue. Pour tout cela, merci à tous !

Pour terminer cette partie, nous dirions encore un grand MERCI aux membres de notre famille proche, qui ont toujours cru en nous et qui nous ont aussi soutenue tout au long de notre formation à la HEPvs.

Résumé

Pour commencer cette synthèse de notre mémoire de fin d'études, nous avons trouvé pertinent de ressortir les deux citations suivantes :

- « Le spectacle de l'injustice m'accable, mais c'est probablement parce qu'il éveille en moi la conscience de la part d'injustice dont je suis capable ». (Georges Bernanos)
- « Il est effrayant de penser que cette chose qu'on a en soi, le jugement, n'est pas la justice. Le jugement, c'est le relatif. La justice, c'est l'absolu. Réfléchissez à la différence entre un juge et juste ». (Victor Hugo *L'homme qui rit*)

Ces extraits démontrent bien notre profond besoin de faire régner la justice en toute occasion, et plus précisément dans le cadre du métier que nous avons choisi, soit l'enseignement. C'est d'ailleurs de ce sentiment prononcé pour la justice qu'est née notre envie de réaliser le travail qui suit. Nous pensons, en tant que future enseignante, qu'il est important d'être juste avec les enfants qui nous sont confiés dans nos classes, c'est pourquoi nous avons voulu approfondir le thème de l'évaluation, et plus particulièrement le sujet des biais de l'évaluation. Bon nombre de personnes ne savent pas vraiment de quoi il s'agit lorsqu'on évoque ce phénomène de cette façon, c'est pourquoi nous pensons qu'il est nécessaire de faire comprendre aux lecteurs, que les biais de l'évaluation sont tous les facteurs qui s'immiscent dans le processus de l'évaluation, et plus précisément, dans l'acte de correction, et qui influencent de façon plus ou moins importante la notation de l'enseignant.

Outre le fait de savoir qu'il existe de tels phénomènes dans l'acte évaluatif, il est intéressant de se pencher sur la nature de ces effets, ainsi que sur l'état de conscience qu'en ont les enseignants qui pratiquent l'évaluation au quotidien. C'est sur ces points-là que notre étude a été menée, afin de pouvoir clarifier et corriger certaines idées reçues concernant les compétences des instituteurs à juger les performances des apprenants. Ainsi, c'est grâce aux entretiens effectués dans le cadre de cette recherche, que l'on pourra mieux comprendre l'approche et les représentations de quatre enseignants, aux profils différents, en matière d'évaluation et en ce qui concerne ces éléments qui biaisent leur manière de noter les élèves.

Enfin, ce que nous avons cherché à démontrer, à travers cette étude, est le fait qu'effectivement, il existe bien des causes qui font commettre certaines erreurs aux enseignants, lorsque ces derniers évaluent, corrigent et notent les apprenants, mais qu'ils en sont conscients, qu'ils y réfléchissent, et surtout, qu'ils essaient de les gérer au mieux, par souci de professionnalisme et par respect pour les élèves qu'ils ont dans leur classe, et qui attendent un jugement correct et justifié de leur titulaire. C'est donc à travers les parties d'analyse et d'interprétation que nous avons pu démontrer cette inquiétude partagée par les quatre personnes interviewées dans le cadre de cette recherche.

Mots – clés :

Evaluation (sommativité / formative)

Objectivité

Biais de l'évaluation

Fiabilité

Notation

Validité

Correction

Gestion des biais

Sommaire

I.	Introduction	6
II.	Partie théorique	7
1.	Problématique	7
2.	Cadre conceptuel	11
2.1.	Evaluation	11
2.1.1.	Définitions	11
2.1.2.	Types d'évaluation	13
2.2.	Biais	17
2.2.1.	Définitions	17
2.2.2.	Types de biais	19
2.2.3.	Liens avec l'évaluation	23
2.3.	Objectivité / Validité / Fiabilité	24
2.3.1.	Définitions	24
2.3.2.	Liens avec l'évaluation	25
2.4.	Représentation	27
2.4.1.	Définition	27
2.4.2.	Liens avec l'évaluation	27
3.	Question de recherche	27
4.	Hypothèses	27
5.	Dispositif méthodologique	28
5.1	Méthodologie	28
5.2	Echantillon choisi	29
III.	Partie empirique	30
1.	Données	30
2.	Analyse des données et interprétation des résultats	30
3.	Vérification des hypothèses et réponse à la question de recherche	46
4.	Analyse critique	47

IV.	Conclusion	48
1.	Apports formatifs	49
2.	Prolongements et perspectives	49
V.	Références bibliographiques	51
VI.	Annexes	54
1.	Annexe I : Guide d'entretien I	54
2.	Annexe II : Guide d'entretien II	56
3.	Attestation d'authenticité	58

I. Introduction

Mise en perspective de la recherche et explication de l'éclairage choisi

C'est dans le cadre de notre formation à la Haute Ecole Pédagogique du Valais que le travail suivant a été réalisé. Afin de démontrer notre capacité à approfondir un sujet et à le comprendre, à mener une recherche, à réfléchir, et à analyser, il nous est demandé de construire un mémoire sur un thème lié au domaine de l'éducation.

C'est donc après plusieurs réflexions que notre choix s'est porté sur le vaste thème de l'évaluation scolaire. Par la suite, il nous a fallu faire une sélection un peu plus précise par rapport à ce large sujet. Pour ce faire, de nombreuses lectures ont été effectuées et cette phase exploratoire a abouti à un aspect très précis de l'acte évaluatif, soit, les effets liés à la pratique de l'évaluation sommative. Notre question de départ fut donc la suivante : « Quelles sont les représentations des enseignants par rapport aux biais relatifs à l'évaluation sommative ? », et par après, nous nous sommes demandée s'il existait des différences dans les conceptions des jeunes enseignants et des enseignants plus expérimentés, concernant ces phénomènes.

Notre choix s'est donc arrêté sur ce thème pour plusieurs raisons. Tout d'abord, l'acte évaluatif a une grande importance dans le cadre scolaire et on peut affirmer qu'il est vraiment complexe, entre autre, puisqu'il concerne plusieurs acteurs : l'enseignant, les élèves, les parents, la direction et, à plus grande échelle, la société. Chacun a donc des représentations différentes de cet objet. Ensuite, depuis des dizaines d'années, on peut remarquer que l'acte évaluatif fait polémique (Hadji, C., 1990, p. 13) et soulève de nombreuses questions et critiques de la part des acteurs concernés et de la société en général (Meyer, G., 2007, p. 15-16). On entend souvent parler d'injustice, de manque de compétences des enseignants, ou encore de favoritisme. Les enseignants sont pourtant formés à l'activité d'évaluer, mais la plupart des gens remettent sans arrêt en question les capacités, la fiabilité et l'objectivité des instituteurs, dans ce domaine.¹ Et enfin, le fait de se pencher sur un sujet comme celui-là nous permet de prendre conscience de l'état actuel des connaissances autour de ce phénomène, de découvrir si les enseignants en ont conscience et de comprendre leur manière de gérer ces biais qui font partie de l'évaluation scolaire.

Ainsi, les buts poursuivis par cette étude sont les suivants : relever les biais qui peuvent fausser le processus évaluatif, découvrir si les jeunes enseignants et les enseignants plus expérimentés ont conscience de l'existence de ces effets, de comparer les représentations de ces deux types d'acteurs scolaires, et d'entrevoir leur façon d'appréhender et de gérer ce phénomène dans leur quotidien.

Contexte de la recherche

Au niveau du contexte institutionnel de la recherche, on peut remarquer que la thématique choisie, soit les effets liés à la pratique de l'évaluation sommative, s'inscrit directement dans le domaine de l'évaluation scolaire. Bien que certains de ces phénomènes se retrouvent aussi dans l'évaluation de type formatif, nous nous arrêterons uniquement sur l'évaluation notée officiellement², puisque c'est celle-ci que nous avons décidé d'approfondir dans ce travail, et parce que, c'est celle-là même qui est tant remise en question par le sens commun. Par rapport au contexte géographique, il s'agit des cantons du Valais et de Vaud, puisque les entretiens de

¹ Source : Discussion avec des étudiants et des parents d'élèves que nous avons en soutien scolaire.

² L'évaluation sommative.

recherche seront menés auprès d'une jeune enseignante chargée de l'enseignement de certaines disciplines dans sept classes de 4P, 5P et 6P, de deux enseignantes plus expérimentées qui enseignent en Valais, et enfin, auprès d'un jeune enseignant travaillant actuellement dans le canton de Vaud. Mais, il nous faut préciser que ce dernier individu a suivi sa formation en Valais. Ce choix est en lien direct avec notre formation professionnelle, puisque celle-ci est suivie dans ce même canton. Enfin, pour le contexte linguistique et culturel de ce mémoire, nous avons décidé de réaliser notre recherche dans la partie francophone du Valais et dans le canton de Vaud, puisque les enseignants interviewés travaillent dans des écoles publiques du Bas-Valais et du village de Villars-sur-Ollon.

Plan du travail

Pour clore cette introduction, nous allons exposer brièvement le plan du travail qui suit. Dans un premier temps, nous expliquons la problématique liée au sujet choisi pour ce mémoire, nous définissons les concepts qui s'y rapportent, puis nous élaborons la question de recherche et quelques hypothèses, et nous terminons la partie théorique du travail en décrivant le dispositif méthodologique. Dans un second temps, nous décrivons la recherche auprès du personnel enseignant, récoltons, analysons et interprétons les données nécessaires pour pouvoir confirmer ou infirmer nos hypothèses de départ, et nous achevons la partie empirique du travail en portant un jugement critique sur l'étude effectuée. Enfin, dans un dernier temps, nous prenons un peu de recul pour élaborer un bilan du travail réalisé, afin d'en ressortir les apports formatifs, les prolongements et les perspectives possibles.

II. Partie théorique

1. Problématique

Pour élaborer ce travail, nous avons choisi de nous positionner du point de vue pédagogique et de porter un regard dans cette direction, sur les effets liés aux pratiques de l'évaluation sommative. Merle nous dit, d'ailleurs, que l'évaluation représente « un idéal de mesure exacte des compétences scolaires » (Merle, P., 1998, p. 5). On fait donc intervenir deux acteurs principaux dans cette procédure : l'enseignant et l'élève, et c'est le rôle du premier de porter un jugement sur les compétences scolaires du second. Ainsi, nous pouvons nous rendre compte que les biais de l'évaluation scolaire relèvent du champ de la pédagogie.

Au niveau des connaissances actuelles, plusieurs études ont été menées autour de ce thème. Des chercheurs se sont déjà intéressés à la notation, aux performances des élèves, mais également à la façon de noter des enseignants, ou encore à l'implication de l'établissement scolaire dans la notation des apprenants. Nous avons alors retenu quelques-unes de ces recherches qui nous paraissaient pertinentes par rapport à ce mémoire. Tout d'abord, il nous faut citer les travaux effectués par Laugier et Weinberg, en 1936, dans le cadre de la commission française pour l'enquête Carnégie sur les examens et concours (Merle, P., 1998, p. 7). Ces études docimologiques ont été réalisées dans le but d'étudier et d'analyser « le degré d'incertitude de la notation professorale » (Merle, P., 1998, 5) et de chercher « à définir des moyens pour la limiter ». D'autres recherches, poursuivant cette même direction, ont été également menées entre 1960 et 1970, notamment par H. Piéron (Merle, P., 1998, p. 5). On peut donc s'apercevoir que le sujet, qui nous préoccupe actuellement, était déjà un point qui faisait débat, un peu avant les années quarante.

Entrons, à présent, un peu plus en profondeur sur l'étude menée par la commission française de l'enquête Carnégie, en 1936, sur le baccalauréat. Pour effectuer cette recherche, Laugier et Weinberg (1936) ont sélectionné un lot de cent copies provenant d'un échantillon représentatif d'étudiants et concernant seulement les examens écrits des six branches scolaires suivantes : latin, composition française, anglais, mathématiques, philosophie et physique. Ces copies ont été soumises à la correction de cinq personnes, ainsi que celle du correcteur du baccalauréat, soit, six individus, au total. Ainsi, la commission de l'enquête Carnégie a pu mesurer l'écart qui pouvait se manifester lors d'une correction réalisée par plusieurs personnes. Les résultats de cette étude ont démontré qu'il existait bien des différences de jugement d'un individu à l'autre. Voici, ci-dessous, les chiffres qui ont été établis :

Tableau 1. – Ecart maximum, écarts moyens et écarts les plus fréquents lors d'une expérience de multi-correction menée sur 100 copies du baccalauréat (session 1930)

Disciplines	Ecart maximum	Moyenne des écarts	Écarts les plus fréquents
Version latine	12 points	2,97	5
Composition française	13	3,29	6 et 7
Anglais	9	2,24	4
Mathématiques	9	2,05	4
Philosophie	12	3,36	5 et 7
Physique	8	1,88	4

Source : Commission d'enquête Carnégie (Merle, P., 1998, p. 8)

C'est donc à partir de cette expérience de multi-correction que la commission française de l'enquête Carnégie a pu mesurer les écarts qui existaient entre les notes attribuées par les six correcteurs sélectionnés. Ainsi, au niveau des résultats obtenus, on peut remarquer que les branches où l'écart de notation est le plus important sont le latin, la composition française et la philosophie, qui sont des disciplines plus difficiles à juger, puisque le panel de réponses aux questions est souvent plus large que celui des branches mathématiques et physique.

Après cette recherche, Laugier et Weinberg (1936) ont conclu que pour obtenir une notation correcte, il fallait prendre au moins septante-huit correcteurs pour la composition française, treize correcteurs pour les mathématiques, cent vingt-sept pour la philosophie, etc... (Merle, P., 1998, p. 9) ; et ceci, afin d'arriver à une moyenne plus juste. On peut donc, encore une fois, constater que les branches dites « plus subjectives »³ nécessitent une correction de la part d'un plus grand nombre de personnes, que les branches plus précises, comme les mathématiques. C'est pourquoi les deux chercheurs cités ci-dessus ont achevé leur travail en précisant que la « recherche d'une note vraie était, en quelque sorte, une utopie » (Merle, P., 1998, p. 9).

Passons, à présent, à une autre étude qui a été réalisée par la section anglaise de la commission Carnégie, cette fois. Cette enquête a été menée sur les examens oraux dans la branche « anglais » (Merle, P., 1998, p. 10). Les chercheurs ont pu constater que s'il existait bien des différences dans la notation d'épreuves écrites, il y en avait d'autant plus dans les interrogations orales (Merle, P., 1998, p. 11). Ils ont remarqué que le « face-à-face » candidat-expert était une variable importante

³ Composition française, philosophie.

à ne pas négliger, puisque celle-ci influençait considérablement la notation. De plus, les individus chargés d'effectuer cette étude se sont rendu compte du fait que les correcteurs d'un examen oral jugeaient différemment les candidats, lorsqu'ils évaluaient, une seconde fois, la performance de ces derniers.

La conclusion de cette enquête a été le fait que le correcteur, quel qu'il soit, ne reste pas « fidèle à lui-même » (Merle, P., p. 11), puisque sa notation varie quelque peu, lors d'un second regard sur la performance des candidats. Les chercheurs ont alors pu définir que la notation d'une épreuve orale se faisait en deux temps : il y a tout d'abord une approche globale où l'expert juge la performance générale de l'étudiant ; puis, il y a un deuxième regard, plus analytique cette fois, afin d'évaluer le candidat selon des critères précis.

Une autre étude a été menée en 1984, pour répondre à la question posée par le recteur de l'Académie de Lille concernant les différences de notation entre les correcteurs. Celui-ci se demandait si le fait que les enseignants se réunissent et discutent des critères qu'ils allaient fixer, avant la correction des copies du baccalauréat, impliquait forcément que la notation établie par ces mêmes différents professeurs soit juste et vraie (Merle P., 1996, p. 217-218). Il avait émis l'hypothèse que si un établissement obtenait un taux de réussite plus bas aux épreuves du baccalauréat, c'est parce que les professeurs avaient noté plus sévèrement les étudiants que leurs collègues, qui corrigeaient les examens dans les autres académies du pays. Ainsi, le recteur pensait que certains enseignants étaient plus exigeants que d'autres, et que les réunions pour l'élaboration de critères de réussite, avant les épreuves du bac, n'apportaient aucune assurance que la notation soit la même pour tous. Cette hypothèse fut plutôt mal reçue par l'ensemble des professeurs.

Le chercheur B. Simler s'est alors attelé à mener une étude par rapport à la notation établie pour les examens du baccalauréat de 1984, dans les académies de Lille, Nancy et Poitiers (Merle, P., 1996, p. 217-224). Il s'est intéressé au phénomène des différences de notation, en se posant les questions suivantes :

- « La réglementation relative à l'organisation des épreuves n'est-elle pas identique pour tous les correcteurs ? »
- « Les désaccords professoraux sur la manière de corriger les copies ne sont-ils pas le signe d'une impossibilité de comportement collectif qui prendrait la forme d'une sévérité ou d'une indulgence commune à chaque correcteur ? »
- « Le recrutement national et la formation des enseignants n'assurent-ils pas des compétences professorales identiques ? » (Merle, P., 1996, p. 218).

De plus, comme susmentionné, le recteur de l'Académie de Lille avait lui-même remis en question la façon de noter des correcteurs lillois, puisque ces derniers avaient l'air d'être plus sévères que les professeurs des autres académies, « notamment pour l'épreuve de sciences économiques et sociales (SES) du baccalauréat B. » (Merle, P., 1996, p. 218).

Pour réaliser cette recherche, Simler (1985) a alors procédé en récoltant les examens de SES du baccalauréat, évalués par les correcteurs lillois et les a soumis à des correcteurs d'autres académies. Huit autres correcteurs ont donc été désignés pour noter 120 copies de la session SES de 1984 de l'Académie de Lille, afin qu'une comparaison puisse être établie par rapport à la notation faite par les correcteurs de Lille. Si l'on s'arrête à une approche globale, les résultats de cette étude ont démontré qu'il existait bien une différence de notation entre les correcteurs de l'Académie de Lille et ceux des autres institutions. Ces premiers

semblent plus sévères que les autres et le taux de réussite s'en retrouve influencé. Mais si l'on analyse la notation faite par chaque correcteur, et copie par copie, on se rend compte que les experts de Lille ne sont pas forcément toujours les plus sévères (Merle, P., 1996, p. 222-224). Il y a des différences de correction selon l'individu qui évalue l'épreuve. On peut alors remarquer qu'il existe simplement une « diversité des lectures professorales » (Merle, P., 1996, p. 224). Ceci prouve le manque de critères d'appréciation des correcteurs⁴, puisqu'il semble y avoir un certain écart entre la manière de noter des personnes désignées pour corriger les examens des étudiants. Mais, il ne faut tout de même pas oublier que « cette incertitude individuelle de la mesure n'exclut pas une certaine cohérence de l'ensemble des correcteurs sur la hiérarchie des compétences des candidats » (Merle, P., 1996, p. 224), cohérence qui « reste cependant assez limitée » (Merle, P., 1996, p. 224).

Enfin, nous avons relevé une dernière recherche qui a été effectuée en 1978, par Noizet et Caverni (De Peretti, Boniface et Legrand, 2005, p. 45), et qui s'approche un peu plus de notre sujet des biais dans l'évaluation. Ces deux chercheurs ont étudié les facteurs qui avaient une influence sur le jugement des enseignants et ont pu souligner les effets qui revenaient le plus souvent et qui « parasitaient » l'acte évaluatif. Mais contrairement aux études précédemment menées, Noizet et Caverni ne sont pas restés sur le bilan fait par leurs prédécesseurs concernant les imperfections liées au domaine de l'évaluation. Ils sont allés chercher plus loin et ont tenté d'analyser les causes de l'incertitude de la notation (Cambon, J., 1979, p. 72). Ainsi l'une des conclusions importantes formulées par ces deux auteurs est le fait que ce n'est pas parce qu'on croit être un bon enseignant, que l'on maîtrise forcément l'acte évaluatif. Ils précisent aussi que ce ne serait pas juste d'axer la formation de ces instituteurs seulement sur le domaine de l'évaluation. Enfin, ils terminent leur travail de recherche sur une note plutôt positive en expliquant qu'il faut bien garder à l'esprit qu'on peut toujours améliorer sa façon d'évaluer, puisque cette compétence fait partie d'un long apprentissage du quotidien (Cambon, J., 1979, p. 73).

Revenons à présent au centre du problème qui nous intéresse dans cette recherche. Ce que nous cherchons à comprendre, en menant ce travail, c'est la manière dont les biais s'immiscent dans l'évaluation scolaire. Nous aimerions prendre conscience de l'ampleur de ce phénomène, découvrir s'il fait partie du quotidien des enseignants, si ces derniers en sont conscients, et surtout, de quelle manière ils gèrent ces effets, pour éviter de tomber dans les pièges liés à l'acte évaluatif. Nous avons donc choisi de nous placer du point de vue pédagogique pour étudier cet élément, puisque nous avons pu découvrir, au fil des lectures effectuées, que les erreurs que les enseignants faisaient en évaluant les apprenants étaient souvent en lien avec ce champ. Par exemple, un enseignant peut mettre une certaine note à un certain élève, parce qu'il a des attentes supérieures par rapport à lui, parce qu'il est fatigué, ou encore, parce qu'il fait un peu de favoritisme pour un des apprenants. Mais il est important de ne pas oublier que ces « erreurs », commises par l'enseignant, sont très souvent inconscientes ou semi-conscientes. L'erreur est humaine, et ce qui semble nécessaire, c'est d'avoir une certaine lucidité sur soi-même.

