



HAL
open science

Mettre l'évaluation sommative au service de la réussite des élèves

Valérian Goth

► **To cite this version:**

Valérian Goth. Mettre l'évaluation sommative au service de la réussite des élèves. Education. 2019.
dumas-02289251

HAL Id: dumas-02289251

<https://dumas.ccsd.cnrs.fr/dumas-02289251>

Submitted on 16 Sep 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Année universitaire 2018-2019

Master MEEF

Mention 1^{er} degré

2^{ème} année

Mettre l'évaluation sommative au service de la réussite des élèves – CM1

Présenté par : Valérian GOTH

Encadré par : Sandrine Meylan

Mots Clefs : évaluation sommative – réussite – égalité des chances

Ce mémoire est dédié à :

Anjay Rambhowan, futur grand professeur – et le meilleur binôme avec lequel il m'était possible de commencer ;

Mélanie Cueto, pour ses conseils précieux, son implication et sa bienveillance. C'est à elle que je dois une telle évolution dans mes pratiques durant cette année ;

Hélène Koch et à Samia Hatri. Pour leur bienveillance également, la qualité de leur accompagnement et leur soutien moral tout au long de l'année. Le tutorat a été un lieu d'échange et d'expression salvateur durant cette année de titularisation ;

Sandrine Meylan, pour avoir accepté mon sujet de mémoire – sans grand lien avec les sciences ! – , son ouverture à nos propositions, son implication dans nos travaux, sa disponibilité sans failles. Les premiers retours sur un travail personnel sont toujours déterminants ; elle m'a donné la confiance d'aller jusqu'au bout ;

à l'équipe d'*Alexandre Dumas*, dans le onzième, et tout particulièrement à Marie-Claire Chitteray. La qualité de son accueil, ses conseils, nos conversations durant l'année... Tout cela a joué à plein pour m'aider à me sentir à l'aise dans un univers inconnu où l'on est, sans cela, facilement livré à soi-même. Il est bien inquiétant, pour les futures générations de stagiaires, que les directeurs d'école soient amenés à disparaître...

Sommaire

Introduction	1
Partie 1 : L'évaluation, une violence ordinaire	2
1. La note, un outil d'une valeur discutable	3
a) Vers une notation objective : le caractère aléatoire de la note	3
b) La note <i>n'est pas</i> l'évaluation	4
2. Pour une évaluation réellement informative...	6
a) Les biais de notation	6
b) L'effet évaluation	7
c) Un outil limité	8
3. ... et formative !	9
a) Une évaluation qui fait obstacle aux apprentissages	9
b) L'évaluation sommative : un maillon d'une chaîne beaucoup plus large	11
Partie 2 : De l'évaluation « au jugé » à la conception d'évaluations sommatives pertinentes : compte-rendu d'une année de travail	13
I. Les problématiques liées à la conception de l'évaluation sommative	14
1. Le contexte	14
2. Analyse	15
a) Une évaluation non congruente	15
b) Aux prises avec la correction : découverte de ce que j'évalue effectivement et de l'inégalité de traitement entre les compétences évaluées	17
c) Cas précis de correction : la question de la justice scolaire	18
d) Bilan	20
II. Les problématiques liées à la conception de l'évaluation sommative	23
1. Limites et points morts des observables	23
2. La grille de compétences : un outil holistique	24
a) Permettre une évaluation réellement congruente	24
b) La notation revisitée	26
c) Une communication de qualité	27
3. Pallier l'effet évaluation	30
Conclusion	33

Introduction

Ce mémoire est né d'un constat. Fraîchement débarqué dans l'univers du professorat, j'ai été profondément marqué de découvrir que l'école, dont le but est de permettre l'égalité des chances, est de fait le premier acteur dans la création des inégalités. Professeur d'une classe de CM1 à l'école Alexandre Dumas, dans le onzième arrondissement de Paris, j'ai pu constater que le nombre d'élèves par classe – 28 dans mon cas –, l'inexpérience et la pression imposée par la masse des programmes amenaient mécaniquement à une forme d'abandon des élèves. Le constat était là : dans l'absolu, *tous* pourraient réussir : il ne s'agit que d'une question de moyens et de temps mis à leur disposition. Dans les faits cependant, l'institution chargée de permettre cette réussite provoque systématiquement des échecs, puisqu'il y a, toujours, des élèves laissés pour compte. A cet égard, le fait que je sois débutant était évidemment un facteur aggravant. Je me suis donc interrogé : sur quelle partie de ma pratique pouvais-je me pencher pour lutter contre les inégalités induites par le système éducatif français? Il m'est apparu que la problématique de « réussir ou échouer » était organiquement liée à celle de l'évaluation sommative, pour plusieurs raisons.

Tout d'abord, c'est elle qui dialectise l'éducation autour de la question de la réussite. Le moment sommatif est en effet celui où l'enseignement ne se suffit plus en tant que pratique et doit être testé, mesuré, quantifié... pour rendre compte de son efficacité. Sans dimension sommative donc, pas de question de « réussite » ou « d'échec ». Il apparaît à cet égard que l'évaluation sommative n'est pas un point anecdotique de la pratique professorale mais bien l'un des principes organisateurs de l'ensemble du système scolaire actuel. Toutes les autres formes d'évaluation – formatives, certificatives, normatives, diagnostiques... la liste est longue – y sont liées d'une façon ou d'une autre, soit qu'elles permettent de mieux savoir quand la déclencher, soit qu'elles en orientent la fonction – sanctionner un niveau d'étude, classer les élèves dans le cadre d'un concours, etc...

Le second point à noter est que l'évaluation sommative, en déterminant le degré de réussite des élèves, entre en contradiction avec les buts de l'école. La réussite de tous est en effet difficilement compatible avec un organe de classement, dont l'un des critères définitoires est la détermination d'une norme en deça de laquelle l'élève sera considéré comme « en situation d'échec ». Il apparaît ainsi que l'appropriation progressive des idéaux démocratiques par l'école, au fil des deux siècles qui ont suivi la révolution, a pour conséquence de la mettre aujourd'hui dans une situation de tension interne. Cette tension peut se traduire par la question suivante : *à quelles conditions l'évaluation sommative favorise-t-elle la réussite des élèves ?*

Cette problématique a guidé ma réflexion durant cette année de PES. Afin d'en rendre compte, je commencerai par présenter les problèmes que soulève l'évaluation sommative. Je montrerai ensuite comment ma pratique a évolué au fil du temps, en partant d'un exemple précis : une évaluation ratée, qui m'a permis de déterminer les indicateurs à prendre en compte pour concevoir une évaluation correcte. Je me pencherai alors sur les transformations que j'ai impulsées à partir de là et sur leurs résultats.

PARTIE 1

L'évaluation, une violence ordinaire

1. La note : un outil d'une valeur discutable.

a) Vers une notation objective : le caractère aléatoire de la note.

Appréciation générale, note sur vingt, lettre placée plus ou moins loin du début de l'alphabet... Voici quelques moyens, parmi la pléthore de ceux existants, utilisés par les professeurs des écoles pour exprimer le résultat d'une évaluation sommative. En effet, s'il est un corollaire inévitable à l'évaluation, et qui lui est organiquement liée, c'est bien *l'expression de son résultat*. Mais pourquoi celle-ci a-t-elle donné naissance à une telle batterie d'outils ? Dans son ouvrage *L'évaluation à l'école*¹, Charles Hadji explique que la note chiffrée a prévalu pendant longtemps. Héritière du système de l'enseignement secondaire, elle poursuivait avant tout un but élitiste : parvenir à classer les élèves en mesurant leur « valeur scolaire ». Avec les différentes étapes de démocratisation de l'école ayant abouti à un cursus unique en 1959 ce type de notation a d'abord été conservé, avant d'être remis en cause puis largement éliminé du premier degré. Pourquoi ?

Hadji note que la docimologie, la « science de la notation », née d'abord précisément pour prouver que le système de notation était fiable, a mis en lumière le caractère absolument arbitraire de celui-ci. Dès 1936, un examen portant sur cent copies du baccalauréat propose des résultats édifiants : en fonction du jury auquel elles sont présentées, les copies présentent jusqu'à treize points d'écart en composition française... et neuf points en mathématiques ! *La réussite à l'examen apparaît donc comme extrêmement fluctuante en fonction du jury*. Autre expérience, encore plus édifiante : en 1967, quinze copies du baccalauréat ayant reçu la même note sont confiées à quarante examinateurs différents pour une nouvelle correction, sans que ceux-ci soient au courant de la note initiale. Résultat : chaque copie reçoit quarante notes différentes... Et il y a évidemment de grandes inégalités de notes au sein du tas de copies de chaque évaluateur. Il apparaît donc que la note chiffrée, sensée refléter un résultat objectif, est en fait extrêmement aléatoire.

Et ce n'est pas tout. Parmi les évaluateurs initiaux de l'expérience ci-dessus, il a été proposé à certains de recorriger les mêmes copies après un délai donné – échelonné de douze à dix-neuf mois. Sur les 210 copies ayant ainsi fait l'objet d'une nouvelle correction, 92 ont obtenu une note différente de la première fois ! Il existe donc non seulement une fluctuation importante de la note en fonction de l'évaluateur, *mais également en fonction du temps pour le même évaluateur* ! Nous pourrions facilement multiplier les exemples : toutes les expériences de corrections multiples vont dans le même sens...

Ce problème de validité de la note soulève la question de la *justice scolaire*, puisque ce sont bien les différentes évaluations jalonnant le parcours de l'élève qui infléchiront son évolution au sein du système scolaire et, *in fine*, détermineront son intégration dans la société. Il interroge également la question de la *mesure* : *sous quels critères et dans quelles conditions une*

¹ *L'évaluation à l'école, « pour la réussite de tous les élèves »*, Charles Hadji, Ed. Nathan, Paris, 2015.

note peut-elle exprimer une mesure JUSTE de ce que l'on évalue ? Ces questions soulevées par la docimologie depuis un siècle sont toujours aussi pertinentes aujourd'hui : en effet, si la notation chiffrée a été largement remplacée dans le primaire par d'autres types de notation, les problématiques sont sensiblement les mêmes. Ainsi que le note de Vecchi : « Si les terminologies ont changé, les pratiques restent les mêmes. « Evaluations » ou « contrôles », il s'agit toujours de classer et de donner des notes. »² Qu'est-ce qui permet en effet, de définir la limite précise entre un *tb* ou un *b* sur une évaluation sommative ? (ou la limite entre un *acquis* et un *en cours d'acquisition* ... la question concerne tout type de notation) En vérité, la transition de la note chiffrée – Hadji parle de *score* – à la note de synthèse n'est un élément de résolution du caractère aléatoire de la notation que si elle amène une plus grande objectivité dans l'évaluation. Puisque le but de l'évaluation est de mesurer le degré d'acquisition des connaissances des élèves et de le donner à voir aux parents³, la question principale à laquelle il faut répondre est de « *savoir comment enregistrer et constater le plus objectivement possible un état et un niveau de progression chez les élèves* ». Il s'agit également de restituer cette constatation, non seulement dans une logique de communication avec les parents mais pour permettre aux apprenants eux-mêmes de se ressaisir de leurs apprentissages : « la juste place de l'évaluation est celle d'une prise d'information qui éclaire le professeur sur les effets de son activité didactique, et l'élève sur ses apprentissages. »⁴ La problématique fondamentale qui découle de tout cela est donc : *à quelles conditions la note est-elle un outil réellement informatif ?* Cela demande de se pencher de plus près sur la nature même de la note.

b) La note n'est pas l'évaluation

A partir de quand, sur une copie, une compétence ou connaissance est-elle considérée comme acquise ? Que sanctionne un *b* en haut d'une copie où plusieurs exercices ont été réalisés avec plus ou moins de succès ? Est-ce qu'il signifie que les compétences mises en jeu sont globalement acquises ? Que certaines le sont et d'autres non ? Ou bien encore est-ce qu'en faisant la « moyenne » des exercices réussis et ratés, ou de l'ensemble des réponses justes et erronées, le professeur arrive à un résultat médian ? Autrement dit, à quelle *logique* renvoie la note de synthèse en haut de la feuille et de quelles informations est-elle le vecteur ?

Comme le note André Gagneux : « l'évaluation ne se résume pas à la synthétisation (note, couleur, lettre), mais doit apporter de réelles informations, utiles tant à ceux qui apprennent qu'à ceux qui les aident à apprendre. »⁵ En fait, il est important de comprendre que la notation *n'est pas* l'évaluation, mais uniquement le moyen d'en rendre compte de façon à rendre celle-ci « lisible » et « compréhensible par les familles »⁶. Or Hadji fait remarquer que le seul point sur lequel la note informe, c'est en général sur le classement de l'élève ou sur son degré de réussite.

² *Evaluer sans dévaluer, et évaluer les compétences*, Gérard de Vecchi, Hachette, Paris, 2011

³ Circulaire de rentrée 2014, URL : [https://www.education.gouv.fr/pid285/bulletin_officiel.html?cid_bo=79642], consulté le 29 janvier 2019

⁴ Hadji, *op.cit.*, p.99

⁵ *Evaluer autrement*, André Gagneux, PUF, coll. Education et formation, Paris, 200, p.30

⁶ Circulaire de rentrée 2014, *art.cit.*

Pris dans la tourmente de « l'évaluationniste », le professeur ne fait souvent de la note *qu'une sanction d'un certain niveau de performance à des exercices scolaires*⁷, ce qui revient à faire d'elle un score et non pas une synthèse des acquis⁸. Prise dans ce sens, la note n'est plus un outil au service de l'élève et des parents mais bien un moyen de classement et de sélection dont la valeur informative est extrêmement faible.

Comment y échapper ? Hadji fait remarquer qu'il n'y a pas de « juste mesure » d'un travail scolaire pour la simple raison que les objets de l'évaluation – les connaissances, capacités et compétences – font encore aujourd'hui l'objet de débats houleux quant à leur définition. Il ajoute même que: « la poursuite de la note vraie est chimérique tout simplement parce que, contrairement à ce que laisse croire la notation, l'évaluation n'est pas une opération de mesure ».⁹ Mesurer consiste en effet à « attribuer un nombre à un objet selon des règles logiquement acceptables »¹⁰ ; c'est ce que fait la note en attribuant un *score* à des exercices servant au professeur *d'indicateurs*. Pour autant, ce *n'est pas* ce que fait l'évaluation, qui est une action d'un ordre *qualitatif* et non *quantitatif*. Il est par exemple notable que deux 12/20 ou deux *en cours d'acquisition* obtenus par deux élèves différents lors d'une évaluation sommative peuvent signifier deux niveaux d'acquisition *très différents* pour le professeur. L'un des élèves peut par exemple s'en être bien sorti en termes de score tout en faisant montre d'une lacune importante sur une notion précise de la leçon. L'autre peut avoir globalement saisi toutes les notions mais manquer d'entraînement et donc faire des fautes dues à la nécessité d'une consolidation plus importante. Dans les deux cas, le seul symptôme observé sera la *note*, qui permettra d'apprécier la « valeur scolaire » de l'élève ; pour autant, ces deux notes égales ne signifient absolument pas la même chose en termes d'acquisition. Il est donc clair que si la notation est bel et bien une opération de mesure, elle est *parallèle* à l'évaluation et n'en rend pas compte *automatiquement* de façon juste.

