

Quel est l'impact de l'évaluation par les pairs sur l'amélioration des productions écrites des élèves du secondaire inférieur issus du qualifiant en langues et en sciences ?

Auteur : Bastin, Louise

Promoteur(s) : Fagnant, Annick

Faculté : Faculté de Psychologie, Logopédie et Sciences de l'Éducation

Diplôme : Master en sciences de l'éducation, à finalité spécialisée en enseignement

Année académique : 2025-2026

URI/URL : <http://hdl.handle.net/2268.2/26908>

Avertissement à l'attention des usagers :

Tous les documents placés en accès ouvert sur le site le site MatheO sont protégés par le droit d'auteur. Conformément aux principes énoncés par la "Budapest Open Access Initiative"(BOAI, 2002), l'utilisateur du site peut lire, télécharger, copier, transmettre, imprimer, chercher ou faire un lien vers le texte intégral de ces documents, les disséquer pour les indexer, s'en servir de données pour un logiciel, ou s'en servir à toute autre fin légale (ou prévue par la réglementation relative au droit d'auteur). Toute utilisation du document à des fins commerciales est strictement interdite.

Par ailleurs, l'utilisateur s'engage à respecter les droits moraux de l'auteur, principalement le droit à l'intégrité de l'oeuvre et le droit de paternité et ce dans toute utilisation que l'utilisateur entreprend. Ainsi, à titre d'exemple, lorsqu'il reproduira un document par extrait ou dans son intégralité, l'utilisateur citera de manière complète les sources telles que mentionnées ci-dessus. Toute utilisation non explicitement autorisée ci-avant (telle que par exemple, la modification du document ou son résumé) nécessite l'autorisation préalable et expresse des auteurs ou de leurs ayants droit.



Université de Liège

Faculté de Psychologie, de Logopédie et des Sciences de l'Éducation

Quel est l'impact de l'évaluation par les pairs sur l'amélioration des productions écrites des élèves du secondaire inférieur issus du qualifiant en langues et en sciences ?

Mémoire présenté par Bastin Louise et Guyot Léa

En vue de l'obtention du grade de Master en Sciences de l'Éducation

À finalité spécialisée en Enseignement

Promotrice : Fagnant Annick

Lecteurs : Monseur Christian

Superviseure : Goffin Christelle

Schillings Patricia

Année académique 2025-2026

Remerciements

Ce mémoire marque l'aboutissement de notre parcours de Master en sciences de l'éducation et représente le fruit de deux années de réflexion, de collaboration et de travail. Sa réalisation n'aurait pas été possible sans le soutien, l'accompagnement et l'investissement de nombreuses personnes que nous souhaitons sincèrement remercier.

Nous adressons tout d'abord nos remerciements à notre promotrice, Madame Fagnant, pour avoir encadré ce travail de recherche et pour les conseils et remarques qu'elle nous a apportés tout au long de sa réalisation.

Nous souhaitons exprimer notre gratitude à Madame Goffin, superviseure, pour son accompagnement, sa disponibilité, son écoute et la qualité de ses conseils. Son soutien, sa bienveillance et ses retours constructifs ont constitué une aide précieuse à chaque étape de ce mémoire.

Nous remercions également les lecteurs de ce mémoire, Monsieur Monseur et Madame Schillings, pour le temps consacré à sa lecture ainsi que pour l'intérêt porté à notre recherche.

Nous adressons également nos remerciements à la direction, aux enseignants et aux élèves de l'établissement scolaire qui ont accepté de prendre part à cette recherche. Leur accueil, leur disponibilité et leur implication tout au long du projet ont été essentiels à la réalisation de cette étude.

Enfin, nous souhaitons remercier chaleureusement nos familles, nos proches ainsi que toutes les personnes qui nous ont soutenues et encouragées tout au long de ces deux années. Leur présence, leur patience et leurs relectures attentives nous ont permis de mener ce projet à son terme.

Table des matières

Introduction	1
Revue de la littérature	3
1. Les fonctions principales de l'évaluation	3
1.1 Fonction sommative.....	3
1.2 Fonction formative	4
1.3 Fonction de soutien aux apprentissages	5
2. L'évaluation par les pairs	7
2.1 Définition.....	7
2.2 La mise en œuvre de l'évaluation par les pairs.....	8
2.2.1 Inscrire l'activité dans une démarche complète d'apprentissage	9
2.2.2 Former les élèves et construire des critères clairs	10
2.2.3 Travailler la qualité des feedbacks	11
2.3 Les avantages de l'évaluation par les pairs	13
2.4 Les limites de l'évaluation par les pairs	14
3. L'évaluation par les pairs dans chaque discipline	15
3.1 L'évaluation par les pairs dans l'enseignement des sciences.....	16
3.1.1 Le contexte de l'enseignement scientifique.....	16
3.1.2 Le rôle du feedback dans les apprentissages scientifiques	17
3.1.3 Les types de feedbacks en sciences	17
3.1.4 La mise en œuvre du dispositif.....	18
3.1.5 Avantages et limites	19
3.2 L'évaluation par les pairs dans l'enseignement des langues	20
3.2.1 Le contexte de l'enseignement linguistique	20
3.2.2 Le rôle du feedback dans les apprentissages linguistiques.....	21

3.2.3	Les types de feedbacks en langues.....	22
3.2.4	La mise en œuvre du dispositif.....	22
3.2.5	Avantages et limites	23
	Questions de recherche	24
	Méthodologie	27
1.	L'échantillon.....	27
2.	La mise en œuvre du dispositif	27
2.1	Structuration du dispositif et posture des enseignants	27
2.2	Préparation des élèves à l'évaluation par les pairs.....	28
2.3	Phases du dispositif d'évaluation par les pairs	28
3.	Les outils exploités	29
3.1	Les tâches	29
3.2	Les grilles de critères	30
4.	Types d'analyses	32
4.1	Analyse de l'évolution des productions.....	32
4.2.	Analyse de la qualité des feedbacks.....	33
4.3.	Analyse du lien entre la qualité des feedbacks et les améliorations	35
	Présentation et discussion des résultats	38
1.	Sous-question 1 :Quelles modifications sont observées entre les versions initiales et finales des productions écrites à la suite du processus d'évaluation par les pairs ?.....	38
1.1	Les modifications observées dans les productions de sciences.....	38
1.2	Les modifications observées dans les productions d'anglais	40
1.3	Analyse complémentaire de l'évolution des productions initiales en anglais ..	42
1.3.1	Objectif de l'étude	42
1.3.2	Présentation des résultats.....	42
1.4	Discussion à propos de la sous-question 1	43

2. Sous-question 2 : Quelles caractéristiques présentent les feedbacks formulés par les élèves lors de l'évaluation par les pairs ?	44
2.1 Les caractéristiques des feedbacks produits en sciences	45
2.2 Les caractéristiques des feedbacks produits en anglais	47
2.3 Discussion à propos de la sous-question 2	49
3. Sous-question 3 : Quels liens peut-on observer entre les feedbacks reçus et les modifications apportées aux productions écrites ?	51
3.1 Les liens entre les feedbacks et les modifications en sciences	51
3.2 Les liens entre les feedbacks et les modifications en anglais	53
3.3 Discussion à propos de la sous-question 3	55
4. Sous-question 4 : Ces résultats se manifestent-ils de manière différente dans les cours de sciences et d'anglais ?	56
Limites et perspectives.....	60
Conclusion.....	64
Bibliographie	66
Table des illustrations.....	69
Annexes.....	70
Annexe 1 – Trame.....	70
Annexe 2 – Tâches sciences.....	75
Annexe 3 – Tâches anglais	77
Annexe 4 - Grille de critères sciences.....	79
Annexe 5 – Grille de critères anglais	80
Annexe 6 – Outil d'analyse des modifications des productions	81
Annexe 7 – Outils d'analyse de la qualité des feedbacks.....	83
Annexe 8 – Outil d'analyse du lien entre les feedbacks et les modifications des productions	90
Annexe 9 : Outil d'analyse comparative des premiers jets en anglais	92

Annexe 10 : Tableau de synthèse : qualité des feedbacks.....	94
Annexe 11 : Tableau de synthèse : modifications des productions.....	95
Annexe 12 : Tableau de synthèse : lien entre feedbacks et modifications	97
Annexe 13 : Tableau de synthèse des résultats par discipline.....	99
Annexe 14 : Grille d'utilisation des intelligences artificielles génératives	106
Annexe 15 : Documents des apports individuels	108

Introduction

L'évaluation joue un rôle central dans l'enseignement. Son influence se fait sentir sur la manière dont les élèves apprennent, la façon dont les enseignants exercent, ainsi que sur les interactions au sein de l'école (Fagnant, 2023; Lafontaine & Toczec-Capelle, 2023). Depuis plusieurs années, la question de l'évaluation revient régulièrement dans les discussions pédagogiques, en particulier autour de la recherche de formats alternatifs aux évaluations classiques qui sont vues comme stressantes et peu utiles (Mottier-Lopez, 2023; Allal & Mottier-Lopez, 2005). L'évaluation conçue comme soutien d'apprentissage cherche à combiner les fonctions formatives et sommatives en plaçant l'élève au centre de son apprentissage (Nicol & Macfarlane-Dick, 2006). Comme le soulignent Girouard-Gagné et Durand (2025), l'évaluation devient un véritable levier d'apprentissage lorsqu'elle est pensée pour accompagner et guider l'élève dans la construction de ses connaissances. Parmi les dispositifs qui s'inscrivent dans cette perspective, figure l'évaluation par les pairs qui consiste à amener les élèves à porter un jugement sur les productions de leurs camarades à partir de critères explicites et à leur formuler un feedback susceptible de soutenir leurs apprentissages. Elle encourage les élèves à être actifs dans leurs apprentissages et à développer leur capacité à donner et recevoir du feedback (Bourgeois & Laveault, 2015; Reinholz, 2016).

Cependant, malgré son intérêt reconnu, l'évaluation par les pairs reste encore peu étudiée dans le contexte de la Fédération Wallonie-Bruxelles, surtout dans l'enseignement secondaire qualifiant (Mottier-Lopez, 2023). La plupart des recherches disponibles proviennent d'autres pays comme les Etats-Unis, le Royaume-Uni, le Canada ou la Suède (Anker-Hansen & Andrée, 2019; Ketonen et al., 2020). Il est dès lors difficile de savoir comment cette méthode est réellement mise en œuvre dans nos classes. C'est pourquoi, nous avons choisi de concevoir et de mettre en place ce dispositif au sein de notre propre contexte scolaire afin d'en explorer les modalités et les effets.

Notre étude porte sur l'évaluation par les pairs dans le secondaire inférieur qualifiant, et ce, dans deux disciplines différentes : en anglais et en sciences. Même si ces matières ont des approches pédagogiques différentes et des objectifs d'apprentissage spécifiques, il est possible d'y appliquer des méthodes d'évaluation similaires (Boilevin & Brandt-Pomares, 2011; Lou et al., 2016). Notre objectif est de comparer l'impact du feedback par les pairs sur l'amélioration des productions écrites dans ces deux contextes. Nous avons formulé la question de recherche

suivante : Quel est l'impact de l'évaluation par les pairs sur l'amélioration des productions écrites des élèves du secondaire inférieur issus du qualifiant en langues et en sciences ?

Ce thème de mémoire n'a pas été choisi au hasard. L'évaluation est un sujet qui revient souvent dans nos discussions professionnelles et soulève de nombreuses questions : Pourquoi évaluons-nous ? Pour les élèves, pour les enseignants ou pour les parents ? Que voulons-nous réellement mesurer ? (Fagnant, 2023; Lou et al., 2016). Ces réflexions ont éveillé chez nous un réel intérêt pour ce sujet. L'évaluation par les pairs nous semble prometteuse car elle favorise l'autonomie des élèves, leur sens critique et leur capacité à réfléchir sur leurs propres apprentissages (Philippakos, 2017; Reinholz & Dounas-Frazer, 2016).

Dans ce mémoire, nous commencerons par poser le cadre théorique nécessaire à la compréhension de notre recherche. Nous aborderons notamment les différentes fonctions de l'évaluation ainsi que les spécificités de l'évaluation et du feedback par les pairs. Nous présenterons ensuite la démarche de recherche adoptée, en précisant la question qui guide notre travail ainsi que les choix méthodologiques effectués. Enfin, nous exposerons les résultats obtenus. Ceux-ci feront l'objet d'une analyse et d'une mise en perspective avec les apports de la littérature. Cela nous permettra de mieux comprendre l'impact de l'évaluation par les pairs sur les apprentissages des élèves, notamment à travers une comparaison entre les contextes d'enseignement en langues et en sciences.

Revue de la littérature

Dans cette partie, nous présentons les fondements théoriques qui sous-tendent notre recherche. Nous aborderons tout d'abord les différentes fonctions de l'évaluation en distinguant les approches sommatives, formatives et celles qui visent spécifiquement à soutenir les apprentissages. Nous nous intéresserons ensuite à l'évaluation par les pairs en présentant ses caractéristiques, ses avantages, ses limites ainsi que les conditions favorisant sa mise en œuvre. Une attention particulière sera accordée au feedback, considéré comme un élément central de ce dispositif et comme un levier essentiel de régulation et d'amélioration des apprentissages. Enfin, nous nous pencherons sur les spécificités de l'évaluation par les pairs dans les disciplines investiguées, en mettant en évidence les particularités de sa mise en œuvre dans les cours de sciences et d'anglais.

1. Les fonctions principales de l'évaluation

L'enseignement, l'apprentissage et l'évaluation forment un tout cohérent où l'évaluation occupe une place importante. Cette dernière permet de guider les apprentissages et de soutenir les choix de l'enseignant (Lafontaine et Toczec-Capelle, 2023).

L'évaluation peut être définie comme « une activité de recueil de diverses informations considérées comme pertinentes par rapport à des objectifs visés, en vue de porter un jugement et de prendre une décision » (Amigues et Zerbato-Poudou, 1996 cités par Mottier-Lopez, 2023, p.13). Cette définition attire l'attention sur les décisions que l'évaluation permet de prendre.

Dès la fin des années 1960, Scriven et Bloom (cités par Lafontaine et Toczec-Capelle, 2023) ont proposé de distinguer les deux fonctions principales de l'évaluation : sommative et formative. Cette dernière, dont la définition s'est élargie au fil des années, dépasse aujourd'hui ce cadre et s'inscrit dans une dynamique continue de soutien aux apprentissages. Nous reviendrons plus en détails sur cette conception élargie de l'évaluation formative au point 1.2.

1.1 Fonction sommative

Parmi les différentes fonctions d'évaluation, la fonction sommative est la plus utilisée à l'école.

En effet, elle est utilisée pour faire le point sur ce qu'un élève a appris à la fin d'une période de formation, que ce soit un semestre, une année ou un cycle. Elle permet à l'enseignant de vérifier si les objectifs ont été atteints et donne généralement une note, chiffrée ou non (Allal, 2008 cité par Fagnant, 2023).

Autrement dit, l'évaluation sommative peut se définir comme « un bilan certifiant les compétences et les connaissances acquises par l'apprenant à la fin d'une étape intermédiaire [...] ou d'une étape finale [...] d'un cursus de formation » (Allal, 2008 cité par Mottier-Lopez, 2023, p.312).

Mais l'évaluation sommative ne sert pas uniquement à faire des bilans. Elle peut aussi prendre d'autres rôles comme celui de classer les élèves, de les orienter dans des filières, voire de limiter leur accès à certaines formations. Ce système de sélection est d'ailleurs fort présent dans notre système scolaire belge. Selon Crahay (2012, cité par Mottier-Lopez, 2023), dans notre système, l'évaluation sert surtout à départager les élèves et ne garantit pas à tous les mêmes chances de réussite.

1.2 Fonction formative

Contrairement à la fonction sommative, la fonction formative s'exerce durant le temps des apprentissages afin de soutenir la progression de chaque élève par une régulation constante des pratiques de l'enseignant et des stratégies de l'apprenant. Comme le souligne Scallon (1988, cité par Lafontaine et Toczec-Capelle, 2023, p.11), cette démarche se définit comme un « processus d'évaluation continue ayant pour objectif d'assurer la progression de chaque individu dans une démarche d'apprentissage, avec l'intention de modifier la situation d'apprentissage ou le rythme de cette progression, pour apporter (s'il y a lieu) des améliorations ou des correctifs appropriés ».

Dès lors, cette modalité ne se limite pas à une simple collecte de données mais constitue une aide réelle visant à adapter les activités pédagogiques aux besoins spécifiques du groupe (Allal, 2008 cité par Lafontaine et Toczec-Capelle, 2023). En ce sens, elle assure une fonction diagnostique tout en favorisant le développement de l'autonomie chez l'élève par l'acquisition de compétences en autoévaluation (Sadler, 1989 ; Scallon, 2004, cités par Bourgeois et Laveault, 2015).

Cette perspective, initialement définie par Bloom, a été élargie par plusieurs auteurs (Allal, 1979 ; Perrenoud, 1998, cités par Allal et Mottier Lopez, 2005) qui envisagent désormais l'évaluation comme une composante intrinsèque des situations d'apprentissage. Elle s'éloigne des tests formels pour mobiliser des outils diversifiés, transformant l'élève en acteur de son propre parcours pendant que l'enseignant module son enseignement. Audibert (1980, cité par Allal et Mottier Lopez, 2005, p.268) illustre cette dynamique en affirmant que l'évaluation formative « se réalise au jour le jour et permet à l'enseignant et à l'élève d'adapter leur action

respective à la situation d'enseignement et d'apprentissage en cause. C'est alors pour eux une occasion privilégiée de prise de conscience de leur vécu, de l'objectivation en acte ».

L'efficacité de ces pratiques sur les acquis scolaires a été démontrée par les travaux de Black, Harrison et Lee (2001) ainsi que ceux de Black et William (1998, cités par Reinholz, 2016). Cet impact positif repose sur la pertinence des retours d'information et la capacité de l'enseignant à réorienter son action pour encourager la régulation des apprentissages. Dans le quotidien de la classe, cette démarche peut prendre des aspects formels (grilles d'évaluation) ou informels (observations directes).

Elle s'apparente alors à un dialogue constant entre les acteurs de la classe (Mottier Lopez, 2023, cité par Lafontaine et Toczec-Capelle, 2023). Plusieurs chercheurs soutiennent cette dimension communicationnelle : Cardinet (1988) l'appréhende comme un échange sur les objectifs et les obstacles, Ouellette (1990) comme une construction issue des interactions éducatives, et Hadji (1989) comme un système de transactions entre les attentes et les réactions de chacun vis-à-vis des résultats (cités par Allal et Mottier Lopez, 2005).

Enfin, Del'Guidice (1999, cité par Allal et Mottier Lopez, 2005) met en évidence que l'intégration de l'évaluation formative, dans les situations d'enseignement, produit non seulement des effets bénéfiques sur les apprentissages immédiats mais favorise aussi leur transfert vers d'autres contextes. Ce constat rejoint celui de Black et William (1998, cités par Reinholz, 2016) quant à l'efficacité de ces pratiques sur le long terme. Il convient de rappeler que cette approche, loin d'être réservée aux seuls élèves en difficulté, s'inscrit dans une logique d'équité pour la réussite de tous (Mottier Lopez, 2023 cité par Lafontaine et Toczec-Capelle, 2023). En tant que levier dynamique et ajustable, l'évaluation formative s'affirme comme un pilier essentiel d'un enseignement au service de l'élève.

1.3 Fonction de soutien aux apprentissages

Comme expliqué précédemment, pendant longtemps, l'évaluation a été divisée entre l'évaluation formative et l'évaluation sommative. Cependant cette classification a ses limites. De nombreux chercheurs comme Allal (2008), De Ketele (2010) ou Hadji (1989, cités par Mottier-Lopez, 2023) ont critiqué cette vision car elle ne permet pas de saisir pleinement la richesse de l'évaluation. Bennet (2011, cité par Lafontaine et Toczec-Capelle, 2023) renforce cette idée en disant que ces deux fonctions sont souvent liées dans la réalité. Selon lui, vouloir les séparer n'a pas beaucoup de sens. En effet, même une évaluation considérée comme

sommative peut donner des informations sur les difficultés des élèves. Ces données aident alors l'enseignant à ajuster son enseignement pour mieux répondre aux besoins des étudiants.

Face à ce constat, un courant de recherche s'est développé.

Depuis les années 1990, des auteurs comme Black et William (1998), Crahay, Mottier-Lopez et Marcoux (2019) ou encore Laveault et Allal (2016, cités par Lafontaine et Toczec-Capelle, 2023) ont cherché une conception de l'évaluation plus intégrative.

C'est suite à cela qu'arrive la notion d'évaluation soutien d'apprentissage au début des années 2010. Laveault et Allal (2016, cités par Lafontaine et Toczec-Capelle, 2023, p.14) la définissent comme « la collecte et l'interprétation d'informations issues d'évaluations dont l'utilisation intentionnelle permet aux enseignants et aux élèves, qu'ils agissent individuellement ou de manière interactive, de prendre des décisions qui ont un impact positif sur l'enseignement et l'apprentissage ». Autrement dit, ce qui compte, c'est que l'évaluation soit utilisée pour soutenir les apprentissages.

Suite au changement de perspective entre la vision classique de Bloom et la vision élargie, il y a eu des conséquences sur les rôles dans la classe. Alors que dans la vision classique l'enseignant est seul à piloter l'évaluation, la vision élargie insiste sur la co-construction du processus évaluatif. L'enseignant garde un rôle de guide mais il soutient l'élève dans le développement de ses capacités d'autorégulation, c'est-à-dire sa capacité à planifier, surveiller et ajuster ses propres stratégies d'apprentissage en fonction des objectifs visés (Nicol & Macfarlane-Dick, 2006). Cela passe par des démarches comme l'autoévaluation, l'évaluation par les pairs ou la co-évaluation où enseignant et élèves discutent ensemble des critères, des résultats et des pistes d'améliorations (Allal, 1999 ; Nunziati, 1990 ; Vial, 1995 cités par Allal et Mottier Lopez, 2005).

Cependant, faire évoluer les pratiques d'évaluation ne se fait pas si facilement. Les recherches montrent qu'il existe parfois un écart entre les instruments proposés et les pratiques réelles en classe (Allal et Mottier Lopez 2005). De fait, certains enseignants mixent inconsciemment les différentes fonctions de l'évaluation (Laveault et Allal, 2016 cités par Fagnant, 2023). Les plus expérimentés parviennent à combiner les fonctions formative et sommative de manière fluide. Par contre, beaucoup d'enseignants croient toujours que l'évaluation a un seul but de certification et d'orientation. Beaucoup pensent impossible de mélanger les deux fonctions (Abernot, 2013 ; Black et William, 1998 ; Mottier-Lopez, 2015 ; Perrenoud, 1991 cités par Mottier-Lopez, 2023). Pourtant, les recherches montrent que lorsque les évaluations sont

pensées comme soutien aux apprentissages, leurs effets sont particulièrement bénéfiques, surtout pour les élèves en difficultés (Black et William, 1998 cités par Fagnant, 2023).

Il existe des éléments clés pour rendre l'évaluation réellement efficace. Par exemple, il faut expliquer clairement les objectifs d'apprentissage, fournir des feedbacks précis, éviter la comparaison entre élèves et leur donner l'opportunité d'utiliser leurs feedbacks pour avancer. L'élève ne doit pas être passif mais être au contraire intégré activement au processus d'évaluation. Cela peut se faire via des moments de réflexion sur ses propres productions, via des échanges avec les pairs ou via des dialogues avec l'enseignant. Ces interactions permettent à l'élève de se rendre compte de son niveau actuel et de trouver des moyens pour atteindre le niveau visé (Black et William, 1998 cités par Fagnant, 2023).

En conclusion, la fonction de soutien aux apprentissages propose une vision intégrée de l'évaluation où l'évaluation formative alimente l'évaluation sommative, et inversement. Cette perspective est de plus en plus présente aujourd'hui. Cependant, mettre en œuvre une telle évaluation intégrée demande un changement en profondeur. Il faut prendre en compte la forme de l'évaluation mais aussi les croyances des enseignants et des élèves (Fagnant, 2023). Cela demande donc encore du temps avant que l'évaluation soutien d'apprentissage ne soit présente dans les pratiques quotidiennes des enseignants. Parmi ces pratiques à encourager, l'évaluation par les pairs occupe une place importante. Elle permet aux élèves de s'impliquer activement dans le processus d'évaluation en donnant et recevant des feedbacks sur les productions. Cette démarche, lorsqu'elle est bien encadrée, favorise la régulation des apprentissages mais aussi le développement de compétences métacognitives, c'est-à-dire la capacité des élèves à réfléchir sur leurs propres apprentissages et à prendre du recul sur leurs productions notamment lorsqu'ils analysent celles de leurs pairs (Bourgeois & Laveault, 2015) ; cette démarche favorise également le développement des compétences sociales. Elle s'inscrit donc pleinement dans la logique de l'évaluation soutien d'apprentissage.

2. L'évaluation par les pairs

2.1 Définition

L'évaluation par les pairs se définit comme une pratique où des élèves sont amenés à juger la qualité des productions ou des performances de leurs camarades (Topping, 1988 cité par Marion et al., 2022). Cette méthode s'applique à divers contextes scolaires allant de travaux écrits à des présentations orales ou encore des projets collaboratifs (Karami et Rezaei, 2015 cités par Mumpuni et al., 2022). Cette approche privilégie une perspective d'évaluation plaçant l'élève

comme acteur de son propre apprentissage. Elle s'inscrit plus spécifiquement dans une logique formative visant à soutenir et réguler les apprentissages en cours (Allal et Mottier-Lopez, 2005).

Dans cette perspective, l'évaluation ne se limite pas à porter un jugement sur les productions mais vise à fournir des informations permettant aux élèves d'ajuster leurs apprentissages en cours de réalisation, dans une logique de régulation (Allal et Mottier-Lopez, 2005).