Pour terminer cette partie concernant la problématique, nous dirions que nous avons pu nous rendre compte que la subjectivité avait une part plus ou moins importante dans le domaine de l'évaluation et que cela pouvait provoquer des écarts de notation assez conséquents. Mais, nous avons pu constater que ce

⁴ En sciences économiques et sociales.

phénomène se produisait, pour la plupart du temps, à l'insu des correcteurs, qui commettaient, inconsciemment, des erreurs de jugement. Ainsi, nous concluons ce point en rappelant l'importance de la réflexion et de la remise en question, dans le métier d'enseignant, car il est nécessaire d'avoir une vision plutôt lucide de sa pratique au quotidien.

2. Cadre conceptuel

Dans cette partie du travail, nous allons développer les concepts principaux en lien avec notre thème de recherche. Tout d'abord, nous nous concentrerons sur le vaste concept d'évaluation, en établissant des liens avec quelques sous-concepts liés à notre thème et à notre question de recherche. Puis, nous entrerons plus précisément dans les biais liés à l'acte évaluatif, afin de les décrire et de démontrer leur importance dans la notation des enseignants. Ensuite, nous consacrerons quelques paragraphes aux qualités de l'évaluation, et ceci, dans le but de les mettre en relation avec l'évaluation scolaire. Ce sont donc les notions d'objectivité, de validité et de fiabilité des notes établies par les instituteurs qui seront abordées, puisqu'il s'agira de démontrer qu'il est indispensable de garantir une certaine cohérence dans sa façon d'attribuer une valeur aux travaux des élèves, pour que la notation soit juste et pertinente. En dernier point, nous expliquerons le terme de « représentation », car notre travail s'articule autour des représentations des enseignants concernant les biais de l'évaluation. Nous allons donc procéder dans un ordre logique en partant du concept d'évaluation et, par la suite, nous reviendrons toujours à ce premier concept central, qui, nous le rappelons, constitue le contexte institutionnel de notre étude, également.

2.1. Evaluation

2.1.1. Définitions

Ce large concept a été défini à plusieurs reprises, par bon nombre d'auteurs. Nous avons retenu les quelques définitions qui nous paraissent les plus pertinentes.

Selon Gilbert de Landsheere (1979), l'évaluation est une « estimation, par une note, d'une modalité ou d'un critère considéré dans un comportement ou un produit » (De Landsheere, G., 1979, p. 122). Il explique que l'évaluation pédagogique est une procédure qui permet de définir si les apprenants ont atteint des objectifs, fixés au préalable, ou non. Ensuite, il ajoute que l'acte évaluatif vise à mesurer le comportement de l'élève, par exemple, de manière quantitative (De Landsheere, G., 1979, 122-123). Il termine sa définition de l'évaluation en introduisant les notions « d'évaluation subjective » et « d'évaluation objective » (De Landsheere, G., 1979, p. 123). Par rapport à ces deux notions, l'auteur explique que l'appréciation est un élément plus ou moins subjectif, car il qualifie l'ensemble des performances d'un individu, dans la globalité, alors que le fait de juger des éléments très précis, sur lesquels on décide de focaliser son attention, relève plus de l'objectivité (De Landsheere, G., 1979, 20).

Selon Claude Guillon (2004), IEN information et orientation, l'évaluation est un élément plutôt équivoque qui sous-entend les notions d'estimation et de valeur. Le mot valeur implique, toujours selon lui, le concept d'échanges et de négociations, puisqu'elle doit inclure les conditions dans lesquelles elle est réalisée ; tandis que le

terme d'estimation se réfère à « la mesure, la recherche de chiffres, de résultats, d'indicateurs, d'évolutions » (Guillon, C., 2004, p. 5). Ainsi, pour cet auteur, on voit que ce vaste concept regroupe deux notions importantes : la mesure et l'interaction entre l'évaluateur et l'évalué. Il s'agit donc, selon lui, d'un processus qui se déroule dans le champ pédagogique, et qui permet de mettre un chiffre sur une performance réalisée.

Pour Laurent Talbot (2009), les sciences humaines cherchent à rompre avec le sens commun qui rattache l'évaluation au mot « contrôle ». Il affirme qu'évaluer correspond au fait « de s'interroger sur ce qui est en train de se passer, tout en prenant en compte le contexte dans lequel se développe cette conduite » (Talbot, L., 2009, p. 41). Il précise également que l'évaluation est une sorte de jugement, d'appréciation qui qualifie un produit – résultat, par rapport à des critères fixés au préalable, soit les attentes de l'enseignant concernant le travail demandé à ses élèves. Il relève aussi le fait que les référents doivent obligatoirement être des éléments observables, afin de juger l'apprenant de manière plus objective. Nous avons encore relevé un autre dire de cet instituteur et inspecteur de l'Education nationale⁵, selon lequel il défend le fait que la personne qui évalue effectue un acte complexe, où plusieurs éléments externes interviennent, et que malgré ceci, l'évaluateur s'efforce de lutter contre une certaine subjectivité. Le but serait de « juger » un produit fini, mais l'enseignant ne peut faire abstraction de l'évolution et des progrès de l'apprenant, lorsqu'il lui attribue une valeur chiffrée (Talbot, L., 2009, p. 41).

Pour Meyer, évaluer est un des rôles de l'enseignant. Elle relève les actions suivantes à réaliser dans ce cadre : « définir ce que l'enseignant veut observer chez l'enfant, imaginer ou choisir les tests significatifs qui lui permettront d'observer ce qu'il a décidé d'évaluer, confronter ce qu'il observe (la performance réalisée par chaque enfant) à ce qu'il souhaitait (la performance attendue), communiquer au moyen d'un code (note, lettre appréciation, points...) le résultat de cette mesure, l'écart entre la performance attendue et la performance réalisée » (Meyer, G., 2007, p.9). Dans les caractéristiques ressorties par cette auteure, on peut découvrir, en premier point, la notion de critères, puisque l'enseignant doit déterminer des éléments sur lesquels il va focaliser son regard, durant le moment d'évaluation.

Critères d'évaluation

Si l'on s'arrête donc, à présent un peu plus en détails, sur la notion de critère, on peut s'apercevoir que ce concept se retrouve au centre de notre sujet. Pour Meyer (2007), toujours, le critère est un élément qui permet de déterminer si l'objectif fixé pour mesurer telle ou telle compétence est acquis ou pas. Par exemple, si l'on demande à l'élève de réaliser un exercice, dans un examen, dans lequel il doit transférer des connaissances apprises, on dira que l'exercice est réussi, et que le critère est atteint, s'il a répondu juste à au moins 70% des questions. Par contre, pour un travail où il s'agit de remettre des savoirs mémorisés par cœur, on exigera peut-être de l'apprenant qu'il ne fasse aucune erreur pour confirmer que le critère est atteint, cette fois-là. A

⁵ L'éducation nationale en France.

partir de sa conception du critère, l'auteure nous ressort quelques qualités que cet élément doit posséder. Pour elle, le critère doit être univoque, observable, il doit faire référence à des conditions présentes, et on doit définir, pour lui, le seuil de tolérance accepté (Meyer, G., 2007, p. 25-26).

Par rapport à ce concept de critère toujours, on peut découvrir que Hadji (1990) formule, lui aussi, sa propre définition. Pour lui, cet élément peut être défini comme le « caractère ou la propriété d'un objet permettant de porter sur lui un jugement de valeur ». Il ajoute que c'est ce qui permet d'établir si un objectif est atteint ou non (Hadji, C., 1990, p.186). On voit donc que, pour cet auteur, le critère est essentiel dans l'évaluation, puisque c'est lui qui indiquera le seuil d'acquisition de la compétence travaillée, chez l'élève. Cette définition se distingue tout de même de la première élaborée par Meyer, puisque pour cette dernière, les notions d'objectif et de critère semblent être plutôt proches et signifier pratiquement la même chose. Par contre, pour Hadji, le critère est quelque chose que l'on formule par rapport à un objectif posé. Selon lui, l'un est la conséquence de l'autre, puisque si le critère est acquis, l'objectif est atteint.

Par rapport à cette deuxième définition, Galiana est encore plus précis, et approfondit un peu plus le concept. Ce professeur explique que ce sont les critères d'évaluation qui permettent de déterminer le « niveau de maîtrise » d'un objectif. Donc, il y aurait différents niveaux d'atteinte des critères ; par exemple, « excellent », « bon », « satisfaisant » et « insatisfaisant », et cela indiquerait si l'élève a largement atteint l'objectif, ou alors, s'il y arrive, mais qu'il faut veiller, à lui faire retravailler certains savoirs pour répondre, avec un peu plus de marge, à cette exigence (Galiana, D., 2005, p. 49).

Ainsi, après avoir lu de nombreuses définitions dans les ouvrages consultés pour l'élaboration de ce travail, nous avons tenté d'esquisser notre propre explication de ce large concept d'évaluation. Nous pensons que l'évaluation est un processus permettant d'estimer, d'appliquer une valeur à un acte réalisé par autrui, grâce à la formulation et à l'élaboration de critères spécifiques précis. Dans le domaine pédagogique, nous pourrions définir l'évaluation scolaire, de manière globale, comme la procédure empruntée par les enseignants pour qualifier, de manière quantitative, les compétences des apprenants.

2.1.2. Types d'évaluation

Il existe plusieurs types d'évaluation qui ont chacune, une fonction bien précise. On peut citer, notamment, les deux grandes familles d'évaluations suivantes : les évaluations sommatives, qui regroupent l'évaluation certificative et informative, et les évaluations formatives, qui englobent l'évaluation diagnostique, pronostique, formatrice, réflexive, ou encore l'auto-évaluation. Il y a également d'autres sortes d'évaluations, comme l'évaluation critériée et l'évaluation normative. Par exemple, on peut parler d'évaluation sommative critériée ; c'est une sorte d'évaluation sommative en particulier. Et pour les normatives, on parle d'évaluation à référence normative, lorsqu'on compare les notes des élèves d'une même classe, ou encore d'une

classe par rapport à une autre. Mais ce sont surtout les évaluations formatives et sommatives, qui sont le plus souvent utilisées dans le domaine scolaire.

Nous allons, à présent, définir plus en détails ces deux types d'évaluation, afin de bien comprendre leurs différences et leurs enjeux, même si nous nous concentrerons, par la suite, principalement sur le sommatif, puisque c'est dans ce cadre-là, que nous menons notre recherche par rapport aux éléments qui viennent biaiser la notation des enseignants. De plus, c'est cette sorte d'évaluation-là qui pose problème, puisqu'on évalue l'apprenant de manière officielle et institutionnelle. Dès lors, les erreurs de jugement ont un impact plus important sur l'élève, que l'évaluation formative, et nous pensons que c'est ceci qui alimente la polémique autour de l'évaluation scolaire.

*L'évaluation **sommative***

L'évaluation sommative est une évaluation qui intervient à la fin d'un apprentissage. Elle sert à « mesurer » les compétences et les connaissances de l'élève à un moment donné, et plus particulièrement, à la fin d'une séquence d'enseignement dans une branche définie. Elle est caractérisée par l'attribution d'une note ou d'une appréciation qui sert à situer l'apprenant par rapport à son apprentissage, et également, par rapport à ses camarades de classe. Il nous faut encore préciser que ce type d'évaluation est dit « officiel », puisque la note obtenue par l'enfant ou par l'étudiant sera inscrite dans un bulletin scolaire et communiquée à l'extérieur⁶. Selon Talbot (2009), elle est utile pour effectuer un « bilan des performances des élèves et de leur situation au terme d'un cursus d'enseignement, ou de formation, donné ». Elle peut aboutir à un classement et à des informations communiquées à d'autres (Talbot, L., 2009, p. 99).

Selon De Landsheere (1979), l'évaluation sommative est synonyme d'évaluation – bilan ; « elle intervient donc après un ensemble de tâches d'apprentissage constituant un tout, correspondant, par exemple, à un chapitre de cours, à l'ensemble du cours d'un trimestre, etc... » (De Landsheere, G., 1979, p. 128). Il soutient sa définition en citant deux exemples d'évaluation sommative : les examens périodiques et les interrogations d'ensemble. Il achève, alors, son explication de ce concept en précisant que cette sorte d'évaluation est dite « publique », puisque la note attribuée au travail de l'élève sera transmise aux parents de ce dernier (De Landsheere, G., 1979, p. 128).

Quant à Ian Smith (2009), il pense que l'évaluation sommative sert à « trier et classer les élèves à l'intention des établissements d'enseignement post-secondaire et des employeurs, à déterminer les cours que devraient suivre les élèves à l'école, et à faire connaître le niveau de réussite des élèves à ceux-ci et à leurs parents » (Smith, I., 2009, p. 9). Pour cet auteur, cette évaluation peut être définie comme un jugement de l'apprentissage de l'élève et son objectif serait de mesurer le « rendement scolaire ». Il ajoute qu'elle a lieu à la fin d'une séquence d'enseignement et qu'elle donne des renseignements sur le

⁶ Résultat communiqué aux parents et/ou au directeur.

niveau de réussite de l'apprenant dans une branche donnée et l'indication permettant de savoir si ce dernier est prêt à passer à l'étape suivante (Smith, I., 2009, p. 10).

Ian Smith (2009) garde un regard plutôt critique sur le fait d'évaluer « sommativement » les élèves et porte un jugement plus favorable sur l'évaluation dite « formative ». Pour lui, un tel type d'évaluation accentue plutôt les erreurs commises par les apprenants et n'a pas un but très formateur. Cependant, il revient, plus loin dans son ouvrage, sur le fait qu'il est tout de même indispensable de noter officiellement les élèves et étudiants, afin que ceux-ci puissent se situer dans leur apprentissage, et que les parents et les enseignants connaissent le niveau de ces derniers. Il termine alors sa définition en disant qu'il est important de trouver un bon équilibre entre évaluation sommative et évaluation formative.

L'évaluation formative

Passons à présent à l'évaluation formative, afin de mieux comprendre les différences existantes entre les types formatif et sommatif. L'évaluation formative est une évaluation qui peut intervenir à plusieurs moments durant une séquence d'enseignement. Elle sert à informer l'élève sur ses compétences à un moment donné de la séquence d'enseignement, et donne également des renseignements à l'enseignant, afin qu'il puisse réguler et adapter sa façon de transmettre les savoirs et de mettre l'accent sur les éléments qui n'ont pas été acquis par les apprenants. Cette évaluation vise à donner un bilan intermédiaire de l'état des connaissances des élèves, mais elle n'est, en général, pas accompagnée d'une qualification, car elle n'est pas « officielle », soit elle ne figurera pas dans le bulletin scolaire de l'enfant. Elle reste donc dans le cadre de la classe, entre l'enseignant et l'apprenant.

Selon Talbot (2009), les enseignants commencent à comprendre qu'il ne faut pas se servir de l'examen – contrôle, à tout moment. Il affirme que ces derniers « ne se cantonnent plus dans des pratiques d'évaluation sommative et ont bien compris le rôle essentiel de l'évaluation formative qui permet, à tout instant, de savoir où en sont les apprenants » (Talbot, L., 2009, p. 87). Il explique aussi que ce type d'évaluation peut, parfois, être informel, et qu'elle sert à accompagner l'élève dans son apprentissage, contrairement à l'évaluation sommative qui résonne plutôt comme une sorte de « contrôle » ou de « sanction ». Il continue sa définition de l'évaluation formative en disant que c'est « un ensemble de procédures, plus ou moins formalisées par le maître, qui a pour ambition d'adapter son action pédagogique et didactique en fonction des progrès ou au contraire, des difficultés d'apprentissage » (Talbot, L., 2009, p. 87).

Selon Dominique Galiana (2005), l'évaluation formative est une aide à l'apprentissage des apprenants. Elle leur permet de compléter leurs connaissances ou de changer leurs représentations de certains savoirs. Il rappelle aussi le fait que les informations récoltées durant l'évaluation formative donnent des pistes importantes à l'enseignant pour qu'il puisse réguler son enseignement et offrir les meilleures conditions possibles aux apprenants (Galiana, D., 2005, p. 21). Il

explique également que le déroulement des évaluations formatives peut être très différent selon les données que l'on cherche à récolter par rapport à l'état des connaissances des élèves.

Selon Ian Smith (2009), l'évaluation formative est là pour « soutenir les élèves et les aider à apprendre en leur donnant de la rétroaction » (Smith, I., 2009, p. 9). Il pense qu'elle permet aux élèves de s'améliorer, d'évoluer, de prendre conscience de l'état de leurs compétences, afin de tendre aux attentes de l'enseignant et de répondre aux objectifs que celui-ci a fixés par rapport à un savoir donné (Smith, I., 2009, p. 10).

Selon De Landsheere (1979), l'évaluation formative est présente après chaque activité d'apprentissage et sert à informer les élèves de leurs connaissances actuelles sur un savoir travaillé, afin de leur donner l'occasion de se rendre compte de ce qu'il leur reste encore à acquérir pour atteindre les critères demandés par l'enseignant. Il pense que « l'évaluation formative fait partie intégrante du processus éducatif normal », et précise que « les erreurs sont alors considérées comme des moments dans l'apprentissage et non comme des faiblesses » (De Landsheere, G., 1979, p. 125).

Christine Tagliante (2001), quant à elle, associe directement le terme « d'objectif » à ce type d'évaluation. Elle précise que ce processus n'aurait aucun sens, si l'on ne posait pas un objectif à atteindre, au préalable. Puis, elle définit l'évaluation formative comme une démarche « entièrement intégrée à l'apprentissage » qui peut être utilisée en début, en cours ou à la fin d'une séquence d'enseignement, et qui concerne un savoir enseigné. Cette évaluation est, selon elle, plus axée sur l'apprenant et sert de repère pour l'enseignant, par rapport aux acquis des élèves. Elle termine sa définition en ajoutant le fait que l'évaluation formative permet à l'enseignant de « réguler son enseignement », et à l'enfant de « gérer son apprentissage » (Tagliante, C., 2001, p. 11-13, 123).

A présent, pour récapituler et synthétiser les différences existantes entre l'évaluation sommative et l'évaluation formative, nous avons trouvé utile de reprendre une partie du tableau synoptique créé par Talbot (2009). Cela nous donnera également plus de précisions sur ces deux types d'évaluation.

Tableau 2. – **L'évaluation : approche synoptique**

	Evaluation formative	Evaluation sommative
Fonction	Fournir des informations relatives aux compétences et aux connaissances en construction à l'évalué et à l'évaluateur	Fournir des informations relatives aux compétences et aux connaissances au terme d'une unité d'enseignement – apprentissage
Moment de Passation	Au cours de l'unité d'enseignement – apprentissage	Au terme de l'unité d'enseignement – apprentissage

Type de comparaisons effectuées	Intra – élève	Inter – élèves
Finalités	Pédagogiques et didactiques → Améliorer le processus d'apprentissage et d'enseignement → Diagnostiquer les réussites futures	Plutôt sociales → Améliorer le résultat → Orienter → Faire un bilan
Destinataires	→ Le professeur → L'élève	→ Le professeur → Le conseil de classe → L'institution globale → Les parents
Enseignement pratiqué	Individualisé et individualisant (prise en compte des styles et des stratégies d'apprentissage)	Collectif
Questionnement	- Comment s'effectue l'apprentissage des élèves ? - Quelles sont les difficultés rencontrées en cours d'apprentissage ? - Quelles modifications dans l'enseignement sont nécessaires ?	- Quel est le niveau ? - Quels sont les résultats d'apprentissage construits ?
Objets	L'élève : - ses connaissances - ses démarches - sa motivation - ses difficultés - Le processus d'apprentissage (global)	- La performance indicatrice d'un niveau atteint - Le produit final

Source : Talbot (2009) « L'évaluation formative » (p. 101-102)

Pour conclure ce concept, nous aimerions préciser que le grand point, qui différencie les types formatif et sommatif de l'évaluation, est le fait que le premier reste dans le cadre de la classe, entre l'apprenant et l'enseignant, tandis que la deuxième sorte d'évaluation se traduit par le fait de communiquer les informations concernant les résultats scolaires des élèves, à l'extérieur de la classe (Talbot, L., 2009, p. 99).

2.2. Biais

2.2.1. Définitions

Les biais de l'évaluation sont des éléments qui interviennent dans la démarche de l'évaluation et qui exercent une certaine influence sur l'évaluateur et sur sa façon de juger. Ils peuvent être de nature diverse et proviennent de plusieurs facteurs. Nous définirons seulement quelques-uns de ces phénomènes, dans la partie consacrée aux

divers types de biais, car ce sont les éléments que nous avons retrouvés le plus souvent dans les lectures effectuées.

