

ÉVALUATION & NOTATION – Aspects institutionnels et historiques

1. Le professeur et l'évaluation

1.1. L'article L311-1 du *Code de l'éducation* énonce¹ :

La scolarité est organisée en cycles pour lesquels sont définis des objectifs et des programmes nationaux de formation comportant une progression annuelle ainsi que des critères d'évaluation.

Pour assurer l'égalité et la réussite des élèves, l'enseignement est adapté à leur diversité par une continuité éducative au cours de chaque cycle et tout au long de la scolarité.

L'enseignant doit ainsi procéder à des *évaluations*. L'article L912-1, issu de l'article 47 de la loi Fillon², précise plus complètement les missions de l'enseignant :

Les enseignants sont responsables de l'ensemble des activités scolaires des élèves. Ils travaillent au sein d'équipes pédagogiques ; celles-ci sont constituées des enseignants ayant en charge les mêmes classes ou groupes d'élèves ou exerçant dans le même champ disciplinaire et des personnels spécialisés, notamment les psychologues scolaires dans les écoles. Les personnels d'éducation y sont associés.

Les enseignants apportent une aide au travail personnel des élèves et en assurent le suivi. *Ils procèdent à leur évaluation*. Ils les conseillent dans le choix de leur projet d'orientation en collaboration avec les personnels d'éducation et d'orientation. Ils participent aux actions de formation continue des adultes et aux formations par apprentissage.

Ils contribuent à la continuité de l'enseignement sous l'autorité du chef d'établissement en assurant des enseignements complémentaires.

Leur formation les prépare à l'ensemble de ces missions.

Notons que la loi parle maladroitement d'évaluation *de l'élève*, alors qu'il s'agit d'évaluer les *travaux* de l'élève, c'est-à-dire ses « performances », regardées éventuellement comme exprimant de façon plus ou moins fidèle ses « compétences ». Dans une circulaire ministérielle du 13 juillet 2000 relative aux procédures disciplinaires, sur laquelle on revient plus loin, on lit ainsi, plus correctement³ : « Il convient (...) de distinguer soigneusement les punitions relatives au comportement des élèves de l'évaluation de leur travail personnel. » C'est à ce point de vue – celui de l'évaluation des travaux de l'élève (ou de la classe) – que l'on s'en tiendra ici, même quand quelque instance (conseil de classe, etc.) s'autorise du résultat d'une telle évaluation pour prononcer un vœu ou pour prendre une décision touchant la personne de l'élève.

1.2. Les programmes de 1995 des classes de collège sont peu disertés à propos d'évaluation. Mais ils soulignent au moins deux points importants.

1.2.1. Le document d'accompagnement du programme de 6^e en vigueur jusqu'en 2004-2005 précisait ceci :

Il convient (...) de distinguer les activités d'apprentissage qui, en général, conduisent à travailler simultanément plusieurs compétences, des activités d'évaluation qui sont, entre autres, l'occasion de vérifier la maîtrise des diverses compétences exigibles. Les compétences exigibles constituent le noyau qui doit être acquis par chaque élève mais les activités mathématiques dans la classe vont bien au delà de celles-ci.

Les « activités d'évaluation » ont pour objet de « vérifier la maîtrise » de certains « compétences », et notamment des compétences dites « exigibles », « noyau » de ce qui doit devenir les « acquis » des élèves ; par contraste, les « activités d'apprentissage », les « activités mathématiques dans la classe » sont plus complexes (elles « conduisent à travailler simultanément *plusieurs* compétences ») et « vont bien au delà ».

1.2.2. L'introduction du programme du cycle central (5^e et 4^e) reprend la distinction précédente entre « activités de formation » et « travaux d'évaluation » :

Les activités de formation, distinctes des travaux d'évaluation portant sur les compétences exigibles, seront aussi riches et diversifiées que possible. Elles seront aussi l'occasion de mobiliser et de consolider les acquis antérieurs dans une perspective élargie.

¹ Voir <http://www.legifrance.gouv.fr/WAspad/UnCode?code=CEDECATL.rcv>.

² Loi n° 2005-380 du 23 avril 2005 d'orientation et de programme pour l'avenir de l'école.

³ BO spécial numéro 8 du 13 juillet 2000.

Quant au programme de 3^e, il ajoute à ces principes une mise en garde contre l'inflation subreptice des exigences en matière d'autonomie didactique dans les activités d'évaluation :

On favorise (...) le développement des capacités d'initiative des élèves sans exigence prématurée d'autonomie lors des évaluations.

1.2.3. Sous le titre « Compétences et activités de formation », l'*Introduction générale pour le collège* figurant en avant-propos du nouveau programme de 6^e entré en vigueur en 2005-2006 confirme une distinction précédemment notée. On y lit ainsi ⁴ :

Il importe de bien garder à l'esprit que *la liste des compétences, si elle fixe les objectifs à atteindre, ne détermine pas pour autant les moyens pédagogiques à utiliser pour cela.*

L'explicitation par le programme des diverses « compétences » à acquérir est commentée en des termes qui soulignent la dialectique des activités de formation et des activités d'évaluation :

La définition de ces compétences vise donc à clarifier les attentes, à préciser les priorités et à fournir des repères dans le but d'aider les enseignants dans leur travail de programmation et dans la mise au point des évaluations qui permettent d'en baliser la réalisation.

La dialectique formation/évaluation peut se construire en prenant appui sur les indications apportées expressément par le programme, comme l'indique encore cette glose :

Le programme décrit, pour chaque contenu, les compétences élaborées dans chacune des classes du collège. Les commentaires qui les accompagnent apportent un éclairage supplémentaire sur les conditions de leur apprentissage.

1.2.4. Le même texte semble corriger quelque peu l'affirmation – qui pourrait être interprétée de façon excessivement prudente – énoncée par le programme de 3^e actuel relative à des attentes exagérées en matière d'autonomie didactique de l'élève ; ses rédacteurs écrivent en effet :

L'évaluation de la maîtrise d'une compétence par les élèves ne peut pas se limiter à la seule vérification de son fonctionnement dans des exercices techniques. Il faut aussi s'assurer que les élèves sont capables de la mobiliser d'eux-mêmes, en même temps que d'autres compétences, dans des situations où leur usage n'est pas explicitement sollicité dans la question posée.

1.2.5. Dans un développement consacré à l'évaluation, l'*Introduction* déjà citée précise encore ceci :

L'évaluation (qui ne se réduit pas au contrôle noté) n'est pas un à-côté des apprentissages. Elle doit y être intégrée et en être l'instrument de régulation, pour l'enseignant et pour l'élève. Elle permet d'établir un constat relatif aux acquis de l'élève, à ses difficultés. Dans cette optique, le travail sur les erreurs constitue souvent un moyen efficace de l'action pédagogique. L'évaluation ne doit pas se limiter à indiquer où en est l'élève ; elle doit aussi rendre compte de l'évolution de ses connaissances, en particulier de ses progrès.

Il y a là un concentré de la doctrine contemporaine « moyenne » en matière d'évaluation, doctrine sur laquelle on revient plus en détail dans la suite de cette notice.

1.3. Le programme de la classe de 2^{de} s'inscrit dans la continuité des principes gouvernant les programmes de collège de 1995, même si les « compétences exigibles » y deviennent, plus modestement, des « capacités attendues ».

1.3.1. Une contrainte nouvelle est ici prise en compte : la classe de 2^{de} est une classe *carrefour*. L'*Introduction* du programme s'en explique en ces termes :

Pour chaque chapitre, les capacités attendues, en nombre volontairement limité, constituent la base commune sur laquelle se fonderont les programmes des années ultérieures.