De là découlent trois conclusions qui peuvent faire office de prémices nécessaires à une évaluation réellement *informative* et non pas purement *clivante* : 1 – il est important de comprendre que la note *n'est pas* l'évaluation mais la mesure d'un score à des épreuves prises comme indicateurs *d'un degré d'acquisition*. La note ne rend donc compte *que de ce dont ces indicateurs rendent compte ; de plus, elle contient ses limites dans sa forme même (une lettre, un nombre... ne donnent pas les nuances nécessaires à l'appréciation complète du niveau d'assimilation d'une notion)* 2 – il est donc nécessaire de *rendre clair*, pour les élèves comme pour les parents, ce dont ces indicateurs rendent compte, afin de rendre lisible ce à quoi renvoie la note : ni à la pure « valeur scolaire », ni à l'acquisition réelle et définitive d'une compétence, mais à la réussite à un moment T d'une certaine action dans des conditions données. 3 – il incombe au professeur de faire en sorte que la note ne renvoie *effectivement qu'à ces seuls indicateurs rendus visibles pour les parents et les élèves*. Nous allons voir que cela est beaucoup plus complexe qu'il

⁷ « Les trous noirs de l'évaluationniste », Pierre Frackowiak, URL : [\[http://www.cafepedagogique.net/lexpresso/Pages/2009/03/FrackLestrousnoirs.aspx\]](http://www.cafepedagogique.net/lexpresso/Pages/2009/03/FrackLestrousnoirs.aspx), consulté le 18 février 2019

⁸ Hadji, *op.cit.*

⁹ *Ibid.*, p. 25

¹⁰ *Ibid.*

n'y paraît et se traduit par un travail conséquent en termes de *conception* et de *mise-en-œuvre* de l'évaluation.

2. Pour une évaluation réellement informative...

a) Les biais d'évaluation

Imaginons un instant une évaluation sommative parfaitement réalisée. Les exercices, équilibrés avec attention et expertise, renvoient chacun à des notions précises. Sans être trop simples ni traîtres, ils forment des indicateurs clairs qui permettent d'apprécier l'acquisition globale de compétences et de connaissances ciblées. Le barème, explicité, sépare chacune d'elle dans un tableau à la fois précis et bien organisé, pour éviter une note globale ne donnant aucune information. Il n'en reste pas moins que cette évaluation va être corrigée par un professeur, c'est-à-dire, *par un être humain, traversé par des émotions, fruit de conceptions diverses et contradictoires, animal social, sujet à la subjectivité et aux sautes d'humeur, bref, tout ce qu'il y a de moins « objectif »*. Or, toutes les études¹¹ montrent que les professeurs sont influencés par divers biais qui viennent contrarier la juste notation : il ne suffit donc pas que l'évaluation-objet soit bien conçue mais encore de prendre conscience de certains effets pervers inévitables.

Le CAIRN montre par exemple que les stéréotypes scolaires et sociaux jouent un grand rôle dans l'appréciation d'une copie¹² : les élèves issus d'un milieu socio-économique plus élevé ont tendance à être mieux notés ; de même un « bon » élève sera surévalué par rapport à un élève « moyen », même si les réponses sont sensiblement les mêmes. Plus encore : un élève redoublant ou souvent en retard en cours aura tendance à récolter des notes plus mauvaises qu'un élève de l'âge « normal » ou « à l'heure » ; et un élève dont le comportement ne convient pas au professeur subira le même sort. En d'autres termes, les conceptions du correcteur sur son élève interviennent au moment de l'évaluation, même à un niveau inconscient. On parle pour cela d'effet « d'assimilation », car la performance de l'élève est évaluée en collusion avec un certains nombres de données extérieures qui jouent grandement sur la note. Il ne s'agit plus d'évaluer une performance, encore moins l'acquisition de compétences et connaissances, mais de statuer sur la « valeur globale » de l'élève telle que la perçoit le correcteur.

D'autres causes, purement techniques, interviennent également dans le processus.¹³ Ainsi, l'ordre d'une copie dans le paquet viendra mécaniquement jouer sur son évaluation par le correcteur : les premières copies ont tendance à être surévaluées (sauf la première !) et les dernières du paquet plus sévèrement notées. L'effet de contraste est également inévitable : toute

¹¹ « Docimologie critique : des difficultés de noter des copies et d'attribuer des notes aux élèves » Institut de Recherche sur l'Éducation de Bourgogne, URL: [http://iredu.u-bourgogne.fr/images/stories/Documents/Cours_disponibles/Demeuse/Cours/p6.4.pdf], consulté le 12 mars 2019 ; « L'évaluation par les notes : quelle fiabilité et quelles réformes ? », CAIRN, URL : [<https://www.cairn.info/revue-regards-croises-sur-l-economie-2012-2-page-218.html>], consulté le 14 mars 2019 ; Hadji, *op. cit.*

¹² « L'évaluation par les notes : quelle fiabilité et quelles réformes ? », *art.cit.*

¹³ Hadji, *op.cit.*, pp.18-20

copie vient après ou avant une autre ; ainsi, une « bonne » copie sera globalement surévaluée après une « mauvaise » et inversement. Enfin, le simple fait que la correction soit un processus temporel intervient dans l'équation : une erreur du même type a tendance à être plus sévèrement notée dans la première partie de la copie que dans la seconde, du fait que le correcteur n'a pas les mêmes attentes au début d'une copie qu'au fur et à mesure qu'il la parcourt effectivement et prend acte de ce qui s'y trouve ; d'une certaine façon, il opère un rééquilibrage entre ses attentes et la réalité des réponses données, ce qui montre bien que *même des critères extrêmement objectifs de notation évoluent au contact effectif des objets à noter.*

b) L'effet évaluation

Un deuxième point très important à noter est que l'évaluation en elle-même n'est pas neutre. Malgré les témoignages de professeurs pour qui l'échec n'a aucun impact sur la motivation générale de l'élève¹⁴, l'impact psychologique de l'évaluation sommative apparaît comme déterminant pour la majorité des élèves.¹⁵ D'une part, le simple fait qu'il s'agisse d'un moment particulier de vérification, hors de la continuité temporelle « normale » de la classe, a un impact sur la qualité des réponses apportées : les élèves, stressés, peuvent voir leurs performances fortement impactées par la confrontation avec la copie et avec l'image qu'ils se font de leurs propres capacités. Bien au contraire, certains peuvent parvenir à faire illusion sur leur maîtrise réelle en répondant correctement par hasard ou en s'appuyant sur des compétences qui ne sont pas celles visées par l'exercice. Comme le note Hadji : « Dans un examen ou un concours, on n'évalue pas seulement les connaissances, mais aussi la gestion de la tension nerveuse (stress), la stratégie (impasses, courses aux points), la chance... et même l'ingéniosité (tricher sans se faire prendre !). »¹⁶ Ainsi, d'une façon ou d'une autre, l'évaluation sommative en « faisant événement » vient modifier les performances qu'elle est sensée enregistrer, parfois de façon très marquée.

De plus, il n'existe pas d'exercice permettant d'informer parfaitement sur l'acquisition d'une compétence ou d'une connaissance à un moment T¹⁷ : les QCM donnent globalement un résultat de 10/20 si on y répond au hasard, les exercices à trou permettent de noter la capacité à restituer exactement la leçon mais en aucun cas d'évaluer sa réelle compréhension ou la capacité à manipuler les notions qu'elle sous-tend, une multiplication ratée peut-être due à une méconnaissance d'une table mais la technique être totalement acquise... Nous pouvons étendre ce principe à la majorité des exercices. Il est donc nécessaire de faire acte des limites de la pratique sommative et de borner très clairement *ce que l'on évalue effectivement* : d'une part pour qu'élèves et parents sachent *ce sur quoi la note informe*, d'autre part pour se concentrer soi-même sur les indicateurs de l'acquisition d'une compétence ou d'une connaissance les plus pertinents.

¹⁴ De Vecchi, *op.cit*, p.21

¹⁵ De Vecchi, Hadji, Gagneux, *opus cités*

¹⁶ De Vecchi, *op.cit*, p.77

¹⁷ *Ibid*, pp.82-85

c) Un outil limité

Il apparaît donc que l'évaluation sommative, du fait de son essence temporelle et de sa corrélation avec l'activité de correction dont elle ne saurait se libérer, *provoque* des effets qui vont *jouer sur le résultat de cette correction*. De plus, son caractère « spécial », tout-à-fait perçu par les élèves, amène à une modification de leur performance qui peut l'empêcher d'informer de façon juste sur le degré d'acquisition des compétences et connaissances envisagées. Enfin, les exercices mêmes évaluent non pas les compétences et connaissances visés mais des compétences considérées comme *indicateurs suffisants*. La question de *ce que ces indicateurs indiquent effectivement* doit par conséquent être pensée si le professeur ne veut pas se retrouver à avoir évalué tout autre chose que ce qu'il visait. Prendre en compte l'imperfection intrinsèque de l'objet « évaluation sommative » est donc absolument fondamental si l'on ne veut pas qu'il annule lui-même sa pertinence. Il en découle :

1 – qu'il est nécessaire de mettre en place des outils permettant de pallier au *maximum* à cette imperfection, de compenser le plus possible « l'effet évaluation », de façon à *effectivement évaluer objectivement l'ensemble des copies qui nous sont confiées et à se rapprocher le plus possible d'une réelle évaluation des compétences et connaissances visées*.

2 – que ces outils passent notamment par la mise-en-place de grilles d'évaluations très précises, explicitant les critères d'évaluation, aussi bien pour répondre à la nécessité informative décrite plus haut que pour *permettre au correcteur une certaine objectivité dans la correction*. Nous notons donc qu'il s'agit bel et bien d'assumer la divergence entre notation et évaluation, en prenant acte du fait que certains exercices peuvent être réussis par les élèves *même si le professeur croit déceler que certaines compétences ne sont pas acquises*. En d'autres termes, la recherche d'objectivité dans la notation amène l'évaluation sommative à se montrer comme objet subjectif et imparfait par rapport à ses propres objectifs : la note ne peut pas pleinement *et* « être lisible » *et* « rendre compte de l'acquisition des compétences »¹⁸ dans le même temps.

3 – La subjectivité et les effets de biais ne pouvant être totalement effacés – ou prenant du moins acte qu'il serait difficile au correcteur lui-même d'en juger – et ayant enregistré l'incapacité de l'évaluation sommative à remplir tous les objectifs de la pratique évaluative tels que définis par la circulaire de Luc Ferry, il apparaît qu'il faut replacer ce moment de l'évaluation dans un contexte plus large. Le sommatif n'est qu'un « moment » de la pratique évaluative, et il faut faire attention de ne pas tomber dans l'écueil d'en faire le *seul* temps de vérification mais bien *un* parmi d'autres. Et, s'il est important que l'évaluation sommative véhicule des informations aux parents et aux élèves sur le niveau d'acquisition des compétences afin de ne pas être simplement un outil de classement, elle doit avant tout informer le *professeur*, afin de guider la suite de son enseignement. En d'autres termes, la pratique sommative doit être pensée à l'aune de l'action évaluative dans son ensemble, action dont la finalité n'est pas le *classement* ni

¹⁸ Circulaire 2014, *Art.cit.* Il apparaît donc que la présence de la note est *en soi* un empêchement à la délivrance la plus précise possible du résultat de l'évaluation, qui devrait en fait passer par des commentaires détaillés et personnalisés pour chaque élève et évacuer la sanction du « score ».

la *vérification* mais bien *l'information* pour permettre *l'acquisition* par l'élève des compétences et connaissances du programme.¹⁹

3. ... et formative !

a) Une évaluation qui fait obstacle aux apprentissages

Nous venons de voir que l'effet évaluation provoquait des effets à la conséquence paradoxale : le professeur n'évalue plus les compétences et connaissances des élèves mais les effets de l'évaluation elle-même ! Pour autant, cela resterait un problème raisonnable s'il ne s'accompagnait d'un autre corollaire, qui est de faire obstacles aux apprentissages eux-mêmes. Comme le fait remarquer Gérard de Vecchi, la très grande majorité des témoignages d'adultes sur leur vie d'élèves porte sur des appréciations négatives liées certes à l'attitude du professeur envers eux, mais aussi évaluations diverses qui ont jalonné leur parcours.²⁰ Génératrices de stress, elles pouvaient mener les élèves à développer une anxiété envahissante, dépassant largement le seul moment de l'évaluation proprement dite.

En fait, il apparaît que l'impact des notes et des appréciations est fondamental pour l'élève : c'est uniquement à leur aune qu'il obtient un retour sur sa « valeur scolaire », c'est-à-dire, finalement, sur sa valeur intrinsèque puisque la grande majorité du temps de vie d'un enfant s'organise autour de l'école. C'est là le sens global des appréciations de Hadji concernant l'évaluation sommative : d'un côté, la « valeur scolaire » est un critère sans pertinence, puisqu'il s'agirait d'une sorte de donnée renvoyant à l'essence de l'élève et que le contexte de classe, notamment l'évaluation, permettrait de mettre en lumière. Or, il montre bien que cette « valeur » est en fait le fruit du contexte ; pire, celui-ci la *créé* en en organisant les conditions d'existence.²¹ De l'autre côté, il note que cette valeur, pure projection d'un cadre contextuel et des agents qui l'incarnent – les acteurs de l'éducation scolaire et en premier lieu le professeur des écoles – a *tout de même un effet sur celui sur qui est plaquée cette projection*²². Comme il le dit : « *on peut peser sur des performances simplement en désignant un élève comme bon ou mauvais, les sujets inscrivant leur conduite dans le système d'attente spécifiquement engendré.* »²³ De Vecchi le note de son côté : « Les élèves se sentent sans cesse jugés en tant que *personnes*, même quand ce n'est pas l'intention du maître qui ne veut juger que leur *travail*. Et, en même temps, l'enseignant doit amener chaque enfant et adolescent à avoir confiance en lui et à devenir autonome ! »²⁴ Il apparaît donc que l'aspect sommatif de l'évaluation à l'école, du fait qu'il émet un jugement de valeur sur l'élève en définissant sa plus ou moins bonne adéquation avec les attentes du système scolaire, a un impact sur l'image globale qu'il se fait de lui en tant que personne.