Nous allons, à présent, exposer les différentes définitions du concept de biais, proposées par plusieurs auteurs. Nous pourrions constater que certains facteurs se retrouvent chez plusieurs d'entre eux. Par contre, les explications et/ou les dénominations varient parfois quelque peu, comme nous le découvrirons dans la partie, qui traite des facteurs, située plus bas.

Selon Talbot (2009), les biais sont des éléments qui viennent perturber l'acte évaluatif et que la docimologie s'est attelée à définir et à théoriser. Il explique que la pratique de l'évaluation amène forcément des questionnements sur la façon de juger des enseignants, c'est pourquoi il est important que ces derniers restent transparents et qu'ils prennent le temps de bien définir leur façon de faire, afin de « rendre leur pratique moins contestable » (Talbot, L., 2009, p. 103). Ceci les aidera à mieux pratiquer l'acte évaluatif, même si cela ne permettra pas d'éviter tous les biais liés au fait de juger autrui.

Selon Pierre Merle (1996), il existe des effets qui interviennent dans le processus de l'évaluation, et ces derniers peuvent expliquer l'incertitude et les variations dans la notation établie par l'enseignant. Ce sociologue a classé ces phénomènes en trois catégories : les effets liés à l'ordre de correction, puis ceux qui sont relatifs aux statuts scolaire et social de l'élève, et enfin ceux qui sont en lien avec le contexte de l'école et les interactions qui ont lieu en classe (Merle, P., 1998, ch. 2). Ainsi, il précise que l'incertitude de la notation est due à divers facteurs qui font varier la façon de juger de l'enseignant.

Selon Christine Tagliante (2001), les biais dans l'évaluation sont des « critères parasites de la fidélité » (Tagliante, C., 2001, p. 27) de la notation. Elle pense que les enseignants ont conscience de ces éléments perturbateurs, mais qu'ils ne savent pas forcément les citer. Elle dit qu'il est indispensable de les connaître pour prendre conscience de leurs conséquences sur les évaluations.

Gilbert de Landsheere (1979), quant à lui, commence sa définition des effets en disant qu'ils sont la conséquence à une cause. Si l'on adapte son explication au domaine de l'évaluation, cet auteur dirait que c'est « un phénomène » qui est produit par une cause, un facteur venant perturber le processus de l'évaluation (De Landsheere, G., 1979, p. 99).

Enfin, André de Peretti, Boniface et Legrand (2005) parlent des biais de l'acte évaluatif comme des « effets intempestifs ou pervers » qui s'immiscent dans les évaluations (De Peretti, Boniface et Legrand, 2005, p. 51). Ces trois auteurs insistent sur le fait qu'il est nécessaire de repérer et de prendre conscience de ce phénomène, afin d'effectuer son rôle d'enseignant – évaluateur de manière plus juste et plus professionnelle.

2.2.2. Types de biais

Nous allons, maintenant, expliquer les biais principaux soulignés par plusieurs auteurs, dont nous avons consulté les ouvrages. Ceux que nous avons retenus sont les suivants : l'effet de contamination, l'effet de stéréotypie, l'effet Pygmalion, l'effet de halo, l'effet d'ordre, l'effet de contraste, la fatigue et l'effet Posthumus. Comme nous pourrions le constater, les définitions de certains de ces biais diffèrent légèrement selon les auteurs.

Effet de contamination

Talbot définit l'effet de contamination comme un phénomène qui vient fausser la notation de l'enseignant. Si un enseignant a connaissance des notes obtenues par un enfant, auparavant, lors d'années précédentes, par exemple, cela peut avoir une influence sur sa façon d'évaluer ce même écolier. Si celui-ci a un bulletin scolaire très mauvais, l'enseignant risque d'en faire une généralité et de le noter plus sévèrement que s'il n'avait pas eu accès à ces informations (Talbot, L., 2009, p. 106). Il ajoute que l'effet d'assimilation peut être apparenté à un effet de contamination, puisque ce phénomène se produit lorsque l'évaluateur assimile la dernière note obtenue par l'enfant, au nouvel examen qu'il est en train de corriger et de noter (Talbot, L., 2009, p. 106).

Lorsque Tagliante explique l'effet de contamination, elle dit que, si un enfant a très souvent des bonnes notes, on peut être influencé dans sa correction, et que, même si ce même élève fait un examen très mauvais, on le notera quand même un peu moins bas, qu'un autre apprenant qui aurait toujours des résultats plutôt médiocres (Tagliante, C., 2001, p. 28).

De Landsheere, quant à lui, ne décrit pas réellement l'effet de contamination, mais il parle de l'effet Larsen, qui pourrait s'apparenter à cet effet-là. Ce qui se cache derrière l'intitulé « effet Larsen », c'est le fait de « contaminer » sa façon de juger l'élève, à cause d'un accès à des informations concernant le passé scolaire de ce dernier (De Landsheere, G., 1979, p. 102).

En résumé, L'effet de contamination est la conséquence liée à diverses causes qui viendraient « contaminer » le processus de l'évaluation. Ainsi le statut scolaire de l'élève et l'établissement, qu'il a fréquenté, ont une influence sur la notation de l'enseignant. Mais, il peut également s'agir d'un très bon élément présent dans une copie qui est globalement plutôt médiocre, qui va fausser la perception et le jugement de l'évaluateur. Il existe plusieurs effets que l'on peut classer dans les effets de contamination, comme l'effet de **Larsen** et l'effet **d'assimilation**. L'effet Larsen est le fait que l'évaluateur biaise sa notation à cause d'informations et de précisions qu'il a obtenues, concernant ses élèves. Ainsi, en allant jeter un œil dans le passé scolaire d'un apprenant, on pourrait facilement changer la perception que l'on a de cet enfant et donc, modifier notre façon de le juger. L'effet de Larsen fait partie des effets de contamination. Et pour l'effet d'assimilation, c'est le fait de prendre en compte une note faite

antérieurement et d'assimiler celle-ci à l'examen fait récemment (Tagliante, C., 2001, p. 28 / Talbot, L., 2009, p. 106 / Merle, P., 1998, p. 27 / De Landsheere, G., 1979, p. 100 et 102).

Effet de stéréotypie

Pour Talbot, l'effet de stéréotypie est lié aux statuts scolaire et social de l'enfant. Il dit que les enseignants ont tendance à être plus souples dans leur notation, lorsqu'il s'agit de classes ayant la réputation d'être plutôt bonnes. Ceci concerne donc le statut scolaire des apprenants. Et pour le statut social de ces derniers, l'auteur affirme aussi, que les classes composées d'enfants, venant d'un milieu socio-économique aisé, reçoivent aussi, en général, de meilleures notes à leurs examens (Talbot, L., 2009, p. 105-106).

Merle parle également des biais qui sont liés au statut scolaire et social de l'élève pour désigner l'effet de stéréotypie. Plusieurs chercheurs ont réalisé des études autour de ce thème et se sont aperçus que les enseignants possédaient certaines informations sur le parcours scolaire et sur l'origine sociale des apprenants de leur classe. Ainsi, ces derniers ont pu démontrer que ces éléments connus par l'enseignant pouvaient avoir une certaine influence sur la notation établie. Le niveau et le statut scolaire, l'origine sociale, l'apparence physique et le sexe ont un effet sur la variabilité de la notation. Puis, Merle revient sur le contexte scolaire en soulignant le fait que l'établissement scolaire, la classe et les caractéristiques propres à chaque élève ont effectivement un impact sur la façon d'évaluer de l'enseignant (Merle, P., 1998, p. 23-27).

Pour l'effet de stéréotypie, Tagliante dit que l'enseignant note différemment les apprenants selon la manière dont il juge chacun des enfants (Tagliante, C., 2001, p. 28).

En résumé, l'effet de stéréotypie, c'est l'effet qu'ont nos stéréotypes par rapport à un élève, sur notre façon de l'évaluer. En fait, un enseignant risque de juger un apprenant toujours, plus ou moins, de la même manière, et ceci, même si les performances de ce dernier sont différentes (Beauté, J., 1995 / Tagliante, C., 2001, p. 28).

Effet Pygmalion

Talbot définit l'effet Pygmalion⁷ par le fait que les attentes d'un enseignant par rapport à ses élèves a une influence sur sa notation vis-à-vis des examens, auxquels il doit attribuer une valeur. Si ce dernier pense que l'enfant aura un avenir scolaire prometteur, il le notera mieux que s'il a des attentes plutôt basses par rapport à ce même élève (Talbot, L., 2009, p. 109).

De Landsheere nomme « effet Pygmalion », l'effet produit, sur un sujet, par le fait de prédire quelque chose le concernant. Ainsi, si l'on transpose ce phénomène au domaine pédagogique, on dirait que, si l'enseignant fait une prédiction positive concernant l'avenir de l'apprenant et croit au fait qu'il est bon, cette croyance aura tendance à

⁷ Résultats démontrés par l'étude de deux chercheurs, Rosenthal et Jacobson, effectuée en 1968.

se réaliser et les notes de l'enfant seront plus élevées. (De Landsheere, G., 1979, p. 99-104).

En résumé, l'effet Pygmalion, aussi appelé effet « Rosenthal », est une sorte de prédiction qui a un impact sur la façon de percevoir la réalité. Si l'on pense qu'un apprenant a un grand potentiel, on aura tendance à lui attribuer une note plus favorable, afin que la prédiction que l'on a faite, soit la réussite de cet élève, se réalise (Talbot, L., 2009, p. 109 / De Landsheere, G., 1979, p. 104).

Effet de halo

Sous l'appellation « effet de halo », Talbot entend le fait que certains éléments relatifs à la forme d'une copie, comme la présentation, l'orthographe ou encore le soin, peuvent influencer de manière plus ou moins importante la notation de l'évaluateur (Talbot, L., 2009, p. 106).

Concernant l'effet de halo, Tagliante se demande, si le fait de devoir évaluer une forme qui présente bien, sans pour autant que le fond soit extraordinaire, n'aurait pas une certaine influence sur la notation de l'enseignant. Ainsi, si un élève très extraverti, qui s'exprime facilement, se montre à l'aise lors d'un examen oral, mais qu'au fond, ce qu'il dit n'est pas très recherché, on aura tendance à le juger plus favorablement qu'un autre qui serait très timide et stressé devant la classe, et ceci, bien que ses dires soient pertinents (Tagliante, C., 2001, p. 28). D'ailleurs, l'auteure dit qu'il existe une « variable de débordement ». Selon elle, dans un examen, il peut y avoir une accumulation de beaucoup de choses négatives, comme le soin, l'écriture, l'orthographe, une mauvaise ponctuation, des oublis, etc., et malgré le fait que l'enseignant ait passé outre tout ceci, un énième élément peut, tout à coup, être « la goutte d'eau qui fait déborder le vase » et avoir une influence sur la notation (Tagliante, C., 2001, p. 28).

De Landsheere définit l'effet de halo en disant que c'est un phénomène caractérisé par une « contamination affective » ; par exemple, si l'on trouve un enfant gentil ou poli, on risquerait de lui mettre une meilleure note. Mais il poursuit son explication en relevant le fait que la littérature expérimentale parle d'effet de halo pour qualifier « l'influence de la qualité de l'orthographe ou de l'écriture ».

En résumé, c'est le fait de concentrer son attention sur des éléments non essentiels d'un examen. Par exemple, si l'enseignant se met à donner plus d'importance au soin, à la présentation ou encore à l'orthographe d'une copie, au lieu de juger réellement les connaissances de l'élève, on parle d'effet de halo (Talbot, L., 2009, p. 106 / Tagliante, C., 2001, p. 28).

Effet d'ordre

Talbot explique qu'il aurait été démontré, par des chercheurs ayant travaillé sur le thème des facteurs biaisant les évaluations, que « les copies placées dans le premier tiers d'un ensemble à corriger sont mieux notées que lorsqu'elles sont placées dans le dernier tiers »

(Talbot, L., 2009, p. 105). Pour lui, il y a donc bien un effet d'ordre, qui intervient et qui influence la notation des instituteurs.

Merle (1996), quant à lui, cite, dans son ouvrage, une étude qui a été effectuée par J.-J. Bonniol (Merle, P., 1998, p. 16-17) et qui a démontré que l'ordre de correction des examens a une influence sur la notation faite par l'enseignant. Ce chercheur a pu prouver qu'une copie était mieux notée si elle était corrigée après un examen plutôt faible, et que cette même copie était sanctionnée d'une note plus basse si elle se trouvait juste après un très bon examen.

Au niveau de l'effet d'ordre, Tagliante affirme que les enseignants sont généralement plus souples en début de correction d'une série d'examens, qu'à la fin (Tagliante, C., 2001, p. 27).

En résumé, c'est l'effet qui est dû à l'ordre de correction des copies. Il semblerait que l'on soit plus indulgent lorsqu'on commence à corriger des examens, que lorsqu'on termine la série de copies. (Tagliante, C., 2001, p. 27 / Talbot, L., 2009, p. 105). Par contre, d'autres auteurs expliquent cet effet de manière inverse et disent que l'on est plus sévère en début de correction, et que c'est la lassitude, lorsqu'on corrige un tas d'examens, qui pousse les enseignants à être plus souples en fin de correction. Mais appréhendé dans ce sens-là, ce biais ressemble plutôt à la fatigue ; biais que nous décrivons ci-dessous. Enfin, dans la définition de Merle, on peut voir que cet auteur parle d'effet d'ordre, mais que la façon dont il le définit relève plutôt d'un autre biais que l'on va découvrir ci-dessous, soit l'effet de contraste.

Effet de contraste

Par rapport à l'effet de contraste, Talbot entend le fait qu'un examen peut être mieux ou moins bien noté selon l'examen corrigé juste avant. Selon lui, une copie quelconque sera plus mal jugée dans le cas où elle vient juste après un excellent examen (Talbot, L., 2009, p. 105).

Par « effet de contraste », Tagliante entend le fait « qu'une copie moyenne souffre d'être corrigée juste après une excellente production » (Tagliante, C., 2001, p. 27). Ainsi, si on corrige un très bon examen, puis, un examen moyen, ce dernier obtiendra certainement une note plus mauvaise que s'il avait été corrigé seul, ou après un examen très médiocre.

De Landsheere explique, l'effet de contraste de la manière suivante : « Déformation de la perception due à une perception immédiatement précédente » (De Landsheere, G., 1979, p. 100).

En résumé, l'effet de contraste, également nommé effet consécutif, est la conséquence d'une perception faussée par deux éléments très différents. Donc, si on corrige un examen très bon, et que l'on évalue, juste après, une copie très mauvaise, on aura tendance à noter cette deuxième de manière plus sévère (Talbot, L., 2009, p. 105 / Tagliante, C., 2001, p. 27).

Fatigue

Tagliante pense que les enseignants peuvent être influencés, dans leur notation, par l'état de fatigue dans lequel ils se trouvent (Tagliante, C., 2001, p. 27-28).

De Peretti, Boniface et Legrand (2005) relèvent les effets de fatigue et d'ennui. Les trois auteurs ne définissent pas ce facteur avec précision, mais dans leur ouvrage, ils affirment que cet élément a une réelle influence sur la notation des enseignants (De Peretti, Boniface et Legrand, 2005, p. 51-52).

En résumé, La fatigue peut être un élément qui biaise la notation des enseignants. On évalue certainement plus « équitablement », si l'on est bien reposé que si l'on a un certain état de fatigue (Tagliante, C., 2001, p. 27 / (De Peretti, Boniface et Legrand, 2005, p. 51-52).

Effet Posthumus

Lorsqu'il parle d'effet Posthumus, Talbot fait référence aux recherches qui ont été effectuées par K. Posthumus, en 1947, et qui ont démontré que, quel que soit le niveau des apprenants présents dans une même classe, on retrouve toujours, en fin d'année scolaire, un maximum d'élèves qui se situent dans la moyenne, et peu d'écopiers très doués ou très faibles. Ainsi, tout enseignant aurait tendance à réadapter sa manière d'évaluer, selon cette loi découverte par ce chercheur. La notation de l'instituteur s'en trouve donc sensiblement modifiée par rapport à la valeur réelle des travaux réalisés par les apprenants (Talbot, L., 2009, p. 108-109).

Ainsi, on peut dire que cet effet est caractérisé par le fait que l'enseignant a généralement tendance à adapter sa notation aux résultats obtenus durant l'année, par ses élèves, afin que la plupart des apprenants se situent dans une moyenne, et qu'il y ait très peu de bons et très peu de faibles. La notation des instituteurs peut, ainsi varier, chaque année, en fonction de la classe qu'ils ont, alors que cela ne devrait pas être le cas. (Galiana, D., 2005, p. 69-70 / Talbot, L., 2009, p. 108-109)

Pour conclure ces divers types de biais, nous aimerions préciser que nous avons choisi de traiter de ces facteurs-là en particulier, car nous avons pu remarquer que ce sont ceux qui ressortaient le plus souvent des lectures effectuées sur ce thème. De plus, ces derniers s'inscrivent vraiment dans le champ de la pédagogie, champ que nous avons choisi de privilégier dans cette étude. Et enfin, nous pensons, à ce stade de notre travail de recherche, que ce sont ces biais-là dont les enseignants ont le plus conscience.

2.2.3.Liens avec l'évaluation

Les biais de l'évaluation sont en lien direct avec le domaine de l'évaluation scolaire. Ils s'immiscent dans la procédure d'estimation des performances des apprenants et ont des conséquences plus ou moins importantes sur la note obtenue par ces derniers.

Comme nous nous concentrons sur les biais liés à la pédagogie uniquement, ce sont surtout les effets qui ont trait au domaine scolaire, exclusivement, et plus précisément à la relation enseignant – élève, ainsi qu'à l'acte même de correction⁸, qui ont été relevés ci-dessus.

Bien entendu, pour revenir au contexte de notre travail, il nous faut préciser que ces divers biais peuvent se retrouver, autant dans les évaluations formatives, que dans le type sommatif, mais comme nous l'avons déjà expliqué plus haut, nous restons dans le domaine du sommatif, puisque ce sont les notes attribuées dans ces évaluations-là qui posent problème, en général, de par leur statut officiel.

2.3. Objectivité / Validité / Fiabilité

2.3.1. Définitions

Objectivité

Dominique Galiana (2005) s'est posé plusieurs questions par rapport à ce terme d'objectivité dans le domaine de l'évaluation. Il s'est demandé si « noter, c'est être juste » ?, si cela « permet d'être objectif ? », si l'on « note toujours les élèves de la même manière » ?, si « certains facteurs auraient une influence sur la façon dont on note une copie » ?, ou encore « comment fait-on pour éviter l'influence de ces facteurs » ? Cet auteur poursuit son questionnement en rappelant que la docimologie étudie les éléments qui s'immiscent dans l'acte évaluatif, et il souligne l'importance de la **fidélité**, de la **validité** et de la **sensibilité** de la notation. Selon Galiana, si on est fidèle, dans sa manière d'évaluer, on attribue la même note au même élève, que l'on corrige sa copie aujourd'hui ou dans un mois. Si l'on veut que son jugement soit valide, on doit estimer le travail demandé, selon des critères précis, et ne pas juger autre chose. Et enfin, l'évaluateur devrait pouvoir être rigoureux lorsqu'il donne une appréciation, afin que l'on puisse percevoir la sensibilité dans la notation. Pour ce dernier point, l'enseignant est donc sensé remarquer et relever des différences entre deux examens très proches. Ainsi, Galiana (2005) pense que l'on pourra tendre à plus d'objectivité si l'on respecte ces trois éléments (Galiana, D., 2005, p. 65-69).

Selon Gilbert de Landsheere (1979), l'objectivité est le « caractère de ce qui donne une image non déformée des organismes et des choses, ou de ce qui les décrit et les juge, sans parti pris » (De Landsheere, 1979, p. 212). Il ajoute que cette notion est « définie comme le degré de concordance entre plusieurs observations » (De Landsheere, 1979, p. 212). Puis, il précise que, presque à chaque fois que l'on doit évaluer autrui, il y a une part de subjectivité qui se glisse dans notre manière de juger. Pour tendre à une certaine objectivité, il faut fixer des critères de correction précis et univoques. Quant à la subjectivité, il se sert d'une citation, pour expliquer que c'est le caractère de ce qui ne peut être justifié et contrôlé directement par « observation et mesure » (De Landsheere, 1979, p. 212).

⁸ Conditions de correction.

Validité

Selon Tagliante (2001), un test est valide s'il « mesure exactement et exclusivement ce qu'il est censé mesurer. Une correspondance entre le contenu du test et l'objectif visé, et entre l'objectif du test et l'objectif de l'apprentissage » (Tagliante, C., 2001, p. 26), doit donc être identifiable. Elle précise aussi que cette qualité est très difficile à obtenir, mais qu'il est possible d'y arriver, si on pose des objectifs observables et univoques dans ses évaluations.

Landsheere (1979) définit la validité comme « la qualité d'un test ou de toute autre mesure qui semble correcte et adéquate à l'objet mesuré » (De Landsheere, G., 1979, p. 322). D'ailleurs, Guillon (2004) parle également de « qualité » que chaque évaluation se doit de posséder.

