

Rapport à madame la ministre de l'Éducation nationale

Pratiques d'évaluation en physique-chimie dans le cycle terminal de la voie générale

N° 23-24 313A - décembre 2024

*Inspection générale de l'éducation,
du sport et de la recherche*

Pratiques d'évaluation en physique-chimie dans le cycle terminal de la voie générale

Groupe physique-chimie de l'IGÉSR
Année scolaire 2023-2024

Jean Aristide CAVAILLÈS

Yannick ALMÉRAS

Cécile BRUYÈRE

Pierre DESBIOLLES

Antoine ÉLOI

Bruno JEAUFFROY

Delphine PAILLER

François VANDENBROUCK

*Inspecteurs généraux de l'éducation,
du sport et de la recherche*

SOMMAIRE

Synthèse	1
Introduction	5
1. Les enjeux de l'évaluation de contrôle continu en voie générale et technologique	5
1.1. Des enjeux multiples qui font parfois oublier la dimension formative de l'évaluation	5
1.2. Les enjeux de l'évaluation vus par les professeurs – pour eux-mêmes et pour leurs élèves	6
2. Des pratiques pédagogiques peu diversifiées et dominées par les plus « traditionnelles ».....	8
2.1. Des évaluations formelles classiques très présentes et pilotées par l'examen	8
2.2. Les évaluations destinées à réguler des apprentissages.....	11
2.3. Des pratiques d'évaluation qui touchent à l'autonomie, la prise d'initiative, l'oralité et au collectif, jugées moins faciles à mettre en œuvre et moins prioritaires.....	12
2.4. Un accompagnement des élèves avant, pendant et après l'évaluation qui doit être amélioré..	13
2.5. Les pratiques d'évaluation des professeurs vues par eux-mêmes	14
2.6. Comparaison avec d'autres enseignements de spécialité	16
3. Le pilotage des pratiques d'évaluation	17
3.1. Le projet d'évaluation, un outil national de cadrage du contrôle continu déployé de manière inégale en académie.....	17
3.2. Des projets d'évaluation qui laissent peu de part aux aspects pédagogiques	18
3.3. Un effet très limité sur les pratiques d'évaluation en physique-chimie	19
4. L'équité de l'évaluation et son pilotage.....	20
4.1. Une équité qui semble aller de soi pour les équipes rencontrées.....	20
4.2. Une analyse comparée des moyennes de contrôle continu attribuées par différents enseignants qui doit tenir compte du profil de leurs élèves.....	21
4.3. Une analyse de l'écart entre les notes de contrôle continu et les notes obtenues aux épreuves terminales du baccalauréat par les élèves de différents enseignants	23
4.4. Une analyse d'écart entre contrôle continu et contrôle terminal à partir de données académiques	25
4.5. Les observables de nature qualitative utiles au pilotage de l'équité.....	27
Conclusion	28
Annexes	31

SYNTHÈSE

Durant l'année scolaire 2023-2024, le groupe physique-chimie de l'IGÉSR a mené une étude consacrée à l'évaluation des élèves dans l'enseignement de spécialité physique-chimie du cycle terminal de la voie générale du lycée.

Objectifs

L'étude avait trois objectifs principaux :

- Observer les pratiques pédagogiques d'évaluation des enseignants ;
- Analyser les procédures qui sont, ou non, mises en place à différentes échelles (échelles académique, de l'établissement, de l'équipe pédagogique) pour s'assurer de l'égalité de traitement des élèves dans l'attribution des moyennes de contrôle continu, c'est-à-dire pour en contrôler l'équité ;
- Étudier sur des exemples concrets les éléments observables quantitatifs et qualitatifs sur lesquels le pilotage des pratiques d'évaluation peut s'appuyer pour en garantir à la fois l'équité et l'efficacité formative.

Méthodologie

Les membres du groupe se sont rendus dans 19 établissements où ils ont rencontré les équipes de direction, les professeurs de physique-chimie engagés dans l'enseignement de spécialité et des groupes d'élèves des classes de première et de terminale ayant choisi cet enseignement. Les chefs d'établissement ont été sollicités en amont pour fournir un ensemble de données quantitatives (rendues anonymes) sur l'évaluation des élèves dans leur établissement. Un sondage a par ailleurs été adressé aux professeurs de physique-chimie des établissements visités portant sur leurs pratiques et sur leur propre perception de l'évaluation.

Les visites ont été organisées avec l'aide des IA-IPR des académies concernées et ont été menées en présence de certains d'entre eux, sauf pour les échanges avec les professeurs où cette présence aurait pu biaiser les réponses des enseignants.

Des échanges formalisés consacrés au pilotage académique des pratiques d'évaluation des enseignants ont par ailleurs été organisés avec des équipes d'IA-IPR.

Les enjeux de l'évaluation de contrôle continu en voie générale et technologique vue par les professeurs

L'évaluation des élèves en contrôle continu remplit des fonctions variées et porte de multiples enjeux. Les professeurs interrogés par sondage perçoivent très bien son utilité pour contrôler les connaissances et le degré de compréhension des élèves. Ils placent cependant au second plan l'amélioration des apprentissages des élèves ainsi que la régulation que cette évaluation peut induire sur leur propre enseignement. Ils la conçoivent encore moins souvent comme un outil au service de la motivation des élèves.

L'enjeu certificatif des notes de contrôle continu en physique-chimie au baccalauréat est limité¹. L'importance qui leur est accordée vient surtout de ce que ces notes figurent dans le dossier de candidature des élèves dans la procédure Parcoursup. Le groupe a pu observer que cet enjeu d'orientation a tendance à occulter tous les autres dans l'esprit des élèves, de leurs parents et même des professeurs.

Des pratiques pédagogiques peu diversifiées et dominées par les plus « traditionnelles »

Des pratiques observées très uniformes

Dans l'ensemble, les pratiques pédagogiques d'évaluation des enseignants de physique-chimie rencontrés restent très traditionnelles. Les situations d'évaluation sont peu variées et se limitent souvent à des devoirs surveillés conçus sur le modèle des épreuves terminales du baccalauréat et dont les notes constituent l'essentiel des moyennes. Sauf exceptions, la notation de ces devoirs s'effectue à partir d'un barème chiffré

¹ Pour un enseignement de spécialité, seules les notes de contrôle continu de la classe de première des élèves ne poursuivant pas l'enseignement de physique-chimie interviennent dans la note finale de l'élève au baccalauréat, à hauteur de 8 % du total. Les élèves qui suivent l'enseignement de physique-chimie en terminale ne sont évalués au baccalauréat que par les notes des épreuves terminales.

et pratiquement jamais par l'analyse des degrés de maîtrise de compétences identifiées, même pour l'évaluation des questions à prise d'initiative, qui restent rares elles aussi.

Les activités expérimentales, lorsqu'elles sont évaluées, le sont plus fréquemment « par compétences », comme aux épreuves d'évaluations des capacités expérimentales (ECE) du baccalauréat. C'est cependant presque toujours le compte rendu écrit de l'activité qui est évalué : les explications orales ou la maîtrise des gestes expérimentaux ne le sont que très rarement. De plus, le poids des activités expérimentales dans les moyennes reste le plus souvent inférieur aux 20 % recommandés par l'IGÉSR dans le guide de l'évaluation des apprentissages².

Les situations d'évaluations réalisées à l'oral sont, elles aussi, exceptionnelles. La préparation du Grand oral, qui pourrait en fournir l'occasion, n'apparaît pas comme une incitation suffisante.

Les tentatives de mise en place de modalités moins classiques comme l'évaluation d'activités de groupe, l'évaluation par les pairs ou l'autoévaluation restent très isolées, malgré leur intérêt pédagogique.

Un accompagnement de l'évaluation insuffisant

Le groupe regrette que l'accompagnement des élèves par les professeurs, avant, pendant et après les évaluations soit généralement limité. Ainsi, les objectifs et attendus d'une évaluation ne sont pratiquement jamais explicités en amont aux élèves, pas plus que les critères de réussite. Pendant l'évaluation, les pratiques de différenciation, ou de « coups de pouces », ne sont pas mises en place malgré leur intérêt formatif. Après les évaluations, les retours faits aux élèves se limitent à une note, le plus souvent assortie de commentaires succincts. Ils sont rarement accompagnés de conseils opérationnels ou de pratiques de remédiation.

Une acculturation à renforcer

La faible diversité des pratiques d'évaluation observée et les échanges avec les enseignants font apparaître des insuffisances dans la culture des professeurs en ce qui concerne les modalités d'évaluation. Un effort de formation important est nécessaire pour leur présenter des techniques d'évaluation plus diverses et pour les aider à mieux intégrer l'activité d'évaluation dans leur pédagogie, par un accompagnement plus actif et mieux construit des élèves.

Il importe en particulier que les enjeux et le poids de chaque évaluation soient mieux explicités aux élèves, de façon à éviter qu'elles soient toutes identiquement perçues comme un élément décisif des chances de succès dans la procédure Parcoursup. Ces explicitations peuvent aider les élèves à vivre leurs évaluations de façon plus sereine et moins stressante.

Le pilotage des pratiques d'évaluation

Un engagement variable des chefs d'établissement

Le pilotage de l'évaluation des élèves par les chefs d'établissement est très variable. Si certains d'entre eux font de cette question un thème de réflexions et d'actions suivies (réunions annuelles, analyse des résultats des élèves dans les différentes disciplines, bilans des résultats obtenus au baccalauréat, etc.), ils sont beaucoup plus nombreux, parmi ceux que le groupe a rencontrés, à déléguer complètement ce pilotage aux équipes disciplinaires.

Des projets d'évaluation sans réel impact sur les pratiques pédagogiques

Les projets d'évaluation que le groupe a pu étudier³ sont à l'image de la diversité des pilotages en établissement. Dans leur majorité, ces textes abordent essentiellement les questions qui relèvent de la vie scolaire (gestion des absences et des stratégies d'évitement, nombre d'évaluations minimum devant figurer dans les moyennes, etc.) et proposent, sur ce plan, un cadrage apprécié des enseignants. Il est beaucoup moins fréquent que les projets d'établissement abordent des questions de nature pédagogique, spécifiques à chaque discipline.

² Ministère de l'éducation nationale et de la jeunesse, novembre 2023, Guide de l'évaluation des apprentissages et des acquis des élèves au lycée général et technologique, <https://eduscol.education.fr/document/5470/download?attachment>.

³ Tous les établissements visités, sauf un, ont été en mesure de présenter un projet d'évaluation.

Les enseignants déclarent que l'élaboration et la mise en place des projets d'évaluation n'ont eu que très peu d'impact sur leurs pratiques : les annexes disciplinaires reprennent le plus souvent ce que les enseignants faisaient avant l'écriture de ces projets. Il en va de même des documents mis en ligne sur Éduscol, en particulier du guide de l'évaluation des apprentissages rédigé par l'IGÉSR2, qui, lorsqu'il n'est pas complètement ignoré, est parfois cité, même si ses recommandations sont peu suivies.

Le rôle essentiel des IA-IPR dans le pilotage de l'évaluation des élèves

Le groupe estime que le pilotage des pratiques d'évaluation en physique-chimie par les IA-IPR est essentiel, à la fois par les formations qui peuvent être proposées dans les plans de formations académiques, par les ressources qui sont souvent mises en ligne par les IA-IPR dans les sites académiques, et surtout par les conseils individuels ou collectifs qu'ils fournissent aux enseignants. Les réunions d'équipe sont un moment bien adapté au développement de la culture des équipes sur la question de l'évaluation des élèves en même temps qu'elles peuvent être l'occasion d'une analyse collective des pratiques s'appuyant sur des éléments quantitatifs et qualitatifs. Elles peuvent également s'appuyer sur les éléments statistiques disponibles à l'échelle académique ou de l'établissement, même si le groupe estime que ces éléments doivent être maniés avec précaution (cf. infra). Elles permettent enfin de sensibiliser les professeurs aux différents enjeux de l'évaluation, particulièrement à l'enjeu formatif.

L'équité de l'évaluation

L'équité au sein de l'établissement

Pour la plupart des équipes d'enseignants rencontrées, l'équité de l'évaluation entre les élèves semble aller de soi. Les équipes sont conscientes de l'enjeu d'équité mais elles estiment invariablement que leurs pratiques d'évaluation sont suffisamment harmonisées pour être équitables. Elles sont souvent surprises lorsqu'on les interroge sur les observations qui fondent ce sentiment. La comparaison raisonnée des moyennes de chacun, la construction de celles-ci et les poids apportés à chaque type d'évaluation, de même que les critères d'évaluation, font très rarement l'objet d'échanges organisés entre les enseignants. Les réunions que le groupe a tenues avec des professeurs pour ce travail ont révélé qu'ils ont une idée en général très floue de la façon dont leurs collègues élaborent leurs moyennes, des modalités d'évaluation qu'ils mettent en œuvre, des observables de réussite, etc. Pourtant, le plus souvent, les équipes de physique-chimie travaillent plutôt bien ensemble : elles harmonisent leurs progressions, organisent des devoirs communs, partagent des textes de travaux pratiques, etc. Mais cette collégialité, dont on ne peut que se féliciter, ne se retrouve pas dans les pratiques d'évaluation qui restent le pré carré de chacun.

Dans la plupart des cas, donc, il apparaît que les équipes ont besoin du pilotage du chef d'établissement et des IA-IPR pour aborder de façon structurée la question de l'équité des évaluations portées par les différents professeurs. Ce pilotage peut s'appuyer sur des éléments de deux natures : les éléments quantitatifs d'une part (moyennes, résultats obtenus au baccalauréat, nombres d'évaluations par période, coefficients accordés à chaque modalité d'évaluation, etc.) et les éléments qualitatifs, d'autre part, qui portent sur la façon dont les équipes concrétisent leurs intentions d'équité dans leurs modalités d'évaluation. Ces deux catégories d'indicateurs ont été étudiées.

L'harmonisation des niveaux de notation des professeurs d'une même équipe pédagogique

Le groupe a analysé le cas concret d'un des établissements visités dans lequel trois professeurs enseignent la spécialité physique-chimie en classe de terminale : il a cherché à comparer les résultats d'évaluation des élèves de ces enseignants afin de détecter d'éventuelles situations anormales, à partir de l'étude des moyennes des élèves de ces enseignants. Les analyses du groupe, détaillées dans le rapport, montrent que la comparaison simple des moyennes est en elle-même insuffisante. Il est indispensable de prendre en compte les caractéristiques des élèves de chaque groupe pour que cette comparaison ait un sens. Ainsi, comme les données récoltées le confirment, la moyenne obtenue en physique-chimie en classe de terminale est très fortement corrélée au choix par les élèves de la deuxième spécialité et des options mathématiques⁴.

⁴ Par exemple, les élèves ayant choisi la spécialité mathématique avec l'option mathématiques expertes ont une moyenne en physique-chimie supérieure de plus de 5 points à celle des élèves ayant choisi pour seconde spécialité sciences de la vie et de la Terre (SVT), sans l'option mathématiques complémentaires.

La composition des groupes d'élèves de chaque enseignant doit donc impérativement être prise en compte si l'on veut comparer le niveau de notation des enseignants entre eux.

Le rapport montre que les analyses à mener pour identifier les pratiques d'évaluation hors normes sont relativement complexes, et il ne semble pas raisonnable d'imaginer qu'une équipe de direction d'un établissement puisse les mener de façon systématique. Le groupe pense cependant qu'il serait utile de procurer aux chefs d'établissement un ensemble de données nationales sur les biais statistiques qu'ils pourraient prendre en compte dans l'analyse d'une situation qui leur semble anormale.

Les comparaisons des écarts entre les moyennes annuelles de terminale et les notes obtenues aux épreuves terminales de spécialité sont sans doute plus faciles à interpréter dans la mesure où les différents facteurs statistiques ont des effets comparables sur les notes de contrôle continu et sur des épreuves terminales, comme cela a pu être vérifié. Dans ce cas également, une analyse experte et mesurée des écarts est indispensable. On cherche en effet à comparer des évaluations à vocation purement certificatives (les épreuves terminales) avec des notes qui portent une dimension formative essentielle et qui ont une influence sur la motivation des élèves. Ainsi, il n'est pas étonnant qu'un professeur ayant une classe dont les résultats sont faibles soit tenté de surnoter légèrement les élèves les plus faibles pour ne pas les démotiver. Cette tendance ressort des données analysées par le groupe.

La comparaison raisonnée des niveaux de notation de deux professeurs est difficile à partir des simples données de moyennes de contrôle continu ou de notes aux épreuves terminales. Des écarts très marqués d'un professeur par rapport aux autres, qui se répètent sur plusieurs années, peuvent être des indicateurs d'une anomalie dont l'instruction doit s'appuyer sur une bonne connaissance des principaux biais statistiques.

