

SOMMAIRE

| | |
|---|-----|
| Remerciements | ii |
| Avant-propos | iii |
| Liste des figures | iv |
| Liste des annexes..... | iv |
| Introduction..... | 1 |
| Chapitre 1. Problématique..... | 4 |
| 1.1 <i>Définition et importance de l'objet de recherche</i> | 4 |
| 1.1.1 Raison d'être de l'étude | 4 |
| 1.1.2 Présentation du problème..... | 4 |
| 1.1.3 Intérêt de l'objet de recherche | 5 |
| 1.2 <i>État de la question</i> | 6 |
| 1.2.1 Bref historique..... | 6 |
| 1.2.2 Champs théoriques et concepts..... | 7 |
| 1.2.3 Résultats de recherches, théories et synthèses | 15 |
| 1.3 <i>Question de recherche et objectifs de recherche</i> | 16 |
| 1.3.1 Identification de la question de recherche | 16 |
| 1.3.2 Objectifs de recherche | 16 |
| Chapitre 2. Méthodologie | 17 |
| 2.1 <i>Fondements méthodologiques</i> | 17 |
| 2.1.1 Type de recherche | 17 |
| 2.1.2 Type d'approche..... | 18 |
| 2.1.3 Type de démarche | 19 |
| 2.2 <i>Nature du corpus</i> | 20 |
| 2.2.1 Récolte des données..... | 20 |
| 2.2.2 Procédure et protocole de recherche | 21 |
| 2.2.3 Échantillonnage | 22 |
| 2.3 <i>Méthodes et/ou techniques d'analyse des données</i> | 23 |
| 2.3.1 Transcription..... | 23 |
| 2.3.2 Traitement des données..... | 24 |
| 2.3.3 Méthodes et analyse | 24 |

| | |
|---|----|
| Chapitre 3. Analyse et interprétation des résultats..... | 25 |
| 3.1 <i>Compte-rendu des pratiques par entretien</i> | 25 |
| 3.1.1 Enregistrement 1 - Josette | 25 |
| 3.1.2 Enregistrement 2 – Bernard | 27 |
| 3.1.3 Enregistrement 3 – Cosette | 28 |
| 3.1.4 Enregistrement 4 – Élisabeth..... | 29 |
| 3.2 <i>Comparaison des résultats en lien avec les axes de recherche</i> | 30 |
| 3.2.1 Les pratiques effectives..... | 30 |
| 3.2.2 Les biais évaluatifs/arrangements..... | 39 |
| Conclusion..... | 50 |
| Références bibliographiques | 53 |
| Annexes : | I |

Introduction

Motivations

Ce travail s'inscrit dans le champ des sciences de l'éducation et traite de la thématique de l'évaluation. La raison d'être de cette étude provient de motivations professionnelles provenant de la théorie (cours suivis à la HEP-BEJUNE) et de la pratique (stages et remplacements). En effet, nous avons pu constater que le processus évaluatif est une tâche complexe à réaliser, d'autant plus que la plupart des enseignants y sont quotidiennement confrontés. Plusieurs éléments nous ont permis d'arriver à ce constat. Tout a commencé lors d'une série de cours théoriques traitant de l'évaluation. Nous avons constaté que l'acte évaluatif peut être biaisé par des éléments externes à l'évaluation et peut engendrer des conséquences néfastes pour les élèves. Nous nous sommes alors intéressée au caractère objectif et subjectif de l'évaluation. Ensuite, nous avons eu l'opportunité de réaliser nous-mêmes une évaluation de mathématiques lors d'un stage. Nous avons alors pu exercer l'entièreté du processus d'évaluation, de la conception jusqu'à la notation, en passant la correction. À ce moment précis, nous avons constaté que le processus évaluatif, et particulièrement la phase de correction, n'est pas aisé. Ayant connaissance du caractère subjectif que peut porter une évaluation, nous voulions être le plus juste et équitable possible dans nos décisions. En cours de correction, nous nous sommes posé une multitude de questions comme l'attribution des points, l'élaboration du barème ou encore la pondération des exercices. Il est donc primordial d'être au clair avec les critères d'évaluation ainsi que d'être à l'aise avec les décisions prises. Après cette expérience, nous avons profité de la pratique d'enseignants en les questionnant sur le sujet durant nos stages. Ils ont, sans aucune exception, souligné certaines difficultés qu'ils ont rencontrées lors de l'acte correctif. Ces conversations nous ont également fait prendre conscience que les pratiques évaluatives peuvent être très différentes d'un enseignant à l'autre. C'est ainsi que nous sommes intéressée à leurs pratiques lors de l'acte correctif. Le sujet de l'évaluation au second cycle nous interpelle particulièrement, étant donné que nous nous destinons à enseigner dans ces degrés-là. Nos recherches vont nous être bénéfiques pour notre future pratique professionnelle, puisqu'elles nous permettront d'élargir nos connaissances sur le

sujet, c'est-à-dire d'obtenir un panel élargi des pratiques évaluatives. Cet apport nous permettra donc de sélectionner et tirer profit des pratiques qui nous semblent les plus adéquates.

Contexte du champ d'étude

L'évaluation est une tâche des plus importantes, compte tenu des enjeux qu'elle véhicule. La réussite ou l'échec de l'élève est basé presque essentiellement sur sa mise en œuvre. La note conditionne donc la réussite ou l'échec scolaire de l'élève. Dans la plupart des cas, les apprentissages ne sont pas visibles aux yeux de la société. Par contre, ce qu'elle voit, c'est la note que l'élève rapporte à la maison. Il est alors nécessaire pour tout enseignant d'utiliser l'évaluation avec précaution. En effet, elle peut avoir un impact conséquent sur l'estime de soi de l'élève, ainsi que sur son cursus scolaire.

Question de départ et sujet précis

À partir de notre question de départ "**Comment les enseignants évaluent-ils?**", nous avons rapidement focalisé notre recherche sur un acte précis de l'évaluation : **la correction**. Nous avons ensuite recentré notre question pour nous intéresser plus précisément au canton du Jura et aux degrés sept et huit HarmoS. Comme mentionné précédemment, l'acte évaluatif est complexe et les pratiques évaluatives peuvent varier d'un enseignant à l'autre. Nous nous demandons alors comment les enseignants corrigent effectivement les productions de leurs élèves. De plus, nous nous intéressons aux facteurs qui peuvent influencer le jugement et les prises de décisions des enseignants dans l'acte correctif.

Plan de travail du mémoire

Notre travail de Bachelor est divisé en trois parties principales : la problématique, la méthodologie et l'analyse/l'interprétation des résultats.

Dans la problématique, nous commençons par présenter l'importance de l'objet de recherche, ainsi que d'exposer le problème à l'origine de la réflexion. Ensuite, nous exposons un bref historique de l'évaluation, en mettant l'accent sur l'évolution de la docimologie. Ensuite, différents concepts liés à notre recherche sont définis (évaluation, jugement professionnel, objectivité/subjectivité, biais évaluatifs et

arrangements). Puis, nous exposons différents résultats et synthèses menés par des chercheurs renommés. Dans la dernière partie de la problématique, nous définissons la question de départ, les questions et les objectifs de recherche.

La méthodologie précise la méthode, l'approche et la démarche de recherche retenues pour réaliser notre récolte de données. Les différents points de la récolte de données et de la méthode d'analyse, tels que l'outil, l'échantillonnage, les règles de transcription, la manière de traiter et analyser les données, sont explicités en détail.

En ce qui concerne l'analyse et l'interprétation des résultats, deux parties séparent ce chapitre. La première expose un compte-rendu chronologique des faits (actes correctifs des enseignants) par entretien. Dans la seconde partie, les pratiques effectives des enseignants sont comparées entre elles. Puis, plusieurs thématiques liées aux biais évaluatifs et aux arrangements sont traitées.

Pour terminer, nous présentons de manière synthétique les principaux résultats en lien avec la question de recherche. Puis, nous effectuons une autoévaluation critique de la démarche en évoquant les limites ou les difficultés rencontrées, ainsi que les apports personnels et professionnels d'un tel travail. Finalement, nous suggérons une proposition d'élargissement de la réflexion.

Chapitre 1. Problématique

1.1 Définition et importance de l'objet de recherche

1.1.1 Raison d'être de l'étude

Après avoir suivi les cours sur l'évaluation et la régulation des apprentissages en sciences de l'éducation à la HEP-BEJUNE de Porrentruy, nous nous sommes intéressée à nous diriger vers un sujet de mémoire professionnel relatif à cette thématique. En effet, nous nous sommes passablement questionnée sur sa mise en application : est-ce que les notes sont vraiment nécessaires? Est-il possible de proposer une alternative au système de notation? Est-ce que les notes donnent-elles réellement du sens aux apprentissages? Comment refléter véritablement les connaissances et les compétences des élèves à travers l'évaluation? Comment rendre les outils de l'évaluation valides, fidèles et justes? Comment rendre l'évaluation moins subjective? Nous avons immédiatement remarqué que l'évaluation, et plus particulièrement le système de notation scolaire et la correction, est une thématique très complexe dans le monde de l'enseignement.

Lors d'un stage en 8^e HarmoS, nous avons eu l'opportunité de créer et corriger une évaluation de mathématiques. Pendant la correction, il nous a parfois été difficile d'attribuer tout ou partie des points dans les exercices et surtout de pouvoir justifier nos décisions. C'est à ce moment-là que nous avons décidé d'étudier ce qu'il se passe dans l'acte évaluatif. En d'autres termes, nous nous intéressons à la manière dont les enseignants corrigent, c'est-à-dire interprètent des données et prennent des décisions par rapport aux productions de leurs élèves.

1.1.2 Présentation du problème

Lorsque l'enseignant établit une évaluation sommative, il recueille des données, observe des informations appropriées aux objectifs d'apprentissage. Finalement, il interprète ces données et informations et prend certaines décisions (attribution de point, élaboration d'un barème, pondération des exercices, etc.). Dans cette dernière partie du processus évaluatif, l'enseignant va porter un jugement sur les productions de ses élèves, c'est-à-dire juger le degré d'atteinte des objectifs d'apprentissage de

ses élèves. Mais comment s'y prend-il? Quels facteurs peuvent influencer ses choix? Sur quoi se base-t-il pour les justifier? L'enseignant doit alors développer une compétence à bien savoir évaluer pour être le plus juste et le plus équitable possible. Cependant, plusieurs facteurs, appelés biais évaluatifs ou arrangements, interviennent dans le processus évaluatif et influencent plus ou moins fortement le jugement de l'enseignant dans l'interprétation et la prise de décision des productions des élèves. Il est alors primordial pour tout enseignant d'en prendre conscience et d'adopter des mesures et des stratégies pour en diminuer les effets.

1.1.3 Intérêt de l'objet de recherche

L'évaluation scolaire est un acte pédagogique complexe, quotidien dans la vie professionnelle de tous les enseignants et fait partie de la réalité scolaire actuelle. En effet, les enseignants utilisent continuellement divers moyens pour récolter des informations qui leur serviront à évaluer leurs élèves. Comme le mentionnent Mottier Lopez et Allal (2010), ces différentes démarches utilisées peuvent être sous forme de contrôles écrits, de productions ou traces d'activités d'apprentissage large (recherche, résolutions de problèmes, etc.), d'un recensement de travaux quotidiens (exercices, fiches, devoirs, etc.) ou encore d'observations et informations provenant d'interactions avec les élèves et leurs parents. Dans ce mémoire professionnel, nous allons plus particulièrement nous intéresser au contrôle écrit (évaluation sommative), où la justification auprès des parents est plus sensible, puisqu'elle est certificative.

L'évaluation est une thématique complexe, nous nous en sommes rendu rapidement compte lors de la correction de productions d'élèves durant un stage. De la définition du barème jusqu'à la mise de la note finale, en passant par l'attribution des points, il nous a parfois été difficile de prendre des décisions et de porter un jugement qui soit le plus juste et équitable possible. En effet, le résultat chiffré d'une évaluation sommative peut avoir des conséquences relativement sérieuses sur les élèves, qu'il soit bon ou mauvais. Dans le cas d'un résultat relativement mauvais, ou inférieur à ce que l'élève espérait, celui-ci peut alors manifester toutes sortes d'émotions et de sentiments, comme la tristesse, le dégoût, le mécontentement, la honte, le désarroi, l'injustice ou encore la colère, qui peuvent avoir un réel effet négatif sur l'estime de soi. C'est pourquoi les enseignants se doivent de développer une compétence à bien

savoir évaluer. Il doit alors interpréter le plus objectivement possible les données et prendre des décisions qui reflètent au plus près les compétences et connaissances de leurs élèves.

Ainsi, en étant nous-mêmes confrontée à cette problématique et afin d'améliorer notre pratique professionnelle, nous aimerions approfondir nos connaissances sur la thématique des pratiques évaluatives, en découvrant l'acte évaluatif de manière pragmatique. De plus, nous souhaiterions également y soulever certains aspects problématiques, notamment celles des biais évaluatifs et des arrangements qu'utilisent les enseignants lors des corrections des évaluations des élèves.

1.2 État de la question

1.2.1 Bref historique

À partir du XVIIe siècle, l'évaluation scolaire, déjà appelée "appréciation", "notation" ou "contrôle", a vu le jour dans les écoles. Elle s'est développée massivement au XIXe siècle avec le début de la scolarité obligatoire (Perrenoud, 1989).

Après sa généralisation, la question de la notation scolaire fait l'objet de débats, notamment au sujet de la subjectivité des notes. C'est en 1922 que le psychologue Henri Piéron crée une nouvelle science : la docimologie (du grec *dokime*, qui signifie "épreuve"). Elle désigne l'étude des examens, des pratiques de notation, de l'influence des facteurs subjectifs et des examinateurs. Piéron (1951) lui-même la définit ainsi : "étude systématique des examens : modes de notation, variabilité interindividuelle et intra-individuelle des examinateurs, facteurs subjectifs, etc.)".