Fiabilité

La fiabilité est la caractéristique d'un élément auquel on peut se référer, se fier. Un résultat est fiable, s'il correspond à des référents clairs, univoques et approuvés. Une notation est fiable, si elle a été établie d'après des critères précis, justes et en lien avec les objectifs de l'examen et les objectifs d'apprentissage. On peut donc remarquer que le terme de « fiabilité » est en lien avec la notion de « validité », vue précédemment, puisque tout semble tourner autour de critères précis à définir.

D'après Galiana (2005), pour que l'évaluation de l'enseignant soit la plus objective et fiable possible, elle doit être stable, fidèle et valide. Ainsi, pour qu'une note soit dite « stable », il faudrait obtenir la même appréciation de la part de tous les enseignants qui pourraient corriger une copie donnée. Pour que la notation soit fidèle, il faut que celle-ci reste la même qu'elle soit établie dans la journée ou trois mois plus tard, soit qu'elle soit constante. Et pour qu'elle soit valide, elle ne doit pas se laisser perturber par les effets liés à l'évaluation. Encore une fois, on s'aperçoit, ici, que les concepts « d'objectivité », de « validité » et de « fiabilité » sont très proches.

Selon De Landsheere (1979), le terme fiabilité est parfois assimilé au concept de fidélité. Il différencie, ensuite, ces deux notions en précisant que « la fidélité désigne principalement la qualité constante d'une mesure » (De Landsheere, G., 1979, p. 137), alors que la fiabilité correspondrait plutôt à un dispositif qui ne présenterait aucun défaut. On utilisera donc l'adjectif « fidèle » pour définir la notation, et le qualificatif « fiable », lorsqu'on traitera de « la qualité de l'instrument qui en assure son fonctionnement constant » (De Landsheere, G., 1979, p. 137).

2.3.2. Liens avec l'évaluation

Nous allons, à présent, développer les liens que nous avons définis entre les concepts d'objectivité, de validité et de fiabilité, et notre domaine de prédilection : l'évaluation.

Grâce aux quelques définitions citées plus haut, on s'aperçoit que la question de l'objectivité est plutôt bien présente dans l'acte évaluatif. Elle est un élément central de ce domaine scolaire et est bien souvent remise en question par bon nombre de personnes. Comme déjà relevé dans l'introduction de ce travail, l'enseignant est bien souvent critiqué par la société, dans son rôle d'évaluateur. Si l'on cherche plus en détails la cause de ces nombreuses insatisfactions de la part des parents ou des autres personnes concernées, on peut remarquer que c'est bien l'objectivité qui fait parfois défaut dans la notation de certains enseignants. C'est probablement à cause de certains abus dans ce domaine, que les gens remettent autant en question cette fonction de l'instituteur et aussi, ses compétences.

On peut alors s'interroger sur les raisons qui poussent les individus à rabaisser le corps enseignant et sur la signification réelle du mot « objectivité ». Pour certains, « être objectif » semble signifier « attribuer la note juste ». Mais on peut aussi se demander s'il existe réellement une « note juste ». C'est la question que s'est posée Dominique Galiana (2005), lorsqu'il aborde les concepts de notation, d'objectivité et de justice. Il pense qu'il n'existe pas de « note réelle » ; pour lui, « la note est relative aux exigences de l'évaluateur, et tous les évaluateurs n'ont pas les mêmes exigences par rapport aux mêmes objectifs et à la même copie » (Galiana, D., 2005, p. 68). La note qu'il appelle « réelle » représente la valeur que mériterait une copie, si l'on pouvait être complètement objectif, et si aucun biais ne venait s'immiscer dans l'acte évaluatif. En fait, pour Galiana, attribuer la note réelle correspondrait à mesurer de la manière la plus exacte une copie. D'ailleurs, le rôle de l'enseignant est de s'approcher « au plus près » de cette valeur impossible à trouver (Galiana, D., 2005, p. 66). Il continue sa réflexion en précisant que la notation dépend également de chaque situation d'évaluation. Cette dernière peut différer selon les années, par exemple.

Passons, à présent, aux liens entre les notions de validité et de fiabilité, par rapport au concept d'évaluation. Comme souligné à plusieurs reprises, la validité et la fiabilité sont des qualités qui doivent caractériser la notation des enseignants pour que celle-ci soit la plus juste possible (Galiana, D., 2005, p. 67 / Talbot, L., 2009, p. 44). Ces deux termes sont donc en lien direct avec le domaine de l'évaluation scolaire. Si l'on parvient à évaluer de manière valide et fiable, en posant des critères cohérents et en mesurant des éléments précis, on pourra être plus professionnel dans le métier d'enseignant – évaluateur. Ainsi, on tendra à satisfaire un peu plus la société et à réduire les reproches qu'elle fait très souvent.

Par rapport à l'évaluation, Guillon (2004) dit que toute évaluation devrait répondre aux trois qualités suivantes : la justesse, la fidélité et la validité. Les conditions d'évaluation doivent donc être justes et surtout égales pour tous les élèves, et la notation doit être fiable et valide, et permettre de mesurer des éléments de référence précis (Guillon, C., 2004, p. 36). Ainsi, la façon de juger de l'enseignant correspondra plus à la réalité et pourra répondre à des critères de fiabilité.

2.4. Représentation

2.4.1. Définition

Selon Durkheim (1898), la société, comme chaque individu, possède ses propres représentations ; on constate donc qu'il y a des représentations dites « collectives » et d'autres, que l'on nomme « individuelles ». Et bien qu'elles soient distinctes, il est possible de les comparer. Pour cet auteur, ces phénomènes relèvent de la conscience de l'être humain, et cette conscience traduit la réalité sous forme de représentations, qui reflètent les croyances et conceptions des gens. En fait, les représentations individuelles constituent la réalité que se construit chaque personne, et peut être une réduction partielle de ce qui est réel. Emile Durkheim ajoute que ces éléments sont présents tout au long de notre vie, et que nous sommes influencés par nos représentations passées et présentes. Ainsi, on peut constater qu'elles peuvent évoluer et changer, mais parfois, il est difficile de les modifier, d'autant plus, quand elles sont bien ancrées en nous. Et par rapport aux représentations collectives, l'auteur précise que ces phénomènes sont « extérieurs aux consciences individuelles », et que, par conséquent, « elles ne dérivent pas des individus pris isolément », mais de l'association de ces personnes (Durkheim, E., 2006, p.13).

2.4.2. Liens avec l'évaluation

Par rapport au lien que l'on peut faire avec notre domaine de prédilection, on constate, dans la suite de ce travail, que nous avons tenté de recueillir les représentations, qu'avaient certains enseignants, concernant l'évaluation, et plus précisément, leurs conceptions du phénomène de biais dans l'évaluation. C'est donc le fait de prendre conscience que nous rechercherions les croyances et opinions des personnes choisies pour notre recherche, qui nous a poussée à être plus au clair avec ce concept de représentation, et à le définir grâce à des lectures effectuées autour de cette notion.

3. Question de recherche

Notre question de recherche a évolué à plusieurs reprises, depuis le début de notre travail. Nous avons commencé par élaborer un questionnement autour du thème des biais dans l'évaluation sommative et, au fil des lectures effectuées et des entretiens de réflexion vécus avec notre directrice de mémoire, nous avons retenu la question de recherche suivante :

«Quels sont les biais dont les enseignants ont le plus conscience et de quelle façon les gèrent-ils ? »

4. Hypothèses

Par rapport à la question de recherche susmentionnée, il n'existe pas beaucoup d'hypothèses pouvant y répondre. Nous avons tout de même réussi à en ressortir quelques-unes, que nous avons listées ci-dessous :

Hypothèse n°1 : Les biais dont les enseignants ont le plus conscience sont les suivants : l'effet de stéréotypie, l'effet de halo, la fatigue et l'effet Posthumus.

Hypothèse n°2 : Les enseignants expérimentés sont plus conscients des biais qui s'immiscent dans le processus d'évaluation que les jeunes enseignants.

Hypothèse n°3 : Les jeunes enseignants font attention aux éléments qui peuvent biaiser leurs évaluations, et ils essaient de les éviter par tous les moyens.

5. Dispositif méthodologique

5.1. Méthodologie

Le dispositif choisi pour mener ce travail de recherche est la méthode de l'entretien. Nous avons sélectionné cette manière de procéder pour deux raisons principales. Premièrement, les données seront récoltées ainsi, car c'est la méthode que nous avons jugée la plus pertinente par rapport aux réponses que nous attendons des entretiens qui seront menés. Et deuxièmement, cette procédure est plus riche et permettra de mieux approcher les représentations des enseignants par rapport au domaine de l'évaluation, et plus particulièrement par rapport aux biais dans l'évaluation sommative. La souplesse de ce dispositif est très intéressante ; on peut obtenir des réponses variées et parfois surprenantes. Grâce aux discussions qui seront effectuées, nous découvrirons probablement des aspects auxquels nous n'avions pas forcément pensés. Ces aspects changeront peut-être nos propres représentations.

Si l'on se concentre, à présent sur cette méthode spécifique de recherche, il nous faut préciser qu'il existe divers types d'entretiens : l'entretien libre, l'entretien semi-directif et l'entretien directif. Le premier n'a pas de structure, il sert à récolter un point de vue sur un sujet, mais l'interviewé est totalement libre ; il n'a aucune directive quant à ce qu'il doit dire durant l'entretien. Le deuxième type porte bien son nom, puisqu'il n'est pas complètement dirigé et qu'il permet à l'interlocuteur de donner son opinion, mais cette fois, sur des éléments un peu plus précis. Il s'agira donc de donner des points ou des questions à la personne interrogée, afin que celle-ci puisse s'exprimer et développer sa conception concernant les éléments discutés. Enfin, par rapport au dernier style d'entretien, l'interviewé est dirigé, ne peut faire aucun ajout personnel et n'a pas la possibilité de s'éloigner un peu de la question posée pour apporter quelque chose en plus. L'entretien est donc entièrement conduit par le chercheur.

La méthode de l'entretien a fait l'objet de diverses définitions élaborées par plusieurs auteurs, et il nous a paru utile d'en relever quelques-unes. Pour De Landsheere (1979), l'entretien libre, ou non directif, est un « échange de vues libre entre un intervieweur et un interviewé » (De Landsheere, G., 1979, p. 110). Pour Blanchet et Gotman (2007), l'entretien est une « technique d'enquête » qui permet de mettre l'intervieweur et l'interviewé sur un pied d'égalité. Ces deux chercheurs précisent que ce dispositif est un « type de rapport social et verbal appartenant au langage diplomatique, à la fois antérieur et externe à la constitution des sciences sociales et de leurs outils méthodologiques, qui désigne une conversation d'égal à égal » (Blanchet, et Gotman, 2007, p. 7). Blanchet et ses collègues (2005) donnent également une autre définition de l'entretien. Pour eux, c'est « un événement de parole dans lequel une personne A extrait une information d'une personne B, information qui était contenue dans la biographie de B » (Blanchet, et al., 2005, p. 81-82).

Finalement, De Ketele et Roegiers (2009) insistent sur les caractéristiques propres à la méthode de l'entretien. Ils soulignent le fait que c'est un moyen de récolter des données, qui a un sens seulement si plusieurs interviews sont réalisées, et dans le cas où on n'utilise que ce dispositif méthodologique pour mener sa recherche. Les deux auteurs ajoutent qu'il faut aussi faire une sélection, quant à l'échantillon de personnes qui seront interrogées pour l'étude effectuée. Puis, ils relèvent aussi le fait qu'il existe plusieurs types d'entretiens : libre, semi-dirigé et dirigé. Et enfin, ils définissent ce terme d'interview ou d'entretien de la manière suivante : « Méthode de recueil d'informations en des entretiens oraux, individuels ou de groupes, avec plusieurs personnes sélectionnées soigneusement, afin d'obtenir des informations sur des faits et des représentations, dont on analyse le degré de pertinence, de validité et de fiabilité en regard des objectifs du recueil d'informations ».

Pour notre recherche, il s'agira donc d'entretiens semi-dirigés, puisque nous poserons des questions aux enseignants choisis, afin qu'ils donnent leur avis sur les sujets abordés. Ces éléments sont présentés dans nos guides d'entretiens, placés en annexes.

5.2. Echantillon choisi

Pour mener notre travail de recherche, nous avons décidé de fixer certains critères pour sélectionner les personnes qui participeraient à nos entretiens. Tout d'abord, nous avons réfléchi à l'échantillon d'individus que nous désirions choisir pour élaborer notre projet. C'est donc après quelques moments de réflexion que nous avons pensé à interroger des enseignants avec une expérience assez grande dans le domaine de l'enseignement. Nous voulions procéder ainsi, car, comme nous souhaitions travailler sur le sujet de la conscientisation des erreurs qui existent dans l'acte évaluatif, nous savions qu'il nous fallait trouver des personnes possédant un certain recul et une certaine lucidité sur leur pratique professionnelle.

Puis, après avoir effectué plusieurs lectures et découvert que passablement d'études avaient été réalisées autour de l'objet des biais dans l'évaluation, nous avons cherché un peu plus l'innovation. C'est donc en prenant un deuxième temps de réflexion que nous avons eu une nouvelle idée. Nous nous sommes dit qu'il serait intéressant d'interroger, dans un premier temps, certains enseignants très expérimentés, comme prévu au départ, et, dans un deuxième temps, d'aller recueillir les représentations d'enseignants avec moins d'expérience dans l'enseignement. Ainsi, en dernière partie de notre travail de mémoire, nous désirions comparer les opinions et la vision des jeunes enseignants avec celles des enseignants un peu plus expérimentés, afin de découvrir s'il existe des différences entre ces deux types d'individus.

Nous avons donc choisi quatre enseignants avec lesquels nous voulions mener des entretiens individuels. Les deux premières personnes sélectionnées ont plus de quinze ans d'expérience dans le domaine de l'enseignement. Nous les avons rencontrées durant les stages effectués dans le cadre de notre formation à la HEPvs. Nous avons eu le temps de discuter longuement avec ces deux enseignantes, et nous avons eu un bon contact avec elles. De plus, ces deux personnes avaient une vision lucide et très intéressante de la profession d'enseignante. Enfin, elles travaillent toujours, actuellement en Valais, et pour nous, il était plus logique

d'interviewer des personnes qui enseignent dans le canton où nous nous formons nous-même, en ce moment.

Pour les deux autres individus sélectionnés, nous avons cherché à retrouver des anciens étudiants de la Haute Ecole Pédagogique du Valais, qui avaient trouvé une place d'enseignement depuis moins de trois ou quatre ans, afin de les classer dans la catégorie des jeunes enseignants à interviewer pour notre recherche. Nous avons donc pris contact avec deux jeunes gens qui travaillent dans la région. Le premier enseigne dans le canton de Vaud depuis deux ans, en 5-6P, mais a tout de même suivi sa formation à la HEPvs. Et la deuxième jeune femme choisie travaille comme enseignante de 4P, 5P et 6P, cette année. Elle a sept classes différentes, dans lesquelles elle mène, notamment, les cours de mathématiques, français, et géographie, et réalise donc des évaluations sommatives dans ces disciplines-là. Par contre, elle était titulaire d'une classe l'an passé. Ces deux personnes ont donc une expérience de maximum deux ou trois ans dans l'enseignement.

Cet échantillon a été choisi de manière stratégique, pour deux raisons. La première est le fait qu'il est important de recueillir les représentations d'au moins trois ou quatre personnes pour obtenir un plus large panel d'idées et d'opinions concernant le sujet des effets dans le domaine de l'évaluation. Et la deuxième raison de notre choix pour ces quatre individus est, comme susmentionné, le fait qu'il est intéressant de pouvoir découvrir les diverses manières de gérer les biais de l'évaluation, chez deux types d'enseignants différents : des enseignants expérimentés et des enseignants peu expérimentés.

III. Partie empirique

1. Données

Le guide d'entretien que nous avons réalisé, et que nous avons utilisé auprès de l'échantillon d'individus sélectionnés pour cette étude, est placé en annexes I et II de ce document. Nous avons élaboré deux guides différents : un pour les jeunes enseignants et un pour les enseignantes expérimentées. Le guide créé pour ces dernières contient simplement quelques questions supplémentaires concernant l'évolution de ces personnes au sein de la profession enseignante. Et au niveau des données récoltées lors des entretiens de recherche menés dans le cadre de ce travail, elles se trouvent sous forme de retranscriptions, dans le corpus de données, dans un document à part.

2. Analyse des données et interprétation des résultats

Passons à présent à l'analyse des données récoltées lors des entretiens de recherche, ainsi qu'à l'interprétation des résultats obtenus. Nous avons choisi de lier ces deux parties, car il nous semblait plus logique de procéder de cette manière. Pour passer de notre corpus de données à l'analyse qui suit, nous avons retenu les éléments essentiels ressortis dans les entretiens, et nous avons synthétisé les réponses apportées par les enseignants dans quatre tableaux correspondant aux thèmes de notre guide d'entretien. A la suite de l'analyse de chaque point, une partie consacrée à l'interprétation a été rédigée.

Evaluation

Lorsqu'on interroge les enseignants, choisis pour cette recherche, sur le concept d'évaluation, ils répondent ainsi :

Questions autour du thème de l'évaluation	
Enseignants	Réponses
<p><i>Enseignant A</i></p> <p><i>Enseignant expérimenté</i></p>	<p>Pour cet enseignant, évaluer, c'est contrôler ce que sait l'enfant, afin de lui dire à lui, principalement, où il en est. Il pratique l'évaluation sommative et formative dans sa classe, mais dit avoir une nette préférence pour le deuxième type d'évaluation. Il pense qu'une bonne évaluation, c'est une évaluation pensée en même temps que la préparation du cours, car il faut que ce dernier soit fait en fonction de ce que l'on veut contrôler, à la fin. Au niveau des avantages qu'il trouve à l'évaluation, il a cité le fait que cela permet de savoir si les enfants ont compris les savoirs travaillés. Enfin, concernant les risques ou inconvénients perçus dans l'évaluation, cet enseignant ressort le fait que cela provoque du stress chez les élèves, et que le danger principal serait de démontrer systématiquement à l'apprenant, qu'il n'est pas capable de retenir les savoirs travaillés en classe.</p>
<p><i>Enseignant B</i></p> <p><i>Jeune enseignant</i></p>	<p>Pour cet enseignant, l'évaluation sert à voir si les objectifs, qu'on a fixés, sont atteints ou pas. Il pratique des évaluations de types formatif et sommatif dans sa classe. Pour lui, une bonne évaluation est une évaluation qui repose sur des objectifs précis et qui est régie par des critères clairs, qui permettent d'évaluer de la manière la plus objective possible. Il fait parfois des évaluations sommatives en cours de séquence, sur un objectif en particulier, mais c'est surtout à la fin du thème qu'il fait passer un test final à ses élèves. Par rapport aux avantages que cet enseignant voit dans l'évaluation, il cite le fait que cela permet de situer l'élève par rapport à ses apprentissages, et cela donne aussi un repère à l'enseignant, pour qu'il puisse reprendre les éléments non acquis avec les apprenants qui ont encore des difficultés. Et pour les inconvénients, il explique qu'il a la crainte de ne pas réussir à voir la progression des apprenants, et donc de ne pas pouvoir les accompagner de manière appropriée.</p>
<p><i>Enseignant C</i></p> <p><i>Enseignant expérimenté</i></p>	<p>Pour cet enseignant, l'évaluation a deux grandes parties : l'évaluation notée, sommative, pour sanctionner les apprentissages, pour situer les enfants par rapport aux connaissances acquises, et pour donner un repère aux parents. Et il y a l'évaluation formative, qui permet à l'enseignant de mieux comprendre les erreurs des élèves et de trouver des remédiations à proposer à ces derniers. Pour cet enseignant, une bonne évaluation permet à l'élève de progresser dans ses apprentissages et de s'améliorer. Au niveau des avantages que cet enseignant perçoit dans l'évaluation, il relève que cela permet de situer l'élève dans son apprentissage, de voir s'il a</p>

	compris les savoirs, et que cela donne un repère aux parents. Et pour les inconvénients, l'enseignant précise qu'il peut y avoir des risques sur l'estime de soi de l'apprenant, dans le cas où il aurait très souvent des mauvaises notes. Il y a également le fait que l'évaluation peut-être assez frustrante pour certains qui auraient bien étudié, et dont la note obtenue ne reflèterait pas le travail fourni.
<i>Enseignant D</i> <i>Jeune enseignant</i>	Pour cet enseignant, évaluer permet d'orienter les élèves. Il pratique des tests formatifs et sommatifs dans sa classe. Pour lui, une bonne évaluation est une évaluation qui correspond aux objectifs demandés dans le plan d'études, et qui permet d'évaluer, de manière précise, l'élève. Pour élaborer ses examens, il part des objectifs qu'il doit faire atteindre aux enfants, et il s'aide également d'exercices ou de tests qu'il a déjà faits avec eux. Pour les avantages qu'il attribue à l'évaluation, il cite le fait que cela permet de savoir où en sont les apprenants et de voir si notre enseignement correspond à ce qu'ils ont retenu. Et au niveau des risques que cet enseignant perçoit, il ressort le fait qu'il est difficile d'adapter les consignes aux élèves, dans les évaluations, et il explique aussi que le danger, auquel il faut faire attention, est le respect des objectifs fixés. Il pense qu'il est important de garder ces repères-là et de ne pas s'en écarter dans les évaluations que l'on fait passer aux enfants.

