

problème de fond<sup>5</sup>, comme celui de la promotion des “faux compétents” dont parle Dubus (2006). Elle ne s’est jamais vraiment posé de questions sur le barème et sa validité, d’où le décalage avec sa remarque sur le fait que les élèves comprennent le barème : si elle ne le comprend que partiellement, comment les élèves le pourraient-ils ?

### 5.3 Portrait de C. (Allemand et anglais, 20 ans d’expérience dans l’enseignement)

Au niveau de l’alignement curriculaire, C. fait attention de reprendre ce qui a été fait en classe (10) et semble s’aligner également avec le CECR<sup>6</sup>, car elle travaille en compétences (14). Le travail en compétences ne prouve pas la cohérence de l’alignement curriculaire élargi ni l’alignement de l’objet avec le reste, mais peut être le signe d’un alignement partiel entre objectifs, tâches et tâches d’évaluation : les objectifs langagiers contiennent en général à la fois la compétence visée (par exemple “écrire un texte”) et l’objet d’apprentissage (p.ex. le passé) et la compétence remplit donc déjà en elle-même une partie des critères. Cependant, C ne cite jamais les objectifs, mis à part sous la forme d’une référence aux compétences (14), qui sont des notions très vagues quand il s’agit d’évaluer l’acquisition d’un objet d’apprentissage. En effet, lorsqu’il faut équilibrer le poids des parties, cela ne se fait pas en fonction des objectifs ou des critères, mais des tâches: *“et, entre les exercices, d’avoir un équilibre, si possible.”* (38) On ne peut pas en déduire que chaque objectif comprend le même nombre d’exercices.

*“Donc là il y a environ, pour- de nouveau plutôt les classes plus... plus jeune et CG: 6-7-8 exercices pour avoir le plus de variété possible. Parce que si on les envoie dans un grand exercice et puis ils se plantent, ben ils se plantent sur toute la longueur. Tandis que s’il y a plusieurs exercices, ils peuvent se rattraper sur une difficulté ou une compétence.”* (14) Cela démontre une envie de donner la possibilité aux élèves de réussir en variant les tâches, mais un point reste dans l’ombre: s’agit-il ici de valider plusieurs fois le même objectif, comme le prône de Ketele, cité par Dubus (2006), ou de multiplier les objectifs ?

C. estime que *“quand tu fais un test, c’est le reflet finalement de ce que les élèves ont appris avec toi.”* (100) A nouveau, ce n’est pas une preuve d’alignement effectif, mais montre qu’elle fait un lien entre enseignement et évaluation, tout comme à la ligne 10. Malgré tout,

---

<sup>5</sup> Quant aux “40 points sur 60”, qui valent en fait 66% et non 60%, nous les avons attribués à la volonté de faire un exemple simple et rapide plutôt qu’à une méconnaissance du barème.

<sup>6</sup> Le CECR est le Cadre Européen Commun de Référence pour les langues. Il a été ratifié en 2001 et promeut l’approche communicative. Les compétences (sous leur forme la plus simple: compréhension orale, compréhension écrite, expression orale et expression écrite), y ont une place de choix.

on trouve chez elle un signe de rupture d'alignement au niveau des tâches d'évaluation : *“Donc ça peut aller jusqu'à, pas du très très facile mais quand-même, heu, voilà, des exercices plus faciles et puis, pour pouvoir faire la différence quand-même avec une bonne, voire très bonne note, des exercices plus... bien plus difficiles.”* (12) Elle ne précise pas si ces exercices “bien plus difficiles” ont été entraînés et si les élèves ont les outils pour les résoudre, ou s'ils contiennent du nouveau. Tout comme Em, elle semble persuadée que son rôle est de différencier les bons élèves des autres, ce qui pourrait expliquer cette habitude de mettre des exercices “bien plus difficiles”.