Plus tard, Barbier (1983), définit deux démarches de nature différente, mais dont le développement est étroitement lié : une démarche de lecture du fonctionnement des pratiques d'évaluation : c'est la docimologie "critique" et une démarche d'amélioration du fonctionnement de ces pratiques : c'est la docimologie "prescriptive".

La première révèle le caractère négatif du système de notation. Henri Piéron et d'autres précurseurs la développent dès les années 1920. Toujours selon Barbier (1983), leurs travaux présentent une caractéristique commune : ils ne remettent pas en cause l'existence de la notation, mais montrent son inadéquation et ses limites.

En constatant une variabilité ou une imprécision des jugements de valeur dans la notation, ils s'attaquent aux problèmes causés par la subjectivité du système de notation.

Progressivement, la docimologie est entrée dans une phase constructive (deuxième démarche selon Barbier, 1983) : la docimologie prescriptive. Des chercheurs s'attèlent à améliorer et perfectionner le système de notation et les actes évaluatifs. Giraudeau (2014) soulève deux directions : l'amélioration de la notation (mesure du travail de l'élève) et l'harmonisation des distributions de notes entre différents évaluateurs. Dans ce travail, nous allons nous intéresser à la première direction, qui vise à améliorer la "justesse" de la notation. Elle s'appuie sur trois notions : la fidélité, la validité et la sensibilité, que Giraudeau (2014, pp. 29-30) décline ainsi :

Pour qu'une note soit fidèle, il est nécessaire qu'un ou plusieurs évaluateurs puissent attribuer des valeurs chiffrées strictement identiques à des élèves présentant des types de travaux identiques ou comparables. Pour qu'une note scolaire soit valide, on doit être en mesure de déterminer une valeur absolue qui représente le travail de l'élève, indépendamment de facteurs secondaires ou halogènes. [...] Enfin, la question de la sensibilité nous renvoie à la capacité à distinguer parmi les élèves ceux qui ont progressé de ceux qui ne l'ont pas. [...] Il est incontournable d'élaborer et d'utiliser de façon stricte un barème (fidélité), de mettre en place le plus souvent possible une procédure d'anonymat (validité), d'utiliser des épreuves critériées, permettant la comparaison du travail de l'élève à ce qui est attendu dans son apprentissage et non au reste de la classe (sensibilité).

1.2.2 Champs théoriques et concepts

Évaluation

L'étymologie du terme "évaluation" laisse déjà apparaître des contradictions. En effet, selon Huver et Springer (2011), en ancien français, *avaluer* signifie simultanément "faire une évaluation" et "fixer la valeur de". Il signifie également

"déterminer approximativement par une appréciation la valeur de quelque chose" ainsi que "fixer le prix, la valeur de quelque chose". Cette signification porte un double sens, puisqu'évaluer est à la fois une estimation subjective et approximative ainsi qu'une mesure objective et précise.

De Landsheere (1979, p. 111) suppose que l'évaluation est "une estimation par une note, d'une modalité ou d'un critère considéré dans un comportement ou un produit". Il indique qu'elle est une procédure qui permet de définir l'atteinte ou non des objectifs fixés au préalable ou non. De plus, il suggère que l'appréciation peut contenir un caractère plus ou moins subjectif.

Selon Bonniol et Vial (2009, p. 48), "l'évaluation appartient à la catégorie des objets de recherche relevant des processus de jugement, elle est alors inscrite dans un champ théorique où le modèle de la décision occupe une place privilégiée".

De Ketele (2010) propose une définition complète du processus d'évaluation, exprimée de manière opérationnelle. Le processus évaluatif permet à l'enseignant de recueillir un ensemble d'informations pertinentes dans le choix, valides pour l'exploitation et fiables dans le recueil. Puis, il confronte, par une démarche adéquate, cet ensemble d'informations à un ensemble de critères pertinents dans le choix, valides dans l'opérationnalisation et fiables dans l'utilisation. Ensuite, il attribue une signification aux résultats de cette confrontation (processus d'attribution de sens). Finalement, il fonde une prise de décision cohérente avec la fonction visée par l'évaluation.

Hadji (1989, pp. 21-22) propose un éventail de significations du verbe "évaluer" :

Évaluer peut signifier, entre autres : vérifier, juger, estimer, situer, représenter, déterminer, donner un avis... Vérifier ce qui a été appris, compris, retenu. Vérifier les acquis dans le cadre d'une progression. Juger un travail en fonction des consignes données ; juger du niveau d'un élève par rapport au reste de la classe ; juger selon des normes préétablies. Estimer le niveau de compétences d'un élève. Situer l'élève par rapport à ses possibilités, ou par rapport aux autres ; situer la production de l'élève par rapport au niveau

général. Représenter par un nombre le degré de réussite d'une production scolaire en fonction de critères variant selon les exercices et le niveau de la classe. Déterminer le niveau d'une production. Donner un avis sur les savoirs ou savoir-faire maîtrisés par un individu ; donner un avis qui concerne la valeur d'un travail.

En ce qui concerne les fonctions de l'évaluation, il en existe trois grandes catégories : l'évaluation formative, l'évaluation diagnostique/pronostique et l'évaluation sommative/certificative. La première fonction relève davantage de la régulation des apprentissages alors que les deux autres fonctions relèvent davantage du contrôle des apprentissages.

L'évaluation formative consiste à vérifier si les élèves progressent en fonction d'objectifs d'apprentissage déterminés au préalable. Elle se déroule pendant les apprentissages, durant une séquence d'enseignement. Ainsi, elle informe à la fois l'enseignant et l'élève sur l'atteinte ou non des objectifs et permet donc à l'enseignant de réguler les apprentissages.

Concernant l'évaluation pronostique, elle intervient pour prédire l'aptitude à réaliser des tâches. Elle permet ainsi de sélectionner et/ou d'orienter les élèves en fonction de leurs facultés présumées à suivre tel ou tel cursus. L'évaluation diagnostique, quant à elle, fait le point de la situation de départ et permet de vérifier les prérequis.

Quant à l'évaluation sommative (fonction qui va être traitée plus en détail dans ce travail professionnel), elle vérifie l'atteinte ou non d'objectifs d'apprentissage à la fin d'un apprentissage et revêt donc un caractère de bilan. Elle est caractérisée par l'attribution d'une note ou d'une appréciation. Albrecht (1991), cité dans Talbot (2009), mentionne que l'évaluation sommative permet de vérifier les acquis, en sanctionnant la performance et en rejetant l'erreur, contrairement à l'évaluation formative, où l'erreur est un moyen pour comprendre la cause et la signification, ainsi que de prendre appui sur elle pour améliorer l'enseignement. Selon Talbot (2009), elle permet de dresser un bilan des performances des élèves et de leur situation au terme d'un cursus d'enseignement. Si l'évaluation sommative débouche sur la délivrance d'une reconnaissance institutionnelle, elle devient certificative.

Principales étapes de l'évaluation

Une fois l'évaluation des élèves récoltée, l'enseignant doit encore exploiter les données, c'est-à-dire les interpréter et parvenir à un jugement. Pour y parvenir, l'enseignant doit effectuer trois étapes successives (Chabot J.-M., 2004) : la correction, la pondération et la notation. Pour la première étape, la correction, l'enseignant identifie chaque élément de réponse dans l'explication de l'élève. Il va ensuite porter sur chacun d'eux un jugement binaire "correct ou incorrect". La deuxième étape, la pondération, consiste à affecter une valeur numérique à chaque élément de réponse en fonction de sa pertinence. En ce qui concerne la troisième étape, la notation, l'enseignant attribue à chaque élève une appréciation globale, le plus souvent sous forme de note, sur sa performance. Cette appréciation résulte en principe d'une totalisation simple et arithmétique des points correspondants aux éléments de jugement corrects et incorrects présents dans la réponse de l'élève.

Jugement professionnel

Les notions d'appréciation, de jugement ou encore d'attribution de sens ressortent clairement dans toutes définitions de l'évaluation citées ci-dessus. Lafortune et Allal (2008, p. 27), reprennent la définition du jugement professionnel de Lafortune (2006), qui la décline ainsi :

Le jugement professionnel est un processus qui mène à une prise de décision, laquelle prend en compte différentes considérations issues de son expertise (expérience et formation) professionnelle. Ce processus exige rigueur, cohérence et transparence. En ce sens, il suppose la collecte d'informations à l'aide de différents moyens, la justification du choix des moyens en lien avec les visées ou intentions et le partage des résultats de la démarche dans une perspective de régulation.

Comme le mentionne Lafortune et Allal (2008), dans le domaine de l'enseignement, le jugement professionnel implique une confrontation de sources d'informations variées et une combinaison de ces informations. Ainsi, lorsque l'enseignant corrige, il émet un jugement sur les productions de ses élèves.

Selon Scallon (2004), cité par Durand & Chouinard (2006), le jugement indique la maîtrise totale, la maîtrise partielle ou la non-maîtrise d'une compétence en fonction d'observations recueillies tout au long d'un apprentissage. Le jugement aboutit à une décision qui peut avoir des répercussions déterminantes sur l'avenir des élèves, tant scolaire que professionnel.

Objectivité/subjectivité

Comme le mentionne Jorro (2000), les docimologues posent trois questions qui permettent d'étudier les champs de variabilité d'une évaluation : la question de fidélité, la question de validité et la question de sensibilité de la note.

Pour la fidélité de la note : est-ce que la note reflète-t-elle le travail attendu? Ce critère de fidélité de la note permet d'estimer la variation de la note attribuée à la même copie à plusieurs mois d'intervalle par un même évaluateur. La différence entre les notes mesure donc la fidélité de l'évaluation.

En ce qui concerne la validité de la note : est-ce que la note correspond à ce que l'enseignant cherche à évaluer? Ce critère de validité estime si la note porte sur les objectifs fixés au départ. Par exemple, l'enseignant évalue les objectifs fixés, mais peut également, et parfois de manière inconsciente, évaluer des critères externes, comme l'orthographe.

Et pour la sensibilité de la note : la note ne couvre-t-elle pas des éléments singuliers dans la copie, y a-t-il surestimation?, sous-estimation? Ce critère de sensibilité évalue si l'enseignant estime finement les différences entre deux copies proches.

Durand & Chouinard (2006) dégagent quant à eux deux conditions essentielles pour qu'un instrument de mesure reflète "un portrait représentatif et juste des apprentissages des élèves" : la fidélité et la validité. L'enseignant s'appuie sur ces deux qualités des instruments utilisés lorsqu'il interprète les données des productions des élèves. Pour qu'une évaluation soit fidèle, les résultats ne doivent pas varier en fonction du lieu, du moment et de l'enseignant-correcteur. Si cette condition n'est pas remplie, le but de l'évaluation est alors biaisé et les objectifs faussés. Pour qu'une évaluation soit valide, les données recueillies doivent s'appuyer sur les critères d'évaluation précis et mesurer effectivement les objectifs annoncés. En d'autres

termes, les données récoltées doivent précisément correspondre à ce que l'enseignant veut mesurer.

De multiples chercheurs, scientifiques, enseignants et autres remettent en question les approches de Jorro (2000) et Durand & Chouinard (2006) inhérentes à l'évaluation des élèves. Des arrangements et erreurs de mesure peuvent en découler directement (voir point suivant).

Biais évaluatifs et arrangements

Mottier Lopez (2013), s'appuie sur les études d'Amigues et Zerbato-Poudou (1996), de Leclercq, Nicaise et Demeuse (2004) et de Bressoux et Pansu (2003) pour attirer l'attention sur les multiples facteurs qui influencent la notation. L'intérêt de ces études est d'avoir mis à jour des variations importantes et systématiques dans les pratiques de notation. Ces résultats démontrent un manque d'objectivité et de fiabilité de l'évaluation scolaire, mais également d'équité sociale. Ces facteurs sont appelés biais évaluatifs et arrangements.

Selon Allal (2009), les biais évaluatifs sont des erreurs de mesure aléatoires. Ce sont des variables qui interviennent dans le processus d'évaluation et qui influencent plus ou moins fortement l'évaluateur et donc la manière dont il va analyser les données. Ils peuvent être liés à des facteurs individuels (caractéristiques de l'élève comme son sexe ou encore le niveau socioéconomique de sa famille) et/ou à des facteurs contextuels (niveau général de la classe, caractéristiques de l'école, composition socio-économique du quartier, etc.).

Leclercq, Nicaise et Demeuse (2004) identifient quelques exemples de biais mis en évidence par la docimologie critique. Tout d'abord, la loi de Posthumus indique que la distribution des notes résulte d'une sorte de modèle type, communément admis. De Landsheere (1980, p 224) la définit ainsi :

Formulée dès 1947, la loi de Posthumus peut s'exprimer de la façon suivante : "un enseignant tend à ajuster le niveau de son enseignement et ses appréciations des performances des élèves de

façon à conserver d'année en année, approximativement la même distribution (gaussienne) de notes.

L'évaluation ne dépend donc pas des performances particulières de l'élève dans l'absolu, mais de ses performances par rapport à celles de ces camarades. Ainsi, un "bon élève" dans une classe pourrait devenir un "mauvais élève" s'il était avec d'autres camarades, "meilleurs" que les élèves de la première classe. C'est ce qu'on appelle également la courbe de Gauss, qui se réfère à une norme (voir schéma à la page suivante), caractérisée par une minorité de faibles et d'excellents niveaux, et une majorité de niveaux moyens.

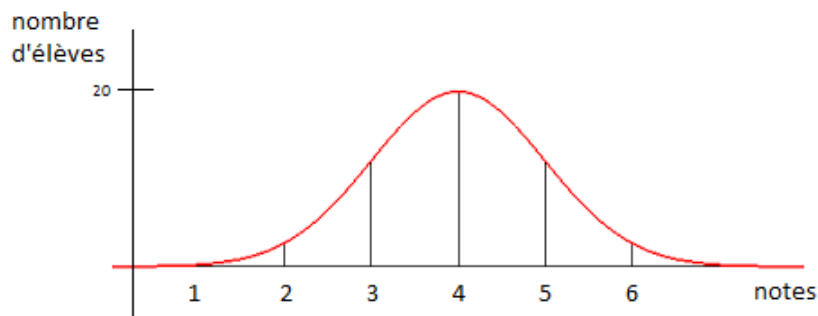


Figure 1 : Courbe de Gauss

Ensuite, ces chercheurs constatent d'autres biais évaluatifs : l'effet de stéréotypie (ou d'inertie), l'effet de halo, l'effet de contraste, l'instabilité du correcteur et les différences entre correcteurs.