Passons, à présent, à l'interprétation de cette analyse. Par rapport au concept d'« évaluation », on peut remarquer que les quatre enseignants le définissent en relevant sa qualité, soit, le fait de permettre de situer l'élève dans ses apprentissages. On peut dire que la signification qu'ils donnent à cet élément concerne plutôt l'objectif ou le but de l'évaluation. Si l'on fait un lien avec les auteurs cités dans notre cadre conceptuel, on se rend compte que les enseignants interrogés ont une vision plutôt similaire de celles de ces chercheurs, mais que ceux-ci, contrairement aux personnes interviewées dans la présente étude, insistent bien sur les notions de mesure et d'estimation, lorsqu'ils définissent le concept d'évaluation. Ce n'est que dans un deuxième temps, que les auteurs expliquent la fonction de cet élément. Nous pensons que la confusion faite par les enseignants entre la définition et la fonction de l'évaluation est tout à fait normale, dans le sens où ces individus sont des gens du terrain et qu'ils savent expliquer les choses de manière plus pratique, plus concrète, peut-être.

Ensuite, on retrouve les termes « formatif » et « sommatif », lorsque les personnes interrogées expliquent les types d'évaluations qu'elles pratiquent dans leur classe. Les deux enseignantes, ayant plus de quinze ans d'expérience, éprouvent une nette préférence pour l'évaluation formative, qu'elles trouvent plus utile et constructive pour l'élève. On peut relever le fait que, à leurs débuts dans l'enseignement, cela ne leur posait pas de problème de mettre une note sur les travaux des apprenants, mais qu'avec le temps et l'expérience, elles ont de plus en plus de mal à comprendre la signification et l'importance accordée à la valeur chiffrée qui est inscrite sur l'examen, et qui indique la mesure des performances des enfants. Par contre, pour les deux jeunes enseignants, le fait de noter les évaluations ne semblent pas leur poser de difficulté ou de souci, pour l'instant. Ceci nous a paru surprenant, au départ, car nous pensions que, plus l'on avait

d'expérience dans le domaine de l'évaluation sommative, moins on se posait de questions sur la procédure. Or, durant les entretiens, nous avons senti une réelle remise en question des deux enseignantes expérimentées par rapport à leur façon d'attribuer une note aux apprenants. Nous pouvons donc retenir de ceci que, plus on a du métier, plus il semble qu'on s'interroge et qu'on remette le système des notes en cause.

Au niveau de l'élaboration des évaluations, nous avons pu retrouver, chez les individus interrogés, l'importance des notions de critère et d'objectif. Le fait de poser des points de référence sur lesquels s'appuyer pour construire les évaluations semble constituer un geste professionnel aux yeux de ces quatre intervenants. Par ailleurs, dès que les interviewés expliquent leur conception de ce qu'est une bonne évaluation, on se rend compte que les quatre focalisent leur point de vue sur des aspects légèrement différents, bien que la réponse des enseignants C et D se ressemblent sensiblement. Pour ces derniers, une bonne évaluation est un moyen qui permettra de situer l'élève dans ses apprentissages, et de le faire progresser. Pour eux, une évaluation bien faite correspond donc aux qualités citées dans leur définition même de ce concept. L'enseignant A, quant à lui, s'arrête plutôt sur le moment où l'évaluation devrait être construite, soit au début d'une séquence, en même temps que le cours est pensé et préparé. Enfin, pour l'enseignant B, une bonne évaluation peut être qualifiée ainsi, seulement si elle reflète l'objectivité. Il ajoute que l'on doit absolument se baser sur des objectifs et des critères précis pour construire une évaluation idéale, et que ce n'est qu'à ces conditions que l'on pourra être vraiment professionnel dans sa façon d'évaluer. Si l'on fait un lien avec la notion de critère que nous avons définie dans notre cadre conceptuel, on peut voir notamment chez Meyer et Hadji, que ce terme est un concept-clé dans le domaine de l'évaluation. C'est grâce à cet élément que l'on peut déterminer si un objectif est atteint ou pas. On constate donc que ce jeune enseignant a associé le fait de poser des critères précis, aux notions de professionnalisme et d'objectivité. C'est ce que nous retiendrons également, essentiellement, de notre rencontre avec cette personne.

Par rapport aux risques ou inconvénients qu'implique le fait d'évaluer, on s'aperçoit très nettement, dans les réponses données, que les deux enseignantes, ayant un certain nombre d'années d'expérience, citent immédiatement les effets que cela peut avoir sur l'élève. Elles pensent que cela peut nuire à certains enfants, peut-être plus faibles, en expliquant que cela peut être pesant pour eux, et que cela peut avoir des conséquences sur leur estime d'eux-mêmes. Elles sont donc bien centrées sur l'apprenant. Pour les deux enseignants « fraîchement » sortis de la HEP, on remarque qu'il y a surtout un souci de professionnalisme dans leur façon de répondre, puisque l'enseignant B dit que le risque principal, pour lui, est de ne pas voir la progression de ses élèves, à travers l'évaluation, tandis que l'enseignant D précise que le danger se situe plutôt autour du fait de s'éloigner des objectifs fixés au début. On peut alors retenir le fait que l'expérience permet de se décentrer et de placer l'apprenant vraiment au centre de ses préoccupations, que ce soit dans l'évaluation ou dans l'enseignement en général.

Objectivité

Lorsqu'on interroge les enseignants, choisis pour cette recherche, autour du concept d'objectivité, ils répondent ainsi :

Questions autour du thème de l'objectivité	
Enseignants	Réponses
<p><i>Enseignant A</i></p> <p><i>Enseignant expérimenté</i></p>	<p>Pour lui, l'acte de correction n'est pas un moment agréable. Pour corriger de manière plus objective, il essaie de cacher les prénoms des élèves et de faire abstraction de l'auteur de la copie qu'il est en train de corriger. Mais il avoue que, tout de même, après un certain temps, on reconnaît l'écriture des apprenants. Il y a aussi le fait que l'on risque d'être plus sévère si on pense que tel ou tel enfant aurait dû faire juste un certain exercice. Cet enseignant trouve donc que ce n'est pas facile de corriger. Puis, il enchaîne en expliquant qu'un bon correcteur, pour lui, c'est une personne qui prend l'élève en compte, même si cela n'est pas très objectif. Par rapport aux risques perçus par cet enseignant, dans l'acte même de correction, il relève le fait que, lorsqu'on corrige depuis un certain temps, on n'est plus aussi attentif qu'au début de la correction. Alors il faudrait s'arrêter un peu, pour continuer un peu plus tard. Par rapport à la notion d'objectivité, cet enseignant répond immédiatement que, selon lui, c'est impossible, et que si on voulait vraiment être objectif, il faudrait que ce soit une machine qui prépare l'évaluation, et qui la corrige également. Puis, il dit qu'il pensait, auparavant, qu'il y avait certaines disciplines dans lesquelles on pouvait être plus objectif, mais aujourd'hui, il a une autre vision de ceci et est conscient que même en mathématiques, on peut être subjectif, notamment quand on demande à l'apprenant d'écrire un raisonnement. Et pour les choses que l'on pourrait mettre en place pour tendre à plus d'objectivité, l'enseignant propose d'être plusieurs correcteurs, ou encore d'essayer de faire abstraction du nom de l'élève, lorsqu'on corrige une copie.</p>
<p><i>Enseignant B</i></p> <p><i>Jeune enseignant</i></p>	<p>Cet enseignant ne fait pas toutes les corrections au même moment, car il ne veut pas être trop fatigué et risquer de mal corriger les examens de ses élèves. De plus, il précise que lorsqu'on corrige une série de copies depuis un certain temps, on est soit un peu plus souple, soit un peu plus sévère, au bout d'un moment. Ainsi, il explique qu'être objectif, pour lui, c'est préparer une grille de corrigé précis pour noter les évaluations. Puis, pour cet enseignant, un bon correcteur est une personne qui ne laisse pas de place à la subjectivité et qui pose des critères à observer. Au niveau des risques qu'il voit dans l'acte même de correction, il cite la fatigue, l'influence du nom de l'élève (auteur de la copie corrigée), et l'influence de l'examen qu'on vient de corriger avant (si ce dernier était très bon ou très mauvais). Pour cet enseignant, c'est plus difficile d'évaluer en expression orale et écrite, par exemple, ou encore lorsqu'on pose des questions ouvertes aux apprenants. Enfin, par rapport aux choses que l'on pourrait mettre</p>

	<p>en place pour tendre à plus d'objectivité, cet enseignant propose de poser des critères précis, ou encore de demander l'avis d'un deuxième correcteur sur les corrections et les valeurs inscrites sur les examens.</p>
<p><i>Enseignant C</i></p> <p><i>Enseignant expérimenté</i></p>	<p>Au niveau de l'acte de correction, cet enseignant dit que ce n'est pas un moment agréable pour lui et qu'il a un rapport trop subjectif à la valeur de la note. Pour lui, un bon correcteur, c'est une personne qui est capable de s'arrêter dans ses corrections s'il est fatigué, pour reprendre plus tard. Puis, par rapport aux risques que cet enseignant voit dans l'acte de correction, il ressort la fatigue, la mauvaise préparation de l'examen (barème et nombre de points non réfléchis). Ensuite, lorsqu'on aborde l'objectivité, il affirme que si l'on voulait être objectif, il faudrait mettre tous les élèves au même barème. Il enchaîne en précisant que dès qu'on demande un travail d'expression écrite aux apprenants, on ne peut plus être dans l'objectivité. Pour lui, dans ce domaine du français, on est forcément influencé par certains éléments. Si on sait qu'un élève a plus de facilité, on risque d'être plus sévère avec lui dans la correction et dans la notation. Donc ce n'est pas objectif. L'enseignant ajoute qu'il trouve qu'il n'y a pas d'objectivité dans les notes. Il pense que les branches dans lesquelles il est plus difficile d'être objectif sont l'expression écrite, le dessin, etc..., et même, parfois, les mathématiques, lorsqu'il y a un raisonnement à évaluer. Enfin, les éléments que l'on peut mettre en place pour tendre à plus d'objectivité sont le fait d'être pointu dans le cadre qu'on pose ; ainsi, il faut bien préparer l'examen, les critères d'évaluation, le nombre de points prévus pour chaque exercice.</p>
<p><i>Enseignant D</i></p> <p><i>Jeune enseignant</i></p>	<p>Cet enseignant dit que, lorsqu'il corrige, il évite de regarder les noms des auteurs des copies qu'il doit noter. Il ajoute qu'il procède en corrigeant exercice par exercice chez tous ses élèves. Puis, il explique qu'un bon correcteur est une personne qui va regarder les réponses des apprenants et qui va chercher à comprendre les erreurs qui ont été commises. Ensuite, les risques, qu'il perçoit dans l'acte de correction, sont les éléments qui pourraient venir perturber le moment où l'on corrige, comme le téléphone, quelqu'un qui vient nous déranger, l'humeur qu'on a, ou encore l'influence du nom de l'auteur de la copie que l'on corrige. Puis, lorsqu'on lui demande ce que signifie « être objectif », cet enseignant répond que, pour lui, c'est être le plus neutre possible, se référer à des objectifs, se tenir à un corrigé préparé rigoureusement. Il ajoute que l'on ne peut pas être complètement objectif, sauf si on fait une évaluation qui contient des réponses justes ou fausses. D'après lui, dès qu'on demande à l'apprenant de développer sa réflexion, dans une évaluation, on ne peut pas être objectif. Et lorsqu'on demande à cet enseignant, s'il y a des branches dans lesquelles il serait plus difficile d'être neutre, il répond que cela dépend plutôt du type de test fait ; par exemple, si les questions sont plutôt ouvertes, il n'est pas facile d'évaluer objectivement. Enfin, au niveau des éléments que l'on peut mettre en place pour tendre à plus d'objectivité, cet</p>

	enseignant propose d'avoir un canevas de correction très précis, avec des critères clairement définis à observer, afin de laisser le moins de hasard possible, et d'essayer de cacher le nom des élèves, pour ne pas être influencé dans sa correction.
--	---

Passons, à présent, à l'interprétation de cette analyse. Par rapport à ce concept d'objectivité, on s'aperçoit que, pour les quatre personnes, il n'est pas possible d'être totalement objectif dans ses évaluations, en règle générale. Pour y tendre, ils utilisent plusieurs petites astuces pour ne pas tomber dans la subjectivité. Deux enseignants proposent de faire abstraction des auteurs des examens, en expliquant que cet élément peut avoir une influence sur leur notation. Puis, trois des quatre individus interrogés ressortent l'importance de la notion de « critère » dans la mise en place d'une évaluation. On voit, donc, chez ces personnes, le souci de tendre à une certaine objectivité, en essayant de déterminer des éléments précis à juger et en préparant rigoureusement leurs examens. Peut-être, pourrait-on également faire un lien avec les concepts de validité et de fiabilité décrits dans notre cadre conceptuel, car lorsque les enseignants insistent sur le terme de « critères » précis à déterminer et de « corrigé » à construire, on peut penser que, finalement, s'ils accordent autant de place à ces éléments, c'est parce qu'ils représentent une source valide et fiable, qui leur permettra de noter les apprenants de manière plus juste et consciencieuse. C'est ce qu'affirme également Galiana (2005), lorsqu'il dit que, pour que l'évaluation soit objective, il faut qu'elle soit fiable et valide (Galiana, D., 2005, p. 65-69).

Ensuite, par rapport aux risques perçus par les enseignants dans l'acte même de correction, on peut déjà voir l'émergence de certains biais que nous avons relevés dans notre cadre conceptuel. Par exemple, plusieurs ont parlé de la fatigue, lorsqu'ils corrigeaient une série d'examens à la suite. Il y a aussi l'effet Pygmalion, quand l'enseignant C explique que, si l'on sait qu'un élève a plus de facilité, on risque d'être plus sévère dans la note qu'on lui attribue, car on pense qu'il aurait dû faire juste. Puis, on retrouve également l'effet de contraste, en particulier chez l'enseignant B, qui le souligne en précisant qu'un des risques majeurs de la correction est le fait de corriger un examen moyen à la suite d'une très bonne copie, car le contraste est si important, que l'on est forcément influencé dans sa notation.

Au niveau de la notion d'objectivité même, les représentations des enseignants diffèrent quelque peu, en particulier dans les compléments qu'ils ajoutent à leur définition de ce concept. Pour l'enseignant A, on s'aperçoit que le premier mot qui lui vient à l'esprit, lorsqu'on parle d'objectivité, dans le domaine de l'évaluation, est « impossible », et d'ailleurs, il n'existe pas de disciplines où l'on puisse être totalement objectif, bien que le sens commun semble laisser penser que les branches scientifiques sont moins discutables. L'enseignant C a, d'ailleurs, une idée assez proche de celle-ci, puisque, comme on peut le voir, il n'y a, pour lui, pas de possibilité d'être objectif dans la notation. Par contre, être objectif, selon cette même personne, semble plutôt être synonyme de mettre tout le monde au même niveau, soit de noter tous les apprenants de la même manière. Cet élément est intéressant, car il revient aussi chez l'enseignant A, qui pense que ce n'est pas objectif de prendre l'enfant en compte dans son évaluation, même s'il le fait, car il est convaincu que cela donne une dimension plus humaine au correcteur. Pour l'enseignant B, on voit que c'est plutôt le contraire, car ce qui ressort de ses propos est le fait que l'objectivité, dans l'évaluation, s'obtient par l'élaboration de critères précis et une préparation rigoureuse des examens. C'est donc en ne laissant aucune place à la subjectivité, que l'on pourra être un bon correcteur. Et

pour l'enseignant D, l'objectivité s'apparente à la neutralité. Pour tendre à ceci, il semble penser, tout comme ses collègues également interrogés dans cette recherche, que plus les critères posés sont précis et l'examen bien préparé, moins la subjectivité peut se « glisser » dans l'acte évaluatif.

Cet élément est très intéressant, car il nous ramène au cœur même de notre problématique. On distingue très clairement chez les deux jeunes enseignants interviewés, que le moyen de tendre à plus d'objectivité réside dans l'acte même de fixer des critères précis à contrôler dans le cadre d'une évaluation. Nous avons pu, lors des entretiens avec ces personnes, remarquer une grande assurance lorsqu'il s'agissait de parler du cadre de référence qu'ils posaient dans leurs évaluations. Par contre, chez les enseignantes plus expérimentées, certes, le terme de critère était évoqué comme moyen de tendre à plus d'objectivité, mais cela semblait être moins efficace que l'affirmaient les deux jeunes. Les enseignants A et C⁹ avaient une vision qui paraissait peut-être moins utopique que celle des deux anciens étudiants de la HEP. D'ailleurs, nous devons avouer que nous partagions cette même conception, puisque nous sommes encore novice dans la profession, et nous pensions que les critères résolvaient tout. A présent, nous prenons un peu plus de recul, en remettant en question ceci. Ce que nous retiendrons donc, ici, c'est que « poser des critères » ne permet pas forcément une objectivité totale.

Enfin, de manière générale, on découvre que les interviewés voient le concept d'objectivité comme une chose très difficile à atteindre, qui est proche des notions de neutralité et d'impartialité, et qui ne laisse pas de place au hasard. Et par rapport aux disciplines, les deux enseignantes, avec beaucoup d'années d'expérience, affirment que l'on ne peut être objectif dans aucune branche, un autre dit que c'est impossible, en particulier en expression orale et écrite, et pour le dernier, l'objectivité ne dépendrait pas de la discipline, mais plutôt du type de test que l'on propose aux élèves. On peut, alors, faire un lien avec les études menées par Laugier et Weinberg (1936) ou encore par B. Simler (1985), que nous avons relevées dans notre problématique, et qui démontraient que les épreuves du Baccalauréat comportant des questions plus ouvertes, qui demandaient une certaine réflexion de la part des étudiants, étaient plus discutables que les autres. Ce que nous pouvons retirer de ceci est le fait que, certes, on peut, de prime abord, croire que ce sont les disciplines qui impliquent plus de subjectivité dans la notation, mais en y réfléchissant bien, on se rend compte que l'incertitude dans la notation dépend plutôt du type de questions posées dans une évaluation, puisque même en mathématiques, il y avait des écarts dans le jugement des différents correcteurs. Ainsi, si les examens sont composés principalement de questions ouvertes, il est plus difficile de tendre à cette neutralité.

⁹ Plus expérimentés.