Elle rompt également clairement l'alignement curriculaire au niveau du barème, car elle utilise le barème fédéral pour tous les tests à l'exception du vocabulaire, bien qu'il ne lui semble pas adapté: *“d'un coup j'avais une moyenne qui était beaucoup trop haute, parce que j'ai fait un test que j'estimais équilibré, et avec ce barème, ils étaient... ben tant mieux pour eux, si y a de ça une fois... Mais voilà, je le trouve un peu trop large, et j'aime pas faire, en fait, exprès des questions justement vaches ou difficiles, d'augmenter un peu artificiellement la difficulté pour avoir une moyenne qui est acceptable au niveau aussi des collègues ou du gymnase ici, quoi. C'est pour ça que j'ai une certaine réticence...”* (80). Elle voit donc le problème que pose un test rendu “artificiellement” plus difficile pour avoir une “moyenne qui est acceptable” au niveau social, mais se range tout de même à la pratique fortement encouragée par l'établissement et renforcée par les files d'allemand et d'anglais. Pour le vocabulaire, le seuil est plus haut, à 70% ou 65% selon les classes, mais c'est de nouveau une convention qui le justifie: *“on a une politique commune pour les langues, anglais et allemand”* (72).

Pour ce qui est de la correction, C. corrige exercice par exercice et même parfois question par question (56) et note tout sur le corrigé. (58) Elle essaie par là de minimiser les biais et d'être la plus impartiale possible. Elle travaille par critères pour l'expression orale et écrite, mais attribue des points aux critères: *“-E: D'accord, il y a une grille avec les critères... -C: Voilà exactement. Et puis qui est formulée en paroles et puis là, c'est après aussi des points.”* (45-46) A la fin, c'est le total de points qui compte. C'est donc d'une certaine manière une forme hybride entre évaluation critériée et évaluation pondérée, ce qui peut venir d'un mélange entre l'envie de juger les élèves sur l'atteinte des critères tout en gardant la mesure à portée de main pour se rassurer sur la “justice” que son jugement représente, car les points sont mieux appréhendés et plus “simples” à justifier.

De manière générale, C. semble chercher à tester ce qu'elle a enseigné, mais cherche parfois plus la simplicité de la mesure et des tâches qu'une recherche de cohérence approfondie dans l'alignement curriculaire, notamment en ce qui concerne les objectifs d'apprentissage.

#### **5.4 Portrait de T. (Maths, 2 ans d'expérience dans l'enseignement)**

T. a un côté très pratique. Elle préfère quand le nombre de points de son évaluation est un multiple de 5, parce que c'est plus facile à transformer en note quand on utilise le barème fédéral (76). Elle distribue des feuilles de révision à ses élèves indiquant quel exercice correspond à quel sujet (14), et elle utilise ces feuilles elle-même pour avoir une vue d'ensemble de ce qui a été traité (24). Les tâches d'apprentissage et celles d'évaluation sont donc corrélées.

En revanche, T. ne parle pas des objectifs d'apprentissage. Au contraire, elle affirme même *“J'ai refusé d'appeler ça objectifs parce que je voulais justement pouvoir sortir un peu de tout ça”* (14). Elle dit qu'elle aime bien *“tester ce qui sera utilisé [dans la suite du programme]”* (26), mais il n'est pas clair si ces connaissances correspondent tout-à-fait aux objectifs d'apprentissage de la séquence en cours. Il y a donc une rupture d'alignement curriculaire à ce niveau-là.

T. essaye de donner aux élèves les moyens de réussir (fiches de révision, notamment) et d'être équitable: *“Je déteste avoir le sentiment d'avantager quelqu'un parce que l'élève est studieuse ou studieux ou parce qu'elle est agréable, ou à l'inverse désavantager quelqu'un parce que l'élève fait tout sauf ça”*. (84)

D'un autre côté, pour avoir une très bonne note, elle estime qu'un élève doit être allé plus loin que ce qui a été fait au cours: *“Parce qu'il faut que, pour que l'élève ait 6, il faut qu'il ait compris plus que... enfin, il faut qu'il fasse plus que répéter ce qu'on a déjà fait”* (28). Au contraire, elle *“essaye d'avoir toujours une partie du test qui soit un peu originale”* (28). Quand elle décrit le test qu'elle nous montre, on a même le sentiment qu'il y a plusieurs parties originales: *“typiquement ça c'est... on l'a pas tellement travaillé sous cette forme-là. Euh, l'exercice 1. L'exercice 2 c'est vraiment du copier-coller de ce qui a été fait en exercice. Et puis l'exercice 3 aussi. Euh, le 1, c'est le fait de... On l'avait fait pour les asymptotes, mais pas tellement pour les tableaux de signes, par exemple.”* (28)

Ses évaluations prennent parfois une forme normative. D'une part parce qu'elle évalue parfois la pertinence d'une réponse en fonction des réponses des autres élèves: *“puis finalement quand tu vois qu'il y en a plein qui ont très bien compris, il n'y a pas de raison d'être gentil*

*pour ceux qui ont fait faux.*” (92). D’autre part parce qu’elle estime qu’elle peut ajouter des pièges quand elle considère que le sujet est facile : *“Si c’est un sujet facile, il y aura peut-être des petits pièges.”* (58). De plus, comme mentionné plus haut, elle accorde relativement peu d’importance aux objectifs d’apprentissage. Cela renforce encore cette tendance normative.