Plus qu'un simple jugement de valeurs, l'évaluation par les pairs est une dynamique d'apprentissage mutuel. Elle met l'accent sur la co-construction des savoirs étant donné que les élèves développent en duo ou en groupe des compétences métacognitives, à savoir la capacité des élèves à réfléchir sur leurs propres apprentissages, et adopter une posture critique. Plusieurs travaux confirment d'ailleurs que l'interaction est un levier fondamental dans l'efficacité de la rétroaction entre pairs (Kasch et al., 2021 ; Ge, 2019 ; Sanchez-Marti et al., 2019 ; Roy et Michaud, 2018 ; Steen-Utheim et Wittek, 2017 ; Van Gennip et al., 2010 ; Yang, 2016 cités par Marion et al., 2022).

Dans cette perspective, l'évaluation entre pairs peut être considérée comme un espace de régulation des apprentissages où les échanges contribuent à ajuster progressivement les productions des élèves (Allal et Mottier-Lopez, 2005).

Ce processus de régulation implique que les élèves soient capables d'interpréter les informations reçues et de les transformer en actions d'amélioration, ce qui constitue un enjeu central dans une perspective d'évaluation pour apprendre (Mottier-Lopez, 2023).

2.2 La mise en œuvre de l'évaluation par les pairs

Comme nous l'avons souligné, l'évaluation par les pairs offre des bénéfices considérables pour l'apprentissage, à condition d'être menée avec rigueur. En revanche, une préparation insuffisante peut engendrer des biais ou des retours d'information peu constructifs (Mumpuni et al., 2022).

Pour garantir la pertinence de ce dispositif, plusieurs facteurs cruciaux doivent être coordonnés, notamment la préparation des élèves, l'organisation logistique et la qualité des échanges (Brookhart, 2015 cité par Mumpuni et al., 2022). Le développement de l'apprenant à une réelle autonomie évaluative ne s'improvise pas. Elle exige un encadrement structuré de l'enseignant, dont le rôle est de réunir les leviers nécessaires à une pratique fructueuse (Stefanou et al., 2004 cités par Bourgeois et Laveault, 2015).

Dans cette optique, Sluijmans (2002) et Topping (2009, cités par Bourgeois et Laveault, 2015) identifient des prérequis spécifiques à développer chez les élèves pour transformer cette activité en un outil pédagogique utile. Les élèves doivent apprendre à s'approprier les objectifs et les critères d'apprentissage, à porter un jugement de valeur pertinent sur un travail et, surtout, à formuler des pistes de progression permettant de réduire l'écart entre leur production actuelle et les attentes visées.

Ces exigences s'inscrivent dans une conception de l'évaluation où l'élève devient un véritable acteur capable de décrypter les critères et d'interpréter les feedbacks pour les réinvestir dans une logique de régulation de son propre parcours (Allal et Mottier-Lopez, 2005 ; Mottier-Lopez, 2023).

2.2.1 Inscrire l'activité dans une démarche complète d'apprentissage

Pour que l'évaluation par les pairs soit vraiment efficace, elle doit s'intégrer dans un cycle d'apprentissage structuré où chaque étape contribue à l'amélioration des apprentissages.

Kollar et Fischer (2010, cités par Reinholz, 2016) proposent un modèle en quatre étapes qui permet de mieux comprendre comment organiser ce processus :

- Task engagement : les élèves réalisent une tâche (écriture, résolution de problème, ...).
- Feedback provision : un pair (ou plusieurs) fournit des retours sur cette tâche.
- Feedback réception : l'élève prend connaissance du retour, le traite et cherche à le comprendre.
- Révision : il réinvestit ce feedback pour améliorer sa production.

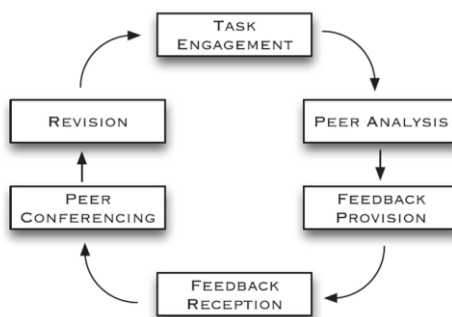


Figure 1. Modèle du cycle de l'évaluation par les pairs (Kollar et Fischer, 2010, cités par Reinholz, 2016).

Ce schéma montre que le feedback entre pairs ne se résume pas à un simple échange de commentaires. Au contraire, il s'agit d'une démarche structurée et complète de régulation des apprentissages.

Le cycle permet de passer d'une évaluation passive à une vraie dynamique d'apprentissage actif. L'élève devient acteur de son parcours, apprend à interpréter les retours et à ajuster ses actions. Ce processus contribue aussi au développement de l'autorégulation, de la pensée critique et de la collaboration qui sont des compétences centrales dans la perspective de l'évaluation soutien d'apprentissage.

L'étape de révision, en particulier, constitue un moment clé de ce processus. Elle ne consiste pas seulement à appliquer les commentaires reçus mais suppose que l'élève interprète ces retours, évalue leur pertinence et décide comment les intégrer dans sa production. Cela demande à l'élève de prendre du recul sur son texte. Cette « décentration » est essentielle pour que l'élève puisse améliorer efficacement son texte à partir des retours reçus (Scardamalia et Bereiter, 1998 cités par Bourgeois et Laveault, 2015).

Dans cette démarche, l'enseignant joue un rôle central dans ce cycle en guidant chaque étape, en assurant un climat de confiance et en surveillant la qualité des échanges. Des moments collectifs peuvent aussi être prévus pour discuter des difficultés rencontrées ou des freins éventuels (Huisman et al., 2018 ; Robinson et al., 2013 cités par Marion et al., 2022).

En conclusion de ces divers conseils, l'enseignant a donc un rôle central dans la réussite de l'évaluation par les pairs. Il doit prévoir des moments de formation, fournir des grilles de critères clairs, encourager une posture réflexive, créer un climat sécurisant et valoriser les rétroactions de qualité. Il peut aussi proposer des temps collectifs. Ces éléments sont essentiels pour renforcer la motivation des élèves et instaurer des habitudes efficaces en termes de feedbacks.

2.2.2 Former les élèves et construire des critères clairs

La formation des élèves est un élément indispensable. Elle améliore la qualité des feedbacks et renforce le sentiment de compétences des élèves (Sluismans et al., 2001 ; Panadero et al., 2016 ; Brookhart, 2015 cités par Mumpuni et al., 2022). Cette formation peut prendre la forme de discussions autour des objectifs et critères ou encore d'activités ciblées telles que l'analyse de copies types ou la co-construction des critères d'évaluation avec les élèves (Eiluius et al., 2019 ; Patchan et al., 2016 cités par Marion et al., 2022 ; Bourgeois et Laveault, 2015). En associant les élèves à ces démarches, on favorise leur appropriation des critères et on développe

leur capacité à juger la qualité d'un texte (Sadler, 1989 ; Orsmond et al., 2002 cités par Leenknecht et Prins, 2017).

Par ailleurs, ce travail sur les critères prépare aussi les élèves à recevoir des rétroactions et à les mobiliser pour réviser leurs textes. Comprendre ce qui est attendu et être capable de proposer des pistes d'améliorations, c'est aussi se donner les moyens de transformer les commentaires reçus en leviers de progrès (Sluijsmans, 2004 ; Topping, 2009 cités par Bourgeois et Laveault, 2015).

Enfin, la répétition de ces entraînements est une condition nécessaire pour installer durablement ces compétences. En multipliant les occasions d'évaluer et d'être évalué, les élèves enrichissent leur expérience, affinent leur jugement et deviennent progressivement plus compétents dans cet exercice complexe (Cui et al., 2021 cités par Reinholz, 2016).

2.2.3 Travailler la qualité des feedbacks

Les feedbacks jouent un rôle clé dans le cadre de l'évaluation soutien d'apprentissage. En effet, ils sont une composante essentielle car ils permettent la régulation des apprentissages.

2.2.3.1 Définition

Le feedback est souvent défini comme une information fournie à quelqu'un à propos de sa performance que ce soit par un enseignant, un parent ou même un ordinateur (Hattie et Timperley, 2007 cités par Lafontaine et Toczec-Capelle, 2023).

Si cette définition paraît simple, elle ne montre pas réellement la complexité et la portée des feedbacks. En effet, comme l'a montré Dweck (2002, cité par Lafontaine et Toczec-Capelle, 2023), les retours ne se limitent pas à corriger ou améliorer une performance. Ils influencent aussi profondément la perception que les élèves ont d'eux-mêmes ainsi que leur motivation. Cette influence dépasse donc le cadre de l'apprentissage strict et touche des dimensions affectives et identitaires (Calone et Lafontaine, 2018 cités par Lafontaine et Toczec-Capelle, 2023).

2.2.3.2 Directions du feedback

Le feedback est, de manière large, une double boucle de rétroaction entre deux personnes. On retrouve trois types de direction des feedbacks : de l'enseignant vers l'élève, de l'élève vers l'enseignant et entre élèves.

La rétroaction par les pairs place l'élève dans un rôle actif de régulateur. Plusieurs recherches en soulignent les bénéfices cognitifs et métacognitifs à condition qu'elle soit bien encadrée.

Elle encourage notamment le développement de l'esprit critique, une compréhension approfondie des contenus et un recul sur sa propre production (Marion et al., 2022 ; Reinholz, 2016).

Cependant, l'efficacité de cette rétroaction dépend fortement des conditions de mise en œuvre. Sans consignes explicites ou formation préalable, les élèves reproduisent les schémas d'interaction traditionnels orientés de l'enseignant vers l'élève (Bourgeois et Laveault, 2015). Il est alors nécessaire de les sensibiliser à l'intérêt et au fonctionnement de l'évaluation par les pairs pour en exploiter tout le potentiel.

C'est à partir de ce constat que notre mémoire se concentre sur cette forme de feedback.

2.2.3.3 Mise en œuvre des feedbacks

Pour que le feedback puisse pleinement jouer son rôle dans l'amélioration des apprentissages, plusieurs recommandations pratiques sont à suivre.

Tout d'abord, il est essentiel que le feedback mette en évidence les attentes de l'enseignant, les éléments réussis ou à améliorer et les moyens pour y parvenir (Lafontaine et Toczec-Capelle, 2023).

Ensuite, le feedback doit intervenir rapidement après l'activité pour pouvoir être exploité par l'élève (Gibbs et Simpson, 2004 cités par Lafontaine et Toczec-Capelle, 2023). De plus, les recherches soulignent que l'efficacité du feedback dépend fortement de la régularité. Une pratique occasionnelle, même si elle est bien faite, ne produira que des effets limités (Lafontaine et Toczec-Capelle, 2023).

Un feedback efficace doit également être clair, limité aux critères définis et personnalisé (Nicol et Macfarlane-Dick, 2006 ; Galand, 2006 cités Lafontaine et Toczec-Capelle, 2023). Il faut donc favoriser des retours sur des critères précis plutôt que sur l'élève. Cette exigence est d'autant plus importante lorsqu'il s'agit de retours entre pairs liés d'amitié. Ils sont souvent influencés par leurs liens et formulent alors des commentaires vagues ou centrés sur des éléments négligeables (Nilson, 2003 cité par Bourgeois et Laveault, 2015).

Le feedback gagne en impact lorsqu'il est perçu comme utile par l'élève. Enseigner explicitement l'utilité du feedback aux élèves peut donc être pertinent (Lafontaine et Toczec-Capelle, 2023).

Enfin, les feedbacks prennent tout leur sens lorsqu'ils s'inscrivent dans une évaluation pensée comme un soutien aux apprentissages, donc dans une démarche où l'erreur est reconnue comme

opportunité d'apprentissage et où les retours servent à guider les progrès (Lafontaine et Toczek-Capelle, 2023). C'est pourquoi l'intégration du feedback dans une logique d'évaluation formative ou d'évaluation soutien d'apprentissage est aujourd'hui considérée comme un levier indispensable pour améliorer les apprentissages (Black et William, 2009 ; William, 2011 cités par Leenknecht et Prins, 2017).

2.3 Les avantages de l'évaluation par les pairs

L'efficacité de l'évaluation par les pairs est confirmée par plusieurs méta-analyses qui montrent que l'évaluation par les pairs est plus bénéfique que celle réalisée uniquement par l'enseignant (Double et al., 2020 ; Li et al., 2020 ; Sanchez et al., 2017 cités par Fagnant, 2023). En plus, cette approche présente de nombreux autres avantages.

En effet, un premier avantage de cette pratique est l'implication active des élèves dans leur apprentissage. Quand les élèves évaluent le travail de leurs camarades, ils ne sont plus de simples apprenants mais deviennent des observateurs et conseillers. Cela les aide à repérer ce qu'ils maîtrisent bien et ce qui peut être amélioré, autant chez les autres que dans leur propre travail. En formulant des commentaires, ils apprennent à réfléchir et adopter une posture réflexive. Cela les pousse aussi à améliorer leurs compétences métacognitives (Topping, 2009 ; Bloxham et West, 2004 ; Orsmond et al., 2002 cités par Mumupuni et al., 2022). À ce sujet, Garcia-Debanc (1986, cité par Bourgeois et Laveault, 2015) souligne que si les élèves ont souvent du mal à repérer les défauts de leurs propres productions, ils éprouvent généralement plus de facilités à les identifier dans celles de leurs pairs. L'évaluation par les pairs devient alors l'occasion de prendre du recul sur ses propres écrits à travers l'analyse du texte de quelqu'un d'autre. Ainsi, elle bénéficie non seulement à l'élève évalué qui reçoit des retours utiles pour améliorer son travail, mais aussi à l'élève évaluateur qui développe plus finement ses propres compétences d'écriture.

Les bénéfices ne s'arrêtent pas seulement au développement cognitif. De fait, l'évaluation par les pairs favorise aussi les compétences sociales. Les élèves apprennent à s'exprimer avec respect, à formuler des critiques bienveillantes, à écouter les conseils et à en tirer parti. Ce type d'interaction améliore leur communication, leur collaboration et renforce leur sentiment d'empathie envers les autres (Topping et al., 2000 cités par Leenknecht et Prins, 2017). Tout cela permet de réduire les stéréotypes et les élèves sont plus motivés à coopérer (Harrison et al., 2015 ; Karami et Rezaei, 2015 cités par Mumupuni et al., 2022). Cela renforce aussi leur confiance en eux (Ion et al., 2017 ; Van Popta et al., 2017 cités par Marion et al., 2022).

Pour l'enseignant aussi, cette pratique est un avantage. Elle permet de diversifier les sources de feedback, de mieux cerner les représentations des élèves et d'alléger sa charge de correction (Rico-Jun et al., 2019 ; Wu et Schunn, 2021 cités par Marion et al., 2022). Mais ce n'est pas uniquement un outil de délégation. Cela devient une véritable ressource qui soutient l'apprentissage et l'autonomie des élèves. L'enseignant peut alors se concentrer davantage sur le suivi et l'accompagnement des étudiants.

Dans une perspective d'évaluation formative, ces effets s'inscrivent dans une logique où l'évaluation soutient directement les apprentissages en cours et permet aux élèves d'ajuster progressivement leurs productions (Allal et Mottier-Lopez, 2005; Mottier-Lopez, 2023).

En conclusion, l'évaluation par les pairs ne se limite pas à un échange de commentaires. Elle offre une véritable dynamique d'apprentissage collectif qui permet aux élèves de progresser. Elle développe à la fois les compétences cognitive, sociale, réflexive et communicationnelle.

2.4 Les limites de l'évaluation par les pairs

Comme expliqué précédemment, l'évaluation par les pairs a de nombreux avantages. Malgré tout, elle présente aussi un certain nombre de limites.

Tout d'abord, la qualité des feedbacks est un premier frein. De nombreux élèves n'ont pas les compétences nécessaires pour fournir des commentaires pertinents (Carless et Boud, 2018 ; Double et al., 2020 ; Bader et al., 2019 cités par Marion et al., 2022). Le manque de confiance en soi ou en ses pairs augmente cette difficulté. Certains élèves se sentent non qualifiés pour évaluer (Cheng et Warren, 1997 ; Falchikov, 1995 ; McDowell, 1995 ; Rust et al., 2003 cités par Leenknecht et Prins, 2017).

Parfois aussi, les conseils donnés sont incompris ou ignorés par le destinataire, ce qui limite leur utilité (Cui et Schunn, 2024). Les élèves doutent de la validité des jugements reçus (Gaynor, 2020 ; Lira-Gonzalez, 2021 ; Wu et Schunn, 2021 cités par Marion et al., 2022). Dans une perspective d'évaluation formative, cette difficulté est problématique car le feedback ne joue pleinement son rôle que s'il permet à l'élève de réguler ses apprentissages (Allal et Mottier-Lopez, 2005). Or, cette régulation suppose que l'élève soit capable de comprendre, interpréter et utiliser les informations reçues, ce qui n'est pas toujours acquis et constitue une limite importante du dispositif (Mottier-Lopez, 2023).

Ensuite, le manque de clarté ou de transparence des critères d'évaluation peut nuire à la qualité de l'évaluation effectuée (Carless, 2006 ; Papinczak et al., 2007 cités par Leenknecht et Prins, 2017). Il est courant que l'enseignant fournisse aux élèves une grille d'évaluation descriptive

pour guider leur analyse que ce soit sur leur propre travail ou celui d'un pair. Cependant, cette pratique peut être contre-productive si les critères sont formulés de manière trop abstraite. Cela va rendre leur compréhension difficile et peu significative pour les élèves (Bourgeois et Laveault, 2015). Par ailleurs, ceux-ci peuvent percevoir ces critères comme incohérents ou mal adaptés, ce qui nuit à la qualité du feedback (Mumupuni et al., 2022).

Ce manque de clarté limite également la capacité des élèves à mobiliser les feedbacks dans une logique de régulation. En effet, sans une compréhension explicite des critères et des attentes, les élèves ne parviennent pas à transformer les feedbacks en améliorations concrètes de leurs productions (Allal et Mottier-Lopez, 2005).

Les dynamiques relationnelles jouent aussi un rôle. La rétroaction peut prendre la forme d'un monologue plutôt que d'un échange constructif (Ajiawi et Boud, 2017 ; Er et al., 2021 cités par Marion et al, 2022).

Au niveau du vocabulaire, l'usage de mots agressifs et négatifs peut impacter l'estime de soi des apprenants et freiner leur engagement (Boud, 2000 cité par Leenknecht et Prins, 2017). Il est donc préférable de privilégier des conseils et explications plutôt que des jugements directs, sans pour autant tomber dans des biais de notations où l'élève a l'envie de préserver l'harmonie en évitant des jugements sévères (Mumpuni et al., 2022).

Enfin, il faut veiller à ce que les élèves perçoivent l'enjeu de la rétroaction, sinon ils risquent de se limiter à des commentaires courts et peu exigeants intellectuellement (Chi et Wylie, 2014 cités par Marion et al., 2022).

En conclusion, si l'évaluation par les pairs présente beaucoup d'opportunités d'apprentissage, elle a besoin d'une préparation rigoureuse, d'une formation explicite auprès des élèves ainsi que d'un cadre clair pour garantir la validité et la productivité des échanges.

3. L'évaluation par les pairs dans chaque discipline

Après avoir présenté les fondements théoriques de l'évaluation par les pairs, cette partie se penche sur sa mise en œuvre et ses effets dans le cadre des cours de langues et de sciences.

Afin de permettre une comparaison rigoureuse entre les deux disciplines, l'analyse est structurée en cinq volets similaires : le contexte disciplinaire, le rôle du feedback, les types de feedbacks, la mise en œuvre du dispositif et les bénéfices et limites.

3.1 L'évaluation par les pairs dans l'enseignement des sciences

L'évaluation par les pairs, bien qu'applicable à de nombreuses disciplines, revêt une pertinence particulière dans l'enseignement des sciences, notamment lors d'activités d'investigation et de rédaction scientifique. En effet, l'écriture scientifique ne se limite pas à la restitution de connaissances : elle occupe une place centrale dans l'apprentissage des sciences, en amenant les élèves à expliciter leur démarche, à relier des données à des hypothèses et à élaborer une argumentation cohérente.

Cette partie du travail se penche donc sur ce que l'évaluation par les pairs peut apporter au développement du raisonnement scientifique et des compétences d'investigation, en s'appuyant sur les particularités de la manière dont les sciences sont construites et enseignées.

3.1.1 Le contexte de l'enseignement scientifique

Les travaux de Boilevin et Brandt-Pomares (2011) montrent que l'introduction des démarches d'investigation transforme en profondeur les pratiques d'enseignement en sciences. Il ne s'agit plus seulement de transmettre des savoirs stabilisés, mais de placer les élèves dans des situations où ils doivent problématiser, formuler des hypothèses, élaborer des protocoles et interpréter des résultats.

Dans cette logique, l'activité scientifique scolaire devient structurée autour d'une chaîne argumentative : question – hypothèse – protocole – données – interprétation – conclusion (Boilevin & Brandt-Pomares, 2011). L'apprentissage repose donc sur la capacité à expliciter les liens logiques entre ces différentes composantes.

Dans cette perspective, l'écriture scientifique ne consiste pas à relater une manipulation ou une simple restitution de résultats. Elle est un outil de structuration de la pensée et d'explicitation des démarches. Elle sert à construire un raisonnement articulé. L'écriture devient un instrument de structuration cognitive.

Dans cette optique, l'évaluation par les pairs trouve un terrain particulièrement favorable en sciences : les productions écrites donnent accès aux raisonnements des élèves et rendent visibles leurs modes de pensée. L'évaluation peut ainsi porter sur la cohérence du processus intellectuel ayant conduit à ce résultat.

La discipline scientifique se prête donc particulièrement bien à un dispositif où les élèves sont amenés à analyser, discuter et améliorer des raisonnements explicites.

3.1.2 Le rôle du feedback dans les apprentissages scientifiques

Dans une perspective formative, le feedback vise à réduire l'écart entre la production de l'élève et les attentes disciplinaires. Montpied (2016) et ses collègues soulignent que l'évaluation formative doit être intégrée à la conception même des séquences d'enseignement. Dans le cadre de démarches d'investigation, cette intégration prend tout son sens, puisque les activités expérimentales produisent des traces écrites exploitables pour des régulations successives.

Ainsi, le feedback n'intervient pas seulement pour corriger, mais pour orienter l'activité cognitive de l'élève.

En sciences, le feedback joue plusieurs rôles spécifiques. Il contribue d'abord à la clarification conceptuelle en permettant d'identifier des confusions ou des interprétations erronées. Il favorise également l'amélioration du raisonnement scientifique, notamment en attirant l'attention sur les ruptures logiques ou les conclusions insuffisamment étayées par les données (Montpied et al, 2016).

Dans son étude en classe de physique au secondaire, Ketonen (2020) montre que la mise en place d'un dispositif structuré d'évaluation par les pairs contribue à rendre les élèves plus attentifs aux critères scientifiques et à la cohérence de leurs explications. Le rôle d'évaluateur amène les élèves à adopter une posture réflexive et métacognitive, en mobilisant explicitement les normes du raisonnement scientifique (Ketonen, 2020).

On retrouve ici des points communs importants entre les différents travaux. Montpied (2016) et ses collègues mettent surtout l'accent sur le rôle de la régulation dans les apprentissages, tandis que Ketonen (2020) montre plus précisément comment cette régulation aide les élèves à mieux comprendre les critères scientifiques attendus. L'évaluation par les pairs peut ainsi être vue comme un moyen d'aider les élèves à s'approprier progressivement les règles et les exigences du raisonnement scientifique.

3.1.3 Les types de feedbacks en sciences

L'évaluation par les pairs en sciences peut porter sur plusieurs dimensions qu'il convient de distinguer explicitement. L'un des enjeux majeurs réside dans la nature des critères mobilisés.

Une première distinction concerne le contenu et la forme. Le contenu renvoie à l'exactitude scientifique, à la pertinence des concepts mobilisés et à la validité des explications. La forme concerne la clarté de la rédaction, l'organisation du texte et la qualité argumentative. Comme le montrent Boilevin et Brandt-Pomares (2011), la structuration du discours est indissociable de la qualité du raisonnement scientifique. Il faut donc veiller à ce que les élèves ne focalisent pas leur évaluation sur des éléments superficiels (orthographe, présentation), au détriment de l'analyse argumentative.

Une seconde distinction oppose la restitution des connaissances au raisonnement. Il est possible que l'évaluation se limite à vérifier si la réponse attendue est conforme, ou bien à s'assurer de la cohérence de la démarche et de la capacité à articuler données et conclusions (Boilevin & Brandt-Pomares, 2011).

Dans le cadre de l'évaluation par les pairs, l'objectivation des critères apparaît déterminante. Selon Lepareur (2016), la mise à disposition de critères explicites favorise une focalisation sur les processus cognitifs mobilisés plutôt que sur une simple vérification formelle. Les élèves sont alors amenés à porter leur attention sur la cohérence du protocole, la pertinence des données sélectionnées et la solidité de l'argumentation (Lepareur, 2016).

3.1.4 La mise en œuvre du dispositif

Les recherches s'accordent pour montrer que l'efficacité de l'évaluation par les pairs repose sur un cadrage précis et un étayage méthodologique explicite. Ketonen (2020) souligne l'importance d'un dispositif structuré, incluant des critères clairement définis, et un accompagnement progressif des élèves dans l'apprentissage du rôle d'évaluateur. De leur côté, Montpied et al. (2016) insistent sur la nécessité d'intégrer ces critères dès la conception de la séquence afin d'assurer une cohérence entre l'évaluation formative et l'évaluation sommative, ce qui contribue à renforcer la légitimité du dispositif aux yeux des élèves.