La comparaison des établissements entre eux

Les IA-IPR et les chefs d'établissement disposent d'un ensemble très complet d'indicateurs globalisés à l'échelle de l'établissement qui permettent de situer un établissement relativement aux autres. Certains de ces indicateurs ont été analysés par le groupe et leur interprétation connaît les mêmes difficultés que les données internes à l'établissement : il est indispensable de prendre en compte les spécificités de l'établissement pour réaliser des comparaisons pertinentes. Le facteur statistique le plus important semble être l'indice de position sociale (IPS), même si les effets de composition des groupes de spécialité (non étudiés par le groupe) pourraient jouer un rôle à l'échelle des établissements.

Il convient cependant d'envisager avec prudence les effets des indices globalisés, qui doivent être distingués des indices personnels. Ainsi l'IPS moyen d'un établissement est-il un indicateur insuffisamment précis ; il faudrait disposer, au minimum, d'un indice IPS moyen des élèves ayant suivi la spécialité physique-chimie pour pouvoir faire des comparaisons fiables.

Le pilotage pédagogique et les observations qualitatives restent l'approche à privilégier

Devant la difficulté d'analyse des données chiffrées, le groupe en est venu à la conclusion qu'il importe avant tout, en établissement, de s'assurer qu'une stratégie collective du contrôle de l'équité est mise en place par et dans les équipes. Plusieurs pratiques peuvent concourir à cet effort : organisation régulière, sur un rythme annuel, de réunions d'échanges formalisés entre les professeurs, attention portée à l'harmonisation de la construction des moyennes, réflexion commune sur les attendus et les critères de réussite, organisation de devoirs communs, avec corrections croisées, etc. Cette collégialité des approches est, pour le groupe, le meilleur moyen de s'assurer que les élèves sont évalués avec l'équité qui leur est due.

Introduction

Face aux nouveaux enjeux associés à l'évaluation de contrôle continu des élèves, particulièrement dans le cycle terminal des lycées généraux et technologiques, le groupe physique-chimie de l'inspection générale de l'éducation, du sport et de la recherche (IGÉSR) a souhaité avoir une vision précise des conditions dans lesquelles cette évaluation est pratiquée, accompagnée, pilotée et ressentie, dans la discipline physique-chimie, par l'ensemble des acteurs de terrain, qu'il s'agisse des professeurs, des inspecteurs territoriaux, des équipes de direction ou des élèves.

Afin de limiter son champ d'investigation d'un sujet aussi large, le groupe a choisi de se concentrer sur l'évaluation des élèves des classes de première et de terminale de la voie générale dans le cadre des enseignements de la spécialité physique-chimie. Avec l'aide des inspecteurs pédagogiques régionaux (IA-IPR) de 12 académies⁵, les inspecteurs généraux du groupe se sont rendus dans 19 établissements⁶ placés dans des contextes territoriaux et sociaux différents. Au cours de ces visites, et suivant un protocole précis⁷, les membres du groupe ont rencontré les professeurs de physique-chimie, l'équipe de direction et des panels d'élèves des classes de première et de terminale ayant choisi l'enseignement de spécialité physique-chimie. Les IA-IPR ayant organisé les visites ont assisté à ces échanges, sauf lors des rencontres avec les professeurs de façon à ne pas introduire de biais.

Les visites étaient précédées par l'envoi aux professeurs rencontrés d'un questionnaire général sur les pratiques et les objectifs donnés à l'évaluation⁸, transmis par les chefs d'établissement. Ceux-ci étaient également sollicités pour fournir au groupe des données statistiques anonymisées tant du point de vue des élèves que des enseignants. Ces données ne sont disponibles dans aucune des banques statistiques auxquelles le groupe a communément accès (base centrale de pilotage, données publiques Parcoursup, etc.).

Le présent rapport rappelle tout d'abord les nouveaux enjeux de l'évaluation de contrôle continu en physique-chimie qui résultent de la réforme de l'organisation des enseignements au lycée et du baccalauréat. Il propose ensuite une description des pratiques pédagogiques d'évaluation telles qu'elles ont été observées et propose des améliorations qu'il serait possible de leur apporter. Les conditions d'un pilotage efficace de ces pratiques, aux niveaux de l'établissement et de l'académie, font l'objet de la partie suivante. Le rapport se termine par une analyse des facteurs favorables à l'équité de l'évaluation de contrôle continu et au contrôle de cette équité. Des annexes complètent le corps du rapport, essentiellement sur des aspects quantitatifs ou de méthodologie.

Les analyses qualitatives et quantitatives que le groupe a réalisées dessinent un paysage complexe et contrasté des pratiques de l'évaluation en physique-chimie. Le groupe estime que ces analyses peuvent être utiles aux corps d'inspection territoriaux et aux chefs d'établissement pour le pilotage de l'évaluation des élèves et de son équité dans leurs académies et au sein de leur établissement. Elles permettent également de relativiser la lisibilité des notes brutes de contrôle continu sur lesquelles se fondent en partie les établissements de l'enseignement supérieur, via Parcoursup, pour classer les élèves. Le rapport propose des pistes pour s'assurer au mieux de la qualité et de l'équité de ces évaluations.

1. Les enjeux de l'évaluation de contrôle continu en voie générale et technologique

1.1. Des enjeux multiples qui font parfois oublier la dimension formative de l'évaluation

Le double enjeu du baccalauréat et du dossier Parcoursup

La réforme du lycée général et technologique de 2018 a placé la construction du parcours d'orientation de l'élève et sa réussite dans le supérieur au cœur de ses objectifs. En plus de donner aux élèves le bagage de

⁵ Aix-Marseille, Bordeaux, Corse, Limoges, Montpellier, Nancy-Metz, Nice, Reims, Rennes, Strasbourg, Toulouse, Versailles.

⁶ La liste complète est fournie dans l'annexe 1.

⁷ Le protocole est détaillé dans l'annexe 2.

⁸ Le questionnaire figure dans l'annexe 3.

connaissances et de compétences dont ils ont besoin pour leurs poursuites d'études, les années de lycée doivent les aider à construire leur projet d'orientation avec des choix raisonnés d'enseignements de spécialité et une constitution cohérente de leur dossier Parcoursup.

Tout en conservant son rôle de validation des acquis du lycée, l'obtention du baccalauréat reste une étape incontournable au sein de ce processus. Le poids des notes de contrôle continu (CC) dans le calcul final de la note du baccalauréat, porté à 40 % du total des points, renforce l'importance des notes obtenues en classe tout au long du cycle terminal en leur conférant aussi une valeur certificative. De plus, avec le passage des épreuves terminales au mois de juin, les notes de contrôle continu obtenues dans la première partie de l'année de terminale sont les éléments quantitatifs principaux appelés à figurer dans le dossier Parcoursup des élèves. Dans le quotidien de la classe, chaque note est donc porteuse d'un double enjeu : celui de l'obtention du baccalauréat et celui de la poursuite du parcours d'orientation par l'incidence qu'elle a sur le contenu du dossier examiné par les établissements de l'enseignement supérieur. Ce double enjeu est bien perçu par les professeurs ainsi que par les élèves et leurs familles ; il génère parfois, particulièrement en terminale, une forte inquiétude à propos des notes de contrôle continu, essentiellement à cause de leur visibilité dans la procédure Parcoursup.

Un enjeu très limité dans les choix de spécialité

Le groupe a constaté que dans la construction même du parcours d'orientation des élèves, les notes de contrôle continu ne portent qu'un enjeu limité et n'influencent que très peu les choix d'enseignements de spécialité à la fin des classes de seconde ou de première. Ainsi, les élèves rencontrés déclarent massivement avoir choisi l'enseignement de physique-chimie par goût ou parce qu'ils savent que les connaissances de physique-chimie leur seront utiles dans leur parcours futur. Les notes interviennent très peu dans ce choix.

Ce sont les élèves qui n'ont pas de projet d'orientation abouti qui déclarent prendre aussi en compte les notes pour leurs choix d'enseignement de spécialité à la fin de la classe de première, mais c'est alors essentiellement pour des raisons stratégiques de réussite au baccalauréat. Selon les cas, ils ne poursuivent pas en terminale l'enseignement de spécialité dans lequel ils ont les notes les plus faibles, pour ne pas se pénaliser au moment des épreuves terminales, ou ils abandonnent au contraire celui pour lequel ils ont de bonnes notes, mais peu d'appétence, afin de maximiser la note de contrôle continu de leur enseignement de spécialité non poursuivi pour le bac⁹.

Un enjeu formatif qui passe au second plan

Dans les échanges qu'il a menés, le groupe a constaté que l'enjeu formatif de l'évaluation de contrôle continu est peu perçu par les professeurs ou les élèves et passe au second plan. La prise de conscience de cet enjeu par les professeurs serait de nature à les inciter à varier leurs pratiques d'évaluation (voir infra) en considérant l'erreur comme un élément essentiel du processus d'apprentissage. D'une façon générale, les enseignants rencontrés qui explicitent clairement les objectifs des évaluations sont rares. Cette explicitation pourrait contribuer à dédramatiser les évaluations dont le but est purement formatif. Elle pourrait également provoquer une mobilisation positive pour certaines évaluations, de nombre limité, dont le résultat intervient avec un poids important dans la moyenne de l'élève.

1.2. Les enjeux de l'évaluation vus par les professeurs – pour eux-mêmes et pour leurs élèves

Le groupe a réalisé un sondage auprès de l'ensemble des professeurs de physique-chimie des établissements qu'il a visités. Au total, 109 réponses ont été obtenues. Deux questions concernaient différents critères d'utilité de l'évaluation des élèves, d'une part du point de vue des professeurs et d'autre part du point de vue des élèves tel que se le représentent les enseignants. Les élèves n'ont pas été questionnés directement.

Les résultats sont reportés sur les figures n° 1 et n° 2.

Figure n° 1 : réponses des professeurs à la question : « À quoi [cela] sert-il pour un professeur d'évaluer les acquis de ses élèves ? »

⁹ Voir note 1.

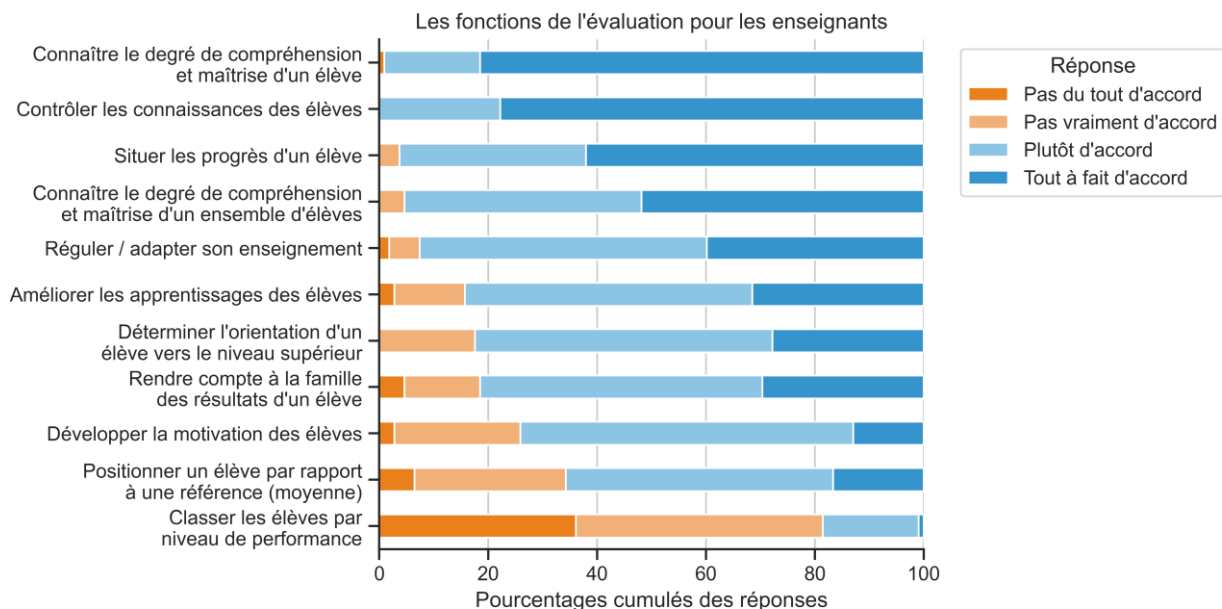
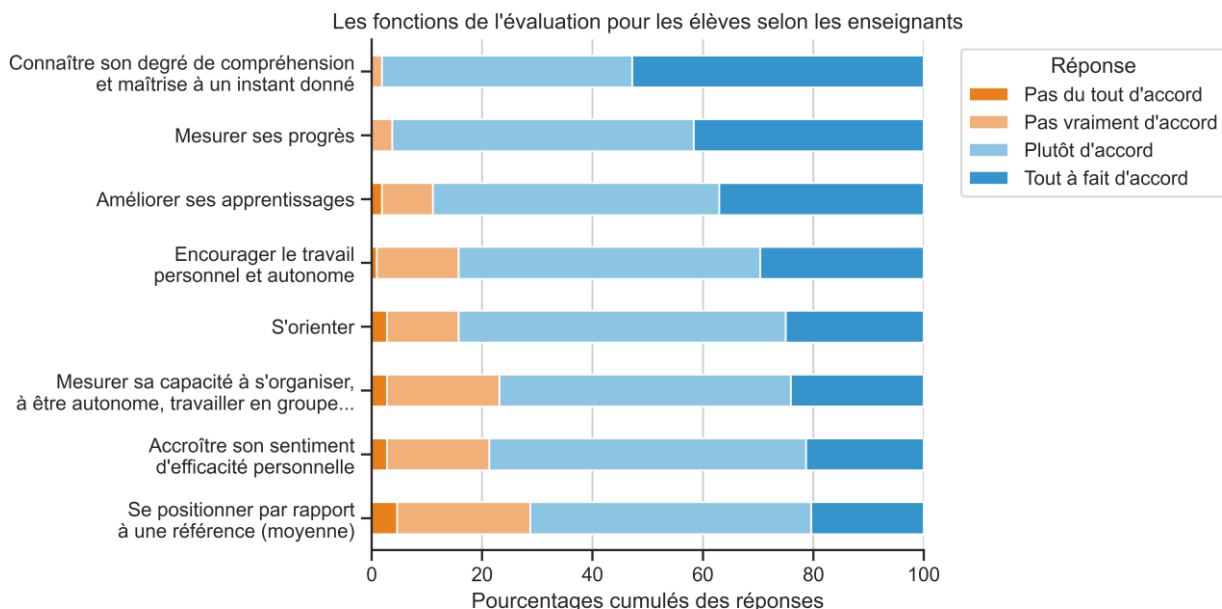


Figure n° 2 : réponses des professeurs à la question : « À quoi sert l'évaluation pour un élève ? »



L'utilité diagnostique, pour le professeur comme pour l'élève (selon les professeurs), est clairement perçue et mise en avant par les réponses « Connaître le degré de compréhension et maîtrise d'un élève », « Situer les progrès d'un élève » et « Contrôler les connaissances des élèves » avec lesquelles la très grande majorité des professeurs est plutôt ou tout à fait d'accord.

Il est étonnant de constater qu'environ 18 % des enseignants ne sont pas vraiment ou pas du tout d'accord pour considérer que l'évaluation est, pour eux, un moyen d'améliorer les apprentissages des élèves, alors que l'on pourrait s'attendre à un consensus beaucoup plus large sur cette finalité.

Le rôle que peut jouer l'évaluation dans la motivation des élèves est encore moins fréquemment perçu par les professeurs, qui sont 25 % à n'être pas vraiment ou pas du tout d'accord avec cette finalité.

Ainsi, l'utilité *a posteriori* de l'évaluation, c'est-à-dire des usages pédagogiques que l'on en fait une fois qu'elle a eu lieu, est-elle loin d'être unanimement reconnue. Les réponses au questionnaire vont dans le sens des échanges que le groupe a eus avec les professeurs et les élèves, dans lesquels il est fréquemment apparu que les retours faits aux élèves sur les évaluations sont peu développés et faiblement opérationnels (cf. infra).

2. Des pratiques pédagogiques peu diversifiées et dominées par les plus « traditionnelles »

Toutes les équipes pédagogiques de physique-chimie rencontrées par le groupe ont bien pris la mesure des enjeux de certification et d'orientation de l'évaluation des élèves ; elles les évaluent avec sérieux et sont en mesure de leur attribuer une note moyenne de contrôle continu qui est raisonnablement représentative de leur niveau en physique-chimie. Dans les échanges, il apparaît que les élèves et leurs parents, en grande majorité, font confiance aux enseignants sur ce point. Les situations de contestation de notes, si elles existent, restent rares.