### **5.5 Portrait de R. (Maths, 10 ans d’expérience dans l’enseignement)**

R. a la volonté de tout tester quand elle fait une évaluation: *“Puis après je regarde si j’ai l’impression qu’il y a tout. Enfin, est-ce qu’on a parcouru toute la matière? Est-ce que tout ce qui a été vu, il y a une partie d’un exercice qui s’y porte”* (7). Elle est donc attentive à l’alignement curriculaire, au moins sur ce point-là, dans le sens où elle a la volonté d’évaluer ce qui a été entraîné. En revanche, elle ne dit pas si elle fait attention à ce que les tâches qu’elle choisit n’évaluent vraiment *que* ce qui a été appris. De même, elle ne se préoccupe pas de la difficulté de ses évaluations. Au contraire, quand on lui pose la question, elle répond : *“Non. Je réfléchis plutôt au style d’exercice.”* (9) et quand on lui parle de piège, elle dit qu’elle n’en met plus beaucoup parce que *“j’ai l’impression que les pièges, ils se les créent déjà eux-mêmes, sans qu’on en rajoute encore.”* (17). À l’inverse, elle ne fait en général pas d’arrangements pour améliorer les notes : *“Je suis pas fan des bonus, comme je suis pas fan d’enlever la moins bonne note ou je sais pas quoi.”* (35). Il y a donc chez elle une volonté de tester ce qui a été enseigné et de le rendre ni plus difficile, ni plus facile que ça ne l’est. Sur ce point aussi, l’alignement curriculaire est respecté. Seule exception à ce principe : elle cherche si elle ne peut pas ajouter une demi-note aux élèves qui sont juste à 3.5: *“En général, les gens qui sont à la limite en-dessous, en général, je reprends la copie et j’essaye de chercher si y’a pas un demi-point que je pourrais...”* (111).

R. se soucie également d’équité dans ses évaluations: elle corrige exercice par exercice, parce que *“ça [lui] semble plus équitable”* (81). Elle compare ses corrections pour s’assurer d’avoir noté tous les élèves de la même manière (87). Elle serait même favorable à ce que les enseignants coordonnent plus leur enseignement entre eux, pour que les élèves soient tous à la même enseigne (121).

Elle peut se montrer très compréhensive à l’égard des élèves. Elle ne pénalise pas les mauvaises présentations: *“j’y prête pas forcément attention parce que il y a des gens qui écrivent mal, je sais pas; c’est naturel qu’ils écrivent mal. Puis voilà. Maintenant, y’a des fois où je me dis quand-même ‘mais non, mais là c’est cochon’ puis j’arrive pas à sanctionner”* (5). Elle tolère, dans une certaine mesure, les fautes de calcul : *“Mais il y a une tonne de*

*fautes de calcul. Alors bon, on va pas toutes les sanctionner parce que sinon tu enlèves tous les points.*” (17) et *“y a parfois une petite faute de calcul mais qui n’a rien à voir avec l’exercice. Et qui change à peu près rien à l’exercice. Alors, en tous cas avec les maths renforcées, ben moi j’appelle ça une faute d’inattention. Je pénalise pas. Pour autant que l’exercice reste identique sans la faute.”* (31).

En même temps, R. est parfois déçue par les efforts des élèves: *“C’est vrai que la classe que j’ai ici, ils sont pas... Je trouve qu’ils se donnent pas beaucoup de peine et ils sont pas forcément très doués non plus.”* (17). A-t-elle des attentes trop élevées ?