Pour l'effet de stéréotypie, l'enseignant normalise la notation établie. Selon Crahay (2007), quand bien même un élève montre une augmentation ou une diminution de ses performances, l'enseignant aura tendance à lui attribuer une note semblable aux précédentes. Il connaît suffisamment son élève pour avoir une idée de ses compétences actuelles et va donc lui octroyer des notes relativement stables.

L'effet de halo, quant à lui, se produit lorsque l'appréciation est influencée par des caractéristiques inhérentes à l'élève, comme son aspect physique, sa tenue vestimentaire, son milieu social, etc. Selon De Landsheere (1992, p. 49), "l'effet de halo présente un caractère affectif accusé. Souvent, on surestime les réponses d'un élève de belle allure, au regard franc, à la diction agréable... Soit pour des raisons de lisibilité, soit pour des raisons nettement affectives, l'écriture pourrait aussi influencer le correcteur". L'enseignant va donc être attiré et concentrer son attention sur des éléments extérieurs et pleinement superflus d'une copie.

Concernant l'effet de contraste, lorsque l'enseignant corrige une série d'épreuves, son appréciation peut dépendre des précédentes. En effet, la copie qui suit une excellente copie va être désavantagée et inversement. Jorro (2000, p. 67) apporte une définition légèrement différente de l'effet de contraste ainsi : "l'évaluateur procède à un jugement comparatif lorsqu'il rencontre une copie qu'il considère comme représentative, celle-ci joue un rôle d'ancre c'est-à-dire qu'elle place l'évaluateur dans un processus de différenciation positive ou négative". Dans tous les cas, lorsqu'il y a effet de contraste, il y a comparaison des copies lors de la correction.

L'instabilité du correcteur provient de variations internes, comme la fatigue, distraction, saute d'humeur, ennui ou encore hasard. En effet, un même correcteur peut, à des moments différents, donner des appréciations différentes à une même copie.

Concernant les différences entre correcteurs, on remarque qu'en faisant corriger la même copie par plusieurs correcteurs, de fortes variations d'appréciation peuvent apparaître.

Trois autres biais d'évaluation sont ajoutés à cette liste : l'effet de contamination, l'effet Pygmalion et l'effet de flou.

Pour l'effet de contamination, l'avis des collègues ou la connaissance des antécédents de l'élève a un rôle influent sur le jugement de l'enseignant. Par

exemple, une production "moyenne" qui a été présentée comme venant d'un élève faible puis venant d'un bon élève. Dans le premier cas, elle a été sous-évaluée, et surévaluée dans le second cas.

Concernant l'effet Pygmalion, Rosenthal & Jacobson (1968) émettent l'hypothèse que la perception que l'on a d'autrui peut induire des attentes plus ou moins positive envers lui. Les attentes d'un enseignant par rapport à ses élèves peuvent donc avoir une influence positive ou négative sur son appréciation. Par exemple, si l'attente envers un élève est positive, l'enseignant doit être aveuglément convaincu que son élève possède les capacités nécessaires à la réussite.

En ce qui concerne l'effet de flou, les objectifs poursuivis et les critères de notation ne sont pas toujours définis avec précision. Cette tendance peut ainsi laisser une grande marge de manœuvre à l'enseignant lorsqu'il corrige.

Ces biais évaluatifs cités ci-dessus sont des arrangements évaluatifs. Selon Merle (2012), cité par Mottier & al. (2012), les arrangements évaluatifs sont des "compromis que l'enseignant effectue plus ou moins consciemment par rapport à des contraintes interpersonnelles et/ou institutionnelles, notamment pendant le processus d'interprétation et de prise de décision en cours de correction des productions des élèves. Ces arrangements peuvent être vus comme fonctionnels, mais ils sont rarement assumés par les enseignants qui semblent les accomplir à défaut de ne pas pouvoir/savoir faire autrement".

1.2.3 Résultats de recherches, théories et synthèses

Mottier Lopez & Allal (2008) ont observé les caractéristiques du jugement professionnel en évaluation par les enseignants. Leurs recherches portent sur les démarches d'évaluations sommatives et pronostiques et impliquent des enseignants genevois de huitième année HarmoS. Les résultats de cette recherche mettent en avant les procédures utilisées par ces enseignants pour établir une note dans le livret scolaire et les ajustements évaluatifs. L'analyse des auteures montrent que "les décisions d'ajustements prennent sens par rapport au cursus de l'élève, à sa progression en cours d'année, à sa participation en classe, à sa situation familiale,

mais également au regard des exigences et attentes, y compris implicites, du système scolaire" (p. 476). De plus, elle montre que "des significations sont attribuées aux notes, en plus de leur définition en terme de degré d'atteinte des objectifs" (p. 476), qui jouent un rôle essentiel en cas d'incertitude (hésitation entre deux notes). Dans leurs recherches, les enseignants ont exposé des valeurs, que la profession tend à valoriser, pour justifier les procédures d'évaluation et les ajustements évaluatifs.

1.3 Question de recherche et objectifs de recherche

1.3.1 Identification de la question de recherche

Notre question de départ est la suivante : comment les enseignants évaluent-ils? Cette question nous amène à la question de recherche : comment les enseignants corrigent les évaluations de leurs élèves et quels facteurs peuvent influencer leur jugement dans l'acte correctif?

Nous avons l'intention de découvrir ce qu'il se passe effectivement dans l'acte correctif pour ainsi dégager l'impact des choix des enseignants sur la correction des productions des élèves dans une évaluation sommative.

1.3.2 Objectifs de recherche

L'objectif principal de recherche est d'identifier la pratique d'enseignants lorsqu'ils corrigent les productions de leurs élèves et de repérer les facteurs qui peuvent influencer cette correction. Il est alors nécessaire de comprendre comment les enseignants justifient leurs choix quant à l'attribution de points, notes, barèmes, etc. La récolte de ces données nous permettra de réaliser un inventaire et une comparaison de plusieurs pratiques lors de la correction et de dégager des facteurs qui peuvent influencer plus ou moins fortement le processus intégral de l'acte correctif.

Chapitre 2. Méthodologie

2.1 Fondements méthodologiques

Dans ce chapitre, nous allons présenter nos choix et réflexions effectués quant aux méthodes et techniques de récolte de données. Pour notre mémoire professionnel, la méthode de récolte de données choisie est la méthode qualitative, le type d'approche est l'approche inductive et le type de démarche est la démarche compréhensive. Ces différents termes sont décrits ci-dessous.

2.1.1 Type de recherche

Pour un travail de recherche, il existe deux types de recherche : elle est soit quantitative, soit qualitative. Selon Giroux (1998), la recherche quantitative met l'accent sur la collecte et l'analyse de données chiffrées. Ces données peuvent donc être converties en chiffres, en données mesurables. Pour ce faire, les outils utilisés pour cette méthode de recherche sont généralement des entretiens directifs, des grilles précises ou encore des questionnaires.

Selon Giroux (1998), la recherche qualitative met, quant à elle, l'accent sur la compréhension des données. Elle regroupe une multitude de techniques d'investigation. Ces techniques permettent d'obtenir un aperçu du comportement et des perceptions d'une population donnée. Cette méthode permet de traiter le sens et l'observation d'un phénomène social en milieu naturel. Les outils utilisés sont principalement des entretiens semi ou non directifs, des grilles ouvertes ou encore des entretiens de groupe.

Le but premier de la recherche qualitative, selon Poisson (1991, p. 12), est de "comprendre des phénomènes sociaux, soit des groupes d'individus ou des situations sociales", donc de comprendre certains principes du comportement humain et les raisons qui régissent ce type de comportement. L'investigation peut être alors faite sous forme d'étude de cas. Cette méthode, au lieu d'effectuer des recherches sur un grand panel de personnes, procède à un nombre restreint de cas considérés comme significatifs en fonction d'un objectif précis et restreint. Cette

manière d'opérer a pour but, selon Poisson (1991, p. 12), "de comprendre en profondeur d'une situation sociale, d'un groupe de personnes, d'un individu, d'une institution, d'un phénomène historique ou de toute autre réalité sociale". Cependant, puisque l'étude de cas se limite à un nombre limité de cas, les données récoltées ne sont pas généralisables et n'ont pas pour objectif d'aboutir à des certitudes. Avec l'étude de cas, les éléments majeurs d'un phénomène particulier sont décrits et interprétés par les chercheurs, dans le but d'en faciliter la compréhension. Elle dépend donc non seulement des auteurs de la recherche, mais également de la population choisie qui participe à la récolte de données, ainsi que de l'objet d'étude.

Le chercheur s'intéresse donc à connaître et comprendre les facteurs qui influencent un certain aspect du comportement de l'individu dans son milieu naturel. C'est un modèle interprétatif qui recueille des processus qui se développent au sein d'une population.

Selon Fortin (2010), elle sert à comprendre les perceptions et les sentiments des personnes. De plus, elle peut être utilisée pour comprendre un domaine encore inconnu ou encore des éléments où la recherche antérieure pourrait être insuffisante ou pas applicable à une situation particulière.

Nos recherches, qui se focalisent sur la verbalisation et la précision des pensées d'enseignants, ainsi que sur les ressentis et les émotions de ces derniers lors de l'acte correctif, visent donc à comprendre leur manière d'aborder la correction. Elles ne cherchent pas à dégager un pourcentage ou des données chiffrées. Le type de recherche choisie pour ce travail est donc qualitatif.

2.1.2 Type d'approche

Il existe deux types d'approches pour une recherche : l'approche déductive et l'approche inductive. L'approche déductive passe du général au particulier. Le chercheur utilisant cette approche formulerait des hypothèses basées sur des éléments théoriques, des concepts, des principes et investiguerait sur le terrain pour

valider ses propos. Il se base donc sur des affirmations générales d'un échantillon représentatif de la population pour en dégager des faits particuliers.

Au contraire, l'approche inductive part du particulier pour atteindre le général. En d'autres termes, elle se base sur des observations générales et limitées qui vont mener à des hypothèses, des théories ou modèles scientifiques. Cette approche s'appuie sur un ensemble de faits pour en tirer une affirmation générale. En effet, selon Blais et Martineau (2006, pp. 3-4), la définition de l'induction est la suivante : il s'agit d'un "type de raisonnement qui consiste à passer du spécifique vers le général; cela signifie qu'à partir de faits rapportés ou observés (expériences, événements, etc.), le chercheur aboutit à une idée par généralisation et non par vérification à partir d'un cadre théorique préétabli".

Puisque notre travail ne s'appuie pas sur des hypothèses, nous allons utiliser l'approche inductive. Nous allons donc aboutir à une idée globale en utilisant la généralisation tout en sachant que les résultats ne seront pas exhaustifs, mais s'appliquent uniquement à ce travail de Bachelor, dans un contexte donné.

2.1.3 Type de démarche

Plusieurs types de démarche sont utilisés en recherche. Une démarche peut être compréhensive, explicative ou descriptive. La démarche descriptive est surtout associée à la recherche quantitative. Le but principal de cette étude est de définir les caractéristiques d'une population ou d'un phénomène. La démarche explicative se base sur des hypothèses qu'elle tente de prouver. Son objectif est de tester, d'élaborer ou d'enrichir une théorie. La démarche compréhensive, quant à elle, cherche, comme son nom l'indique, à comprendre un groupe d'individus, ses croyances, sa manière de vivre, de fonctionner et de s'adapter. Elle vise à observer sur le terrain et à comprendre un groupe humain, ses croyances, sa façon de vivre et de s'adapter.

Dans le cadre de notre travail de mémoire, le type de démarche retenu est la démarche compréhensive. En effet, nos recherches visent à comprendre la manière dont les enseignants corrigent les évaluations de leurs élèves.

2.2 Nature du corpus

2.2.1 Récolte des données

En ce qui concerne la récolte des données, la technique choisie se rapproche de l'auto-explicitation selon Vermersch (2014, 2007).

Pour Vermersch (2014), la technique de l'auto-explicitation est une technique où le chercheur est au centre de l'étude et explicite son propre vécu par écrit. Cette technique consiste à adresser à soi-même des demandes. Par exemple, comme le mentionne Vermersch (2007), il s'agit de poser des questions à soi-même, de se faire de vraies demandes, de se parler intérieurement ou à voix haute pour s'exprimer une requête, avec l'intention ferme d'attendre de soi une réponse à la question que l'on s'est adressée. La technique que nous avons décidé d'utiliser se rapproche de ce modèle, mais diffère tout de même sur quelques points. En effet, c'est un sujet externe (et non pas le chercheur) qui va lui-même auto-expliciter son vécu, verbaliser et préciser sa manière de procéder, ses pensées, justifier ses choix et ses décisions, ou encore mettre en mots ses ressentis, ses émotions.

Dans l'auto-explicitation, "l'informateur se trouve seul, mais pour autant toutes les attentions relationnelles qui sont mobilisées par un intervieweur doivent aussi être présentes. Commencer une auto-explication sans avoir vérifié que l'on soit bien d'accord avec soi-même pour la faire, que l'on soit au clair avec le but poursuivi ou avec la décision de ne pas s'en donner immédiatement, renouveler cet accord avec soi-même quand les difficultés surviennent, etc. tout cela doit faire l'objet d'une négociation avec soi-même, au risque de rester coincé si ce n'est pas fait, ou si c'est mal fait" (Vermersch, 2007, p. 19). Il est important d'avoir à l'esprit que cette activité peut susciter des peurs chez l'informateur : peurs de ne pas se souvenir, peur de ne pas savoir écrire, peur d'être seul devant ce travail ou encore peur de ne pas

dépasser une description superficielle et inutile. Ces "attentions relationnelles", dans le cas précis de nos recherches, ont pu être évincées. En effet, un protocole de verbalisation (voir annexe 1, p. I) distribué au préalable a permis aux participants d'être au clair avec la manière de procéder, ainsi qu'avec les objectifs du travail à réaliser.