L'objectivation des critères constitue ainsi une condition centrale. Des critères explicites et compréhensibles permettent d'expliquer les attentes disciplinaires, de stabiliser des normes communes de jugement et de réduire l'arbitraire perçu dans l'évaluation. Ketonen (2020) montre d'ailleurs que la structuration préalable de ces critères améliore la qualité des feedbacks produits par les élèves.

Cependant, la seule présence de critères ne suffit pas. L'évaluation par les pairs suppose un apprentissage spécifique : savoir analyser une production, argumenter son jugement et proposer

des pistes pertinentes d'amélioration. Sans un accompagnement adapté, le dispositif risque de se limiter à des commentaires superficiels.

Dans cette perspective, l'enseignant conserve un rôle central de régulation, garant de la validité scientifique et de la qualité des échanges.

3.1.5 Avantages et limites

Les études mettent en évidence plusieurs effets positifs de l'évaluation par les pairs en sciences, sous réserve que le dispositif soit structuré et accompagné. Cette évaluation contribue d'abord à une meilleure compréhension des critères scientifiques et des attentes disciplinaires, ce qui favorise le développement de la métacognition.

Les travaux de Montpied et al. (2016) montrent par ailleurs que même des élèves, ayant rencontré des difficultés dans la réalisation de l'activité, peuvent produire des évaluations pertinentes lorsqu'ils disposent de critères explicites. Cela suggère que le rôle d'évaluateur peut constituer une opportunité d'apprentissage, y compris pour les élèves les plus fragiles.

De son côté, Ketonen (2020) met en évidence un engagement accru des élèves ainsi qu'une prise de conscience plus fine des exigences du raisonnement scientifique. Dans ces conditions, l'évaluation ne se limite plus à un moment de contrôle : elle devient une situation d'apprentissage à part entière et peut être envisagée comme un espace de construction collective du savoir.

Cependant, plusieurs limites et tensions sont également mises en évidence dans la littérature. La qualité des feedbacks peut être inégale, certains élèves peinant à dépasser des commentaires descriptifs ou superficiels. L'hétérogénéité des retours constitue ainsi un premier point de vigilance, d'autant plus que l'évaluation entre pairs peut entraîner un risque de validation d'erreurs conceptuelles si les échanges ne sont pas régulés.

Par ailleurs, le dispositif peut générer une surcharge cognitive pour certains élèves, notamment lorsque les attentes ne sont pas suffisamment explicitées ou que les critères restent trop complexes. Un autre risque réside dans une focalisation excessive sur des aspects formels au détriment de l'analyse du raisonnement scientifique, voire dans une instrumentalisation du dispositif sans réelle exploitation pédagogique des retours produits (Lepareur, 2016).

Comme le rappellent Boilevin et Brandt-Pomares (2011), développer des pratiques davantage centrées sur l'investigation nécessite un cadre institutionnel et professionnel adapté. Dans cette perspective, l'évaluation par les pairs en sciences ne constitue pas une solution miracle, mais un dispositif exigeant, qui suppose une démarche formative cohérente, explicitement outillée et régulée par l'enseignant.

3.2 L'évaluation par les pairs dans l'enseignement des langues

Si l'évaluation par les pairs peut être mobilisée dans de nombreuses disciplines, elle apparaît particulièrement pertinente en cours de langues, notamment dans le cadre d'activités de production écrite. En effet, l'écriture en langues étrangères constitue à la fois un objet d'apprentissage et un levier de développement des compétences linguistiques.

Cette partie du travail se concentre dès lors sur les apports de l'évaluation par les pairs dans le développement des compétences en langues étrangères, en s'appuyant sur les spécificités de ce contexte disciplinaire.

3.2.1 Le contexte de l'enseignement linguistique

Dans l'enseignement des langues étrangères, les activités de production écrite occupent une place centrale, dans la mesure où elles contribuent au développement des compétences linguistiques. Écrire en langues étrangères ne consiste pas uniquement à organiser des idées mais implique également la mobilisation de ressources linguistiques en cours d'acquisition, telles que le lexique, la grammaire et les structures discursives (Chen, 2024). L'écriture est dès lors envisagée comme un processus marqué par des phases de rédaction, de relecture et de réécriture successives.

Dans ce cadre, l'évaluation joue un rôle fondamental. Ritonga et al. (2022) soulignent que l'évaluation en cours de langues ne se limite pas à mesurer les acquis mais qu'elle soutient l'apprentissage en permettant aux apprenants d'identifier leurs forces et faiblesses linguistiques. Elle remplit ainsi une fonction formative en guidant à la fois l'enseignement et les apprentissages. Cette conception de l'évaluation met l'accent sur le processus de progression plutôt que sur le produit final, ce qui correspond aux objectifs de l'enseignement de l'écriture en langues étrangères.

L'évaluation par les pairs s'inscrit pleinement dans cette logique formative. En permettant aux élèves de recevoir des retours réguliers sur leurs productions écrites et d'analyser celles de leurs pairs, elle favorise un apprentissage continu et réflexif (Ritonga et al., 2022). Chen (2024) et Lou et al. (2016) montrent que la production de plusieurs versions d'un texte, accompagnée de

feedbacks entre pairs, constitue un levier important pour l'amélioration des compétences rédactionnelles en langues étrangères. Les apprenants sont ainsi amenés à réviser leurs textes en tenant compte des commentaires reçus, ce qui renforce leur implication dans le processus d'écriture.

Le cours de langues constitue un contexte particulièrement propice à l'évaluation par les pairs en raison de la place centrale accordée à l'interaction et à la collaboration. Lou et al. (2016) soulignent que le travail de groupe et les échanges autour des productions écrites favorisent l'interaction, la négociation du sens et la prise de conscience des normes linguistiques. Dans cette perspective, l'évaluation par les pairs ne se limite pas à une activité d'évaluation mais devient un outil d'apprentissage à part entière, intégré au processus d'acquisition de la langue étrangère.

3.2.2 Le rôle du feedback dans les apprentissages linguistiques

L'évaluation par les pairs repose sur un échange de feedbacks entre apprenants qui endossent tour à tour le rôle de donneur et de receveur. Ces deux postures mobilisent des processus cognitifs complémentaires (Dennen & Burner, 2008 cités par Cui & Schunn, 2024). En cours de langues, ce double rôle revêt une importance particulière dans la mesure où il contribue directement au développement de la compétence linguistique.

Lorsqu'ils formulent un feedback, les élèves sont amenés à analyser la production écrite d'un pair à partir de critères explicites tels que la clarté, la structure ou la correction linguistique. Cette activité développe des compétences critiques et analytiques et favorise une meilleure compréhension des normes de qualité propres à l'écriture en langues étrangères (Thompson, 2002 ; Nicol et al., 2014 ; Berggren, 2015, cités par Cui & Schunn, 2024). Chen (2024) souligne par ailleurs que le rôle de donneur de feedback peut être particulièrement bénéfique, y compris pour les apprenants dont le niveau linguistique est plus faible, car il les conduit à réfléchir explicitement sur l'usage de la langue.

Le rôle de receveur de feedback est tout aussi important. Les retours formulés par les pairs permettent aux élèves de prendre du recul par rapport à leur propre production et de repérer des problèmes qu'ils n'avaient pas identifiés seuls (Ferris, 2003 ; Trautmann, 2006 ; Nicol et al., 2014, cités par Cui & Schunn, 2024). Dans l'apprentissage d'une langue étrangère, cette prise de distance est essentielle car les apprenants ont souvent une conscience partielle de leurs erreurs linguistiques ou discursives.

Ainsi, en cours de langues, le feedback entre pairs remplit une fonction à la fois formative et métacognitive. Il aide les élèves à mieux comprendre leur propre fonctionnement linguistique et à développer des stratégies d'amélioration, ce qui correspond aux objectifs de l'évaluation formative décrits par Ritonga et al. (2022).

3.2.3 Les types de feedbacks en langues

Les feedbacks échangés entre pairs peuvent porter sur différents aspects des productions écrites. Fagley et Witte (1981, cités par Cui & Schunn, 2024) distinguent les feedbacks de surface qui concernent la grammaire, le lexique, l'orthographe ou la ponctuation, et les feedbacks de sens qui portent sur le contenu, l'organisation du texte et la cohérence de l'argumentation.

Dans les cours de langues, cette distinction est particulièrement importante. Chen (2024) montre que les apprenants en langues étrangères ont tendance à se concentrer prioritairement sur les aspects linguistiques locaux tels que les erreurs grammaticales ou lexicales car celles-ci sont plus visibles et plus faciles à identifier. Les commentaires portant sur l'organisation du texte ou le développement des idées sont, quant à eux, moins fréquemment pris en compte bien qu'ils soient essentiels à la qualité de la production écrite.

Plusieurs études soulignent toutefois que les feedbacks portant sur le sens sont généralement plus bénéfiques pour l'amélioration des textes (Kollar & Fischer, 2010 ; Cho & MacArthur, 2011, cités par Cui & Schunn, 2024). Cela ne signifie pas que les feedbacks de surface soient inutiles. Cui et Schunn (2024) précisent qu'ils peuvent constituer une aide précieuse, en particulier pour les apprenants débutants ou intermédiaires, pour qui la correction grammaticale et lexicale restent un enjeu majeur.

En cours de langues, l'enjeu n'est donc pas d'exclure un type de feedback au profit d'un autre mais de trouver un équilibre entre commentaires sur la forme linguistique et ceux sur le sens. Cet équilibre doit être soutenu par des critères explicites et un accompagnement pédagogique afin d'éviter que l'évaluation par les pairs ne se limite à une simple correction d'erreurs linguistiques.

3.2.4 La mise en œuvre du dispositif

Plusieurs auteurs insistent sur le fait que l'efficacité de l'évaluation par les pairs dépend fortement des conditions de mise en œuvre. En cours de langues, une formation explicite au feedback apparaît comme indispensable. Ritonga et al. (2022) soulignent que les apprenants doivent être guidés dans l'utilisation de critères clairs et adaptés aux composantes de la

compétence linguistique telles que le contenu, l'organisation, la grammaire, le vocabulaire et la mécanique de la langue.

Chen (2024) et Lou et al. (2016) montrent également que, sans accompagnement, les élèves ont tendance à produire des feedbacks superficiels ou peu exploitables. L'utilisation de grilles d'évaluation, de check-list ou de supports structurés permet de guider l'attention des élèves et de favoriser des commentaires plus équilibrés portant à la fois sur la forme et sur le sens.

Par ailleurs, la mise en œuvre de l'évaluation par les pairs en cours de langues doit tenir compte de la dimension socio-affective de l'apprentissage linguistique. Chen (2024) et Lou et al. (2016) soulignent que la formulation de feedbacks peut être influencée par des facteurs culturels tels que la peur de critiquer et d'être jugé, la recherche de l'harmonie du groupe ou la notion de « face ». Dans ce contexte, le rôle de l'enseignant est essentiel pour instaurer un climat sécurisant et bienveillant dans lequel le feedback est perçu comme une aide à l'apprentissage plutôt que comme un jugement.

3.2.5 Avantages et limites

Les recherches montrent que l'évaluation par les pairs présente de nombreux bénéfices en cours de langues. Elle favorise le développement de la compétence linguistique, tant pour les receveurs que les donneurs de feedback, et contribue à renforcer la conscience linguistique, la métacognition et l'autonomie des apprenants (Chen, 2024 ; Ritonga et al., 2022). Cette évaluation permet également de renforcer le lien entre lecture et écriture dans la mesure où l'analyse des productions des pairs mobilise des compétences de compréhension qui soutiennent indirectement les compétences rédactionnelles (Stotsky, 1983 ; Jouhar & Rupley, 2021, cités par Cui & Schunn, 2024).

Toutefois, plusieurs limites sont également identifiées. Sans formation adéquate, les feedbacks peuvent rester superficiels ou se concentrer excessivement sur les erreurs linguistiques locales (Chen, 2024). De plus, des facteurs affectifs peuvent freiner l'engagement des élèves dans l'évaluation par les pairs, notamment lorsqu'ils hésitent à formuler des commentaires critiques (Lou et al., 2016).

Ces éléments soulignent que, si l'évaluation par les pairs constitue un outil particulièrement pertinent en cours de langues, son efficacité dépend étroitement du cadre pédagogique dans lequel elle est mise en œuvre. Une formation explicite, des critères clairs et un accompagnement de l'enseignant apparaissent comme des conditions essentielles pour que ce dispositif contribue réellement à l'amélioration des productions écrites des élèves.

Questions de recherche

Au regard de la littérature mobilisée et de la problématique formulée, notre étude s'articule autour d'une question de recherche principale, déclinée en quatre sous-questions. La question principale qui guide cette étude est la suivante : **Quel est l'impact de l'évaluation par les pairs sur l'amélioration des productions écrites des élèves du secondaire inférieur issus du qualifiant en langues et en sciences ?**

Les sous-questions visent à examiner les effets de l'évaluation par les pairs sur les productions écrites des élèves, les caractéristiques des feedbacks qu'ils produisent, les liens entre ces feedbacks et les révisions apportées aux productions, ainsi que les éventuelles différences observées selon le contexte disciplinaire.

Les différents apports théoriques présentés dans la revue de la littérature ont guidé la construction des sous-questions de recherche présentées ci-après. Celles-ci permettent d'explorer l'évaluation par les pairs sous plusieurs angles complémentaires, afin de mieux comprendre les mécanismes par lesquels cette évaluation peut soutenir l'amélioration des productions écrites des élèves.

1. Sous-question 1

Dans une perspective d'évaluation formative, le feedback constitue un levier essentiel de régulation des apprentissages. Les travaux de Black et William (1998, cités par Bourgeois et Laveault, 2015) montrent que l'intégration de feedbacks dans le processus d'apprentissage permet aux élèves d'identifier les écarts entre leur production actuelle et les attentes visées. De même, Hattie et Timperley (2007 cités par Lafontaine et Toczec-Capelle, 2023) soulignent que l'efficacité du feedback repose sur sa capacité à orienter l'élève vers des actions concrètes d'amélioration. Dans cette optique, le feedback par les pairs peut favoriser un processus de révision des productions, tel que décrit par Sadler (1989), en soutenant l'appropriation des critères d'évaluation par les élèves.

Dans cette perspective, notre première sous-question s'intéresse aux évolutions observées dans les productions écrites après le dispositif d'évaluation par les pairs : **Quelles modifications sont observées entre les versions initiales et finales des productions écrites des élèves à la suite du processus d'évaluation par les pairs ?**

2. Sous-question 2

La littérature met en évidence que la qualité du feedback constitue un facteur déterminant de son efficacité. Hattie et Timperley (2007 cités par Lafontaine et Toczec-Capelle, 2023) distinguent différents niveaux de feedback (tâche, processus, autorégulation) tandis que Nicol et Macfarlane-Dick (2006) soulignent l'importance de feedbacks clairs, précis et orientés vers l'amélioration. De leur côté, Orsmond et al. (2002, cités par Mumpuni et al., 2022) montrent que des feedbacks élaborés, incluant des justifications et des suggestions concrètes, ont un impact plus important sur les apprentissages que des commentaires généraux.

Dans cette lignée, notre deuxième sous-question porte sur les caractéristiques des feedbacks produits par les élèves : **Quelles caractéristiques présentent les feedbacks formulés par les élèves lors de l'évaluation par les pairs ?**

3. Sous-question 3

Les recherches consacrées au feedback montrent que son efficacité dépend non seulement de sa qualité, mais également de la manière dont il est mobilisé par le destinataire. Les travaux de Hattie et Timperley (2007 cités par Lafontaine et Toczec-Capelle, 2023), de Nicol et Macfarlane-Dick (2006) ainsi qu'Orsmond et al. (2002, cités par Mumpuni et al., 2022) suggèrent que des feedbacks précis, justifiés et orientés vers l'action sont davantage susceptibles d'être pris en compte lors de la révision des productions.

Dans cette perspective, notre troisième sous-question vise à mettre en relation les feedbacks reçus et les évolutions observées dans les productions écrites : **Quels liens peut-on observer entre les feedbacks reçus et les modifications apportées aux productions écrites ?**

4. Sous-question 4

Plusieurs recherches suggèrent que l'efficacité du feedback peut varier en fonction du contexte disciplinaire. Les attentes en matière de production écrite diffèrent en effet selon les disciplines, tant au niveau des contenus que des formes attendues. Anker-Hansen et Andrée (2019) mettent en évidence que les pratiques d'évaluation et les modalités de feedback peuvent être influencées par les spécificités des disciplines. Dans cette perspective, les productions en langues mobilisent davantage des compétences liées à la structuration du discours et à l'argumentation, tandis qu'en sciences l'accent est davantage mis sur la précision des contenus et la validité des explications.

Dans cette lignée, notre quatrième sous-question vise à examiner si les résultats observés diffèrent selon la discipline concernée : **Ces résultats se manifestent-ils de manière différente dans les cours de sciences et d'anglais ?**

Méthodologie

Afin d'apporter une réponse à notre question de recherche, une étude de cas a été mise en œuvre au sein d'une classe de troisième année de l'enseignement secondaire qualifiant, dans le cadre des cours de sciences et d'anglais. Dans cette section, nous présenterons tout d'abord l'échantillon sur lequel repose notre recherche. Nous décrirons ensuite l'intervention menée et les outils mobilisés. Enfin, nous préciserons les méthodes d'analyse utilisées pour examiner les résultats de cette intervention.

1. L'échantillon

Le public cible de cette recherche est constitué d'élèves du secondaire inférieur relevant de l'enseignement qualifiant en Fédération Wallonie-Bruxelles. Ce choix se justifie par le fait que cette population demeure relativement peu investiguée dans les travaux portant sur l'évaluation par les pairs.

Le dispositif de recherche a été mis en œuvre au sein d'une classe de troisième année de l'enseignement secondaire qualifiant, relevant de l'enseignement libre en Belgique francophone. Cette classe a été sélectionnée sur base volontaire. L'échantillon se compose d'un groupe de onze élèves appartenant à une même année scolaire, ce qui permet de favoriser la comparabilité des contextes.

À la suite d'échanges avec les enseignantes volontaires, il est apparu que l'évaluation par les pairs n'était pas mise en œuvre dans leurs pratiques habituelles. Le dispositif proposé constituait dès lors une première expérience de cette méthode d'évaluation pour les élèves, tant en cours de sciences qu'en cours d'anglais. Les autres enseignants de l'établissement n'ont pas été sollicités dans le cadre de cette recherche.

2. La mise en œuvre du dispositif

2.1 Structuration du dispositif et posture des enseignants

Dans le cadre du dispositif, nous avons adopté une posture d'observation. Ce sont les enseignants titulaires qui ont décidé de mener les activités. Ils ont eu un rôle d'encadrement et de guidance, tout en adoptant une position en retrait lors des activités d'évaluation. Dans cette logique, les élèves occupaient une place centrale dans l'activité. Ce choix s'inscrit dans les

fondements théoriques de l'évaluation par les pairs, qui repose sur une conception du rôle de l'enseignant caractérisée par un décentrement au profit des élèves (Reinholz, 2016).

Par ailleurs, afin d'assurer la cohérence du dispositif et d'accompagner les enseignants dans sa mise en œuvre, nous avons élaboré une trame d'accompagnement basée sur les apports de la littérature. Celle-ci visait à guider la structuration des activités et à harmoniser les pratiques d'évaluation par les pairs dans les deux disciplines (voir annexe 1).

2.2 Préparation des élèves à l'évaluation par les pairs

Préalablement à la mise en œuvre, une formation des élèves au feedback constructif a été organisée, conformément aux recommandations de la littérature. Cette formation était commune aux cours de sciences et de langues et animée conjointement lors d'une même séance. Elle visait à distinguer feedback constructif et jugement, ainsi qu'à expliciter les critères d'un retour pertinent. Elle a pris la forme d'un jeu au cours duquel les élèves devaient identifier la nature de différents commentaires, avant de co-construire une série de critères à mobiliser lors de la rédaction de feedback.

2.3 Phases du dispositif d'évaluation par les pairs

Le dispositif s'inscrit dans une logique d'évaluation formative par les pairs et s'appuie sur le modèle du *assessment cycle* proposé par Reinholz (2016). Il a été organisé en deux cycles successifs, chacun comprenant cinq phases : une phase préparatoire, une production individuelle, une évaluation par les pairs, une réception du feedback et une révision.

La phase préparatoire visait à expliciter les critères de réussite et à accompagner leur appropriation par les élèves, condition essentielle au développement du jugement évaluatif (Bloxxham & West, 2004 ; Sluijsmans et al., 2004). Chaque cycle débutait ensuite par une production individuelle, permettant aux étudiants de mobiliser ces critères dans une première mise en situation.

Les élèves étaient ensuite amenés à analyser la production d'un pair à l'aide d'une grille d'évaluation. Cette activité, centrale dans le dispositif, les engage dans un processus de lecture critique et favorise la compréhension des attentes ainsi que le développement de compétences argumentatives (Noroozi, 2023). Elle donnait lieu à la rédaction d'un feedback écrit, pour lequel une attention particulière était portée à la précision, à l'argumentation et à l'orientation vers l'amélioration.

Les élèves prenaient ensuite connaissance du feedback reçu, afin d'identifier des pistes de régulation de leur production. Cette phase de réception du feedback s'inscrit dans une logique de régulation des apprentissages, dans laquelle l'élève ajuste progressivement son travail en fonction des écarts identifiés (Girouard-Gagné & Durand, 2025). Enfin, une phase de révision leur permettait de retravailler leur texte initial en tenant compte des commentaires formulés, et ce, dans une perspective d'amélioration. Cette étape permet de transformer le feedback en actions concrètes d'amélioration et de soutenir les apprentissages.

2.4 Adaptations du dispositif

Le dispositif de recherche prévoyait initialement la réalisation de deux cycles complets d'évaluation par les pairs tant en sciences qu'en anglais.

Si les deux cycles en sciences ainsi que le premier cycle en anglais ont pu être menés conformément au protocole, le second cycle d'anglais a été interrompu à la suite des mouvements de grève ayant affecté l'enseignement secondaire en Fédération Wallonie-Bruxelles en fin d'année scolaire. Seules les productions initiales ont ainsi pu être recueillies, les étapes d'évaluation par les pairs, de feedback et de réécriture n'ayant pas pu être réalisées.

Afin de valoriser les données collectées malgré cette interruption, une analyse complémentaire a été mise en place. Celle-ci consiste à comparer les productions initiales du premier cycle (A1) et du second cycle (A2), qui mobilisent une même compétence d'expression écrite et reposent sur des attentes similaires. Les deux séries de productions ont été analysées à l'aide de la grille critériée utilisée par les élèves tout au long du dispositif, et d'un document de codage commun élaboré afin d'harmoniser les analyses réalisées (voir annexe 9). Cette analyse ne vise donc pas à évaluer l'effet d'un second cycle d'évaluation par les pairs, mais à documenter l'évolution des productions écrites entre deux moments distincts de la recherche.

3. Les outils exploités

3.1 Les tâches

Dans le cadre du dispositif, les tâches proposées dans les cours de sciences et de langues ont été choisies en concertation avec les enseignants titulaires afin de s'inscrire au mieux dans le programme et dans les apprentissages en cours. En sciences, il s'agissait de la rédaction de deux rapports de laboratoire portant sur des phénomènes physiques et chimiques (voir annexe 2). Il est à noter que les élèves n'avaient, pour la plupart, jamais été confrontés à ce type d'exercice

auparavant, ce qui impliquait un accompagnement particulier. Ce type de production mobilise en effet plusieurs compétences propres aux démarches d'investigation, telles que la représentation du dispositif expérimental, la formulation d'observations et l'interprétation des résultats (Boilevin & Brandt-Pomares, 2011). En langues, les élèves devaient réaliser deux productions écrites sur le thème de la description physique, la première sous forme de devoir et la seconde sous forme d'évaluation de fin de chapitre (voir annexe 3). Contrairement au cours de sciences, ce type de tâches s'inscrivait davantage dans les pratiques habituelles de la discipline. Ce choix de tâches visait à garantir la pertinence pédagogique des activités tout en limitant les perturbations du déroulement habituel du cours.

Dans les deux disciplines, les tâches retenues présentent une caractéristique commune essentielle : elles ne conduisent pas à une réponse unique attendue. La rédaction d'un rapport de laboratoire, tout comme la production d'un texte en langues, permet en effet une diversité de réponses tant dans la manière de structurer les propos que dans les choix opérés par les élèves. Dans le cas des productions écrites en langues, cette ouverture se manifeste notamment par la variété des choix lexicaux, organisationnels et stylistiques qui participent à la richesse des productions (Chen, 2024 ; Lou et al., 2016). Cette caractéristique distingue ces tâches d'exercices plus fermés et favorise la production de travaux variés.

Cette diversité des productions rend ces tâches particulièrement propices à la mise en œuvre d'une évaluation par les pairs. Elle permet aux élèves d'analyser le travail d'un camarade en s'appuyant sur différents aspects de la production, qu'il s'agisse de la représentation du dispositif expérimental, de la formulation des observations et de l'interprétation des résultats en sciences ; en langues, du choix du vocabulaire, de l'organisation du texte et du respect des règles linguistiques. Les élèves sont ainsi amenés à formuler des jugements argumentés et à proposer des pistes d'amélioration, ce qui favorise la production de feedbacks plus développés et contribue à une perspective d'évaluation formative (Montpied et al., 2016; Ketonen, 2020).

3.2 Les grilles de critères

Dans le cadre du dispositif, des grilles de critères ont été élaborées afin de soutenir l'évaluation par les pairs, tant pour les rapports de laboratoire en sciences que pour les productions écrites en langues. L'objectif était de proposer des outils à la fois cohérents avec les apports de la littérature et adaptés aux contraintes du terrain.

Dans chaque discipline, les critères ont été définis en fonction des spécificités des apprentissages visés.