¹⁹ Circulaire 2014, *Art.cit.*

²⁰ De Vecchi, *op.cit.*

²¹ Hadji, *op.cit.*, p.28

²² De Vecchi arrive au même constat en prenant pour exemple l'imortance de l'effet pygmalion chez les animaux sociaux, *op.cit.*, p.21

²³ Hadji, *op.cit.*, p.24

²⁴ De Vecchi, *op.cit.*, p.8

Dans le meilleur des cas, cela peut stimuler son ego et l'amener à vouloir reproduire indéfiniment des performances égales afin de maintenir son image dans la durée. Bien souvent, le constat, déjà discutable, est plus problématique puisqu'il installe durablement des failles narcissiques chez un individu encore très dépendant du regard de l'autre, failles qui vont potentiellement entraver sa réussite et son développement. A cet égard, certains témoignages sont édifiants : « quand les notes tombent, c'est comme si leur personne rétrécissait, et n'étaient plus vues qu'à travers ces chiffres rouges ou noirs écrits sur une copie. Et qui prennent souvent des proportions dramatiques. »²⁵ Cette *violence* de l'évaluation²⁶, inhérente au système scolaire, ne doit pas être écartée si l'on veut pouvoir y pallier de façon efficace en vue de permettre une réelle « *formation de la personne et du citoyen* »²⁷. C'est ainsi que l'on peut comprendre Sylvain Connac lorsqu'il affirme que « si à la fin d'une année scolaire, un élève a – je ne dis pas grandi – mais au moins n'a pas *diminué*, alors on peut considérer qu'on a réussi notre travail. »²⁸

Enfin, et de façon peut-être moins dramatique, de Vecchi observe l'impact global de l'évaluation sommative sur la formation de l'élève. Il apparaît que les enfants comprennent très tôt l'importance donnée à l'évaluation et à la notation dans le système scolaire et s'y adaptent en conséquence. Ainsi, plutôt que de se concentrer sur l'apprentissage des connaissances et l'acquisition des compétences, le « métier d'élève » tourne-t-il bien souvent autour d'une question simple : *comment avoir de bonnes notes aux évaluations ?*²⁹ On voit d'emblée comme tout est faussé par ce glissement. L'école n'étant plus pensée que dans l'optique de « la préparation à la prochaine épreuve »³⁰, les élèves organisent l'ensemble de leurs apprentissages en fonction de cela. Il en découle que la mémorisation des leçons n'est effectuée que pour permettre la meilleure performance le jour J, souvent au détriment d'une maîtrise sur le long terme ; que le développement de stratégies de réussite prime sur la recherche de la maîtrise réelle ; et que l'évaluation, dont le travail est d'enregistrer le degré d'intégration des savoirs et savoir-faire, produit une distorsion de l'ensemble du système dont elle se veut le témoin. Or, et comme le rappelle à toutes fins utiles de Vecchi : « la finalité de l'école devrait être d'apprendre. » Il importe donc de replacer la question de *l'apprentissage* au centre des pratiques et de réfléchir aux moyens de mettre l'évaluation *au service* de l'acquisition des savoirs, plutôt

²⁵ *Ibid.*, p.69, cité d'Odile Omblard, directrice adjointe d'Okapi, dans une interview donnée au journal Lacroix le 02/06/04

²⁶ Violence déjà reconnue dans la circulaire de 2014 : « Aujourd'hui, de l'école primaire au lycée, certaines pratiques d'évaluation rendent difficilement compte des progrès des élèves et de la nature de leurs erreurs. Elles induisent parfois des classements perçus comme stigmatisants et dévalorisants ; elles peuvent donc porter atteinte à l'estime de soi et générer un découragement face au travail. Cette « pression » est révélée lors des évaluations du programme Pisa : les élèves français ne répondent pas et manquent d'initiative par peur de se tromper et de perdre des points », *art.cit.*

²⁷ Troisième domaine du socle commun de connaissances, de compétences et de culture.

²⁸ « L'altruisme et la fraternité par la coopération avec les élèves », URL : [<https://www.youtube.com/watch?v=Giiim9M1Xaeo>], consulté le 25/03/2018. Je rapproche ces termes de ceux tenus par une directrice d'école maternelle du onzième arrondissement lors de mon stage de M1, qui disait « La première mission du professeur, ça n'est même pas l'éducation, c'est d'assurer la sécurité de l'enfant ; sécurité physique et sécurité *psychologique*. »

²⁹ De Vecchi, *op.cit.*, p.15

³⁰ *Ibid.*

que ceux-ci soient au service de l'évaluation. Comme le note Gagneux, « l'examen ne doit pas perturber, au point de paralyser les apprentissages... qui seront nécessaires à sa réussite ! »

b) L'évaluation sommative : un maillon d'une chaîne beaucoup plus large

A cet égard, un autre rappel de Gagneux est intéressant. Il réaffirme en effet qu' « avant d'être celui qui opère un bilan sommatif à des fins de certification sociale, l'enseignant est celui qui évalue quotidiennement les prestations et produits de ses élèves pour mieux accompagner les progressions. »³¹ Or, c'est un point important : au-delà de la question des notes, qui servent avant tout à remplir les livrets à l'heure et à rendre compte aux parents du travail fait en classe, il est important de se demander à *quoi nous sert réellement l'évaluation sommative*. Si on la considère comme l'une des formes particulières du processus global d'évaluation qui est la tâche du professeur, alors une esquisse de réponse apparaît.

Hadji note en effet que l'évaluation est une activité qui prend son sens dans le cadre d'une dynamique de régulation. Or, « réguler signifie intervenir dans le déroulement d'un processus pour (...) le conduire vers un but. L'action enseignante relève de ce (...) cas : on veut accompagner l'élève vers la maîtrise de connaissances et de compétences. »³² Il apparaît donc que l'évaluation sommative, qui enregistre précisément si le *but* a été atteint, est un point nodal du processus d'apprentissage, puisqu'elle permet de prendre acte de la réussite de l'enseignement et d'ouvrir vers des pistes de remédiation adaptées. Comme l'indique la circulaire de 2014, « *L'évaluation (...) est au service des apprentissages et doit permettre à chaque élève de progresser par une connaissance objective de ses acquis et aux professeurs de mieux adapter les aides et les approfondissements aux besoins constatés.* »³³ Autrement dit, si on considère le processus d'évaluation comme la somme des divers types d'évaluations menées au sein de la classe, l'évaluation sommative n'est pas la *finalité* d'un apprentissage mais une *étape* en vue d'un développement au long cours, adapté aux besoins et aux difficultés détectés. C'est donc, *avant même d'être un outil de classement des élèves, un outil indicateur pour le professeur.*

Comme le remarque de Vecchi, s'échapper du temps T est de plus un moyen de *mieux apprécier* le degré d'acquisition *réel* d'une compétence. Il prend en effet note d'un écueil récurrent, dont ma pratique et les discussions que j'ai eues confirment le caractère répétitif. Il arrive souvent que des élèves parviennent à réussir parfaitement des exercices précis pour lesquels ils ont été formatés, mais soient incapables de réutiliser les compétences et connaissances impliquées dans un autre contexte. C'est ici la question du transfert qui se pose. D'après le chercheur, transférer consiste précisément à « savoir réutiliser la compétence sans la réintervention du professeur »³⁴, ce qui correspond à la meilleure définition possible d'une compétence *acquise*. Or, cela ne peut être vérifié lors de l'évaluation sommative, qui cible la capacité à restituer des compétences et connaissances dans un contexte ciblé – exercices déjà

³¹ Gagneux, *op.cit.*, p.22

³² Hadji, *op.cit.*, p.83

³³ Circulaire 2014, *art.cit.*

³⁴ De Vecchi, *op.cit.*, p.81

vus, révisions la veille *en vue de* réussir, le tout à la sortie d'un apprentissage de plusieurs jours... Ce n'est donc *qu'une fois l'évaluation sommative réussie et terminée* que le moment sommatif « réel » peut avoir lieu ; ce n'est que là que le professeur, au fil de l'année, peut apprécier la capacité des élèves à réinvestir les compétences et connaissances acquises dans divers contextes où elles sont requises, sans qu'il ait à guider leur attention vers elles.

Il apparaît enfin que replacer l'évaluation sommative dans le contexte plus large de l'activité évaluative globale du professeur des écoles, activité qui est le *fondement* de sa pratique puisqu'elle permet de réguler en permanence son enseignement en vue de permettre les apprentissages, est un acte *beaucoup plus logique et juste* que de la considérer comme un objet autonome à la durée de vie limitée. En effet, il arrive fréquemment que les élèves ne comprennent certains points que lors de la *correction* des évaluations, voire même, qu'ils ne les comprennent que bien plus tard dans l'année ! Dès lors la question d'un moment T comme *seul moment sommatif* sur une notion se pose : est-il juste de remplir le LSU avec des données qui ne sont peut-être plus vraies au moment où le professeur les communique, mais dont la transmission va avoir un impact effectif sur la façon dont l'élève est perçu par ses parents, l'institution scolaire, et *in fine*, lui-même ? Comme le note de Vecchi, c'est là l'un des écueils inacceptables de l'évaluation ponctuelle qu'elle ne laisse *aucune* chance à l'élève de se rattraper.³⁵ A l'inverse, replacé dans un temps plus long, le sommatif peut devenir une pratique au service de la validation des acquis par les élèves, à charge alors pour le professeur d'observer finement le moment où il lui paraîtra le plus indiqué de soumettre une nouvelle évaluation à ceux qui semblent pouvoir la réussir. A charge aussi pour lui, évidemment, de donner à cette évaluation la forme la plus adéquate, au vu de ce qui a été dit dans toute cette partie.

C'est cette question de la forme qui va maintenant faire l'objet d'un compte rendu analytique. Dans la partie qui suit, je proposerai un retour d'expérience permettant d'apprécier l'évolution de mes pratiques sommatives durant l'année. Puisqu'il est dit qu'on n'apprend bien que par l'expérience, et qu'une vraie chute veut mieux que dix schémas pour apprendre à marcher, la première partie sera consacrée au compte-rendu d'une évaluation qui m'a été très instructive par les *erreurs* qu'elle comporte. Je me pencherai ensuite sur les différentes façons dont mes copies de « contrôles » ont évolué à la suite de cette expérience. Enfin, je rendrai compte des diverses pratiques que j'ai mises en place et qui *entouraient* l'évaluation sommative proprement dite, en vue de pallier ses divers effets structurels.

³⁵ *Ibid.*, p.62

PARTIE 2

***De l'évaluation « au jugé » à la conception
d'évaluations sommatives pertinentes : compte-
rendu d'une année de travail***

I. Les problématiques liées à la conception de l'évaluation sommative

L'exemple sur lequel je vais me pencher maintenant a été fondamental dans l'évolution de ma réflexion, en ceci qu'il m'a mis aux prises directes avec presque tous les écueils que l'on peut rencontrer lorsqu'on conçoit une évaluation sommative. Il illustre de façon claire, à mon sens, la difficulté que peut avoir un débutant à évaluer des compétences dans une matière faisant intervenir des savoirs purs.

1. Le contexte

Durant la période 2, j'ai longuement travaillé sur les gaulois en histoire. La séquence, transdisciplinaire, a été conçue en vue de permettre la construction de compétences précises. Il s'agissait pour les élèves d'apprendre à construire leurs savoirs par eux-mêmes en développant une méthode d'étude de documents. Cette méthode était basée sur l'observation fine, permettant l'émission d'hypothèses en vue de répondre à une question. Je leur ai par exemples demandé en début de séquence de me donner leurs croyances concernant le mode de vie des gaulois en répondant aux questions : « où vivaient-ils ? » et « comment se nourrissaient-ils ? ». Après avoir noté leurs réponses initiales au tableau, je leur ai demandé de travailler sur des fiches de recherche (annexe n°1).

Leur travail était le suivant : réunis par binômes, les élèves devaient observer attentivement les photographies présentées et noter dans la grille toutes les informations qu'ils tiraient de cette observation. Après l'avoir remplie, je leur demandais de répondre à nouveaux aux questions initiales – « où vivaient-ils » et « comment se nourrissaient-ils » - au moyen d'une phrase. L'idée était de les amener à mettre en lien les différentes informations prélevées afin de tirer une conclusion générale de leurs observations. Ils allaient ensuite accrocher leur fiche d'activité au tableau, en vue d'un temps de mutualisation orale durant lequel les différents groupes ont pu dialoguer autour de leurs observations et de leurs conclusions. Après discussion, les élèves en sont arrivés à formuler une conclusion collective que j'ai transcrite au tableau (« Les Gaulois vivent dans des fermes, ce sont des agriculteurs »). J'ai ensuite impulsé un court débat autour de la question des habitations (« vous me dites d'après la fiche 1 que les Gaulois vivent dans des fermes mais aussi dans des villes d'après la fiche 2 ; qu'en est-il ?) afin de susciter la réflexion et l'émission d'hypothèses, ce qui m'a permis d'introduire les savoirs dont ils ne pouvaient disposer – concernant les *vicus* et les *oppidums*. La dernière phase était un moment d'institutionnalisation durant lequel les élèves étaient invités à reformuler ce qu'ils avaient fait et appris, aussi bien en termes de méthode qu'en termes de savoirs. La leçon, que je leur ai distribuée le lendemain, réunissait les différentes images sur lesquelles ils avaient travaillé ainsi que la leçon finale conçue en commun. Je leur ai également donné à chacun un

exemplaire photocopié de leurs feuilles de recherche afin de garder une trace mémorielle du travail effectué.

Par cette pratique, j'essayais de les mettre au plus près des méthodes de travail de l'historien et de l'archéologue, qui tirent leur savoir de l'interrogation systématique des sources primaires à leur disposition. L'idée était de leur faire adopter une posture active les poussant à construire eux-mêmes leur savoir, et ce grâce à l'intégration d'une méthodologie mettant en jeu l'analyse des données relevées ; il s'agissait également de les laisser se confronter aux limites de cette méthodologie et aux questions qu'elle soulève.

Le déroulement que je viens de décrire a servi de modèle pour les trois premières séances. Les deux dernières, quant à elles, ont porté sur la question de la *source* : l'objectif était de pousser un peu plus loin la méthode d'analyse de documents en y intégrant un questionnement sur sa provenance. Les élèves ont appris d'une part à identifier la source d'un document en se référant à sa légende, d'autre part à s'interroger sur sa pertinence historique au vu de cette identification (*cf* annexe n°2, qui présente un exercice mené dans ce cadre).

Toute la séquence portait donc sur des compétences liées à l'analyse de documents. Ces compétences étaient tirées des programmes : il s'agissait pour les élèves de « formuler des hypothèses », « se poser des questions », « identifier un document et comprendre pourquoi il doit être identifié », « extraire des informations pertinentes pour répondre à une question » et, les élèves travaillant par binômes, « travailler en commun pour faciliter les apprentissages individuels ».

Ces compétences étant réellement CE QUE les élèves apprenaient à maîtriser, mon évaluation aurait naturellement dû porter dessus. Or, au final, les exemples que j'ai trouvés dans les manuels et sur internet, mais également mon expérience, m'ont poussé à me focaliser sur les *contenus*, et j'ai produit une évaluation qui m'a posé de nombreux problèmes à la correction. Je renvoie ici le lecteur à l'annexe n°3 qui donne à voir un exemplaire de cette évaluation : il s'agit d'un document pivot pour aborder la suite de ce mémoire.

2. Analyse

a) Une évaluation non congruente

La première chose qui m'a frappé lors de la correction est la disjonction entre l'objet d'apprentissage initial et les points effectivement évalués. Alors même que l'intégralité de ma séquence portait sur les compétences suscitées et faisait preuve d'une certaine cohérence à cet égard, les questions que je propose manquent cruellement de congruence avec ces compétences. Il apparaît ainsi que la méthode d'analyse travaillée durant les trois premières séances ne fait l'objet d'*aucune* évaluation, sous quelque forme que ce soit. Quant à la méthode d'analyse des sources, objet de la seconde phase de ma séquence, elle est bien reprise... dans la question bonus ! Ce qui revient à dire que j'ai travaillé d'arrache-pied pour faire intégrer des compétences

très précises, le tout au travers de deux méthodologies d'analyse ciblées... et que j'ai balayé tout cela au moment de l'évaluation, comme si ma séquence n'avait servi à rien. Sur quoi porte donc effectivement mon évaluation ? Si l'on regarde les différents types de questions, il apparaît que je ne suis en train de me concentrer *que* sur la vérification de l'apprentissage des différentes leçons. Autrement dit, tout se passe comme si les séances n'avaient été qu'une façon plus ou moins folklorique d'amener aux *connaissances* qu'elles permettaient de faire émerger, connaissances qui auraient été l'unique objet de mon enseignement.