La méthode choisie peut également s'apparenter à l'entretien d'explicitation, mis au point par Vermersch. L'entretien d'explicitation est une méthode singulière dont le but est de verbaliser la pratique consciente ou non d'une personne. Il peut permettre de mettre en mot une action singulière, une action spécifique, ce qui s'est réellement produit ou une conscientisation des événements contenus dans une action, entre autres. Cette technique permet de faire émerger avec plus de précisions les représentations, les modes de fonctionnement d'une personne, les attitudes face à un sujet. Cependant, dans notre technique particulière, le chercheur ne se retrouve pas en présence du sujet lors de l'explicitation.

2.2.2 Procédure et protocole de recherche

Dans le courant du mois de juin 2015, nous avons établi une liste d'enseignants susceptibles de pouvoir participer à nos recherches. Au final, quatre enseignant(e)s ont été contacté(e)s par e-mail. Dans cet e-mail, nous leur avons présenté un résumé de notre recherche et avons expliqué les conditions de participation, ainsi que la manière dont allaient être traitées les données récoltées. Pour finir, nous leur avons demandé s'ils/elles nous donnaient leur accord quant à leur participation à notre étude. Nous avons reçu des réponses positives très rapidement. De plus, nous nous sommes engagée à respecter les principes du code d'éthique. En septembre, nous avons commencé activement notre récolte de données. En effet, nous avons communiqué par e-mail des informations plus précises quant à la récolte de données, telles que l'avancement du travail, les objectifs de la recherche, ainsi que la méthodologie avec laquelle nous allons récolter les données. Ce mail a été accompagné d'un protocole de verbalisation. Ce protocole, contenant plusieurs points, a pour but de guider les participants dans l'auto-explicitation de leur correction. Les différents points rappellent donc des éléments à ne pas oublier lors

de la verbalisation, comme les éléments "techniques" à adopter lors de la correction (numérotation des copies dans l'ordre de correction, précision de l'exercice qui est en train d'être corrigé, etc.), la manière de procéder à la verbalisation, à la précision des pensées, à la justification des choix et des décisions, ainsi qu'à l'explicitation des émotions et ressentis. Nous avons demandé aux participants de prendre connaissance de ce protocole de verbalisation et de ne pas hésiter à nous contacter en cas d'éventuelles questions ou remarques. Les quatre auto-explicitations se sont déroulées dans le courant des mois de septembre et octobre 2015 lors de la correction des évaluations sommatives respectives des enseignant(e)s. Les enregistrements ont duré entre 45 et 75 minutes et ont été, pour chacun d'entre eux, enregistrés de manière sonore et transcrits intégralement pour en faciliter l'analyse. Les données récoltées ont bien évidemment été rendues anonymes (changement des prénoms des enseignant(e)s, des élèves et du nom de la ville ou du village de l'école).

2.2.3 Échantillonnage

Pour créer l'échantillon de nos recherches, nous avons sélectionné une population de référence, c'est-à-dire des personnes enseignant au primaire, et plus particulièrement au cycle 2, dans l'arc jurassien. Le sexe, l'ancienneté, l'âge de ces derniers ne sont ici pas pris en compte. De ce fait, nous avons profité de certains contacts établis au fil de nos expériences professionnelles, notamment les stages et les remplacements, pour trouver un échantillon et récolter des données.

Comme le mentionne Blanchet et Gotman (1992, p. 48), "définir la population, c'est sélectionner les catégories de personnes que l'on veut interroger et déterminer les acteurs dont on estime qu'ils sont en position de produire des réponses aux questions que l'on se pose".

Au fil des recherches d'échantillonnage, nous avons décidé de réduire le panel aux enseignants primaires de septième et huitième HarmoS. De plus, il nous semblait pertinent de focaliser les évaluations sommatives des élèves sur une branche par souci de pertinence et de comparabilité : les mathématiques. Nous avons donc

demandé à des personnes enseignant les mathématiques de participer à nos recherches.

2.3 Méthodes et/ou techniques d'analyse des données

Cette partie est consacrée à l'explication du traitement du corpus, qui est, selon Blanchet et Gotman (2001), l'ensemble des enregistrements qui ont été par la suite transcrits de manière littérale, ainsi que son analyse. Premièrement, nous allons définir nos choix quant aux règles de transcription des données récoltées. Deuxièmement, nous allons décrire les procédés, opérations de traitement des données. Finalement, nous allons expliquer le type d'analyse choisi pour notre travail de mémoire.

2.3.1 Transcription

Dans un premier temps, les enseignants sélectionnés se sont auto-enregistrés de manière sonore. Puis, pour faciliter la lecture et l'analyse des données, les enregistrements ont été transcrits scrupuleusement et intégralement. Il s'agit de faire l'inventaire et mettre par écrit toutes les informations orales recueillies au cours de la collecte de données, donc de faire passer le contenu verbal recueilli oralement de l'enregistrement sonore à un support écrit. Cette étape permet d'analyser plus facilement et plus en détail les données.

Techniquement, lors de la transcription, nous avons écouté les enregistrements et transcrit en même temps les propos des enseignants. Les pauses ont été signalées par une barre oblique "/". Les mots parasites tels que "ben", "heu", "hum" ont été transcrits fidèlement. Les négations ont été écrites comme nous les entendions (erreur de syntaxe à l'oral). Nous avons utilisé les points de suspension "..." à la fin des phrases non terminées par les enseignant(e)s. Les rires des enseignants, les bâillements, les bruits annexes (par exemple un bruit de sonnette), ou lorsque les dires étaient inaudibles ont également été indiqué entre parenthèses.

2.3.2 Traitement des données

Traiter des données consiste, pour le chercheur, à transformer les données brutes en données élaborées par différentes techniques et méthodes. Après avoir écouté les enregistrements et les avoir transcrits, nous les avons relus un à un pour nous remémorer les différents aspects abordés. Durant cette première lecture, nous avons réalisé une description chronologique des faits concrets pour chaque enseignant. Ce compte-rendu reflète au plus près et le plus fidèlement possible le déroulement d'une correction d'une série d'évaluation pour les quatre participants.

Ensuite, nous avons repris les enregistrements transcrits de manière plus approfondie pour en ressortir les éléments principaux.

2.3.3 Méthodes et analyse

Le but de l'analyse de contenu est d'attribuer une signification aux données récoltées, ce que nous avons utilisé lors de notre travail. Selon Fortin (2010), l'analyse de contenu traite des données de type qualitatif pour découvrir les thèmes notables, éléments importants ou encore les tendances. Pour ce travail de mémoire, nous avons décomposé l'analyse en deux temps. Le premier temps consiste à donner un compte-rendu, une description détaillée par entretien de ce qui se passe effectivement lors de la correction de chaque participant. Les actes sont exposés de manière chronologique. Le deuxième temps consiste à analyser et comparer en détail les différentes manières de corriger, ainsi que les biais évaluatifs et les arrangements qui peuvent intervenir dans la correction.

Pour réaliser la deuxième partie, nous utilisons un tableau comparatif, qui nous permet de comparer et d'analyser certains actes évaluatifs spécifiques.

Chapitre 3. Analyse et interprétation des résultats

Après avoir récolté les quatre enregistrements sonores, nous allons à présent exposer nos données de manière synthétique dans un tableau comparatif (voir annexe 2, pp. II et III). Puis, à l'aide de ce tableau, nous allons pouvoir comparer et analyser différents éléments de la correction. Les propos relatifs sont illustrés par des verbatim extraits des enregistrements sonores que les enseignants ont réalisés. Il convient de préciser que les verbatim des enseignants ont été modifiés pour des raisons de clarté et de lisibilité (ajout de la ponctuation, suppression des pauses, des répétitions, des hésitations). Ce chapitre reprend les différents thèmes de notre problématique et a pour but de répondre à notre question de recherche qui est : comment les enseignants corrigent les évaluations de leurs élèves et quels facteurs peuvent influencer leur jugement dans l'acte correctif?

3.1 *Compte-rendu des pratiques par entretien*

Généralement, nous avons remarqué que les pratiques des enseignant(e)s consulté(e)s diffèrent d'un enseignant à l'autre. Dans les sous-chapitres suivants, nous allons présenter brièvement les participants. Puis, nous décrivons le plus fidèlement possible le déroulement de chaque enregistrement. Les actes décrits sont insérés de manière chronologique, ce qui permet de se rendre compte concrètement comment se déroule la correction.

3.1.1 Enregistrement 1 - Josette

Josette a huit années d'expérience dans l'enseignement primaire. Elle est enseignante de septième HarmoS. Elle occupe son poste à 100 %. Lors de l'enregistrement, elle corrige les évaluations sommatives des thèmes 1 et 2 de mathématiques.

Josette indique que l'évaluation porte sur les thèmes 1 et 2. Les bases du thème 1 sont évaluées de manière approfondie, puisqu'elles ont été vues et terminées en classe. Celles du thème 2 ont seulement été commencées en classe, mais sont également évaluées. Elle signale qu'elle préfère réaliser plusieurs évaluations de ces

deux thèmes durant l'année pour que les élèves n'oublient pas ces notions. De plus, elle informe qu'elle évalue toujours deux thèmes à la fois, mais peut aller jusqu'à quatre thèmes. Ensuite, elle entre dans le vif du sujet : la correction. Elle indique que les points totaux des exercices sont attribués à l'avance. Elle signale qu'il lui arrive de compter des points pour la compréhension de consigne. Avec un corrigé de chaque exercice, elle le compare avec les réponses des élèves. Elle corrige un exercice sur l'évaluation du premier élève, elle met les points, puis elle corrige l'exercice du deuxième élève, et ainsi de suite. Elle corrige donc le même exercice, un élève après l'autre, puis passe au prochain exercice. Au moment de corriger l'exercice 2, elle fait part de son ennui de le corriger, puisque les réponses peuvent différer d'un élève à l'autre. Donc, elle passe à l'exercice 3, par "flemme" (selon ses mots). À la fin de la correction de l'évaluation d'un élève, elle opte pour le comptage des points obtenus pour savoir où l'élève en est. Elle indique également qu'il est possible qu'elle attribue un point bonus dans l'évaluation. Dans le cas où tout est juste et que le point bonus est juste, elle met la note 6+. Dans cette évaluation, il n'y a pas de point bonus. Une fois la correction de toutes les évaluations terminée, elle réalise le barème. Le barème, pour cette évaluation-là, est d'attribuer une note pour chaque total de points (par exemple, 22 et 21 $\frac{1}{2}$ points donnent la note 6, 21 donnent 5 $\frac{3}{4}$, ...). Elle mentionne qu'il peut varier en fonction de l'évaluation. Par exemple, elle utilise le calcul "nombre de points de l'élève x cinq ÷ par le nombre de points totaux + 1" pour les évaluations de mathématiques où il y a peu de points. Après avoir réalisé le barème, elle prend sa calculatrice et compare la note avec la formule "nombre de points de l'élève x cinq ÷ par le nombre points totaux + 1" pour légitimer son barème et pour se sentir rassurée. Une fois la comparaison réalisée, elle regarde si les deux résultats correspondent. Dans cette évaluation, après avoir comparé les résultats, elle remarque qu'ils diffèrent à certains moments. Elle décide alors de garder la ligne de son barème de départ. Une fois la note attribuée, elle l'introduit pour chaque élève dans un programme informatique et l'inscrit également dans le journal de classe. Finalement, elle fait la moyenne de classe.

3.1.2 Enregistrement 2 – Bernard

Bernard a sept années d'expérience dans l'enseignement primaire. Il est principalement enseignant de huitième HarmoS, mais donne également quelques heures de co-enseignement avec des classes de quatrième, cinquième, sixième et septième. Il occupe son poste à 85 %.

Lors de l'enregistrement, il corrige les évaluations sommatives du thème 2 de mathématiques. Bernard indique en premier lieu le thème de l'évaluation et fait un inventaire du matériel qu'il emploie lors de la correction (les copies d'évaluation, un stylo, une calculatrice). Les points pour chaque exercice sont définis avant la correction. Sur ce point précis, il mentionne qu'il collabore avec une enseignante d'un autre village. En effet, ils s'échangent les évaluations, les supervisent mutuellement et font part de remarques quant à l'attribution des points, entre autres. Ce deuxième regard permet aux deux enseignants d'être le plus juste possible. Pour commencer, il prend les deux ou trois évaluations des meilleurs élèves pour vérifier si les calculs sont corrects et pour voir assez vite si l'évaluation est adaptée aux élèves. Il commence alors la correction du premier exercice du meilleur élève, en s'aidant de sa calculatrice. Une fois le premier exercice de la première copie corrigé, il prend la copie suivante et compare avec la première copie, et ainsi de suite. Selon lui, cette méthode lui permet de gagner du temps. Dans la classe de Bernard, il y a deux élèves à besoins particuliers. Leur évaluation est différenciée par rapport aux autres élèves. Le premier élève n'a pas de diagnostic clairement défini, mais Bernard différencie ces évaluations d'un commun accord avec le conseiller pédagogique. Pour le deuxième élève, qui selon lui est assez spécial, Bernard indique qu'il va corriger sa copie totalement séparément et en dernier lieu. Lors de la correction d'un exercice, si ce qu'a réalisé l'élève n'est pas clair, n'a pas de sens ou alors a été bien maîtrisé, il met une annotation sur la copie. Ce commentaire permet de personnaliser la correction et plaît beaucoup aux élèves. Dans le cas où la réponse est fausse, il écrit la bonne réponse en-dessous de celle de l'élève. Une fois toutes les copies corrigées, Bernard compte le total des points par élèves. Puis, il réalise en direct une grille d'évaluation (barème), où chaque total de points correspond à une note. D'un

commun accord avec sa duettiste, les notes sont au demi-point (il passe, par exemple, de 6 à 5½). Les éventuels demi-points dans les exercices sur comptés vers le haut (par exemple, si un élève a 37 ½ points, il lui compte 38 points). Lors de l'attribution des notes par rapport aux points, il regarde tout d'abord le nombre de points maximum et minimum des élèves, pour pouvoir plus ou moins se situer. De plus, pour savoir si son barème est trop sévère ou trop indulgent, il compare son barème avec la formule officielle "nombre de points de l'élève x cinq ÷ par le nombre points totaux + 1". Finalement, il inscrit les notes dans son journal de classe.