Le document d'accompagnement précise :

Ce programme est celui d'une *classe de détermination*. Il s'agit principalement de constituer une base de connaissances exploitables pour les années ultérieures quelle que soit l'orientation choisie mais aussi de faire percevoir modestement quelques volets scolaires de l'activité mathématique pour fonder le choix d'orientation. En conséquence, *les capacités attendues sont volontairement limitées.*

1.3.2. Ce même document comporte une section intitulée « Évaluation et orientation » qui explicite encore la problématique précédente. Rappelant que, « en seconde de détermination, les élèves ont à préciser leur choix d'orientation » et que « les enseignants jouent un rôle important pour les aider et ensuite prendre les décisions que leur confient les textes réglementaires », les rédacteurs écrivent :

⁴ BO hors série n° 5 du 9 septembre 2004.

Le processus d'orientation est à mener tout au long de l'année. La communication des objectifs à atteindre et leur évaluation régulière permettent aux élèves de mieux se situer.

L'organisation et le suivi du travail personnel des élèves, évoqués dans le paragraphe précédent, sont des éléments incomparables d'information au service d'un jugement conscient et éclairé, et d'une action pertinente pour la réussite de chaque élève.

Ainsi s'introduit un thème neuf : l'évaluation, essentielle à la régulation de la formation, est également articulée à l'*orientation*, processus crucial s'il en est.

1.3.3. À cette novation correspond un renouvellement des remarques des programmes du collège sur la distinction entre activités de formation et activités d'évaluation : celle-ci, interne à l'effort pour acquérir les « compétences exigibles », est maintenant doublée d'une distinction entre formation « de détermination » et formation « au-delà ». L'*Introduction* du programme de 2^{de} la présente ainsi :

Le programme qui suit est écrit dans le cadre d'une seconde de détermination. Il est composé de trois grands chapitres : statistique, calcul et fonctions, géométrie. De plus, un ensemble de thèmes d'études est proposé, dans lequel l'enseignant pourra puiser au gré du questionnement et des motivations de ses élèves ; ces thèmes, entourant le contenu du chapitre, permettent de faire vivre l'enseignement au-delà de l'évaluation sur les capacités attendues et de prendre en compte dans une certaine mesure l'hétérogénéité des classes. L'enseignant a toute liberté pour choisir les thèmes au-delà de ces propositions.

L'affirmation concernant le rôle possible de cette « formation au-delà » dans la gestion de « l'hétérogénéité des classes » est reprise de façon concise dans le document d'accompagnement du programme :

La proposition de travailler quelques thèmes autour d'un chapitre, au delà de l'évaluation, peut être un moyen de mieux gérer cette hétérogénéité.

Au-delà de cet ensemble de principes et de préceptes (qui relèvent du registre de la technologie didactique), les textes officiels commentent avec une certaine prolixité les dispositifs d'évaluation.

2. Les dispositifs d'évaluation dans les textes officiels

2.1. Le document qui accompagne le programme de 2^{de} comporte cette notation :

Dans les multiples formes que peut prendre l'évaluation, apparaissent les devoirs de contrôle : le nombre de ceux-ci doit être réduit (trois, voire deux, par trimestre suffisent) ; on y intégrera plusieurs composantes de l'activité mathématique signalées par le programme. La moyenne annuelle de ces contrôles, souvent déterminante pour l'orientation de chaque élève, sera pondérée par un bilan chiffré des divers travaux (cahier de statistique, travail sur ordinateur...) portant sur les sujets difficiles à évaluer lors de ces contrôles.

2.1.1. À côté de l'évaluation comme outil de formation est évoquée ici le processus de fabrication d'une « moyenne annuelle », qui est l'intermédiaire entre l'évaluation interne à la classe et la décision d'orientation qui sera prise dans le cadre de l'établissement. On notera que l'exigence quantitative touchant les « devoirs de contrôle » est modeste : trois par trimestre suffisent.

2.1.2. Le passage précédemment cité mentionne « les multiples formes que peut prendre l'évaluation ». Cette multiplicité est liée à la pluralité des types de travaux demandés aux élèves, travaux qui feront (ou non) l'objet d'une évaluation. L'ancien programme de 6^e comportait à cet égard l'indication suivante :

Le travail personnel des élèves en classe, en étude ou à la maison, est essentiel à leur formation. Il a des fonctions diversifiées :

- la résolution d'*exercices d'entraînement*, combinée avec l'étude du cours, permet aux élèves d'affermir leurs *connaissances de base* et de les mettre en œuvre sur des exemples simples ;
- les travaux individuels de *rédaction* sont nécessaires au développement des capacités d'expression écrite et de la *maîtrise de la langue* ;
- les *devoirs de contrôle*, courts et peu nombreux, permettent de vérifier les acquis des élèves.

Seuls les devoirs de contrôle semblent devoir être à tout coup *évalués* (généralement sous la forme de l'attribution d'une *note*).

2.1.3. Le nouveau programme de 6^e est plus direct, puisqu'il parle non seulement d'*évaluation* mais encore d'évaluation *sommative* (qualificatif sur lequel on va revenir), par quoi il faut sans doute entendre ici une évaluation dont le résultat entre dans la composition de la « note annuelle » (ou « trimestrielle ») :

L'évaluation sommative, en mathématiques, est réalisée sous trois formes complémentaires :

- des interrogations écrites courtes dont le but est de vérifier qu'une notion ou une méthode sont correctement assimilées ;
- des devoirs de contrôle courts et peu nombreux qui permettent de vérifier, de façon plus synthétique, la capacité des élèves à utiliser leurs acquis, à la suite d'une phase d'apprentissage ;
- certains devoirs de contrôle peuvent être remplacés par un bilan trimestriel qui est l'occasion de faire le point sur les acquis des élèves relatifs à une longue période d'étude.

2.1.4. Dans l'*Introduction commune à l'ensemble des disciplines scientifiques* précédant les programmes des classes de 5^e et de 4^e qui entreront en vigueur, respectivement, en 2006-2007 et 2007-2008, un rappel est fait – sous le titre « L'évaluation comme repère des apprentissages » – de la place de l'évaluation dans la conduite de la formation. On notera la référence aux trois moments en lesquels une évaluation (ou du moins une analyse) est appelée à intervenir :

Vérifier les acquis fait partie intégrante de l'action pédagogique.

L'évaluation est un outil indispensable au professeur dans la conduite de son enseignement, à différents moments de son apprentissage.

En début, comme en cours d'apprentissage, le repérage des acquis, des difficultés et des obstacles permet d'adapter les supports et les modalités de l'enseignement.

Le bilan terminal permet de mesurer la maîtrise qu'a chaque élève des savoirs et des savoir-faire visés et, si nécessaire, d'envisager des activités de remédiation.

2.2. Des précisions analogues à celles qui précèdent peuvent être trouvées, au lycée, dans les programmes des classes de 1^{re}, qui reprennent explicitement des formulations présentes déjà dans les programmes auxquels ils ont succédé.

[Dans l'acquisition de la démarche mathématique,] les travaux proposés en dehors du temps d'enseignement, à la maison ou au lycée, jouent aussi un rôle primordial ; ils ont des fonctions diversifiées :

- la résolution d'exercices d'entraînement, combinés avec l'étude du cours, permet aux élèves d'affermir leurs connaissances de base et d'évaluer leur capacité à les mettre en œuvre sur des exemples simples ;
- l'étude de situations plus complexes sous forme de préparation d'activités en classe ou de problème à résoudre et à rédiger alimente le travail de recherche, individuel ou en équipe, et permet aux élèves d'évaluer leur capacité à mobiliser leurs connaissances dans des secteurs variés ;
- les travaux individuels de rédaction (solution d'un problème, mise au point d'exercices étudiés en classe, rapport de synthèse sur un thème d'étude, analyse critique d'un texte...) visent essentiellement à développer les capacités de mise au point d'un raisonnement et d'expression écrite ; vu l'importance de ces objectifs, ces travaux de rédaction doivent être fréquents mais leur longueur doit rester raisonnable ;
- les devoirs de contrôle, peu nombreux, combinent des exercices d'application directe du cours (voire des questions de cours), des problèmes plus synthétiques, comportant des questions enchaînées de difficulté progressive et permettant aux élèves de vérifier leurs résultats, et des problèmes plus ouverts (susceptibles d'amener l'élève à choisir un modèle mathématique approprié, à émettre une conjecture, à expérimenter à travers des exemples ou des contre-exemples, à construire un raisonnement) ;
- l'exploitation de documents, individuelle ou en équipe, contribue au développement des capacités d'expression écrite (rédaction d'un rapport) ou orale (mise au point d'un exposé).