Pourquoi ai-je agi ainsi ? En début d'année de PES, on avance beaucoup au jugé, parfois presque « à l'aveugle ». On est peu sûr de sa légitimité et, surtout, nous manquons de repères. Il s'agit de trouver des modèles afin de pouvoir construire une normativité à laquelle se référer lorsqu'on voudra concevoir une « bonne » évaluation. Je me suis donc inspiré de manuels, de recherches internet et d'évaluations diverses déjà existantes pour concevoir une évaluation qui fasse « professeur ». Le problème est que, ce faisant, j'en suis arrivé précisément à évaluer la *seule chose* que nous n'avions pas travaillée en classe : l'apprentissage de la leçon par l'élève lui-même. Lorsque j'agis de la sorte, il est évident que je ne suis pas en train de mettre en œuvre une évaluation bienveillante, permettant à l'élève d'apprécier l'intégration des compétences travaillées et me permettant à moi de voir dans quelle mesure mon enseignement a été adapté aux élèves. Bien au contraire, je tombe dans l'écueil d'une évaluation-sanction, qui ne peut que reproduire mécaniquement les inégalités : mon évaluation valide et sanctionne un *état de fait* dans le rapport à l'apprentissage par cœur, sans y apporter la moindre modification. De plus, elle crée une congruence artificielle entre cette validation et le fait d'avoir « compris la leçon », puisqu'elle clôt la séquence ; les élèves obtenant un retour négatif à cette épreuve se retrouvent donc avec un constat d'échec alors même qu'ils peuvent avoir travaillé et compris les notions enseignées. Or, nous avons vu avec de Vecchi l'importance que ce constat avait sur l'image de l'élève et sa motivation : en agissant de la sorte, je n'évalue pas – mes enseignements – mais je *classe* – mes élèves. Bien plus grave, je *mets en place une dynamique d'échec chez les élèves que je classe mal*, comme cela a été montré en partie I !

A la lumière de ce problème, il m'est donc apparu clairement une chose simple : la première question à se poser lorsque l'on conçoit une séquence, *quel que soit le champ disciplinaire concerné*, est celle de l'évaluation. *Qu'est-ce que j'évalue en fin de séquence ?* Cette question doit être en congruence totale avec une autre : *Quel est mon objectif en termes d'apprentissages ?* Il ne s'agit pas tant d'un débat entre évaluation par compétences et évaluation sur l'acquisition des savoirs *en soi* mais bien de se demander *ce qui a été le but de notre enseignement* durant la séquence. L'évaluation ne se fabrique pas au jugé, n'a pas pour but de ressembler à une « bonne évaluation ». En fait, elle n'a pas pour but de ressembler à *quoi que ce soit*. La forme qu'elle prend doit tout simplement être la plus adéquate possible en vue d'une évaluation objective et bienveillante, permettant l'enregistrement des *réelles* capacités des élèves. Le deuxième point est qu'elle doit être conçue pratiquement en même temps que la fiche de séquence ; du moins l'évaluateur doit-il avoir clairement en tête *ce qu'il évaluera*. Ces deux conclusions, qui paraîtront sans doute évidentes à tout professeur expérimenté, sont

absolument *fondamentales* à prendre en compte si l'on veut que nos évaluations soient réellement des outils au service de l'élève et non des moyens avérés de cristallisation des inégalités. Je montrerai plus bas comment j'ai pris acte de ces faits dans la construction de mes évaluations suivantes.

b) Aux prises avec la correction : découverte de ce que j'évalue *effectivement* et de l'inégalité de traitement entre les compétences évaluées

Les considérations ci-dessus sont nées des problèmes que j'ai rencontrés lorsqu'est venu le moment de la correction. Si les questions que je posais dans l'évaluation avaient fait l'objet d'une réflexion – globalement, le but final était de traverser l'ensemble des leçons, de couvrir tous les points importants –, la façon dont je devais les corriger était resté dans l'angle mort de ma conscience. Avançant au jugé, je n'avais pas encore réalisé les difficultés qui allaient m'être posées. J'avais certes eu plusieurs fois, durant la période 1, à réfléchir longuement sur la façon d'organiser mon barème au moment de corriger les copies – ce qui occasionnait des après-midi pour le moins laborieuses. La difficulté atteint cependant ici un point tel, que la nécessité d'une scission avec les méthodes qui avaient été les miennes jusqu'à présent devint évidente.

Face aux copies, je me suis retrouvé aux prises avec une question très concrète : *quelle importance donner à chaque réponse dans son rapport à la « note » finale, c'est-à-dire à l'appréciation générale ?* (j'utilisais encore à l'époque les *ab, b, tb*). Autrement dit, et étant attendu que je ne comptais pas mettre en place une note sur 20 dont nous avons vu l'inanité dans la partie 1, il s'agissait de trouver une façon d'organiser mon système de validation des acquis à partir des questions posées dans l'évaluation. *Or, je ne m'étais pas posé en amont la question de la nature des acquis à évaluer*, puisque j'avais, comme je l'ai dit, créé mon évaluation d'après un seul principe : vérifier la mémorisation des différentes informations contenues dans la leçon. Me voilà donc, *a posteriori*, en train d'analyser mon évaluation pour y identifier *ce que j'évalue effectivement* ! Il apparaît vite que trois pôles de compétences sont mis en jeu dans cette épreuve :

- La question 1. a) porte sur des compétences « stratégiques » de réussite à une épreuve : l'élève fait jouer ensemble les éléments dont il se rappelle et ses capacités de déduction pour replacer le plus correctement possible l'ensemble des noms de vêtements. Le support visuel et le fait que les noms soient donnés rendent ce type d'exercice (pourtant assez répandu) assez impropre à évaluer quoi que ce soit de pertinent, si ce n'est la capacité de l'élève à mettre en place des stratégies pour gagner des points dans une copie.
- Les questions 1.b) à 6 évaluent l'apprentissage brut des éléments de la leçon et la capacité à les restituer correctement. L'évaluation ne porte donc pas tant sur une compétence que sur des connaissances.
- La question bonus évalue la compétence à analyser la pertinence d'un document à partir de sa source, compétence qui a été travaillée en classe. L'élève doit être capable : a) d'identifier la source, ce qui implique d'avoir compris le rôle de la

légende, b) d'en conclure la pertinence, ce qui implique d'avoir retenu quels éléments confèrent sa pertinence à un document d'une part (mémorisation pure), et d'être capable d'articuler ces éléments à sa propre réflexion d'autre part (capacité de transfert).

Plusieurs observations découlent de ce découpage. D'une part, deux des trois pôles n'ont pas fait l'objet d'un travail en classe (les deux premiers). De ces deux pôles, seul celui portant sur la mémorisation pure de connaissances peut être considéré comme pertinent, et encore, *si et seulement si* les conditions de cette pertinence sont réunies (nous reviendrons sur ce point plus bas). Quant au premier pôle, il s'agit en fait d'une tautologie puisqu'il n'évalue que la capacité d'un élève à s'en sortir face à une question d'évaluation, et ce précisément lorsqu'il n'a pas intégré les connaissances ou les compétences requises. On pourrait m'objecter que la question 1.a) demande tout de même d'avoir appris sa leçon mais les faits montrent que de nombreux élèves ont su s'en sortir sur cet exercice tout en échouant totalement aux questions de cours, tandis que l'inverse n'est pas vrai. Je me retrouve donc avec trois pôles d'évaluation dont l'un n'est en soi pas du tout pertinent pour évaluer les acquisitions des élèves, mais qui doit tout de même entrer en ligne de compte dans la notation finale, et un autre qui ne peut-être pertinent que si on met en place les moyens pour cela, étant attendu qu'il correspond à un travail fait à la maison.

L'autre point à prendre en compte concerne l'importance de ces trois pôles en termes d'espace et d'importance relative. Il apparaît en effet que le pôle sur les compétences stratégiques n'occupe même pas une question entière, tandis que celui sur la restitution de connaissances représente la très grande majorité de la copie. Quant à la compétence d'analyse de la pertinence d'une source, elle occupe un exercice entier, permet un réinvestissement et un transfert de la compétence (le chaudron correspond à un cas « classique » tandis que les « jouets représentant un objet historique » n'ont pas été étudiés) mais elle est évaluée dans un exercice « bonus », donc désignée comme n'ayant pas la même importance que les autres. Autrement dit, je me retrouve avec une copie qui évalue très inégalement chaque compétence, ce qui pose un problème pour la pertinence de l'appréciation finale. Je dois donc me livrer à un travail d'équilibriste dans ma correction, de manière à compenser les manques d'une façon qui redonne leur juste place aux compétences et connaissances évaluées tout en respectant le principe de justice scolaire.

c) Cas précis de correction : la question de la justice scolaire.

Je vais maintenant me pencher sur une copie en particulier : il s'agit de Benjamin, plutôt situé dans la moyenne haute de la classe. Son cas va me servir d'exemple car il illustre bien les problèmes que j'ai rencontrés en corrigeant.

Benjamin a bon à la première partie de la question 1, ou il faut replacer du vocabulaire sur un schéma. Il ne répond cependant pas à la deuxième partie de la question, qui demande une action cognitive différente (restitution d'un fait appris). Comment évaluer cette question dans ce cas là ? Je ne peux pas le sanctionner alors qu'une partie de l'exercice est juste ; pour autant, le

fait qu'il ait faux à la question finale m'empêche de valider entièrement la mémorisation du vocabulaire. Je me retrouve dans le cas d'une disjonction avérée entre *ce que je veux évaluer*, à savoir l'apprentissage d'une partie précise de la leçon portant sur les vêtements des gaulois, et une logique « à points » qui s'impose à moi *de facto*, du fait que mon exercice est en deux parties faisant chacune intervenir une action cognitive différente. En d'autres termes, je me vois contraint de « mettre des points » tout en ne pouvant valider l'apprentissage, ce qui fait glisser mon évaluation d'une vérification des acquis vers un test de performance répondant à une logique quantitative. Même si ma copie n'est pas notée sur 20, l'ambiguïté est là. Je me rends compte que cet exercice me renvoie à une question fondamentale, que je ne me suis pas posée en amont : celle de *l'objectif* de mon évaluation. *Qu'est-ce que je veux faire lorsque j'évalue ?* Lié à cela, *quelle logique, quel objectif président à la conception des exercices ?* Ne sachant encore comment mener ma correction, je continue d'avancer dans la lecture de la copie, en gardant ces questions en tête.

Le second point problématique se présente au détour des questions 2 à 6 : il apparaît vite que Benjamin n'a pas appris sa leçon. Or, nous l'avons vu plus haut, mon devoir était très déséquilibré : la grande majorité des questions porte sur la restitution pure de connaissances, soit sur la vérification *d'un seul* fait. Ceci a pour conséquence l'apparition d'un biais de correction inévitable : à chaque nouvelle question à laquelle il répond mal, Benjamin m'apparaît un peu plus comme « n'ayant pas fait d'effort », n'ayant « vraiment » rien appris... l'effet de surenchère dû à la répétition des réponses erronées peut mener à vouloir sanctionner au lieu d'évaluer – d'autant plus face à certaines réponses qui semblent témoigner d'une réelle désinvolture (« *où vivent les Gaulois ? – Ils vivent en Amérique* ») où lorsqu'il donne deux fois la même réponse à deux questions différentes dans l'idée que celle-ci doit être bonne dans l'un des deux cas. Or, il est bien évident que je ne peux pas me permettre de sanctionner un élève pour la non-acquisition *d'un* point d'évaluation (l'apprentissage de la leçon) dont il est de *ma* responsabilité d'avoir étendu l'évaluation à l'excès. Dès lors, que faire ?

J'ai fini par regrouper les questions en trois groupes, qui ne répondent pas *totalem*ent à une logique de séparation par pôles de compétences mais qui permettent de rééquilibrer vers une évaluation bienveillante, de se rapprocher des programmes par rapport à la version initiale, et de palier à mes erreurs de conception. Les élèves sont donc évalués sur trois items : « habillement », qui correspond à la question 1, « mode de vie », questions 2 à 6, et « analyse de documents », qui correspond à la question bonus, chacun d'eux évalué d'après la notation LSU (acquis, en cours d'acquisition, non acquis) et le tout synthétisé en une appréciation (*ab, b, tb*) accompagnée si besoin d'un commentaire. Cette répartition présente trois avantages pratiques : d'une part, cela permet de compenser l'effet de sanction qui naitrait sans cela du fait d'avoir faux aux questions de cours – puisqu'elles sont plus nombreuses que les autres – en les ramenant à *un* bloc d'évaluation sur trois. D'autre part, cela me permet de noter « acquis » à un élève pour le bloc « habillement » s'il a bon à la question 1.a) mais qu'il ne se rappelle pas du mot « fibule » : puisque je n'évalue pas ici une compétence précise mais selon une logique thématique, le fait de ne pas se rappeler d'un mot alors que l'on a bon à tout le reste n'est plus

aussi pénalisant. A l'inverse, une logique de classement des questions en fonction de la compétence mise en jeu aurait été injuste pour les élèves ayant eu bon à la question 1.b) et faux, totalement ou en partie, à la question 1.a). Enfin, cela me permet de redonner une réelle importance à la question bonus puisqu'elle est la seule à être reliée à une compétence explicitement travaillée en cours.

Cette solution de compensation conserve toutefois quelques problèmes notables. D'une part, il n'y a pas de congruence parfaite, nous l'avons dit, entre les trois items de correction retenus et les compétences mises en jeu, ce qui fausse l'appréciation finale : les élèves, comme les parents, ne sont pas *réellement* informés par la note. En effet, même si la congruence compétence/questions est correcte concernant les groupements « mode de vie » et « analyse de documents », le pôle « habillement » continue de poser problème puisqu'il compte dans l'appréciation finale sans pour autant évaluer un élément réellement important. A l'inverse, je ne peux me permettre de sanctionner de façon trop importante dans l'appréciation finale un élève obtenant un NA au pôle « analyse de documents » puisqu'il a d'abord été présenté comme une question bonus. Or, il s'agit d'une compétence réellement importante dans mon évaluation. Je me retrouve donc avec un objet composite dont il est très difficile de supprimer toute l'ambiguïté. La question des élèves ayant parfaitement appris leur leçon mais qui auraient, par exemple, mal intégré la compétence mise en jeu dans la question bonus et opéré des confusions dans l'exercice 1.a) se trouve également posée : eux qui auraient naturellement été les mieux classés dans une logique scorée, peuvent se retrouver ici pénalisés alors même qu'ils ont rempli le contrat didactique (apprendre sa leçon pour l'évaluation), ce qui tendrait *également* à l'injustice.

Je me retrouve donc avec un système finale de notation très composite : trois items, évalués par A, ECA, NA, une appréciation globale – *tb, b, ab* – et un commentaire pour préciser la notation lorsque c'est nécessaire.

d) Bilan

Je me suis donc retrouvé face à une évaluation dans laquelle les contenus enseignés durant la séquence étaient absents ou sous-représentés, du fait d'un manque de questionnement en amont. J'ai dû opérer une compensation très compliquée – et imparfaitement satisfaisante – pour rattraper le déséquilibre entre les compétences et connaissances que j'évaluais effectivement, et me suis tout de même retrouvé au final avec un item problématique puisque n'évaluant rien de pertinent (l'item « habillement »). La question de *ce que j'évalue*, donc de *ce que j'ai enseigné*, a été insuffisamment posée ; il apparaît ici qu'il est absolument fondamental de l'avoir en tête en permanence lorsqu'on conçoit son évaluation.

La question de l'objectif poursuivi derrière chaque exercice est également fondamentale : *est-ce que je fonctionne selon une logique scorée ou est-ce que je privilégie la vérification de l'acquisition d'une compétence/connaissance ?* Se la poser face à chaque exercice peut aider à éviter les écueils de ma question 1, qui m'empêche de vérifier clairement l'acquisition du vocabulaire, tout en induisant un fonctionnement « à points » - l'inverse, en un mot, de ce qu'une

évaluation « juste » est sensée être. Une évaluation ne peut pas être un objet conçu au jugé, ou trop rapidement. En l'absence d'une réflexion bien structurée, elle tombera forcément dans les écueils qui ont été les miens ici, en induisant des biais dans la correction, en mettant à mal la justice scolaire par un déséquilibre dans l'évaluation des différentes compétences/connaissances, en générant opacité et ambiguïté (évalue-t-on selon une logique thématique, une logique de points, une logique par compétences...) dans l'information finale transmise par la note/appréciation à l'élève et à ses parents.