3.1.3 Enregistrement 3 – Cosette

Cosette a onze années d'expérience dans l'enseignement primaire. Elle est enseignante de huitième HarmoS. Elle occupe son poste à 100 %. Lors de l'enregistrement, elle corrige les évaluations sommatives du thème 2 de mathématiques.

Lors de l'enregistrement, elle corrige les évaluations sommatives du thème 2 de mathématiques. Cosette mentionne tout d'abord qu'elle sélectionne les épreuves des élèves qui ont un régime différencié (trois élèves). Pour ces trois élèves, elle signale que leur barème est un peu plus sévère que celui des autres élèves, puisqu'il y a moins de chose à réaliser. Elle corrige alors le premier côté de page (recto) pour chacun de ces élèves. Une fois le recto des évaluations différenciées corrigées, elle passe aux autres copies. Elle les prend comme elles sont empilées (pas de tri). Comme pour les trois premiers élèves, elle corrige chaque recto et passe à un autre élève. Elle passe ensuite au verso de chaque copie. D'une manière générale, lorsqu'elle corrige, elle entoure ce qui lui semble étrange et biffe ce qui est faux. Si une preuve est écrite fausse, elle note en-dessous la manière dont elle souhaiterait que ça soit écrit. Il est possible qu'elle écrive des commentaires par rapport à ce que l'élève a réalisé. Si un élève efface sa preuve ou son résultat, il lui arrive qu'elle tienne compte des choses effacées qui sont potentiellement justes. Une fois tous les exercices corrigés pour une copie, elle compte le total de points. Ensuite, elle établit le barème, basé sur la formule "nombre de points de l'élève x cinq ÷ par le nombre points totaux + 1", qui représente donc le pourcentage des points justes et qui donne

une note au dixième. Finalement, elle informe qu'il est probable qu'elle ne compte pas cette évaluation, puisqu'elle n'est pas très satisfaite des résultats obtenus.

3.1.4 Enregistrement 4 – Élisabeth

Élisabeth a dix années d'expérience dans l'enseignement primaire. Elle est enseignante de septième HarmoS. Elle occupe son poste à 50 %. Lors de l'enregistrement, elle corrige les évaluations sommatives du thème 2 de mathématiques.

Élisabeth corrige les évaluations portant sur le thème 2. Pour commencer, elle indique qu'une fois l'évaluation préparée, elle totalise le nombre de points. Elle mentionne qu'elle note toujours le nombre de points obtenus par les élèves et non pas le nombre de fautes. Le barème est toujours réalisé avant de faire le contrôle. Le barème est réalisé avant la correction (ici, la note est basée sur le total de points obtenus). Elle mentionne également qu'il se peut qu'elle corrige le barème en fonction du résultat des élèves. Les éventuels demi-points sont comptés vers le haut (par exemple, si un élève a $32 \frac{1}{2}$ points, elle lui compte 33 points). Elle prend ensuite les évaluations dans l'ordre inverse de l'ordre dans lequel elles ont été rendues. La première copie correspond donc à l'évaluation du dernier élève qui l'a rendue et la dernière correspond à celle du premier élève qui l'a rendue. Elle corrige tous les exercices d'une copie en une seule fois, compte le total de points obtenu, inscrit la note et passe à la copie suivante. Après avoir corrigé la sixième copie, elle se questionne sur le nombre de points accordé à un exercice (12 points au total). Elle indique qu'il y a beaucoup de points à cet exercice par rapport aux autres et que si certains élèves ont 0 point, ils risquent de largement baisser leur note. Elle prend donc la décision de modifier l'attribution des points pour cet exercice-là (6 points au lieu de 12). À ce moment-là, elle redéfinit le barème en fonction des changements et reprend les copies déjà corrigées. Puis, elle reprend la correction. Finalement, elle récapitule le nombre de points et la note finale par élève, puis indique la moyenne de classe.

3.2 Comparaison des résultats en lien avec les axes de recherche

3.2.1 Les pratiques effectives

Dans les sous-chapitres suivants, nous allons dégager des éléments pragmatiques qui ont été observés dans les enregistrements. Ces actes, décrits ci-dessous, ressortent de la verbalisation des enseignants au moment de la correction.

3.2.1.1 Domaine évalué

D'une manière générale, aucun enseignant n'a décliné ou explicité clairement tous les critères d'évaluation, que ce soit dans l'évaluation sommative ou dans l'enregistrement. Cependant, au cours de l'écoute, quelques critères d'évaluation ont été exposés. En résumé, les quelques éléments évalués évoqués sont les suivants : compréhension de consignes et méthode d'application ou développement.

| Critères d'évaluation |
|--|
| <ul style="list-style-type: none">- J'ai compté un point pour ça parce que c'est vrai que les consignes c'est quelque chose que j'évalue (Josette).- Ça m'embête parce que les résultats sont justes, mais je leur ai vraiment dit que ce qui était important c'était de me montrer ce qui se passait dans leur tête. [...] Mais mathématiquement, si je veux être vraiment très rigoureuse, eh bien non c'est faux. [...]. Ça fait vraiment partie de l'objectif, alors c'est faux (Josette).- Génial! Là elle a compris la consigne. C'est bien qu'elle puisse faire un peu des points (Bernard).- Je vais lui mettre 2 points. Je veux quand même, s'il me pose la question, lui dire que j'aimerais bien avoir un petit développement (Bernard).- On va lui mettre aussi un point pour ça. Ça lui fait donc 1 point sur 4 pour la méthode (Cosette).- Par contre, j'observe toujours le développement. Si le développement est juste et la réponse est fausse, à ce moment-là, je mets un demi-point (Élisabeth). |

Josette mentionne également que les fautes d'orthographe ne sont pas prises en compte dans cette évaluation :

Je ne compte pas les fautes d'orthographe dans la mesure où on peut lire.

Pour Cosette, l'orthographe n'est pas non plus prise en compte dans la correction, mais elle indique qu'elle la corrige à titre informatif.

Nous avons également observé que deux enseignants auraient tendance à prendre certaines mesures particulières pour leurs élèves, en prenant en compte des éléments qui ne sont pas directement liés à l'évaluation, comme le mérite, l'encouragement, une aide due à certaines difficultés. Voici quelques exemples :

Mesures particulières

- Ce n'est vraiment pas évident de trouver la limite entre ceux qui méritent quand même au moins 1 point et ceux qui ne le méritent pas (Bernard).
- Bonne idée! Là c'est ce que j'écris comme annotation pour Brice. Mais mauvaise application. Je peux mettre un point ici, il le mérite quand même (Bernard).
- Je vais faire compter pour la simple et bonne raison que Bertrand il mérite 6, parce qu'il bosse comme un diable et puis qu'il maîtrise vraiment son sujet (Bernard).
- Elle le mérite avec ce qu'elle bosse. Ça c'est plus encourageant (Bernard).
- S'ils ont fait quelques calculs corrects, ça me permet aussi de de les encourager en leur donnant un peu quelque chose (Bernard).
- Il peut m'arriver parfois d'être indulgent avec des élèves qui ont des difficultés à s'exprimer entre autres. Là, j'accepte une justification même si elle n'est pas aussi limpide que pour d'autres (Bernard).
- Finalement j'aurais presque envie de lui mettre 6 sur 6 parce qu'elle a trouvé le plus difficile (Cosette).

3.2.1.2 *Choix du type d'exercice*

Les enseignants ont une grande marge de manœuvre lorsqu'ils définissent les types d'exercice qui vont être inclus dans leur évaluation, tout comme l'attribution des points par exercice. Tout d'abord, ils ont la liberté de choisir la part d'exercices à tâche simple et la part d'exercices à tâche complexe (situation-problème). Ce choix peut permettre aux enseignants d'attribuer des points "facilement" à leurs élèves ou encore de vérifier si les bases sont acquises. Voici les propos de deux enseignants à ce sujet :

| Attribuer des points "facilement" |
|---|
| <ul style="list-style-type: none">- Ils font aussi des exercices qui sont relativement faciles aussi pour qu'ils puissent faire quelques points (Bernard).- L'exercice c) il est vraiment bien allé. Bon c'était un peu le but ça, je le savais de toute façon (Bernard). |
| Vérifier si les bases sont acquises |
| <ul style="list-style-type: none">- Ils sont un petit peu tombés dans le piège, mais c'est normal. Il faut aussi un petit peu les tester. On ne peut pas faire que du facile (Bernard).- Ils ont souvent fait la même faute hein, mais c'était fait exprès. J'ai aussi mis un calcul qui ressemblait à la bonne réponse, mais qui ne l'était pas, pour les obliger aussi à calculer et pas juste à se contenter du minimum d'observation (Bernard).- Il faut aussi de temps en temps qu'il y ait quelque chose d'un petit peu plus compliqué quand même, pour voir s'ils ont vraiment tout compris (Cosette). |

D'autre part, certains enseignants font le choix de mettre un certain type d'exercices pour leur faciliter la tâche lorsqu'ils corrigent et attribuent des points. Nous avons constaté que la plupart des enseignants ont plus de facilité à corriger et à attribuer des points pour les exercices à tâche simple que pour les exercices à tâches complexes. En voici quelques exemples :

Faciliter la correction et l'attribution des points

- Là, c'est vraiment noir-blanc. Il n'y a pas de milieu, il n'y a pas de demi-point. C'est un exercice qui est facile, donc c'est l'aligné juste : un point, l'aligné faux : zéro point (Josette).
- Là j'ai fait en sorte qu'il n'y ait pas plusieurs solutions possibles sauf erreur, pour que je ne sois pas embêté dans l'attribution des points. Donc là, 1 point par réponse juste (Bernard).
- J'aime bien mettre quelques exercices comme ça dans les contrôles parce que ça me facilite aussi la tâche (Bernard).
- C'est vraiment compliqué de donner des points à des problèmes (Bernard).

3.2.1.4 Attribution des points

L'attribution des points peut s'avérer une tâche difficile à réaliser. Nous avons pu observer cette difficulté dans certains enregistrements :

| Difficultés d'attribuer les points | |
|---|---|
| - | "Parce que c'est vrai que ce n'est pas toujours évident d'attribuer des points à des exercices pour pas trop influencer le résultat des élèves, que ce soit pour tirer vers le bas ou plutôt tirer vers le haut, mais essayer d'avoir un test le plus représentatif possible de leurs capacités" (Bernard). |
| - | Il faut quand même réfléchir à l'attribution des points et à la difficulté des exercices. On ne peut pas mettre que du dur ou que du facile (Bernard). |
| - | C'est vraiment compliqué de donner des points à des problèmes (Bernard). |
| - | "Ça va être un peu plus compliqué à corriger pour mettre les points, parce que comme il faut biffer des intrus, c'est un peu complexe de calculer" (Cosette). |

Ensuite, nous avons également remarqué que l'attribution des points peut varier en fonction de l'ensemble des élèves, pour créer une sorte d'égalité des chances. Bernard fait parfois en sorte d'équilibrer les points des exercices pour ne pas déprécier certains élèves ou au contraire pour ne pas en valoriser d'autres. Il le mentionne ainsi :

Donc là, c'est clair que je mets $\frac{1}{2}$ par réponse juste. J'attends huit réponses différentes finalement, mais je ne peux pas mettre 8 points pour un exercice comme ça, c'est beaucoup. Pour ceux qui ne comprendraient pas, ils perdent énormément, et pour ceux qui comprennent, ils sont tout de suite dans le haut du classement. Ce n'est pas évident, mais faut quand même essayer de bien juger où on met les points et puis ne pas en mettre trop d'un coup dans un exercice sauf s'il est vraiment conséquent et si c'est un exercice accessible à tous peut-être.

Même chose pour Élisabeth qui a décidé de complètement modifier l'attribution des points pour un exercice. En effet, après avoir réalisé que certains élèves avaient eu beaucoup de difficultés pour un exercice, elle a décidé d'attribuer la moitié des

points. Le nombre de points pour cet exercice est passé de 12 à 6 points ($\frac{1}{2}$ point par réponse correcte au lieu de 1 point). Voici son explication :

Je me rends compte qu'avec les deux dernières corrections, cet exercice totalise un nombre de points de 12 points, ce qui est vraiment beaucoup par rapport aux autres exercices. Je me rends compte avec les deux dernières copies que ces deux élèves perdent énormément de points sur cet exercice ce qui va vraiment baisser leur note. Je suis en train de me demander si je ne vais pas compter plutôt 6 points à la place de 12 points pour cet exercice-là. [...]. Après réflexion, j'ai effectivement mis plutôt 6 points à l'exercice 3 à la place de 12 points. Je trouve que c'est plus équilibré par rapport au reste du contrôle. [...]. Les points sont mieux répartis dans le contrôle. Je trouve que c'est plus équitable, que c'est plus juste. Voilà pourquoi j'ai effectué ce changement."

Josette, quant à elle, attribue parfois plus ou moins de points en fonction du degré d'entraînement des élèves en disant :

Je m'étais moins de poids à cet exercice-là, parce qu'on l'a moins entraîné en classe.