2.2.1. L'ancien programme de 2^{de} imposait des exigences plus explicites encore ; on comparera la formulation qui y est adoptée (ci-après) avec celles des programmes actuellement en vigueur :

Les *devoirs de contrôle, peu nombreux*, combinent des exercices d'application directe du cours et des problèmes plus synthétiques, comportant des questions enchaînées de difficulté progressive et permettant aux élèves de vérifier leurs résultats. Ils doivent être suffisamment *courts* pour permettre à la grande majorité des élèves d'étudier l'ensemble des questions posées et de *rédiger posément* la solution qu'ils proposent.

2.2.2. Une note du groupe des mathématiques de l'Inspection générale de l'éducation nationale datée du 27 mars 1997 et intitulée *Les travaux écrits des élèves en mathématiques au collège et au lycée* a synthétisé et repris certains éléments touchant l'évaluation. Dans une section ayant pour titre « L'évaluation en temps limité », les rédacteurs de cette note écrivent notamment :

Il convient de garder un rapport correct entre l'évaluation et la formation : c'est l'évaluation qui est au service de la formation, et non le contraire. En particulier, il ne faut pas négliger, par un choix judicieux des épreuves, le rôle formateur de l'évaluation.

Il convient de faire se côtoyer deux types d'épreuves écrites d'évaluation :

- les interrogations écrites courtes (10 à 20 min) dont le but est de vérifier qu'une notion, une méthode ou une démonstration est correctement assimilée. On peut en prévoir une par chapitre du cours (soit une par quinzaine en moyenne) ;
- les devoirs de contrôle (de 30 min en 6^e à 3 ou 4 h en terminale) sont peu fréquents (2 à 3 par trimestre) et doivent rester de difficulté et de longueur raisonnables. Ils ne doivent en aucun cas déborder du programme de la classe, ni faire appel à des notions ou des méthodes qui n'y sont pas étudiées.

2.2.3. La même note comporte une section intitulée « La correction des copies et la notation » que l'on reproduit ici.

Les objectifs de formation poursuivis à travers les travaux écrits (à la maison et en classe) doivent être communiqués et régulièrement rappelés aux élèves. C'est en rapport avec ces objectifs que la correction et la notation des copies doit prendre son sens : la clarté des raisonnements, la qualité de la rédaction et le soin apporté à la présentation jouent un rôle essentiel. Dans cette optique, il convient d'annoter les copies par des appréciations écrites, des conseils, des remarques constructives.

La pertinence du calibrage de la notation constitue un objectif important : il convient d'éviter tant la sur-notation, génératrice d'illusion, que la sous-notation, génératrice de découragement.

On a là, apparemment, l'unique mention du problème de la *notation chiffrée*.

3. Le foisonnement des « évaluations »

3.1. Longtemps le terme d'évaluation est demeuré inconnu dans les usages qui sont aujourd'hui les siens dans l'institution scolaire. Le *Dictionnaire de la langue pédagogique* de Paul Foulquié, paru en 1971, l'ignore, alors que cet ouvrage traite longuement de la *note*.

3.1.1. Un récent rapport ministériel brosse l'introduction du mot dans les termes suivants – on notera le lien fondateur entre « évaluation » et « objectifs pédagogiques »⁵.

La première promotion officielle du mot apparaît dans la revue du ministère *L'éducation* du 13 juin 1974, sous la signature de Lucien Géminard et Victor Marbeau, respectivement inspecteur général, représentant permanent au Conseil de l'Europe et inspecteur d'académie, chargé de mission aux affaires internationales du ministère. L'article rend compte de différentes conférences internationales : sont évoquées les critiques faites aux examens traditionnels, la volonté de promouvoir des systèmes de contrôle continu, d'unités capitalisables, de définitions de programmes par objectifs, de procédures d'évaluation formative et d'auto-évaluation. La question de la définition des objectifs pédagogiques est ensuite reprise par René Haby dans le *Courrier de l'Éducation* du 12 avril 1976 : le ministre souligne notamment la priorité qu'il convient de donner « aux objectifs sur les programmes au sens étroit du terme ». En 1977, Lucien Géminard organise un séminaire sur l'évaluation ; à cette occasion est publiée, à destination de tous les inspecteurs généraux, une brochure intitulée *Notions générales sur l'évaluation*. Enfin, consécration suprême, la pédagogie par objectifs et, avec elle, l'évaluation entrent en 1977 dans les programmes nationaux qui définissent « des objectifs à atteindre à la fin du cycle, objectifs formulés en termes de comportement et de compétence (savoir et savoir-faire) » (*Contenus de formation à l'école élémentaire, cycle préparatoire*, MEN, CNDP, 1977).

3.1.2. En cette matière comme en d'autres, la France apprend les doctrines venues d'Outre-Atlantique. À titre d'illustration, on s'arrêtera un instant, dans ce qui suit, sur un ouvrage que nous citerons dans sa troisième édition, parue en 1976 (les deux premières éditions sont de 1965 et de 1971, respectivement) : sous le titre *Measurement and Evaluation In Teaching*, il est dû à Norman E. Gronlund, professeur de psychologie de l'éducation à l'Université de l'Illinois ; il aura une sixième et dernière édition en 1990. La préface présente le schéma que le contenu proposé développera en un peu moins de 600 pages :

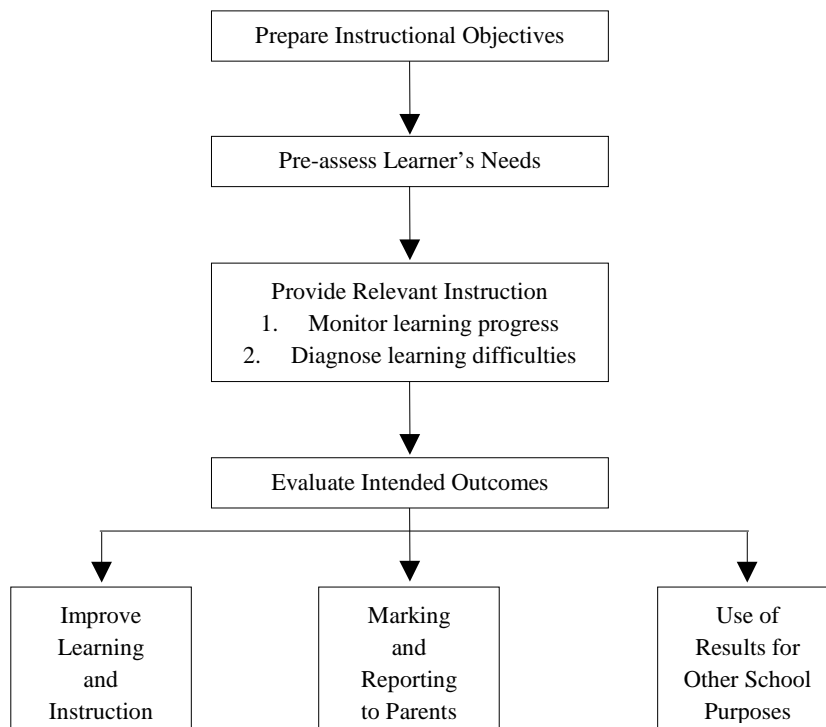
... classroom evaluation is viewed as an integral part of the teaching-learning process. It involves three fundamental steps: (1) identifying and defining the intended learning outcomes, (2) constructing or selecting tests and other evaluation instruments that are relevant to the specific outcomes, and (3) using the results to improve learning and instruction.