J'aimerais lier ces considérations à un dernier point, qui me semble mis en lumière par cette expérience. Lorsqu'il parle de l'évaluation, de Vecchi note un phénomène assez surprenant, que j'ai d'abord trouvé difficile à comprendre : le fait que pour beaucoup d'élèves, *l'évaluation ne fasse pas sens*³⁶. Que veut-il dire par là ? Je pense que ce constat peut être mieux compris si l'on considère que c'est au détour de ce genre d'expérience, où une évaluation est *en effet* en décalage avec ce qu'elle est sensée évaluer, qu'un élève peut développer cette sensation de confusion. En considérant les trois violences successives que constituent un sujet d'évaluation non congruent avec les conceptions que l'on a construites, une note et un commentaire médiocres, et le fait de devoir faire signer cette évaluation par ses parents, il me semble qu'on peut comprendre que l'élève, psychologiquement secoué, s'installe dans un sentiment durable de confusion quant à sa capacité à comprendre les attendus scolaires. S'il ne peut analyser ce qui n'allait pas dans la façon dont il a été évalué, il semble logique qu'il en arrive à percevoir l'ensemble du système évaluatif comme aléatoire. De plus, comme d'autres élèves réussissent effectivement, il a de grandes chances d'en arriver à la conclusion que c'est *lui* qui n'est pas « suffisant ». Indépendamment de la pertinence – qu'il n'a pas forcément les clés pour percevoir – ou de la facilité de l'évaluation, c'est la situation évaluative qui devient donc génératrice de stress *en soi*, enfermant les élèves dans un cercle vicieux : l'évaluation ne fait plus jamais sens, du simple fait qu'elle n'a, un jour, *en effet pas eu de sens*³⁷.

Or, menée ainsi, elle devient le point où se noue la mise à l'écart potentiellement définitive de certains élèves. Henri Laborit, grand médecin et penseur de la logique comportementale, note qu'une situation ayant un impact émotionnel anxiogène a tendance à générer toujours le même type de comportement : l'affrontement, la fuite ou la paralysie.³⁸ Si le premier est adéquat – il faut que l'élève accepte d'être mis à l'épreuve et affronte la situation évaluative pour s'en sortir – les deux dernières sont inadaptées et vont l'enfermer dans l'échec. En guise de fuite, je repense ainsi à Madany ou à Paul, deux élèves dont le système de défense consistait à refuser purement d'affronter l'évaluation. Devant leur copie, je les ai vus énervés, dépités, figés dans une expression de refus et des comportements ritualisés – gonfler les joues, souffler, le regard fixé sur un objet invisible, lorsque je leur demandais ce qui n'allait pas. Leur réponse, arrachée à force de patience, était d'ailleurs édifiante : « j'y arrive pas [sic] » ; « je comprends pas » ... Or, ils n'avaient parfois pas lu plus loin que la première question ! Il faut

³⁶ De Vecchi, *op.cit.*

³⁷ *La stratégie du choc*, Naomi Klein, Actes Sud, Arles, 2008 [trad] ; *Eloge de la fuite*, Henri Laborit, Gallimard, Paris, 1976 ; *La colombe assassinée*, Henri Laborit, Grasset, Paris, 1983

³⁸ Henri Laborit, *op.cit.*

croire que celle-ci suffisait à raviver tout un circuit de comportements acquis. Le réel problème n'était pas la copie elle-même mais la perspective de l'échec, l'image d'eux-mêmes que la situation évaluative risquait d'actualiser.

Quant au troisième comportement évoqué, la paralysie, je l'ai rencontrée chez Alicia, Sara, Mohamed... qui se retrouvaient victimes de trous de mémoire continus face à des copies trop compliquées... La vérité est qu'ils étaient pris de panique ! Ce n'était pas les questions en soi qui posaient problème, mais bien le fait de se retrouver face à une situation associée à un événement psychologiquement traumatique. Il est bien évident que plusieurs facteurs peuvent être la cause du trauma initial : le fait de ne pas avoir su sa leçon un jour, des échecs répétés aux évaluations, une relation médiocre avec un professeur, voire des facteurs familiaux... Cependant, le fait de s'être retrouvé face à des évaluations mal conçues, s'il n'est que rarement évoqué, apparaît comme potentiellement déterminant. Fabriquer des évaluations congruentes est donc un enjeu fondamental dans la lutte contre les facteurs d'échec.

Ce n'est pas tout. Henri Laborit fait remarquer que chaque itération d'une situation anxiogène non seulement réactive un comportement de défense donnée, *mais renforce la probabilité future de recours à ce comportement*.³⁹ Il apparaît donc que construire une évaluation pertinente, qui ne fasse pas obstacle aux élèves est absolument nécessaire mais qu'il s'agit aussi *de mobiliser les moyens pour casser ces systèmes de défense inadéquats* et pour installer la possibilité chez l'élève *d'affronter* les situations évaluatives. J'en conclus que je dois non seulement produire une évaluation « congruente » mais également mettre en place autour une pédagogie permettant de l'aborder sereinement, notamment pour les élèves ayant déjà automatisé des comportements inadaptés, si je veux bel et bien favoriser la réussite de chacun.

La façon dont je m'empare de l'évaluation sommative est donc loin d'être anodine : soit j'en fais le lieu où se noue la confiance des élève en leurs capacités, soit j'en fais l'une des causes directes de la mise-en-échec d'une partie d'entre eux. Il apparaîtra sans nul doute au lecteur qu'une fois ces considérations mises-en-lumière, il était très difficile de conserver mon mode de fonctionnement. J'ai donc cherché comment impulser une transformation en profondeur du processus évaluatif mis en place jusqu'alors. Le chapitre suivant rend compte de mes expériences dans ce cadre.

³⁹ Henri Laborit, *Ibid.*

II. Concevoir une évaluation au service des apprentissages

A la lumière de tout ce qui vient d'être dit, il me semble pouvoir tirer d'une évaluation pertinente trois caractéristiques principales : elle est informative – claire, lisible, elle délivre des informations aux parents, aux élèves et au professeur sur le degré d'assimilation des notions ciblées – formative – elle aide aux apprentissages et est à leur service – et réellement sommative – c'est-à-dire qu'elle vise à évaluer les notions ciblées et rien d'autre, de façon claire, pertinente et efficace. Le travail doit donc se faire à trois niveaux : celui de la conception de la copie, celui de la préparation des élèves au moment sommatif de façon à déverrouiller les blocages potentiels, et celui de l'intégration de l'évaluation sommative dans le contexte plus large de la formation continue.

1. limites et points morts des observables

Concernant la conception de l'évaluation, j'ai deux interrogations importantes : « est-ce que mon évaluation est « bien » conçue ? » et « est-ce qu'une évaluation bien conçue favorise la réussite des élèves ? » Les indicateurs vont donc logiquement découler de ces interrogations. Concernant la première question, je conclus d'après ce que nous venons de voir qu'une évaluation est bien conçue si :

- elle évalue l'acquisition de compétences et non selon une logique à points
- elle évalue clairement les compétences ciblées, en ne faisant pas obstacles à leur réinvestissement par les élèves et n'opérant pas de confusion entre plusieurs compétences dans les exercices
- elle ne met pas d'obstacles à la notation juste de l'élève
- les compétences et connaissances attendues sont clairement établies et communiquées

Si l'on reprend la typologie de Houssaye, qui définit l'apprentissage comme un système de relations entre trois acteurs que sont le professeur, l'objet d'enseignement et l'élève, mes indicateurs principaux se situent donc majoritairement entre les deux premiers – c'est-à-dire au niveau de la façon dont je conçois l'évaluation d'après les divers indicateurs que j'ai pu définir. L'avantage de ce point est qu'il m'est assez facile de déterminer objectivement si les changements que j'apporte à l'évaluation sommative vont dans le sens défini ; en ce sens, je peux noter moi-même si j'ai fait du « bon » travail.

Ce fait comporte toutefois un point mort important, à savoir, qu'il m'est difficile de définir dans quelle mesure ces indicateurs sont réellement efficaces pour permettre la réussite des élèves de façon plus probante. En effet, il est complexe de juger de l'efficacité de mes transformations à ce niveau là pour trois raisons :

- d'une part, les changements dans la forme de l'évaluation se jouent à un niveau profond, et c'est toute ma façon de noter qui se transforme avec les modifications apportées, comme nous allons le voir dans les pages qui suivent. Comparer une note obtenue avec l'ancien système arbitraire et avec le système que j'ai mis en place par la suite est donc sans fondements, en ceci qu'une note plus « juste » n'implique pas forcément une note supérieure.

- deuxièmement, il me faudrait, pour une comparaison réellement parlante, pouvoir comparer deux évaluations portant sur les mêmes notions, puisqu'un élève ayant par exemple échoué à une évaluation portant sur la soustraction dans le premier modèle de mon évaluation peut parfaitement réussir une autre évaluation portant sur la multiplication sans que cela soit imputable à la nouvelle logique de conception de mes copies. Or, même si j'avais opéré de la sorte, les données auraient été biaisées, puisque les élèves auraient dès lors été confrontés deux fois d'affilée aux mêmes notions, ce qui inclut un biais dans les performances de la seconde évaluation. Je n'aurais dans ce cas pas seulement évalué le fonctionnement de mon évaluation mais également les effets qu'un temps supplémentaire de réflexion, une première correction et le fait d'être mis une seconde fois face à des questions en partie semblables ont sur la performance.

- Enfin, il serait présomptueux de vouloir opérer une évaluation objective de ma réussite à l'aune de la performance des élèves, alors même que le nombre de facteurs qui entrent en compte dans la réussite d'un devoir est tellement important qu'il pourrait faire l'objet d'un mémoire de recherche à part. Il est notable que, dès lors que l'on travaille dans le cadre d'un système socialisé complexe tel que celui de la classe, une variable seule est difficilement isolable.

Pour toutes ces raisons, s'il est clair que je peux évaluer de façon assez claire les avantages et limites de mes expériences en vue de rendre l'évaluation plus congruente avec les objectifs de la circulaire de 2014 – rendre l'évaluation plus claire, lisible, et au service des apprentissages –, il sera difficile pour moi d'évaluer leur efficacité à l'aune de la réussite effective des élèves. Tout au plus puis-je constater que mes évaluations ne vont en effet plus dans le sens d'un arbitraire de notation, d'un manque de cohérence entre ce qui est enseigné et ce qui est évalué, d'une opacité de la note, qu'elles ne font plus obstacle à la validation des compétences dont elles tentent d'évaluer l'acquisition ; en un mot, qu'elles sont effectivement plus justes et cohérentes... ce qui est déjà un point important ! Ces limites une fois fixées, voici maintenant les évolutions apportées.

2. La grille de compétences : un outil holistique

a) permettre évaluation réellement congruente

Afin d'atteindre les objectifs définis ci-dessus, je me suis approprié un outil découvert chez ma PEMF, Mélanie Cueto, lors des séances d'observation dans sa classe. Mélanie fait figurer en haut de chaque copie une grille de compétences précises ; pour chaque compétence est indiqué le ou les exercices correspondant(s) et une colonne dans laquelle inscrire le niveau

d'acquisition de la compétence – acquis, en cours d'acquisition, non acquis – après correction de la copie. Loin d'être un simple moyen de clarifier les notions évaluées pour les parents, cette grille implique e fait de nombreuses transformations dans la façon d'aborder la construction de mon devoir sommatif.

L'annexe n°4 présente une évaluation de géométrie que j'ai conçue grâce à une telle grille. Le premier point à noter est que sous cette forme, les compétences et connaissances sur lesquelles portent l'évaluation sont bien définies pour moi. Ceci me permet de structurer mon évaluation de façon claire et pertinente : les exercices ne sont pas le fruit de conceptions plus ou moins aléatoires mais d'une réflexion précise. Ils suivent désormais cette chaîne de construction mentale : « *qu'est-ce que j'ai enseigné aux élèves ? → telle compétence, telle connaissance → quels gestes et actions permettent de rendre compte de l'acquisition de ces connaissances et compétences au mieux ?* » De la réponse à cette dernière question naissent à la fois la grille et les exercices. Cette forme de conception a l'avantage notable de resserrer l'évaluation au plus près des connaissances et compétences effectivement apprises durant la séquence. Elle permet de plus de dresser une série d'observables précis qui réduisent d'une part l'arbitraire de la correction et d'autre part les amalgames entre plusieurs compétences et/ou connaissances dans les exercices. Le cas de l'exercice 1 de l'évaluation d'histoire, qui mêle restitution de faits appris et utilisation de stratégies logiques, a ainsi peu de chances de se reproduire, puisque face à chaque exercice je suis amené à me demander s'il évalue bien la compétence désirée et uniquement celle-là.

Globalement, l'utilisation de cet outil amène donc à un recul dans la construction de l'évaluation et à une certaine technicité dans la conception des différents exercices. Ceux-ci ne sont plus jamais repris tels quels d'un manuel ou d'internet, mais modifiées ou conçus de A à Z pour correspondre à mes observables. Cette congruence directe entre construction d'une grille d'observables et exercices permettant l'évaluation de ceux-ci permet également d'éviter les répétitions inutiles – je pense ici aux 6 questions de l'évaluation d'histoire qui n'allaient toutes que dans le sens de la vérification des savoirs – comme les oublis importants – je n'ai *jamais* évalué la compétence d'analyse de documents qui était au cœur de ma séquence sur les gaulois.

Si l'on regarde l'évaluation de géométrie, on peut ainsi noter qu'il y a sept exercices pour neuf compétences évaluées, là où un nombre presque équivalent de questions dans l'évaluation précédente renvoyait à seulement trois pôles de compétences. La courbe a été inversée en termes d'efficacité ; aucun exercice n'est superflu, et chaque compétence renvoie à un ou plusieurs exercices précis – ainsi les exercices 2 et 3 renvoient-ils tous deux à la compétence « je sais identifier un polygone » mais en couvrant deux aspects différents de cette compétence.

La question bonus, quant à elle, correspond à une compétence que j'ai faite travailler durant la séquence mais pour l'étude de laquelle le temps a manqué. Face aux impératifs temporels des programmes et de l'année de PES, j'ai décidé de ne pas retarder le moment de l'évaluation ; il n'aurait cependant pas été correct de donner à cette compétence le même poids qu'aux autres. La question bonus prend donc ici un sens réel en ceci qu'elle résulte d'un compromis entre le calendrier des programmes et l'avancée effective dans la séquence ; il ne

s'agit ni de « donner des points » - nous verrons plus bas que cette notion n'a d'ailleurs plus de sens dans le cas qui nous occupe - ni de minimiser un exercice de peur que les élèves le ratent - ce qui avait présidé à mettre mon exercice sur la pertinence des sources en bonus dans l'évaluation sur les gaulois.