### **5.6 Portrait de F. (Maths, 20 ans d’expérience dans l’enseignement)**

F. fait attention à l’alignement curriculaire sur l’ensemble de ses séquences : *“Préparer les évaluations, c’est juste une partie, n’est-ce pas, de tout un cycle qu’on engage avec les élèves, dans lequel j’essaie de me faire une idée assez claire sur le savoir-faire que je cherche à développer chez les élèves, d’abord. Et puis, ensuite en général, comment je vais tester cette acquisition du savoir-faire”* (6)

F. souhaite faire en sorte que les élèves ne se sentent jamais distancés. Ceci non seulement pendant l’évaluation, mais aussi pendant le cours: *“Je veux éviter autant que possible, comme d’ailleurs en général, dans les cours aussi, que l’élève ait le sentiment qu’il est totalement dépassé par les exigences.”* (16) et *“Mais, enfin, nous savons, enfin, tu le sais aussi entre-temps que les étudiants peuvent vivre un véritable enfer si en mathématiques ils sont largués. Et une de mes préoccupations permanentes, c’est de faire en sorte que, en dépit de,... et puis de ne lâcher personne, donc ceci aussi a son implication au niveau de l’évaluation.”* (52). Il met donc dans ses évaluations des questions dont le but est de mettre l’élève en confiance: *“Je prévois en règle générale, une certaine partie qu’on peut qualifier de mise en confiance”*. Il cherche aussi à récompenser ses élèves quand il a le sentiment que ceux-ci ont fait un effort particulier: *“cette envie de marquer le moment où l’élève, manifestement, a fait un travail plus important [...] Alors il a pu m’arriver de surnoter un tout petit peu certaines personnes parce que je voulais, à un moment donné, qu’il sente...”* (52).

F. fait d’autres arrangements, toujours dans le sens des élèves. Il tient compte des problèmes qui ont été bien réussis pour affiner la pondération de ses problèmes: *“j’aurais tendance à gonfler un peu les problèmes qui se sont bien passés pour... pour, disons, ne pas pénaliser inutilement.”* (22). Il compte certains points de son barème en bonus: *“Et, ensuite, j’applique quelque chose qui est assez proche de ce qu’on appelle, je crois, l’échelle fédérale [...] C’est*

*toujours un petit peu enveloppant dans le sens que, si l'élève était capable de répondre à toutes les questions, il aurait en général une note légèrement supérieure à 6: 6,3; 6,4; quelque chose comme ça.*"

Par ailleurs, F. se soucie des apprentissages de ses élèves. Il fait confiance à son jugement professionnel: *"Il est clair que, à un moment donné, cet aspect "compréhension générale" du sujet, qui n'est pas si simple à articuler, où l'on sent que, dans l'ensemble, il sait comment ça se passe. Et il est un peu maladroit au niveau technique, j'en tiens compte [...] j'ai plutôt tendance à ne pas aller trop loin et de garder, non pas un certain flou artistique, mais une certaine marge d'appréciation très liée à ces aspects que je viens de développer un tout petit peu."* (32). Pour autant, en accord avec la remarque de Crahay (2003) mentionnée plus haut, il met dans ses évaluations des questions qui lui permettent de faire la distinction entre les bons et les très bons élèves: *"Ce ne sont pas des problèmes qui pour autant nécessitent de très grands développements, mais où il faut faire preuve d'une élasticité mentale un tout petit peu plus importante."* (18). On peut postuler ici que l'"élasticité mentale" que recherche F. est en fait une qualité personnelle de l'élève qui n'aurait pas été travaillée en classe. Si tel est le cas, ces tâches constituent une rupture d'alignement curriculaire car les élèves n'ont pas appris à les réaliser.

En somme, F. est un enseignant qui se soucie à la fois des apprentissages de ses élèves, et de leur motivation à apprendre.

## 6 Analyse transversale

Dans cette partie de l'analyse, nous traitons des similitudes et différences entre les manières de faire des différents interviewés. Il s'est agi pour nous d'une part de trouver sur quels points un consensus semblait régner et d'essayer d'en comprendre l'origine. D'autre part, les avis partagés nous ont également menés à des hypothèses quant à leur raison d'être et leur signification. Nous séparons ici encore l'analyse selon nos trois focii : la construction de l'épreuve, sa correction et les échelles.

### 6.1 Construction des épreuves

Dans cette partie de l'analyse, notre attention s'est portée sur la présence et la cohérence perceptible de l'alignement curriculaire.