En sciences, la grille (voir annexe 4) s'organise autour de trois critères : le schéma de l'expérience, les observations et la conclusion. Ce choix s'appuie sur les travaux portant sur l'écriture scientifique et l'évaluation des rapports de laboratoire ; ceux-ci mettent en évidence l'importance de structurer l'analyse autour des étapes clés de la démarche expérimentale, à savoir la mise en place de l'expérience, le recueil de données et leur interprétation (Anker-Hansen & André, 2019). Plus précisément, le critère du schéma de l'expérience vise à évaluer la capacité des élèves à représenter, de manière claire et structurée, le dispositif expérimental en rendant compte des éléments essentiels et leur organisation. Le critère des observations porte sur la qualité du relevé des résultats, en termes de précision, de pertinence et de fidélité aux données recueillies. Enfin, le critère de la conclusion évalue la capacité des élèves à interpréter les résultats obtenus, à établir des liens avec la question de départ et à formuler un raisonnement scientifique cohérent. Ce choix a été réalisé en concertation avec l'enseignante titulaire de la classe, afin de cibler des éléments jugés essentiels dans l'évaluation des productions des élèves. Le fait de limiter le nombre de critères vise également à faciliter l'appropriation de la grille par les étudiants et à éviter une surcharge cognitive. Cette orientation rejoint les recommandations issues des travaux sur les rubriques en écriture scientifique, qui insistent sur la nécessité de proposer des outils simples, centrés sur les apprentissages essentiels (Newstadt & Godley, 2019).

En langues, la grille (voir annexe 5) comprend cinq critères principaux qui distinguent différentes dimensions de l'écriture. Le critère de contenu porte sur la pertinence et le développement des idées en lien avec la consigne. L'organisation concerne la structuration du texte et la cohérence des idées, ainsi que l'usage de connecteurs logiques. Le vocabulaire renvoie à la richesse lexicale, à la précision des mots choisis et à la variété des formulations. La grammaire évalue la correction des structures syntaxiques, des accords et des temps verbaux. Enfin, l'orthographe prend en compte la ponctuation et le respect des conventions de l'écrit. L'ensemble de ces critères permet de couvrir les dimensions linguistiques, discursives et formelles de la production écrite, en cohérence avec les objectifs d'apprentissage visés (Jacobs et al., 1981).

Dans les deux disciplines, chaque critère est décliné en quatre niveaux de performance décrits de manière progressive. Le choix d'une échelle à quatre niveaux vise à proposer une gradation

suffisamment fine pour distinguer les différents niveaux de maîtrise, tout en conservant une structure lisible pour les élèves. Par ailleurs, l'absence de niveau intermédiaire neutre conduit les étudiants à prendre position, de manière plus explicite, dans leur évaluation ce qui peut favoriser leur engagement et la qualité des feedbacks formulés.

Une attention particulière a été portée à la formulation des descripteurs, afin qu'ils soient à la fois compréhensibles par les élèves et suffisamment précis pour guider leur analyse. Dans le cadre de l'évaluation par les pairs, la clarté des critères constitue un élément essentiel dans la mesure où elle conditionne la qualité des feedbacks produits. Des descripteurs explicites permettent ainsi d'orienter l'attention des élèves vers les éléments pertinents et de favoriser des commentaires plus argumentés. (Montpied et al., 2016).

Enfin, les grilles ont été présentées aux élèves en amont des activités de production afin qu'elles puissent également servir d'outil d'anticipation des attentes. Elles ont ensuite été mobilisées lors des activités d'évaluation par les pairs, au cours desquelles chaque élève devait analyser le travail d'un camarade et formuler un feedback constructif.

4. Type d'analyse

Cette section présente le type d'analyse mobilisé pour examiner les données recueillies dans le cadre du dispositif. L'analyse s'articule autour de trois volets complémentaires : l'évolution des productions écrites entre la version initiale et la version révisée, la qualité des feedbacks formulés par les élèves et le lien entre le feedback reçu et les améliorations observées.

Dans un premier temps, les données ont fait l'objet de deux codages distincts. Dans un second temps, les deux codages ont été confrontés et mis en commun afin d'aboutir à un codage consensuel avant de comparer les résultats obtenus en sciences et en anglais. Cette démarche visait à renforcer la fiabilité et la rigueur de l'analyse en confrontant les interprétations issues des deux codages. Les divergences éventuelles ont ainsi pu être discutées jusqu'à l'obtention d'un consensus, limitant les biais liés à l'interprétation individuelle et garantissant une analyse la plus objective et cohérente possible.

4.1 Analyse de l'évolution des productions

Le premier volet de l'analyse porte sur l'évolution des productions écrites des élèves, en comparant les versions initiales et les versions révisées à la suite de la réception des feedbacks.

L'objectif est de mettre en évidence les modifications apportées par les élèves et d'examiner dans quelle mesure celles-ci traduisent une amélioration des productions grâce à un document cadre garantissant la même analyse selon la discipline (voir annexe 6).

Dans cette perspective, une attention particulière est portée aux transformations observées entre les deux versions des productions. Il s'agit d'identifier la présence ou l'absence de modification, ainsi que leur ampleur. Trois situations ont ainsi été distinguées : l'absence de modification, lorsque la production reste inchangée ; la modification partielle, lorsque certaines transformations sont apportées ; et la modification plus complète, lorsque la production a été retravaillée de manière approfondie. Afin de préciser ces évolutions, les modifications ont également été caractérisées selon leur type : ajout, suppression ou reformulation.

L'ensemble de ces éléments permet ainsi d'appréhender, de manière fine, l'évolution des productions des élèves, en mettant en évidence la nature des transformations opérées.

4.2. Analyse de la qualité des feedbacks

Le second volet de l'analyse porte sur la qualité des feedbacks formulés par les élèves dans le cadre de l'évaluation par les pairs. L'objectif est d'examiner dans quelle mesure les commentaires produits sont précis, constructifs, alignés sur les critères d'évaluation et susceptibles de soutenir les apprentissages.

Afin d'analyser les feedbacks, une grille spécifique a été élaborée (voir annexe 7). Celle-ci repose sur plusieurs dimensions identifiées dans la littérature, telles que la précision du commentaire, sa justification, son caractère constructif, sa clarté ainsi que son orientation vers l'amélioration.

En effet, la précision du feedback permet de cibler explicitement les écarts entre la production de l'élève et les attentes, ce qui constitue une condition essentielle pour soutenir l'apprentissage (Sadler, 1989 ; Hattie & Timperley, 2007 cités par Lafontaine et Toczec-Capelle, 2023). De plus, la mobilisation explicite des critères d'évaluation favorise une meilleure compréhension des attentes et contribue à la cohérence des jugements entre pairs (Panadero & Jonsson, 2013 ; Sluismans et al., 2004).

La dimension constructive du feedback, à travers la formulation de propositions d'amélioration, apparaît également centrale. En effet, elle permet d'orienter l'élève vers des actions concrètes de révision et de soutenir l'autorégulation (Hattie & Timperley, 2007 cités par Lafontaine et

Toczek-Capelle, 2023 ; Nicol & Macfarlane-Dick, 2006). Par ailleurs, la clarté et la lisibilité du message conditionnent la compréhension du feedback par le destinataire, tandis que la qualité du ton influence son acceptation et son appropriation (Panadero, 2016 ; Topping, 1998).

Enfin, la présence de justifications ou de preuves constitue un élément essentiel du feedback, dans la mesure où elle permet de rendre explicite le jugement porté et d'en renforcer la pertinence (Sadler, 1989 ; Topping, 1998).

L'ensemble de ces dimensions a ainsi été intégré dans la grille d'analyse afin de rendre compte, de manière structurée, de la qualité des feedbacks produits par les élèves.

Chaque feedback a été codé à partir de cette grille selon une échelle à quatre niveaux (de 0 à 3), allant d'un critère absent à un critère fortement présent. Ce choix permet de rendre compte de la variabilité des productions des élèves tout en conservant une structure d'analyse lisible. Le score total obtenu pour chaque feedback constitue une estimation globale de sa qualité.

L'analyse a été réalisée à l'échelle de chaque commentaire constituant le feedback. Ainsi, lorsqu'un feedback comporte plusieurs commentaires portant sur différents aspects du travail, chacun d'eux a été analysé séparément à l'aide de la grille afin d'en évaluer la qualité de manière fine. Cette approche permet de prendre en compte la diversité des commentaires au sein d'un même feedback.

Afin d'interpréter les scores obtenus, des seuils ont été définis pour distinguer différents niveaux de qualité des feedbacks. Il s'agit d'un choix méthodologique propre à cette recherche, établi en cohérence avec l'échelle de codage utilisée. Cette catégorisation vise à faciliter l'interprétation des résultats, sans constituer une norme absolue, en distinguant trois niveaux de qualité. Les feedbacks obtenant un score de 0 à 6 sont considérés comme étant de faible qualité, car les critères retenus sont peu présents et les commentaires restent généralement vagues ou peu exploitables. Les scores de 7 à 12 correspondent à des feedbacks de qualité moyenne, dans lesquels les critères sont présents, mais de manière partielle ou insuffisamment développée. Enfin, les scores de 13 à 18 renvoient à des feedbacks de bonne qualité, caractérisés par des commentaires précis, justifiés, constructifs et appuyés sur les critères d'évaluation. Cette démarche s'inscrit dans la perspective des travaux de Nicol et Macfarlane-Dick (2006), qui soulignent l'existence de degrés dans la qualité des feedbacks.

Enfin, les résultats de cette analyse visent à mieux comprendre la capacité des élèves à produire un feedback de qualité et à identifier les dimensions nécessitant un accompagnement spécifique.

Ils permettent également d'évaluer, dans quelle mesure, le dispositif mis en place favorise le développement de compétences liées à l'évaluation et à la régulation des apprentissages.

4.3. Analyse du lien entre la qualité des feedbacks et les améliorations

Le troisième volet de notre analyse porte sur le lien entre les feedbacks et les améliorations apportées. En effet, si l'analyse des feedbacks et des productions permet d'apporter des éclairages, elle ne permet pas, prise seule, de rendre compte de l'impact du dispositif d'évaluation par les pairs sur les productions écrites des élèves.

Dès lors, il apparaît nécessaire de mettre en relation ces deux dimensions afin de mieux comprendre les processus à l'œuvre dans la révision des productions. L'objectif est d'examiner, dans quelle mesure, les feedbacks reçus influencent les modifications apportées par les élèves et de déterminer si une amélioration des productions peut être associée à la qualité des commentaires formulés par les pairs.

Dans cette perspective, un outil de mise en relation (voir annexe 8) a été élaboré sous la forme d'un tableau de synthèse permettant de croiser, pour chaque élève, les informations issues de l'analyse des feedbacks (qualité des commentaires) et celles issues de l'analyse des productions (nature des modifications). Ce dispositif vise à structurer l'analyse en facilitant l'identification de tendances.

Toutefois, les élèves ne mobilisent pas tous les feedbacks de la même manière. Afin de rendre compte de cette diversité, une approche par profil a été privilégiée. Celle-ci permet d'identifier différentes manières d'utiliser les feedbacks dans la révision des productions et de mieux comprendre les liens entre la qualité des commentaires reçus et les améliorations observées.

Afin de structurer cette analyse, une typologie de profils d'élèves a été définie. Celle-ci s'inspire des travaux Ketonen et al. (2020) qui mettent en évidence différents parcours d'élèves (« pathways ») dans le cadre de l'évaluation par les pairs. Les auteurs montrent que les effets du feedback ne sont pas uniformes et dépendent notamment de l'engagement des élèves, de la qualité du feedback reçu et des bénéfices qui en découlent.

Dans le cadre de la présente recherche, cette typologie a été adaptée afin de correspondre aux variables analysées, à savoir la qualité des feedbacks, leur prise en compte par les élèves et la qualité des modifications apportées aux productions écrites. Les profils ont ainsi été construits

à partir de ces trois dimensions, dans une perspective qualitative visant à rendre compte de la diversité des trajectoires d'apprentissage plutôt qu'à établir des relations strictement causales.

Six profils ont été définis :

- Le premier profil regroupe les élèves ayant reçu des feedbacks de bonne qualité qu'ils ont su mobiliser, de manière pertinente, dans la révision de leur production. Les modifications apportées sont majoritairement cohérentes avec les suggestions reçues et contribuent à une amélioration globale du texte.
- Le deuxième profil concerne les élèves ayant reçu des feedbacks de bonne qualité, mais n'en exploitant qu'une partie. Les modifications apportées demeurent limitées ou incomplètes au regard des suggestions formulées, traduisant une appropriation partielle des feedbacks.
- Le troisième profil correspond aux situations dans lesquelles les feedbacks reçus sont de faible qualité, ce qui se traduit par des modifications limitées et une faible amélioration de la production écrite.
- Le quatrième profil regroupe les élèves ayant reçu des feedbacks de qualité moyenne, mais dont les révisions conduisent malgré tout à une amélioration satisfaisante de leur production. Ces situations témoignent d'une capacité des élèves à tirer parti de feedbacks imparfaits.
- Le cinquième profil concerne les élèves ayant reçu des feedbacks de qualité moyenne, mais dont l'appropriation demeure limitée. Les modifications apportées restent partielles et les améliorations observées sont modestes au regard du potentiel offert par les feedbacks.
- Le sixième profil rassemble les situations caractérisées par une appropriation différenciée des feedbacks selon leur qualité. Il s'agit de cas atypiques ou plus complexes, dans lesquels les élèves exploitent de manière variable les feedbacks reçus. Ce profil permet de rendre compte de la diversité des processus d'appropriation observés et de préserver la richesse des données qualitatives en intégrant des trajectoires qui ne s'inscrivent pas dans les tendances générales.

4.4 Analyse comparative des résultats en sciences et en anglais

À l'issue du double codage des trois analyses, des tableaux de synthèse ont été élaborés pour chacun des volets étudiés : la qualité des feedbacks, la nature des modifications apportées aux productions et le lien entre les feedbacks reçus et les modifications réalisées.

Le premier tableau présente, pour chaque discipline et chaque tâche, les scores moyens obtenus selon les différents critères d'évaluation de la qualité des feedbacks, ainsi que la répartition des élèves selon le niveau de qualité de ceux-ci. (voir annexe 10)

Le deuxième tableau synthétise les caractéristiques des modifications apportées aux productions en indiquant leur nombre total, leur répartition selon leur nature ainsi que les parties des productions les plus concernées par ces corrections dans chaque discipline. (voir annexe 11)

Le troisième tableau met en évidence le degré de correspondance entre les feedbacks formulés et les modifications effectivement réalisées, tout en présentant une répartition des élèves selon différents profils caractérisant leur manière de prendre en compte les feedbacks. (voir annexe 12)

Enfin, les analyses réalisées en sciences et en anglais ont été synthétisées dans des tableaux récapitulatifs propres à chaque discipline (voir annexe 13). Ces tableaux ont constitué le support de la comparaison des résultats entre les deux disciplines, en permettant d'identifier les principales convergences et divergences observées.

Présentation et discussion des résultats

1. Sous-question 1 : Quelles modifications sont observées entre les versions initiales et finales des productions écrites à la suite du processus d'évaluation par les pairs ?

Afin d'examiner les modifications apportées par les élèves à leurs productions après la réception des feedbacks de leurs pairs, les différentes versions des travaux ont été comparées à l'aide du document cadre (voir annexe 6) et du tableau de synthèse présentant les résultats du double codage (voir annexe 11). L'analyse porte sur la nature des modifications réalisées, sur les éléments des productions concernés ainsi que sur l'ampleur des changements observés.

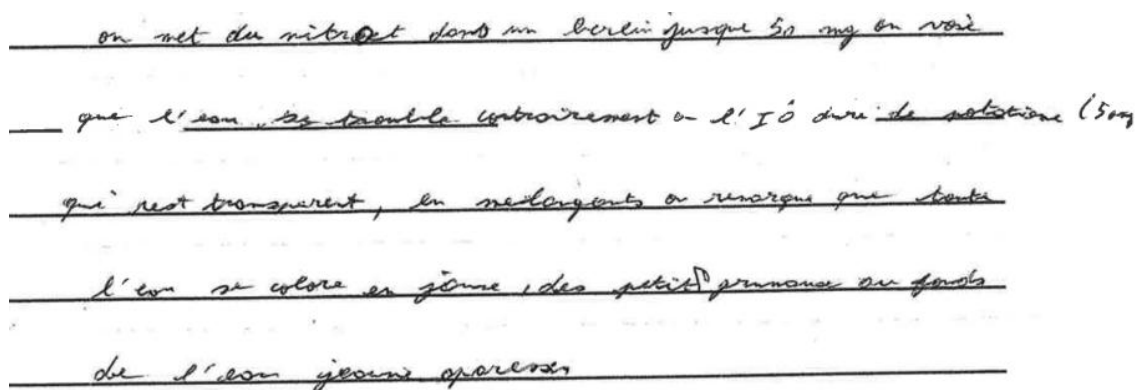
1.1 Les modifications observées dans les productions de sciences

1.1.1 Nature des modifications

En analysant les productions de sciences, il est évident que les élèves ont apporté diverses modifications à leurs rapports entre la version initiale et la version finale.

La première tâche porte sur l'analyse de 11 productions, tandis que la deuxième en compte 9, en raison de l'absence de certains élèves. Pour la première tâche, les modifications observées se répartissent entre les ajouts (48,4%) et les reformulations (41,9%), tandis que les suppressions restent peu fréquentes (9,7%). La deuxième tâche présente une tendance similaire : les ajouts représentent 57,2% des changements, les reformulations 33,3% et les suppressions seulement 9,5%. Ces résultats indiquent que les élèves privilégient l'ajout d'informations et la reformulation d'éléments existants plutôt que la suppression de contenus.

Les ajouts concernent notamment des précisions apportées aux observations ou aux schémas, tandis que les reformulations visent souvent à clarifier ou à préciser certaines informations scientifiques.



on met du nitrate dans un bocal jusqu'à 50 mg on voit
que l'eau se trouble uniformément on l'a ôté de la solution (50g
qui rest transparent, en mélangeant on remarque que toute
l'eau se colore en jaune, des petites particules au fond
de l'eau jaune apparaît

Figure 2. Extrait de la version initiale avant reformulation (élève 9, tâche 1).

On verse 0,5g de nitrat dans un bérin rempli
 jusqu'à 50 ml d'eau, l'eau se trouble grace
 au mélange eau (50ml) + nitrat (0,5g). Ensuite
 on répète la même opérations avec 0,6g
 d'IO dure de potassium et 5 ml d'eau dans
 un autre bérin. Après on mélange les
 deux substance (IO) dure + nitrat et
 on obtient une substance jaune.

Figure 3. Extrait de la version finale après reformulation (élève 9, tâche 1).

1.1.2 Parties du travail concernées

Les modifications observées touchent principalement les observations et les schémas. Lors de la première tâche, 11 modifications concernent les observations et 10 les schémas, contre seulement 4 pour les conclusions. Une répartition relativement similaire est observée dans la deuxième tâche, avec 6 modifications portant sur les observations, 6 sur les schémas et 5 sur les conclusions.

Ces résultats montrent que les élèves interviennent davantage sur les parties descriptives et visuelles de leur rapport scientifique. Les observations font fréquemment l'objet d'ajouts ou de reformulations visant à préciser les phénomènes observés ou à intégrer un vocabulaire scientifique plus adéquat. De même, les schémas sont régulièrement retravaillés, notamment par l'ajout de légendes ou de précisions graphiques. Les conclusions, bien que modifiées dans certaines productions, semblent faire l'objet d'un nombre plus restreint de révisions.

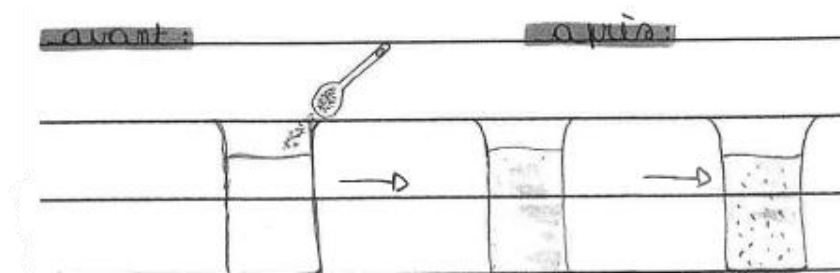


Figure 4. Extrait de la version initiale illustrant un schéma (élève 2, tâche 1).

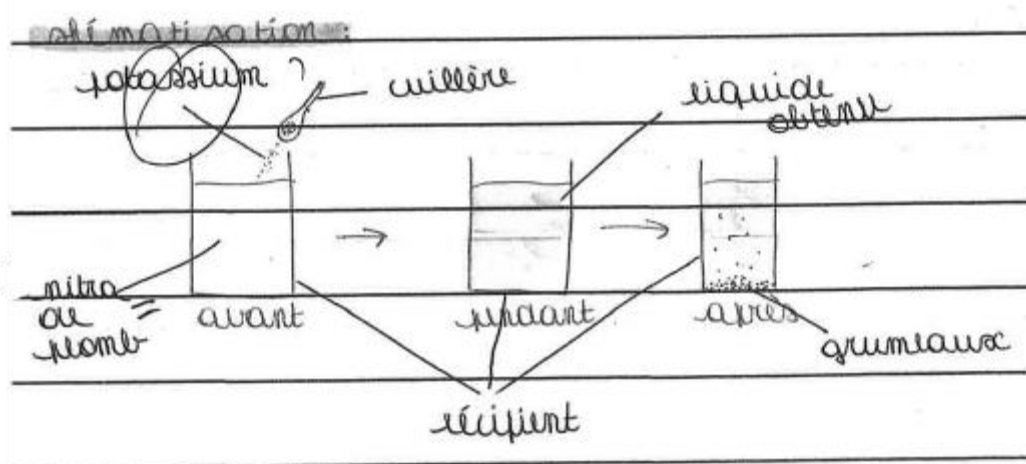


Figure 5. Extrait de la version finale illustrant le même schéma (élève 2, tâche 1).

1.1.3 Ampleur des modifications

L'analyse montre également une variabilité dans l'ampleur des modifications apportées d'un élève à l'autre.

Lors de la première tâche, la majorité des élèves réalise entre 2 et 3 modifications, tandis que certains apportent jusqu'à 4 ou 5 changements. La majorité des élèves a donc procédé à plusieurs changements entre les deux versions de son rapport.

Dans la deuxième tâche, les profils apparaissent plus contrastés. Un élève n'a apporté aucune modification à sa production après la réception du feedback. 3 élèves ont réalisé une seule modification, 3 autres en ont effectué 3 et 2 élèves en ont apporté 4. Bien que le nombre total de changements soit légèrement inférieur à celui observé lors de la première tâche, la majorité des élèves a néanmoins retravaillé sa production à la suite du processus d'évaluation par les pairs.

Dans l'ensemble, les résultats mettent en évidence une activité de révision des productions scientifiques après la réception des feedbacks.

1.2 Les modifications observées dans les productions d'anglais

1.2.1 Nature des modifications

L'analyse des productions d'anglais met en évidence que les élèves ont également apporté diverses modifications à leurs textes entre la version initiale et la version finale. En raison de l'absence de certains élèves, l'analyse porte sur 8 productions, contre 11 initialement prévues pour cette première tâche. Au total, 32 modifications ont été recensées dans ces 8 productions,

dont 17 reformulations (53,1%), 11 ajouts (34,4%) et 4 suppressions (12,5%). Les reformulations constituent ainsi le type de modification le plus fréquent.

Les reformulations concernent principalement la reformulation de phrases, la correction d'expressions ou l'amélioration de certaines formulations. Les ajouts permettent, quant à eux, d'enrichir le contenu des productions ou d'apporter des précisions supplémentaires. Les suppressions restent relativement peu nombreuses et concernent essentiellement le retrait d'éléments jugés inutiles ou redondants.

Ces résultats montrent que les élèves privilégient majoritairement la réécriture et l'enrichissement de leurs productions plutôt que la suppression de contenu existant.

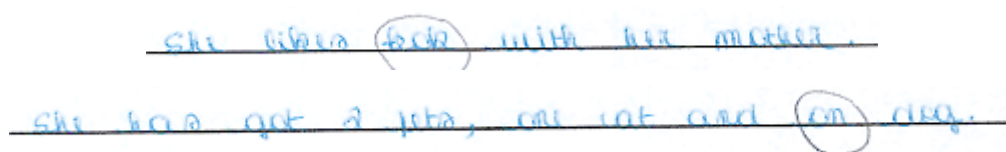


Figure 6. Extrait de la version initiale avant reformulation (élève 2, tâche 1).

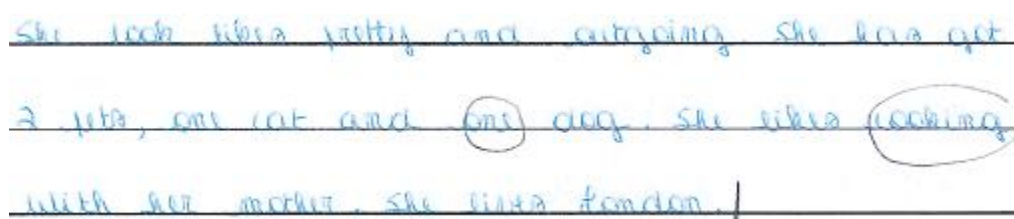


Figure 7. Extrait de la version finale après reformulation (élève 2, tâche 1).

1.2.2 Parties du travail concernées

Les modifications observées touchent différentes dimensions de la production écrite. L'organisation du texte constitue la dimension la plus fréquemment retravaillée avec 6 modifications recensées. Viennent ensuite la grammaire qui fait l'objet de 5 modifications, puis le contenu et le vocabulaire avec 4 modifications chacun. Enfin, 3 modifications concernent l'orthographe.

La fréquence relativement élevée des modifications liées à l'organisation suggère que plusieurs élèves ont retravaillé la structure de leur texte, notamment en réorganisant certaines idées ou en améliorant l'enchaînement des informations avec des paragraphes. Les modifications grammaticales témoignent, quant à elles, d'un travail de correction linguistique, tandis que les changements portant sur le vocabulaire visent généralement à préciser ou enrichir le lexique utilisé.