Si l'on compte les items de la grille de l'évaluation de géométrie, on constate qu'il n'y a pas une congruence totale - neuf compétences pour sept exercices. Les deux compétences "surnuméraires" sont « je sais écrire les noms des différentes familles de polygones » et « je sais écrire correctement les noms des différents types de triangles ». Celles-ci sont évaluées dans les exercices 4 et 5, en même temps que les compétences « je sais identifier les différentes familles de polygones » - exercice 4 - ; « je connais les propriétés des différents triangles particuliers et je sais les nommer » et « je sais reconnaître les différents types de triangles en me reportant aux codifications » - exercice 5. En effet, les élèves sont amenés ici à écrire le nom des différents types de polygones et de triangles dans le même moment où ils les reconnaissent. Si mettre en place une grille d'évaluation permet d'éviter la confusion entre différentes compétences au sein d'un même exercice, comme cela a pu être le cas dans l'exercice 1 de mon évaluation d'histoire, il ne s'agit pas pour autant de dissocier artificiellement celles qui sont évaluables en même temps sans risque d'amalgame. L'avantage est de permettre de valider la compétence de reconnaissance des polygones/triangles même lorsque l'orthographe est faussée, et inversement : l'évaluation est ainsi réellement opérée au plus près des différents acquis et ne pénalise pas les élèves de façon injuste.

b) La notation revisitée

Il est à noter que cet équilibrage dans l'évaluation des différentes notions va dans le sens d'une meilleure justice scolaire. En partant des compétences et connaissances enseignées et en tirant les exercices idoines, je réduis le risque de sur-représentation de l'une d'elles au détriment des autres. Ainsi a-t-on vu aux paragraphes ci-dessus la congruence qu'il y avait dans le devoir de mathématiques entre les compétences évaluées et la taille du devoir : globalement 1 compétence = 1 exercice. Mais cela est vrai aussi lorsque ce n'est pas le cas : ainsi d'une évaluation de français conçue à la même période (annexe n°5) où la première compétence de la grille est représentée par trois exercices. Etant donné que je me base sur cette grille pour noter le devoir, ces exercices plus nombreux n'ont pas d'incidence directe le « score » final ; c'est au niveau de l'évaluation de la compétence visée, et seulement d'elle, qu'ils jouent. En créant ainsi plusieurs pôles scorés au sein du devoir, je peux donc moduler la taille de mon évaluation en fonction de mes besoins, tout en conservant la même importance finale à chaque compétence. Ceci permet d'éviter de sur- ou de sous-noter les élèves tout en gardant la congruence nécessaire à mon devoir.

Concernant la note elle-même, s'il est à noter que je n'ai pas su la supprimer, il est cependant clair qu'elle a changé de nature. Elle n'est en effet plus un score, mais bien une synthèse des différents acquis : sa valeur n'est plus sommative mais bien informative. Ainsi un *tb* ou un *b* ne se valent plus en soi, ne comptent plus comme résultat, puisque je me sers

directement des grilles de compétences pour remplir les différents items du LSU. Dans l'évaluation de géométrie présentée, le *b* présenté ne me sert en rien pour définir les acquis des élèves ; il est juste là pour assurer une continuité entre ma nouvelle forme d'évaluation et mes anciennes notations. A l'inverse, chaque compétence de la grille est prise individuellement pour construire le livret scolaire de l'élève, ce qui permet une analyse plus fine et une communication plus claire de ses acquis, tout en allant dans le sens d'une plus grande justice scolaire. Cela est en fait indispensable pour éviter les biais de notation et les effets dus aux représentations inconscientes du professeur sur l'élève : en effet, au moment de remplir la case « étude de la langue » par exemple, mes représentations joueront à plein si le seul outil technique sur lequel m'appuyer est un *b*, un *tb*, *ab*... tandis que si le même devoir me propose des niveaux d'acquisition par observables, j'ai des données solides sur lesquelles m'appuyer. Ceci me paraît d'autant plus important que la forme actuelle du LSU, trop générale, peut facilement encourager à amalgamer les résultats des élèves dans chaque matière plutôt qu'à essayer de les affiner. Ici, les notes de synthèse communiquées au-dessus de la grille ne sont donc qu'une indication pour les élèves et les parents, reprenant leur objectif initial qui est de *communiquer* sur les acquis mais non de les sanctionner.

Deux limites sont à noter toutefois : ma grille de compétences n'est pas en congruence totale avec le LSU, ce qui m'oblige à des aménagements parfois encore imparfaits. Il y aurait à mon sens à rendre le LSU plus fin dans les entrées qu'il propose : sa forme actuelle, qui limite la validation des acquis aux grands domaines de chaque matière, est en contradiction avec l'évaluation fine demandée par les textes. En attendant cette hypothétique transformation, il me faut donc trouver un moyen de rendre compte de mon évaluation de façon complète sans noyer les parents dans des informations trop nombreuses. Je dois également trouver comment donner sens à un degré d'acquisition donné dans les grands domaines : en effet, que signifie un « en cours d'acquisition » en étude de la langue quand les diverses évaluations témoignent d'acquis très différents entre la grammaire, le vocabulaire, la conjugaison... mais aussi au sein même de ces différents domaines ? Il y a donc une contradiction entre mes tentatives actuelles d'évaluation par compétence et le LSU qui répond encore, me semble-t-il, à une logique scorée du fait de sa simplification.

La deuxième limite notable de mon système concerne la note de synthèse : celle-ci est encore le fruit d'une logique de « score » en ce sens que je la tire d'un calcul entre les compétences acquises, non acquises et en cours d'acquisition. Au point actuel de mes lectures, je n'ai pas trouvé de solution permettant d'y remédier de façon satisfaisante. Le fait que ce score ne soit à aucun moment sommatif est cependant un point fondamental dans la transformation opérée, qui me laisse penser que l'objectif est atteint à ce niveau là.

c) Une communication de qualité

Le dernier avantage de cette grille est la valeur informationnelle qu'elle confère à l'évaluation. Nous venons de voir qu'elle me permettait de remplir le LSU de façon plus rigoureuse, or celui-ci est l'outil de communication le plus important pour renseigner sur le

parcours scolaire de l'élève, aussi bien durant l'année que pendant toute sa scolarité. Le fait d'aller vers un remplissage plus précis et plus proche des divers acquis a donc des implications indirectes profondes, qui jouent sur le parcours global de l'élève – n'oublions pas que les professeurs des niveaux supérieurs ont accès au LSU et que sa consultation mènera presque invariablement à la création de biais dans la représentation qu'ils se font des élèves. Biais qui, eux-mêmes, joueront ensuite sur sa notation et donc sur la suite de son parcours...

Plus pragmatiquement, le fait que la grille soit apparente va dans le sens d'un meilleur lien entre l'école et la famille : les parents d'élèves peuvent ainsi mieux comprendre la notation, ce qui est attendu des élèves, ainsi que les différentes lacunes et acquisitions de leur enfant. La note de synthèse ne sort plus *ex nihilo*, au contraire, elle est bien reliée aux différents items de la grille, ce qui permet une lisibilité accrue. De plus, les commentaires que je mets ne sont plus une sentence colorant de façon diffuse l'ensemble du devoir mais peuvent être associés plus facilement à des entrées de la grille. Les parents peuvent aller consulter les différents exercices concernés plutôt que de devoir faire une confiance aveugle au professeur ou explorer la copie de fond en comble. Pour ceux qui n'ont pas le temps ou l'envie de relire le devoir –très sûrement la majorité des parents - , cette grille est un moyen de détailler efficacement ce sur quoi il portait, ce qui permet de briser l'opacité de la logique évaluative et de communiquer plus largement sur ce qu'apprennent leurs enfants dans la classe.

Enfin, réfléchir en amont aux compétences travaillées et aux observables à évaluer me permet de dresser une liste très précise d'*attendus*. Ceci fait de cette grille un outil de communication vers les élèves eux-mêmes : elle me permet de leur communiquer, la veille du devoir ou quelques-jours avant, les compétences et connaissances sur lesquelles ils seront évalués. Mieux encore, ce sont eux qui, se remémorant ce qu'ils ont fait durant la séquence, sont amenés à dresser, sous ma guidance, la liste précise des attendus. Ceci leur permet donc de se ressaisir de l'évaluation, de vérifier s'ils y sont prêts et de s'entraîner de façon ciblée le cas échéant. La grille est également lue à haute voix en même temps que le reste de la copie le jour J et les exercices correspondant à chaque item sont désignés, ce qui permet aux élèves de développer éventuellement des stratégies de résolution, en se concentrant par exemple d'abord sur les notions qu'ils maîtrisent. Il est à noter à ce propos que les demandes pour faire les exercices dans le désordre se sont systématisées avec la mise en place de la grille ; tandis qu'elles restaient marginales auparavant. De même, et alors que cela n'avait pratiquement jamais eu lieu, il arrive fréquemment que les élèves me demandent durant la séquence si tel type d'exercice ou telle connaissance/compétence « sera dans le contrôle », ce qui marque une réappropriation stratégique du rapport à l'évaluation et une réflexion nouvelle sur les attendus.

Je note toutefois trois limites dans mes évaluations pour atteindre ces objectifs. D'une part, des évaluations comme celle de géométrie évaluent trop de compétences et connaissances, ce qui rend la grille peu lisible. Il pourrait être intéressant de penser à organiser les compétences en les regroupant par grands domaines, soit thématiques, comme « les polygones », soit liés au type de compétence, comme « je connais ma leçon », « je reconnais les figures », « je sais les écrire », etc... ce qui permettrait aux parents de conserver une certaine lisibilité. Mon deuxième

écueil est de ne pas avoir écrit les exercices auxquels correspondent les compétences évaluées, ce qui empêche une lecture facile de la structure de l'évaluation. Ce point, dommageable, sera assez facilement remédiable. Le troisième point concerne la façon dont je corrige les exercices liés à une compétence précise : force est de constater que je ne communique pas sur ce qui me mène, depuis une correction, à définir telle notion comme acquise et telle autre comme « en cours d'acquisition ». Concernant ce point-là, des difficultés persistent : je pense que la meilleure solution serait de faire construire aux élèves des grilles permettant l'autocorrection afin que les critères de validation leur apparaissent, à eux, naturellement. Concernant les parents cependant, aucune solution claire et simple ne m'est apparue pour l'instant.

Il est probable que cela soit dû au fait que ma correction, si elle est basée sur des critères précis, contient encore une part de subjectivité : c'est en effet, à mon avis, la plus-value du professeur que sa capacité à finement percevoir à travers des erreurs presque semblables des causes différentes. Ainsi, un élève peut avoir quelques fautes mais qui apparaissent clairement comme de l'inattention, tandis qu'un autre peut en avoir autant, mais qui témoignent d'une lacune sur un point précis. L'évaluation de ce type de différence naît de la connaissance des élèves et de la lisibilité des exercices proposés, qui doivent permettre précisément d'opérer la distinction. Si la conception d'une évaluation selon des critères plus techniques permet d'aider à la justesse de cette observation subjective en éliminant les « parasites », il ne me semble pas qu'elle doive l'occulter ou mener à sa disparition au profit d'une liste de critères totalement objectifs et coupés de toute dimension humaine. Revenons par exemple à l'évaluation de géométrie. Si la question 1 répond à des attendus très précis, dressés en amont – il faut que tous les termes nécessaires de la définition du polygone soient présents – la question trois, elle, laisse une certaine place à une diversité de réponses. J'ai ainsi accepté aussi bien celles qui venaient directement de la leçon – la figure n'est pas fermée ; tous ses côtés ne sont pas des segments – que certaines autres explications n'en reprenant pas exactement les termes mais indiquant une compréhension juste – ainsi de Sara qui me dit que la figure 1 n'est pas tracée à la règle, ce qui est un biais de réflexion juste et indiquant une bonne compréhension de l'ensemble même si la terminologie n'est pas réemployée exactement. De même, pour l'exercice 2, l'oubli d'une seule figure ou le fait de colorier la figure à l'extrémité en bas à gauche de la grille, qui pouvait pousser au doute, ne m'ont pas empêché de sanctionner l'acquisition de la compétence « je sais identifier un polygone » pour peu que le reste de l'exercice soit juste et que l'exercice 3 soit réussi à son tour. Il m'apparaissait en effet qu'une faute d'inattention ou due à un « piège » ne devait pas pénaliser un élève qui me montrait manifestement qu'il avait bien acquis les bases attendues. Il m'apparaît à cet égard que le passage d'une évaluation scorée à une évaluation par compétences a notamment pour effet de permettre une évaluation moins liée au score pur et laissant plus de place à l'observation fine. Ceci ne va peut-être pas dans le sens d'une clarté *absolue* des critères de notation pour les parents, mais est, à mon sens, plus indiquée pour une évaluation *juste* de l'acquisition des compétences par chacun.

3. Pallier l'effet évaluation

Le dernier point sur lequel j'aimerais faire un retour ne concerne pas la conception de l'évaluation en tant que telle mais tout ce qui permet de pallier les effets négatifs générés par le fait de se retrouver en situation d'évaluation, indépendamment de la qualité de celle-ci. J'ai en effet commencé, au fil de l'année, à interroger mes élèves à propos de leur vécu des évaluations, et il est apparu qu'il ne s'agissait pas pour eux d'un exercice anodin. Sachant qu'ils devront le faire signer, et que cela compte pour leur livret, cela génère chez eux une anxiété notable. Certains de mes élèves sont même restés « bloqués » dans une posture d'échec avant que je ne mette en place un certain nombre de procédés pour dédramatiser le moment sommatif, tant celui-ci prenait d'importance sur leur propre vision d'eux-mêmes. Ma prise de conscience s'est faite au fil de mes lectures, dont témoigne la partie I, mais également à certains signes des élèves. Ceux-ci, lorsque je leur annonçais un contrôle, manifestaient verbalement une certaine anxiété, pouvant aller jusqu'à des somatisations – maux de ventre – ou l'adoption d'une posture de repli – mutisme, air perdu, regard fixe. Les questions sur la nature du contrôle, sur les exercices qui allaient le composer, étaient multiples lors de l'écriture des devoirs, prenant à mon avis trop d'importance. Enfin, j'ai fini par effectuer un petit sondage juste avant de distribuer une évaluation sommative prévue afin que les élèves expriment le point auquel ils se sentaient confiants. J'ai été surpris du nombre d'élèves, certains très bons, qui ne se sentaient absolument pas prêts à affronter cette épreuve. Comprenant qu'il y avait là un facteur important à traiter, j'ai décidé de mettre en place plusieurs mesures.

La première est un rituel ponctuel, un moment de « mise-en-contrôle » construit expressément pour faire baisser la pression suscitée par l'évaluation. Pour commencer, j'ai généralisé les sondages, en les transformant un peu. Avant de poser les questions : « qui se sent prêt/ qui est un peu stressé mais de façon gérable / qui est très stressé », je demande « qui a bien révisé/ qui a révisé de façon correcte mais sans tout revoir / qui n'a pas révisé ». Comme les élèves ayant bien révisé sont généralement les plus nombreux, je peux souvent leur faire prendre conscience du décalage qui existe entre leur impression de ne pas être prêt et leur niveau de préparation effectif. Suit un petit atelier de relaxation pour asseoir la détente dans le corps : les élèves font quelques exercices de visualisation et de respiration que j'ai adaptés pour l'occasion. Il s'agit ensuite de créer l'événement et de coaliser le groupe : je distribue les feuilles à chacun, à l'envers, ils doivent attendre que tous les aient reçus pour pouvoir les retourner. Lorsque cela arrive, je leur donne deux minutes pour appréhender l'évaluation, lire les exercices, repérer déjà les points qui vont leur être faciles et ceux qui vont leur demander plus de concentration, de réflexion, etc... Arrive ensuite le moment de la lecture à voix haute, effectuée par moi, et que j'accompagne des explications nécessaires ; puis les élèves, si ces explications n'y ont pas déjà répondu, posent les questions qui leur sont venues quand ils se sont retrouvés face à la feuille. Une fois toutes les questions réglées, et après vérification que personne n'a rien à ajouter, je leur souhaite bonne chance et leur dis de donner le meilleur d'eux-mêmes. Cela rajoute un temps de dix à quinze minutes à l'évaluation ; cependant, je n'ai pas eu à allonger mes séances. Globalement, avec ce type de rituel préparatoire et des exercices mieux conçus, les

élèves finissent aujourd'hui plus vite des évaluations pourtant plus longues qu'auparavant. Concernant les exercices de respiration, j'ai parfois arrêté de les prodiguer ; je constate toutefois que certains élèves les font d'eux-mêmes avant le contrôle, ce qui tend à laisser penser que cette pratique les rassure effectivement. Parmi ceux-là, Alicia est un cas pour qui ce procédé a vraiment bien marché : élève timide et peu sûre d'elle, elle se retrouvait souvent bloquée de panique devant certaines questions. En mettant en place ces techniques lorsque la pression était handicapante, elle a réussi plusieurs fois à débloquer sa mémoire et à compléter les exercices demandés. Je dirais donc que c'est un moment important et positif, même si l'ampleur exacte de ses effets est difficile à cerner. Partant du constat qu'il ne m'oblige à réorganiser la durée de mes séances d'évaluation, j'ai décidé de continuer à le mettre en place.