De plus, nous avons observé que tous les enseignants attribuent des points à l'avance en fonction des étapes à réaliser (en principe pour les situations-problèmes) ou en fonction de la réponse obtenue (en principe pour les tâches simples). Ils peuvent donc avoir des attentes précises quant aux réponses des élèves :

| Points attribués |
|---|
| <ul style="list-style-type: none">- Comme il y avait cinq points à mettre dans le système d'axe, j'ai mis 1 point par point (Josette).- 1 point pour la division et puis 1 point pour la preuve (Bernard).- À l'exercice b), j'ai mis 6 points, donc 1 point pour le vrai ou faux si c'est correct et puis 1 point pour la justification si elle tient la route (Bernard).- J'avais mis 1 point pour le calcul pour la part de la commune, j'avais mis 1 point pour ce que doivent payer les élèves pour le trajet, 1 point pour ce qu'ils doivent payer pour les repas, 1 point pour l'addition totale, 1 point pour la soustraction de ce qu'on a comme argent et 1 point pour la division qui nous donne le résultat. Ça fait bien 6 (Cosette).- J'ai mis un demi-point par nombre trouvé ce qui donne un total de 6 points (Élisabeth). |

3.2.1.5 Barème

Sur les quatre enseignants de notre étude, trois d'entre eux (Josette, Bernard et Josette) constituent leur barème lorsque toutes les copies sont corrigées. La quatrième enseignante (Élisabeth) l'établit avant de commencer la correction.

En ce qui concerne le barème, les manières de procéder des enseignants peuvent différer d'un enseignant à l'autre. Josette, Bernard et Élisabeth fonctionnent de la même manière. En effet, pour chaque total de points, ils attribuent une note :

| Total de points = note |
|---|
| <ul style="list-style-type: none">- 22 points 21 ½ ça fait 6, 21 points 5 ¾, etc. (Josette).- Je fais ma répartition de notes sur le nombre de points, comme ça j'ai directement mon barème et je vois directement la note qu'ils ont faite. [...] Donc là je vais essayer de mettre directement 6 pour les quatre premiers points, ensuite 5 ½, etc. (Bernard).- Pour celui-ci, j'ai fait comme barème 51, 50 et 49 points donnent 6, 48 et 47 points donnent 5 ¾, etc. (Élisabeth). |

Ensuite, pour légitimer leur barème, Josette et Bernard le comparent avec la formule fédérale suivante :

$$\text{note} = (\text{points acquis} / \text{total des points}) \times 5 + 1$$

| Formule fédérale officielle |
|---|
| <ul style="list-style-type: none">- Pour légitimer mon barème, par exemple prends 18 ½, je fais quand même la formule fois 5 divisé par 22 plus 1 et je regarde si ça correspond. Là en l'occurrence ça correspond donc je me sens assez juste en fait dans mon barème (Josette).- C'est mon barème, ça équivaut au barème "nombre de points divisés par le nombre de points total fois 5 plus 1", le barème officiel entre guillemets (Bernard).- Le barème, ça va se passer comme ça : en math, en général, je prends le total des points de l'épreuve et ensuite je le multiplie par le nombre de points qu'a fait l'élève. Ça donne une note au dixième qui est représentative du pourcentage de points justes. |

Nous avons observé plusieurs arrangements que font les enseignants quant à l'attribution du barème ou à sa modification :

Arrangements en faveur de l'élève

- Si j'ai oublié une faute, je ne baisse pas la note et si j'ai mis une faute en trop, alors là je mets la note qu'ils auraient dû avoir (Josette).
- Je n'aurais peut-être pas besoin de faire un barème quand même trop sévère finalement. Je pense que c'est une évaluation où les points sont relativement bien répartis (Bernard).
- Le demi, je le compte vers le haut (Bernard).
- Par contre, il se peut que je corrige le barème en fonction du résultat des élèves. Je verrai à la fin de la correction. [...]. On verra à la fin des corrections si je maintiens ce barème ou pas (Élisabeth).
- J'ai donc choisi de garder le barème que je m'étais fixé au départ, car je trouve qu'il correspond bien, enfin qui montre bien ce que les élèves ont été capables de faire (Élisabeth).

Arrangements en défaveur de l'élève

- Je suis plus sévère, mais pour un contrôle comme ça, ça ne me dérange pas, parce que c'est un contrôle qui est facile, qui reflète bien ce qu'ils doivent savoir. Et je n'ai pas envie d'être trop sévère, mais c'est juste que j'ai une attente assez haute (Josette).
- Peut-être que je vais devoir faire un barème un peu plus sévère (Bernard).

D'une manière générale, nous apercevons la constitution du barème diffère que légèrement d'un enseignant à l'autre. En effet, soit ils utilisent la formule officielle, soit ils attribuent une note pour chaque total de points. Cependant, certains enseignants modifient leur barème en cours de correction en fonction des résultats des élèves. Cette modification peut soit être bénéfique, soit défavorable à l'élève.

3.2.1.7 Interprétation des pratiques effectives

L'acte correctif est une tâche complexe pour l'enseignant. En effet, une multitude de paramètres entre en ligne de compte et doit être gérée par le correcteur. Les principaux paramètres sont, entre autres : le choix du type d'exercice, la pondération des points par exercice, l'attribution des points par critère d'évaluation, le choix et l'élaboration du barème. L'attribution des points nous paraît être la tâche la plus difficile à réaliser et à justifier. Comme l'ont mentionné Durant & Chouinard (2006), pour qu'une évaluation soit valide, les données récoltées doivent s'appuyer sur des critères précis et mesurer effectivement les objectifs de départ. Nous trouvons surprenant qu'aucun enseignant n'ait mentionné les objectifs spécifiques de leur évaluation. Sont-ils instinctivement intégrés dans leur processus d'évaluation? Est-ce qu'avec l'expérience, les objectifs n'ont plus lieu d'être explicités? Ou au contraire, les objectifs poursuivis et les critères de notation ne sont pas définis avec précision (effet de flou)? Il nous est ainsi difficile de savoir si les enseignants ont effectivement et strictement évalué les objectifs de départ, ou si certains critères ont été inclus en cours de route par les enseignants. Si un enseignant ne définit pas clairement les objectifs de départ, il est possible que des critères intrusifs se greffent au domaine évalué, ce qui engendrerait un regard faussé et une modification de la notation. Selon Roegiers (2000), pour qu'il y ait efficacité dans la correction, il est nécessaire de construire des grilles de corrections ciblées. Il conseille de s'aider d'un tableau d'indicateurs pour effectuer les corrections, ce qui lui permet de donner des points de repère pour garantir la fiabilité de la correction. Sans l'utilisation d'une grille de correction, la notation peut dépendre du type d'interprétation du correcteur et peut donc devenir arbitraire.

Nous avons également relevé une pratique intéressante dans l'enregistrement de Bernard. En effet, il indique qu'il collabore avec une enseignante du même degré en s'échangeant les évaluations sommatives et en les supervisant mutuellement. Cette pratique nous semble pertinente, puisque le choix des objectifs, des exercices, l'attribution des points et du barème sont élaborés conjointement. L'enseignant peut avoir une meilleure vue d'ensemble, ce qui peut impliquer une baisse de la subjectivité.

En ce qui concerne le choix du type d'exercice par l'enseignant, nous pouvons nous questionner sur la subjectivité, surtout si le type d'exercice est choisi en vue de faciliter la correction. Le type d'exercice devrait être défini en fonction des critères de départ, donc de ce qu'on attend des élèves, et non pas en fonction de son humeur ou n'importe quel autre critère inhérent au correcteur. Il semblerait que ce choix, certainement fait de manière inconsciente, réconforte les enseignants. Il semble plus aisé de justifier la correction lorsque les réponses attendues sont limitées (c'est soit juste, soit faux). C'est le contraire avec les situations-problèmes, où un développement est demandé et où plusieurs variables peuvent être prises en compte.

Enfin, concernant la constitution du barème, nous pouvons admettre que tous les enseignants établissent leur barème en fonction des points acquis et non en fonction des erreurs. Cette pratique est à promouvoir, puisqu'elle met en avant ce que sait l'élève et donc valorise son travail. Lorsque nous étions élève à l'école primaire, il n'était pas rare de recevoir une évaluation sommative avec le nombre de fautes en haut de la feuille, qui était transformé en note finale. C'était une pratique peu encourageante, mais qui heureusement a évolué.

3.2.2 Les biais évaluatifs/arrangements

Il convient tout d'abord d'indiquer que les différents biais évaluatifs et arrangements cités dans la problématique n'ont pas pu tous être observés, tels que la loi de stéréotypie et les différences entre correcteurs. En effet, ces deux points ne sont pas ressortis des enregistrements puisque dans notre recherche, nous n'avons pris en compte qu'une évaluation par enseignant. Cela ne nous indique pas s'il y a effet de stéréotypie (normalisation de la notation établie) ou non. Puis, pour les différences entre les correcteurs, nous n'avons pas élaboré de comparaison de correction d'une même copie par plusieurs enseignants.

3.2.2.1 État du correcteur

Plusieurs aspects attachés à l'état de l'évaluateur sont ressortis dans trois enregistrements sur quatre. Ces états peuvent modifier la manière dont l'enseignant corrige. Ils peuvent être liés à la fatigue de l'enseignant, à l'ennui/à la motivation de vouloir corriger tel ou tel exercice, à un agacement, à une distraction quelconque, à l'incertitude ou à des difficultés face à des décisions à prendre :

| Fatigue |
|---|
| <ul style="list-style-type: none">- J'en ai marre. J'ai envie d'aller au lit. Courage (Josette).- Je suis fatiguée. Ce n'est pas l'heure pour faire des corrections. Je vais quand même aller jusqu'au bout (Cosette).- Je dois être fatiguée ou alors je ne comprends vraiment pas (Cosette).- Je suis vraiment fatiguée. Faut que j'arrête, parce qu'il est tard (Cosette).- Oulàlà! Ça fait déjà 51 minutes que je parle. Quelle horreur! [...]. Ça m'a dépitée de réaliser que ça faisait 51 minutes que je parlais. [...] Il y a encore du boulot apparemment. Je ne sais pas où j'en suis dans mes corrections, mais je préfère ne pas savoir si ça fait déjà plus de 50 minutes (Cosette). |
| Ennui ou agacement |
| <ul style="list-style-type: none">- Je n'ai pas envie de corriger cet exercice, parce que je pense que chacun à fait à sa sauce et qu'il y aura plusieurs choses qu'il faudra que j'accepte ou pas (Josette).- Il ne me reste plus que mon exercice 2 que je n'ai pas envie de corriger parce que voilà. Mais bon, je m'y mets (Josette).- Ça m'énerve, c'est du "pifaumètre" (Bernard).- Ça m'énerve par exemple un brouillon comme ça (Cosette).- C'est un vrai casse-tête ce truc ! J'ai déjà fait des épreuves avec ce genre de problème, mais je ne me souviens pas d'avoir autant peiné (Cosette). |

| Motivation |
|---|
| <ul style="list-style-type: none"> - Alors, je vais déjà, par flemme, commencer par l'exercice 3. [...]. C'est plus facile à corriger (Josette). - J'aime l'exercice 7, c'est tellement plus facile (Cosette). - Oh youpi! C'est la dernière (Cosette). |
| Distraction |
| <ul style="list-style-type: none"> - Il faut que je m'interrompe dans ma correction parce qu'on a sonné. [...]. J'arrête deux minutes. [...]. Je reprends. Alors j'en étais où? (Josette). - Je vois que la batterie est basse j'espère que... Petite pause de batterie [...]. Ok je reprends la correction du contrôle thème 2 que j'ai dû interrompre ce matin parce que je n'avais plus de batterie (Bernard). - Je vais devoir te quitter un instant pour chercher un post-it (Cosette). |
| Incertitude ou difficulté |
| <ul style="list-style-type: none"> - Je me demande toujours si je suis trop sévère (Josette). - Maintenant ça fait plusieurs années que j'enseigne, mais je ne suis toujours pas trop sûr de la manière que j'attribue des points (Bernard). - Problème de consigne. J'ai eu de la peine à le formuler ça ici (Bernard). |

3.2.2.2 Connaissance de l'élève et de ses antécédents

Dans les enregistrements, nous avons remarqué que les enseignants sont intéressés à connaître le prénom de l'élève avant de corriger une copie. Par exemple, Josette dit : "C'est qui? C'est Julie." ou encore "C'est qui ça? Jocelyne." Tout comme Cosette qui dit "Je passe à ... C'est qui? À Clément." Pour les autres enseignants, ils n'expriment pas explicitement leur envie à connaître le prénom de leur élève. Cependant, nous pouvons en déduire, puisqu'ils font souvent le lien entre la réponse donnée et certaines caractéristiques des élèves. Cette déduction nous porte à croire que, pour tous les participants, la correction ne se fait pas de manière anonyme.

Le fait de connaître le prénom (et donc les antécédents scolaires) de l'élève que l'enseignant corrige peut influencer son jugement. C'est un élément qui se rapporte à

l'effet de contamination et à l'effet pygmalion décrits dans la problématique. Dans les enregistrements, les enseignants mentionnent très clairement la connaissance des antécédents, soit par l'étonnement, soit par la normalité. Quelques exemples sont décrits ci-dessous :

Etonnement

- Il m'épate quand même (Josette).
- Là par exemple, c'est bizarre. C'est un assez bon élève, mais il manque une division (Bernard).
- Génial là, elle a compris la consigne. C'est bien qu'elle puisse faire un peu des points, parce que celle-ci, c'est quand même pas évident les math (Bernard).
- Je suis d'ailleurs étonnée, car c'est une bonne élève en principe en math et il me semble que pour le début de ce contrôle, il y a quand même passablement d'erreurs (Élisabeth).
- Je suis très étonnée du résultat de cette élève qui a fait un bon résultat alors qu'habituellement elle a quelques difficultés en math donc je suis très contente de son résultat (Élisabeth).