Pour autant, les observations du groupe montrent que les professeurs doivent renforcer la dimension formative des évaluations des élèves. En effet, dans le cycle terminal de la voie générale, les pratiques pédagogiques de l'évaluation en enseignement de spécialité physique-chimie s'avèrent très peu diversifiées et dominées par les plus traditionnelles d'entre elles ; elles sont peu développées dans le sens de l'accompagnement des apprentissages et de la valorisation des progrès. Un accompagnement est nécessaire pour les faire évoluer dans le sens de l'enrichissement de ces pratiques et du développement d'une culture autour de l'évaluation comme cela est recommandé par la récente conférence de consensus du Centre national d'étude des systèmes scolaires (CNESCO)¹⁰.

2.1. Des évaluations formelles¹¹ classiques très présentes et pilotées par l'examen

Les contrôles ou devoirs surveillés (DS)

Le groupe observe que la plupart des enseignants pratiquent une évaluation chiffrée qui repose majoritairement sur des devoirs surveillés (DS) de fin de chapitre. Ces devoirs se déroulent sur des durées variables, pendant les créneaux des séances de physique-chimie. Ils durent donc généralement moins de deux heures, hors cas particuliers de temps banalisés dans l'emploi du temps qui autorisent des devoirs communs ou de type « bac blanc », typiquement en terminale.

Dans le calcul de la moyenne d'un élève, les notes de DS sont largement prépondérantes par rapport à celles des autres modalités d'évaluation. Leur structure et leur contenu sont souvent inspirés des sujets des épreuves terminales du baccalauréat, surtout en classe de terminale (voir encadré n° 1).

La notation s'appuie sur un barème détaillé selon les éléments attendus question par question. À une ou deux exceptions notables près, le groupe n'a pas vu d'évaluation écrite par compétences¹², même pour des exercices de type résolutions de problème ou questions ouvertes qui sont d'ailleurs très peu proposés.

¹⁰ CNESCO (2023). Conférence de consensus « L'évaluation en classe, au service de l'apprentissage des élèves » : <https://www.cnesco.fr/evaluation-en-classe>

¹¹ Comme développé par la conférence de consensus du CNESCO précitée, il est possible de distinguer l'évaluation formelle – « évaluation qui dit son nom » (exemple : devoir surveillé) – de l'évaluation informelle – « évaluation qui ne dit pas son nom » (exemple : avis donné à un élève lors d'une circulation en classe). Les enseignants développent souvent des pratiques d'évaluation, de type informel, sans les identifier comme telles.

¹² Évaluation qui attribue aux élèves des niveaux de maîtrises de compétences ou de capacités. Cette évaluation peut être résumée par une note, en fonction de barèmes de conversion.

Encadré n° 1 : la typologie des devoirs surveillés

En classe de première, les DS présentent des contenus moins influencés par les sujets d'épreuves terminales. Certains enseignants ont recours à un assemblage de courts exercices issus de manuels scolaires, contextualisés de façon concise ; d'autres font le choix d'une approche mixte combinant des exercices courts avec un problème plus long – conformément au format de l'épreuve écrite du baccalauréat, avec des documents et une contextualisation plus élaborée. Certains devoirs débutent par un quiz sur le cours. Globalement, les DS sont centrés sur les connaissances de base du cours selon un format choisi par l'enseignant, qui ne change pas, ou relativement peu, au fil de l'année ; la démarche scientifique n'est pas particulièrement mise en exergue. La Banque Nationale de Sujets (BNS) de l'enseignement de spécialité physique-chimie est parfois mobilisée, surtout en classe de première.

En classe de terminale, les devoirs surveillés se rapprochent de la présentation et des exigences de l'épreuve écrite du baccalauréat. Ils s'appuient sur des extraits de sujets de sessions passées ou de manuels scolaires. On trouve donc surtout des questionnements progressifs enchaînant des tâches simples, relativement proches du cours et limitant les points de blocage ; les tâches les plus complexes sont placées en fin d'exercice, certaines étant considérées comme une occasion de gain de « points bonus ». La place des documents et de la contextualisation dans les DS est plus importante qu'en classe de première.

Les enseignants précisent le plus souvent, de façon globale pour chaque exercice, le nombre de points prévus au barème ; quelques-uns vont jusqu'à porter une mention des points, voire des compétences de rattachement, pour chaque question. À de rares exceptions près, il n'y a pas d'éléments de différenciation pédagogique intégrés ou associés aux énoncés.

Les énoncés de DS consultés n'investissent généralement pas les domaines des incertitudes de mesure et n'évaluent pas les capacités numériques liées au langage Python. Ces points sont davantage considérés dans le cadre des activités expérimentales.

Le groupe recommande que la palette des DS proposés par chaque professeur présente une variété suffisante de tâches, avec une progressivité pensée non seulement par séquence mais aussi sur l'année. Cette modalité d'évaluation mérite d'être distinguée de celle des interrogations spécifiques de cours qui ne se placent pas au même moment dans une séquence (cf. infra).

L'évaluation des activités expérimentales

Les enseignants et les élèves rencontrés considèrent que les activités expérimentales réalisées en travaux pratiques (TP) sont un élément essentiel de la formation en physique-chimie. Une majorité de professeurs évaluent ces activités, même si cela est loin d'être systématique. Le poids de ces évaluations dans la moyenne des élèves est faible ; il est le plus souvent inférieur au poids de 20 % préconisé par l'IGÉSR dans le *Guide de l'évaluation*¹³.

Les activités expérimentales sont moins souvent évaluées en classe de première qu'en classe de terminale, sans doute parce que l'horaire disponible y est plus limité.

En classe de terminale, cette évaluation est souvent conçue de façon analogue à celle de l'épreuve d'évaluation des compétences expérimentales (ECE) du baccalauréat. C'est parfois également le cas en classe de première (voir encadré n° 2). L'évaluation par compétences pratiquée dans les ECE est fréquemment prise comme modèle dans les évaluations des activités expérimentales dans le cadre du contrôle continu. Les ECE contribuent donc à diffuser dans les classes les pratiques d'évaluation par compétences, même si celles-ci restent le plus souvent réservées aux seules activités expérimentales.

Le groupe observe cependant que ces évaluations s'appuient presque toujours sur les comptes rendus écrits et plus rarement sur l'observation du geste expérimental par le professeur en cours de TP, contrairement à ce qui se pratique dans les ECE.

¹³ Ministère de l'Éducation Nationale et de la Jeunesse, novembre 2023, Guide de l'évaluation des apprentissages et des acquis des élèves au lycée général et technologique, *op. cit.*

Lorsque l'évaluation par compétences est mise en place, la compétence « Valider » est beaucoup plus rarement évaluée que les autres compétences de la démarche expérimentale¹⁴. Cette compétence donne pourtant tout leur sens aux activités expérimentales et les relie explicitement à un élément majeur de la démarche scientifique. Le travail et l'évaluation de cette compétence doivent être développés.

Très peu d'ECE blanches sont organisées dans les lycées. Néanmoins, lorsqu'on les interroge, les élèves s'estiment bien préparés aux ECE par leurs enseignants.

Encadré n° 2 : la typologie des travaux pratiques

Dans les classes de première et en terminale, les énoncés des activités expérimentales présentent des contenus conformes aux attentes des programmes et à l'esprit de l'enseignement de la physique-chimie. Les sujets de TP se rapprochent de la présentation et des exigences de l'épreuve d'ECE du baccalauréat et sont conçus en s'appuyant sur les étapes de la démarche scientifique.

On trouve ainsi fréquemment dans les sujets une présentation du contexte du TP, s'appuyant sur des documents qui donnent aux élèves les informations nécessaires à la résolution et à la réalisation du TP. Des questions progressives permettent d'aboutir à une conclusion en lien avec le contexte énoncé en début de TP.

Certains énoncés de TP consultés par le groupe investissent le domaine des incertitudes de mesure et mettent parfois en application la notion d'écart normalisé, ce qui permet aux élèves de travailler la compétence « Valider » avec rigueur. Des enseignants intègrent dans leurs TP le développement des capacités numériques en mobilisant le langage Python, l'utilisation de microcontrôleurs ou celle de logiciels d'acquisition et de traitements numériques variés comme Latis Pro.

Les élèves sont évalués sur leur capacité à suivre ou, plus rarement, à proposer, des protocoles. Les énoncés contiennent parfois des appels au professeur répartis pendant toute la durée du TP durant lesquels les élèves s'entretiennent avec le professeur. Ces échanges peuvent être considérés comme un élément de travail de la compétence « Communiquer » et des capacités d'expression orale. Ils sont également favorables à la différenciation pédagogique. Les professeurs évaluent rarement la compétence « Valider ». Ils estiment parfois que la compétence « Réaliser » est difficile à évaluer quand les élèves sont en binôme.

Les élèves rencontrés apprécient généralement l'évaluation des activités expérimentales par compétences, lorsqu'elle est mise en place. Ils la trouvent plus explicite que d'autres modes d'évaluation plus traditionnels : « *On sait quoi travailler* », dit l'un d'entre eux.

Le groupe recommande que les enseignants réalisent plus fréquemment des évaluations des activités expérimentales et qu'ils les incluent dans le calcul de la moyenne globale de l'enseignement à hauteur de 20 % de la note (de même que les ECE constituent 20 % de la note de l'épreuve finale de spécialité physique-chimie).

De courtes interrogations spécifiques de cours

Certains professeurs proposent à leurs élèves des évaluations courtes visant à vérifier l'assimilation de connaissances du cours ou, moins fréquemment, à développer des automatismes dans des savoir-faire importants ou « classiques ». Sans être exceptionnelles, ces évaluations sont plus rares que les devoirs surveillés et les enseignants leur affectent une pondération plus faible que celle des DS. Les modalités de ce type d'évaluations sont assez diverses (voir encadré n° 3).

¹⁴ Les programmes de physique-chimie identifient six compétences travaillées dans le cadre de la démarche expérimentale : s'approprier, analyser / raisonner, réaliser, valider et communiquer.

Encadré n° 3 : la typologie des interrogations spécifiques de cours

Deux approches sont essentiellement suivies pour ces interrogations spécifiques de cours : celle reposant sur des quiz – dont les questionnaires à choix multiples (QCM) – et celle s'appuyant sur de courts exercices peu ou pas du tout contextualisés. Les énoncés se présentent souvent sous la forme de documents lacunaires à compléter individuellement en classe, en un temps plus court que les DS. Quelques enseignants mettent à profit un outil numérique spécifique pour la pratique des QCM. Certains intègrent de telles interrogations dans les textes de DS.

De telles interrogations spécifiques de cours sont, pour le groupe, très favorables aux apprentissages des élèves. L'efficacité formative est moindre si l'interrogation est incluse dans un devoir surveillé qui sollicite l'élève sur un éventail plus large de notions, de façon moins cadrée dans le temps.

Le groupe recommande donc que les enseignants aient un recours plus systématique à ces interrogations, mises en œuvre distinctement des DS, faiblement contextualisées et allégées en texte pour le travail des fondamentaux et des automatismes associés. De telles évaluations ne sont pas nécessairement notées et doivent impérativement s'accompagner d'un retour rapide de l'enseignant aux élèves. Celui-ci est indispensable à la fois pour les élèves et pour les professeurs qui peuvent y trouver des indications précieuses sur l'effectivité de leurs apprentissages.

2.2. Les évaluations destinées à réguler des apprentissages

L'évaluation est souvent qualifiée de diverses manières selon ses fonctions spécifiques avec les termes « diagnostique », « formative », « sommative » ou encore « certificative ». Toutefois, des recherches récentes suggèrent que cette catégorisation est trop rigide et réductrice si l'on souhaite développer des pratiques pédagogiques intégrant l'évaluation de façon efficace en tenant mieux compte de sa finalité éducative^{15 16}.

Les chercheurs invitent à une évaluation qui accompagne davantage les apprentissages par un accroissement de sa fonction formative, un retour vers les élèves toujours présent avec une adaptation de l'enseignement (différenciation) – plutôt que de simples actions correctives *a posteriori* (remédiations) – et un appui sur une participation active des élèves dans l'évaluation. Celle-ci peut alors pleinement jouer un rôle de régulation en ce qu'elle permet aux professeurs d'adapter leurs enseignements et aux élèves d'agir pour combler les lacunes constatées dans leurs apprentissages. L'intérêt de cette « corégulation », qui implique élèves et enseignants, est souligné par plusieurs travaux¹⁷. Des modalités d'évaluation diverses peuvent y contribuer.

Les évaluations réalisées pendant les cours

Les questions non notées en cours – par exemple posées à l'oral, sans support écrit, à la classe, ou proposées sous forme de QCM – constituent des situations d'évaluation informelles qui sont mises en œuvre très fréquemment. Elles sont à distinguer des interrogations spécifiques de cours analysées plus haut. Leur rôle dans la régulation des apprentissages du fait des retours faits aux élèves est peu perçu. Cette pratique implique pourtant positivement les élèves. Il est important que chaque enseignant ait conscience qu'il s'agit d'une situation d'évaluation qui, même si elle n'est pas notée, s'intègre au processus de régulation des apprentissages des élèves et qu'il doit en tirer des conclusions, y compris sur son propre enseignement.

Les évaluations diagnostiques sont en revanche rares et généralement effectuées par des QCM quand elles existent. Les enseignants rencontrés disent « *démarrer à zéro* » en première et en terminale et ne vérifient pas la solidité des acquis antérieurs par des évaluations spécifiques. Le manque de temps ou la simple

¹⁵ Cf. la conférence de consensus du CNESCO citée supra et, en particulier, la note de Lucie Mottier Lopez intitulée *Quelles nouvelles conceptions de l'évaluation en classe ? Vers des évaluations pour apprendre durablement* ; dans J.-F. Chesné, L. Piedfer-Quénéey (coord.) (2023). Conférence de consensus du CNESCO. L'évaluation en classe, au service de l'apprentissage des élèves. Notes des experts ; pp 13-22. https://www.cnesco.fr/wp-content/uploads/2024/07/CNESCO-CC-Eval_Notes-des-experts.pdf

¹⁶ Brigitte Hazard, Dominique Catoir, Jean Aristide Cavaillès, Antonella Durand, Michel Quéré (2021). L'évaluation des élèves au lycée (général, technologique et professionnel). Rapport n° 2021-159, octobre 2021. IGÉSR

¹⁷ Le lecteur peut obtenir des éléments approfondis sur le sujet dans D. Lafontaine, M.-C. Toczec-Capelle (2023). L'évaluation en classe au service de l'apprentissage des élèves : rapport de synthèse. CNESCO - CNAM: https://www.cnesco.fr/wp-content/uploads/2023/04/CNESCO-CC-Eval_LAFONTAINE_TOCZEK-CAPELLE.pdf

méconnaissance de l'intérêt que présente cette modalité sont les principaux obstacles à leur mise en place. Il s'agit pourtant encore une fois d'un levier d'amélioration des apprentissages permettant une adaptation au plus tôt par le professeur de son enseignement (régulation en amont des actions pédagogiques), sans avoir recours à la notation. Il permet aussi aux élèves de réactiver leurs connaissances, de visualiser plus précisément leur horizon de travail et leurs perspectives de progrès et d'insérer un cours particulier dans un ensemble plus large.

Les évaluations formatives (réalisées par exemple en cours sans nécessité de notation) – présentant cette fonction avec un retour vers les élèves explicite dans le but d'une régulation efficiente des apprentissages des élèves – sont aussi peu employées pour des raisons similaires. Ce type d'évaluation est cependant à encourager.

Les exercices et les devoirs à la maison

Si les exercices à la maison sont relativement fréquents, l'activité de devoir à la maison l'est de moins en moins. Les situations observées sont hétérogènes selon les classes. Des enseignants rencontrent des difficultés pour obtenir un engagement constructif des élèves lorsque l'activité n'est pas notée. Ils évoquent notamment une perte du goût de l'effort et des problèmes de recopiage. Le groupe regrette que les enseignants ne mettent pas suffisamment en avant les objectifs précis et motivants associés aux devoirs à la maison selon leur finalité éducative afin que les élèves y voient un bénéfice. Le choix méticuleux des devoirs à la maison, des notions qui y sont abordées et des compétences qui y sont développées est particulièrement important si l'on souhaite que les élèves trouvent leur intérêt à s'y engager.