3.1.3. La table des matières de la première partie de l'ouvrage – *The Evaluation process* – est reproduite ci-après : on y saisit à nouveau le nouage entre évaluation et objectifs de formation⁶.

⁵ Inspection générale de l'éducation nationale & Inspection générale de l'administration de l'éducation nationale et de la recherche, *Les acquis des élèves, pierre de touche de la valeur de l'école ?*, Ministère de l'Éducation nationale, de l'Enseignement supérieur et de la Recherche, Paris, juillet 2005.

1 – The Role of Evaluation in Teaching	3
2. – Preparing Instructional Objectives	28
3. – Relating Evaluation Procedures to Instructional Objectives	60
4. – Validity	79
5. – Reliability and Other Desired Characteristic	105

Ce nouage s'inscrit en un schéma classique et fondamental, que l'on reproduit ci-après.



3.1.4. En réalité, ce schéma va se trouver rapidement surchargé par des apports multiples et variés, dont le premier et le plus célèbre est la distinction entre ce qu'un auteur travaillant sur l'évaluation *des curriculums*, Michael Scriven, baptise (en anglais) *summative evaluation* et *formative evaluation*, notions que l'auteur que nous suivons présente – dans leur cadre d'origine – dans les termes suivants :

... evaluation of the curriculum itself also plays an important role in curriculum development. During the early stages, when new methods and materials are being tried, evaluation data enable the curriculum developer to determine the effectiveness of the new procedures and to identify areas where revision is needed. When the new curriculum program has been fully developed, evaluation data make it possible to determine the degree to which the new curriculum is effective in meeting the instructional objectives for which it was designed. The first type of curriculum evaluation has been called *formative evaluation* and the second *summative evaluation* [M. Scriven, "The Methodology of Evaluation," in *Perspectives of Curriculum Evaluation*, AERA Monograph Series on Curriculum Evaluation, No 1 (Chicago: Rand McNally, 1967)].

3.1.5. On aura noté que la réalité dont il s'agit d'assurer le « développement » n'est nullement un élève mais... un *curriculum* ou, comme on disait en français autrefois, un *cours d'études*. Le même auteur explicite ce fait ainsi :

The main purpose of formative evaluation is to improve the instructional methods and materials so that greater student learning will result. The main purpose of summative evaluation is to appraise the overall effectiveness of a curriculum program.

⁶ S'agissant des notions de *validity* et de *reliability* mentionnées ici, on se limitera à reproduire le propos de l'auteur en tête des chapitres qui leur sont respectivement consacrés. Chapitre 4 : "In selecting or constructing an evaluation instrument the most important question is: To what extent will the results serve the particular uses for which they are intended? This is the essence of validity." Chapitre 5 : "Next to validity, reliability is the most important characteristic of evaluation results. Reliability (1) provides the consistency which makes validity possible, and (2) indicates how much confidence we can place in our results. The practicality of the evaluation procedure is, of course, also of concern to the busy classroom teacher."

3.1.6. L'adaptation de ce couple notionnel au cadre de la classe va se faire en articulation avec un grand nombre de notions auxiliaires. L'auteur du manuel examiné propose ainsi d'abord de distinguer entre cinq « évaluations » dont on reproduit partiellement la description ⁷.

Placement Evaluation. Placement evaluation is concerned with the pupil's entry behavior and typically focuses on questions such as the following: (1) Does the pupil possess the knowledge and skills needed to begin the planned instruction? For example, does the beginning algebra student have a sufficient command of computational skills? (2) To what extent has the pupil already mastered the objectives of the planned instruction? Sufficient mastery might indicate the desirability of the pupil's skipping certain units or of his being placed in a more advanced course. (3) To what extent do the pupil's interests, work habits, and personality characteristics indicate that one mode of instruction might be better than another (e.g., group instruction versus independent study)?

Formative Evaluation. As noted earlier, formative evaluation is used to monitor learning progress during instruction. Its purpose is to provide continuous feedback to both pupil and teacher concerning learning successes and failures. Feedback to students provides reinforcement of successful learning and identifies the specific learning errors that need correction. Feedback to the teacher provides information for modifying instruction and for prescribing group and individual remedial work.

Diagnostic Evaluation. Diagnostic Evaluation is concerned with the pupil's persistent or recurring learning difficulties that are left unresolved by the standard corrective prescriptions of formative evaluation. If a pupil continues to experience failure in reading, mathematics, or other subjects, despite the use of prescribed alternate methods of instruction (e.g., programmed materials, visual aids), then a more detailed diagnosis evaluation is indicated. To use a medical analogy, formative evaluation provides first aid treatment for simple learning problems, and diagnostic evaluation searches for the underlying causes of those problems that do not respond to first aid treatment.

Summative Evaluation. Summative Evaluation typically comes at the end of a course (or unit) of instruction. It is designed to determine the extent to which the instructional objectives have been achieved and is used primarily for assigning course grades or for certifying pupil mastery of the intended learning outcomes.

3.1.7. Une autre distinction doit être notée, qui oppose évaluation « critériée » et évaluation « normée ». L'auteur pris pour guide la présente dans les termes suivants ⁸.

There are two basic ways of interpreting pupil performance on tests and other evaluation instruments. One is to describe his performance in terms of the specific behavior he can demonstrate (e.g., He can type 40 words per minute without error). The other is to describe his performance in terms of the relative position he holds in some known group (e.g., He can type better than 90 per cent of his classmates). The first type of interpretation is called *criterion referenced*; the second is *norm referenced*. Both types of interpretation are useful. Criterion-referenced interpretations enable us to describe what an individual can do, without reference to the performance of others. For example, we can judge pupil performance by comparing it to some absolute standard that has been set (e.g., He can define at least 80 per cent of the terms in the unit). Norm-referenced interpretations enable us to determine how an individual's performance compares to that of others. This might be a classroom group, or some local, state, or national group, depending on the use to be made of the results. Using national norms, for example, we might describe a pupil's performance on a vocabulary test as exceeding that of 76 per cent of a national sample of sixth-graders. It should be noted that with norm-referenced interpretation we are *not* describing what percentage of the vocabulary items the pupil answered correctly, but simply what per cent of the pupils in the norm group he surpassed.

3.2. Bien d'autres distinctions ont été envisagées, dans lesquelles on n'entrera pas ici. Mais ce qui précède doit d'abord attirer l'attention sur ce fait que, en France et en Europe, l'essentiel des efflorescences notionnelles subsumées sous le mot d'évaluation sont venues des États-Unis, parfois sans qu'on prenne même le temps de les traduire.

3.2.1. Alors en effet que l'adjectif *formative*, par exemple, est courant en anglais (l'expression *the formative years* désigne ainsi les « années de formation », ou, plus classiquement, les « années d'apprentissage »), l'adjectif « formatif », qui existe en français, y est d'un emploi des plus rares : un étudiant angliciste se ferait sans doute chapitrer s'il rendait *the formative years* par *les années formatives* ! L'adjectif *sommatif*, lui, est, en français, une création *ex abrupto*. Dans le français des mathématiques, par exemple, on parle traditionnellement de formule *sommatoire* (et non

⁷ On peut traduire *placement evaluation* par « évaluation d'entrée » ou « de niveau » ou « pronostique ».