Ainsi, les révisions observées ne se limitent pas à un seul aspect de l'écriture mais concernent à la fois la structuration du texte, la correction de la langue et le contenu développé par les élèves.

1.2.3 Ampleur des modifications

L'ampleur des modifications varie sensiblement d'un élève à l'autre. Un élève n'a apporté aucune modification à sa production après la réception du feedback. 2 élèves ont effectué 2 modifications, un élève en a réalisé 3 et un autre 4. Par ailleurs, 2 élèves ont apporté 6 modifications à leur texte, tandis qu'un élève a effectué 9 modifications entre les deux versions de sa production.

Ces résultats mettent en évidence une implication variable des élèves dans le processus de révision. Si certains se limitent à quelques ajustements ponctuels, d'autres entreprennent une modification plus approfondie de leur travail. Malgré cette hétérogénéité, la majorité des élèves a procédé à plusieurs modifications à la suite de l'évaluation par les pairs.

Dans l'ensemble, l'analyse des productions d'anglais montre que les élèves ont majoritairement recours à la reformulation et aux ajouts pour réviser leurs textes.

1.3 Analyse complémentaire de l'évolution des productions initiales en anglais

1.3.1 Objectif de l'étude

Le second cycle d'évaluation par les pairs en anglais n'ayant pu être mené à son terme en raison des perturbations de la fin de l'année scolaire, seules les premières versions des productions écrites ont pu être recueillies. Bien que ces données ne permettent pas d'analyser l'effet du feedback entre pairs ni les modifications apportées lors de la révision des textes, elles offrent l'opportunité d'examiner l'évolution des productions initiales entre les deux cycles du dispositif.

Cette analyse complémentaire vise ainsi à déterminer si les élèves abordent la seconde tâche d'écriture avec un niveau de maîtrise différent de celui observé lors du premier cycle. Elle ne cherche pas à mesurer directement l'impact de l'évaluation par les pairs, mais à documenter l'évolution des productions initiales au cours de la mise en œuvre du dispositif.

1.3.2 Présentation des résultats

Les résultats montrent qu'une progression des productions initiales est observée chez 5 élèves sur 9, tandis qu'un élève présente un niveau comparable et que 3 élèves obtiennent un score inférieur lors de la seconde production. Cette évolution apparaît toutefois variable selon les critères évalués. Le vocabulaire constitue le critère présentant l'évolution la plus favorable,

alors que l'organisation des textes demeure globalement stable. Les résultats relatifs à la grammaire, au contenu et à l'orthographe apparaissent plus contrastés, avec des progressions et des régressions selon les élèves.

Ces observations suggèrent qu'une partie des élèves semble avoir amélioré ses productions initiales entre les deux cycles. Toutefois, cette évolution ne peut être attribuée directement au dispositif d'évaluation par les pairs, puisque les feedbacks et la révision de la production finale n'ont pas pu être réalisés lors de ce second cycle. Plusieurs facteurs peuvent avoir contribué aux changements observés, tels que la progression des apprentissages, la familiarisation avec les tâches d'écriture ou encore l'expérience acquise lors du premier cycle d'évaluation par les pairs.

Cette analyse complémentaire apporte ainsi un éclairage supplémentaire à l'étude des évolutions des productions écrites, tout en nécessitant une interprétation prudente quant aux facteurs explicatifs de ces évolutions.

1.4 Discussion à propos de la sous-question 1

L'analyse des productions révèle que la majorité des élèves ont apporté des ajustements à leurs travaux entre la version initiale et la version finale, tant en sciences qu'en anglais. Les changements observés prennent principalement la forme d'ajouts et de reformulations, tandis que les suppressions demeurent plus marginales. Ces résultats suggèrent que les élèves se sont engagés dans un processus de révision de leurs productions à la suite des feedbacks reçus de leurs pairs.

Ces observations rejoignent les travaux de Sadler (1989), selon lesquels le feedback permet aux apprenants d'identifier les écarts entre leur production actuelle et les attentes visées. Dans cette perspective, la phase de révision constitue un élément essentiel du processus d'apprentissage puisqu'elle offre aux élèves l'opportunité d'ajuster leur travail à partir des informations reçues. Les modifications observées entre les deux versions des productions témoignent ainsi d'une prise en compte des retours formulés lors de l'évaluation par les pairs.

Ces résultats s'inscrivent également dans la conception de l'évaluation formative développée par Black et Wiliam (1998 cités par Reinholz, 2016), pour qui le feedback constitue un outil de régulation des apprentissages. En invitant les élèves à revenir sur leur production afin d'y apporter des modifications, le dispositif d'évaluation par les pairs semble avoir favorisé une démarche réflexive et un travail de révision. Cette idée rejoint également les travaux de Hattie et Timperley (2007 cités par Lafontaine et Toczec-Capelle, 2023), qui considèrent le feedback

comme un levier permettant d'orienter l'apprenant vers des actions susceptibles de faire progresser son travail.

Par ailleurs, les résultats obtenus peuvent être mis en relation avec le modèle du *Peer Assessment Cycle* proposé par Reinholz (2016). Selon cet auteur, l'évaluation par les pairs ne se limite pas à la production de feedbacks ; elle s'inscrit dans un processus plus large comprenant la réception des commentaires, leur interprétation et leur utilisation lors de la révision de la production. Les modifications observées dans les travaux des élèves semblent illustrer cette phase d'utilisation du feedback, qui constitue une étape centrale du processus.

Toutefois, l'ampleur des modifications varie sensiblement d'un élève à l'autre. Alors que certains élèves réalisent plusieurs changements entre les deux versions de leur travail, d'autres n'apportent que peu de modifications, voire aucune dans un cas observé en sciences et en anglais. Cette variabilité rappelle que la réception d'un feedback ne conduit pas systématiquement à une révision approfondie de la production. Comme le soulignent Cui et Schunn (2024), les feedbacks peuvent être partiellement compris, interprétés de manière différente ou encore ne pas être exploités par les élèves. Dans le même sens, Allal et Mottier Lopez (2005) mettent en évidence que la régulation des apprentissages dépend non seulement des informations reçues mais également de la capacité de l'apprenant à les mobiliser pour transformer son travail.

Ainsi, les résultats de cette première sous-question montrent que le dispositif d'évaluation par les pairs a donné lieu à un travail de révision des productions dans les deux disciplines étudiées. Les élèves ont effectivement apporté des modifications à leurs réalisations, même si l'ampleur de ces changements demeure variable selon les individus et les tâches réalisées.

2. Sous-question 2 : Quelles caractéristiques présentent les feedbacks formulés par les élèves lors de l'évaluation par les pairs ?

Au-delà des modifications observées dans les productions, il apparaît nécessaire de s'intéresser aux feedbacks eux-mêmes afin de mieux comprendre la nature des commentaires échangés entre les élèves lors du dispositif d'évaluation par les pairs.

Afin d'examiner les caractéristiques des feedbacks produits, ceux-ci ont été analysés à l'aide d'une grille (voir annexe 7) et d'un tableau de synthèse présentant les résultats du double codage (voir annexe 10). L'analyse présentée ci-dessous vise à identifier les principales tendances observées dans les feedbacks formulés par les élèves en sciences et en anglais.

2.1 Les caractéristiques des feedbacks produits en sciences

À partir des différentes dimensions retenues dans la grille d'analyse (précision, référence aux critères, propositions de pistes d'amélioration, clarté, respect et justification), il est possible de dégager plusieurs tendances concernant la qualité et la nature des feedbacks produits.

2.1.1 Des feedbacks généralement précis et en lien avec les critères

Lors de la première tâche, la précision constitue la dimension la mieux évaluée avec un score moyen de 2,56/3. Ce score demeure relativement élevé lors de la deuxième tâche (1,93/3). De même, les feedbacks font régulièrement référence aux critères d'évaluation, avec des scores moyens de 2/3 pour la première tâche et de 2,04/3 pour la seconde.

Ces résultats suggèrent que les élèves parviennent généralement à cibler les éléments attendus dans la production et à formuler des commentaires en lien avec les critères d'évaluation proposés dans leur grille.

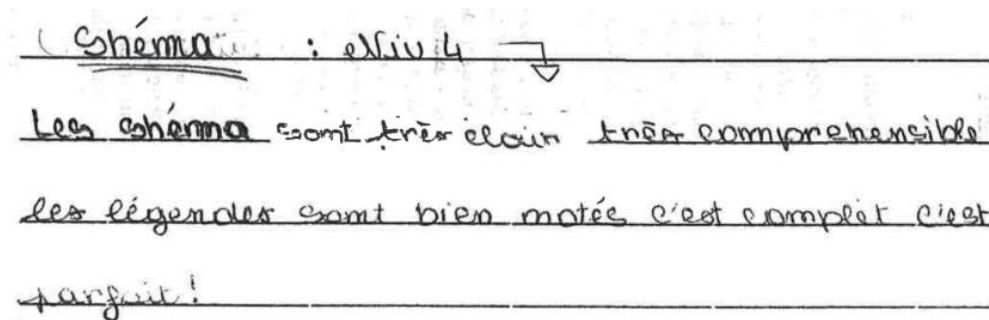


Schéma : élève 4 ↓
Les schéma sont très clairs très compréhensibles
les légendes sont bien notées c'est complet c'est
parfait!

Figure 8. Extrait d'un feedback illustrant un feedback précis (élève 10, tâche 1).

2.1.2 Des feedbacks globalement clairs et respectueux

Les dimensions relatives à la clarté et au respect obtiennent également des scores relativement élevés. La clarté des feedbacks atteint une moyenne de 2,26/3 lors de la première tâche et de 1,54/3 lors de la seconde. Le respect obtient, quant à lui, des scores moyens de 2,15/3 puis de 1,79/3.

Dans l'ensemble, les élèves semblent avoir adopté une posture bienveillante d'évaluateur lors du dispositif d'évaluation par les pairs.

Pour les observations je met 2 car ce n'est pas complet car me voit pas de que si en mélange ni te quantifier...

Figure 9. Extrait d'un feedback illustrant un feedback respectueux, mais peu chaleureux (élève 7, tâche 1).

2.1.3 Des justifications présentes mais des pistes d'amélioration peu développées

Les résultats montrent toutefois des différences importantes selon les dimensions analysées. Les justifications obtiennent des scores moyens de 1,9/3 lors de la première tâche et de 1,21/3 lors de la deuxième tâche. Les élèves expliquent donc parfois leurs remarques, mais ces explications restent peu développées.

Les propositions de pistes d'amélioration constituent la dimension la plus faible dans les deux tâches. Elles obtiennent un score moyen de seulement 0,6/3 lors de la première tâche et de 0,46/3 lors de la seconde. Si les élèves parviennent généralement à identifier les éléments améliorables dans les productions de leurs pairs, ils éprouvent davantage de difficultés à formuler des suggestions concrètes permettant de les actionner directement.

pour les observation: tu es a 3 je te conseil de être précis en globalité mais c'est très bien déjà

Figure 10. Extrait d'un feedback illustrant une suggestion de piste d'amélioration implicite (élève 3, tâche 1).

2.1.4 Une qualité des feedbacks plus élevée lors de la première tâche

Enfin, une légère diminution de la qualité des feedbacks est observée entre les deux tâches. La plupart des dimensions analysées présentent des scores inférieurs lors de la deuxième activité. Alors que la première tâche ne comporte en moyenne aucun feedback de niveau faible et un feedback de niveau bon, la deuxième tâche présente en moyenne un feedback faible et aucun feedback bon.

Cette évolution suggère que la qualité des feedbacks produits par les élèves n'est pas constante et peut varier selon les tâches proposées ou les conditions dans lesquelles l'activité est réalisée.

• conclusion : tu es sûr je trouve que ton raisonnement
 n'est pas justifié ni y précis alors oui
 le chimique est juste mais je n'ai pas
 le pourquoi tu pourrais t'améliorer la dessus
 aussi mais un bon début

Figure 11. Extrait d'un feedback illustrant une justification pertinente (élève 3, tâche 1).

observation : 3 bien
 enemas : 1 pour la participation :)

Figure 12. Extrait d'un feedback illustrant un feedback peu élaboré (élève 2, tâche 2).

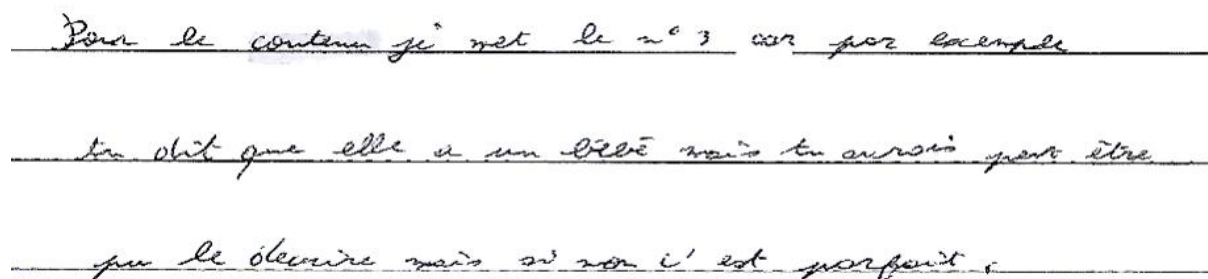
2.2 Les caractéristiques des feedbacks produits en anglais

Comme les résultats observés en sciences, les feedbacks ont été étudiés selon plusieurs dimensions liées à leur qualité, notamment leur précision, leur référence aux critères d'évaluation, leur clarté, leur respect, leur justification ainsi que leur capacité à orienter l'élève vers des pistes d'amélioration. L'analyse met en évidence plusieurs caractéristiques récurrentes dans les commentaires produits par les élèves.

2.2.1 Des feedbacks généralement précis et en lien avec les critères

L'analyse des feedbacks produits dans le cadre de la tâche d'anglais montre que les élèves sont globalement capables d'identifier les éléments à commenter dans les productions de leurs pairs. La précision obtient un score moyen de 2,17/3, tandis que la référence aux critères atteint une moyenne de 1,85/3. Ces résultats indiquent que les commentaires formulés s'appuient généralement sur des éléments spécifiques des productions plutôt que sur des appréciations générales.

Les élèves semblent ainsi en mesure de mobiliser les critères proposés pour guider leur évaluation et cibler certains aspects du texte tels que le contenu, l'organisation, le vocabulaire, la grammaire ou l'orthographe. Les feedbacks produits témoignent donc d'une certaine appropriation du cadre d'évaluation mis à leur disposition.



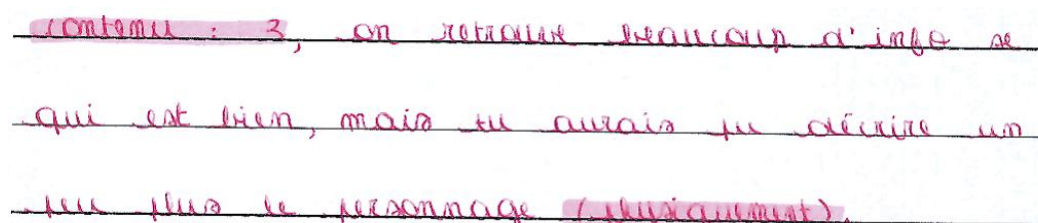
Pour le contenu j'ai noté le 3 car par exemple
tu dis que elle a un bébé mais tu aurais pu être
pu le décrire mais si non c'est parfait.

Figure 13. Extrait d'un feedback illustrant un feedback précis avec une référence modérée aux critères (élève 3, tâche 1).

2.2.2 Des feedbacks globalement clairs et respectueux

Les dimensions liées à la clarté et au respect obtiennent également des valeurs relativement satisfaisantes, avec des moyennes respectives de 1,9/3 et 1,92/3. Dans l'ensemble, les commentaires formulés par les élèves apparaissent compréhensibles et énoncés dans un registre respectueux à l'égard des productions évaluées.

Ces résultats suggèrent que les élèves ont généralement adopté une posture constructive lors de l'activité d'évaluation par les pairs. Les feedbacks formulés permettent le plus souvent d'identifier l'élément concerné ; ils ne contiennent que rarement des remarques susceptibles d'être perçues comme dévalorisantes ou inappropriées.



contenu : 3, on retient beaucoup d'infos
qui est bien, mais tu aurais pu décrire un
peu plus le personnage (physiquement).

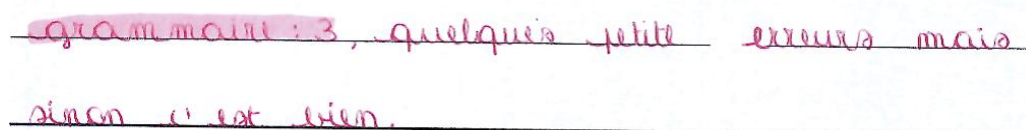
Figure 14. Extrait d'un feedback illustrant un feedback très clair (élève 7, tâche 1).

2.2.3 Des justifications présentes mais des pistes d'amélioration peu développées

Les résultats mettent toutefois en évidence certaines limites dans les feedbacks produits. Si les justifications obtiennent un score moyen de 1,65/3, les propositions de pistes d'amélioration demeurent particulièrement faibles avec une moyenne de 0,57/3. Les élèves expliquent donc parfois leurs observations, mais ils proposent rarement des actions concrètes permettant à leurs pairs d'améliorer leur production ou alors celles-ci sont implicitement proposées.

Cette tendance suggère que les élèves parviennent davantage à identifier les forces ou les faiblesses d'un texte qu'à formuler des recommandations précises pour le retravailler. Ainsi, les

commentaires prennent plus souvent la forme d'un constat ou d'une évaluation que d'un accompagnement explicite vers l'amélioration de la production.



grammaire : 3, quelques petite erreurs mais
sinon c'est bien.

Figure 15. Extrait d'un feedback illustrant une suggestion de piste d'amélioration implicite (élève 7, tâche 1).

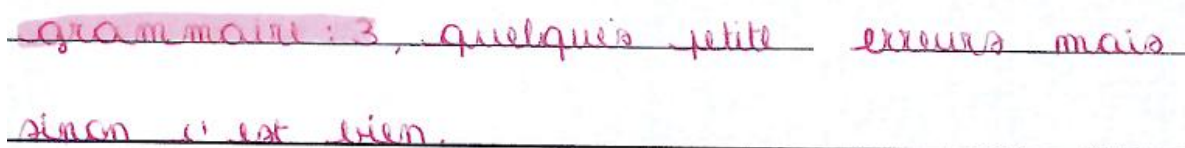
2.2.4 Tendances générales observées dans les feedbacks d'anglais

L'analyse globale des feedbacks montre que les élèves produisent majoritairement des commentaires de qualité intermédiaire.

En moyenne, chaque production a reçu 5 feedbacks, correspondant aux différentes dimensions évaluées (contenu, organisation, vocabulaire, grammaire et orthographe).

Au total, parmi les 8 feedbacks reçus, 7 sont classés dans la catégorie des feedbacks moyens et 1 dans celle des feedbacks faibles. Aucun feedback n'atteint le niveau des feedbacks considérés comme bons.

Ce constat suggère que les élèves parviennent généralement à formuler des commentaires compréhensibles, relativement précis et en lien avec les critères d'évaluation. Toutefois, les faibles scores obtenus pour les propositions de pistes d'amélioration et, dans une moindre mesure pour les justifications, indiquent que les feedbacks restent souvent descriptifs et peu développés.



grammaire : 3, quelques petite erreurs mais
sinon c'est bien.

Figure 16. Extrait d'un feedback illustrant un feedback de qualité moyenne (élève 7, tâche 1).

2.3 Discussion à propos de la sous-question 2

Les résultats obtenus rejoignent en partie les travaux de Nicol et Macfarlane-Dick (2006), qui soulignent l'importance de feedbacks clairs et fondés sur des critères explicites afin de soutenir les apprentissages. Dans notre étude, les élèves semblent effectivement s'appuyer sur les critères mis à leur disposition pour orienter leurs commentaires. Cette mobilisation, relativement

fréquente des critères, peut témoigner d'une appropriation des attentes liées aux tâches proposées et contribuer à rendre les feedbacks plus compréhensibles pour leurs destinataires.

Cette observation peut également être mise en relation avec les travaux de Ketonen et al. (2020), qui montrent que l'explicitation des critères favorise la production de feedbacks de meilleure qualité, particulièrement dans les activités scientifiques. Dans le cadre de cette recherche, les élèves disposaient de grilles d'évaluation et de critères clairement définis pour guider leurs observations. Ces outils ont vraisemblablement facilité l'identification des éléments à commenter et contribué à la précision relativement élevée des feedbacks observés.

Par ailleurs, les résultats pourraient également être influencés par la formation préalable au feedback constructif organisée avant la mise en œuvre du dispositif. Cette préparation visait notamment à familiariser les élèves avec les principes d'un feedback constructif et à les sensibiliser à l'utilisation des critères d'évaluation. Il est donc possible que cette phase de préparation ait favorisé la formulation de commentaires plus ciblés, plus respectueux et davantage ancrés dans les attentes des tâches.

Toutefois, les résultats mettent également en évidence certaines difficultés. Les justifications restent modérément présentes dans les feedbacks et les propositions de pistes d'amélioration obtiennent les résultats les plus faibles dans l'ensemble des tâches analysées. Les élèves semblent donc davantage capables d'identifier un point fort ou un élément à améliorer plutôt que de proposer des actions concrètes permettant à leurs pairs de progresser. Ce constat rejoint les observations d'Orsmond et al. (2002, cités par Mumpuni et al., 2022), selon lesquelles les feedbacks les plus efficaces sont ceux qui comportent non seulement une évaluation de la production, mais également des justifications et des suggestions précises d'amélioration.

Ces résultats peuvent également être rapprochés des travaux de Carless et Boud, 2018 ; Double et al., 2020 ; Bader et al., 2019 cités par Marion et al., 2022, qui soulignent que la production de feedbacks pertinents constitue une compétence complexe nécessitant un apprentissage spécifique. Identifier une faiblesse dans une production apparaît souvent plus accessible que formuler une recommandation précise et exploitable permettant de la corriger.

Enfin, les limites observées peuvent également s'expliquer par le sentiment de légitimité parfois réduit des élèves lorsqu'ils sont amenés à évaluer leurs pairs. Cheng et Warren (1997) ainsi que Falchikov (1995) (cités par Leenknecht et Prins, 2017) mettent en évidence que certains apprenants hésitent à formuler des critiques ou des recommandations détaillées par crainte de se tromper ou de porter un jugement inadéquat sur le travail d'un camarade. Cette réticence

pourrait contribuer à expliquer pourquoi les élèves formulent davantage des constats que de véritables suggestions d'amélioration.

Dans l'ensemble, les résultats montrent que les élèves sont capables de produire des feedbacks globalement pertinents, notamment grâce à leur précision, leur clarté et leur ancrage dans les critères d'évaluation. Néanmoins, les commentaires demeurent souvent descriptifs et peinent à atteindre un niveau plus élaboré caractérisé par des justifications développées et des pistes concrètes d'amélioration. Ces constats suggèrent que si les élèves semblent s'être approprié certaines composantes du feedback constructif, un accompagnement supplémentaire pourrait être nécessaire pour renforcer leur capacité à produire des feedbacks pleinement formatifs.

3. Sous-question 3 : Quels liens peut-on observer entre les feedbacks reçus et les modifications apportées aux productions écrites ?

3.1 Les liens entre les feedbacks et les modifications en sciences

L'analyse du lien entre les feedbacks reçus et les modifications apportées aux productions vise à examiner dans quelle mesure les commentaires formulés par les pairs sont effectivement mobilisés par les élèves lors de la révision de leur travail. Cette analyse a été réalisée à l'aide du document cadre (voir annexe 8) et du tableau de synthèse présentant les résultats du double codage (voir annexe 12). Elle permet d'identifier les types de relations observées entre les feedbacks et les modifications, ainsi que les différentes modalités d'appropriation de ces commentaires.

3.1.1 Les relations observées entre les feedbacks et les modifications

Pour la première tâche de sciences, 31 modifications ont été recensées dans les 11 productions analysées. 41,9% des modifications présentent un lien direct avec les feedbacks reçus, 6,5% montrent un lien partiel et 51,6% n'ont aucun lien identifiable avec les commentaires formulés par les pairs. Plus de la moitié des modifications réalisées ne semblent donc pas directement liées aux feedbacks reçus.

Lors de la deuxième tâche, 21 modifications ont été observées dans les 9 productions analysées. Parmi celles-ci, 52,4% des modifications entretiennent un lien direct avec les feedbacks reçus, 23,8% présentent un lien partiel et 23,8% ne peuvent être reliées aux commentaires formulés par les pairs. Par ailleurs, une des neuf productions n'a fait l'objet d'aucune modification entre la version initiale et la version finale, malgré les feedbacks reçus.

Ces résultats mettent en évidence que certaines modifications peuvent être directement ou partiellement reliées aux feedbacks reçus, tandis que d'autres ne présentent aucun lien identifiable avec les commentaires formulés par les pairs. La comparaison des deux tâches montre également que, si le nombre total de modifications est moins élevé lors de la deuxième activité, les liens directs et partiels y sont proportionnellement plus nombreux, tandis que les modifications sans lien identifiable sont moins fréquentes.

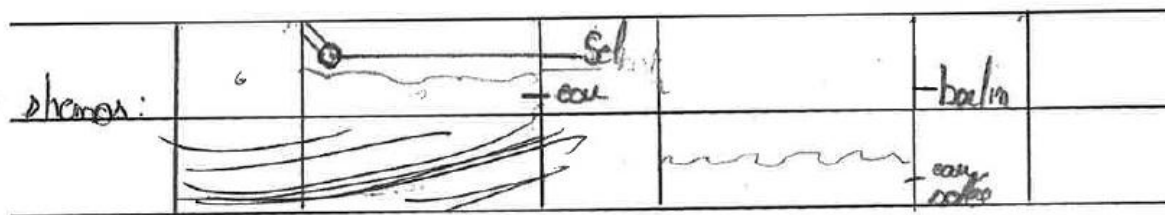


Figure 17. Extrait de la version initiale de la production (élève 1, tâche 2).