Tous les enseignants interrogés basent principalement leurs évaluations sur ce qui a été traité en classe. F., par exemple, dit, quand il parle de la construction de ses épreuves: *“Préparer les évaluations, c'est juste une partie, n'est-ce pas, de tout un cycle qu'on engage avec les élèves, dans lequel j'essaye de me faire une idée assez claire sur le savoir-faire que je cherche à développer chez les élèves, d'abord. Et puis, ensuite en général, comment je vais tester cette acquisition du savoir-faire, de sorte à ce que l'élève... là, je parle encore de la phase préalable au test à proprement parler... euh, comprenez ce que j'attends de lui. Donc, si tu me demandes comment je prépare les tests, ils sont en général un prolongement assez naturel de ce qui a été fait au cours.”* (F. 6)

Cependant, la majorité de nos interviewés a tendance à poser des questions qui sortent un peu de ce qui a strictement été traité au cours. T., par exemple, dit *“Parce qu'il faut que, pour que l'élève ait 6, il faut qu'il ait compris plus que... enfin, il faut qu'il fasse plus que répéter ce qu'on a déjà fait”* (T. 28). F, Em et C. disent de même. Au contraire, R. et L. ne font pas référence au fait de demander plus que ce qui a été fait au cours. L. précise même qu'elle reprend le genre d'exercices vus précédemment : *“Donc euh, l'idée c'est un petit peu de me baser sur euh... puis sur ce qu'on a fait en classe, donc dans, dans le même ordre d'idée.”* (L. 2). Bien que le fait de reprendre le genre d'exercices réalisés en classe puisse participer d'une cohérence de l'alignement curriculaire, il n'en est en rien le garant. En effet, les tâches effectuées en classe comme les tâches d'évaluation peuvent être en décalage avec les objectifs.

Au niveau des objectifs, F. en parle assez clairement (F. 6), même s'il n'utilise pas ce terme: il parle de savoir-faire qu'il veut atteindre. Il est le seul des six qui fasse vraiment le lien entre objet, objectif, tâche et évaluation. L. fait mention d'objectifs (L. 4) sans entrer dans le détail.

T. utilise le mot “*objectifs*” (T. 14), mais pas dans le sens où nous l’entendons ici. Elle parle en réalité d’alignement entre l’évaluation et les tâches: “*je leur distribue une feuille qui leur résume tous les exercices qu’on a faits, qui leur permet... J’ai refusé d’appeler ça “objectifs” parce que je voulais justement pouvoir sortir un peu de tout ça*”. (T. 14)

C. mentionne qu’elle “*travaille en compétences*” (C. 14). Les compétences sont, en langues étrangères, souvent envisagées comme étant principalement l’expression orale, l’expression écrite, la compréhension orale et la compréhension écrite. Néanmoins, dans ce contexte, il nous semble pertinent de l’interpréter comme une expression partiellement alternative pour “objectifs”, puisque la plupart des objectifs langagiers en cours de langue étrangère mentionnent la compétence visée par l’objectif, sans pour autant que cette compétence en soit la composante principale. Cela n’est pas suffisant pour attester d’un alignement curriculaire. T. dit qu’elle fait attention aux habiletés qui seront utiles pour la suite du parcours scolaire. Elle dit: “*J’aime bien tester ce qui sera utilisé*” (T. 26). Dans ce sens, elle est attentive aux objectifs, mais partiellement seulement, puisque le programme d’une année dépasse le strict minimum nécessaire comme base au programme de l’année suivante.

En conclusion, les enseignants parlent peu d’alignement curriculaire. Ils semblent ne faire de liens qu’entre les tâches et l’évaluation, et même à ce niveau, on peut faire l’hypothèse qu’il existe des ruptures chez certains. Seul F. montre qu’il voit le lien entre les quatre piliers de l’alignement curriculaire (objet, objectif, tâche et évaluation), et pourtant, il montre aussi des signes de rupture entre les tâches et l’évaluation. De plus, les références à l’apprentissage sont rares. Peut-être cette absence est-elle liée à un alignement curriculaire lacunaire, en particulier au niveau des objectifs d’apprentissage : en effet, que peut-on dire de l’apprentissage si l’on n’a pas d’objectif ?