Une diversité et un renouvellement des situations d'évaluation de régulation à accroître sur le cycle terminal

Les évaluations signalées dans cette partie 2.2 présentent, pour le groupe, de nombreux avantages formatifs et leur utilisation permettrait de diversifier les pratiques évaluatives des enseignants. Le groupe observe pourtant une certaine résistance de la part d'une partie des enseignants face aux pratiques pédagogiques moins traditionnelles de l'évaluation. Les IA-IPR ont un rôle essentiel à jouer pour surmonter ces résistances en présentant ces pratiques aux enseignants, lors des rendez-vous de carrière, des visites d'accompagnement, des réunions d'équipe ou encore à l'occasion de formations spécifiques. Il convient de ne pas encourager la pratique de ces modalités sous prétexte d'innovation (plusieurs sont connues et pratiquées depuis longtemps) ou sur la base d'arguments purement « théoriques », mais plutôt en insistant sur les bénéfices qu'elles présentent pour les apprentissages des élèves, en s'appuyant sur des exemples concrets. L'emploi d'outils numériques peut également être conseillé pour une mise en œuvre efficace d'évaluations de régulation, comme le montrent certains enseignants. L'intelligence artificielle (IA) peut, par exemple, être mobilisée pour faciliter la création de questionnaires ou la génération de QCM.

2.3. Des pratiques d'évaluation qui touchent à l'autonomie, la prise d'initiative, l'oralité et au collectif, jugées moins faciles à mettre en œuvre et moins prioritaires

Les enseignants rencontrent des difficultés pour la pratique de l'évaluation sur des champs qui touchent aux compétences transversales relatives à l'autonomie, la prise d'initiative, l'oralité et le collectif. Des interrogations existent à propos des méthodes à employer – avec le choix de critères pertinents – et de l'équité de jugement.

L'évaluation par compétences des tâches complexes, des résolutions de problème, des questions ouvertes

Les questionnements ouverts sollicitant des raisonnements plus élaborés sont en retrait : les activités comme les résolutions de problème sont devenues plus rares et l'évaluation par compétences associée n'a que rarement été adoptée ; seuls les enseignants les plus experts s'y investissent. Pourtant, lorsqu'ils y sont exposés, les élèves apprécient l'évaluation par compétences, en complément des notes. Le groupe observe également un recours rare à la démarche inductive. L'accompagnement des équipes sur le terrain par les inspecteurs territoriaux peut contribuer à améliorer ce point, une évolution au niveau des énoncés et les corrigés d'épreuves de baccalauréat pouvant venir en appui en montrant l'exemple, ce qu'ils tâchent de faire depuis plusieurs années.

Une évaluation de l'oral rare et dispersée

Il a déjà été souligné que les évaluations sont le plus souvent centrées sur les écrits (principalement les DS et les comptes rendus de TP) : peu de place est accordée à une évaluation explicite de l'oral. Seuls quelques enseignants prennent parfois un temps, souvent en début de séance, pour un exposé oral d'élève(s) sur un questionnement, la grille d'évaluation pouvant être explicitée avec des critères relatifs non seulement au contenu mais aussi à l'attitude, au niveau sonore, au débit de paroles et au respect du temps imposé.

La préparation au Grand oral n'est pas une incitation suffisante pour que les enseignants pratiquent des évaluations orales. Celles-ci restent exceptionnelles et les établissements qui organisent des Grands oraux blancs le sont tout autant.

Le groupe estime pourtant que les situations pouvant se prêter à une évaluation à l'oral sont nombreuses en physique-chimie. Les activités expérimentales, notamment, se prêtent bien au développement des capacités orales et à leur évaluation. Ces activités peuvent en particulier mobiliser des compétences orales descriptives (d'un protocole, d'un phénomène, d'une courbe expérimentale, etc.) et des compétences d'argumentation lorsqu'il s'agit de valider ou non une hypothèse, par exemple.

Les auditions indiquent que l'oral est plus travaillé dans l'enseignement scientifique, dans lequel les enseignants estiment avoir plus de temps pour mettre en œuvre son évaluation.

L'évaluation collective, de groupe

Les évaluations de groupe sont réalisées surtout via les comptes rendus de TP, qui sont souvent rendus par binôme. Cette évaluation collective gêne parfois les enseignants pour des questions d'équité entre les membres des groupes. Ce sujet mérite d'être un objet d'échange des inspecteurs avec les équipes.

L'évaluation déportée

La recherche montre l'importance d'impliquer les élèves dans l'évaluation ; cela facilite leur compréhension des critères d'évaluation et des attendus. Cette implication des élèves – par exemple par des techniques d'autoévaluation, de co-évaluation ou d'évaluation par les pairs – est très peu observée sur le terrain et gagnerait à être développée ; ces techniques d'évaluation dite déportée sont mal connues et jugées moyennement faisables par la plupart des enseignants alors que certains savent les exploiter. Elles sont davantage observées au collège ou dans d'autres disciplines.

Des formations pourraient être mises en place à propos de l'évaluation déportée au lycée. Il est à souligner que cette pratique ne doit pas consister en l'établissement d'une note par un élève. Il s'agit que celui-ci forme un jugement qualitatif pour des critères d'évaluation bien ciblés par l'enseignant pour éviter toute surcharge cognitive : les critères les plus importants sont à prioriser, ce qui peut dépendre de l'élève. L'enseignant doit également apporter un retour final à l'élève en s'appuyant sur les éléments d'évaluation produits par celui-ci, qu'ils soient corrects ou non.

2.4. Un accompagnement des élèves avant, pendant et après l'évaluation qui doit être amélioré

Avant l'évaluation

Le groupe constate que la plupart des enseignants n'explicitent ni les attendus ni les objectifs préalables à une évaluation. Les élèves ne savent pas suffisamment quoi travailler précisément pour l'évaluation – avec parfois la simple donnée du titre du chapitre à réviser – et se retrouvent pour certains surpris au moment de l'évaluation. Il conviendrait de mieux cibler les savoirs et savoir-faire visés. Le défaut d'explicitation réduit la qualité de préparation de l'évaluation par l'élève et tend à le placer dans une situation de stress, surtout pour les évaluations notées à fort enjeu.

Pendant l'évaluation

Les pratiques de différenciation pédagogique au cours des évaluations sont pauvres. Le groupe a, au mieux, observé des évaluations avec cours autorisé ou encore avec des « coups de pouce » de l'enseignant. Il existe une importante marge de progression sur ce point.

Après l'évaluation

Les élèves prennent le plus souvent connaissance des notes obtenues via leur espace numérique de travail (ENT) avant la restitution de leurs résultats en classe, même si parfois des enseignants prennent soin d'afficher la note sur l'ENT après restitution. Cette restitution comporte généralement des corrigés, des copies annotées et des commentaires généraux en classe. Le groupe a observé que le retour fait aux élèves est fréquemment réduit à la note et comporte peu de commentaires opérationnels sur lesquels l'élève pourrait s'appuyer pour progresser.

Ainsi, l'explicitation des critères d'évaluation fait le plus souvent défaut : la transparence de l'évaluation est à accroître, qu'elle soit notée ou non. L'exploitation post-évaluation avec un accompagnement du travail de l'élève sur les éléments restitués – hiérarchisation des points à revoir comprise – est à renforcer pour assurer un meilleur retour vers les élèves et faciliter ainsi leur progrès.

Le groupe a rarement observé des remédiations constructives qui font suite à des évaluations peu réussies, ce qui réduit leur fonction formative. Les enseignants ont plutôt recours à un jeu sur les barèmes, les coefficients, le calcul de moyenne de période afin d'atténuer l'impact chiffré de ces évaluations. Pour autant, cela ne suffit pas à renforcer les apprentissages et à faire progresser les élèves. Quelques évaluations de type « seconde chance » sont mises en œuvre, visant un retour sur les erreurs commises antérieurement pour permettre une remédiation.

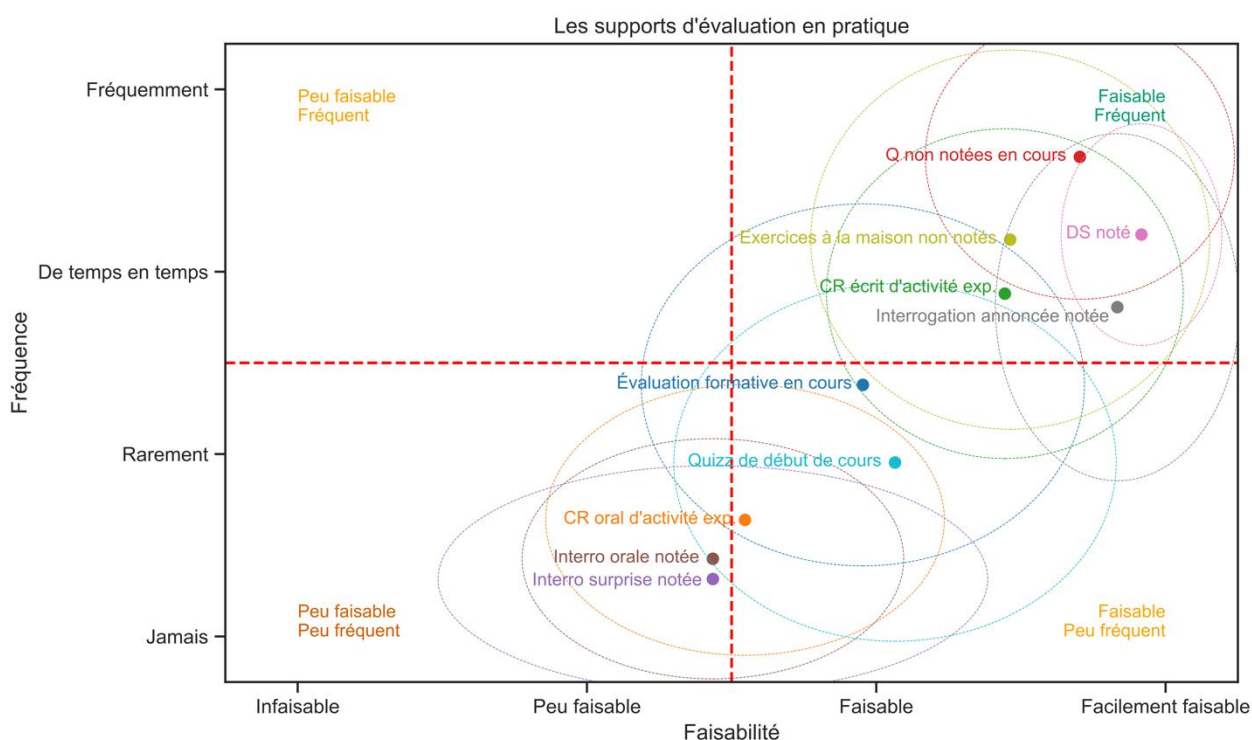
Il convient d'assurer globalement une meilleure intégration pédagogique de l'évaluation : il ne s'agit pas uniquement de sanctionner un travail par une note, mais aussi – et surtout – d'accompagner au mieux chaque élève dans ses apprentissages en lui proposant des tâches, des informations associées et un retour adapté.

2.5. Les pratiques d'évaluation des professeurs vues par eux-mêmes

Le questionnaire distribué aux enseignants¹⁸ les interrogeait sur différentes modalités d'évaluation en leur demandant de les classer selon deux axes : la fréquence de mise en œuvre (en quatre niveaux) et la faisabilité de cette mise en œuvre (en quatre niveaux également). Les résultats sont représentés sur les figures n° 3 et n° 4, où chaque point est associé à une modalité particulière et situé sur le plan en fonction des réponses agrégées des enseignants.

¹⁸ Voir l'annexe 3.

Figure n° 3 : réponses aux questions 3 et 4. Analyse en faisabilité / fréquence des modalités d'évaluation, selon les professeurs. Pour chaque modalité, la moyenne des réponses obtenues à la question 3 est traduite en un indice numérique de faisabilité, celle des réponses obtenues à la question 4 est traduite en un indice de fréquence. Ces deux indices définissent un point dans le plan faisabilité / fréquence. Les ellipses symbolisent la dispersion des réponses : plus elles sont grandes, plus les réponses sont variées.



Les supports d'évaluation vus par les enseignants (figure n° 3)

Les formes d'évaluation les plus courantes, telles que les devoirs surveillés notés, les interrogations prévues et notées, les devoirs à faire à la maison et non notés, les questions posées aux élèves pendant le cours, les comptes rendus classiques de travaux pratiques et les quiz (bien que ces derniers soient jugés un peu moins faisables), sont considérées comme réalisables par les professeurs. Il existe un consensus, tant du côté des élèves que du côté des enseignants, sur la faisabilité des évaluations lorsqu'elles sont planifiées. Cependant, l'évaluation formative en cours semble étonnamment moins faisable, ce qui pourrait être dû à une mauvaise compréhension de la question par les répondants au questionnaire, qui imaginaient peut-être une évaluation écrite formalisée.

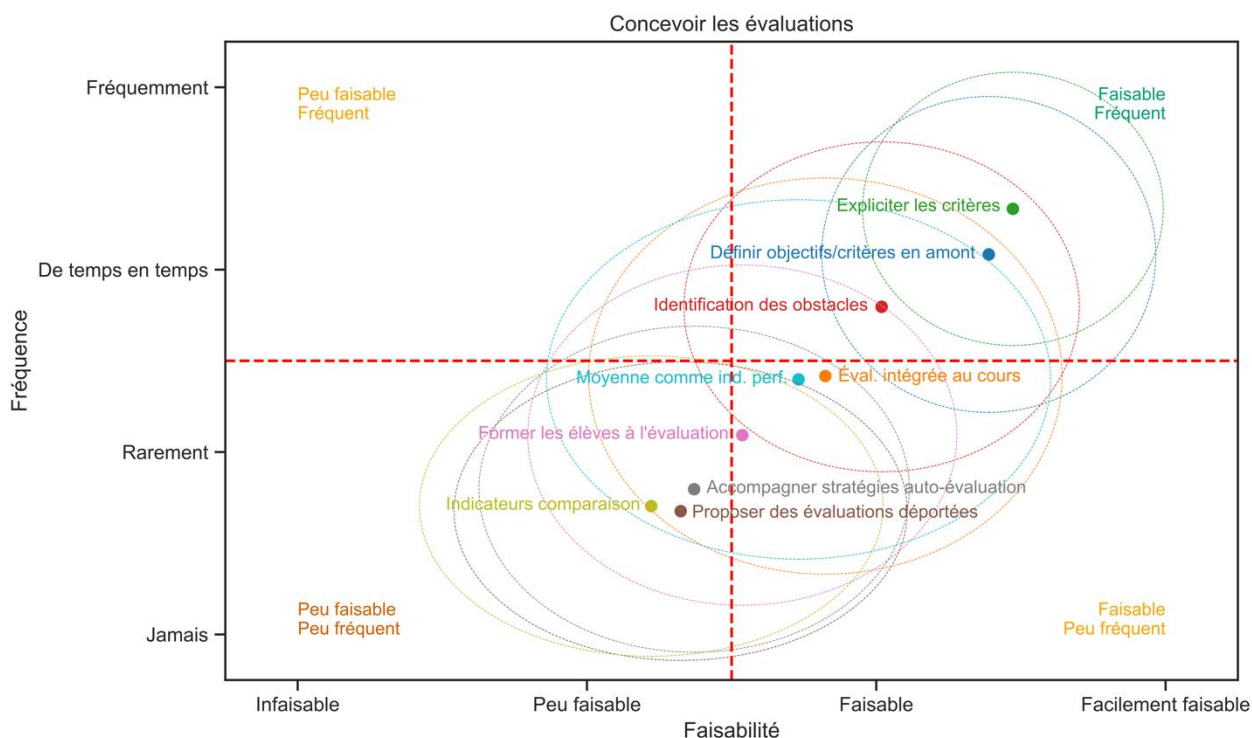
Les pratiques d'évaluation les plus fréquemment déclarées par les enseignants incluent les questions orales pendant le cours, les devoirs et interrogations surveillés notés, ainsi que les exercices à la maison non notés. Il est à noter que les supports d'évaluation utilisés semblent peu variés, selon les professeurs eux-mêmes. Sans surprise, cela reflète la faisabilité perçue : ce qui est jugé moins faisable est généralement moins réalisé. Paradoxalement, certaines modalités jugées faisables, comme les quiz ou les comptes rendus oraux de TP, ne sont jamais mises en œuvre, probablement en raison d'une réticence à s'écarter des habitudes et des usages pour proposer des évaluations innovantes. Ici également, l'IA peut être mobilisée et aider à vaincre ces réticences chez certains.