Le second procédé est de nature plus profonde. J'ai réfléchi à la façon de tisser solidement dans la séquence d'apprentissage elle-même une préparation à l'évaluation sommative. J'ai en effet considéré que celle-ci en était indissociable et, puisque la première mène de toutes façons à la seconde, il était aussi bien de rendre ça très clair et de faire en sorte qu'elle lui serve réellement de préparation. Concrètement, j'ai fait des moments de remobilisation des phases fortes de mes séquences : répétées au début de chaque séance, d'une durée allant souvent jusqu'à dix minutes, elles ne servent pas seulement à rappeler aux élèves ce sur quoi nous sommes en train de travailler mais sont bien un moment de révision dédié. Les élèves peuvent y rappeler à l'oral les notions qu'ils ont intégrées et réciter les leçons qu'ils ont apprises l'un après l'autre : il n'est à aucun moment gênant que plusieurs élèves répètent la même leçon. J'en profite pour rappeler chaque jour, à ce moment-là, quels éléments rappelés par les élèves sont très importants à savoir pour l'évaluation et pour leur parcours scolaire. Durant cette phase, on ne rappelle pas seulement ce qui a été fait à la séance précédente mais tout ce dont on se souvient : ainsi, pour préparer mon évaluation de géométrie, j'ai demandé la définition du segment *tous les jours* aux élèves, en leur rappelant chaque fois que cela était fondamental à connaître et à se rappeler pour le contrôle. Je constate trois effets notables à cette pratique : d'une part, la participation des élèves est massive ; ils voient vraiment ce moment comme un temps permettant de vérifier en direct leurs connaissances et s'y essaient de bon cœur. En quelques séances, les moments de remobilisation ont arrêté de ne faire intervenir que les cinq meilleurs élèves ; bien au contraire, pris au jeu et s'appuyant sur les interventions de leurs camarades, certains de mes élèves les plus en difficulté ou ayant le moins confiance en eux se sont mis à participer. A mon avis, cela est dû au fait que je suis passé d'un temps de vérification, plutôt à tendance punitive, à un moment d'émulation. Le fait est que le seul enjeu qu'ont les élèves est une possible gratification de ma part, mais que l'erreur ne leur provoque aucun retour négatif. Le résultat est probant : là où je devais aller chercher les participants, je dois désormais souvent arrêter ce moment collectif avant d'avoir interrogé tous les volontaires. Le deuxième effet positif est que ce procédé permet de réintégrer dans le temps de classe l'apprentissage des notions. Le troisième est qu'il donne sens aux apprentissages en les liant à l'enjeu évaluatif. Le fait de rendre cet exercice routinier amène la leçon à être sue par un très grand nombre d'élèves : on peut noter que sur la question de la définition du polygone dont j'ai parlé, 23 élèves sur 28 répondent parfaitement. Globalement, cette évaluation est d'ailleurs très bien réussie, la

répartition de ma classe en trois groupes de niveau étant ici démentie par les résultats, sauf concernant l'orthographe des mots, la codification des triangles et la capacité à les reconnaître en mesurant les côtés. Or, il s'agit précisément du type d'exercice qui a été vu le plus tard dans la séquence et a le moins fait l'objet de questions en remobilisation. Je constate donc une corrélation intéressante, qui, associée au fait que mes précédentes évaluations ont toutes laissé apparaître trois groupes de niveau, semble aller dans le sens d'une efficacité de la méthode. Je reste toutefois prudent au vu du nombre de facteurs qui entrent en jeu dans le résultat d'une évaluation sommative. Notons toutefois, et c'est sûrement l'observable le plus parlant, que cela a poussé certains élèves comme Alicia et Sara – dont le profil ressemble à celui de sa camarade – à réviser d'elles-mêmes juste avant les montées en cours de mathématiques ou de français afin de pouvoir intervenir en début de séance, chose qu'elles n'avaient jamais faite auparavant.

Conclusion

Si l'évaluation sommative est un incontournable de la pratique professorale – comme l'atteste le référentiel de compétences des personnels de l'éducation nationale⁴⁰ – elle est surtout un outil pédagogique fondamental. Selon la façon dont elle est mise en place, elle peut mener à la destruction de l'estime des élèves ou accompagner sa croissance ; aider aux apprentissages ou y faire puissamment obstacle ; permettre la mise en place de l'égalité des chances ou être le vecteur de la reproduction sociale. Il est également important de noter que l'évaluation sommative ne saurait être vue comme un objet hors-sol, analysable en soi, isolable d'une pratique générale. En d'autres termes, l'évaluation est à la fois le symptôme et le résultat d'une pédagogie dans son ensemble, et l'une ne saurait être modifiée sans que l'autre le soit également. A cet égard, cette année de travail a été parlante : je suis passé d'une conception vide de l'évaluation à la prise de conscience qu'elle était le moyen par lequel les élèves étaient classés, les inégalités créées, et le professeur enfermé. Cet état de fait me semble inévitable aussi longtemps que l'évaluation n'est pas réinvestie consciemment par l'enseignant en vue de la faire changer de forme.

Les recherches et les expériences que j'ai menées m'ont permis de mettre en place un certain nombre d'outils dont la grille de compétences d'une part, et les préparations diverses au moment sommatif de l'autre, ont été les plus efficaces. J'aurais pu m'attarder sur d'autres formes testées durant l'année, mais la taille limitée de ce mémoire impliquait de faire des choix. Ces deux méthodes m'ont semblé pertinentes car elles sont un bon exemple de l'impact que peuvent avoir un nombre très limité d'outils sur une pédagogie. Si je devais dresser un bilan de ce qu'elles m'ont apporté, je noterais deux grandes victoires : la transformation de la nature de la note, et une conception beaucoup plus pertinente des exercices.

Je perçois cependant certaines limites. Il m'est par exemple difficile d'évaluer précisément l'impact de ces changements sur la réussite effective des élèves, notamment parce que l'évaluation fait intervenir trop de facteurs en même temps, mais aussi parce que mes systèmes de notation ont évolué en même temps que mes évaluations durant l'année. Je pense que c'est là l'écueil principal de ma recherche : s'il est certain que mes contrôles sont mieux conçus et que les élèves vivent mieux les évaluations – deux indicateurs importants –, je ne pourrais toutefois affirmer avec certitude qu'une amélioration des notes découle *directement* de cela. Il est cependant évident qu'une telle amélioration est *favorisée* par ces procédés, ce qui est un bilan à mon avis acceptable pour une première année d'enseignement.

Etant attendu que mon travail s'est uniquement concentré sur l'évaluation sommative, il me semblerait intéressant désormais d'orienter mes recherches vers les autres types

⁴⁰ *Référentiel des compétences professionnelles des métiers du professorat et de l'éducation*, Bulletin officiel, URL : [http://ww2.ac-poitiers.fr/doc/sites/doc/IMG/pdf/mene1315928a_-_ministere.ieur_et_de_la_recherche.pdf], consulté le 26/02/2019

d'évaluation. En effet, s'il est possible d'améliorer la conception d'une évaluation et d'en compenser les effets négatifs – mes deux chantiers principaux cette année –, il n'en reste pas moins que le « contrôle » est et restera un objet limité. Je souhaiterais, dans les années à venir, étudier les pédagogies ayant réussi à le dépasser et à intégrer le sommatif dans la continuité du temps de classe, sans que des moments spéciaux et identifiés lui soient dédiés. Espérons donc que ce mémoire est la première pierre d'un long chemin vers une pédagogie plus juste et de qualité...

Bibliographie :

- *L'évaluation à l'école, « pour la réussite de tous les élèves »*, dir. Alain Bentolila, Préface Philippe Meirieu, auteur Charles Hadil, Ed. Nathan, Coll « Questions d'enseignants ».
- *Evaluer autrement*, André Gagneux, PUF, coll. Education et formation, Paris, 2002.
- *Evaluer sans dévaluer, et évaluer les compétences* Gérard de Vecchi, hachette éducation, hachette, paris, 2011
- *La stratégie du choc*, Naomi Klein, Actes Sud, Arles, 2008 [trad]
- *Eloge de la fuite*, Henri Laborit, Gallimard, Paris, 1976
- *La colombe assassinée*, Henri Laborit, Grasset, paris, 1983

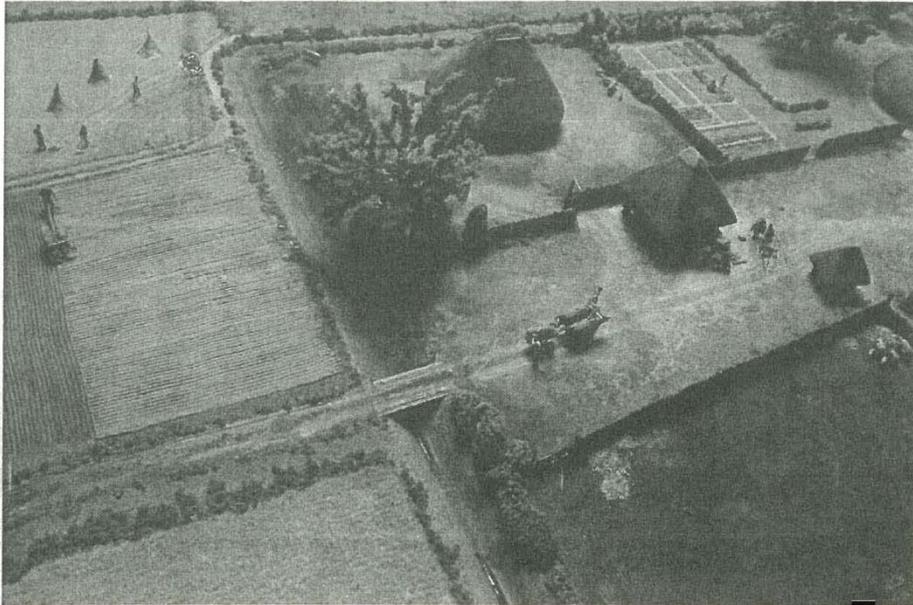
Sitographie :

- « *Docimologie critique : des difficultés de noter des copies et d'attribuer des notes aux élèves* », Institut de Recherche sur l'Éducation de Bourgogne, URL : [http://ired.u-bourgogne.fr/images/stories/Documents/Cours_disponibles/Demeuse/Cours/p6.4.pdf], consulté le 12 mars 2019
- « *L'évaluation par les notes : quelle fiabilité et quelles réformes ?* », CAIRN, URL : [https://www.cairn.info/revue-regards-croises-sur-l-economie-2012-2-page-218.html], consulté le 14 mars 2019
- « *Les trous noirs de l'évaluationnisme* », Pierre Frackowiak, URL : [http://www.cafepedagogique.net/lexpresso/Pages/2009/03/FrackLestrousnoirs.aspx], consulté le 18 février 2019
- Circulaire de rentrée 2014, URL : [https://www.education.gouv.fr/pid285/bulletin_officiel.html?cid_bo=79642], consulté le 29 janvier 2019
- *Référentiel des compétences professionnelles des métiers du professorat et de l'éducation*, Bulletin officiel, URL : [http://ww2.ac-poitiers.fr/doc/sites/doc/IMG/pdf/mene1315928a_ministere.ieur_et_de_la_recherche.pdf], consulté le 26/02/2019

Annexe n°1

FIRMIN *Stimmel* 1810-1880

1. Observe attentivement ce document :



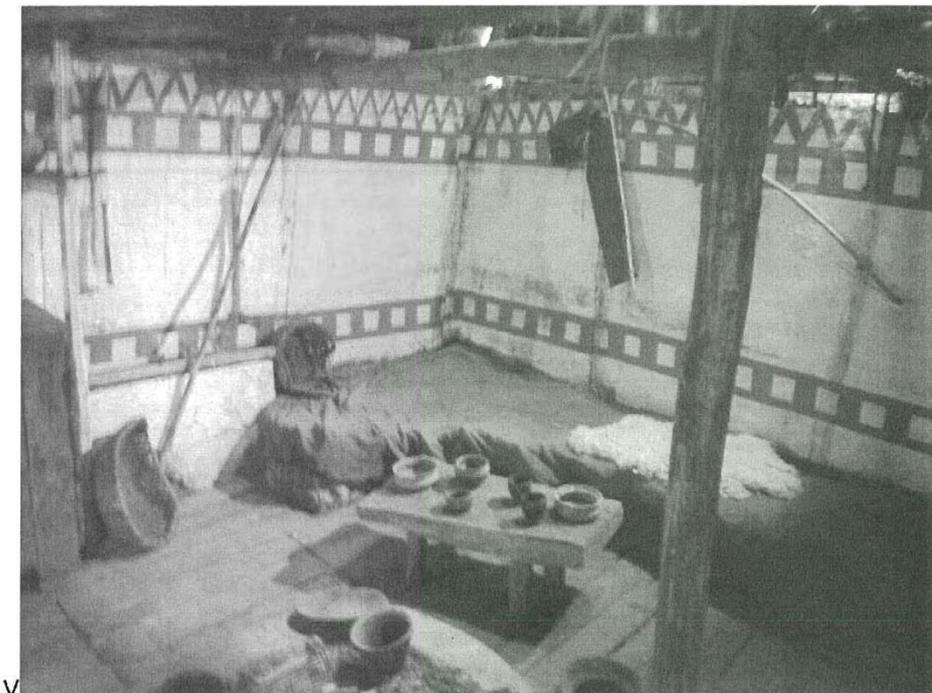
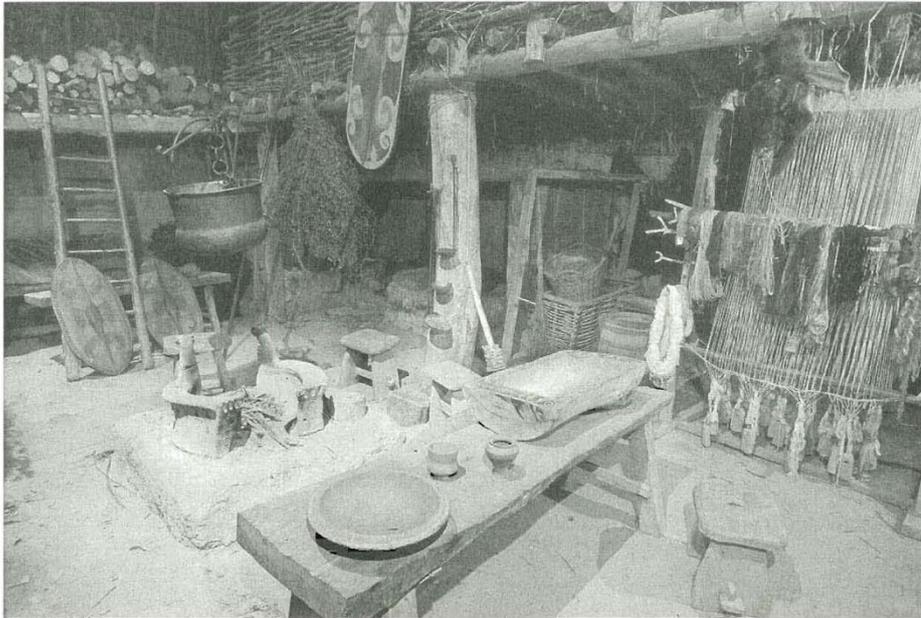
2. Recherche le plus d'informations possibles :

usages, quelques maisons en	potager / jardin
toit de paille / bœuf /	
moutons. Ils se	
nourissent d'agriculture /	
pont de bois / rivière	
enclot / sentier / sol	
de terre / champs / charreue	

Les gaulois vivent dans des plaines.
Ils se nourrissent d'agriculture.