Biais de l'évaluation

Lorsqu'on interroge les enseignants, choisis pour cette recherche, sur les biais de l'évaluation, ils répondent ainsi :

Questions autour du thème des biais de l'évaluation	
Enseignants	Réponses
<p><i>Enseignant A</i></p> <p><i>Enseignant expérimenté</i></p>	<p>Par rapport aux facteurs qui viennent biaiser les évaluations, cet enseignant avait, auparavant, cité le fait de connaître l'auteur de l'examen. Il précise aussi que c'est important de faire attention aux biais de l'évaluation et qu'il faut les prendre en compte. Il avoue qu'il essaie de faire avec et a cessé de vouloir les éviter par tous les moyens. Par rapport aux facteurs rencontrés le plus souvent dans sa pratique, et qui ont une influence sur la notation, il nomme le fait de connaître les élèves et le fait de vouloir adapter l'évaluation à ce qui a été fait durant la séquence travaillée, au lieu de rester sur l'objectif fixé. Il pense aussi que l'on peut facilement être influencé, dans sa façon de juger, par les préjugés que l'on a sur certains des apprenants. Ensuite, par rapport à l'effet Pygmalion, cet enseignant est persuadé que les attentes qu'on a, concernant les enfants, ont une influence sur notre façon de noter. Puis, il dit que le soin est important dans un examen, et que cela peut avoir une certaine influence sur la notation. Quand on parle de l'effet de contraste, l'enseignant avoue qu'il croit qu'effectivement, lorsqu'on compare deux copies, on risque d'être influencé dans sa façon de juger. Puis, il confirme aussi que le regard des collègues peut changer notre notation, si l'on craint le jugement de ces derniers. Enfin, l'enseignant termine en expliquant que sa vision de l'évaluation a évolué au cours de ses années d'expérience, et qu'il a beaucoup plus conscience de ces biais, actuellement.</p>
<p><i>Enseignant B</i></p> <p><i>Jeune enseignant</i></p>	<p>Au niveau des biais relevés par cet enseignant, il y a tout ce qui concerne la forme (orthographe, soin, ...) d'une évaluation, l'influence des prénoms des élèves, la fatigue, le fait de corriger un examen moyen juste après un très bon. Il ajoute qu'il est, selon lui, important de faire attention à ces facteurs, qu'il y a été sensibilisé durant sa formation à la HEPvs, et qu'il sait donc, combien il est nécessaire de fixer des critères précis pour éviter de trop biaiser ses évaluations. Par rapport aux facteurs rencontrés le plus souvent dans sa pratique, cet instituteur cite le fait de connaître les apprenants, le fait de comparer les examens de deux enfants, ainsi que la fatigue. Concernant l'effet de stéréotypie, il pense qu'on peut être influencé par les préjugés qu'on a sur ses élèves. Puis, il affirme que nos attentes, par rapport aux apprenants, ont une influence sur la façon que nous avons de les noter. Ensuite, par rapport à l'effet de halo, il dit que l'orthographe ou le soin ne devraient pas avoir d'influence, si ces éléments ne sont pas mis comme critères d'évaluation. Puis, l'enseignant pense également que l'effet de contraste peut changer la notation. Il confirme aussi que l'on peut modifier sa</p>

	<p>manière de juger, à cause du regard des collègues, en particulier, lorsqu'on est jeune enseignant, comme lui. Enfin, par rapport à l'effet Posthumus, cet enseignant avoue qu'on peut être tenté de modifier son barème de notation, dans le cas où la majorité des apprenants ne se situent pas dans la moyenne. Par exemple, il précise que, lors du dernier examen qu'il avait fait, la plupart des enfants se situaient en dessous de la moyenne, c'est pourquoi, il leur a fait repasser un autre examen pour remonter un peu les notes obtenues par les élèves. Il a tout de même conservé les notes du premier test et a établi la moyenne des deux évaluations pour déterminer la note finale.</p>
<p><i>Enseignant C</i></p> <p><i>Enseignant expérimenté</i></p>	<p>Cet enseignant ressort les biais de l'évaluation suivants : la fatigue et les attentes en lien avec les élèves. Ces facteurs sont inhérents et il faut en avoir conscience, afin de pouvoir s'auto-corriger, en tant qu'enseignant – évaluateur. Il continue en expliquant que si l'on veut être plus juste et objectif dans sa correction, il faut absolument poser des critères d'évaluations clairs et précis. Il propose aussi de faire passer des examens standards aux apprenants, soit des tests que l'on n'a pas fait soi-même et que plusieurs classes devront effectuer. Ensuite, l'enseignant interrogé dit qu'il n'a jamais été sensibilisé au phénomène des biais de l'évaluation, durant sa formation, mais il a découvert ces éléments dans son expérience du métier. Puis, il pense que l'on peut être influencé par les préjugés et les attentes qu'on a envers les élèves. Par rapport à l'effet de halo, il affirme que la présentation, le soin ou l'orthographe peuvent être des petits détails qui influencent la notation de l'enseignant, en particulier dans un texte écrit, également. Enfin, il avoue ne pas être influencé, dans sa façon d'évaluer les apprenants, par le regard de ses collègues. De plus, il ne se laisse pas influencer par l'effet Posthumus et ne modifie pas sa notation pour que la majorité des apprenants se situent dans la moyenne. Et pour terminer, il dit qu'il a plus conscience des biais actuellement, mais qu'il croit que sa vision de ces phénomènes n'a pas beaucoup évolué.</p>
<p><i>Enseignant D</i></p> <p><i>Jeune enseignant</i></p>	<p>Les biais, que cet enseignant cite, sont l'influence du prénom de l'élève et l'humeur. Il pense que l'on peut changer sa notation si l'on connaît les élèves et si on les apprécie ou pas. Il continue en disant qu'il faut faire attention à ces facteurs, en particulier dans les examens où on demande aux apprenants d'écrire une réflexion. Puis, lorsqu'il parle des biais qu'il rencontre le plus souvent dans sa pratique de correcteur, il cite la présentation d'un examen, le soin et l'écriture, car il met beaucoup d'importance sur ces points. Ensuite, par rapport à l'effet de stéréotypie, il pense que cela peut arriver de se laisser influencer par ce biais, que ce soit dans un sens positif ou négatif pour l'enfant. Au niveau des attentes qu'on a envers les apprenants, cet enseignant est persuadé que l'on peut modifier sa façon de juger en fonction de ce phénomène-là. Par rapport à l'effet de contraste, il dit que cela ne lui arrive pas d'être influencé par ce facteur, car il corrige exercice par exercice, mais il conçoit que cela peut être le piège</p>

	classique dans lequel il ne faut pas tomber. Puis, il avoue qu'on peut être tenté de changer sa manière d'évaluer, à cause du regard des collègues. Il a beaucoup ressenti cette influence durant les remplacements effectués.
--	--

Passons, à présent, à l'interprétation de cette analyse. Parmi les biais que les enseignants connaissent déjà, on peut voir qu'il y a l'effet de stéréotypie, l'effet de halo, la fatigue, l'effet Pygmalion, et l'humeur du correcteur, ressortie seulement par l'enseignant D. Ce dernier avoue noter autrement les copies, selon l'état dans lequel il se trouve. Les facteurs relevés ci-dessus sont donc les éléments ressortis, dans un premier temps, par les enseignants interrogés. Par contre, on peut voir que leur réflexion est un peu plus poussée, lorsqu'ils doivent citer les phénomènes, qu'ils rencontrent le plus souvent, et qui biaisent leur notation. Un autre effet émerge de cette réflexion : l'effet de contraste. En général, la plupart des enseignants expliquent qu'ils connaissent un ou deux biais, mais lorsqu'ils doivent relever ceux qui viennent les perturber dans leurs évaluations, au quotidien, ils parviennent à en distinguer un peu plus. Les effets, ressortis par les individus choisis dans cette recherche, ont tous été définis dans notre cadre conceptuel, sauf ce qui concerne l'humeur. Nous avons décidé de ne pas étudier ce biais-là, car il n'avait été souligné que par De Peretti, Boniface et Legrand (2005), qui avaient émis l'hypothèse que ce phénomène biaisait aussi les évaluations.

Si l'on revient, maintenant, sur les biais de l'évaluation relevés par les enseignants interviewés, on se rend compte que la plupart d'entre eux sont conscients d'un certain nombre de ces éléments. Mais, on voit très nettement une meilleure connaissance de ces facteurs chez l'enseignant B, qui a réussi à identifier quatre effets différents, lorsque nous lui avons demandé quelles étaient les choses qui venaient déranger le moment de la correction. On pourrait alors penser que ces savoirs proviennent de la formation que cette personne a reçue à la HEPvs, mais, lorsqu'on regarde les réponses données par l'autre jeune enseignant, on se rend compte que ces éléments ne sont pas aussi présents chez ce deuxième individu. On peut alors émettre l'hypothèse que l'enseignant B était peut-être plus intéressé à ce thème, que l'enseignant D, car nous avons ressenti une réelle implication de cette première personne, dans les réponses apportées lors de l'entretien de recherche mené.

Puis, lorsqu'on aborde plus précisément l'effet de stéréotypie, l'effet Pygmalion, l'effet de contraste, l'effet de halo, et l'effet Posthumus, les avis des enseignants interrogés divergent selon le biais en question. Par exemple, pour l'effet de stéréotypie, on se rend compte que les quatre personnes interviewées sont tout à fait d'accord avec le fait qu'on peut être influencé, dans sa notation, par les préjugés que l'on a par rapport aux apprenants. D'ailleurs, si l'on fait une parenthèse en rappelant ce que nous avons décrit dans notre analyse, on remarque que les quatre individus ont ressorti le fait qu'il leur arrivait, parfois, de modifier leur jugement, en fonction de l'auteur de la copie corrigée. On peut donc s'apercevoir qu'il y a de grande chance de ne pas être objectif, lorsqu'on connaît bien les apprenants. Apparemment, le facteur humain, et peut-être plus particulièrement, la relation entretenue entre l'évaluateur et l'évalué, constitue un biais important dans l'estimation des compétences de l'apprenant, faite par l'enseignant. On peut alors faire un lien avec la recherche que nous avons citée dans notre problématique, et qui avait été effectuée par l'enquête Carnégie, par rapport aux examens oraux d'anglais. Les chercheurs de cette étude avaient démontré qu'il existait un réel écart lorsque les experts se retrouvaient en face

d'un candidat. On en revient donc aux stéréotypes que les enseignants peuvent avoir sur les élèves.

Ensuite, par rapport à l'effet Pygmalion, on s'aperçoit que les quatre personnes pensent, à l'unanimité, que l'on est forcément influencé par les attentes que l'on a concernant les apprenants. Si l'on pense qu'un élève est très doué, on le notera peut-être de manière un peu plus haute, qu'un autre enfant que l'on considère comme moins bon. Inversement, on peut être plus sévère avec un élève qu'on estime fort, car on attend plus de sa part. Nous retirons donc de ceci, le fait qu'effectivement, l'effet Pygmalion a un réel impact sur la notation des enseignants. Nous y ferons donc attention, lorsque nous serons enseignante et que nous aurons notre propre classe.

Puis, pour l'effet de contraste, on peut dire que les enseignants conçoivent que ce phénomène peut avoir une certaine influence sur l'acte évaluatif, mais l'enseignant D précise tout de même que cela ne devrait pas arriver, et que, d'ailleurs, pour contrer ce biais-là, il a développé sa propre méthode de correction, soit procéder en corrigeant exercice par exercice pour chaque examen. Ainsi, il évalue tous les exercices n°1 des copies de ses élèves, puis tous les exercices n°2, etc. Mais si l'on se remémore ce qu'est l'effet de contraste, on se rend compte qu'il peut quand même se glisser dans une telle procédure, par exemple, si l'enseignant vient de corriger un exercice très bien fait, et qu'il en corrige un autre plutôt moyen, à la suite. On retrouvera donc ce facteur, qui sera présent non pas d'un examen à l'autre, cette fois, mais d'un exercice à l'autre. Ce que l'on peut en conclure, c'est qu'il est difficile de trouver un moyen efficace de gérer cet effet, puisqu'il n'est pas aisé d'en avoir conscience.

Par rapport à l'effet de halo, les avis des enseignants sont plutôt partagés. Pour les interviewés, le soin, l'écriture ou encore l'orthographe peuvent avoir une influence considérable dans la notation, mais les enseignants A et C pensent qu'on ne devrait pas se laisser « piéger » par ce phénomène. Quant à l'enseignant B, on voit que ce biais-là n'a pas d'importance, s'il ne fait pas l'objet de critères précis posés. On constate donc que les enseignantes expérimentées disent ne pas accorder d'attention particulière à cet élément concernant la forme, alors qu'honnêtement, nous pensions que cela aurait été le contraire. Nous croyions que les personnes avec plus d'expérience seraient plus pointilleuses sur le soin et l'orthographe, car nous nous souvenons que, durant notre parcours scolaire, les instituteurs plus âgés insistaient beaucoup sur ces points. Nous avons donc été surprise de découvrir l'inverse, et surtout, de constater que le jeune enseignant D était si influencé par cet effet de halo, lors de la correction de copies.

Enfin, par rapport à l'effet Posthumus, on peut remarquer que l'enseignant B, qui a peu d'années d'expérience dans le domaine de l'enseignement, a tendance à se plier à ce biais-là, en particulier, dans les cas où la majorité des apprenants se situent en-dessous de la moyenne. Cette personne fera une autre évaluation afin de remonter la note des enfants, pour ne pas trop les pénaliser. Par contre, pour l'enseignant C, l'effet Posthumus ne l'influence apparemment pas du tout dans sa notation. On peut faire un lien entre ce biais et le fait d'être influencé par le regard des collègues, car on risque parfois de vouloir absolument situer la plus grande partie de ses élèves dans la moyenne, afin de ne pas avoir de remarques de la part d'autres enseignants de l'établissement scolaire. D'ailleurs, ce regard, porté par des personnes extérieures, a été particulièrement relevé par les deux jeunes enseignants, qui ont ressenti une réelle pression sur leurs épaules, en ce qui concerne les évaluations qu'ils réalisaient et les moyennes que les apprenants de leur(s) classe(s) obtenaient.

Ce que l'on peut retenir de ce concept de biais, est le fait que les enseignants connaissent un certain nombre de ces facteurs, auxquels ils ont pratiquement tous été confrontés, et qu'ils sont capables de les identifier dans leur pratique de l'évaluation. On peut aussi se rendre compte que les phénomènes qui sont le plus souvent relevés sont l'effet de stéréotypie, l'effet Pygmalion, la fatigue et l'effet de halo. Concernant l'effet de contraste ou encore l'effet Posthumus, les personnes interrogées ont eu plus de mal à les ressortir, mais ils ont finalement, pour la plupart, reconnu le fait que ces éléments avaient une certaine influence dans l'acte évaluatif. On voit alors que l'on touche à l'inconscient de ces enseignants, dès que l'on aborde un sujet comme celui des biais dans l'évaluation, puisque certains de ces facteurs n'ont pas été trouvés par les interviewés. Ainsi, ce que l'on peut retenir de cette partie consacrée aux biais de l'évaluation, c'est le fait que les enseignants sont conscients de certains phénomènes, et qu'ils les connaissent, plus ou moins bien. Mais, nous avons aussi pu remarquer qu'il restait une part d'inconscient et d'ignorance, que ce soit chez les jeunes enseignants ou chez ceux qui ont plus d'expérience.

Gestion des biais de l'évaluation

Lorsqu'on interroge les enseignants, choisis pour cette recherche, sur leur mode de gestion des biais de l'évaluation, ils répondent ainsi :

<i>Questions autour du thème de la gestion des biais de l'évaluation</i>	
Enseignants	Réponses
<i>Enseignant A</i> <i>Enseignant expérimenté</i>	Par rapport à la gestion des divers biais susmentionnés, cet enseignant dit qu'il imagine qu'il les gère, mais il ne sait pas de quelle façon. Puis, après réflexion, il avoue qu'il préfère les prendre en compte au lieu de les éviter, tout en précisant qu'il ne sait pas s'il fait bien de procéder ainsi. Par rapport à l'effet Pygmalion, il explique que s'il a plus d'attentes par rapport à un apprenant, c'est peut-être que c'est justifié, et que ce dernier peut donner certainement plus. Puis, pour ne pas être influencé par le regard des collègues, il se met d'accord avec les autres enseignants afin d'avoir des tests en commun. Ainsi, les autres ne pourront pas critiquer la manière d'évaluer de cet enseignant. Concernant l'effet de contraste, il dit qu'il ne fait pas attention à ce facteur lorsqu'il corrige, malgré le fait qu'il est persuadé que cela a une influence. Il continue en expliquant qu'il trouve que c'est important d'accorder une attention particulière aux biais de l'évaluation, tout en gardant à l'esprit que ces derniers font partie des aléas du métier ; mais il faut en être conscient. Enfin, il termine en disant qu'il pense gérer les biais en les acceptant, et en les intégrant à sa façon de juger les apprenants. Pour lui, l'idéal serait de ne plus avoir à mettre des notes, car cette manière de faire donne l'impression d'être objective, alors qu'en réalité, ce n'est pas le cas.
<i>Enseignant B</i> <i>Jeune enseignant</i>	Cet enseignant pense qu'il est possible de gérer les biais de l'évaluation, si on en a conscience, bien entendu. Il précise que si on ne connaît pas ces phénomènes, on peut croire que c'est normal et ne pas se remettre en question. Puis, il revient sur

	<p>l'importance de poser des critères pour éviter de trop tomber dans le piège des facteurs qui peuvent fausser la notation. Par rapport à l'effet de stéréotypie et à l'effet Pygmalion, l'enseignant dit qu'il faudrait peut-être cacher les prénoms des apprenants, sur les copies à corriger. Pour le biais de la fatigue, il conseille de ne pas corriger toute une série d'examens d'un coup. Enfin, l'enseignant insiste sur le fait qu'il ne faut pas considérer ces éléments comme des aléas du métier, car il trouve qu'il ne faut pas les accepter. Il pense que si on les connaît, il faudrait essayer de les éviter.</p>
<p><i>Enseignant C</i></p> <p><i>Enseignant expérimenté</i></p>	<p>Au niveau de la gestion des biais, cet enseignant pense qu'il est possible de les gérer et de les atténuer, si l'on en a conscience. Pour parer au biais de la fatigue, cet enseignant préfère laisser ses corrections de côté un moment, afin de ne pas évaluer de manière trop erronée. Puis, cette personne avoue qu'il est nécessaire de prendre l'élève en compte lorsqu'on évalue son travail. Il n'est donc pas trop grave de se laisser influencer par les biais qui pourraient venir de préjugés ou de la relation entre l'enfant et lui. D'ailleurs, il ajoute que le phénomène des facteurs biaisant les évaluations font partie des aléas du métier, et qu'il faut juste essayer de s'améliorer. Il pense que, des fois, un élève peut être gagnant sur quelques dixièmes, et qu'une autre fois, il sera peut-être perdant. A son avis, il n'est pas utile d'exagérer cette problématique. Mais si l'on revient à la gestion même de l'effet de stéréotypie, par exemple, cet enseignant dit qu'on pourrait essayer de cacher les prénoms des auteurs des examens, afin de ne pas tomber dans ce biais-là. Et enfin, pour éviter d'être influencé dans sa notation, il est indispensable, selon lui, de poser des critères clairs.</p>
<p><i>Enseignant D</i></p> <p><i>Jeune enseignant</i></p>	<p>Pour cet enseignant, il est possible de mettre des choses en place pour gérer les biais de l'évaluation, mais, selon lui, la correction parfaite n'existe pas. Par rapport au regard de ses collègues, c'est en allant discuter avec eux, qu'il évite de changer sa façon de juger les apprenants. Puis, pour la fatigue et l'humeur, il dit que l'idéal serait de corriger les copies dans de bonnes conditions, et de laisser parfois les examens de côté, s'il est trop fatigué, ou de ne pas corriger dans les moments où il n'est pas de bonne humeur. Pour éviter de biaiser ses évaluations à cause des préjugés qu'il a par rapport à ses élèves, il essaie de cacher les prénoms des auteurs des copies. Puis, par rapport au phénomène des biais de l'évaluation, en général, il pense que c'est important d'accorder une attention particulière à ces éléments, en particulier pour les degrés dans lesquels il enseigne, car, au canton de Vaud, c'est en 5P et 6P, que l'avenir des apprenants se décide. Ainsi, si on laisse trop les biais intervenir dans ses évaluations et dans sa notation, cela risque d'avoir des conséquences sur la suite du parcours scolaire de l'enfant. Enfin, cet enseignant pense que, pour gérer les biais de l'évaluation au mieux, il faut construire des évaluations qui reposent sur des critères précis et élaborer un corrigé clair qui permettra d'être le plus objectif possible, lors de la correction et de la notation.</p>

Passons, à présent, à l'interprétation de cette analyse. Par rapport à la gestion des divers biais mentionnés plus haut, on s'aperçoit que les enseignants ont recours à diverses méthodes, lorsqu'ils se retrouvent confrontés à ces phénomènes liés au domaine de l'évaluation. Mais avant d'entrer véritablement dans les divers dispositifs mis en place pour la gestion de ces effets, nous pensons qu'il est important de rappeler le fait que les quatre personnes interviewées semblent avoir, apparemment, conscience de ces facteurs qui biaisent leur notation. En tout cas, elles en connaissent quelques-uns. On peut affirmer ceci, puisque, d'une part, elles sont capables de citer ou d'expliquer certains biais présents dans l'acte évaluatif, et d'autre part, elles parviennent à y faire attention, lorsqu'elles corrigent leurs élèves. D'ailleurs, comme le disent les enseignants B et C, ce n'est qu'en ayant conscience des biais de l'évaluation, que l'on peut réussir à les gérer. Cette phrase peut paraître anodine, mais elle fait référence à une vérité, puisque la gestion de toute chose est sensée être un acte conscient. Ainsi, on peut dire que si l'on connaît les facteurs qui peuvent venir fausser la notation, on peut essayer de prendre du recul et de proposer une manière de prendre en compte ou d'éviter ces éléments perturbateurs.

Entrons, à présent, plus précisément dans la gestion des divers types de biais susmentionnés. Par rapport aux réponses données par l'enseignant A, on peut s'apercevoir qu'il y a une sorte de contradiction dans ses propos, car il commence par dire qu'il pense gérer les biais du domaine de l'évaluation, mais qu'il ne sait pas de quelle manière il les appréhende. Et par la suite, on découvre qu'il pense qu'il vaut mieux les prendre en compte que les éviter, pour être plus juste et pour mieux respecter les besoins des apprenants, malgré le fait que cela ne tende pas vraiment vers l'objectivité. Cet enseignant choisit donc de ne pas éviter ces facteurs à tout prix, malgré son souci de justice et de neutralité. On pourrait alors se demander, s'il n'a pas tout de suite osé avouer qu'il n'essayait pas d'éliminer ces phénomènes par tous les moyens, en raison de la crainte d'un certain jugement de notre part, ou peut-être à cause d'un dilemme qu'il ne parvient pas à résoudre lui-même.