⁸ Un *sixth-grader* est un élève de 6^e.

« sommative ») pour désigner ce que les mathématiciens de langue anglaise appellent d'ailleurs *summation formula*.

3.2.2. La disponibilité de l'adjectif « formateur, -trice » permettra, dans les années 1980, de nommer un autre « type » encore d'évaluation : l'évaluation *formatrice*. Née d'une recherche conduite en 1974-1977 dans un lycée marseillais par une équipe emmenée par G. Nunziati, cette conception de l'évaluation aura une large diffusion auprès des professeurs de l'académie d'Aix-Marseille, comme le souligne dès 1986 l'un des chercheurs qui s'en firent les promoteurs ⁹ :

... dans la seule académie d'Aix-Marseille, depuis 1982, 600 enseignants ont suivi cette formation continue assurée par un groupe de formateurs dynamiques, animée par G. Nunziati, et par des universitaires de sciences de l'éducation, alors que plus de 1000 enseignants se sont inscrits sur la liste d'attente, volontaires pour cette formation dans cette même académie.

3.2.3. Au plan de la théorie, l'évaluation « formatrice » découle d'un constat relatif à l'évaluation formative ordinairement théorisée et pratiquée, constat que le même chercheur explicite dans les termes suivants – en précisant au passage la genèse de l'appellation assignée à cette forme « améliorée » d'évaluation formative.

On voit que s'il peut y avoir un accord général sur la spécificité de l'évaluation formative, sur son rôle global et peut-être sur les fonctions qu'elle assure, il n'y a pas toujours un accord sur l'importance relative des différentes boucles ; en particulier la gestion active *par l'élève* de la majeure partie des boucles de régulation qui concerne ses propres activités apparaît souvent secondaire dans la littérature, au point que notre collègue québécois G. Scallon suggérait naguère que nous utilisions à Aix-en-Provence un autre terme pour désigner la régulation assurée par l'élève lui-même et proposait le terme d'« évaluation formatrice » pour la distinguer de l'évaluation formative dont la régulation concerne en priorité les stratégies pédagogiques du maître, et non les stratégies d'apprentissage de l'élève.

3.2.4. En 1990, un spécialiste des questions d'évaluation, Jean Cardinet, présentera ainsi le dispositif auquel conduit la mise en œuvre en classe des préceptes de l'évaluation formatrice ¹⁰.

Appliquant les conceptions de Bonniol (1986), des enseignants français ont essayé de conduire leurs élèves à une prise en charge personnelle de la gestion de leur apprentissage. Pour cela, ils ont mis au point une démarche pédagogique où les élèves apprennent à s'auto-évaluer, mais guidés par l'enseignant, qui précise au départ les objectifs à atteindre, les critères à prendre en compte et les indicateurs permettant de déterminer si la performance est, ou non, satisfaisante (Nunziati, 1984).

Concrètement, Vial (1987) montre comment constituer d'abord une « carte d'étude » avec la classe, en analysant des travaux déjà corrigés, pour y repérer ce qui est attendu de l'élève et les erreurs à éviter. Les élèves ayant ainsi intériorisé la norme d'évaluation rédigent ensuite leur « plan de formation » individuel, base de leurs « contrats de travail » successifs avec l'enseignant. Les pré-requis peuvent alors être complétés, si nécessaires ; les exercices d'entraînement peuvent être centrés sur les opérations qui posent le plus de problèmes à l'élève ; les notions que ce dernier a le sentiment de mal dominer suscitent de sa part une recherche documentaire appropriée.

La discussion maître-élève amène à confronter la logique de la matière et celle de l'apprenant, comme à introduire au moment voulu le métalangage nécessaire pour parler des opérations à effectuer. Ce travail d'exploration en commun permet de modifier, chaque fois que cela paraît utile, la carte d'étude collective et les plans de formation individuels, qui sont ainsi « négociés », au sens où ils sont le résultat d'interactions multiples, l'enseignant apprenant à découvrir ce qui est possible pour les élèves et ce qui fait obstacle pour certains, les élèves essayant de se représenter ce qui est attendu d'eux et le chemin pour y parvenir...

3.3. La complexité et la lourdeur de l'organisation didactique associée aux idées de l'évaluation formatrice, jointes à d'autres faiblesses dans lesquelles on n'entrera pas ici ¹¹, ont sans doute limité sa

⁹ Jean-Jacques Bonniol, « Recherches et formations : pour une problématique de l'évaluation formative », in De Ketele, J.-M. (éd.), *L'évaluation : approche descriptive ou prescriptive ?*, De Boeck-Wesmael, Bruxelles, 1986, p. 119-133.

¹⁰ « Évaluation interne, externe ou négociée ? », in *Hommage à Jean Cardinet*, Delval, Fribourg, 1990, p. 139-156. Le travail de J.-J. Bonniol mentionné dans cet extrait est celui que nous avons cité ci-dessus. Le travail de G. Nunziati a pour titre « Évaluation formative et réussite des élèves » : il est paru dans le numéro 2 de la revue *Collège* publiée par le rectorat de l'académie d'Aix-Marseille. Le travail de M. Vial s'intitule *Un dispositif d'évaluation formatrice en expression écrite au collège* (CRDP, Marseille).

¹¹ Pour une présentation circonstanciée de l'évaluation formatrice et une discussion critique large, voir l'ouvrage de Jean-Jacques Bonniol et Michel Vial, *Les modèles de l'évaluation* (De Boeck Université, 2000), notamment p. 279-298.

durée de vie, même si l'intérêt que lui ont manifesté nombre d'enseignants a laissé plus que des souvenirs dans l'académie d'Aix-Marseille en particulier.

3.3.1. Certains des principes sur lesquels l'évaluation formatrice s'est construite, sinon l'organisation pratique de sa mise en œuvre, ont été durablement intégrés dans la pensée contemporaine de l'évaluation. Ainsi en va-t-il de l'idée que l'évaluation n'est pas confinée dans le *topos* du professeur. D'une façon générale, ses principes permettent d'interroger les modalités de l'évaluation, ainsi que le fait par exemple telle auteure dans ce passage qui ouvre le chapitre de son ouvrage¹² intitulé « Pratiquer l'évaluation formative en classe » : on y notera la mention de ce dispositif emblématique qu'est la « grille d'évaluation », associée ici à l'évaluation « normative », notions sur lesquelles on va revenir.

... quand un enseignant ou quand un élève coche sur une grille d'évaluation normative les critères réussis et les critères échoués dans un produit réalisé, il n'y a pas aide à l'apprentissage : souligner des réussites ne suffit pas pour entraîner de nouvelles réussites dans une situation plus complexe ; souligner des échecs entraîne rarement – voire jamais – la transformation de ces échecs en réussite. Car vérifier ce qui est acquis, ou non acquis, n'est pas apprendre.

3.3.2. L'évaluation « normative » dont il est question ici ne doit pas être confondue avec l'interprétation *norm referenced* d'une évaluation évoquée plus haut : visant à « vérifier ce que chaque élève a acquis », elle se donne pour outil de base le dispositif de la fiche de suivi, individuel ou collectif, interne (destiné aux enseignants et/ou aux élèves) ou externe (destinée aux parents, à d'autres établissements, etc.). Une fiche proposée par l'auteure déjà citée et qui, comme c'est fréquemment le cas, prend la forme d'un tableau assez complexe, a trait ainsi au suivi individuel de la compétence « Rédiger un texte à dominante descriptive » : elle est reproduite ci-après¹³.