FIELD RESEARCH

1. Observe attentively the document:



Carrille et Abelina

2. Recherche le plus d'informations possibles :

Ils cuisinent	Ils font du feu pour se réchauffer
Ils font de la poterie	Ils ont des paniers.
Ils écrasent du blé.	Ils ont pas de cadis
Ils ont des couvertures en laine	Ils ont pas de magasin
Ils ont des verres en bois ou en pierre	Ils ont pas d'électricité
Ils vivent près de la nature	Ils dorment pas sur des lits
Ils jouent au jeu de cartes.	Ils vivent dans des cabanons

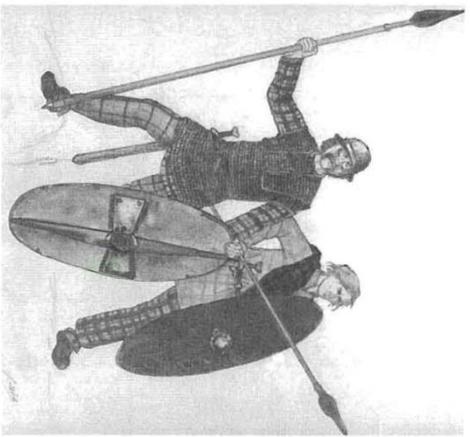
Ils vivent comme les hommes Préhistorique

Annexe n°2



Image de la bande-dessinée Astérix
Ce document est-il pertinent ? Pourquoi ?

Bien parce que ceci est un dessin de bande dessinée.



Ce document est-il pertinent ? Pourquoi ?

Bien parce qu'on me s'est pour d'eu il vient et qu'il y a pas de légende.



Pièce de monnaie de la tribu des Parisii
Ce document est-il pertinent ? Pourquoi ?

Bien parce que ceci est une pièce de monnaie de la tribu des Parisii.

Elise et Meina

Annexe n°3

hamza

th

om

- habillement : acquis.
- mode de vie : acquis
- analyse de documents : acquis

Evaluation d'histoire : les Gaulois

1 Complète le schéma en mettant les noms des habits au bon endroit :

La saie/ La tunique /Les braies/Les brogues



la saie /

la tunique /

Les braies /

les brogues

reconstitution d'une tenue gauloise

Comment s'appelle l'épingle qui sert à tenir le manteau chez les Gaulois ?

elle s'appelle la fibule.

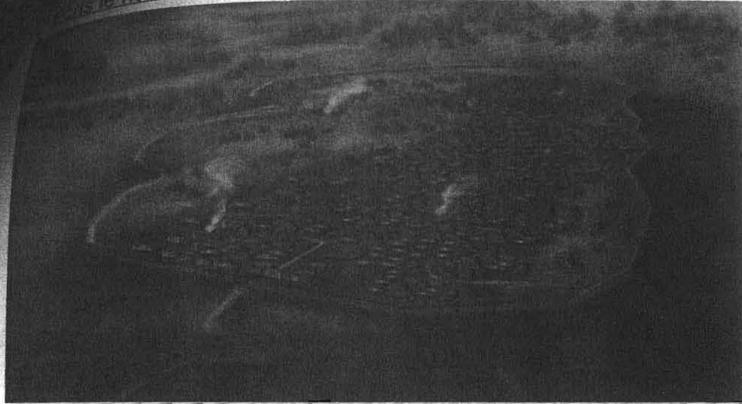
.....

.....

.....

hamza

... le nom de ce type de ville gauloise sur la ligne en pointillés :



opidum.

3 Que font les Gaulois dans ces villes ?

Ils font du commerce et ils y vont pour se protéger. / protéger

4 Où vivent-ils le reste du temps ?

Il vivent dans un vicus. /

5. Comment les gaulois se nourrissent-ils ?

agriculture
élever des animaux

6. Donne deux domaines dans lesquels les Gaulois sont doués :

harvoja
ils sont doués en agriculteurs et sculpteurs.
agriculteur sculpteur

QUESTION BONUS :

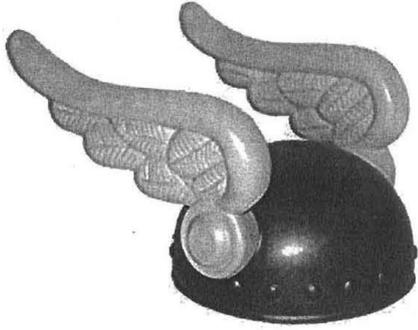


chaudron Gaulois en bronze du 1^{er} siècle avant Jésus-Christ (vestige archéologique)

Ce document est-il pertinent ? Pourquoi ?

oui parce que il y a une légende et
ils disent que c'est un vestige archéologique /

hamza



Jouet de casque gaulois vendu à « Toysrus » (magasin de jouet)

Ce document est-il pertinent ? Pourquoi ?

non, parce que ils disent qu'il est vendu à
toysrus.

Annexe n°4

Jona

TK *JK*

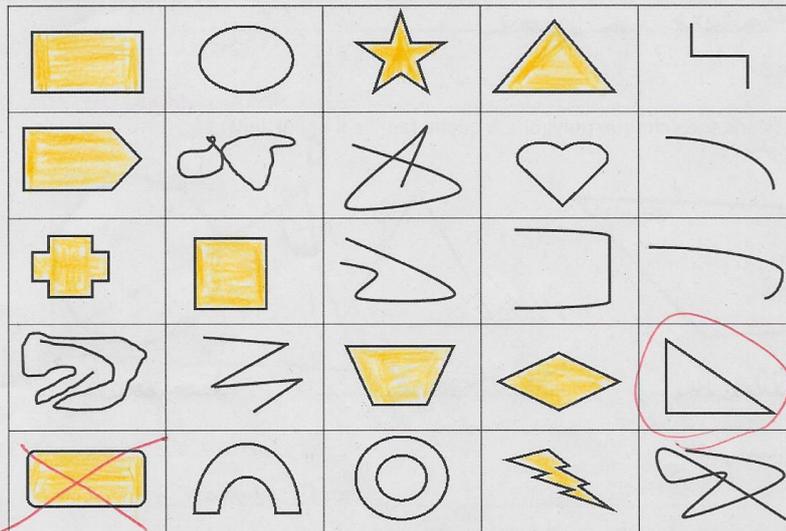
Evaluation de géométrie.

compétences	acquis	En cours d'acquisition	Non acquis
Je sais définir ce qu'est un polygone	X		
Je sais identifier un polygone	X	X	
Je sais identifier les différentes familles de polygones	X		
Je sais écrire les noms des différentes familles de polygones		X	
Je connais les propriétés des différents triangles particuliers et je sais les nommer	X		
Je sais reconnaître les différents types de triangles en me reportant aux codifications	X		
Je sais reconnaître les différents types de triangle en mesurant les côtés			X
Je sais écrire correctement les noms des différents types de triangles	X		
Bonus : je sais dessiner un triangle			X

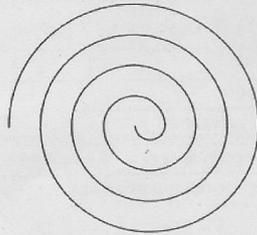
Exercice 1 : Qu'est-ce qu'un polygone ?

Un polygone est une figure géométrique qui a minimum 3 cotés, fermée et ses cotés sont des segments.

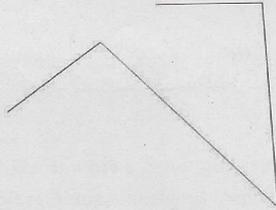
Exercice 2 : Robert Mitchum, le roi des polygones, a éternué très fort : le souffle de son éternuement a été tellement puissant qu'il a mélangé toutes les figures du royaume! Aide-le à remettre de l'ordre en coloriant uniquement les polygones.



Exercice 3 : Maria ne se souvient plus de la définition du polygone. Elle a tracé deux figures. Aide-là à se rappeler de la définition : sous chaque figure, écris s'il s'agit d'un polygone et pourquoi.

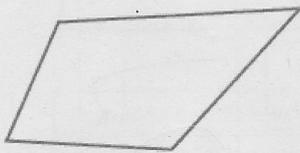


Non, ce n'est pas un polygone parce qu'elle n'est pas fermée, elle a moins de 3 côtés.



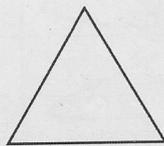
Non, ce n'est pas un polygone parce que la figure n'est pas fermée.

Exercice 4 : écris sous chaque polygone à quelle famille il appartient.

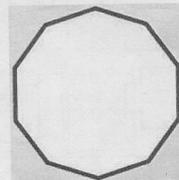


quadrilatère

quadrilatère

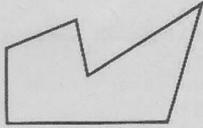


triangle

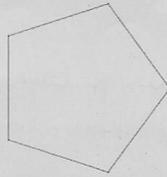


décagone

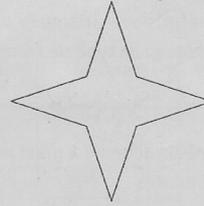
Zona



hexagone

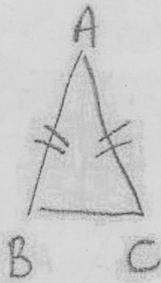


pentagone



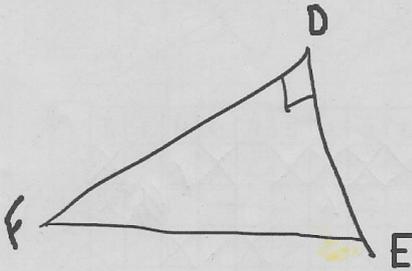
octogone

Exercice 5 : Henri a dessiné trois triangles à main levée : ces petits schémas vont lui permettre de mieux réaliser les vrais triangles car il aura déjà les informations nécessaires pour bien les tracer à la règle. Il a codé leurs particularités. Pour chacun d'eux, dis de quel type de triangle il s'agit et pourquoi :



ABC est un triangle... isocèle

car il a deux côtés égaux



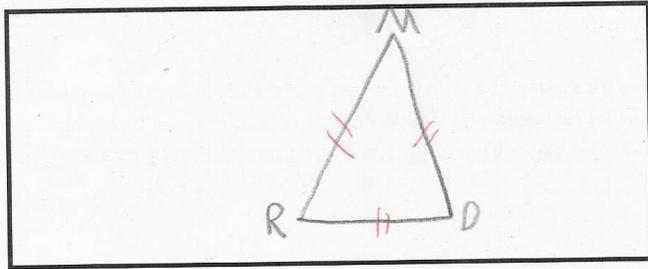
DEF est un triangle... rectangle

car il a un angle droit

→ Son troisième triangle s'appellera IJK. Il s'agit d'un triangle possédant trois côtés égaux. Comment appelle-t-on ce type de triangle ?

Cen appelle ce type de triangle le triangle équilatéral.

Dessine un schéma à main levée du triangle IJK en le codant de façon à aider Henri à connaître ses particularités :



Exercice 6 : Remplis le tableau en t'aidant des instruments de mesure appropriés.

	A	B	C	D	E	F	G	H	I	J
2 côtés de même longueur		X	X	X	X		X	X	X	X
3 côtés de même longueur				X					X	X
1 angle droit	X					X	X	X		
2 côtés de même longueur + 1 angle droit					X					

Exercice BONUS : Trace au dos de la feuille le triangle ABC avec $AB = 6$ cm, $BC = 8$ cm, $AC = 5$ cm

Annexe n°5

Nael

tt!

My

Evaluation de grammaire et de conjugaison

compétences	Acquis	En cours d'acquisition	Non acquis
Je sais conjuguer les verbes en -ER, -IR, -DRE, -EINDRE, -TIRE, aller, faire, dire, pouvoir, vouloir, valoir, avoir, être.	<i>X</i>		
Je sais définir les différents éléments qui composent le groupe nominal		<i>X</i>	
Je sais identifier les différents éléments du groupe nominal	<i>X</i>		
Je sais accorder l'adjectif dans le groupe nominal	<i>X</i>		

Exercice n°1. Recopie les phrases en conjuguant les verbes entre parenthèses au présent de l'indicatif.

1. Je (aller) à la plage. Tu (vouloir) venir ?

Il va à la plage. Tu veux venir?

2. Nous (être) fatigués, nous (aller) dormir.

Nous sommes fatigués, nous allons dormir.

3. Que (dire)-vous ?

Que dites-vous?

4. Que (manger)-nous ce midi ?

Que mangeons-nous ce midi?

5. Nous (éteindre) la lumière.

Nous éteignons la lumière.

Exercice n°2. Complète par *d* ou *t*.

Il pein*t*. - Elle enten*d*. - On répon*d*. - Il rejoin*t*. - Elle mor*d*. - Il crain*t*. - Elle pren*d*. - Il atten*d*. - On attein*t*.

tt!

iel

Exercice n° 3. Complète par s ou ds

Je vend s - Tu pein s - Je prend s - J'attend s - tu étein s - Tu fini s - Tu men s - Tu pon s - Tu crain s

Exercice 4. Qu'est-ce qu'un nom ?

Un nom désigne des personnes, des animaux et des lieux.

Exercice 5. Qu'est-ce qu'un adjectif ?

Un adjectif donne des informations sur le nom.

Exercice 6. De quoi est composé le groupe nominal ?

Il est composé d'un déterminant et d'un nom + et peut être d'un adjectif.

Qu'est-ce qui est obligatoire dans ces trois éléments pour former un groupe nominal correct ?

Il est obligatoire de mettre le déterminant et le nom.

Exercice 7. Souligne seulement les adjectifs.

Chien - vert - torche - grand - méchant - enfant - arbre - bleu - terrible - incroyable

Exercice 8. Recopie ces groupes nominaux en accordant correctement les adjectifs.

Des (grand) chiens → Des grands chiens.

Une (joli) maison → Une jolie maison.

Des citrons (vert) → Des citrons verts.

Des personnes (fatigué) → Des personnes fatigués.

Un homme (blond) → Un homme blond.

el

Exercice 9. Propose un groupe nominal pour chaque nom donné. Ajoute au moins un adjectif.
Attention aux accords !

- monstre : un monstre effrayant ✓
- chevaux : des beaux chevaux ✓
- horloge : une belle horloge ✓