Le schéma pour ton schéma je te mes le n° 2 car c'est plus ou moins complet mais difficile à comprendre et pas assez clair ; peu de détails mais quand même les annotations correctes.

Figure 18. Extrait du feedback reçu par l'élève (élève 1, tâche 2).

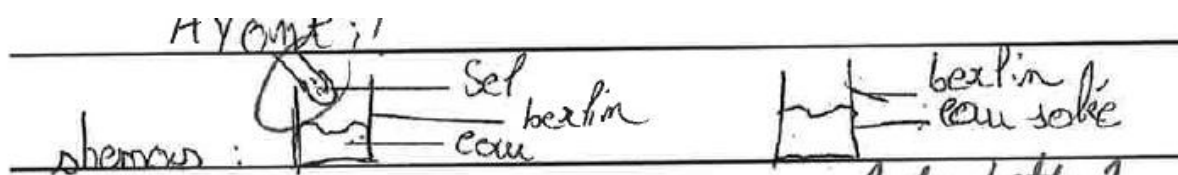


Figure 19. Extrait de la version finale de la production (élève 1, tâche 2).

3.1.2 Les profils d'appropriation des feedbacks

L'analyse met également en évidence une diversité de profils d'appropriation des feedbacks.

Lors de la première tâche, 27,3% des élèves relèvent du profil 1, caractérisé par des feedbacks de bonne qualité accompagnés d'une forte appropriation. 27,3% des élèves appartiennent au profil 2, correspondant à des feedbacks de bonne qualité mais à une appropriation plus limitée. 9,1% des élèves relèvent du profil 3, caractérisé par des feedbacks de faible qualité et une faible appropriation, et 9,1% appartiennent au profil 4, associant des feedbacks de qualité moyenne à

des améliorations satisfaisantes. Enfin, 27,3% présentent un profil 5, marqué par des feedbacks de qualité moyenne et une appropriation limitée.

La répartition observée lors de la deuxième tâche diffère sensiblement. Aucun élève ne relève du profil 1. 11,1% des élèves appartiennent au profil 2, 11,1% au profil 3 et 55,6% au profil 4. 11,1% des élèves présentent un profil 5, caractérisé par des feedbacks de qualité moyenne et une appropriation limitée. Enfin, 11,1% des élèves relèvent du profil 6, caractérisé par une appropriation différenciée selon la qualité des feedbacks reçus.

3.1.3 Synthèse des observations

Dans l'ensemble, les résultats mettent en évidence une appropriation variable des feedbacks par les élèves. Si plusieurs modifications peuvent être directement reliées aux commentaires reçus, d'autres ne présentent qu'un lien partiel, voire aucun lien identifiable. Les différents profils observés montrent que les modalités d'appropriation des feedbacks diffèrent d'un élève à l'autre. Certains semblent mobiliser directement les remarques formulées par leurs pairs, tandis que d'autres n'en reprennent qu'une partie ou choisissent de modifier d'autres aspects de leur production. Un cas met également en évidence une appropriation différenciée des feedbacks selon leur qualité, les commentaires les plus précis et les plus actionnables donnant lieu aux modifications les plus pertinentes.

3.2 Les liens entre les feedbacks et les modifications en anglais

L'analyse du lien entre les feedbacks reçus et les modifications apportées aux productions d'anglais vise à examiner dans quelle mesure les commentaires formulés par les pairs sont mobilisés par les élèves lors de la révision de leur production écrite. Comme pour les sciences, les modifications observées ont été mises en relation avec les feedbacks reçus afin d'identifier des liens directs, partiels ou absents, ainsi que les différentes modalités d'appropriation des commentaires.

3.2.1 Les relations observées entre les feedbacks et les modifications

Pour la tâche d'anglais, 32 modifications ont été recensées dans les 8 productions analysées. Parmi celles-ci, 28,1% sont en lien direct avec les feedbacks reçus, 21,9% présentent un lien partiel et 50,0% n'ont aucun lien identifiable avec les commentaires formulés par les pairs. Ainsi, la moitié des modifications observées ne semblent pas directement liées aux feedbacks reçus. Par ailleurs, une des huit productions n'a fait l'objet d'aucune modification entre la version initiale et la version finale, malgré les feedbacks reçus.

Ces résultats montrent que les feedbacks reçus sont à l'origine d'une partie des modifications apportées aux productions. Toutefois, la moitié des modifications observées ne présentent aucun lien identifiable avec les commentaires formulés par les pairs, ce qui traduit une appropriation variable des feedbacks au sein du groupe.

over and he wears blue. She is a 3 brothers, rocha,
tom en victoire, rocha is 3, tom is 5 en victoire & there

Figure 20. Extrait de la version initiale de la production (élève 9, tâche 1).

Grammaire / Syntaxe: tu as confondu avoir et être
dans une phrase tu dis "elle est 3 frères"

Figure 21. Extrait du feedback reçu par l'élève (élève 9, tâche 1).

She has a 3 brothers, Rocha, Tom and Victore.

Figure 22. Extrait de la version finale de la production (élève 9, tâche 1).

3.2.2 Les profils d'appropriation des feedbacks

L'analyse qualitative met en évidence plusieurs profils d'appropriation des feedbacks. Pour la tâche d'anglais, 50,0% des élèves présentent un profil 5, caractérisé par des feedbacks de qualité moyenne et une appropriation limitée. Les profils 2 et 4 concernent chacun 25,0% des élèves. Aucun élève ne relève des profils 1, 3 ou 6.

3.2.3 Synthèse des observations

Dans l'ensemble, les résultats mettent en évidence une appropriation hétérogène des feedbacks par les élèves. Si certaines modifications peuvent être directement ou partiellement reliées aux commentaires reçus, une proportion importante des changements observés ne présente aucun lien explicite avec les feedbacks formulés. Les profils identifiés montrent également que les modalités d'appropriation diffèrent selon les élèves. Si certains mobilisent une partie des remarques formulées par leurs pairs, d'autres ne les exploitent que de manière limitée ou choisissent de modifier leur production sur d'autres aspects.

3.3 Discussion à propos de la sous-question 3

Les observations réalisées rejoignent les travaux de Sadler (1989), selon lesquels le feedback n'a de valeur formative que s'il permet à l'élève d'identifier l'écart entre sa production actuelle et la production attendue, puis d'agir pour réduire cet écart. Dans notre dispositif, les nombreux liens directs observés entre les feedbacks et les modifications montrent que plusieurs élèves se sont effectivement appuyés sur les commentaires reçus pour réviser leur production.

Les données recueillies sont également cohérentes avec le modèle proposé par Hattie et Timperley (2007 cités par Lafontaine et Toczec-Capelle, 2023), qui considèrent que l'efficacité du feedback réside dans sa capacité à orienter l'élève vers des actions concrètes d'amélioration. Les productions analysées montrent que certains feedbacks ont effectivement conduit à des modifications ciblées, ce qui suggère qu'ils ont joué un rôle de levier dans le processus de révision. Toutefois, cette influence demeure variable selon les élèves et selon les feedbacks reçus.

Les observations réalisées semblent également illustrer le fonctionnement du *assessment cycle* décrit par Reinholz (2016). Après la production initiale, la réception du feedback et la phase de révision ont conduit plusieurs élèves à modifier leur travail. Le fait que certaines modifications puissent être directement reliées aux commentaires formulés par les pairs montre que cette boucle de rétroaction a effectivement été mobilisée par une partie des élèves.

Les résultats constatés lors de la deuxième tâche de sciences apparaissent, par ailleurs, légèrement plus favorables que ceux de la première. Les liens directs et partiels entre les feedbacks reçus et les modifications apportées sont proportionnellement plus nombreux, tandis que les changements sans lien identifiable sont moins fréquents. Cette évolution pourrait traduire une appropriation progressive du dispositif par les élèves.

Cette observation semble aller dans le sens du modèle de Reinholz (2016), selon lequel les bénéfices de l'évaluation par les pairs reposent sur la répétition des différentes étapes du cycle. La possibilité offerte aux élèves de réinvestir les critères d'évaluation, de produire de nouveaux feedbacks et de réviser une nouvelle production en sciences pourrait ainsi favoriser le développement progressif de leurs compétences évaluatives ainsi que de leur capacité à exploiter les commentaires reçus.

Cette interprétation est également cohérente avec le dispositif mis en place dans cette recherche, qui comprenait deux cycles successifs précisément afin de permettre aux élèves de réinvestir

les apprentissages réalisés tant dans leurs productions que dans la formulation et l'exploitation des feedbacks.

Les analyses montrent également que certains feedbacks ne donnent lieu à aucune modification, voire que deux productions n'ont fait l'objet d'aucune révision malgré les commentaires reçus. Ces observations rejoignent les travaux de Cui et Schunn (2024), qui montrent que les élèves n'exploitent pas systématiquement les feedbacks qui leur sont adressés. Les commentaires peuvent être ignorés, jugés peu utiles ou encore mal compris, ce qui limite leur impact sur les productions.

Les résultats de cette troisième sous-question peuvent également être rapprochés de ceux obtenus lors de l'analyse de la qualité des feedbacks. Les commentaires produits par les élèves étaient généralement précis et en lien avec les critères, mais comportaient peu de justifications et surtout très peu de pistes d'amélioration. Cette caractéristique pourrait contribuer à expliquer pourquoi certains feedbacks ont été peu exploités lors de la révision des productions.

L'identification d'un profil, dans lequel la qualité des modifications varie selon la qualité des feedbacks reçus, constitue un résultat particulièrement intéressant. Dans ce cas, les commentaires les plus précis et les plus actionnables sont ceux qui donnent lieu aux améliorations les plus pertinentes. À l'inverse, les feedbacks plus généraux ou moins développés sont peu exploités. Cette observation renforce l'idée selon laquelle la qualité du feedback constitue un facteur susceptible d'influencer son appropriation par les élèves.

Dans l'ensemble, cette troisième sous-question montre que l'évaluation par les pairs peut effectivement favoriser la révision des productions écrites lorsque les feedbacks sont mobilisés par les élèves. Toutefois, cette appropriation demeure variable et semble dépendre à la fois des caractéristiques des commentaires reçus et de la manière dont chaque élève choisit de les intégrer dans son processus de révision. Ces résultats soulignent ainsi que l'efficacité de l'évaluation par les pairs ne repose pas uniquement sur la production de feedbacks, mais également sur leur appropriation par les étudiants.

4. Sous-question 4 : Ces résultats se manifestent-ils de manière différente dans les cours de sciences et d'anglais ?

Cette dernière sous-question vise à adopter une perspective transversale afin d'identifier les éléments communs et les spécificités observés dans les deux contextes disciplinaires. L'objectif n'est plus de présenter de nouveaux résultats, mais de mettre en perspective les analyses précédentes afin de mieux comprendre dans quelle mesure les effets de l'évaluation par les pairs

varient selon la discipline. Cette analyse a été réalisée à l'aide d'un tableau de synthèse reprenant les principaux résultats de trois premières sous-questions pour chacune des deux disciplines (voir annexe 13).

4.1 Des mécanismes communs dans les deux disciplines

Les analyses réalisées dans les cours de sciences et d'anglais mettent en évidence plusieurs mécanismes communs liés au fonctionnement du dispositif d'évaluation par les pairs.

Tout d'abord, dans les deux disciplines, la majorité des élèves a apporté des modifications à sa production entre la version initiale et la version finale. Bien que l'ampleur des révisions varie d'un élève à l'autre, les résultats montrent que les élèves se sont engagés dans un processus de relecture et de révision de leur travail à la suite des feedbacks reçus. Cette observation rejoint les travaux de Sadler (1989) selon lesquels le feedback prend tout son sens lorsqu'il conduit l'apprenant à réviser sa production afin de réduire l'écart entre celle-ci et les attentes visées.

Par ailleurs, les analyses des feedbacks révèlent des tendances similaires dans les deux disciplines. Les commentaires formulés par les élèves sont généralement précis, compréhensibles et s'appuient sur les critères d'évaluation mis à leur disposition. En revanche, les justifications restent relativement peu développées et les propositions de pistes d'amélioration demeurent limitées. Cette difficulté apparaît de manière récurrente, indépendamment du contexte disciplinaire.

Enfin, les résultats montrent que l'appropriation des feedbacks varie fortement d'un élève à l'autre, tant en sciences qu'en anglais. Certains élèves exploitent directement les commentaires reçus pour améliorer leur production, tandis que d'autres n'en mobilisent qu'une partie, voire ne les utilisent pas. Cette variabilité, déjà mise en évidence dans les trois premières sous-questions, souligne que la réception d'un feedback ne conduit pas systématiquement à une révision de la production.

Dans l'ensemble, ces différents constats montrent que les principaux mécanismes associés à l'évaluation par les pairs semblent se retrouver dans les deux disciplines étudiées. Les élèves révisent leurs productions, produisent des feedbacks globalement pertinents et s'approprient ces derniers de manière variable, quel que soit le contexte d'enseignement.

4.2 Des différences liées aux spécificités disciplinaires

Si les mécanismes observés apparaissent largement similaires, les analyses mettent néanmoins en évidence certaines différences qui semblent davantage liées aux caractéristiques des disciplines qu'au fonctionnement du dispositif lui-même.

En sciences, les élèves cherchent à compléter leurs productions, à préciser les phénomènes observés ou à améliorer la validité scientifique de leurs explications. Les ajouts constituent ainsi le type de modification le plus fréquent.

À l'inverse, en anglais, les révisions concernent davantage l'organisation du texte, le vocabulaire, la grammaire ainsi que la reformulation des idées. Les élèves s'attachent principalement à améliorer la qualité de l'expression écrite et la cohérence de leur production. Les reformulations y occupent une place plus importante que dans les productions scientifiques.

Ces différences s'expliquent probablement par les attentes propres à chaque discipline. Comme le soulignent Anker-Hansen et Andrée (2019), les pratiques d'évaluation et les modalités de feedback sont étroitement liées aux objectifs d'apprentissage poursuivis. En sciences, la qualité d'une production repose sur la précision des observations, la rigueur des explications et la validité des contenus scientifiques. En anglais, les critères portent davantage sur l'organisation du discours, la maîtrise de la langue et la qualité de l'argumentation. Les élèves révisent donc prioritairement les éléments qui correspondent aux exigences propres à leur discipline.

Ainsi, les différences observées concernent moins l'efficacité de l'évaluation par les pairs que les objets sur lesquels porte cette évaluation. Le dispositif semble fonctionner selon une logique similaire dans les deux contextes, tout en s'adaptant naturellement aux attentes disciplinaires.

4.3 Ce que cette comparaison apporte à la compréhension de l'évaluation par les pairs

La comparaison entre les deux disciplines apporte un éclairage complémentaire sur le fonctionnement de l'évaluation par les pairs. Alors que les trois premières sous-questions ont permis d'analyser séparément les productions, les feedbacks et leur appropriation, cette mise en perspective montre que les effets observés ne semblent pas dépendre principalement de la discipline enseignée.

Les résultats suggèrent en effet que l'évaluation par les pairs constitue avant tout un processus de régulation des apprentissages reposant sur des mécanismes relativement transversaux. Quel que soit le contexte disciplinaire, les élèves sont amenés à analyser une production, à mobiliser

des critères d'évaluation, à formuler un feedback puis à décider dans quelle mesure ils souhaitent exploiter les commentaires reçus pour améliorer leur propre travail.

Cette comparaison conduit également à nuancer l'idée selon laquelle l'efficacité de l'évaluation par les pairs dépendrait essentiellement de la discipline. Les résultats de cette recherche montrent plutôt que ce sont les objets évalués qui diffèrent, tandis que les mécanismes de production, de réception et d'exploitation des feedbacks restent comparables. Dès lors, l'efficacité du dispositif semble davantage liée à la qualité des feedbacks, à l'explicitation des critères d'évaluation ainsi qu'à l'accompagnement proposé aux élèves plutôt qu'au contexte disciplinaire lui-même.

Dans cette perspective, l'évaluation par les pairs apparaît comme un dispositif pouvant être mobilisé dans différents contextes d'enseignement, à condition d'être adapté aux objectifs spécifiques de chaque discipline et d'accompagner progressivement les élèves dans le développement de leurs compétences évaluatives. Cette dernière sous-question met ainsi en évidence que, malgré des productions et des attentes disciplinaires différentes, les principes qui sous-tendent l'évaluation par les pairs demeurent largement communs.

Limites et perspectives

1. Les limites de l'étude

Comme toute recherche, cette étude présente certaines limites qu'il convient de prendre en considération afin de situer la portée des résultats obtenus. Celles-ci concernent à la fois le dispositif mis en œuvre, l'analyse des données recueillies ainsi que l'interprétation des résultats.

La première limite est d'ordre méthodologique et concerne l'échantillon de l'étude. Cette recherche a été menée auprès d'un nombre restreint d'élèves (11) issus de deux classes d'une même école de l'enseignement qualifiant en Fédération Wallonie-Bruxelles. Si ce contexte a permis de réaliser une analyse approfondie des productions écrites et des feedbacks, il limite toutefois la possibilité de généraliser les résultats à d'autres établissements, niveaux scolaires ou filières d'enseignement.

La durée du dispositif constitue une deuxième limite. Celui-ci s'est déroulé sur une période relativement courte et n'a permis d'observer les effets de l'évaluation par les pairs qu'au travers d'un nombre limité de cycles.

Par ailleurs, le dispositif ne prévoyait pas de temps d'échange entre les élèves à propos des feedbacks formulés. Les commentaires étaient transmis aux auteurs des productions sans qu'un moment de discussion ne soit organisé afin de les clarifier, de les nuancer ou d'en préciser le sens. Cette absence d'interaction a pu limiter la compréhension de certains feedbacks ainsi que leur appropriation par les élèves.

Enfin, le second cycle prévu dans le cours d'anglais n'a pas pu être mené à son terme en raison des grèves survenues durant la période de collecte des données. Cette interruption a réduit les possibilités de comparaison initialement envisagées entre les deux disciplines et n'a pas permis d'observer l'évolution des élèves au fil de plusieurs cycles d'évaluation par les pairs.

L'analyse des données présente également certaines limites. Bien qu'un document cadre ait été élaboré afin d'harmoniser le codage des productions et des feedbacks, une part de subjectivité demeure inhérente à toute analyse qualitative. Certaines décisions ont nécessité une interprétation, en particulier lorsque les feedbacks étaient ambigus ou particulièrement succincts.

Plusieurs difficultés spécifiques sont apparues au cours du codage. Tout d'abord, la distinction entre les justifications et les pistes d'amélioration s'est parfois révélée délicate. Certains commentaires pouvaient être interprétés à la fois comme une explication du jugement formulé

et comme une suggestion implicite visant à améliorer la production, rendant leur catégorisation moins évidente.

Ensuite, certains feedbacks se limitaient à une appréciation très brève, telle que « bien » ou « bon ». Si ces commentaires traduisaient une évaluation positive, ils ne permettaient pas d'identifier précisément les éléments jugés satisfaisants ni d'apprécier leur réelle portée formative. Leur évaluation s'est donc révélée plus complexe au regard des critères retenus.

Une autre difficulté concerne le codage des pistes d'amélioration. Lorsqu'aucune suggestion n'était formulée, cette dimension était systématiquement codée comme absente. Or, cette absence ne traduisait pas nécessairement un manque de qualité du feedback. Dans certaines situations, la production évaluée ne nécessitait tout simplement pas de recommandation particulière. Le codage pouvait alors conduire à attribuer un score plus faible à un feedback pourtant adapté à la qualité de la production. Cette limite est davantage liée au fonctionnement de l'outil d'analyse qu'à la qualité des feedbacks produits par les élèves.

Enfin, plusieurs facteurs extérieurs au dispositif invitent à interpréter les résultats avec prudence. Les améliorations observées entre les versions initiales et finales des productions ne peuvent être attribuées exclusivement à l'évaluation par les pairs. D'autres variables ont également pu influencer les révisions réalisées, notamment la progression naturelle des apprentissages, les apprentissages réalisés au cours de la séquence d'enseignement, le niveau de motivation des élèves ou encore leur familiarisation progressive avec les tâches proposées. Il convient dès lors de considérer les résultats de cette étude comme le produit d'un ensemble de facteurs susceptibles d'interagir. Parmi ceux-ci, l'évaluation par les pairs constitue un élément important, sans pour autant être le seul à expliquer les évolutions observées.

2. Les perspectives de l'étude

Les résultats de cette recherche ouvrent plusieurs perspectives, tant sur le plan de la recherche que sur celui du développement des dispositifs d'évaluation par les pairs.

Une première perspective consiste à approfondir le dispositif mis en œuvre. Une expérimentation menée sur une période plus longue, et comprenant davantage de cycles d'évaluation par les pairs, permettrait d'observer l'évolution progressive des compétences évaluatives des élèves. Il serait notamment intéressant d'examiner si la qualité des feedbacks s'améliore au fil des expériences et si les élèves développent une meilleure capacité à exploiter les commentaires reçus pour réviser leurs productions. Dans cette même perspective, une formation plus approfondie à la formulation du feedback pourrait être intégrée au dispositif.

Celle-ci pourrait accorder une attention particulière à la distinction entre les différents types de commentaires, à la formulation de justifications explicites ainsi qu'à l'élaboration de pistes d'amélioration précises et directement exploitables. Une telle évolution permettrait de mieux comprendre dans quelle mesure un accompagnement plus soutenu favorise le développement des compétences évaluatives. À partir de combien de cycles les élèves produisent-ils des feedbacks plus élaborés ? Le développement de ces compétences suit-il une progression continue ou existe-t-il un seuil à partir duquel elles se stabilisent ?

Une deuxième perspective concerne l'élargissement du contexte de recherche. Cette étude a été réalisée auprès d'un nombre restreint d'élèves issus d'un seul établissement scolaire. Il serait dès lors pertinent de reproduire le dispositif auprès d'un échantillon plus large, incluant plusieurs établissements, différentes filières ainsi que d'autres disciplines d'enseignement. Une telle démarche permettrait d'apprécier la robustesse des résultats obtenus et d'examiner si les effets observés dans cette recherche se maintiennent dans des contextes éducatifs variés ou s'ils diffèrent selon les caractéristiques des élèves, des disciplines ou des établissements.

Au-delà de l'élargissement de l'échantillon, cette recherche invite également à approfondir les connaissances relatives au développement des compétences évaluatives au cours de la scolarité. Il serait intéressant d'examiner comment les capacités à produire, interpréter et mobiliser des feedbacks évoluent en fonction de l'âge et du niveau de scolarisation des élèves. De telles recherches permettraient notamment de mieux comprendre à partir de quel moment les élèves sont en mesure de formuler des feedbacks de qualité et dans quelle mesure ces compétences se développent naturellement au fil de la scolarité ou nécessitent un enseignement explicite. Une autre option serait aussi de faire une distinction plus précise entre les effets liés à la production du feedback et ceux liés à sa réception. Les apprentissages résultent-ils principalement des commentaires reçus ou de l'activité cognitive mobilisée lorsqu'un élève évalue le travail d'un pair ?

Enfin, les évolutions récentes de l'intelligence artificielle générative ouvrent de nouvelles perspectives pour la recherche en évaluation formative. Les outils conversationnels sont aujourd'hui capables de produire des feedbacks personnalisés sur des productions écrites et commencent progressivement à être intégrés dans certains contextes éducatifs. Dès lors, une piste de recherche particulièrement prometteuse consisterait à comparer les effets d'un feedback fourni par un pair à ceux d'un feedback généré par une intelligence artificielle, tant sur l'amélioration des productions écrites que sur le développement des compétences d'autorégulation, de réflexion critique et d'évaluation.

Cette perspective conduit à une interrogation plus fondamentale sur les mécanismes d'apprentissage mobilisés par l'évaluation par les pairs. Si l'intelligence artificielle est capable de produire un feedback de qualité, peut-elle pour autant offrir les mêmes opportunités d'apprentissage que celles qui émergent de l'interaction entre élèves ? Autrement dit, la richesse de l'évaluation par les pairs réside-t-elle principalement dans la qualité du feedback reçu ou dans le processus cognitif et social que représente le fait d'évaluer le travail d'un pair ? À mesure que l'intelligence artificielle s'intègre progressivement aux pratiques éducatives, ces questions apparaissent comme des pistes de recherche particulièrement fécondes pour mieux comprendre ce qui fait la spécificité et la valeur pédagogique de l'évaluation par les pairs.

Conclusion

Cette recherche avait pour objectif d'examiner l'impact d'un dispositif d'évaluation par les pairs sur l'amélioration des productions écrites d'élèves de troisième secondaire de l'enseignement qualifiant, dans deux disciplines distinctes : les sciences et l'anglais.

Dans la continuité des travaux de Black et Harrison (2001), de Nicol et Macfarlane-Dick (2006) ou encore de Hattie et Timperley (2007 cités par Lafontaine et Toczec-Capelle, 2023), qui considèrent le feedback comme un levier essentiel de régulation des apprentissages, les analyses réalisées montrent que le dispositif d'évaluation par les pairs a globalement contribué à l'amélioration des productions écrites des élèves. Les comparaisons entre les versions initiales et finales mettent en évidence des révisions positives, même si l'ampleur des changements varie selon les élèves et les critères d'évaluation mobilisés.