La conception des évaluations par les enseignants (figure n° 4)

Lors de la conception des évaluations, les enseignants déclarent souvent intégrer dès le départ les critères d'évaluations et les partager. Ils affirment également identifier les obstacles de temps en temps. Cependant, l'idée d'impliquer davantage les élèves dans le processus d'évaluation, que ce soit en tant qu'évaluateurs ou concepteurs des évaluations, n'est pas vraiment jugée faisable. De plus, la comparaison entre élèves apparaît comme une pratique jugée peu faisable et elle est rarement mise en œuvre.

La fréquence des évaluations dépend essentiellement de leur faisabilité, c'est-à-dire de la capacité à les intégrer de manière réaliste et efficace dans le déroulement des séquences pédagogiques : sans grande surprise, seul ce qui semble faisable est effectivement mis en place.

Figure n° 4 : réponses aux questions 5 et 6. Analyse en faisabilité / fréquence des caractéristiques de conception des évaluations, selon les professeurs. Pour chaque critère / objectif de conception, la moyenne des réponses obtenues à la question 5 est traduite en un indice numérique de faisabilité, celle des réponses obtenues à la question 6 est traduite en un indice de fréquence. Ces deux indices définissent un point dans le plan faisabilité / fréquence. Les ellipses symbolisent la dispersion des réponses : plus elles sont grandes, plus les réponses sont variées.

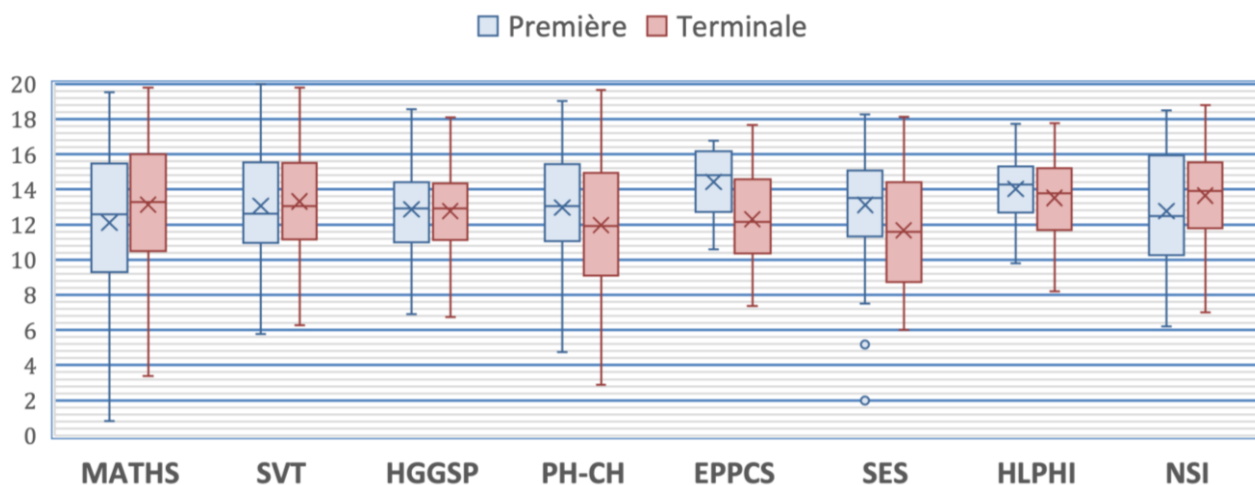


2.6. Comparaison avec d'autres enseignements de spécialité

Le groupe a souhaité tester à partir de données objectives la réputation de sévérité de notation de la discipline physique-chimie en analysant les statistiques, parfois disparates, que les établissements ont pu lui fournir. Le résultat est résumé dans la figure n° 5. On voit que la réputation de sévérité de notation en physique-chimie au lycée est, dans les faits, peu justifiée en classe de première (4^e moyenne parmi les 8 disciplines étudiées). Elle l'est davantage en classe de terminale où la moyenne des étudiants en physique-chimie est la 7^e sur les 8 disciplines de l'étude.

Les données sur lesquelles le groupe a pu travailler concernent seulement quelques centaines d'élèves et n'ont par conséquent qu'une valeur statistique limitée. Il serait souhaitable que les groupes disciplinaires puissent disposer de données de cette nature à l'échelle nationale, afin d'objectiver les comparaisons entre disciplines qui sont parfois énoncées sur la base de réputations d'origines incertaines – avec une finalité qui n'est pas toujours innocente dans un contexte de concurrence entre enseignements disciplinaires. Il y a là un objet de réflexion pour le pôle Données, Intelligence Artificielle et Numérique (DIAN) récemment créé dans l'IGÉSR.

Figure n° 5 : distribution des moyennes de contrôle continu observées en classe de première et de terminale, représentée en diagrammes en boîte¹⁹, pour différents enseignements de spécialité. Les croix représentent les moyennes ; les traits horizontaux, qui séparent les boîtes du deuxième et du troisième quartiles représentent les médianes



Lecture : pour la distribution des notes en classe de première de spécialité EPPCS, la moyenne est voisine de 14,4 et la médiane de 14,8. Le deuxième quartile de la distribution s'étend de 12,7 à 14,8 et le troisième quartile de 14,8 à 16,2.

3. Le pilotage des pratiques d'évaluation

L'importance des moyennes annuelles du contrôle continu dans le processus d'affectation et dans le calcul de la moyenne du baccalauréat²⁰ demande une équité irréprochable. La garantie de l'égalité de traitement des élèves face à l'examen et à l'orientation nécessite donc une plus grande harmonisation des pratiques d'évaluation durant le cycle terminal, tant sur le plan pédagogique (choix de la forme et du rythme de l'évaluation) que sur le processus d'élaboration des moyennes (modalités de calcul des moyennes, de prise en compte des notes, de gestion des absences, de la fraude, des besoins particuliers des élèves, etc.).

Les projets d'évaluation ont notamment été mis en place pour satisfaire à ce besoin ainsi que pour clarifier le rôle joué par les différents acteurs de la communauté éducative (enseignants, chefs d'établissement, IA-IPR de la discipline, collège des IA-IPR, Écoles académiques de la formation continue (EAFC), comités de pilotage ou groupes de travail académiques dédiés à l'évaluation, directions des examens et concours (DEC), etc.) dans le pilotage des pratiques d'évaluation. Ils donnent en premier lieu un cadre national pour outiller les académies et les chefs d'établissement en tenant compte des situations locales. Bien qu'ils constituent aujourd'hui un axe incontournable pour analyser le pilotage de l'évaluation au lycée général et technologique, le groupe a constaté que les équipes de direction et les équipes pédagogiques s'en emparent, en général, de façon très limitée, sans réelle conséquence sur l'harmonisation des pratiques.

3.1. Le projet d'évaluation, un outil national de cadrage du contrôle continu déployé de manière inégale en académie

Une mise en place très suivie

Les projets d'évaluation ont été mis en place par la note de service du 28 juillet 2021²¹ qui précise les modalités d'évaluation des candidats à compter de la session 2022 du baccalauréat général et technologique. Cette note de service définit le projet d'évaluation comme étant la formalisation du cadre de l'évaluation en

¹⁹ Se reporter par exemple à l'article wikipedia https://fr.wikipedia.org/wiki/Boîte_à_moustaches pour la structure et l'interprétation de ces diagrammes en boîte, parfois dénommées « boîtes à moustaches ».

²⁰ En spécialité physique-chimie, les notes de contrôle continu de première interviennent seulement pour les élèves qui abandonnent cette spécialité en terminale.

²¹ <https://www.education.gouv.fr/bo/21/Hebdo30/MENE2121270N.htm>

contrôle continu des élèves des lycées généraux et technologiques, afin d'expliciter la façon avec laquelle sont construites leurs moyennes annuelles prises en compte pour la partie « contrôle continu » du baccalauréat.

Le projet d'évaluation comprend un volet transversal et organisationnel (gestion des absences, de la fraude, etc.) et un volet pédagogique (harmonisation des pratiques, sens de l'évaluation, formes d'évaluation, etc.), dont les académies se sont emparées variablement. Depuis 2021, l'institution s'est dotée d'outils nationaux pour alimenter l'évolution des pratiques d'évaluation au lycée (guide et F.A.Q. sur Éduscol²² : séminaires et formations nationales, etc.). Un groupe de pilotage national, porté par la mission du pilotage des examens (MPE) de la direction générale de l'enseignement scolaire (DGESCO) et l'IGÉSR, a accompagné et coordonné les travaux en académie, avec les doyens des IA-IPR et les DEC.

À la rentrée 2021, les recteurs d'académie ont impulsé les réflexions collectives, à l'échelle académique et au sein des établissements, pour que chaque lycée produise son projet d'évaluation. Les lycées généraux et technologiques (LGT) publics ont, à quelques exceptions près, tous produit leur projet d'évaluation dès la première année, mais l'engagement des lycées privés a été plus inégalement suivi et a pris plus de temps.

Une dynamique accompagnée mais peu entretenue

Depuis, des modalités de suivi et d'accompagnement ont souvent été mises en place pour entretenir la dynamique créée dans les académies. Des séminaires de cadres, inspecteurs et proviseurs ont pu être organisés dans une petite moitié des académies observées. Les EAFC des trois quarts d'entre elles proposent ou ont proposé des formations sur le thème de l'évaluation, soit transversales, soit disciplinaires, à destination des enseignants et, quelques fois, des chefs d'établissement. Une petite partie des académies a pu produire une synthèse des projets d'évaluation de ses établissements, qui a été utilisée pour orienter le pilotage de l'évaluation.

Certaines académies se sont emparées des fiches statistiques permettant à chaque lycée de comparer les moyennes de ses élèves dans chaque matière, en contrôle continu et aux épreuves terminales, avec les moyennes académiques et les moyennes nationales. La réécriture et l'évolution des projets d'évaluation au fil des rentrées scolaires a été explicitement soutenue par certains recteurs.

D'une façon générale, cependant, force est de constater que les projets d'évaluation ont peu évolué depuis leur création et que, dans les établissements, la dynamique est rompue. Le volet pédagogique, qui avait souvent été laissé pour plus tard en 2021 ne s'est que très rarement concrétisé depuis.

Un bilan contrasté

Bien que d'abord perçu par des enseignants comme une contrainte à leur liberté pédagogique, le caractère régulateur du projet d'évaluation est souvent apprécié des équipes rencontrées par le groupe pour gérer avec les familles les conflits sur les notes de contrôle continu, et ainsi apaiser le stress des élèves face à l'évaluation.

Certains établissements se sont réellement emparés des enjeux de l'évaluation au lycée pour harmoniser et expliciter leurs pratiques. Néanmoins, la portée de ce dispositif reste inégale et ces cas se limitent trop souvent aux lycées qui avaient déjà une forte culture de travail collaboratif.

3.2. Des projets d'évaluation qui laissent peu de part aux aspects pédagogiques

La quasi-totalité des équipes des 19 lycées rencontrées par le groupe se sont engagées dans l'élaboration et l'application de leur projet d'évaluation, mais avec des stratégies diverses reflétant autant les spécificités de chaque établissement que celles du pilotage de leur académie. Certaines ont parfois collaboré afin d'harmoniser leurs pratiques à une plus grande échelle que celle d'un seul établissement, en particulier dans le secteur privé sous contrat.

Les disparités de rédaction sont grandes entre établissements, tout comme entre disciplines dans le même lycée. La version la plus souvent rencontrée est très prudente, très générale, et donc minimaliste, ignorant tout aspect pédagogique et se limitant aux aspects réglementaires pour prévenir les stratégies d'évitement

²² <https://eduscol.education.fr/2688/modalites-d-evaluation-pour-le-baccalaureat>

des évaluations par les élèves (gestion des absences et mise en place des épreuves de remplacement), gérer les fraudes et les réclamations des parents.

Certains établissements complètent leurs projets d'évaluation par des annexes disciplinaires, qui ne sont pas toujours diffusées aux familles. Des éléments de cadrage de la rédaction, des conseils et des modèles ont pu être diffusés aux chefs d'établissement par les IA-IPR ou les comités de pilotage académique de l'évaluation. Quoiqu'il en soit, la formalisation écrite des pratiques d'évaluation est particulièrement précieuse en cas de changement d'équipe de direction, ce qui n'est pas rare : depuis la mise en place des premiers projets d'évaluation des changements de direction se sont produits dans plus d'un tiers des lycées visités par le groupe.

Les chefs d'établissement qui ont saisi l'opportunité de la création des projets d'évaluations pour opérer un réel pilotage pédagogique des pratiques d'évaluation sont peu nombreux, parmi les lycées visités par le groupe. Quand c'est le cas, ils ont pu d'une part s'appuyer sur une dynamique volontariste et engagée de leurs équipes. Ils ont pu, d'autre part, solliciter les IA-IPR disciplinaires ou demander des formations d'initiative locale.

Le guide de l'évaluation de l'IGÉSR²³ est parfois apprécié pour les repères qu'il donne, mais il est aussi parfois considéré comme un modèle qui contraint la réflexion locale, quand il n'est pas complètement ignoré.

3.3. Un effet très limité sur les pratiques d'évaluation en physique-chimie

La plupart des équipes de physique-chimie rencontrées affirment que l'élaboration des projets d'évaluation n'a que très peu influencé leurs pratiques dans le domaine de l'évaluation des élèves. Selon elles, la plupart des recommandations qui figurent dans le guide de l'évaluation de l'IGÉSR étaient déjà mises en place avant la rentrée 2021.

Une formation pédagogique des enseignants insuffisante

Dans presque les deux tiers des académies observées, les IA-IPR ont mis en place des actions spécifiques pour accompagner les équipes de physique-chimie : webinaires de coordinateurs et réunions de bassin, réunions d'équipe et diffusion de ressources autrement. Ils auraient cependant souhaité pouvoir proposer aux enseignants un accompagnement plus approfondi en déployant des formations académiques spécifiquement dédiées aux aspects pédagogiques de l'évaluation, prenant en compte les particularités de la discipline (compétences de la démarche scientifique, évaluation de la pratique expérimentale, etc.). Selon la majorité d'entre eux, cet effort de formation n'a pas eu l'ampleur souhaitée.

Une dimension collective encore trop peu répandue

Pour les chefs d'établissement rencontrés, les professeurs de physique-chimie, ainsi que ceux de sciences de la vie et de la Terre (SVT), ont des pratiques collaboratives en général mieux installées que ceux d'autres disciplines, sans doute à cause de l'importance des pratiques expérimentales qui nécessitent davantage de concertation. De nombreuses équipes mutualisent leurs documents, choisissent une progression commune, ont des textes de travaux pratiques identiques. Le groupe a également observé que les professeurs d'une équipe disciplinaire de physique-chimie organisent presque toujours le même nombre de devoirs surveillés (trois par trimestre, le plus souvent, comme le préconisent la plupart des projets d'évaluation).

Malgré cela, dans le domaine de l'évaluation et même dans les équipes qui apparaissent les plus soudées, les pratiques partagées en physique-chimie restent limitées. L'évaluation des élèves est encore considérée par beaucoup d'enseignants comme le domaine le plus personnel de leur pratique professionnelle, on pourrait presque dire le plus intime.

Ainsi les entretiens menés par le groupe ont souvent été la première occasion d'échanges entre les professeurs sur leurs pratiques d'évaluation ainsi que sur la façon dont ils pondèrent les activités pour constituer leur moyenne.

²³ Ministère de l'éducation nationale et de la jeunesse, novembre 2023, Guide de l'évaluation des apprentissages et des acquis des élèves au lycée général et technologique, *op. cit.*

Les devoirs préparés en commun, en dehors des sujets de bac blanc, sont rares. Les énoncés et les barèmes de ces devoirs sont le plus souvent partagés, mais la détermination des critères d'évaluation l'est beaucoup moins fréquemment. De même, les corrections de copies sont rarement réalisées en équipe, chacun corrigeant les copies de ses élèves. Enfin, l'analyse quantitative comparée des résultats obtenus est exceptionnellement observée.

Recommandation 1

Améliorer les pratiques d'évaluation des enseignants et leur pilotage :

Un effort de formation significatif doit être fourni afin de parvenir à une meilleure acculturation des professeurs aux enjeux formatifs de l'évaluation des élèves ainsi qu'aux diverses modalités d'évaluation évoquées dans ce rapport. L'importance de la diversité des activités prises en compte (activités expérimentales, orales, évaluation déportée, etc.) ainsi que l'indispensable accompagnement des élèves par le professeur, avant, pendant et après l'évaluation, seront à souligner dans ces formations. Cet effort, déjà entrepris dans de nombreuses académies, doit être amplifié et accompagné au niveau national, notamment par une action du plan national de formation et par la production de ressources dédiées.