Fiche de suivi individuel de la compétence : Rédiger un texte à dominante descriptive										
Cette fiche est destinée à l'usage interne : enseignants, équipe pédagogique, élèves										
Nom de l'enfant :										
Formes d'écrits	C	Dépliants touristiques			Paysages			Portraits		
Produits évalués	O	en 2	en 4	en 6	≈ 3	≈ 6	≈ 10	≈ 3	≈ 6	≈ 10
Critères d'évaluation	E	parties	parties	parties	lignes	lignes	lignes	lignes	lignes	lignes
	F	environ	environ	environ						
	F									
Critères concernant la globalité du texte :										
<i>Aspect matériel</i>										
• Il y a un titre	1									
• L'écriture est lisible	1									
• Le texte est propre et aéré	1									
<i>Situation d'énonciation</i>										
• L'enfant a introduit la description par une phrase qui en présente le thème général	2									
• L'enfant a utilisé des articulateurs spatiaux	2									
• Les articulateurs sont placés correctement.	2									
<i>Sémantique</i>										
• Les différentes parties de la description se rapportent au thème général	2									
• La description est ordonnée dans l'espace	3									
• L'enfant a utilisé un procédé d'expression pour terminer la description	3									
• L'enfant a évité la répétition des adjectifs utilisés	2									
• Les champs lexicaux sont riches et variés	2									
• Les anaphores sont employées correctement	2									

¹² Geneviève Meyer, *Évaluer : Pourquoi ? Comment ?*, Hachette Éducation, Paris, 1995, p. 110.

¹³ Pour plus d'information sur ces dispositifs, on se reportera à l'ouvrage indiqué.

• Les adjectifs et/ou les verbes sont imagés, évocateurs	2									
• La progression de l'information de phrase en phrase est cohérente et sans rupture.	2									
<i>Morphosyntaxe</i>										
• L'élève a utilisé le même temps au long de la description	2									
• L'ordre des groupes de mots dans chaque phrase est correct par rapport à la progression de l'information	2									
• La plupart des GN sont enrichis par des adjectifs ou des compléments de nom ou des propositions relatives.	2									
Critères concernant chaque phrase :										
• Les énumérations sont correctement ponctuées	1									
• Les mots sont orthographiés correctement	1									
• Les accords du GN sont corrects	1									
• Les accords SV sont corrects.	1									

3.3.3. L'ambition affichée est ici d'estimer la maîtrise d'une *compétence*, c'est-à-dire d'apprécier la capacité construite à effectuer un certain type de tâches – en l'espèce, « rédiger un texte à dominante descriptive » : on ne se contente pas d'évaluer une *performance*, c'est-à-dire la réalisation par l'élève d'une tâche particulière du type en question. Si le souci d'apprécier la maîtrise par l'élève des compétences désignées comme exigibles par le programme est présente dans les textes officiels gouvernant l'enseignement des mathématiques – on a vu plus haut que les rédacteurs du programme de 3^e parlent par exemple de « l'évaluation de la maîtrise d'une compétence par les élèves » –, il convient pourtant de ne passer qu'avec de beaucoup de circonspection d'une performance observable à l'affirmation d'une compétence supposée la permettre et s'exprimer à travers elle. En nombre de cas, au reste, les formulations officielles font montre d'une telle retenue, parlant par exemple, ainsi qu'on l'a vu, de *vérifier* la maîtrise des diverses compétences exigibles, ou de travaux d'évaluation *portant* sur les compétences exigibles, ou encore d'évaluation *sur* les capacités attendues.

4. La question des notes

4.1. L'ouvrage qu'on vient de citer présente une fiche de suivi individuel des « compétences-synthèses » en sciences destinée aux parents et censée servir à l'évaluation de diverses compétences, telles celles d'« utiliser une démarche expérimentale » ou de « mémoriser des connaissances. La fiche précise ceci.

Les performances de l'enfant sont codées de 1 à 4 pour chacune des compétences :

- 1 : activité réussie (notes correspondantes : de 16 à 20)
- 2 : activité moyennement réussie (notes correspondantes : de 10 à 15)
- 3 : activité insuffisamment réussie (notes correspondantes : de 5 à 9)
- 4 : activité non réussie (notes correspondantes : de 1 à 4)

La permanence de la référence à la note chiffrée (et, dans le cas français, de la note sur 20) est ici frappante.

4.1.1. La note a un long et lourd passé scolaire. Historiquement, elle est d'abord un instrument de discipline : la note sanctionne. Dans ses *Souvenirs d'enfance et de jeunesse*, Ernest Renan (1823-1892), qui fut élève du petit séminaire de Saint-Nicolas du Chardonnet créé et alors dirigé par Félix Dupanloup (1802-1878), écrit à ce propos : « Il n'y avait aucune punition dans la maison ; la lecture des notes et les réflexions du supérieur étaient la sanction qui tenait tout en haleine et en éveil. » Lui faisant écho, Dupanloup écrit sans ambiguïté, dans le volume III de son ouvrage *De l'éducation* (1850) :

Les notes (...) doivent être bien commentées. Ce commentaire peut (...) humilier, dompter, écraser ; ou bien, au contraire, consoler, relever, enflammer, et cela sans qu'il soit besoin de longues phrases ; un mot, et souvent quelquefois même un geste, un regard, c'est assez.

4.1.2. On comprend mieux ainsi que, traversant les siècles, le réflexe se soit conservé de recourir à la « mauvaise note », voire au *zéro*, pour fustiger un comportement inadéquat. Une telle pratique est aujourd'hui proscrite. La circulaire du 13 juillet 2000 déjà citée indique plus généralement, à cet égard :

Il convient également de distinguer soigneusement les punitions relatives au comportement des élèves de l'évaluation de leur travail personnel. Ainsi n'est-il pas permis de baisser la note d'un devoir en raison du comportement d'un élève ou d'une absence injustifiée. Les lignes et les zéros doivent également être proscrits.

Un tel interdit a suscité des commentaires « sauvages ». À cet égard, les auteurs d'un ouvrage intitulé *Le droit de la vie scolaire* écrivent¹⁴ :

En vérité, cette réaction démontre une nouvelle fois une profonde méconnaissance du droit. Le juriste, s'il s'en tient au seul exemple du zéro, ne peut que se révéler perplexe. Cette note n'est interdite tout autant qu'elle viendrait sanctionner un comportement répréhensible ou une absence injustifiée. Elle demeure légale si elle apprécie le niveau scolaire de l'élève. La circulaire, vécue comme restreignant le pouvoir de sanction, ne fait ici que rappeler l'un des quatre cas du recours d'ouverture pour excès de pouvoir : le détournement de pouvoir. Bien avant l'été 2000, l'élève qui s'adressait au juge administratif pour demander l'annulation d'une note nulle pour inconduite notoire avait de sérieuses chances de voir son recours aboutir. En effet, le pouvoir (l'attribution du zéro) utilisé dans un but (la sanction) autre que celui pour lequel il est conféré (la pédagogie) est depuis cent vingt cinq ans [CE, 26 novembre 1875, *Pariset*, GAJA n° 4] sanctionné par nos juridictions. Dans cette optique il est probable que seront considérés comme prohibés les zéros pour fraude, pour absence calculées (refus délibéré de participer à un contrôle par exemple). Le travail non rendu sans excuse valable [L'élève excusé ne peut se voir attribué un zéro pour ses absences ou pour n'avoir pas remis un devoir (TA Melun, 4 septembre 2001, *Lebras c/recteur de l'Académie de Créteil*, req. 013518/5)] autorise, en revanche, son emploi [TA Montpellier, 8 juin 2000, req. 981837]. Ces méfaits s'analysent plus en un comportement disciplinairement répréhensible que comme l'aboutissement d'une tâche devant faire l'objet d'une notation. Le juge administratif s'est d'ailleurs indirectement prononcé sur la question en annulant la décision d'une commission d'appel qui s'opposait au passage d'un élève en classe de 1^{re} S en se fondant notamment sur des notes « zéros » venues sanctionner des retards [TA Paris, 14 septembre 2000, *M. Niang*, req. 0010322/7. Dans le même sens, cf. CE, 18 décembre 1968, *Brunne*, Rec. CE, p. 658].