Chez l'enseignant B, on découvre un réel souci de justice envers les apprenants par rapport au domaine de l'évaluation, notamment lorsqu'il revient sur la notion de critère, qui peut s'apparenter aux éléments qui permettront une certaine validité et fiabilité, dans la notation de cet enseignant. On s'aperçoit également que, pour lui, il paraît vraiment essentiel de ne pas aller dans le sens de ces phénomènes et d'essayer de trouver des dispositifs à mettre en place pour les éviter à tout prix. Et lorsqu'on lui demande ce qu'il fait pour gérer ces facteurs, on voit qu'il utilise notamment une méthode spécifique pour parer à l'effet de stéréotypie et à l'effet Pygmalion ; il cache les prénoms des élèves sur les examens, afin de ne pas être influencé dans sa façon de juger les apprenants. Et pour la fatigue, il conseille simplement de ne pas corriger toute une série de copies à la suite et de faire des pauses régulièrement. On peut alors se demander, si le fait que cet enseignant veuille absolument éviter les biais dans l'acte évaluatif ne viendrait pas du fait qu'il n'a pas beaucoup d'années d'expérience, ou alors, du fait qu'il y a été sensibilisé plus profondément durant sa formation à la HEP. Nous émettons cette hypothèse, car nous avons, avant de réaliser ce travail, une vision assez similaire. Nous voulions absolument éliminer ces biais, car nous croyions qu'il fallait procéder de la sorte pour être irréprochable dans le rôle de correcteur. Aujourd'hui, avec un peu plus de recul, nos conceptions sont plus nuancées.

Ensuite, chez l'enseignant C, qui a plus de trente-cinq ans d'expérience, on voit très clairement un souci de prendre en compte l'élève, dans l'évaluation que l'on fait, et dans la notation qu'on lui attribue. Pour lui, ce n'est pas grave de se laisser influencer par les effets de stéréotypie ou Pygmalion, car il pense que le but de l'évaluation n'est pas de sanctionner l'apprenant. On peut, alors s'interroger sur la raison qui pousse cet enseignant à considérer ces facteurs comme éléments moindres. Peut-être, est-ce le nombre d'années d'enseignement qui aide à relativiser et qui permet de ne pas se focaliser sur ce genre de phénomène ? Personnellement, nous pensons, que c'est effectivement l'expérience qui permet de prendre plus de recul par rapport à ces effets liés à l'évaluation. Mais peut-être, pouvons-nous également ajouter que le fait de s'informer sur ce thème, et d'en avoir une meilleure connaissance, aide aussi à dédramatiser l'impact de cet élément dans le quotidien de l'enseignant. En tous les cas, ce que l'on peut retenir de cette rencontre avec l'enseignant C, c'est le fait qu'effectivement, certains biais viennent fausser les évaluations, qu'il est important d'en avoir conscience, afin d'être bien lucide sur soi-même et sur sa pratique, mais qu'il faut aussi pouvoir se dire que ces effets font partie des aléas du métier, et qu'il n'est pas utile d'exagérer cette problématique.

Et par rapport aux méthodes de gestion des biais utilisées par cette même enseignante expérimentée¹⁰, nous pouvons ressortir les moyens suivants : pour la fatigue, elle propose de laisser les corrections de copies de côté durant un moment, afin de ne pas fausser sa notation ; pour l'effet de stéréotypie, elle pense que l'on pourrait éventuellement cacher les prénoms des auteurs des examens, mais elle est bien consciente que l'écriture de chaque apprenant est connue à partir d'un certain temps ; et finalement, elle insiste sur l'importance de poser des critères précis, en relevant le fait que c'est le meilleur moyen de ne pas laisser de place au hasard, et donc, aux facteurs biaisant les évaluations. On voit donc un lien entre ce que cette enseignante pense et ce que l'enseignant B relève également, soit, que la manière principale de contrer ces effets est le fait de fixer des éléments clairs, observables et univoques à évaluer.

Enfin, du côté de l'enseignant D, on s'aperçoit que, contrairement à la position tenue par l'enseignant C, ce jeune enseignant pense qu'il est important de bien faire attention à ce phénomène de biais dans l'évaluation. Il semble y être particulièrement sensible du fait qu'il enseigne en classe de 5-6P, qui sont, sur le canton de Vaud, des années déterminantes pour le futur parcours d'études de l'élève, puisqu'il s'agit du cycle de transition, et que c'est, déjà à ce moment-là, que les apprenants se destinent soit à aller au gymnase, soit à choisir une autre école, soit à se lancer dans un apprentissage. D'après les dires de cette personne interrogée, on peut donc déduire que si l'on ne fait pas attention aux biais et que l'on se laisse tout le temps influencer dans sa notation, cela peut, à la longue, avoir des répercussions sur l'avenir professionnel de l'enfant. Ainsi, cet interviewé donne les méthodes de gestion suivantes : pour ne pas craindre le regard des collègues, il va directement discuter avec eux, s'il a un doute sur sa notation ; pour la fatigue et l'humeur, il propose d'essayer de corriger dans les meilleures conditions possibles, soit au calme et dans un moment où il est de bonne humeur ; et enfin pour l'effet de stéréotypie et l'effet Pygmalion, il pense, tout comme les enseignants B et C, qu'il faudrait faire abstraction des prénoms des auteurs des copies à corriger.

¹⁰ Enseignant C.

Pour conclure cette partie d'analyse et d'interprétation, nous retiendrons donc le fait que les enseignants interviewés sont conscients des biais de l'évaluation, qu'ils en connaissent certains, qu'ils les ont déjà rencontrés ou qu'ils y ont déjà été sensibilisés dans leur formation, mais également le fait qu'ils ont des représentations qui convergent et d'autres qui divergent, par rapport au thème de ce mémoire. Et enfin, nous soulignons aussi le fait qu'ils ont plusieurs idées pour gérer au mieux, ces facteurs, qui viennent influencer la notation de l'enseignant. Les entretiens de recherche nous ont grandement enrichie, car nous avons découvert diverses façons d'appréhender ce phénomène des biais. Nous avons pu mieux comprendre la vision de jeunes enseignants et d'enseignants plus expérimentés, et nous en avons conclu, de manière générale, que, plus on avait d'expérience dans le métier, plus les facteurs biaisant les évaluations semblaient être des éléments gérables et sans gravité significative.

3. Vérification des hypothèses et réponse à la question de recherche

Rappelons, à présent les hypothèses que nous avons posées, en début de travail :

- **Hypothèse n°1** : Les biais dont les enseignants ont le plus conscience sont les suivants : l'effet de stéréotypie, l'effet de halo, la fatigue et l'effet Posthumus.
- **Hypothèse n°2** : Les enseignants expérimentés sont plus conscients des biais qui s'immiscent dans le processus d'évaluation que les jeunes enseignants.
- **Hypothèse n°3** : Les jeunes enseignants font attention aux éléments qui peuvent biaiser leurs évaluations, et ils essaient de les éviter par tous les moyens.

Par rapport à notre première hypothèse, on se rend compte, grâce aux entretiens menés, et à la partie analyse – interprétation, que les biais qui ont été identifiés, de prime abord, par les enseignants interrogés sont : l'effet de stéréotypie, l'effet de halo, la fatigue, l'effet de contraste et l'humeur. Ensuite, dans un deuxième temps, lorsque nous nous sommes arrêtés sur certains biais en particulier, nous avons pu remarquer que les individus interrogés n'avaient pas forcément conscience des autres facteurs qui pouvaient également venir biaiser leurs évaluations, comme l'effet Pygmalion, ou encore l'effet Posthumus. Mais, la plupart d'entre eux ont reconnu qu'effectivement, ces effets pouvaient influencer sensiblement leur manière de noter. Si l'on revient donc à notre hypothèse n°1, on se rend compte que nous avons raison par rapport à l'effet de stéréotypie, l'effet de halo, et la fatigue, car les enseignants ont conscience de ces phénomènes-là, lorsqu'ils évaluent leurs élèves, mais par contre, pour l'effet Posthumus, nous avons pu remarquer qu'aucun des interviewés n'avaient pu identifier ce facteur. Nous pensons que cela est dû au fait que ce biais est probablement le plus difficile à se représenter, car il n'est pas facile de prendre du recul par rapport à sa façon de noter, et de conscientiser le fait que nous modifions notre notation en fonction du niveau des apprenants qui sont dans notre classe. Ainsi, ce que nous pouvons retenir de ceci, c'est le fait que les biais dont les enseignants sont le plus conscients sont les effets liés à la relation enseignant – élève, les effets liés à la forme d'une évaluation, ainsi que les effets liés aux conditions de correction des évaluations.

Par rapport à notre deuxième hypothèse, nous ne pouvons affirmer que les enseignants expérimentés sont plus conscients des biais de l'évaluation, que les enseignants plus jeunes. L'enseignant B, ancien étudiant de la HEP, est la personne qui a su ressortir le plus de biais différents, certainement grâce à sa

formation professionnelle, au cours de laquelle elle a étudié ces phénomènes. Mais l'on remarque que les interviewés plus âgés étaient également bien conscients de ces phénomènes et qu'ils ont eu de la facilité à en parler également. On se rend compte que ces derniers ont découvert ces éléments à travers leur pratique et que, grâce à leur grande expérience du métier, ils sont tout à fait aptes à identifier et à expliquer la plupart des facteurs biaisant les évaluations. Les deux enseignantes plus expérimentées nous ont précisé qu'elles n'avaient pas étudié ce genre de savoirs durant leurs années à l'Ecole Normale, et que c'est vraiment à travers leur pratique de l'évaluation qu'elles ont pu expérimenter ces facteurs. On peut alors retenir le fait que la formation peut avoir une influence sur la conscientisation de ces effets liés à l'acte évaluatif, mais que l'expérience est également un bon moyen de découvrir et de conscientiser ces éléments. Il n'y a donc pas de différence réellement marquée entre les jeunes enseignants et les enseignants plus expérimentés, par rapport à la conscientisation des biais de l'évaluation, en général.

Puis, on peut voir que notre troisième hypothèse était plutôt juste, puisque les deux jeunes personnes, soit les enseignants B et D, ont insisté, à plusieurs reprises, sur l'importance de se rendre compte du fait qu'il y a différents éléments qui viennent fausser la notation, et surtout, qu'il faut absolument essayer de les réduire, voire de les éviter complètement, si possible. Et du côté des enseignants A et C, on a pu remarquer que les avis étaient un peu moins tranchés, et que, ces personnes préféraient faire avec ces facteurs, car elles trouvaient logique de prendre l'élève en compte dans leur notation, et ceci, bien que la notion d'objectivité ne soit pas aussi présente qu'elle devrait l'être dans le contexte de l'évaluation scolaire. Pour revenir aux jeunes enseignants, nous avons remarqué, durant les entretiens, que ces derniers donnaient beaucoup d'importance aux biais de l'évaluation, qu'ils trouvaient que ces éléments pouvaient avoir des impacts conséquents sur les apprenants, et qu'il fallait, selon eux, absolument les éviter et essayer de les supprimer complètement de l'acte évaluatif.

Et si l'on revient, pour terminer, à notre question de recherche qui était « *Quels sont les biais dont les enseignants ont le plus conscience et de quelle façon les gèrent-ils ?* », on peut, à présent, y répondre. Concernant la première partie de notre question, on peut affirmer que les effets, dont les enseignants, interrogés dans cette recherche, ont le plus conscience, sont l'effet de stéréotypie, l'effet de halo, et la fatigue. Bien entendu, d'autres biais ont été évoqués, mais ce sont ces facteurs-là qui ont été identifiés le plus souvent. Et pour la deuxième partie de notre question, on s'aperçoit qu'il y a divers moyens de gérer les biais, mais que les points qui ressortent le plus souvent sont le fait de ne pas corriger une série d'examens à la suite, pour contrer la fatigue, le fait de faire abstraction de l'auteur d'une copie à corriger, pour éviter d'être influencé par les effets Pygmalion ou de stéréotypie, et enfin, le fait de poser des critères bien précis, afin de laisser le moins de place possible à la subjectivité et à ces biais de l'évaluation en général.

4. Analyse critique

Prenons, à présent, une certaine distance par rapport à cette démarche, afin de relever les points positifs et les limites de cette recherche. Tout d'abord, dans les forces que nous relèverions, nous citerions le fait que notre guide d'entretien a été assez bien construit, car il nous a permis d'obtenir des réponses aux questions que nous avions. Ainsi, nous avons pu confirmer ou infirmer nos hypothèses, et nous avons également pu répondre à notre question de recherche. Ensuite, la deuxième chose que nous aimerions ressortir est le fait que notre cadre conceptuel est plutôt complet et qu'il a pu s'enrichir grâce aux auteurs qui ont traité

les divers concepts, décrits dans ce mémoire, dans leurs travaux de recherche autour du domaine de l'évaluation. Et enfin, le dernier point fort sur lequel nous voudrions insister est le fait que ce travail est tout à fait pertinent dans le cadre de nos études à la HEP, car le sujet, que nous avons approfondi ici, s'insère directement dans l'un des domaines scolaires principaux¹¹. Aussi, la démarche, que nous avons menée, nous a permis d'assimiler de nouvelles connaissances, par rapport aux savoirs scientifiques déjà acquis lors des thèmes suivis au cours de notre formation professionnelle¹².

Concernant, les points un peu plus négatifs, maintenant, nous aimerions, dans un premier temps, relever le fait que les données récoltées ne peuvent pas réellement confirmer ou infirmer nos hypothèses, puisque nous n'avons interrogé que quatre enseignants et que, par conséquent, notre échantillon est trop réduit pour pouvoir ressortir des généralités, qui pourraient s'approcher le plus possible de la réalité, concernant les biais de l'évaluation. Nous proposerions, pour remédier à ceci, d'augmenter notre effectif de personnes à interviewer dans le cadre de notre recherche. Ainsi, nos données, notre analyse et notre interprétation seraient probablement plus valide. Et le deuxième point faible que nous relèverions, par rapport à ce travail, est le fait que notre recherche peut paraître un peu négative, au premier abord, puisqu'elle met en avant, un élément plutôt controversé dans le domaine de l'enseignement. Or, il est important que le lecteur prenne du recul par rapport au contenu de ce mémoire, afin de comprendre que notre but final est de mettre en avant le fait que les enseignants sont bien conscients de cette problématique liée à l'acte évaluatif, et qu'ils essaient d'y trouver des solutions en y réfléchissant et en y faisant attention, lorsqu'ils doivent juger les connaissances et compétences de leurs élèves.

IV. Conclusion

Pour terminer ce travail, nous aimerions revenir sur le problème posé par un sujet tel que celui des biais dans l'évaluation. Comme nous l'avons expliqué dans notre résumé, nous avons un grand sens de la justice et c'est une valeur qui est mise au premier rang, dans notre conception de l'enseignement, et de la vie en général. Lorsque nous avons constaté, durant notre parcours scolaire et professionnel, qu'il existait des éléments biaisant la réalité et l'objectivité des individus, nous nous sommes interrogée sur ces phénomènes qui nous dérangeaient profondément, puisqu'ils empêchaient, en quelque sorte, une certaine justice.

Honnêtement, lorsque nous avons décidé de construire notre mémoire sur le thème des biais dans l'évaluation, nous étions, au départ, motivée par le fait de prouver qu'effectivement, les enseignants ne sont pas irréprochables dans leur rôle d'évaluateur, et que leur jugement pouvait être altéré. Nous voulions donc découvrir la vérité dans l'univers un peu secret des évaluations et de la notation, puisque, comme le disait l'enseignant A, « On n'aime pas trop dire ce qu'on a mis comme note »¹³.

Puis, après avoir débuté nos recherches sur le sujet, soit dans des ouvrages, soit auprès de connaissances liées à la profession, nous avons commencé à voir les choses sous un angle légèrement différent. Au fil des entretiens menés, nous découvrions le phénomène des biais sous un nouveau regard, certainement avec plus de recul que nous avions, de prime abord, dans notre condition d'étudiante. C'est

¹¹ Le domaine de l'évaluation.

¹² Formation initiale HEPvs / Thèmes autour de l'évaluation : Thèmes 4.5 (Notions de base en évaluation) et 4.6 (Evaluation formative et différenciation).

¹³ Corpus de données : Retranscription de l'entretien n°1.

alors que nous nous sommes rendu compte que les enseignants, bien qu'élevés au rang de modèles, étaient des personnes comme les autres, qui faisaient des erreurs et qui n'étaient pas parfaites. Nous avons d'ailleurs beaucoup apprécié l'intervention des quatre individus interrogés, qui ont souligné l'importance d'avoir conscience des biais de l'évaluation, et qui pensent que la seule chose à faire, pour être un bon évaluateur, est d'essayer de reconnaître ces éléments, et surtout, de tenter de s'améliorer soi-même.

Enfin, ce que nous retenons de cette recherche, est le fait qu'il existe bien des effets biaisant les évaluations, que les enseignants ont conscience de certains de ces phénomènes, et qu'ils essaient de répondre à leur rôle d'évaluateur du mieux qu'ils peuvent. Ils gèrent ces éléments à leur façon, dans le but, bien sûr, d'accompagner l'élève de manière appropriée. En effet, comme le disent Christine Tagliante (2001) et les enseignants interrogés dans cette étude, ce qui est important, c'est le fait de connaître ces biais pour pouvoir se rendre compte de l'impact qu'ils ont sur notre pratique de l'évaluation sommative. Nous terminerons donc cette partie avec cette petite phrase – conseil : Ne pointe pas du doigt les erreurs de jugement de ton voisin, offre-lui plutôt une chance de s'améliorer !

1. Apports formatifs

En élaborant ce mémoire de fin d'études à la Haute Ecole Pédagogique du Valais, nous avons pu acquérir certaines connaissances et développer plusieurs compétences bien précises. Ces éléments sont présentés dans le bilan formatif du mémoire de fin d'études à la HEPvs, dans un document à part.

2. Prolongements et perspectives

Pour terminer notre mémoire, nous allons présenter quelques prolongements et perspectives possibles à notre recherche. Comme nous l'avons expliqué dans les limites de notre étude, nous pensons que nous avons interrogé trop peu d'enseignants pour pouvoir répondre de manière plus correcte à notre question de recherche et à nos hypothèses. Bien que certaines des représentations des interviewés se soient retrouvées chez plusieurs, on ne peut, à cause de l'effectif réduit de l'échantillon choisi, tirer de réelles conclusions valables pour le reste de la population enseignante.

Ainsi, le premier prolongement que nous proposerions serait de reprendre cette étude, et de choisir un échantillon de personnes plus important. En procédant ainsi, plusieurs options s'offriraient à nous. Par exemple, nous pourrions, soit sélectionner des jeunes individus avec très peu d'expérience dans le métier d'enseignant. Cela pourrait peut-être amener d'autres résultats. Ou alors, on pourrait se concentrer uniquement sur des personnes qui ont plus de trente ans d'expérience dans l'enseignement, et nous obtiendrions peut-être des réponses allant dans le sens des deux enseignants plus expérimentés, que nous avons rencontrés dans le cadre de cette recherche. Cela pourrait être intéressant de faire une différence plus marquée entre ces deux types d'enseignants. En fait, c'est un peu ce que nous avons essayé de réaliser à travers ce travail, mais il est vrai que si l'on prenait le temps de faire des entretiens avec au moins dix jeunes enseignants et dix enseignants qui ont plus d'expérience, puis, dans un deuxième temps, de comparer leurs représentations, cette étude aurait une validité plus grande.

Et le deuxième prolongement que l'on pourrait faire serait de chercher à découvrir, en direct, si les biais de l'évaluation sont des éléments qui influencent réellement

la notation des enseignants, et pour ceci, il faudrait procéder par la méthode de recherche « expérimentation ». On pourrait, par exemple, demander à plusieurs enseignants de corriger et de noter les examens des apprenants d'une même classe, et d'un même degré. Dans cet échantillon de personnes, on prendrait le titulaire de la classe concernée, et d'autres enseignants venant du même établissement et étant eux-mêmes titulaires de classes du même degré, ainsi que d'autres instituteurs encore, qui viendraient d'autres écoles, mais qui s'occuperaient toujours du même degré. Il serait donc intéressant d'observer si, aux yeux de l'échantillon d'individus choisis, tous noteraient les copies des élèves de la même manière ou non. Comme expliqué dans notre problématique, des études similaires ont déjà été menées, mais on pourrait retenter l'expérience, afin de découvrir, si chez nous actuellement, en Suisse, et plus particulièrement en Valais, on retrouve autant de différences que dans les recherches effectuées par Laugier et Weinberg, en France, en 1936. Mais, en réalisant une telle étude, il faudrait être prêt à recevoir de nombreuses critiques, car toute cette mise en œuvre pourrait avoir des conséquences plutôt négatives, puisque cela ne ferait qu'alimenter encore plus les polémiques qui touchent au domaine scolaire, et plus particulièrement aux compétences des enseignants à attribuer des notes aux performances des enfants.

Enfin, au niveau des perspectives que ce travail de mémoire met en avant, nous pouvons citer le fait qu'il nous permet à nous, auteure, de faire évoluer notre pratique de l'évaluation, en tant qu'enseignante, mais qu'il pourra aussi, certainement, modifier le regard que les éventuels lecteurs de ce document pourraient avoir sur le phénomène des biais de l'évaluation. D'un point de vue plus personnel, les perspectives qui s'offrent à nous, après la réalisation de cette recherche, sont le fait de pouvoir être un peu plus au clair sur les deux principaux types d'évaluations, le fait de pouvoir prétendre mieux connaître les « défauts » de l'acte évaluatif, mais surtout, le fait d'être apte à expliquer ces éléments à autrui et de pouvoir sensibiliser, peut-être, mes futurs collègues, à ces facteurs, en leur proposant les diverses pistes, que nous avons pu recueillir, pour gérer ces phénomènes.