Au niveau local, une approche collective des pratiques évaluatives doit être développée, sous l'impulsion des chefs d'établissement, avec l'aide des IA-IPR. La dynamique collective qui a été souvent observée à la rentrée scolaire 2021 lors du travail d'élaboration des projets d'évaluation doit être relancée, pour compléter systématiquement ces derniers avec des lignes directrices relatives aux aspects pédagogiques de l'évaluation propres à chaque discipline.

4. L'équité de l'évaluation et son pilotage

Compte tenu des enjeux de l'évaluation de contrôle continu rappelés plus haut, il est de la responsabilité de l'ensemble du système scolaire d'en s'assurer l'équité. La note de service du 28 juillet 2021²⁴ portant sur les modalités d'évaluation des candidats prévoit divers dispositifs d'harmonisation entre établissements à partir des moyennes constatées en établissement (section 1F de la note de service), mais cela ne vaut que pour les notes du baccalauréat et pas pour les notes utilisées dans Parcoursup. Aucune procédure de régulation de l'équité n'est réglementairement prévue à l'échelle de l'établissement.

Le groupe a souhaité analyser les différentes observables dont l'institution dispose pour évaluer cette équité au sein de l'établissement, d'un professeur à l'autre, ainsi qu'entre les différents établissements pour s'assurer que les notes certificatives ou à enjeu d'orientation renseignent davantage leurs destinataires sur l'élève lui-même que sur son professeur ou son établissement d'origine.

4.1. Une équité qui semble aller de soi pour les équipes rencontrées

Les équipes de professeurs rencontrées sont pratiquement toutes convaincues qu'elles pratiquent une évaluation équitable. Le groupe constate cependant que ces équipes ne se donnent que très rarement les moyens d'étayer cette affirmation par une argumentation objectivée.

La comparaison des moyennes de chacun, la construction de celles-ci et les poids apportés à chaque type d'évaluation de même que les critères d'évaluation font très rarement l'objet d'échanges organisés entre les enseignants. Tout au plus donnent-ils lieu à des discussions informelles, sans trace écrite, par exemple à l'occasion d'une rencontre en salle des professeurs.

Toutes les équipes rencontrées jugent que la situation est satisfaisante et ne voient pas l'intérêt de modifier leurs pratiques à cet égard. Elles ont pourtant conscience, lorsqu'on les interroge, qu'il y a dans cette équité un enjeu important qu'elles sont, au bout du compte, les seules à pouvoir réguler.

L'attitude des chefs d'établissement est plus variable. Certains s'emparent réellement de la problématique et organisent, en général, une ou deux fois par an, des réunions avec chaque équipe pédagogique dans lesquelles des données objectives sont analysées : moyennes de contrôle continu, notes aux devoirs

²⁴ Ministère de l'éducation nationale et de la jeunesse, novembre 2023, Guide de l'évaluation des apprentissages et des acquis des élèves au lycée général et technologique, *op. cit.*

communs et bac blancs, notes aux épreuves terminales du baccalauréat, discussions sur la constitution des moyennes, etc. Dans le cas le plus fréquent, les chefs d'établissement n'interviennent pas du tout auprès des équipes pédagogiques qui sont ainsi seules garantes de l'équité de leurs pratiques.

Le groupe s'est interrogé sur les observables quantitatives dont disposent les chefs d'établissement et sur lesquelles ils pourraient s'appuyer, avec l'aide des inspecteurs territoriaux pour s'assurer que la notation est équitable dans leur établissement. Afin de comparer les pratiques d'évaluation des différents professeurs, les données quantitatives sont essentiellement de deux natures : les moyennes de contrôle continu attribuées aux élèves et, *a posteriori*, l'écart entre ces moyennes et les notes obtenues aux épreuves terminales. Le groupe a pu obtenir ces deux types de données dans les lycées visités et a pu ainsi analyser les usages que l'on peut en faire pour le pilotage de l'équité en établissement. Les paragraphes qui suivent montrent que ces analyses sont complexes et doivent être utilisées avec précaution.

La question de l'équité d'évaluation d'un établissement à l'autre est également essentielle ; elle relève essentiellement du pilotage par les IA-IPR. Ceux-ci disposent pour cela d'un riche éventail de données statistiques fournies par la DGESCO à chaque académie. Le groupe, qui a pu analyser des données de ce type, considère cependant que leur exploitation rigoureuse nécessite des renseignements encore plus fins. La simple comparaison des moyennes de contrôle continu ou des écarts entre ces moyennes et celles des épreuves terminales permet rarement de conclure, tant les biais statistiques (l'IPS individuel en premier lieu) sont importants (cf. infra).

4.2. Une analyse comparée des moyennes de contrôle continu attribuées par différents enseignants qui doit tenir compte du profil de leurs élèves

L'analyse comparée des moyennes posées par différents enseignants nécessite la prise en compte de l'hétérogénéité de composition des groupes d'enseignement de spécialité. En effet, les élèves diffèrent par plusieurs facteurs qui sont corrélés à leurs performances en physique-chimie. En classe de terminale, le groupe s'est concentré sur l'un de ces facteurs : le deuxième enseignement de spécialité et le choix d'options des élèves. Ceux-ci sont très fortement corrélés aux résultats obtenus en physique-chimie, comme le montre la figure n° 6, obtenue à partir des résultats d'élèves issus d'établissements différents et ayant des enseignants différents.

Figure n° 6 : distribution des moyennes de contrôle continu de terminale et des notes de l'épreuve de spécialité physique-chimie obtenues par cinq groupes d'élèves. Les groupes « M / Me » et « M » rassemblent les élèves qui suivent l'enseignement de spécialité mathématiques respectivement avec et sans l'option mathématiques expertes. Les élèves des groupes « SVT / Mco » et « SVT » suivent l'enseignement de spécialité SVT respectivement avec et sans l'option mathématiques complémentaires. Le groupe « Autre » désigne les élèves ayant choisi un autre enseignement de spécialité. L'échantillon total comprend 196 élèves

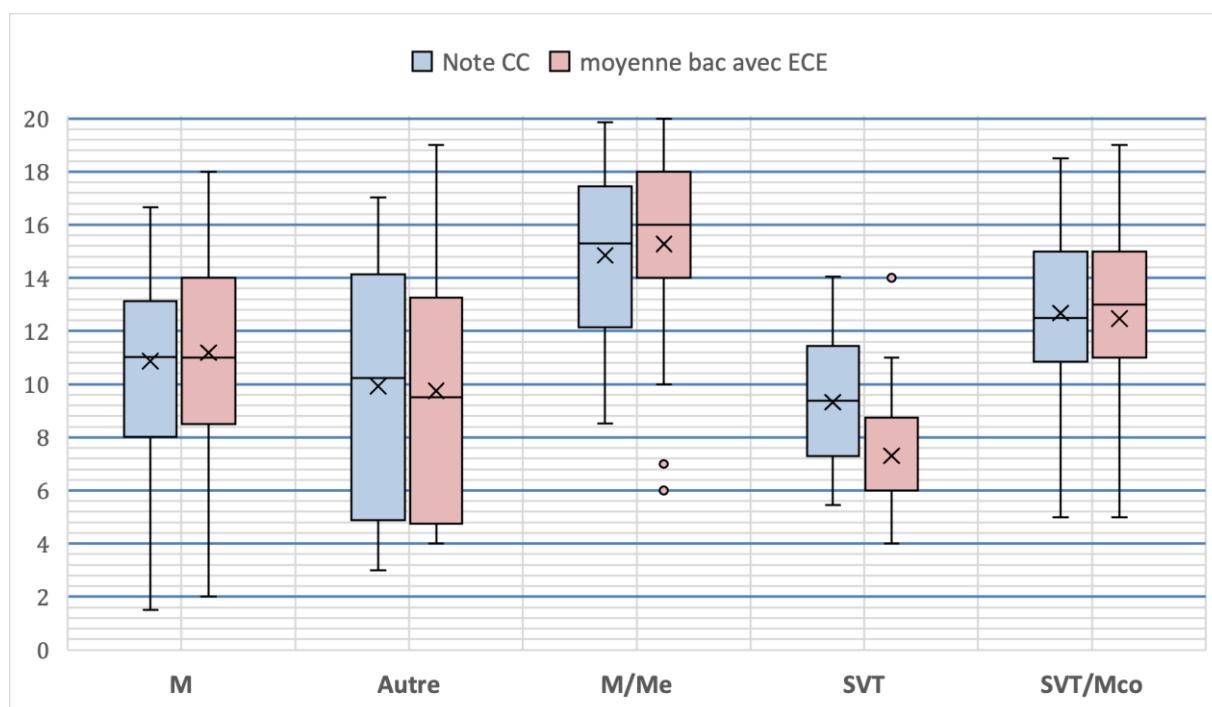


Tableau n° 1 : moyennes de contrôle continu obtenues par les cinq groupes considérés (voir légende de la figure n° 6)

Les moyennes obtenues par les cinq groupes d'élèves concernés sont résumées dans le tableau n° 1 :

	M / Me	M	SVT / Mco	SVT	Autre
Moyenne de contrôle continu en physique-chimie Classe de terminale	14,84	10,87	12,67	9,32	9,93
Écart à la moyenne générale des 196 élèves de l'échantillon étudié	+2,28	-1,69	+0,11	-3,24	-2,63

Selon le choix des enseignements de spécialité et d'option, les résultats moyens varient de près de 5 points sur 20. Il en résulte que, si on compare directement les moyennes de contrôle continu obtenues pour différents professeurs, on peut être amené à des conclusions erronées si l'on ne tient pas compte de la composition des groupes auxquels ils enseignent en termes de second enseignement de spécialité ou d'enseignements optionnels de mathématiques.

La figure n° 7-a) montre à titre d'exemple les moyennes de contrôle continu des élèves de trois professeurs différents d'un même établissement. On pourrait déduire de ces données que le professeur « A », dont les élèves ont une moyenne de 13,26, évalue ses élèves de façon plus généreuse que les deux autres (moyennes de 12,28 et 12,24). Il apparaît cependant que les compositions des classes des trois enseignants sont très différentes : le groupe du professeur « A » est essentiellement constitué d'élèves ayant choisi l'enseignement de spécialité mathématiques, avec ou sans option mathématiques expertes. Il n'est donc pas surprenant que la moyenne de ce groupe soit plus élevée que celle des deux autres.

On peut tenter de compenser le biais lié au second enseignement de spécialité et aux enseignements optionnels de mathématiques en corrigeant les notes des élèves de chaque groupe des écarts des moyennes constatés statistiquement et rapportés dans le tableau n° 1. On obtient alors les résultats de la figure n° 7 b) dans laquelle on voit que les résultats sont plus homogènes entre les trois professeurs, mais où c'est maintenant le professeur « C » qui semble se singulariser en notant plus sévèrement ses élèves que ses deux collègues.

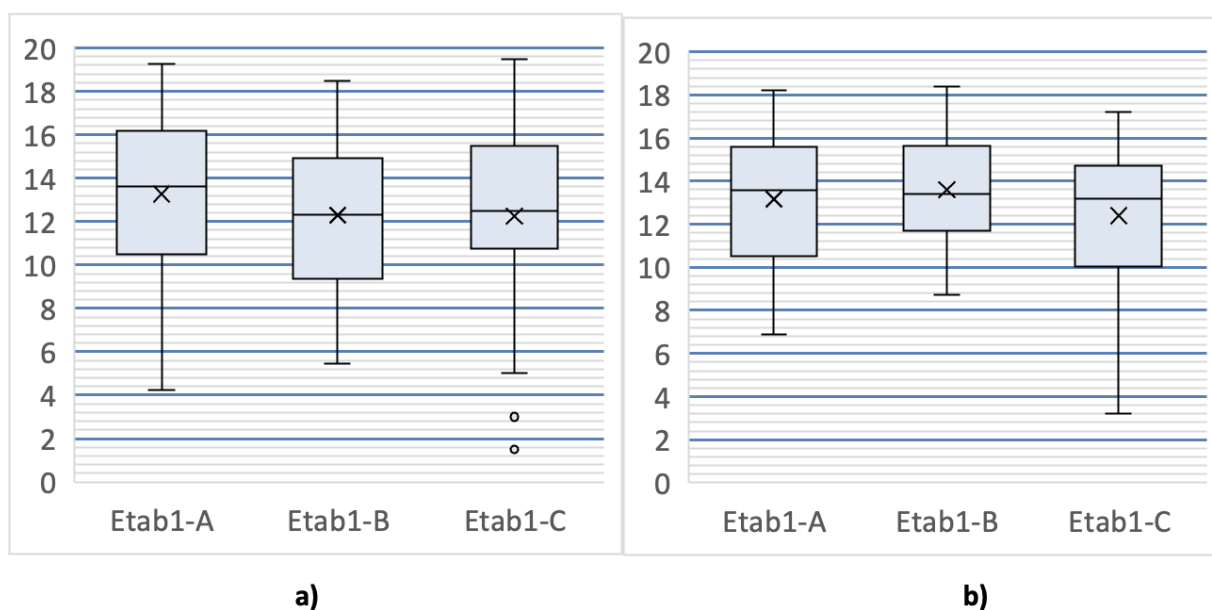
Ainsi, l'analyse des moyennes de notes de différents professeurs doit être très prudente et prendre en compte les facteurs objectifs qui peuvent justifier que la moyenne des notes d'un enseignant présente un écart important avec celles de ses collègues. Même lorsque certains facteurs sont pris en compte, comme le groupe a tenté de le faire, il n'est pas aisé de conclure car les échantillons sur lesquels les moyennes sont réalisées restent limités et peuvent connaître des fluctuations importantes.

La comparaison des moyennes n'est donc pas un indicateur à considérer de façon automatique. Elle peut cependant constituer une alerte, incitant à analyser plus avant les circonstances dans lesquelles les enseignants exercent, de façon à expliquer une partie des écarts éventuellement observés.

Pour accomplir ces analyses, les chefs d'établissement ont besoin de données leur indiquant les impacts statistiques constatés de telle ou telle caractéristique des élèves, à grande échelle, académique ou nationale.

Par ailleurs, si l'on cherche déceler une éventuelle anomalie dans mes pratiques d'évaluation d'un enseignants, les données obtenues pour une seule année scolaire ne sont pas suffisantes. C'est seulement si, sur plusieurs années, un enseignant attribue à ses élèves des notes dont les moyennes, corrigées des biais statistiques, sont systématiquement plus élevées ou plus basses que celles de ses collègues que l'on pourra conclure à une anomalie.

Figure n° 7 : distributions des moyennes de contrôle continu obtenues dans les classes de trois professeurs différents d'un même établissement. a) moyennes brutes ; b) moyenne corrigée du biais lié au second enseignement de spécialité et aux enseignements optionnels de mathématiques²⁵



4.3. Une analyse de l'écart entre les notes de contrôle continu et les notes obtenues aux épreuves terminales du baccalauréat par les élèves de différents enseignants

Dans les comparaisons entre classes différentes, l'écart entre les notes de contrôle continu obtenues en classe et les notes obtenues aux épreuves terminales par cette même classe est une donnée plus facile à

²⁵ La moyenne corrigée est obtenue en soustrayant aux notes des élèves l'écart de moyenne constaté pour le groupe auquel ils appartiennent. On a par exemple soustrait 2,28 points de moyenne aux élèves du groupe « M / Me » et ajouté 3,24 points à ceux du groupe « SVT » (voir tableau n° 1).