4.1.3. Un autre trait que l'histoire permet de mettre en perspective est l'usage ancien d'utiliser, pour noter, le demi-point, voire le quart de point : c'est que, longtemps, les notes sont là, dans le cadre d'une « pédagogie de l'émulation », pour *classer* les élèves. Commentant les *Instructions de 1890* (dont les rédacteurs s'efforcent de combattre l'obsession du classement individuel), les auteurs du rapport *Les acquis des élèves...*, déjà mentionné, écrivent :

En somme, ce système de notation a une triple fonction :

- Il vise à récompenser ou à punir les élèves pour le travail fourni et pour leur comportement scolaire (il est d'usage d'attribuer une note « de conduite » ou « de morale »).
- Il classe et compare les élèves entre eux, afin de susciter l'émulation.
- Il renseigne les autorités scolaires et les parents sur les mérites ou démérites de chaque élève et permet ainsi des sanctions publiques comme les prix, les tableaux d'honneur, les félicitations ou les blâmes, comme aussi le passage en classe supérieure (récompense) ou le redoublement (punition).

Dans ce système, l'enseignant exerce le pouvoir discrétionnaire, sinon arbitraire, d'un juge qui distingue les « bons » et les « mauvais » élèves.

4.2. Dans la première moitié du XX^e siècle, toutefois, un changement va s'amorcer : il est lié aux travaux que le psychologue Henri Piéron (1881-1964) proposera d'englober sous le terme de *docimologie* (du grec *dokimé*, épreuve).

4.2.1. Les premiers résultats obtenus en 1936 dans le cadre de l'enquête Carnegie sur les examens et concours confirmaient clairement ce que d'aucuns supputaient depuis longtemps. Sur un lot de 100 copies de mathématiques du baccalauréat, confiées à 5 correcteurs, et pour lesquelles on disposait donc de 6 notes (en incluant celle obtenue au baccalauréat), les enquêteurs ont calculé les écarts deux à

¹⁴ Yann Buttner, André Maurin, Blaise Thouveny, *Le droit de la vie scolaire*, Dalloz, Paris, 2002, p. 181.

deux, soit 15 écarts pour une copie et 1500 écarts pour le lot de 100 copies. La même procédure a été suivie pour six autres disciplines : les résultats sont consignés dans le tableau suivant ¹⁵.

	Écart moyen	Écart le plus fréquent	Écart maximum
Composition française	3,29	6 et 7	13
Version latine	2,97	5	12
Anglais	2,24	4	9
Mathématiques	2,05	4	12
Philosophie	3,36	5 et 7	12
Physique	1,88	4	8

Piéron ajoutait :

En aucun cas n'a été trouvée une note identique dans un couple de correcteurs, et, en envisageant une coupure au niveau de la note critique de 10, il s'est trouvé que, dans les couples de correcteurs, la note, atteinte ou dépassée avec l'un, restait inférieure avec l'autre dans une proportion de copies atteignant 70 % pour la composition française, 50 % pour la version latine, 47 % pour l'anglais, 36 % pour les mathématiques, 81 % pour la philosophie et 50 % pour la physique.

4.2.2. Les travaux réalisés alors reposaient sur l'idée de la « vraie note » d'un travail scolaire, dont les notes effectivement attribuées s'éloigneraient plus ou moins. Dans cette perspective, Henri Laugier (1888-1973) et Dagmar Weinberg (1897–1946), à qui l'on doit les résultats précédents, se posèrent un problème que Piéron expose dans les termes suivants :

Étant donné le fait de la fluctuation des notations relevant de ce facteur aléatoire qu'est la personnalité du correcteur, on est en droit de penser qu'en multipliant les correcteurs on compensera ces fluctuations. Laugier et D. Weinberg ont appelé valeur « vraie » la moyenne d'un nombre assez grand de notations indépendantes. Ils ont cherché à déterminer le nombre minimum d'examineurs compétents auxquels il faudrait faire appel pour obtenir la notation méritant confiance.

Le nombre minimal répondant aux conditions spécifiées par Laugier et Weinberg est indiqué dans le tableau ci-après, pour l'ensemble des disciplines déjà examinées : fortement variable, il est élevé, même dans le cas des mathématiques.

Composition française	78
Version latine	19
Anglais	28
Mathématiques	13
Philosophie	127
Physique	16

4.3. De multiples travaux approfondiront ces premières observations.

4.3.1. Dès 1965, J.-J. Bonniol met en évidence des « effets d'ordre » dans la correction d'un paquet de copies : les copies figurant dans le premier tiers des copies corrigées sont mieux notées ; une copie venant après une copie faible est sur-notée, et inversement.

4.3.2. Lorsqu'il n'y a pas anonymat des auteurs des copies, on observe des effets du statut scolaire et social de l'élève : l'évaluation d'un travail antérieur de l'élève qu'a pu réaliser le correcteur influe sur l'évaluation de la copie de cet élève. Quand il en a connaissance, le correcteur est influencé de même par l'évaluation de travaux de l'élève qu'ont pu réaliser d'autres correcteurs.

4.3.3. Dans d'autres expériences, on a par exemple confié la correction d'un lot de copies – en fait des copies de bac choisies aléatoirement – à des correcteurs à qui elles étaient présentées, pour certaines comme provenant d'un lycée « bourgeois », pour d'autres d'un lycée de la banlieue ouvrière : la note moyenne des premières copies s'est révélée supérieure à la moyenne des secondes. De tels travaux ont

¹⁵ Henri Piéron, *Examens et docimologie*, PUF, Paris, 1963, p. 20.

de même mis en évidence l'effet de l'origine sociale, du sexe, de l'apparence physique, du contexte de scolarisation, etc.¹⁶

4.4. Les pratiques de notation, longtemps figées en dépit des travaux évoqués, vont, dans la période qui suit Mai-68, faire l'objet d'une circulaire du 6 janvier 1969, qui tente de les modifier en profondeur¹⁷.

4.4.1. La circulaire rappelle notamment l'origine de l'usage contemporain de noter sur 20 et les critiques élevées contre cet usage par les travaux de docimologie :

C'est un texte ancien, l'arrêté du 5 juillet 1890, qui a prescrit que « dans les compositions chaque copie aura sa note chiffrée de 0 à 20 ». Il en résultait un « classement linéaire », les différences entre élèves se chiffrent par points ou même par demi-points et quarts de point. Or, les études docimologiques dont l'origine est antérieure à 1930 et qui se sont multipliées dans les vingt dernières années ne laissent aucun doute sur le caractère illusoire d'un tel raffinement dans la précision de la note et du classement obtenus.