L'étude des feedbacks produits par les élèves révèle également qu'ils sont capables de formuler des rétroactions pertinentes et généralement en lien avec les critères proposés. Toutefois, comme le soulignent Bourgeois et Laveault (2015) ainsi que Marion et al. (2022), un feedback véritablement formatif ne se limite pas à repérer les forces et les faiblesses d'une production. Il doit également orienter l'élève vers des pistes concrètes d'amélioration. Or, les analyses montrent que les justifications demeurent souvent peu développées et que les suggestions d'amélioration constituent la dimension la plus fragile des feedbacks recueillis.

La mise en relation des feedbacks reçus et des modifications apportées aux productions montre que les commentaires des pairs ont effectivement soutenu une partie des révisions réalisées par les élèves. Néanmoins, leur appropriation demeure variable. Certains élèves exploitent directement les remarques reçues, tandis que d'autres ne les mobilisent que partiellement, voire pas du tout. Ces résultats rejoignent les observations de Reinholz (2016), selon lesquelles l'efficacité de l'évaluation par les pairs repose sur un véritable cycle d'apprentissage mobilisant la production, le feedback, la révision et la réflexion. Ils confirment également les travaux plus récents de Cui et Schunn (2024), qui montrent que les élèves ne s'approprient pas systématiquement les feedbacks qui leur sont adressés.

La comparaison entre les cours de sciences et d'anglais apporte un éclairage complémentaire. Les mécanismes de l'évaluation par les pairs apparaissent largement similaires dans les deux disciplines. Les élèves révisent leurs productions, produisent des feedbacks globalement

pertinents et s'approprient ces derniers de manière variable. En revanche, la nature des révisions diffère selon les attentes disciplinaires. En sciences, les modifications portent davantage sur les observations, les schémas et la précision des explications, tandis qu'en anglais elles concernent principalement l'organisation du texte, le vocabulaire, la grammaire et la reformulation des idées. Cette observation rejoint les analyses d'Anker-Hansen et Andrée (2019) qui montrent que les modalités de feedback et les objets d'évaluation sont influencés par les spécificités des disciplines, sans pour autant remettre en question les principes généraux de l'évaluation formative.

Dans l'ensemble, les résultats de cette recherche suggèrent que l'évaluation par les pairs constitue un levier pertinent pour soutenir les apprentissages en production écrite, à condition que les élèves soient accompagnés dans le développement de leurs compétences évaluatives. Au-delà de l'amélioration des productions elles-mêmes, cette étude rappelle que l'évaluation par les pairs s'inscrit dans une conception de l'évaluation au service des apprentissages (Allal & Mottier Lopez, 2005 ; Mottier-Lopez, 2023). Lorsqu'elle est intégrée à un dispositif structuré et accompagnée d'un travail explicite sur les critères et sur la formulation du feedback, elle offre aux élèves l'occasion de porter un regard critique sur les productions, de réviser leur travail et de développer progressivement leur capacité à réguler leurs propres apprentissages.

Les conclusions de cette étude doivent toutefois être interprétées avec prudence. Les effets observés ne sont ni uniformes selon les élèves ni identiques dans toutes les situations d'apprentissage. En outre, le second cycle prévu dans le cours d'anglais n'a pu être mené à son terme, ce qui limite les comparaisons initialement envisagées entre les deux disciplines et invite à considérer les résultats à la lumière de cette contrainte méthodologique.

Bibliographie

Allal, L., & Mottier Lopez, L. (2005). L'évaluation formative de l'apprentissage : Revue de publications en langue française. In OCDE (Éd.), *L'évaluation formative : Pour un meilleur apprentissage dans les classes secondaires* (pp. 265–290). Éditions OCDE.

Anker-Hansen, J., & Andrée, M. (2019). Using and rejecting peer feedback in the science classroom: A study of students' negotiations on how to use peer feedback when designing experiments. *Research in Science & Technological Education*. <https://doi.org/10.1080/02635143.2018.1557628>

Black, P., & Harrison, C. (2001). Self- and peer-assessment and taking responsibility: The science student's role in formative assessment. *School Science Review*, 83(302), 43–49.

Bloxham, S., & West, A. (2004). Understanding the rules of the game: Marking peer assessment as a medium for developing students' conceptions of assessment. *Assessment & Evaluation in Higher Education*, 29(6), 721–733. <https://doi.org/10.1080/0260293042000227254>

Boilevin, J.-M., & Brandt-Pomares, P. (2011). Démarches d'investigation en sciences et technologie au collège : Les conditions d'évolution des pratiques. In M. Grangeat (Éd.), *Les démarches d'investigation dans l'enseignement scientifique : Pratiques de classe, travail collectif enseignant, acquisitions des élèves* (pp. 51–62). École normale supérieure de Lyon, Institut national de recherche pédagogique.

Bourgeois, L., & Laveault, D. (2015). Évaluation par les pairs à l'écrit : Qualité des rétroactions pour soutenir la phase de révision. In *Des études sur la formation à l'évaluation et à l'autoévaluation : Les lieux et les situations de formation à l'école* (pp. 101–117). De Boeck.

Chen, Y. (2024). The effectiveness of peer feedback on second language writing. *Journal of Contemporary Educational Research*, 8(2), 68–79. <http://ojs.bbwpublisher.com/index.php/JCER>

Cui, Y., & Schunn, C. D. (2024). Peer feedback that consistently supports learning to write and read: Providing comments on meaning-level issues. *Assessment & Evaluation in Higher Education*, 49(8), 1168–1181. <https://doi.org/10.1080/02602938.2024.2364025>

Fagnant, A. (2023). *Les pratiques d'évaluation en classe : Des compétences professionnelles pour soutenir l'apprentissage des élèves*. Cnesco-Cnam.

Girouard-Gagné, M., & Durand, M. J. (Éds.). (2025). *Guide pédagogique : Mise en œuvre d'une évaluation en soutien aux apprentissages : Évaluer pour mieux apprendre*. Programme APPRENDRE, Agence universitaire de la Francophonie. <https://programme-apprendre.org>

Jacobs, H. L., Zingraf, S. A., Wormuth, D. R., Hartfiel, V. F., & Hughey, J. B. (1981). *Testing ESL composition: A practical approach*. Newbury House.

Ketonen, L., Häikiöniemi, M., Nieminen, P., & Viiri, J. (2020). Pathways through peer assessment: Implementing peer assessment in a lower secondary physics classroom. *International Journal of Science and Mathematics Education*, 18, 1465–1484. <https://doi.org/10.1007/s10763-019-10030-3>

Lafontaine, D., & Toczec-Capelle, M.-C. (2023). *L'évaluation en classe au service de l'apprentissage des élèves : Rapport de synthèse* (pp. 8–20). Cnesco-Cnam.

Leenknecht, M., & Prins, F. (2017). Formative peer assessment in primary school: The effects of involving pupils in setting assessment criteria on their appraisal and feedback style. *European Journal of Psychology of Education*, 33(1), 101–116. <https://doi.org/10.1007/s10212-017-0340-2>

Lepareur, C. (2016). *L'évaluation dans les enseignements scientifiques fondés sur l'investigation : Effets de différentes modalités d'évaluation formative sur l'autorégulation des apprentissages* [Thèse de doctorat, Université Grenoble Alpes].

Lou, Y., Liu, H., Li, Y., & Yang, J. (2016). A study on guided peer feedback in group work in an English as Foreign Language writing class in China. *Education Journal*, 5(6), 136–141. <https://doi.org/10.11648/j.edu.20160506.12>

Marion, C., Tremblay-Gagnon, D., & Laferrière, T. (2022). Rétroaction par les pairs : Quelles conditions pour un apport optimal aux apprentissages des étudiantes et étudiants ? *Revue internationale du CRIRES*, 6(1), 52–72. <https://doi.org/10.51657/ric.v6i1.51446>

Montpied, P., Moulin, M., Tiberghien, A., Le Hebel, F., & Urgelli, B. (2016). *Contribution d'une évaluation formative pour le développement du savoir dans la classe : Le cas de l'évaluation par les pairs*. Communication présentée au colloque de l'Association pour la recherche en didactique des sciences et des technologies (ARDiST), Lens, France.

Mottier-Lopez, L. (2023). Quelles sont les nouvelles conceptions de l'évaluation en classe ? Vers des évaluations pour apprendre durablement. In Cnesco (Éd.), *L'évaluation en classe, au service de l'apprentissage des élèves : Notes des experts* (pp. 13–22). Cnesco-Cnam.

Mumpuni, K. E., Priyayi, D. F., & Widoretno, S. (2022). How do students perform a peer assessment? *International Journal of Instruction*, 15(3), 751–766. <https://doi.org/10.29333/iji.2022.15341a>

Newstadt, M., & Godley, A. (2019). *Using rubrics to teach science writing*. Writers Who Care. <https://writerswhocare.wordpress.com/2019/05/06/using-rubrics-to-teach-science-writing/>

Nicol, D. J., & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. *Studies in Higher Education*, 31(2), 199–218. <https://doi.org/10.1080/03075070600572090>

- Noroozi, O., & De Wever, B. (Éds.). (2023). *The power of peer learning: Fostering students' learning processes and outcomes*. Springer Nature Switzerland. <https://doi.org/10.1007/978-3-031-29411-2>
- Panadero, E. (2016). Is it safe? Social, interpersonal, and human effects of peer assessment: A review and future directions. In G. T. L. Brown & L. R. Harris (Éds.), *Handbook of social and human conditions in assessment* (pp. 247–266). Routledge.
- Panadero, E., & Jonsson, A. (2013). The use of scoring rubrics for formative assessment purposes revisited: A review. *Educational Research Review*, 9, 129–144. <http://dx.doi.org/10.1016/j.edurev.2013.01.002>
- Philippakos, A. (2017). Giving feedback: Preparing students for peer review and self-evaluation. *The Reading Teacher*, 71(1), 13–22. <https://doi.org/10.1002/trtr.1568>
- Reinholz, D. (2016). The assessment cycle: A model for learning through peer assessment. *Assessment & Evaluation in Higher Education*, 41(2), 301–315. <https://doi.org/10.1080/02602938.2015.1008982>
- Reinholz, D. L., & Dounas-Frazer, D. R. (2016). Using peer feedback to promote reflection on open-ended problems. *arXiv*. <https://arxiv.org/abs/1603.03686>
- Ritonga, M., Tazik, K., Omar, A., & Saberi Dehkordi, E. (2022). Assessment and language improvement: The effect of peer assessment (PA) on reading comprehension, reading motivation, and vocabulary learning among EFL learners. *Language Testing in Asia*. <https://doi.org/10.1186/s40468-022-00188-z>
- Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional Science*, 18(2), 119–144.
- Sluijsmans, D. M. A., Brand-Gruwel, S., van Merriënboer, J. J. G., & Martens, R. L. (2004). Training teachers in peer-assessment skills: Effects on performance and perceptions. *Innovations in Education and Teaching International*, 41(1), 59–78. <https://doi.org/10.1080/1470329032000172720>
- Topping, K. J. (1998). Peer assessment between students in colleges and universities. *Review of Educational Research*, 68(3), 249–276. <https://doi.org/10.3102/00346543068003249>

Table des illustrations

Figure 1 : Modèle du cycle de l'évaluation par les pairs (Kollar et Fischer, 2010, cités par Reinholz, 2016).

Figure 2 : Extrait de la version initiale avant reformulation (élève 9, tâche 1).

Figure 3 : Extrait de la version finale après reformulation (élève 9, tâche 1).

Figure 4 : Extrait de la version initiale illustrant un schéma (élève 2, tâche 1).

Figure 5 : Extrait de la version finale illustrant le même schéma (élève 2, tâche 1).

Figure 6 : Extrait de la version initiale avant reformulation (élève 2, tâche 1).

Figure 7 : Extrait de la version finale après reformulation (élève 2, tâche 1).

Figure 8 : Extrait d'un feedback illustrant un feedback précis (élève 10, tâche 1).

Figure 9 : Extrait d'un feedback illustrant un feedback respectueux, mais peu chaleureux (élève 7, tâche 1).

Figure 10 : Extrait d'un feedback illustrant une suggestion de piste d'amélioration implicite (élève 3, tâche 1).

Figure 11 : Extrait d'un feedback illustrant une justification pertinente (élève 3, tâche 1).

Figure 12 : Extrait d'un feedback illustrant un feedback peu élaboré (élève 2, tâche 2).

Figure 13 : Extrait d'un feedback illustrant un feedback précis avec une référence modérée aux critères (élève 3, tâche 1).

Figure 14 : Extrait d'un feedback illustrant un feedback très clair (élève 7, tâche 1).

Figure 15 : Extrait d'un feedback illustrant une suggestion de piste d'amélioration implicite (élève 7, tâche 1).

Figure 16 : Extrait d'un feedback illustrant un feedback de qualité moyenne (élève 7, tâche 1).

Figure 17 : Extrait de la version initiale de la production (élève 1, tâche 2).

Figure 18 : Extrait du feedback reçu par l'élève (élève 1, tâche 2).

Figure 19 : Extrait de la version finale de la production (élève 1, tâche 2).

Figure 20 : Extrait de la version initiale de la production (élève 9, tâche 1).

Figure 21 : Extrait du feedback reçu par l'élève (élève 9, tâche 1).

Figure 22 : Extrait de la version finale de la production (élève 9, tâche 1).

Annexes

Annexe 1 – Trame

Mise en œuvre du dispositif expérimental

L'évaluation par les pairs

Objectif du dispositif

Ce dispositif vise à analyser l'impact de l'évaluation par les pairs sur :

- la qualité des productions écrites des élèves
- leur capacité à formuler un feedback constructif
- leur appropriation des critères d'évaluation

Le protocole repose sur un cycle d'apprentissage structuré, comprenant une phase de formation, une première mise en pratique et une seconde mise en pratique permettant d'observer une éventuelle progression.

Déroulement détaillé

Heure 1 (Louise & Léa)

Présentation du projet

- Explication du cadre du projet
- Sensibilisation à l'intérêt du feedback
 - Comprendre la différence entre jugement et feedback constructif via un petit jeu.

Heures 2-3 (Enseignantes)

Enseignement théorique par les professeurs.

Heure 4 (Enseignantes + observation Louise & Léa)

Lors de cette séance, l'objectif est de permettre aux élèves de mieux comprendre ce qui caractérise une production écrite de qualité (rapport d'expérimentation ou expression écrite).

L'enseignante commence par poser la question suivante à la classe :

« Selon vous, quels éléments doit contenir un bon rapport d'expérimentation / une bonne production écrite ? »

Les réponses des élèves sont notées au tableau sous forme de mind map (carte mentale).

Une fois les propositions recueillies, l'enseignante guide la discussion afin de :

- Regrouper les idées similaires
- Reformuler certaines propositions pour les rendre plus précises
- Compléter si nécessaire certains éléments essentiels

Exemple du mind-map :



Dans un second temps, la grille d'évaluation est présentée aux élèves. Chaque critère est repris un par un et mis en lien avec les éléments évoqués dans la mind map.

L'enseignante explicite concrètement la signification de chaque critère afin que les élèves en comprennent précisément les attendus. Certains critères peuvent être reformulés afin qu'ils soient compréhensibles pour le plus d'élèves possible.

Heures 5-6 (Enseignantes + observation Louise & Léa)

Dans un premier temps, les élèves réalisent une tâche d'écriture en lien avec la matière travaillée (rapport d'expérimentation ou production écrite). Ils s'appuient sur les consignes données ainsi que sur la grille de critères préalablement travaillée en classe. Ils mobilisent déjà ces critères au moment de rédiger leur production.

Une fois les productions terminées, l'enseignante organise l'évaluation par les pairs. Chaque élève reçoit la production d'un autre élève. Cette attribution se fait de manière aléatoire.

Les élèves relisent attentivement la production reçue en utilisant la grille d'évaluation comme support. Ils sont invités à analyser le travail à partir des critères étudiés et à rédiger un feedback structuré. Ce feedback doit être :

- en lien avec les critères,
- précis et argumenté,
- orienté vers des pistes concrètes d'amélioration.

Pendant cette phase, l'enseignante adopte une posture volontairement discrète. Elle laisse les élèves travailler de manière autonome et n'intervient qu'en cas de difficulté, d'incompréhension d'un critère ou de malentendu entre élèves. Son rôle est d'accompagner et de réguler si nécessaire, tout en favorisant la responsabilisation des élèves.

Dans un troisième temps, les élèves récupèrent leur production initiale (feuille A) ainsi que le feedback rédigé par leur pair (feuille B). Après avoir pris connaissance des remarques formulées, ils rédigent une nouvelle version complète de leur travail sur une autre feuille (feuille C).

Cette feuille C correspond à la version finale révisée. Les élèves y réécrivent l'ensemble de leur production au propre en tenant compte des suggestions reçues. Les éléments modifiés ou ajoutés apparaissent dans une couleur distincte, tandis que les parties conservées sans modification restent rédigées en bleu.

Cette étape de révision est essentielle : elle permet aux élèves de transformer le feedback en action concrète d'amélioration.

Heures 7-8 (Enseignantes + observation Louise & Léa)

Lors de ces deux périodes, le dispositif est reconduit selon la même structure que précédemment afin de permettre aux élèves de réinvestir ce qu'ils ont appris. La séance débute par un rappel des critères d'évaluation afin de réactiver les repères travaillés lors du premier cycle.

Les élèves réalisent ensuite une seconde tâche d'écriture, similaire ou légèrement différente de la première, en s'appuyant sur les consignes et sur la grille de critères. Une fois les productions terminées, une nouvelle évaluation par les pairs est organisée : chaque élève reçoit le travail d'un autre élève, l'analyse à l'aide de la grille et rédige un feedback structuré, précis et orienté vers l'amélioration.

Comme lors du premier cycle, l'enseignante adopte une posture discrète, favorise l'autonomie des élèves et n'intervient qu'en cas de difficulté ou de malentendu.

Enfin, les élèves récupèrent leur production initiale (feuille A) ainsi que le feedback reçu (feuille B) et rédigent une nouvelle version complète sur une troisième feuille (feuille C). Les modifications ou ajouts sont rédigés dans une couleur distincte, tandis que les passages conservés sans changement restent en bleu.

Les rôles en détail :

Léa & Louise :

- Conception du dispositif
- Animation de la séance de formation au feedback
- Observation non participante lors des cycles
- Collecte des données
- Analyse des productions et des feedbacks

Enseignantes :

- Enseignement de la théorie
- Animation des séances de productions écrites
- Animation des séances d'évaluation

Données collectées par Louise & Léa :

- Versions initiales et révisées des productions
- Feedbacks écrits des pairs
- Observations de classe

Annexe 2 – Tâches **sciences**

Tâche 1

LABO – Découverte de différents phénomènes

But du labo : Distinguer un phénomène chimique d'un phénomène physique.

Par groupes de 2/3 élèves, réaliser les 2 expériences proposées ci-dessous, à l'aide du matériel mis à disposition au labo. Vous disposez de 2h de cours pour faire cette activité et rédiger votre rapport.



Expérience A – La pluie d'or

Matériel et réactifs :

- du nitrate de plomb $\text{Pb}(\text{NO}_3)_2$ solide, de l'iodure de potassium KI solide et de l'eau du robinet.
- 2 berlins de 100 mL, une spatule, un verre de montre et une balance.

Mode opératoire :

- Prélever 0,5 g de nitrate de plomb et les dissoudre dans 50 mL d'eau du robinet (berlin 1). Agiter jusqu'à dissolution complète.
- Répéter l'opération en utilisant 0,6 g d'iodure de potassium dans 50 mL d'eau du robinet (berlin 2).
- Verser lentement la solution d'iodure de potassium (berlin 2) dans le berlin contenant le nitrate de plomb (berlin 1).
- Noter vos observations.
- Emettre une conclusion sur le phénomène observé : est-il un phénomène chimique ou physique ? Justifier votre réponse.

Ton rapport de labo (1 par personne et nominatif !) doit comporter :

Le but du labo.

Le schéma de l'expérience réalisée.

Les observations de l'expérience.

La conclusion et explications du phénomène observé : est-il physique ou chimique ?

Tâche 2

LABO - Découverte de différents phénomènes

But du labo : Distinguer un phénomène chimique d'un phénomène physique.

Par groupes de 2/3 élèves, réaliser les 2 expériences proposées ci-dessous, à l'aide du matériel mis à disposition au labo. Vous disposez de 2h de cours pour faire cette activité et rédiger votre rapport.



Expérience B - La mer morte

Matériel et réactifs :

- du chlorure de sodium NaCl (sel de cuisine) et de l'eau du robinet.
- 1 berlin de 100 mL, une spatule, un verre de montre et une balance.

Mode opératoire :

- Prélever 0,5 g de chlorure de sodium et les dissoudre dans 50 mL d'eau du robinet. Agiter jusqu'à dissolution complète.
- Noter vos observations.
- Émettre une conclusion sur le phénomène observé : est-il un phénomène chimique ou physique ? Justifier votre réponse.

Ton rapport de labo (1 par personne et nominatif !) doit comporter :

Le but du labo.

Le schéma de l'expérience réalisée.

Les observations de l'expérience.

La conclusion et explications du phénomène observé : est-il physique ou chimique ?

Annexe 3 – Tâches anglais

Tâche 1 :

English homework N°
U3 : What does she look like ?

Contexte : Dans son dernier email, ton correspondant t'a parlé de son personnage préféré. Tu décides toi aussi de faire la même chose. Tu lui écris un email dans lequel tu lui fais la description de ce personnage et tu y joins une photo.

Tâche : Utilise les images ci-dessous et ton cours pour rédiger une description **précise et complète** de cette personne. N'oublie aucun détail ! Tu peux évidemment inventer les informations que tu ne connais pas.

Rédige un email de 120 mots minimum avec des phrases complètes en anglais à propos des éléments suivants :

1. nom/prénom
2. âge
3. pays ou ville d'origine
4. taille
5. corpulence
6. cheveux
7. yeux
8. oreilles, nez, bouche
9. autre(s) caractéristique(s) éventuelles
10. famille (2-3 phrases)
11. actions habituelles (2-3)

Grammaire : Be/have got/present simple + déterminants adéquats



Tâche 2 :

C. ECRIRE :

/20

Dans son dernier email, ton correspondant t'a parlé de son/sa meilleur(e) ami(e). Tu décides toi aussi de faire la même chose. Tu lui écris un email dans lequel tu lui fais la description de cette personne en t'inspirant **d'une des deux photos**.

Tâche : une description **précise et complète**. N'oublie aucun détail! Tu peux évidemment inventer les informations que tu ne connais pas.

Utilise « to be », « have (got) », des verbes d'actions, des compléments de phrases et, bien sûr, décris-le en détails : nom/prénom, âge, pays ou ville d'origine, physique, caractère, loisirs, ...

3TQ, bilan unit 3



Grille d'évaluation par les pairs – Rapport de laboratoire en sciences

	1	2	3	4
Schéma de l'expérience	Le schéma est difficile à comprendre, incomplet ou incorrect. Les éléments essentiels ne sont pas représentés ou pas légendés.	Le schéma est partiellement correct. Certains éléments importants sont absents ou mal représentés. Les légendes sont incomplètes ou imprécises.	Le schéma est clair et comporte la plupart des éléments nécessaires. Les légendes sont présentes mais parfois peu précises.	Le schéma est clair, soigné et complet. Tous les éléments importants du montage sont représentés et correctement légendés. Les légendes sont précises et scientifiques. Les éléments représentés sont scientifiquement pertinents et cohérents avec le protocole.
Les observations	Les observations sont très incomplètes, confuses ou incorrectes. Peu ou pas de données exploitables.	Les observations sont partielles ou peu précises. Des données manquent ou les unités sont absentes.	Les observations sont globalement correctes et compréhensibles. Quelques imprécisions ou oublis mineurs. Les unités sont généralement présentes.	Les observations sont complètes, précises et rédigées avec un vocabulaire scientifique approprié. Les données sont correctement présentées (tableaux, unités, mesures exactes). Les observations décrivent uniquement ce qui est observé (sans interprétation prématurée) avec un vocabulaire scientifique précis.
La conclusion	La conclusion est incorrecte ou non justifiée.	La conclusion identifie le phénomène mais sans justification claire fondée sur les observations.	La conclusion identifie correctement le type de phénomène mais la justification reste partielle ou peu développée.	La conclusion identifie correctement s'il s'agit d'un phénomène physique ou chimique et justifie cette classification à partir des observations réalisées.

Échelle : 1 = Maîtrise insuffisante / 2 = Maîtrise partielle / 3 = Bonne maîtrise / 4 = Maîtrise complète et justifiée

Grille d'évaluation par les pairs – Expression écrite en langues

	1	2	3	4
Contenu	Idées peu pertinentes ou absentes. Le texte ne répond pas à la consigne .	Idées incomplètes ou certains points hors sujets.	Idées pertinentes mais développement partiel.	Idées pertinentes et complètes. Le texte répond parfaitement à la consigne.
Organisation	Texte désorganisé Absence de paragraphe Connecteurs inexistant	Organisation faible Idées mélangées dans les paragraphes Connecteurs rares	Organisation correcte Paragraphes peu organisés Manque de quelques connecteurs	Organisation correcte Paragraphes cohérents Connecteurs pertinents
Vocabulaire	Vocabulaire très pauvre ou incorrect Beaucoup d'erreurs Texte difficile à comprendre.	Vocabulaire limité Quelques erreurs Mots inadaptés.	Vocabulaire correct Quelques répétitions Manque de précision dans le vocabulaire	Vocabulaire riche et précis Adapté au contexte
Grammaire/syntaxe	Beaucoup d'erreurs Phrases incompréhensibles ou trop simples Grammaire non maîtrisée	Erreurs fréquentes Phrases peu variées Difficile de comprendre le texte	Quelques erreurs grammaticales Phrases simples mais compréhensibles	Phrases correctes et variées Grammaire maîtrisée
Orthographe	Fautes fréquentes Ponctuation absente ou incorrecte Texte difficilement lisible	Nombreuses fautes Ponctuation approximative Texte compliqué à lire	Quelques fautes mineures Ponctuation relativement correcte	Orthographe et ponctuation correctes Texte facile à lire

Échelle : 1 = Maîtrise insuffisante / 2 = Maîtrise partielle / 3 = Bonne maîtrise / 4 = Maîtrise complète et justifié

Annexe 6 – Outil d’analyse des modifications des productions

Analyse des productions A et B

- **Nom de l’élève :**
- **Discipline :** ☐ Sciences ☐ Anglais
- **Tâche :** ☐ 1^{ère} ☐ 2^{ème}

1. Présence et ampleur des modifications

Élément observé		Commentaire
Absence de modification	☐	
Modification partielle	☐	
Modification importante / approfondie	☐	

2. Type(s) de modification(s)

Type de modification		Exemple / extrait de la copie
Ajout	☐	
Suppression	☐	
Reformulation	☐	

3. Partie du travail concernées

Sciences

Partie du travail	Modification observée
Observation	☐
Schéma	☐
Conclusion	☐
Autre : _____	☐

Anglais

Partie du travail	Modification observée
Contenu	☐
Organisation	☐
Vocabulaire	☐
Grammaire	☐
Orthographe	☐
Autre : _____	☐

4. Commentaire global de l'analyse

.....