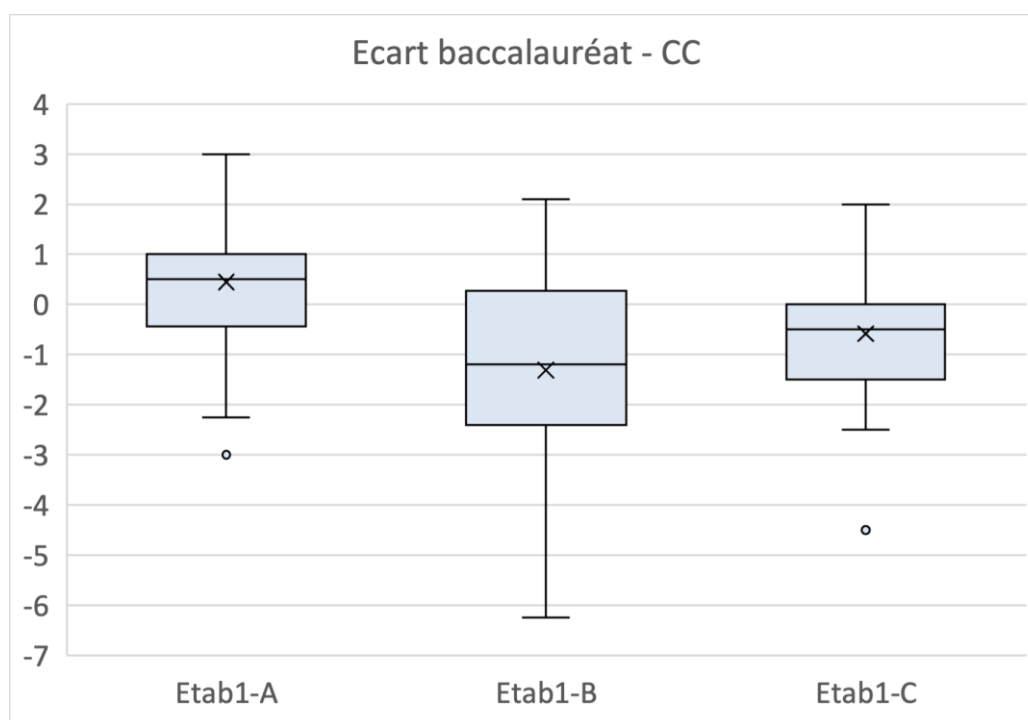
interpréter que les valeurs de ces notes elles-mêmes. En effet, les biais associés aux notes de contrôle continu sont comparables à ceux qui sont liés aux épreuves terminales et se compensent dans la différence²⁶.

C'est le cas du biais lié au second enseignement de spécialité et à celui d'option, visible sur la figure n° 6, où l'on voit que les différents groupes ont des résultats voisins en contrôle continu et en contrôle terminal.

De façon notable, on constate cependant que le groupe « SVT » qui a les notes les plus faibles en contrôle continu de physique-chimie a des notes encore plus faibles aux épreuves terminales de physique-chimie. Cette tendance serait intéressante à observer au niveau national. Elle pourrait s'expliquer par le fait que les professeurs peuvent avoir tendance à surnoter leurs élèves les plus faibles de façon à ne pas les démotiver, pratique à laquelle on peut souscrire d'un point de vue pédagogique dans la mesure où ce « bonus de motivation » reste raisonnable.

Si l'on analyse les écarts contrôle continu-baccalauréat observés entre les professeurs dont les moyennes de contrôle continu ont été étudiées au paragraphe précédent, on obtient les résultats de la figure n° 8. On voit que les élèves du professeur « A » ont, en moyenne, de meilleures notes en contrôle continu qu'au baccalauréat, alors que c'est le contraire pour les professeurs « B » et « C ». Ici encore les écarts restent limités, sauf peut-être dans le cas du professeur « B », mais on ne peut être affirmatif sur le fait que les pratiques de ce professeur sont anormales. Comme cela a déjà été souligné, pour identifier clairement une anomalie dans la façon de noter de ce professeur, il conviendrait d'analyser l'écart entre les moyennes de contrôle continu et les notes obtenues dans les épreuves terminales durant plusieurs années. Si cet écart se confirmait, il serait utile d'en rechercher les causes possibles.

Figure n° 8 : différence entre les notes de baccalauréat et les notes de contrôle continu des élèves de chaque professeur. Un écart positif signifie que la note obtenue au baccalauréat est supérieure à celle de contrôle continu



²⁶ Cette affirmation est vérifiée dans les statistiques limitées dont le groupe a disposé. Elle devrait être validée par une étude statistique à plus grande échelle.

4.4. Une analyse d'écart entre contrôle continu et contrôle terminal à partir de données académiques

Le groupe a pu disposer pour un panel de trois académies métropolitaines d'un ensemble complet de données d'évaluation des élèves, agrégées par établissement²⁷ : la moyenne des notes de contrôle continu (CC) du cycle terminal de la voie générale pour l'année scolaire 2022-2023 et celle du contrôle terminal (CT) du baccalauréat 2023 (épreuve écrite et épreuve d'évaluation des compétences expérimentales). Ce panel concerne des académies qui ont traditionnellement des résultats en physique-chimie dans la moyenne nationale ou au-dessus.

Une absence d'écart significatif des moyennes des CC et CT en physique-chimie

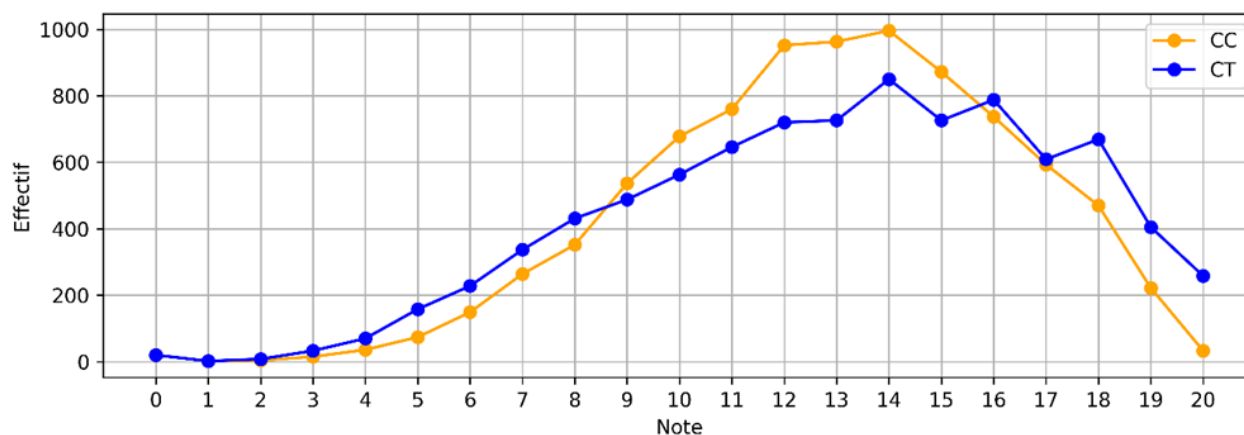
Les moyennes de CC et de CT en physique-chimie sur l'ensemble des établissements des trois académies considérées sont respectivement de 12,84 / 20 et 13,07 / 20. Elles ne montrent globalement aucune incohérence significative.

Ces moyennes se situent au-dessus de celles calculées sur l'ensemble des enseignements de spécialité pour le même panel (ces moyennes sont respectivement de 12,56 / 20 et 12,26 / 20). Le niveau de notation du CC en physique-chimie reste comparable à celui des autres enseignements de spécialité. Le CT se situe à un niveau plus élevé, probablement en raison du profil des élèves qui choisissent l'enseignement de spécialité physique-chimie (niveau académique favorable).

Une dispersion des moyennes du CT plus élevée que celle pour le CC

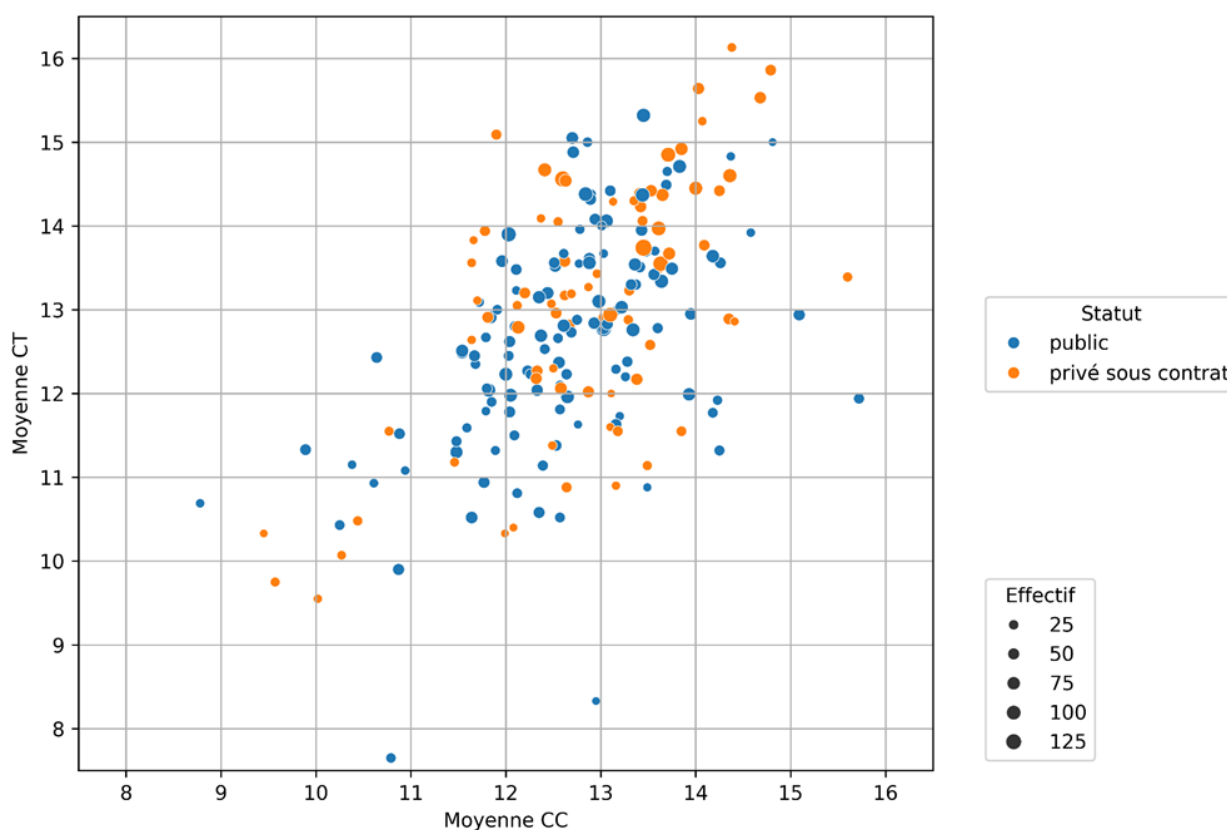
La dispersion des moyennes de physique-chimie sur le panel est un peu plus élevée pour le CT que pour le CC : les écarts-types respectifs sont de 3,98 et 3,37. La figure n° 9 permet de visualiser l'étalement plus important des moyennes pour le CT.

Figure n° 9 : distribution des notes de contrôle continu et de contrôle terminal sur le panel étudié de trois académies



²⁷ Dans cette partie, les statistiques sont donc calculées à partir des distributions de notes par établissement. Les données par élève n'étaient pas connues : par conséquent, les diverses grandeurs telles les moyennes et les écarts-types ne sont pas exactement relatives aux notes élève par élève.

Figure n° 10 : croisement des moyennes de contrôle continu (abscisse) et de contrôle terminal (ordonnée) par établissement avec précision de l'effectif concerné (taille de disque) et du statut (couleur) pour le panel étudié de trois académies



Une faible corrélation entre les moyennes de CC et de CT par établissement

On positionne sur le graphique de la figure n° 10 – avec pour abscisse la moyenne au CC et pour ordonnée la moyenne au CT – chaque établissement du panel (sauf les rares pour lesquels il manque l’une des données).

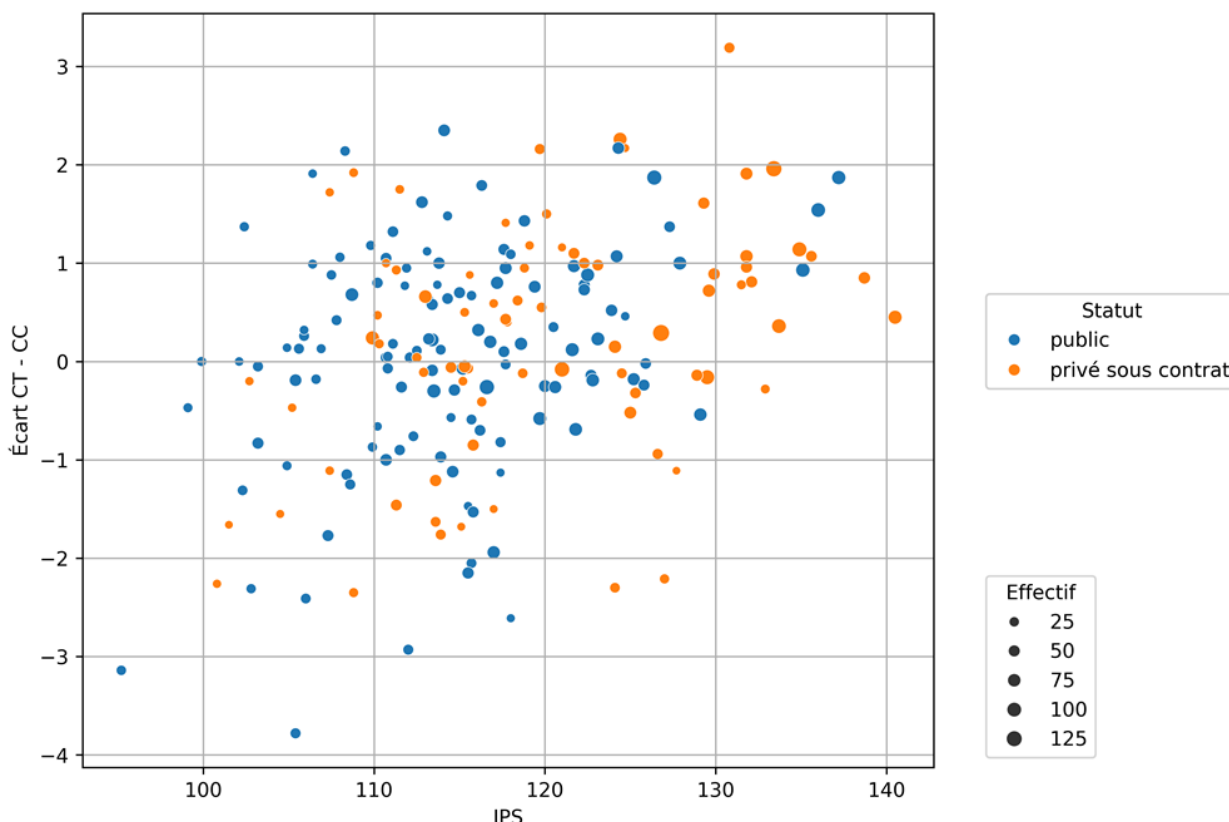
On note qu’un nombre non négligeable d’établissements, qu’ils soient publics ou privés sous contrat, ont des écarts significatifs entre leurs moyennes aux CC et CT (sur le panel de 205 lycées, la différence dépasse d’un point respectivement en faveur du CT et du CC pour 40 et 33 lycées). Les établissements privés se situent plus souvent dans le domaine des moyennes supérieures, avec un CT au-dessus du CC, ce qui est en lien avec le profil de leurs élèves (cf. infra).

Aucune corrélation entre les moyennes aux CC et CT ne ressort. Le groupe souligne qu’il n’y a pas lieu de tendre vers une corrélation linéaire entre ces moyennes et qu’il s’agit plutôt d’être attentif à la variabilité des écarts des moyennes : il convient qu’un regard soit porté par les inspecteurs territoriaux sur les établissements où les écarts peuvent interroger en analysant les éventuelles évolutions au fil des années.

Un focus selon l’indice de positionnement social (IPS) moyen par établissement pour la physique-chimie

Le graphique de la figure n° 11 permet de visualiser l’« écart CT - CC » – c’est-à-dire la moyenne de CT moins la moyenne de CC en physique-chimie – en fonction de l’IPS moyen sur les élèves de lycée GT pour chaque établissement du panel pour l’année scolaire 2022-2023 (données Archipel). Certains établissements ne sont pas pris en compte car cet IPS est manquant sur Archipel.

Figure n° 11 : crois ement de l'indice de positionnement social (IPS ; abscisse) et de l'écart entre les moyennes de contrôle terminal et de contrôle continu (ordonnée) par établissement avec précision de l'effectif concerné (taille de disque) et du statut (couleur) pour le panel étudié de trois académies



On n'observe pas de corrélation marquée entre l'écart CC-CT et l'IPS dans cette analyse réalisée par établissement. Néanmoins, il apparaît que les établissements aux populations scolaires de milieux sociaux les plus défavorisés (IPS plus faible) présentent plus souvent une moyenne de CT en physique-chimie inférieure à celle du CC : ceci est à mettre en relation avec la bienveillance plus importante précédemment soulignée du CC par rapport au CT. Pour les établissements avec les populations scolaires de milieux sociaux les plus favorisés, c'est le contraire : le CC a une moyenne en deçà de celle du CT, ce qui est probablement en relation avec le niveau d'exigence plus élevé dans les évaluations pratiquées par les enseignants pendant l'année.