4.4.2. La circulaire rappelle ainsi les premiers efforts pour se déprendre du point de vue ancien :

Nos méthodes d'appréciation du travail scolaire sont depuis longtemps en évolution. D'anciennes instructions ont eu pour objet d'exercer une influence régulatrice en ramenant à des proportions raisonnables le rôle des compositions dans la vie scolaire. C'est ainsi qu'une circulaire du 31 août 1928 notait que : « Certes l'émulation est un facteur essentiel de travail et de progrès, mais à condition que ne se développe pas chez les enfants un esprit d'âpreté et de lutte qui risque de les inciter à des procédés douteux et où il faut voir l'une des raisons de la fraude aux examens. Il importe de maintenir à ces exercices le caractère d'une saine compétition sportive ; il faut n'y voir au surplus qu'un simple épisode de la vie scolaire et non le but unique des efforts d'un trimestre. Faute de quoi le travail perdrait sa régularité et son calme. À des surmenages passagers succéderaient fatalement des périodes de relâchement et d'indolence. »

4.4.3. Dès 1890, la critique des classements et de la compétition pour les « places » est formulée très officiellement. Les « compositions » vont donc être remplacées par ce qu'on nomme aujourd'hui des « contrôles ». La circulaire indique à cet égard :

En substituant à la « composition » l'exercice de contrôle – dépouillé de cérémonial mais mieux compatible avec la régularité et le calme du travail – et en corrigeant les procédés usuels de notation, on sera tout naturellement conduit à éliminer ces « places » proclamées, qui provoquent chez tant d'élèves tantôt une anxiété aussi nuisible à leur équilibre général qu'à leur développement intellectuel, tantôt une indifférence plus ou moins résignée ou rétive, tantôt la dérision, parfois des vanités ridicules ou un esprit de rivalité quelque peu agressive ou mesquine, et qui sont aussi à l'origine de bien des conflits familiaux, accablants pour l'enfance, irritants pour l'adolescence.

4.4.4. Pour les rédacteurs de la circulaire, « les trois notions essentielles de notre système d'appréciation des résultats scolaires, notions de composition, de note, de classement, doivent faire l'objet d'une triple révision, de trois réformes indissolublement liées, celles des procédés de notation servant d'instrument aux deux autres. » Ces réformes s'appuient sur des principes tout modernes – les contrôles y participent visiblement d'une évaluation formative –, que le texte explicite en ces termes :

Les travaux scolaires les plus formateurs sont ceux où la préoccupation de la note s'efface : maître et élèves avancent ensemble dans la découverte d'un texte, d'un raisonnement, d'une expérience scientifique, d'une activité sportive, d'une donnée de géographie humaine, etc., et ce n'est qu'à regret que le fil est interrompu pour permettre les contrôles cependant nécessaires. Une pédagogie véritablement active réussit d'ailleurs, sans difficultés, à inclure le contrôle dans le champ même de l'élaboration des connaissances. Le contrôle permet en effet au maître d'orienter de manière plus efficace les directions de son action. Sans doute l'élève a-t-il besoin de voir son travail apprécié, ses efforts motivés et sa progression jalonnée. Sans doute les parents comme les autorités scolaires ont-ils besoin d'informations précises. Il faut cependant éluder l'obsession de la note, presque aussi pernicieuse que l'obsession de la « place », comme l'ont observé depuis longtemps bien des maîtres expérimentés.

4.4.5. C'est en ce point que la circulaire dévoile son objet central :

À cet effet, il est bon d'abord de prendre conscience de la relativité de la note, et par suite d'écarter les procédés dont la précision apparente est trompeuse. La notation chiffrée de 0 à 20 peut être abandonnée sans

¹⁶ Pour une synthèse sur tous ces points, on se reportera à l'ouvrage de Pierre Merle, *Sociologie de l'évaluation scolaire* (PUF, coll. « Que sais-je ? », 1998).

¹⁷ La circulaire est parue dans le *BO* n° 2 du 9 janvier 1969.

regret. Une échelle convenue d'appréciation, libérée d'une minutie excessive, sera moins prétentieuse. En indiquant la zone dans laquelle l'élève se situe, on cerne déjà la réalité d'assez près, on évite de multiplier systématiquement des différences qui ne seraient pas confirmées par d'autres correcteurs, ni par le même correcteur à une autre époque. Des appréciations globales telles que « très satisfaisant », « satisfaisant », « moyen », « insuffisant », « très insuffisant » auxquelles on peut faire correspondre, si on le juge bon, les symboles A, B, C, D, E, ou 1, 2, 3, 4, 5, constituent donc un système non pas plus rudimentaire que le système traditionnel, mais plus rationnel et mieux adapté aux données. Il sera bien entendu utile à l'élève que cette appréciation globale s'accompagne d'annotations plus détaillées, concernant par exemple, l'orthographe, l'ordre, le vocabulaire, la syntaxe, la précision, l'habileté, les facultés de raisonnement, l'invention, le sens artistique, etc.

4.4.6. La circulaire, signée du ministre de l'Éducation nationale de l'époque, Edgar Faure (1908-1988), se fait alors plus injonctive :

Dès maintenant, il est recommandé aux chefs d'établissement et aux enseignants, professeurs et instituteurs :

1. de substituer à la notion de composition traditionnelle celle d'exercices de contrôle divers, faits en classe, en un temps limité et présentant les caractères ci-dessus décrits ;
2. de substituer à l'échelle de notation traditionnelle de 0 à 20 une échelle simplifiée d'appréciation globale du type ci-dessus défini, ou d'un type analogue ;
3. d'exclure en général les classements par rang, établis et annoncés par le maître.

Deux de ces trois points passeront dans les mœurs scolaires ; le troisième, la notation en lettres, ou, plus exactement, l'usage d'une échelle de notation réduite, ne parviendra pas à se substituer au système traditionnel de notation sur 20.

Annexe. – La « note de vie scolaire »

1. Il s'agit d'une note dont l'institution est inscrite dans l'article 32 de la loi du 23 avril 2005 dite loi « d'orientation et de programme pour l'avenir de l'école » (ou loi Fillon).

2. On trouvera une explication de ce dispositif d'évaluation dans la circulaire n° 2006-105 du 23 juin 2006 (<http://www.education.gouv.fr/bo/2006/26/MENE0601604C.htm>). On a reproduit ci-après l'essentiel de ce qu'en dit le site ÉduSCOL (<http://eduscol.education.fr/D0111/noteviescolaire.htm>).

La note de vie scolaire

La note de vie scolaire est une mesure concrète prévue par la loi d'orientation et de programme pour l'avenir de l'École, qui fait de l'éducation à la responsabilité un enjeu de formation avec pour maîtres mots : l'éducation au respect (respect des autres, respect du règlement) et l'apprentissage des valeurs de la République.

Un enjeu de formation

L'objectif est de donner aux élèves des repères de comportement qui leur permettent de s'intégrer dans la société et d'être, à l'âge adulte, de vrais citoyens. C'est aussi de marquer le lien entre la scolarité, la vie scolaire et la vie sociale.

Cible

La note de vie scolaire concerne les élèves de collège, de la sixième à la troisième. Elle est mise en place à la rentrée 2006.

Contenu

La note de vie scolaire évalue d'abord :

- * l'assiduité,
- * le respect du règlement intérieur.

Des points supplémentaires peuvent être attribués pour prendre en compte :

- * la participation de l'élève à la vie de l'établissement et aux activités organisées ou reconnues par l'établissement,
- * l'obtention de l'attestation scolaire de sécurité routière de premier ou de second niveau et l'obtention de l'attestation de formation aux premiers secours.

Modalités d'attribution

La note de vie scolaire est trimestrielle.

Le rôle de coordination du chef d'établissement

Le chef d'établissement recueille :

- * les propositions du professeur principal qui a préalablement consulté les membres de l'équipe pédagogique de la classe,
- * l'avis du conseiller principal d'éducation.

Il fixe ensuite la note qui est communiquée au conseil de classe.

Note de vie scolaire et diplôme national du brevet

Le diplôme national du brevet prend en compte la note de vie scolaire. Celle-ci est calculée sur la base des notes trimestrielles obtenues en classe de 3^e. Elle est affectée du coefficient 1.

3. La création de cette note, la légitimité de son objet, sa définition même ont été fortement contestées. Pour un exemple de débat, on pourra examiner les arguments de deux prises de position publiées dans le numéro 451 des *Cahiers pédagogiques* consacré au thème de *la sanction* (http://www.cahiers-pedagogiques.com/pedagogie_dossiers.php3).