## **6.2 Correction**

L’élément le plus frappant nous semble être la manière de corriger les épreuves: tous les enseignants interrogés ont déclaré corriger leurs épreuves exercice par exercice. C. corrige encore plus horizontalement, “*même parfois une question pour tous les élèves*” (C. 56). Tous semblent s’accorder sur le fait que c’est une question d’égalité et de justice. D’une part, on se souvient mieux de la manière dont on a corrigé précédemment: “*Donc je fais vraiment chaque question pour chaque élève, pour que j’aie exactement l’idée en tête, pour être le plus égal possible*” (C. 56). D’autre part, cela sert à anonymiser les copies, parce qu’une fois qu’on a tourné la première page, on ne voit plus le nom de l’élève qu’on est en train de

corriger: *“Surtout que ça me permet aussi vu que l'exercice 1 est sur la page des prénoms, quand je fais l'exercice 2 je me souviens pas que je corrige tel prénom. Et j'aime bien avoir cette version un peu anonyme. Parce que voilà. Je déteste avoir le sentiment d'avantager quelqu'un parce que l'élève est studieuse ou studieux ou parce qu'elle est agréable, ou à l'inverse désavantager quelqu'un parce que l'élève fait tout sauf ça.”* (T. 84)

Une autre manière de garantir autant que possible l'égalité au niveau de la correction est l'utilisation d'un corrigé évolutif. C, Em. et L. préparent un corrigé à l'avance et y ajoutent peu à peu les réponses inattendues afin de pouvoir s'y référer si elles réapparaissent. C'est également un moyen de pouvoir facilement justifier l'attribution des points auprès des élèves: *“Aussi pour pouvoir argumenter si après il dit “oui, mais j'ai mis ça...” lors de la correction en commun, pour pouvoir dire “mais ça va pas, parce qu'il y a telle et telle chose qui est fausse.”* (C. 60) Il nous paraît relativement logique que ce soient les enseignantes de langues qui utilisent ce procédé de manière systématique, puisqu'il y a dans une évaluation la plupart du temps des questions ouvertes ou dont les réponses sujettes à interprétation. Un corrigé pensé à l'avance et évolutif devient facilement nécessaire si l'on veut garder une trace de ce qui a été accepté ou non.

Pour pondérer leurs épreuves, les enseignants en mathématiques procèdent tous de la même manière: ils font un corrigé, puis le décomposent en étapes. Pour chaque étape, ensuite, ils attribuent un certain nombre de points: *“Donc en gros la factorisation c'est un point, le fait de donner l'ensemble de définition un demi point, le tableau de signes euh voilà. Pour chaque petit truc il y a un demi point”* (T. 64). Ce nombre de points, d'ailleurs, semble en général être attribué de manière subjective: T. dit que *“[c]haque étape ne vaut pas forcément le même nombre de points”* Suite à cela, B postule que *“[...] Par exemple là, cette factorisation, tu mets 1 point. Si tu donnais la même chose à factoriser en première année”*. T confirme que *T - Ce serait beaucoup plus.*” (T. 80-82) Elle laisse entendre qu'une même opération ne vaudra pas le même nombre de points selon le niveau de l'apprenant mais n'explique pas en fonction de quoi elle décide de la valeur de l'opération pour un niveau en particulier.

Après l'attribution des points, F. parle très ouvertement des arrangements évaluatifs. Il en fait quand il veut récompenser un élève qui a fait un effort particulier: *“cette envie de marquer le moment où l'élève, manifestement, a fait un travail plus important [...]. Alors il a pu m'arriver de surnoter un tout petit peu certaines personnes parce que je voulais, à un moment donné, qu'ils sentent...”* (F. 52). Il en fait également quand il pense que les points qu'il a attribués aux différentes étapes de son évaluation ne reflètent pas le niveau général de compréhension de l'élève: *“Il est clair que, à un moment donné, cet aspect “compréhension*

*générale” du sujet, qui n’est pas si simple à articuler, où l’on sent que, dans l’ensemble, il sait comment ça se passe. Et il est un peu maladroit au niveau technique, j’en tiens compte.”* (F. 32)