4.5. Les observables de nature qualitative utiles au pilotage de l'équité

Si les données quantitatives disponibles doivent, comme on l'a vu, être mobilisées par le chef d'établissement, leur analyse est délicate et ces données ne méritent d'être considérées que comme des éléments d'alerte. Le groupe physique-chimie est convaincu qu'elles ne sont qu'une partie de l'analyse d'équité de notation que doit réaliser un chef d'établissement.

L'important est surtout de vérifier que les équipes pédagogiques s'assurent elles-mêmes, collectivement, que, dans leurs pratiques d'évaluation, tout est fait pour garantir l'équité de notation des élèves.

Il est donc, selon le groupe, essentiel que le chef d'établissement contrôle que cette dynamique collective est en place. Les visites réalisées montrent cependant qu'elle est pratiquement absente dans les établissements, même si les chefs d'établissements nous disent que les équipes de physique-chimie travaillent bien ensemble (voir le paragraphe 3.3).

Une véritable prise en charge collective de l'équité d'évaluation par les équipes peut s'appuyer sur les éléments suivants, qui peuvent être impulsés par le chef d'établissement et soutenues par les IA-IPR. Ces éléments constituent des observables de nature qualitative pour ces derniers, comme pour les

chefs d'établissements :

- organisation régulière, sur un rythme annuel, de réunions d'échanges formalisés – avec compte rendu – autour des pratiques d'évaluation au sein de l'équipe pédagogique ;
- attention portée à l'harmonisation de la construction des moyennes. La part relative des évaluations des compétences expérimentales, dont on sait qu'elles sont généralement généreusement notées, doit être commune aux enseignants ;
- harmonisation du nombre d'évaluations essentiellement sommatives (DS). Le groupe a observé que c'est généralement le cas, car ce point est fréquemment imposé par le projet d'évaluation de l'établissement ;
- partage des attendus et des critères d'évaluation dans les devoirs sommatifs ;
- partage des attendus dans les évaluations d'activités expérimentales ;
- organisation de devoirs sommatifs communs avec partage des attendus et des barèmes et corrections croisées des copies. La double ou triple correction d'un échantillon de quelques copies, peut faciliter ces partages.

Ainsi, aux côtés du chef d'établissement, les IA-IPR ont un rôle essentiel à jouer dans le pilotage pédagogique de l'équité au niveau de l'établissement et au niveau académique, entre établissements. Ils peuvent aborder le sujet avec les enseignants au cours des rendez-vous de carrière individuels, des réunions d'équipe, ou encore à l'occasion de formations spécifiques. Ils peuvent également mettre des ressources à disposition de leurs enseignants. Le groupe a pu constater que les équipes d'IA-IPR sont très engagées sur ce sujet et qu'elles ont dès à présent mis en place ce type d'actions dans la plupart des académies visitées.

Recommandation 2

Piloter l'équité d'évaluation en s'appuyant sur un usage raisonné des données statistiques :

Le pilotage de l'équité ne peut pas se passer d'une analyse, par le chef d'établissement et les inspecteurs, des résultats obtenus par les élèves, en contrôle continu et en contrôle terminal. Ces données doivent être rassemblées de façon systématique, chaque année sous la même forme, pour être à même d'identifier des tendances longues.

Ces données sont avant tout des indicateurs d'alerte permettant de repérer des situations hors norme. Les conclusions que l'on peut en tirer nécessitent une analyse approfondie qui doit prendre en compte, avec prudence et discernement, les différents biais statistiques. Pour identifier ces biais, le groupe recommande de conduire des études statistiques au niveau national permettant de quantifier l'influence de certains facteurs individuels sur les moyennes de contrôle continu et sur les notes des épreuves terminales obtenues dans chaque enseignement de spécialité (IPS, second enseignement de spécialité, options choisies, etc.). Ces données devront être accompagnées de conseils méthodologiques sur leur exploitation.

Conclusion

Recommandation 3 :

Développer une approche collective de l'évaluation :

Le groupe recommande que les chefs d'établissement et les IA-IPR organisent des échanges entre professeurs d'une même équipe pédagogique au sujet de l'évaluation. Ces échanges pourraient prendre la forme de réunions organisées annuellement et animées par le chef d'établissement ou de réunions d'équipe consacrées aux pratiques d'évaluation mises en place de façon ponctuelle par l'IA-IPR, à l'occasion d'une visite dans l'établissement. De telles pratiques se rencontrent déjà mais elles devraient être généralisées.

Au-delà des recommandations qui ont été faites dans le corps de ce rapport, le groupe estime qu'il est prioritaire de développer une approche collective des pratiques d'évaluation au sein des établissements.

L'analyse des données recueillies et les entretiens réalisés par le groupe montrent que, dans la grande majorité des cas, les enseignants évaluent leurs élèves avec compétence et sérieux, même si des progrès importants sont possibles dans l'utilisation des évaluations pour améliorer les apprentissages.

Les pratiques d'évaluation des enseignants restent très individuelles ; ils se concertent rarement avec leurs collègues sur ce sujet, même dans les équipes qui, par ailleurs, travaillent bien ensemble. De ce fait, les enseignants manquent d'occasions pour prendre du recul sur leurs pratiques personnelles et pour, éventuellement, les remettre en question. Ce manque de concertation est, de plus, défavorable à une approche harmonisée des évaluations de contrôle continu qui, seule, est à même de garantir une équité de traitement entre les élèves des différents professeurs.

Il n'est pas souhaitable d'uniformiser les pratiques d'évaluation, mais certains aspects de celles-ci doivent être partagés ; c'est, par exemple, le cas des critères de réussite, ou de la pondération des éléments qui interviennent dans le calcul des moyennes.

C'est seulement au prix de cette dynamique collective que l'institution pourra garantir qu'elle s'assure, au mieux, de la qualité des évaluations qu'elle produit.

Pour la mission,

Jean Aristide CAVAILLÈS

Annexes

Annexe 1 :	Liste des établissements visités.....	33
Annexe 2 :	Protocole de visites et d'entretiens	34
Annexe 3 :	Questionnaire envoyé aux professeurs.....	35

Liste des établissements visités

Le groupe tient à remercier vivement les équipes de direction des établissements visités pour leur accueil et pour les données qu'ils ont pu nous faire parvenir.

Il remercie chaleureusement l'ensemble des IA-IPR rencontrés au cours de ce travail, ainsi que celles et ceux qui ont accompagné les membres du groupe en établissements : Emmanuelle Laage (Versailles) , Isabelle Delarue (Versailles) , Isabelle Delattre (Rennes), Pierre Labarbe (Rennes), Sophie Cognac (Nancy-Metz), Tanguy Saibi (Nancy-Metz), Laurent Arer (Nancy-Metz), Olivier Michel (Reims), Antoine Maigret (Reims), Brigitte Abisset et Sandrine Cances (Toulouse), Pierre Rigat (Aix-Marseille), Michel Barde (Aix-Marseille), Guillaume Copin (Corse), David Boyer (Limoges), Claudine Ageorges (Limoges), Romain Salvan (Montpellier), Pierre Nass et Anne-Laure Schmitt (Strasbourg).

	Lycée	Ville (département)	Académie	IPS moyen LGT (source : Archipel)
LG	Frédéric Mistral	Avignon (84)	Aix-Marseille	111
LGT	Perier	Marseille 8ème (13)	Aix-Marseille	104
LG	Stendhal	Aiguillon (47)	Bordeaux	100
LGT	Fesch	Ajaccio (20)	Corse	113
LG	Bernard Palissy	Saint-Léonard-de-Noblat (87)	Limoges	115
LPO	Suzanne Valadon	Limoges (87)	Limoges	98
LPO	Jules Fil	Carcassonne (11)	Montpellier	103
LPO	Henri Vogt	Commercy (55)	Nancy-Metz	103
LGT	Louis Majorelle	Toul (54)	Nancy-Metz	111
LGT	Robert Schuman	Metz (57)	Nancy-Metz	93
LGT privé	Jeanne d'Arc	Nancy (54)	Nancy-Metz	105
LGT	Thierry Maulnier	Nice (06)	Nice	102
LG	Georges Clémenceau	Reims (51)	Reims	119
LPO	Marcellin Berthelot	Questembert (56)	Rennes	107
LGT	Charles de Gaulle	Vannes (56)	Rennes	118
LGT	Jean Monnet	Strasbourg (67)	Strasbourg	95
LGT	Toulouse Lautrec	Toulouse (31)	Toulouse	110
LGT	Eugène Ionesco	Issy-les-Moulineaux (92)	Versailles	132
LGT	Charles de Gaulle	Poissy (78)	Versailles	124

Protocole de visites et d'entretiens

Objectifs :

- observation des pratiques d'évaluation en enseignement de spécialité physique-chimie ;
- étude des formes d'évaluation mises en œuvre, de la constitution des moyennes, de la prise en compte des aspects formatifs de l'évaluation, de la gestion et du pilotage de l'équité d'évaluation ;
- analyse du ressenti des élèves sur l'évaluation en PC ;
- échanges sur les pratiques d'accompagnement des enseignants par les IA-IPR en ce qui concerne l'évaluation des élèves.

Méthodologie :

- réunions des équipes de professeurs de PC, rencontre de panels d'élèves et de l'équipe de direction dans plusieurs établissements ;
- échanges préalables (ou a posteriori) avec les IA-IPR des académies concernées sur le thème de l'évaluation ;
- demande de données et documents (professeurs et élèves rendus anonymes) :
 - projet d'évaluation de l'établissement,
 - pour chaque groupe de spécialité, en classes de première et de terminale, le relevé des moyennes de contrôle continu trimestrielles (ou semestrielles) et annuelles de chaque élève de l'enseignement de spécialité physique-chimie, avec indication des autres enseignements de spécialités suivis,
 - Pour chaque groupe de spécialité de terminale, pour les élèves de l'année 2022-2023, le relevé des moyennes de contrôle continu reportées dans les dossiers Parcoursup de chaque élève ayant suivi l'enseignement de spécialité physique-chimie, accompagné de la note obtenue par ces élèves en épreuve de spécialité physique-chimie du baccalauréat de la session 2023, ainsi que de l'indication des autres enseignements de spécialités suivis.
- envoi préalable d'un questionnaire aux professeurs de physique-chimie de l'établissement visité sur les objectifs de l'évaluation, la faisabilité et la fréquence des pratiques, la faisabilité et la fréquence des modes de conception des évaluations.

Questionnaire envoyé aux professeurs

Le questionnaire envoyé aux enseignants comportait, en plus des questions habituelles d'identification de l'académie et de l'établissement, six questions réparties en deux sections précisées ci-dessous. Les réponses étaient anonymes.

Les fonctions de l'évaluation

Dans cette section, il était demandé aux enseignants d'apprécier des réponses possibles aux questions suivantes avec quatre niveaux : pas du tout d'accord, pas vraiment d'accord, plutôt d'accord, tout à fait d'accord.

Question 1 : À quoi sert-il pour un professeur d'évaluer les acquis de ses élèves ?

Réponses possibles :

- Contrôler les connaissances des élèves ;
- Positionner un élève par rapport à une référence (moyenne) ;
- Connaître le degré de compréhension / maîtrise d'un élève ;
- Connaître le degré de compréhension / maîtrise d'un ensemble d'élèves ;
- Situer les progrès d'un élève ;
- Réguler / adapter son enseignement ;
- Déterminer l'orientation d'un élève vers le niveau supérieur ;
- Classer les élèves par niveau de performance ;
- Développer la motivation des élèves ;
- Rendre compte à la famille des résultats d'un élève ;
- Améliorer les apprentissages des élèves.

Question 2 : À quoi sert l'évaluation pour un élève ?

Réponses possibles :

- Se positionner par rapport à une référence (moyenne) ;
- Connaître son degré de compréhension / maîtrise à un instant donné ;
- Mesurer ses progrès ;
- Mesurer sa capacité à s'organiser, à être autonome, travailler en groupe... ;
- S'orienter ;
- Accroître son sentiment d'efficacité personnelle ;
- Encourager le travail personnel et autonome ;
- Améliorer ses apprentissages.

L'évaluation en pratique

Question 3 : Quels supports pour l'évaluation – analyse de faisabilité

Dans cette question, il était demandé aux enseignants d'apprécier la faisabilité de certaines modalités d'évaluations, selon quatre niveaux : infaisable, peu faisable, faisable, facilement faisable.

Modalités d'évaluation :

- Évaluation formative pendant le cours ;
- Compte-rendu à l'oral d'une activité expérimentale ;
- Compte-rendu à l'écrit d'une activité expérimentale ;
- Questions aux élèves (non notées) pendant le cours ;

- Interrogation « surprise » notée ;
- Interrogation orale notée ;
- Devoir écrit programmé, surveillé et noté ;
- Interrogation annoncée à l'avance et notée ;
- Exercices d'entraînement non notés (éventuellement en ligne) à faire à la maison ;
- Quiz en début de cours (par écrit ou par voie numérique).

Question 4 : Quels supports pour l'évaluation – analyse de fréquence

Dans cette question, il était demandé aux enseignants d'apprécier la fréquence à laquelle ils mettaient en place certaines modalités d'évaluations, selon quatre niveaux : jamais, occasionnellement (moins de trois fois par trimestre), de temps en temps (au moins trois fois par trimestre), fréquemment (plus de cinq fois par trimestre)

- Modalités d'évaluation ;
- Évaluation formative pendant le cours ;
- Compte-rendu à l'oral d'une activité expérimentale ;
- Compte-rendu à l'écrit d'une activité expérimentale ;
- Questions aux élèves (non notées) pendant le cours ;
- Interrogation « surprise » notée ;
- Interrogation orale notée ;
- Devoir écrit programmé, surveillé et noté ;
- Interrogation annoncée à l'avance et notée ;
- Exercices d'entraînement non notés (éventuellement en ligne) à faire à la maison ;
- Quizz en début de cours (par écrit ou par voie numérique).

Question 5 : Concevoir une évaluation – analyse de faisabilité

Dans cette question, il était demandé aux enseignants d'apprécier la faisabilité de certains critères ou objectifs pouvant guider la conception d'évaluations, selon quatre niveaux : infaisable, peu faisable, faisable, facilement faisable.

Critères / objectifs de conception :

- Définir, dès la conception d'une séquence pédagogique, les objectifs d'apprentissage et les critères d'évaluation associés ;
- Intégrer une évaluation comme composante à part entière du cours ;
- Définir les critères d'évaluation et en informer explicitement les élèves ;
- Identifier les obstacles potentiels (nombre d'exercices ou d'items à traiter, lisibilité du support, formulation des consignes, résultats qui conditionnent la poursuite de l'évaluation) dans une situation d'évaluation et les supprimer pour qu'elle soit accessible à tous les élèves ;
- Proposer des évaluations entre pairs et auto-évaluation ;
- Former les élèves à l'évaluation ;
- Accompagner les élèves dans l'élaboration individuelle de stratégies d'auto-évaluation ;
- Proposer aux élèves des indicateurs qui leur permettent de se comparer aux autres ;
- Choisir la moyenne (éventuellement pondérée) des évaluations notées comme indicateur de performance.

Question 6 : Concevoir une évaluation – analyse de fréquence

Dans cette question, il était demandé aux enseignants d'apprécier la fréquence à laquelle ils basaient leurs conceptions d'évaluation sur certains critères ou objectifs, selon quatre niveaux : jamais, occasionnellement

(moins de trois fois par trimestre), de temps en temps (au moins trois fois par trimestre), fréquemment (plus de 5 fois par trimestre)

Critères / objectifs de conception :

- Définir, dès la conception d'une séquence pédagogique, les objectifs d'apprentissage et les critères d'évaluation associés ;
- Intégrer une évaluation comme composante à part entière du cours ;
- Définir les critères d'évaluation et en informer explicitement les élèves ;
- Identifier les obstacles potentiels (nombre d'exercices ou d'items à traiter, lisibilité du support, formulation des consignes, résultats qui conditionnent la poursuite de l'évaluation) dans une situation d'évaluation et les supprimer pour qu'elle soit accessible à tous les élèves ;
- Proposer des évaluations entre pairs et auto-évaluation ;
- Former les élèves à l'évaluation ;
- Accompagner les élèves dans l'élaboration individuelle de stratégies d'auto-évaluation ;
- Proposer aux élèves des indicateurs qui leur permettent de se comparer aux autres ;
- Choisir la moyenne (éventuellement pondérée) des évaluations notées comme indicateur de performance.