Normalité

- Je sais que Juliane, elle aurait tendance à confondre encore les axes 1 et axes 2 (Josette).
- Ça ne m'étonne pas vraiment qu'elle a eu de la peine dans cet exercice (Bernard).
- Bérénice, le numéro 3, elle est en net progrès. [...]. Alors là on voit qu'elle est bien suivie ça ne m'étonne pas (Bernard).
- Génial il a tout compris c'était qui? Brigitte. Ça ce n'est pas surprenant (Bernard).
- Ensuite on passe à Billy, numéro 5, notre étourdi (Bernard).
- Ce n'est pas extraordinaire. Il pourrait faire mieux, mais il traverse un peu une mauvaise passe ces temps-ci. On peut comprendre (Bernard).
- C'est correct. Ça lui correspond, j'ai l'impression (Bernard).
- C'est dommage, mais ça lui arrive assez souvent de perdre un peu des points au contrôle, parce qu'il ne fait pas tout en entier (Bernard).
- Cyril, ce n'est pas vraiment son truc le calcul réfléchi (Cosette).
- Elle s'en sort bien Lola en math depuis un certain nombre de semaines (Cosette).
- Clovis c'est un bon élève aussi en math. Donc pour l'instant, il le prouve bien (Cosette).
- Cassie, la meilleure élève de la classe. Ce qui présage quelque chose d'intéressant (Cosette).
- J'ai de la peine à comprendre, parce que normalement c'est un élève qui comprend bien ce qu'on fait, qui obtient de bons résultats. Donc là vraiment je suis étonnée (Élisabeth).
- C'est une bonne élève donc ça correspond à ce à quoi je m'attendais chez elle (Élisabeth).

Nous avons également remarqué que les quatre enseignants ont parfois tendance à s'appuyer sur la connaissance des antécédents des élèves pour justifier l'erreur de l'élève :

Justification d'une erreur

- Pour moi là, c'est un manque de concentration plus que vraiment il n'a pas compris, surtout qu'en classe, je peux voir que c'est bon (Josette).
- Pour moi ça c'est manque de relecture (Josette).
- Je sais que c'est juste un oubli. Il n'a juste pas relu (Josette).
- Là encore un qui s'est un peu dépêché et qui se n'est pas rendu compte que la première servait à vérifier qu'on avait fait juste (Bernard).
- C'est de l'inattention à mon avis ça (Cosette).
- Cette élève n'a vraiment pas bien écouté quand on a donné les consignes (Élisabeth).
- Vraiment dommage par rapport à l'écoute des consignes. D'ailleurs, ce n'est pas la première fois que ça arrive pour cette élève (Élisabeth).

3.2.2.4 Comparaison entre les élèves/globalité de la classe

Dans les enregistrements, nous avons remarqué que les enseignants ont tendance à comparer les réponses ou les résultats entre les élèves. Cette manière de procéder peut influencer le jugement et donc la manière de corriger des enseignants.

Dans un premier temps, nous avons soulevé l'effet de contraste décrit par Leclercq, Nicaise et Demeuse (2004), qui est un jugement comparatif lorsque l'enseignant rencontre une copie qu'il considère comme représentative. En effet, Bernard utilise la copie des meilleurs élèves comme base de correction :

| Effet de contraste |
|---|
| <ul style="list-style-type: none">- Ce que je fais quand je commence la correction d'un contrôle de math, c'est que je prends les deux ou trois contrôles des meilleurs élèves. Ça me permet de voir déjà si les calculs sont corrects et de voir assez vite si je suis dans le juste avec les élèves ou pas.- Je reprends avec une élève, la numéro 1, qui fait toujours des assez bons résultats. Ça me permet toujours de me situer pour voir si c'est faisable ou pas, si tout va bien.- En fait en corrigeant le contrôle d'un élève qui travaille bien, qui a généralement des bons résultats, ça me fait directement mon corrigé. Je peux comparer, ça va beaucoup pour vite dans la correction.- Je vais recomparer les bons élèves au début pour faire la bonne solution.- Je vais comparer les deux contrôles des deux élèves qui se débrouillent bien, comme ça, ça me permet de vérifier que ça joue. |

Ensuite, d'autres éléments de comparaison ont apparu dans les propos des enseignants :

Éléments de comparaison

- C'était quoi le nombre de points minimum? Ça j'avoue que je triche un petit peu, mais ça m'arrive de regarder pour voir un petit peu où je me situe (Bernard).
- Ce qui ne me réjouit pas tellement, parce que si la meilleure élève de la classe fait 28,5 sur 36, comme ça de tête, ça ne doit pas faire un résultat plus haut que 5 (Cosette).
- Ça arrive des fois qu'ils ne comprennent pas la consigne, alors je regarde si vraiment ce n'était pas clair du tout ou s'il y en a juste deux ou trois qui se sont plantés. Et là, globalement l'exercice a quand même été compris comme je voulais qu'il soit fait [...]. Donc ma fois, je suis navré pour mon bon élève qui a commencé par se planter (Bernard).
- Yes, il a réussi! C'est la preuve que c'était possible. Ouf! (Bernard).

Il arrive parfois que certains enseignants jugent de manière globale la réussite ou non d'un exercice particulier ou alors de l'ensemble de la copie. Ils regardent alors la réussite ou l'échec dans sa moyenne, c'est un regard global et non individuel. Si bien qu'un élève aurait très bien pu réussir un exercice alors que les autres élèves pas. Inversement, un élève aurait pu rater l'épreuve alors que la majorité l'aurait réussie. Il est possible que les enseignants en fassent une généralité :

Regard global

- C'est bien, je suis contente. Je ne pensais pas que ça irait si bien quand même (Josette).
- Globalement, c'est vraiment un bon contrôle. C'est épatant, c'est assez homogène. C'est vrai qu'on a quand même bien travaillé, bien réviser, à par ce thème 2 là, mais le calcul mental un peu moins. Sinon, on a bien travaillé. Je ne suis pas toujours aussi contente (Josette).
- D'après ce que j'ai vu, ça a posé pas mal de problèmes à l'ensemble des élèves (Cosette).
- Quelle catastrophe cette épreuve (Cosette).
- C'est des élèves en huitième qui ne maîtrisent pas du tout le livret, qui ne maîtrisent pas du tout les soustractions, qui ne maîtrisent pas du tout (Cosette).
- Alors c'est vrai que si pour un contrôle, je vois que c'est la catastrophe, pour moi ça veut dire que l'on n'a pas assez travaillé en classe. Je les évalue pour quelque chose qui ne maîtrise pas donc ce n'est pas le but. Ce que je veux c'est qu'ils maîtrisent (Josette).
- Ça me permet de viser toujours assez dans le juste par rapport à une courbe de Gauss qu'on devrait toujours retrouver (Bernard).
- Sinon, je risque d'avoir un peu trop de bonnes notes. Ça sera dangereux (Bernard).
- Ce que je peux faire aussi pour essayer d'appliquer une courbe de Gauss, ce que je fais des fois, c'est que je mets une coche à côté du nombre de points qu'ils ont fait (Bernard).

Pour finir, nous avons observé que deux enseignants sur quatre établissent une moyenne de classe :

Josette : Moyenne de classe : 5,28, j'ai arrondi à 5,3, voilà donc en réalité, c'est des thèmes qui vont bien, qui sont globalement assez acquis.

Élisabeth : On arrive donc à une moyenne de classe de 4,6 pour ce contrôle.

3.2.2.5 Interprétation des biais évaluatifs/arrangements

Selon les indications ci-dessus, nous pouvons en déduire que les biais évaluatifs décrits par la docimologie critique sont présents dans les évaluations, que nous le voulions ou non.

De nombreux exemples nous indiquent que l'état du correcteur peut modifier son jugement. En effet, certains enseignants ne cachent pas leur fatigue, leur ennui ou leur agacement à réaliser cette tâche. Sont-ils effectivement plus sévères dans un tel cas de figure? Ou au contraire, sont-ils plus cléments lorsqu'ils sont reposés et de bonne humeur? Nous ne pouvons malheureusement pas y répondre. Mais nous pouvons affirmer que parfois, il est préférable de différer la correction plutôt que de s'acharner sur une évaluation. Nous remarquons également qu'il est possible que des enseignants éprouvent des difficultés ou sont incertains face à leur pratique. Malgré leurs années d'expérience, certains facteurs peuvent poser problème, comme le degré de sévérité ou encore l'attribution des points. Ces incertitudes nous semblent être favorables dans le monde de l'enseignement, puisqu'il implique une constante remise en question. Cependant, elles ne doivent pas être un frein à la correction et augmenter le caractère subjectif.

Selon Allal (2009), différents biais évaluatifs sont liés à des facteurs individuels (de l'élève). Dans les données récoltées, nous avons décelé quelques éléments qui pourraient s'apparenter aux biais suivants : l'effet de stéréotypie, l'effet de halo, l'effet de contamination ou encore l'effet Pygmalion. En effet, les enseignants explicitent parfois très clairement leur étonnement face à une réponse d'un élève ou, au contraire, se réconfortent dans l'idée qu'ils avaient de l'élève en question. Lorsqu'un enseignant dit : "Je sais que Juliane, elle aurait tendance à confondre encore les axes 1 et axes 2.", n'est-elle pas d'emblée influencée par l'idée qu'elle a des capacités de son élève? Il leur arrive parfois même de justifier l'erreur d'un élève en imaginant la cause de l'erreur ou en s'appuyant sur les antécédents de l'élève. Ce sont des indices qui nous montrent bien que les enseignants ont connaissance des antécédents de leurs élèves et en tiennent compte dans la correction.

Dans les processus d'évaluation, nous avons décelé une certaine comparaison entre les élèves. Par exemple, l'effet de contraste, défini par Jorro (2000), est clairement

représenté dans un enregistrement en particulier. En effet, en corrigeant et en se basant sur les copies des deux meilleurs élèves (copies représentatives), il effectue un jugement comparatif. Il va ainsi entrer dans un processus de différenciation positive ou négative.

Comme le mentionnent Leclercq, Nicaise et Demeuse (2004), la loi de Posthumus indique une distribution des notes selon une norme. Nous avons constaté qu'un enseignant applique la courbe de Gauss dans son processus d'évaluation. La courbe de Gauss est une référence utilisée consciemment ou non par les enseignants. Elle favorise une évaluation au service de la sélection et d'une comparaison des uns par rapport aux autres. En utilisant une courbe de Gauss pour déterminer les notes des élèves, l'enseignant n'évalue pas des maîtrises individuelles par rapport à des critères précis. Il va mettre en évidence une distribution des performances (bonnes, moyennes ou mauvaises), donc une comparaison (les performances des uns se définissent par rapport aux performances des autres). Cette pratique va à l'encontre d'une pédagogie qui prône la réussite de tous. Le but de l'école n'est-il pas d'amener les élèves à réussir dans leur cursus scolaire? Si tel est le cas, nous pensons que cette pratique est à éviter.

Il en va de même pour la moyenne de classe. Pour quelles raisons certains enseignants indiquent-ils la moyenne de classe aux élèves et à leurs parents? À notre avis, la moyenne de classe est uniquement une information destinée à l'élève et à ses parents. Elle permet de repérer la position de chaque élève en fonction des résultats de tous les élèves de la classe et laisse donc apparaître une comparaison entre les pairs. Nous nous questionnons sur cette pratique d'un point de vue pédagogique.

Conclusion

Présentation synthétique des principaux résultats

Pour conclure notre démarche de recherche, qui avait pour finalité de découvrir et comprendre la manière dont les enseignants corrigent, ainsi que de déceler les facteurs qui peuvent influencer leurs choix, voici ci-dessous la synthèse de notre recherche. En effet, après avoir rédigé un cadre théorique, mis en place une méthodologie et récolté des données, nous allons maintenant revenir sur des éléments importants et susceptibles d'amener des réponses à la première question de recherche suivante :

Comment les enseignants corrigent les évaluations sommatives de leurs élèves?

Tout d'abord, les données récoltées nous ont permises de découvrir de multiples façons de procéder lors de l'acte correctif. Les pratiques évaluatives peuvent s'avérer bien différentes lorsque nous comparons la manière dont les enseignants corrigent. En effet, les quatre enseignants qui ont participé à notre recherche ont montré une certaine diversité et une richesse dans leurs pratiques évaluatives.

Plusieurs éléments entrent en ligne de compte dans le processus évaluatif. L'enseignant doit prendre une multitude de décisions. Dans la phase de correction, l'enseignant opère un jugement sur la réponse donnée. Il détermine si la réponse est correcte ou incorrecte, selon une référence (par exemple un corrigé).

Concernant la phase de pondération, l'enseignant détermine des données attendues (éléments de réponse dans un exercice) et les compare avec les données de l'élève. Il pondère ensuite chaque élément de réponse. Malheureusement, la manière concrète de pondérer n'est pas explicitée dans les propos des enseignants.

Lors de la phase de la notation, les enseignants totalisent simplement les points obtenus par l'élève pour en faire une note. Dans tous les cas, un barème précis les aide à la définir. Deux pratiques ont été relevées pour l'élaboration d'un barème : soit le total de points correspond à une note, soit il est constitué avec la formule officielle "nombre de points de l'élève x cinq ÷ par le nombre de points totaux + 1".

Cependant, ces deux pratiques se mettent d'accord sur un même point : l'enseignant comptabilise le nombre de points corrects et non pas le nombre d'erreurs.

La deuxième question de recherche, visant à déterminer certains critères pouvant influencer le jugement et les prises de décisions des enseignants, est la suivante :

Quels facteurs peuvent influencer le jugement des enseignants dans l'acte correctifs?

Nous pouvons mettre en avant que tous les participants de notre recherche sont influencés par des biais évaluatifs et des ajustements de manière consciente ou inconsciente lors de la correction. En effet, nous remarquons que certains facteurs, comme l'état du correcteur, la connaissance de l'élève et de ses antécédents, ainsi que la comparaison entre les élèves (globalité), peuvent amener à une certaine subjectivité dans le jugement de l'enseignant. La plupart des biais évaluatifs et des arrangements évoqués dans la problématique sont donc visibles dans les pratiques des enseignants. Cependant, il est important de préciser que ces différents éléments ne prouvent en rien l'existence du côté objectif ou subjectif dans l'évaluation. Il reflète simplement des indications utiles à la compréhension de l'acte correctif. Cependant, au terme de notre démarche de recherche, d'autres questionnements ont apparu : dans quelle mesure les différents biais évaluatifs et arrangements sont-ils néfastes à l'élève? La connaissance de certaines spécificités de l'élève lors de l'évaluation ne pourrait-elle pas être un facteur pour le faire progresser? Il est probable que l'ajustement de l'acte évaluatif en fonction des besoins de chacun et non pas des performances ne rend-il pas l'évaluation plus juste?