L. parle de changements dans la pondération en rapport avec la difficulté d’un exercice. Elle ne précise cependant pas les conditions pour que celui-ci soit considéré comme “trop difficile”. C’est le cas de la plupart de nos interviewés, qui nous parlent d’exercices ou de test “*trop difficiles*” (L. 20 ; Em. 286 ; C. 92 ; F 24) ; L. parle aussi d’un test avec un exercice “*peut-être trop dur*”(L. 62), dans tous les cas sans expliquer sur quels critères ils se basent. S’agit-il de tâches se situant au-delà de la zone proximale de développement des élèves ? S’agit-il de tâches faisant appel à des habiletés qui n’ont pas été exercées en classe ? S’agit-il simplement d’exercices qu’une trop grande partie des élèves n’a pas réussi à faire ? R. l’explique très peu et de manière normative : “*Mais ça m’arrive de modifier le nombre de points de l’exercice parce qu’on se rend compte qu’il y en a plein qui ont fait telle erreur, ou qui sont passés par autre-chose et puis voilà*” (R. 69). T. parle également d’arrangements normatifs : “*C’est le genre de petites choses que, une fois, le premier tu te dis que tu es gentil, puis finalement quand tu vois qu’il y en a plein qui ont très bien compris, il n’y a pas de raison d’être gentil pour ceux qui ont fait faux.*” (T. 92) Cette correction normative se retrouve notamment chez L, qui dit que “*quand je mets une note, ça m’arrive d’être un peu parce qu’ici on est au demi donc on peut pas mettre- donc des fois ça m’arrive d’être un peu qu’y ait 5-5½ ou alors 4-4½ et là de me dire ‘ok ben je prends quelqu’un qui a 4, quelqu’un qui a 4½’ puis je me dis ‘ok, est-ce que ça vaut...’*” (L. 80). Elle précise quand-même : “*Ce qui me va pas très bien d’ailleurs.*” (L. 84)

Les arrangements évaluatifs font partie intégrante de l’acte d’évaluation et sont des outils sociaux. Pour Dubus (2006), il y a un risque dans toute évaluation de ne pas promouvoir des élèves compétents et d’en promouvoir des incompetents, ce qui menace l’efficacité de l’épreuve. Nous pouvons lier ce risque avec les règles implicites et explicites qui stipulent qu’on peut faire des arrangements lorsqu’ils sont à l’avantage de l’élève : qu’en est-il si ces arrangements poussent à promouvoir plus d’incompétents ?

### **6.3 Echelles**

Tous les interviewés utilisent le barème fédéral, même si tous ne le comprennent pas de la même manière. Chez les enseignants en mathématiques, la formule utilisée pour transformer les points en note est systématiquement : (nombre de points obtenus/nombre de points

total) $\times 5 + 1$ . Les enseignants en langues choisissent parfois de fixer le seuil de suffisance à une valeur autre que 60%, puis appliquent une échelle affine entre le 1 et le 4, et une autre échelle affine entre le 4 et le 6. À noter que les arrondis font que toutes ces échelles ne sont pas aussi affines qu'elles veulent bien le dire, mais personne ne semble beaucoup s'en émouvoir.

On trouve aussi certains enseignants qui font leur calcul en se basant sur un "nombre de points total" légèrement inférieur à ce qui est réellement présent dans leurs évaluations, afin de favoriser les élèves. F. (maths) dit le faire souvent, tandis que R. (maths) et Em (anglais) disent le faire très rarement, quand elles ont le sentiment a posteriori que leurs évaluations étaient "trop difficiles".

Par ailleurs, beaucoup d'enseignants choisissent le barème fédéral sans se demander pourquoi ils le choisissent. C'est, pour eux, le barème "normal", et il n'est pas remis en question: c'est celui que tout le monde utilise, et c'est aussi celui que les élèves comprennent, disent-ils.

Seuls les enseignants en langues, comme mentionné plus haut, le modifient parfois. Mais même là, ils considèrent le barème fédéral comme "normal", et les autres barèmes comme "pour les tests plus faciles". C. fait exception, car elle exprime son désaccord avec le barème fédéral, qui lui semble trop gentil. Néanmoins, elle l'utilise tout de même pour des raisons de hiérarchie et de cohérence dans la file. L'alignement curriculaire n'est donc pas du tout pris en compte dans le choix des échelles chez nos interviewés.

## 7 Conclusion

Notre question de départ, “comment les enseignants préparent-ils leurs évaluations, en pondèrent-ils les différents éléments, les corrigent-ils, choisissent-ils les barèmes et construisent-ils leurs notes ?” a été discutée en profondeur avec un petit nombre d’enseignants, nous permettant de confirmer notre hypothèse selon laquelle les enseignants auraient tendance à faire des ruptures dans l’alignement curriculaire. Il s’agit en effet d’un nombre limité d’interviewés, ils présentent souvent leurs pratiques comme des habitudes communes aux membres de leurs établissements. Ces habitudes communes font par ailleurs figure d’autorité et il est possible qu’elles rassurent les enseignants sur le bien-fondé de leurs pratiques personnelles.