V. Références bibliographiques

Ouvrages

- Beauté, J. (1995). *Les courants de la pédagogie contemporaine*. Lyon : Chronique sociale.
- Black, P., William, D. (1998). « Assessment and Classroom Learning ». *Assessment in Education. Principles. Policy and Practice*. Vol.5, n°1, pp. 7-73. Cité par : Smith, I., (2009). *L'évaluation et l'apprentissage*. Clés pédagogiques. pp. 12. Montréal : Chenelière Education.
- Blanchet, A., (et al.) (2005). *Les techniques d'enquête en sciences sociales*. Paris : Dunod.
- Blanchet, A., Gotman, A. (2007). *L'enquête et ses méthodes. L'entretien*. Sociologie 128. Paris : Armand Colin.
- De Ketele, J., Roegiers, X. (2009). *Méthodologie du recueil d'informations. Fondements des méthodes d'observation, de questionnaire, d'interview et d'étude de documents*. 4^e édition. Méthodes en sciences humaines. Bruxelles : De Boeck Université.
- De Landsheere, G. (1979). *Dictionnaire de l'évaluation et de la recherche en éducation*. Paris : PUF.
- De Peretti, A., Boniface, J. & Legrand J., (2005). *Encyclopédie de l'évaluation en formation et en éducation. Guide pratique. Pédagogies outils*. Issy-Les-Moulineaux : ESF éditeur.
- Galiana, D. (2005). *Mémento de l'évaluation. Analyser et améliorer sa pratique de l'évaluation. Guide méthodologique*. Dijon : Educagri éditions.
- Guillon, C. (2004). *Les évaluations scolaires. Ressources formation. Enjeux du système éducatif*. Paris : Hachette Education.
- Hadji, C. (1990). *L'évaluation, règles du jeu. Des intentions aux outils. Collection pédagogies*. 2^e édition. Paris : ESF.
- Hartog, P., Rhodes, E. C. (1936). *An examination of Examinations*. Londres: Macmillan. Cité par : Merle, P. (1998). *Sociologie de l'évaluation scolaire. Le point des connaissances actuelles*. Que sais-je ? pp. 10. Paris : Presses Universitaires de France.
- Jones, R. (2000). *Méthodes de recherche en sciences humaines. Méthodes en sciences humaines*. Bruxelles : De Boeck Université.
- Jorro, A. (2000). *L'enseignant et l'évaluation. Des gestes évaluatifs en question. Pratiques pédagogiques*. Bruxelles : De Boeck Université. Page Web. Accès : <http://books.google.ch/books?hl=en&lr=&id=mlIx-pefrsAC&oi=fnd&pg=PA5&dq=repr%C3%A9sentations+enseignants+%C3%A9valuation&ots=pSLUUGZeJA&sig=dJMHTfBczJ3PvNyvB6R3ZpeVsLY#v=onepage&q&f=false> (consulté le 10 janvier 2011).
- Laugier, H., Weinberg, D., (1936). *Commission française pour l'enquête Carnégie sur les examens et concours. La correction des épreuves écrites au baccalauréat*. Paris : Maison du livre. Cité par : Merle, P. (1998). *Sociologie de l'évaluation scolaire. Le point des connaissances actuelles*. Que sais-je ? pp. 7. Paris : Presses Universitaires de France.
- Long, D. *Introduction à la recherche. La problématique de recherche*. Canada : Université de Moncton. Centre de Recherche et de Développement en Education (CRDE).
- Long, D. *Introduction à la recherche. Les étapes de la recherche*. Canada : Université de Moncton. Centre de Recherche et de Développement en Education (CRDE).
- L'Hôte, M. (1987). *Les notes à l'école. L'école des parents*. Paris : Syros.

- Mancovsky, V. (2006). *L'évaluation informelle dans l'interaction de la classe. Savoir et formation*. Paris : L'Harmattan.
- Merle, P. (1996). *L'évaluation des élèves. Enquête sur le jugement professoral*. Paris : PUF.
- Merle, P. (1998). *Sociologie de l'évaluation scolaire. Le point des connaissances actuelles. Que sais-je ?* Paris : Presses Universitaires de France.
- Meyer, G., (2007). *Evaluer, pourquoi ? Comment ? Livrets d'évaluation. Fiches de suivi individuel. Fiches de suivi collectif. Profession enseignant*. Paris : Hachette.
- Noizet, G., Caverni, J. (1978). *Psychologie de l'évaluation scolaire*. Paris : PUF. Cité par : De Peretti, A., Boniface, J. & Legrand J., (2005). *Encyclopédie de l'évaluation en formation et en éducation. Guide pratique. Pédagogies outils*. Issy-Les-Moulineaux : ESF éditeur.
- Perrenoud, P. (1999). *L'évaluation des élèves. De la fabrication de l'excellence à la régulation des apprentissages. Entre deux logiques. Pédagogies en développement*. Bruxelles : De Boeck Université.
- Piéron, H. (1963). *Examens et docimologie*. Paris : PUF. Cité par : Merle, P. (1998). *Sociologie de l'évaluation scolaire. Le point des connaissances actuelles. Que sais-je ?* pp. 5. Paris : Presses Universitaires de France.
- Piéron, H., Reuchlin, M. & Bacher, F. (1962). Données comparatives sur la distribution des notes aux écrits et oraux du baccalauréat d'organisation classique, *Biotypologie*. 23, pp. 1-16. Cité par : Merle, P. (1998). *Sociologie de l'évaluation scolaire. Le point des connaissances actuelles. Que sais-je ?* pp. 11. Paris : Presses Universitaires de France.
- Quivy, R. et Campenhoudt, L. (2006). *Manuel de recherche en sciences sociales. 3^e édition*. Paris : Dunod.
- Scallon, G. (2004). *L'évaluation des apprentissages dans une approche par compétences. Pédagogies en développement*. Bruxelles : De Boeck Université.
- Simler, B. (1985). *Comportement de notation au baccalauréat. Une étude comparée dans trois académies : Lille, Nancy – Metz, Poitiers*. Etude ronéotypée. Cité par : Merle, P. (1996). *L'évaluation des élèves. Enquête sur le jugement professoral*. pp. 217-224. Paris : PUF.
- Smith, I., (2009). *L'évaluation et l'apprentissage. Clés pédagogiques*. Montréal : Chenelière Education.
- Sutter, cité par : De Landsheere, G. (1979). *Dictionnaire de l'évaluation et de la recherche en éducation*. pp. 212. Paris : PUF. Tagliante, C. (2001). *Techniques de classe. L'évaluation*. Paris : CLE International.
- Talbot, L. (2009). *L'évaluation formative. Comment évaluer pour remédier aux difficultés d'apprentissage. Sciences humaines et sociales. U Collection*. Paris : Armand Colin.

Revues

- Bonniol, J. (1965). Les divergences de notation tenant aux effets d'ordre de la correction. *Cahiers de psychologie*, n°8, pp. 181-188. Cité par Merle, P. (1998). *Sociologie de l'évaluation scolaire. Le point des connaissances actuelles. Que sais-je ?* pp. 16-17. Paris : Presses Universitaires de France.
- Cambon, J. (1979). *Revue française de pédagogie*, vol. 49, n°49. pp. 72-73. Psychologie de l'évaluation scolaire. Persee. Revues scientifiques. Page Web. Accès : http://www.persee.fr/web/revues/home/prescript/article/rfp_0556-7807_1979_num_49_1_2192_t1_0072_0000_2 (consulté le 1 septembre 2010).
- Christinat, C. & Delévaux, O., HEPL, Lausanne, (2010). Evaluer et différencier : des enseignants spécialisés relèvent le défi. *Educateur. Ser, 05/2010, No spécial 2010*, 27-29.
- De Ketele, J. (1993). *Revue française de pédagogie*, vol. 103, n° 103. pp. 59-80. L'évaluation en paradigmes. Persee. Revues scientifiques. Page Web. Accès :

- http://persee.fr/web/revues/home/prescript/article/rfp_0556-7807_1993_num_103_1_1298 (consulté le 22 juin 2010).
- Durkheim, E. (1898). Représentations individuelles et représentations collectives. *Cahiers de psychologie politique. Revue d'informations, de réflexion et de recherche.* n° 8. Janvier 2006. Page Web. Accès : <http://lodel.irevues.inist.fr/cahierspsychologiepolitique/index.php?id=1084> (consulté le 10 janvier 2011).
 - Enright, M., Favre Marmet A. & Tessaro W. (2009, septembre). Le jugement professionnel en évaluation. Les caractéristiques du jugement professionnel. *Educateur.* Ser, 9/2009, 27-30.
 - Perrenoud, P (2010, mai). Enseigner : un métier sous contrôle ?. Le contrôle, un sujet qui dérange. *Educateur.* Ser, 05/2010, 26-27.
 - Tochon, F. (1992). A quoi pensent les chercheurs quand ils pensent aux enseignants. Les cadres conceptuels de la recherche sur la connaissance pratique des enseignants. *Revue française de pédagogie.* n° 99. pp. 89-113. Page Web. Accès : http://www.inrp.fr/publications/edition-electronique/revue-francaise-de-pedagogie/INRP_RF099_7.pdf (consulté le 10 janvier 2011).

Sites Internet

- *Introduction à la recherche de l'Université de Moncton, Canada.* Page Web. Accès : http://www3.umoncton.ca/templates/udem_gen_01.cfm?CFID=137177&CFTOKEN=24407998&user_id=546&page=15725&template=42&resultat=0&order_num=&mot_recherche=&write=0&student_id=0&debut=0&curr_page=1 (consulté le 23 août 2010).
- Merle, P. *A quoi servent les notes ? Education. L'école en questions > L'école en questions.* Sciences Humaines.com. Page Web. Accès : http://www.scienceshumaines.com/a-quoi-servent-les-notes-_fr_14909.html (consulté le 19 février 2010).
- Perrenoud, P. (1989). *Bulletin de l'Association des enseignants et chercheurs en éducation (Paris)*, n° 6, pp. 19-31. Vers une sociologie de l'évaluation. Faculté de psychologie et des sciences de l'éducation. Université de Genève. Page Web. Accès : http://www.unige.ch/fapse/SSE/teachers/perrenoud/php_main/php_1989/1989_01.rtf (consulté le 21 juin 2010).

Documents académiques (institution HEPvs)

- *Document des critères pour le mémoire à la HEPvs.* Réseau Recman. St-Maurice : HEPvs. Accès : NetStorage/DriveR@DATA/3_FI_GA/3.13_Examen_Abschluss/Session 2011-2012 (consulté le 26 janvier 2011).
- *Référentiel de compétences de la HEP Valais.* Réseau Recman. St-Maurice : HEPvs. Accès : NetStorage/DriveR@DATA/3_FI_GA/3.12.3_Portfolio/St-Maurice_Guide de l'etudiant/3. Referentiel de competences (consulté le 26 janvier 2011).

VI. Annexes

1. Annexe I : Guide d'entretien I

Guide d'entretien pour les jeunes enseignants

Informations générales

Date / Heure :

Lieu :

Nom de l'interviewer : Alizée Buthey

Code de l'interviewé : Enseignant B / D

Thèmes et questions à traiter
Présentation
Lieu et degré d'enseignement / Formation / Expérience / Autre...
Evaluation
Pour vous, qu'est-ce que l'évaluation scolaire ?
Quel(s) type(s) d'évaluation pratiquez-vous, en général, dans votre classe ?
Qu'est-ce qui fait une bonne évaluation, selon vous ?
Comment élaborez-vous vos évaluations ? <ul style="list-style-type: none">- Comment construisez-vous vos évaluations ?
Quels sont les avantages de l'évaluation ?
Quels sont les risques que vous percevez au moment où vous évaluez les élèves ?
Objectivité
Comment gérez-vous l'acte de correction ? <ul style="list-style-type: none">- Que se passe-t-il derrière la correction ?- Quels sont les enjeux de la correction ?
Pour vous, qu'est-ce qu'un bon/mauvais correcteur ?
Quels sont les risques que vous percevez au moment où vous corrigez ?
Selon vous, que signifie « être objectif » ? <ul style="list-style-type: none">- Que pensez-vous de l'objectivité dans le domaine de l'évaluation ?
Pensez-vous que l'on puisse être complètement objectif ? <ul style="list-style-type: none">- Si oui, de quelle manière s'y prend-on ?- Si non, dans quelle situation est-ce impossible ?- Comment procéderiez-vous pour être le plus objectif possible ?
Dans quel cas l'objectivité est-elle plus difficile à atteindre ? <ul style="list-style-type: none">- Et dans ces cas, comment procédez-vous pour l'atteindre ?
Est-ce qu'il existe certaines choses que l'on peut mettre en place pour être le plus objectif possible ? <ul style="list-style-type: none">- Que mettez-vous en place pour que la note corresponde le plus à la réalité ?

Biais de l'évaluation

Quand vous corrigez un examen, est-ce que vous avez l'impression qu'il y a certains facteurs ou éléments qui empêchent l'objectivité ?

- Si oui, lesquels ?
- Si non, pourquoi pensez-vous ceci ?
- Connaissez-vous certains de ces biais ?
- Si oui, pourriez-vous les nommer et/ou les définir ?

Ces facteurs portent le nom de « biais » de l'évaluation. Pour vous, que représentent-ils, autrement dit, quelle signification leur donnez-vous ?

- Avez-vous déjà entendu parler de ces phénomènes de correction ?
- Est-ce que vous en connaissez d'autres ?

A votre avis, dans votre pratique de correcteur, quels sont les biais rencontrés le plus souvent ?

Questions sur divers types de biais :

- Est-ce que vous pensez que l'on peut être influencé par certains préjugés que l'on pourrait avoir par rapport à un élève ?
- Est-ce que l'on n'a pas plus tendance à oublier des erreurs dans l'examen d'un élève plutôt doué ?
- Pensez-vous que les attentes que l'on a par rapport à un certain élève ont une influence également sur la notation ?
- Et au niveau du soin, de la présentation et/ou de l'orthographe d'une copie, pensez-vous que ces éléments peuvent influencer la note ?
- Si l'on corrige un très bon examen, puis un examen plus faible, est-ce que l'on n'a pas tendance à noter encore plus sévèrement la copie plus mauvaise ?
- Pensez-vous que vous êtes influencé par le regard de vos collègues sur votre façon d'attribuer des notes à vos élèves ?
 - o N'êtes-vous pas tenté de modifier votre notation, si vous avez une majorité de très bonnes notes ou de très mauvaises notes ?

Gestion des biais de l'évaluation

Pensez-vous qu'il soit possible de gérer ces facteurs ou biais ?

- Si oui, de quelle manière ?
- Si non, pour quel(s) motif(s) ?

Par rapport aux divers biais mentionnés tout à l'heure, comment gérez-vous chacun de ces facteurs ?

Est-ce important, selon vous, de porter une attention particulière à ces biais, ou faut-il accepter que ces phénomènes font partie des aléas du métier ?

- Pour quelle raison pensez-vous ceci ?

Quel(s) conseil(s) donneriez-vous, par rapport à la gestion de ces types de biais, à un jeune enseignant qui vient de commencer à enseigner et donc à évaluer ses élèves ?

- Pour permettre à un jeune collègue de tendre vers l'objectivité quels conseils donneriez-vous ?

2. Annexe II : Guide d'entretien II

Guide d'entretien pour les enseignants expérimentés

Informations générales

Date / Heure :

Lieu :

Nom de l'interviewer : Alizée Buthey

Code de l'interviewé : Enseignant A / C

Thèmes et questions à traiter
Présentation
Lieu et degré d'enseignement / Formation / Expérience / Autre...
Evaluation
Pour vous, qu'est-ce que l'évaluation scolaire ?
Quel(s) type(s) d'évaluation pratiquez-vous, en général, dans votre classe ?
Qu'est-ce qui fait une bonne évaluation, selon vous ?
Comment élaborez-vous vos évaluations ? <ul style="list-style-type: none">- Comment construisez-vous vos évaluations ?
Quels sont les avantages de l'évaluation ?
Quels sont les risques que vous percevez au moment où vous évaluez les élèves ?
Objectivité
Comment gérez-vous l'acte de correction ? <ul style="list-style-type: none">- Que se passe-t-il derrière la correction ?- Quels sont les enjeux de la correction ?
Pour vous, qu'est-ce qu'un bon/mauvais correcteur ?
Quels sont les risques que vous percevez au moment où vous corrigez ?
Selon vous, que signifie « être objectif » ? <ul style="list-style-type: none">- Que pensez-vous de l'objectivité dans le domaine de l'évaluation ?
Pensez-vous que l'on puisse être complètement objectif ? <ul style="list-style-type: none">- Si oui, de quelle manière s'y prend-on ?- Si non, dans quelle situation est-ce impossible ?- Comment procéderiez-vous pour être le plus objectif possible ?
Dans quel cas l'objectivité est-elle plus difficile à atteindre ? <ul style="list-style-type: none">- Et dans ces cas, comment procédez-vous pour l'atteindre ?
Est-ce qu'il existe certaines choses que l'on peut mettre en place pour être le plus objectif possible ? <ul style="list-style-type: none">- Que mettez-vous en place pour que la note corresponde le plus à la réalité ?
Biais de l'évaluation
Quand vous corrigez un examen, est-ce que vous avez l'impression qu'il y a certains

facteurs ou éléments qui empêchent l'objectivité ?
<ul style="list-style-type: none"> - Si oui, lesquels ? - Si non, pourquoi pensez-vous ceci ? - Connaissez-vous certains de ces biais ? - Si oui, pourriez-vous les nommer et/ou les définir ?
<p>Ces facteurs portent le nom de « biais » de l'évaluation. Pour vous, que représentent-ils, autrement dit, quelle signification leur donnez-vous ?</p> <ul style="list-style-type: none"> - Avez-vous déjà entendu parler de ces phénomènes de correction ? - Est-ce que vous en connaissez d'autres ?
<p>A votre avis, dans votre pratique de correcteur, quels sont les biais rencontrés le plus souvent ?</p>
<p>Questions sur divers types de biais :</p> <ul style="list-style-type: none"> - Est-ce que vous pensez que l'on peut être influencé par certains préjugés que l'on pourrait avoir par rapport à un élève ? - Est-ce que l'on n'a pas plus tendance à oublier des erreurs dans l'examen d'un élève plutôt doué ? - Pensez-vous que les attentes que l'on a par rapport à un certain élève ont une influence également sur la notation ? - Et au niveau du soin, de la présentation et/ou de l'orthographe d'une copie, pensez-vous que ces éléments peuvent influencer la note ? - Si l'on corrige un très bon examen, puis un examen plus faible, est-ce que l'on n'a pas tendance à noter encore plus sévèrement la copie plus mauvaise ? - Pensez-vous que vous êtes influencé par le regard de vos collègues sur votre façon d'attribuer des notes à vos élèves ? <ul style="list-style-type: none"> o N'êtes-vous pas tenté de modifier votre notation, si vous avez une majorité de très bonnes notes ou de très mauvaises notes ?
<p>Questions pour les enseignants plus expérimentés :</p> <ul style="list-style-type: none"> - Par rapport à votre pratique, est-ce que votre vision du phénomène des biais de l'évaluation a changé par rapport à celle que vous aviez à vos débuts ? - Pensez-vous que vous avez plus conscience des biais de l'évaluation actuellement, ou que vous en avez toujours eu conscience ?
<p>Gestion des biais de l'évaluation</p>
<p>Pensez-vous qu'il soit possible de gérer ces facteurs ou biais ?</p> <ul style="list-style-type: none"> - Si oui, de quelle manière ? - Si non, pour quel(s) motif(s) ?
<p>Par rapport aux divers biais mentionnés tout à l'heure, comment gérez-vous chacun de ces facteurs ?</p>
<p>Est-ce important, selon vous, de porter une attention particulière à ces biais, ou faut-il accepter que ces phénomènes font partie des aléas du métier ?</p> <ul style="list-style-type: none"> - Pour quelle raison pensez-vous ceci ?
<p>Quel(s) conseil(s) donneriez-vous, par rapport à la gestion de ces types de biais, à un jeune enseignant qui vient de commencer à enseigner et donc à évaluer ses élèves ?</p> <ul style="list-style-type: none"> - Pour permettre à un jeune collègue de tendre vers l'objectivité quels conseils donneriez-vous ?
<p>Questions aux enseignants plus expérimentés :</p> <ul style="list-style-type: none"> - Par rapport à votre évolution dans le métier depuis vos débuts, pensez-vous que vous gérer les biais de l'évaluation de manière différente ?

3. Attestation d'authenticité

*Je certifie que ce mémoire constitue un travail original et j'affirme en être l'auteur.
Je certifie avoir respecté le code éthique et la déontologie de la recherche en le réalisant.*

Lieu, date : St-Maurice, le 21 février 2011

Signature : Alizée Buthey