Autoévaluation critique de la démarche

Nous réalisons que le nombre de participants à notre recherche est assez faible. En effet, il aurait été plus intéressant d'avoir un échantillon plus conséquent pour pouvoir observer et comparer plus de pratiques, ce qui aurait enrichi de manière non négligeable nos résultats. Cependant, les quatre enregistrements récoltés se reposent sur des pratiques effectives, dictées par des enseignants sur le terrain. Ce sont des discours certes, mais ils reflètent effectivement ce qui se retrouve dans leurs pratiques.

Le domaine de l'évaluation est un aspect que nous allons retrouver fréquemment tout au long de notre future profession. Ce travail de recherche nous a permis d'élargir nos connaissances sur différentes pratiques évaluatives. Elles nous ont apporté certaines clés qui nous seront utiles pour notre avenir professionnel. Au début de ce travail, nous avions des doutes liés à la volonté d'être juste et équitable. Avec les différentes pratiques observées, nous nous sentons rassurée, puisque nous avons découvert certaines difficultés et certains doutes chez les enseignants. En outre, nous pourrions sélectionner et utiliser les procédés qui nous semblent les plus pertinents pour notre future pratique professionnelle. Ce travail nous a ainsi permis de développer une réflexion personnelle sur le sujet, de se poser les bonnes questions, de développer un sens critique à travers notre pratique, ainsi que de prendre les bonnes décisions, de se positionner et d'assumer nos choix.

Proposition d'élargissement de la réflexion

Au moment de définir la méthodologie que nous allons utiliser dans ce travail de recherche, deux options étaient ressorties. La première option était de récolter les enregistrements, les transcrire et les analyser (option finalement choisie). La deuxième option était de récolter des enregistrements, utiliser les propos des enseignants pour se focaliser sur des éléments précis dans la correction pour en faire un guide d'entretien, s'entretenir avec les enseignants concernés, transcrire les entretiens et les analyser. De ce fait, cette deuxième option pourrait être une proposition d'élargissement de la réflexion, puisqu'elle permettrait d'aller plus en détail dans la problématique. En effet, nous pourrions cibler et questionner les enseignants sur des moments-clés, des pratiques complexes lors de la verbalisation leur pratique évaluative. Cette méthode impliquerait davantage les participants et permettrait une meilleure compréhension et un approfondissement des propos des enseignants (peut-être pas assez complets, mal compris ou mal interprétés dans la première option).

Références bibliographiques

Ouvrages

Allal, L. (2009). Chapitre 1. Pratiques évaluatives des enseignants face aux méthodologies des recherches évaluatives portant sur le système scolaire. In L. Mottier Lopez & M. Crahay, *Évaluations en tension* (pp. 29-45). Bruxelles : De Boeck Supérieur.

Blanchet, A. & Gotman, A. (1992). *L'enquête et ses méthodes : l'entretien*. Paris : Nathan.

Blanchet, A. & Gotman, A. (2001). *L'enquête et ses méthodes. L'entretien*. Paris : Armand Colin.

Bonniol, J.-J. & Vial, M. (2009). *Les modèles de l'évaluation : textes fondateurs avec commentaires*. Bruxelles : De Boeck Supérieur.

Crahay, M. (2007). *Peut-on lutter contre l'échec scolaire?* Bruxelles : De Boeck Supérieur.

De Landsheere, G. (1979). *Dictionnaire de l'évaluation et de la recherche en éducation*. Paris : PUF-Presses Universitaires de France.

De Landsheere, G. (1980). *Examens et évaluation continue*. Précis de docimologie. Bruxelles : Labor, Paris : Nathan.

De Landsheere, G. (1992). *Évaluation continue et examens*. Précis de Docimologie. Bruxelles : Labor, Paris : Nathan.

Durand, M.-J., & Chouinard, R. (2006). *L'évaluation des apprentissages : de la planification de la démarche à la communication des résultats*. Montréal : Hurtubise HMH.

Fortin, M.-F. (2010). *Fondements et étapes du processus de recherche. Méthodes quantitatives et qualitatives*. Montréal : Chenelière Éducation.

- Giraudeau, C. (2014). Notation scolaire : les apports de la psychologie. In C. Giraudeau & G. Chasseigne (EPU). *Psychologie, éducation et vie scolaire* (pp. 29-30). Paris : Publibook.
- Giroux, S. (1998). *Méthodologie des sciences humaines. La recherche en action*. Québec : Éditions du Renouveau Pédagogique.
- Hadji, C. (1989). *L'évaluation, règles du jeu*. Paris : ESF.
- Huver, E. & Springer, C. (2011). *L'évaluation en langue-Nouveaux enjeux et perspectives*. Paris : Didier.
- Jorro, A. (2000). *L'enseignant et l'évaluation : des gestes évaluatifs en question*. Bruxelles : De Boeck Supérieur.
- Lafortune, L. & Allal, L. (2008). *Jugement professionnel en évaluation : pratiques enseignantes au Québec et à Genève*. Québec : Presses de l'Université du Québec.
- Leclercq, D., Nicaise, J. & Demeuse, M. (2004). Docimologie critique : des difficultés de noter des copies et d'attribuer des notes aux élèves. In M. Demeuse, *Introduction aux théories et aux méthodes de la mesure en sciences psychologiques et en sciences de l'éducation* (pp. 273-292). Liège : Les Éditions de l'Université de Liège.
- Mottier Lopez, L. & Allal, L. (2010). Chapitre 15. Le jugement professionnel en évaluation : quelles triangulations méthodologiques et théoriques ? In Léopold Paquay & al., *L'évaluation, levier du développement professionnel ?* (pp. 237-249). Bruxelles : De Boeck Supérieur.
- Piéron, H. (1951). *Vocabulaire de la psychologie, publié avec la collaboration des travailleurs scientifiques*. Paris : Presses Universitaires de France.
- Poisson, Y. (1991). *La recherche qualitative en éducation*. Québec : Presses de l'Université du Québec.
- Roegiers, X. (2000): *L'école et l'évaluation*. Bruxelles : De Boeck Supérieur.
- Talbot, L. (2009). *L'évaluation formative : comment évaluer pour remédier aux difficultés d'apprentissage*. Armand Colin.

Vermersch, P. (2014). Le dessin de vécu dans la recherche en première personne. Pratique de l'auto-explicitation. In Depraz N., *Première, deuxième, troisième personne* (pp. 195-233). Bucarest : Zeta Books.

Articles

Barbier, J.-M. (1983). Pour une histoire et une sociologie des pratiques d'évaluation en formation. *Revue française de pédagogie*, (63), pp. 47-60.

Blais, M. & Martineau, S. (2006). L'analyse inductive générale : comment extraire le sens derrière les données brutes. *Recherches qualitatives*, 26(2), pp. 1-18.

Chabot, J.-M. (2004). Evaluer, corriger, pondérer, noter, classer... *Revue pratique*, (54), 311-2.

De Ketele, J.-M. (2010). Ne pas se tromper d'évaluation. *Revue française de linguistique appliquée*, XV(1), pp. 25-37.

Mottier Lopez, L., & Allal, L. (2008). Le jugement professionnel en évaluation: un acte cognitif et une pratique sociale située. *Schweizerische Zeitschrift für Bildungswissenschaften*, 30(3), pp. 465-482.

Mottier Lopez, L. & al. (2012). La modération sociale : un dispositif soutenant l'émergence de savoirs négociés sur l'évaluation certificative des apprentissages des élèves. *Questions Vives* [En ligne], 6(18), pp. 159-175, mis en ligne le 26 mai 2014, consulté le 13 août 2015 (<http://questionsvives.revues.org/1235>).

Mottier Lopez, L. (2013). De la mesure à l'évaluation collaborative en éducation. *Revue française d'administration publique*, (148), pp. 993-952.

Perrenoud, P. (1989). L'évaluation entre hier et demain. Genève : Service de la Recherche Sociologique et Faculté de Psychologie et des Sciences de l'Éducation. Université de Genève.

Vermersch, P. (2007). Bases de l'auto-explicitation. *Revue Expliciter*, (69), pp. 1-31.

Annexes :

Annexe 1 : Protocole de verbalisation

Haute Ecole Pédagogique - BEJUNE



PROTOCOLE DE VERBALISATION

Voici ci-dessous quelques informations relatives au protocole de verbalisation lors de la correction des évaluations sommatives de mathématiques de vos élèves. Il est très important, pour des questions d'organisation et de pertinence pour mon étude, de corriger vos évaluations **comme vous avez l'habitude** de procéder. Il est évident que toutes les données récoltées, ainsi que votre identité, resteront strictement **anonymes et confidentielles**.

1. Avant la correction, numérotez les copies **dans l'ordre de correction**.
2. En commençant de corriger chaque évaluation, **nommez le numéro de la copie** (ou si vous le souhaitez, vous pouvez simplement nommer le prénom de l'élève, en sachant que ces données resteront confidentielles).
3. Précisez à chaque fois **l'exercice qui vous êtes en train de corriger et le moment où vous tournez une page**. Verbalisez également **vos faits et gestes** en lien avec l'acte de correction (par exemple : reprise d'une copie, retour sur un exercice déjà corrigé, moment de pause, ...)
4. Pendant l'acte correctif même, **verbalisez ET précisez le plus fidèlement possible TOUTES vos pensées**. Ce point est le plus important pour mon étude : plus vos propos seront précis et fidèles, meilleure sera l'étude. N'ayez aucune crainte de dire le fond de votre pensée, même si elle peut vous sembler inadéquate. Il n'y aura **AUCUN jugement** sur ce que vous pourrez dire. Pour rappel, l'objectif de cette étude ne porte pas sur les "bonnes" ou "mauvaises" pratiques évaluatives, mais sur les pratiques effectives des enseignants.
5. **Justifier**, dans la mesure du possible, vos choix et vos décisions (par exemple : pourquoi avoir mis 2 points sur 3 ?, la moitié des points ?, pourquoi avoir changé le barème (si tel est le cas) ?, ...).
6. Verbaliser également vos ressentis ou vos émotions (joie, tristesse, agacement, colère, satisfaction, surprise, ...) lors de la correction.

Annexe 2 : Tableau comparatif des actes correctifs des enseignants (pp. II-III)

| | Josette | Bernard | Cosette | Elisabeth |
|-----------------------------------|--|--|---|--|
| Points par exercice | attribution des points par exercice avant la correction | attribution des points par exercice avant la correction | - | attribution des points par exercice avant la correction |
| Demi-point | - | sont comptés vers le haut | - | sont comptés vers le haut |
| Corrigé | utilisation d'un corrigé réalisé par elle-même | utilisation des copies des deux ou trois meilleurs élèves en comparant les résultats | utilisation d'un corrigé réalisé par elle-même | - |
| Ordre des copies | - | correction des 2-3 meilleurs élèves d'abord correction des évaluations différenciées à la fin | correction des évaluations différenciées au début, puis au hasard | dans l'ordre inverse dans lequel elles ont été rendues |
| Déroulement | correction et attribution des points de l'exercice 1 du premier élève, puis celui du deuxième, etc. Puis, passe à l'exercice 2, etc. | correction de l'exercice 1 du premier élève, puis celui du deuxième, etc. Puis, passe à l'exercice 2, etc. | correction des exercices sur le recto du premier élève, puis celui du deuxième, etc. Puis, passe au verso, etc. | correction de tous les exercices d'une copie en une seule fois, comptage du total des points, inscription de la note, puis passe à la copie suivante |
| Comptage des points totaux | comptage des points lorsqu'une copie est corrigée | comptage des points lorsque toutes les copies sont corrigées | comptage des points lorsqu'une copie est corrigée | comptage des points lorsqu'une copie est corrigée |
| Points bonus | Il lui arrive de donner des points bonus | - | - | point bonus attribué si l'élève à besoin particulier réalise correctement un exercice alors qu'il devait en faire qu'une partie |
| Constitution du barème | lorsque toutes les copies sont corrigées | lorsque toutes les copies sont corrigées | lorsque toutes les copies sont corrigées | avant la correction |

| | Josette | Bernard | Cosette | Elisabeth |
|--|---|---|--|--|
| Barème | chaque total de points correspond à une note (mais compare son résultat avec la formule officielle) | chaque total de points correspond à une note (mais compare son résultat avec la formule officielle) | formule : "nombre de points de l'élève x cinq , par le nombre points totaux + 1" qui donne une note au dixième | chaque total de points correspond à une note |
| Notes | au quart de point | au demi-point | au dixième | au quart de point |
| Notation des notes | introduction des notes dans un programme informatique et dans le journal de classe | inscription des notes dans le journal de classe | - | - |
| Moyenne de classe | oui | - | - | oui |
| Evaluation différenciée | - | Oui, correction à la fin des séries de copie | Oui, correction au début des séries de copie | Oui, pas d'ordre spécifique pour la correction |
| Annotations/commentaires | -met une croix ou entoure si la réponse est incorrecte -met un vu lorsque la réponse est correcte -souligne (vague) si l'orthographe d'un mot est incorrect | -commentaires possibles (si quelque chose n'est pas clair ou n'a pas de sens ou est bien maîtrisé) -écrit la bonne réponse en-dessous d'une réponse incorrecte | -commentaires possibles -entoure ce qui lui semble étrange -biffe ce qui est faux -écrit la bonne preuve en-dessous d'une preuve incorrecte -met un vu lorsque la réponse est correcte | -commentaires possibles -biffe ce qui est incorrect et entoure parfois le point de l'exercice (a), b), c), ...) -met un vu lorsque la réponse est correcte |
| Changements lors de la correction | - | hésite à changer le barème (le faire plus sévère) | -décide de ne pas faire compter la note -hésite à ne pas faire compter la note aux élèves qui ont un régime différencié -change le nombre de point pour un exercice | décide de changer le barème |
| Preuve | compte des points pour les preuves | compte des points pour les preuves | compte des points pour les preuves | compte des points pour les preuves |
| Autres | | | tient compte des réponses ou preuves effacées | |