S’il ne nous est pas possible de prétendre à une compréhension totale et véritable des interviewés, certains éléments peuvent néanmoins être relevés. Malgré les différences au niveau des branches et des années d’expérience, il semble y avoir un certain consensus dans la manière de concevoir, corriger et noter les évaluations chez les enseignants du secondaire II qui ont participé à ce travail. Les points saillants de cette analyse nous semblent être la rupture d’alignement quasi-omniprésente entre la pondération et l’échelle ainsi que parfois entre les items de l’épreuve et la pondération. De plus, en amont, les objectifs paraissent peu voire pas définis et aucun des enseignants interviewés ne montre de signes d’analyse de l’objet d’apprentissage, ce qui rend l’alignement curriculaire difficile. Nous faisons l’hypothèse que ces fréquentes ruptures d’alignement – par ailleurs inconscientes – sont notamment due à des pratiques évaluatives rarement questionnées, justifiées par des phrases du type “c’est comme ça qu’on fait”, à une opacité sur les critères d’évaluation ou encore à un manque de connaissances au niveau des barèmes et des échelles mis en œuvre.

Les interviewés ont tendance, dans l’ensemble, à mettre dans leurs évaluations ce qui a été vu en classe. Mais plusieurs d’entre eux, selon leurs dires, mettent facilement une question légèrement différente de ce qui a déjà été traité, afin de distinguer les bons des très bons élèves. Cela constitue probablement une rupture entre les tâches d’apprentissage et les tâches d’évaluation. Les raisons et effets de cette pratique pourraient également être investigués plus en profondeur : pourquoi veulent-ils faire cette distinction entre les élèves et quelles en sont les conséquences ?

Nous avons dégagé une certaine envie d’être “équitable” des portraits, notamment au niveau de la correction : on n’a pas envie de désavantager des élèves par rapport à d’autres et certaines stratégies entrent alors en jeu. Mais par rapport à quoi doit-on être équitable ? La

réponse nous semble souvent être “par rapport à la norme” : on compte juste une même réponse chez tout le monde. Il serait naturellement bon d’étudier la notion d’“équité” chez les enseignants, car il est probable qu’elle n’a pas la même signification pour chacun et qu’elle peut donner lieu à de nouveaux biais.

Le jugement professionnel apparaît peu dans ces entretiens. On voit beaucoup de points, de mesure, et ceux-ci servent à justifier la note. On peut interpréter cette absence de plusieurs manières : premièrement, on peut penser que les enseignants, consciemment ou inconsciemment, évitent autant que possible de s’éloigner de la mesure afin de pouvoir justifier chaque étape de correction ou de notation. Deuxièmement, on peut postuler que les enseignants évitent volontairement de s’y référer au cours des entretiens, le considérant par exemple plus difficilement explicable que la mesure. Nous avons trouvé notamment des signes de la première hypothèse dans plusieurs entretiens, bien que nous n’ayons pas directement parlé de cela avec les interviewés. Il y a en particulier des références à la manière de justifier les points et la note vis-à-vis des élèves. Ces références se rapportent toutes à la mesure, par exemple le comptage de points ou le barème fédéral. Nous n’avons donc pas la possibilité de confirmer ou d’infirmer l’une ou l’autre hypothèse. Quelles conséquences ce centrage sur la mesure a-t-il sur la considération des apprentissages et des objectifs ? Peut-on ainsi savoir quelles lacunes chaque élève a ?

Tous ces points mériteraient une étude plus approfondie et des entretiens plus pointus. Nous avons beaucoup appris de l’étude de ces six entretiens, mais n’avons pas toujours su réagir directement à certains éléments qui nous ont par la suite intrigués.

Il serait également intéressant de comparer les points de vue des enseignants avec ceux des élèves : que veut dire un 4 ou un 6 pour les élèves ? Les enseignants font-ils vraiment passer le message désiré lorsqu’ils veulent “motiver” les élèves ? Cette étude n’est en effet qu’un modeste premier pas dans l’étude de la notation au secondaire II. Mais elle ouvre un questionnement à nos yeux passionnant.

[MCours.com](http://MCours.com)