.....

.....

Annexe 7 – Outils d’analyse de la qualité des feedbacks

	Définition	0	1	2	3
1	Capacité du commentaire à cibler un élément concret, identifiable et précis du travail évalué.	Le commentaire est totalement vague, général, ou ne porte pas sur le travail (ex. “c’est bien”, “à revoir”).	Le commentaire évoque un aspect du travail, mais de façon très générale et imprécise, sans point identifiable.	Le commentaire identifie un élément du travail, mais sans grande précision (ex. une partie large, un aspect global).	Le commentaire cible précisément une phrase, une partie identifiée, un argument, un choix méthodologique ou une étape du travail.
2	Utilisation des critères fournis dans la grille d’évaluation pour formuler le retour.	Aucun critère n’est mentionné et le commentaire n’y fait pas allusion.	Le commentaire semble se rapporter à un critère, mais sans le nommer ou le décrire clairement.	Le commentaire mentionne un critère, mais de manière superficielle ou indirecte.	Le commentaire fait explicitement référence à un critère précis (ex. “selon le critère de clarté...”) et l’utilise pour argumenter l’évaluation.
3	Capacité du feedback à proposer des pistes d’amélioration exploitables.	Aucune suggestion d’amélioration n’est donnée.	Suggestion implicite ou non actionnable (“tu peux améliorer”).	Suggestion existante mais générale ou incomplète (ex. “développer l’introduction”).	Suggestion précise, utile, actionnable et adaptée au problème identifié (ex. “clarifie ta thèse dès la première phrase de l’intro”).
4	Clarté, organisation du texte et qualité linguistique du commentaire.	Commentaire difficile ou impossible à comprendre.	Compréhensible globalement, mais désorganisé ou maladroit.	Message clair, compréhensible, quelques maladresses ou imprécisions.	Commentaire parfaitement clair, structuré et facile à comprendre.
5	Degré de respect, bienveillance et adéquation du ton employé.	Ton agressif, moqueur, blessant ou ironique.	Ton neutre mais parfois sec, distant ou maladroit.	Ton globalement respectueux, mais manque de chaleur ou de formulation positive.	Ton pleinement respectueux, encourageant, constructif et bienveillant.

	Définition	0	1	2	3
6	Présence d'éléments qui appuient le jugement formulé.	Aucun argument ni exemple.	Une justification implicite ou très générale.	Exemple ou justification présent mais succinct.	Justification claire, précise, appuyée sur un exemple concret du travail.

Analyse de la qualité des feedbacks

- **Nom de l'élève évaluateur :**
- **Nom de l'élève évalué :**
- **Discipline :** ☐ Sciences ☐ Anglais
- **Tâche :** ☐ 1^{ère} ☐ 2^{ème}

Extrait du feedback analysé n°1

.....

Commentaire n°1 :

Critère	Score (0–3)	Commentaire
Précision	☐ 0 ☐ 1 ☐ 2 ☐ 3	
Référence aux critères	☐ 0 ☐ 1 ☐ 2 ☐ 3	
Proposition de pistes d'amélioration	☐ 0 ☐ 1 ☐ 2 ☐ 3	
Clarté et lisibilité du message	☐ 0 ☐ 1 ☐ 2 ☐ 3	
Respect	☐ 0 ☐ 1 ☐ 2 ☐ 3	
Justifications	☐ 0 ☐ 1 ☐ 2 ☐ 3	

- **Score du commentaire n°1 : ... / 18**
- **Niveau de qualité du feedback**

Niveau	Intervalle	Coche
Faible	0 – 6	☐
Moyen	7 – 12	☐

Niveau	Intervalle	Coche
Bon	13 – 18	☐

- **Commentaire global :**

.....

Extrait du feedback analysé n°2

.....

Commentaire n°2 :

Critère	Score (0–3)	Commentaire
Précision	☐ 0 ☐ 1 ☐ 2 ☐ 3	
Référence aux critères	☐ 0 ☐ 1 ☐ 2 ☐ 3	
Proposition de pistes d'amélioration	☐ 0 ☐ 1 ☐ 2 ☐ 3	
Clarté et lisibilité du message	☐ 0 ☐ 1 ☐ 2 ☐ 3	
Respect	☐ 0 ☐ 1 ☐ 2 ☐ 3	
Justifications	☐ 0 ☐ 1 ☐ 2 ☐ 3	

- **Score du commentaire n°2 : ... / 18**

Niveau de qualité du feedback

Niveau	Intervalle	Coche
Faible	0 – 6	☐
Moyen	7 – 12	☐
Bon	13 – 18	☐

- **Commentaire global :**

.....

Extrait du feedback analysé n°3

.....

Commentaire n°3 :

Critère	Score (0–3)	Commentaire
Précision	☐ 0 ☐ 1 ☐ 2 ☐ 3	
Référence aux critères	☐ 0 ☐ 1 ☐ 2 ☐ 3	
Proposition de pistes d'amélioration	☐ 0 ☐ 1 ☐ 2 ☐ 3	
Clarté et lisibilité du message	☐ 0 ☐ 1 ☐ 2 ☐ 3	
Respect	☐ 0 ☐ 1 ☐ 2 ☐ 3	
Justifications	☐ 0 ☐ 1 ☐ 2 ☐ 3	

- **Score du commentaire n°3 : ... / 18**

Niveau de qualité du feedback

Niveau	Intervalle	Coche
Faible	0 – 6	☐
Moyen	7 – 12	☐
Bon	13 – 18	☐

- **Commentaire global :**

.....

Extrait du feedback analysé n°.....

.....

Commentaire n°..... :

Critère	Score (0–3)	Commentaire
Précision	☐ 0 ☐ 1 ☐ 2 ☐ 3	
Référence aux critères	☐ 0 ☐ 1 ☐ 2 ☐ 3	
Proposition de pistes d'amélioration	☐ 0 ☐ 1 ☐ 2 ☐ 3	
Clarté et lisibilité du message	☐ 0 ☐ 1 ☐ 2 ☐ 3	
Respect	☐ 0 ☐ 1 ☐ 2 ☐ 3	
Justifications	☐ 0 ☐ 1 ☐ 2 ☐ 3	

- **Score du commentaire n°..... : ... / 18**

Niveau de qualité du feedback

Niveau	Intervalle	Coche
Faible	0 – 6	☐
Moyen	7 – 12	☐
Bon	13 – 18	☐

- **Commentaire global :**

Synthèse du feedback complet

Nombre de commentaires

Score total

Score moyen (optionnel)

Niveau global de qualité du feedback

Niveau	Intervalle	Coche
Faible	0 – 6	☐
Moyen	7 – 12	☐
Bon	13 – 18	☐

• **Commentaire final :**

.....

.....

.....

Annexe 8 – Outil d’analyse du lien entre les feedbacks et les modifications des productions

Analyse du lien entre **feedbacks** et **modifications**

- Nom de l’élève :
- Discipline : ☐ Sciences ☐ Anglais
- Nombre total de feedbacks reçus :
- Nombre total de modifications reçues :

Synthèse des feedbacks reçus

Feedback	Porte sur	Score (0–18)	Catégorie : faible, moyenne, bonne
F1			
F2			
F3			

Synthèse des modifications reçues

Modification	Nature de la modification
M1	
M2	
M3	

Lien entre les feedbacks et les modifications

Feedback	Modification	Type de relation : directe, partielle ou absente

Conclusion du tableau :

.....

.....

.....

Interprétation et profil

Cocher le profil :

- Profil 1 – Feedback efficace et bonne appropriation
- Profil 2 – Feedback de qualité mais appropriation limitée
- Profil 3 – Faible feedback et faible impact
- Profil 4 – Feedback moyen mais amélioration correcte
- Profil 5 – Feedback moyen et modifications moyennes
- Profil 6 - Appropriation différenciée des feedbacks selon leur qualité

Justification du profil :

.....

.....

.....

Annexe 9 : Outil d'analyse comparative des premiers jets en anglais

Analyse comparative des premiers jets (Anglais)

- Nom de l'élève :
- Discipline : Anglais
-

Comparaison des résultats

Critère	A1	A2	Observation
Contenu			☐ Progression ☐ Stabilité ☐ Régression
Organisation			☐ Progression ☐ Stabilité ☐ Régression
Vocabulaire			☐ Progression ☐ Stabilité ☐ Régression
Grammaire			☐ Progression ☐ Stabilité ☐ Régression
Orthographe			☐ Progression ☐ Stabilité ☐ Régression
Score global			☐ Progression ☐ Stabilité ☐ Régression

Éléments d'évolution observés

Progression(s) constatée(s)

.....

Stabilité constatée

.....
Difficultés persistantes

.....
Nouvelles difficultés observées

.....
Conclusion :

.....
.....
.....
.....

Annexe 10 : Tableau de synthèse : qualité des feedbacks

Sciences :

Critères	Tâche 1 (score moyen /3)	Tâche 2 (score moyen /3)
Précision	2,56	1,93
Critères	2	2,04
Améliorations	0,6	0,46
Clarté	2,26	1,54
Respect	2,15	1,79
Justifications	1,9	1,21
Répartition des élèves selon le niveau	0 faible (0,0%) 10 moyens (90,9%) 1 bon (9,1%)	1 faible (11,1%) 8 moyens (88,9%) 0 bon (0,0%)

Anglais :

Critères	Tâche 1 (score moyen /3)
Précision	2,17
Critères	1,85
Améliorations	0,57
Clarté	1,9
Respect	1,92
Justifications	1,65
Répartition des élèves selon le niveau	1 faible (12,5%) 7 moyens (87,5%) 0 bon (0,0%)

Annexe 11 : Tableau de synthèse : modifications des productions

Discipline	Nombres de productions analysées	Modifications totales	Ajouts	Suppressions	Reformulations
Sciences – tâche 1	11	31	15 (48,4%)	3 (9,7%)	13 (41,9%)
Sciences – tâche 2	9	21	12 (57,2%)	2 (9,5%)	7 (33,3%)
Anglais – tâche 1	8	32	11 (34,4%)	4 (12,5%)	17 (53,1%)

Sciences :

Sciences – tâche 1	
Nombre total de modifications	31
Observation	11
Schéma	10
Conclusion	4

Sciences – tâche 2	
Nombre total de modifications	21
Observation	6
Schéma	6
Conclusion	5

Anglais :

Anglais – tâche 1	
Nombre total de modifications	32
Contenu	4
Organisation	6
Vocabulaire	4
Grammaire	5
Orthographe	3

Annexe 12 : Tableau de synthèse : lien entre feedbacks et modifications

Sciences :

	Tâche 1	Tâche 2
Nombre de productions	11	9
Nombres de modifications	31	21
Lien direct	13 (41,9%)	11 (52,4%)
Lien partiel	2 (6,5%)	5 (23,8%)
Lien absent	16 (51,6%)	5 (23,8%)

	Tâche 1	Tâche 2
Nombre d'élèves	11	9
Profil 1	3 (27,3%)	0 (0,0%)
Profil 2	3 (27,3%)	1 (11,1%)
Profil 3	1 (9,1%)	1 (11,1%)
Profil 4	1 (9,1%)	5 (55,6%)
Profil 5	3 (27,3%)	1 (11,1%)
Profil 6	0 (0,0%)	1 (11,1%)

Anglais :

	Tâche 1
Nombre de productions	8
Nombres de modifications	32
Lien direct	9 (28,1%)
Lien partiel	7 (21,9%)
Lien absent	16 (50,0%)

	Tâche 1
Nombre d'élèves	8
Profil 1	0 (0,0%)
Profil 2	2 (25,0%)
Profil 3	0 (0,0%)
Profil 4	2 (25,0%)
Profil 5	4 (50,0%)
Profil 6	0 (0,0%)

Annexe 13 : Tableau de synthèse des résultats par discipline

Sciences

N° de l'élève	Tâche 1				Tâche 2			
	Score qualité des feedbacks		Améliorations perçues A vs B	Lien entre améliorations et feedbacks	Score qualité des feedbacks		Améliorations perçues A vs B	Lien entre améliorations et feedbacks
1	F1	14	0 ajout 0 suppression 2 reformulations	2 directs 0 partiel 0 absent Profil 1	F1	14	2 ajouts 2 suppressions 1 reformulation	3 directs 0 partiel 2 absents Profil 4
	F2	13			F2	9		
	F3	10			F3	9		
	Moyenne	12,3			Moyenne	10,6		
2	F1	11	1 ajout 0 suppression 2 reformulations	2 directs 0 partiel 1 absent Profil 1	F1	4	1 ajout 0 suppression 0 reformulation	0 direct 0 partiel 1 absent Profil 3
	F2	14			F2	4		
	F3	13			F3	8		
	Moyenne	12,6			Moyenne	5,3		
3	F1	11	1 ajout 0 suppression 1 reformulation	1 direct 0 partiel 1 absent Profil 2	F1	9	2 ajouts 0 suppression 1 reformulation	2 directs 1 partiel 0 absent Profil 4
	F2	11			F2	8		
	F3	15			F3	8		
	Moyenne	12,3			Moyenne	8,3		
4	F1	7	2 ajouts 1 suppression 2 reformulations	1 direct 1 partiel 3 absents Profil 5	F1	12	2 ajouts 0 suppression 1 reformulation	0 direct 2 partiels 1 absent Profil 4
	F2	13			F2	9		
	F3	10			F3	5		
	Moyenne	10			Moyenne	8,6		
5	F1	10	2 ajouts 1 suppression 0 reformulation	1 direct 0 partiel 2 absents Profil 5	F1	12	0 ajout 0 suppression 1 reformulation	1 direct 0 partiel 0 absent Profil 5
	F2	12			F2	12		
	F3	14			F3	9		
	Moyenne	12			Moyenne	11		

6	F1	10	0 ajout 0 suppression 1 reformulation	0 direct 0 partiel 1 absent Profil 3	F1	9	0 ajout 0 suppression 0 reformulation	0 direct 0 partiel 0 absente Profil 2
	F2	7			F2	11		
	F3	8			F3	9		
	F4	11						
	Moyenne	9			Moyenne	9,6		
7	F1	14	0 ajout 0 suppression 3 reformulations	1 direct 1 partiel 1 absent Profil 2	F1	5	2 ajouts 0 suppression 1 reformulation	1 direct 1 partiel 1 absent Profil 5
	F2	13			F2	5		
	F3	11			F3	10		
					F4	12		
	Moyenne	12,6			Moyenne	8		
8	F1	10	2 ajouts 0 suppression 0 reformulation	1 direct 0 partiel 1 absente Profil 4	F1	10	1 ajout 0 suppression 0 reformulation	1 direct 0 partiel 0 absent Profil 4
	F2	10			F2	7		
	F3	9			F3	9		
	Moyenne	9,6			Moyenne	8,6		
9	F1	9	2 ajouts 1 suppression 2 reformulations	3 directs 0 partiel 2 absents Profil 2	F1	9	2 ajouts 0 suppression 2 reformulations	3 directs 1 partiel 0 absent Profil 4
	F2	13			F2	14		
	F3	13			F3	9		
	Moyenne	11,6			Moyenne	10,6		
10	F1	14	2 ajouts 0 suppression 0 reformulation	1 direct 0 partiel 1 absent Profil 1	Absent pour la tâche 2			
	F2	15						
	F3	14						
	Moyenne	14,3						
11	F1	11	3 ajouts 0 suppression 0 reformulation	0 direct 0 partiel 3 absents Profil 5	Absent pour la tâche 2			
	F2	12						
	F3	10						
	Moyenne	11						

Anglais

N° de l'élève	Tâche 1					Tâche 2
	Score qualité des feedbacks		Améliorations perçues A vs B	Lien entre améliorations et feedbacks	Score version A /20	Score version A /20
1	F1	14	1 ajout 0 suppression 1 reformulation	1 direct 0 partiel 1 absent Profil 2	11	13
	F2	14				
	F3	12				
	F4	10				
	F5	10				
	Moyenne	12				
2	F1	5	2 ajouts 1 suppression 3 reformulations	3 directs 1 partiel 0 absent Profil 4	15	17
	F2	10				
	F3	10				
	F4	12				
	F5	8				
	Moyenne	9				
3	F1	14	1 ajout 0 suppression 1 reformulation	0 direct 2 partiel 0 absent Profil 2	17	20
	F2	12				
	F3	15				
	F4	10				
	F5	11				
	Moyenne	12,4				
4	F1	8	0 ajout 0 suppression 0 reformulation	0 direct 0 partiel 0 absent Profil 5	13	15
	F2	12				
	F3	11				
	F4	11				
	F5	8				
	Moyenne	10				

5	Absent pour la tâche 1					
6	F1	12	2 ajouts 0 suppression 1 reformulation	2 directs 0 partiel 0 absent Profil 4	13	10
	F2	4				
	F3	4				
	F4	9				
	F5	4				
	Moyenne	6,6				
7	F1	14	3 ajouts 2 suppressions 4 reformulations	0 direct 2 partiels 0 absent Profil 5	13	14
	F2	12				
	F3	4				
	F4	11				
	F5	10				
	Moyenne	10,2				
8	Absent pour la tâche 1					
9	F1	7	1 ajout 0 suppression 5 reformulations	1 direct 1 partiel 0 absent Profil 5	10	8
	F2	11				
	F3	8				
	F4	12				
	F5	12				
	Moyenne	10				
10	Absent pour la tâche 1					
11	F1	4	1 ajout 1 suppression 2 reformulations	2 directs 1 partiel 0 absent Profil 5	14	14
	F2	12				
	F3	14				
	F4	10				

	F5	13				
	Moyenne	10,6				

Annexe 14 : Grille d'utilisation des intelligences artificielles génératives

ANNEXE 1. Grille d'utilisation des intelligences artificielles génératives pour le mémoire et le travail préparatoire à la recherche (TPR)

Dans ce cours, l'utilisation des IA Génératives pour ...	Utilisation autorisée	Utilisation de l'étudiant	Nom de/des IA générative(s) utilisée(s)
Aider à la rédaction du plan et de la structure du travail	✓	Oui	Chat GPT
Vérifier l'orthographe et la grammaire	✓	Non	
Reformuler le contenu de votre travail ou améliorer le style	✓	Oui	Chat GPT
Traduire une autre langue	✓	Non	
Aider à la planification et gestion de projet	✓	Non	
Obtenir des informations générales sur un sujet, expliquer et approfondir la compréhension de concepts	✓	Oui	Chat GPT
Faire un brainstorming et évaluer ses propres idées, pour trouver d'autres perspectives ou des contre-arguments	✓	Oui	Chat GPT
Aider à la recherche documentaire	✓	Oui	Chat GPT
Aider à la synthèse de la littérature	✓	Oui	Chat GPT
Aider à la formulation de questions de recherche et d'hypothèses	✓	Oui	Chat GPT
Aider à l'analyse de données (identification d'analyses à réaliser)	✓	Non	
Aider à l'interprétation des résultats	✓	Non	
Aider au respect des normes APA de la liste des références	✓	Oui	Chat GPT
Aider à la programmation de code et au débogage	✓	Non	
Générer un feedback critique et révision	✓	Non	
Créer du contenu multimédia, par exemple des images, des vidéos, des animations ou du son (mais toujours expliquer l'utilisation des IA)	✓	Non	

Générer le contenu du travail à partir de mots clés, d'un thème ou d'une question de départ sans vérification et appropriation personnelle	✗	Non	
Reproduire ou reformuler du contenu existant sans citer la source (plagiat)	✗	Non	
Falsifier, créer des données	✗	Non	
Générer des références bibliographiques sans en contrôler l'existence ou l'exactitude	✗	Non	
Utiliser dans une IA des données non anonymisées issues de l'étude (participants, résultats, verbatim, etc.)	✗	Non	

Je, soussigné Louise BASTIN & Léa GUYOT.....
(nom et prénom), déclare sur l'honneur avoir complété la grille
avec exactitude et honnêteté, en toute bonne foi.

Légende :

✓ = L'utilisation des IA Génératives est autorisée (ce qui ne veut pas dire qu'elle est encouragée. Voir ci-dessous)

✗ = L'utilisation des IA Génératives n'est pas autorisée, les infractions entraîneront des sanctions.

n.a = Non-applicable pour ce cours

L'étudiant remplit sa colonne par « Oui » ou par « Non » en fonction de s'il a utilisé ou non une IA générative.

***Attention :** le fait que certaines pratiques soient autorisées ne signifie pas que l'on attende ou que l'on encourage l'utilisation d'IA Génératives pour cette évaluation. Dans de nombreuses situations, vous obtiendrez probablement de meilleurs résultats sans utiliser d'IA Génératives.

Cette grille est une adaptation du « Tableau d'utilisation des IA Génératives » de la Faculté des Arts et des Sciences Sociales (FASoS) de l'Université de Maastricht.

Annexe 15 : Documents des apports individuels

Critères d'auto-évaluation du fonctionnement du groupe 	Cote (entre 0 et 5)	Justification (avec éléments concrets)
Vision claire et partagée par chacune de ce que nous avons à faire dans le projet	5	Dès le début du projet, nous avons défini conjointement les objectifs à atteindre. Cette vision commune nous a permis d'établir une ligne directrice claire, de structurer notre démarche de recherche et de progresser de manière cohérente tout au long de la réalisation du mémoire.
Contribution de chacun à l'avancée du projet, responsabilité collective envers le projet	5	Nous avons tout d'abord élaboré ensemble notre question de recherche afin de construire une problématique commune et de définir une ligne directrice pour notre travail. Nous avons ensuite réalisé des recherches documentaires de manière individuelle, en nous appuyant à la fois sur des sources communes et sur des ressources spécifiques à nos disciplines respectives (sciences et anglais). Ces recherches ont ensuite été mises en commun afin de construire une revue de la littérature partagée. Nous avons également toutes les deux lu les sources retenues et réalisé des fiches de lecture individuelles, ce qui nous a permis de croiser nos analyses et de compléter mutuellement notre compréhension des différents concepts mobilisés. La méthodologie a également été réfléchi conjointement. Nous avons discuté ensemble de la construction du dispositif de recherche, des cycles d'apprentissage, ainsi que de l'élaboration des outils de collecte et d'analyse des données. La mise en œuvre du dispositif a ensuite été réalisée dans nos classes respectives : Louise

		dans le cadre du cours de sciences et Léa dans le cadre du cours d'anglais. L'analyse des données a été réalisée dans un premier temps individuellement, chacune analysant les données récoltées dans sa discipline et dans la discipline de l'autre. Nous avons ensuite comparé nos analyses, discuté des interprétations proposées et réalisé un codage commun afin d'assurer une cohérence dans le traitement des données. Concernant la rédaction du mémoire, nous avons d'abord établi un plan commun afin de suivre une même ligne directrice et d'assurer une cohérence entre nos différentes parties. Les sections ont ensuite été réparties entre nous, tout en maintenant un travail collaboratif grâce aux relectures croisées, aux échanges réguliers et aux discussions visant à améliorer le contenu du mémoire. Enfin, l'ensemble du travail a fait l'objet de relectures communes ainsi que d'une relecture externe afin de garantir la qualité et la cohérence du document final.
Répartition équilibrée des tâches entre nous, en fonction des compétences, aspirations et disponibilités de chacune	5	Chacune s'est investie dans les différentes étapes du projet, tout en veillant à maintenir une démarche collaborative constante. Bien que certaines tâches aient été réparties entre nous pour des raisons organisationnelles, notamment la mise en œuvre du dispositif dans nos classes respectives, nous avons cherché à construire et à analyser ensemble l'intégralité du travail. Les échanges réguliers, les comparaisons entre nos disciplines et les relectures croisées nous ont permis de garantir une cohérence dans notre démarche, d'enrichir nos interprétations et de limiter les biais liés à une analyse individuelle.
Procédure collective de travail bien définie, avec une répartition adéquate des moments de travail collectifs et des moments de travail individuels, une gestion du temps appropriée au projet et aux impératifs de chacune, un respect du planning	4	Nous avons anticipé les différentes étapes en fixant des délais à respecter. Ces échéances ont globalement été tenues. Le travail alternait entre des moments individuels et des temps à deux permettant des ajustements réguliers et un bon suivi.
Utilisation adéquate et efficace des moyens de communication entre nous	5	Nous avons privilégié le travail en présentiel ce qui facilitait grandement les échanges d'idées et la prise de décisions rapides.

		Pour assurer un bon suivi, nous communiquions aussi régulièrement via Messenger afin de nous tenir informées de l'avancée du projet. Un document partagé nous servait de support central : nous y avons noté le plan du travail, les sources collectées, ainsi que la version en cours de rédaction. Ce fonctionnement nous a permis de suivre l'évolution du projet en temps réel, de coordonner nos contributions plus facilement et d'éviter les répétitions. Rédiger dans un même espace nous a aussi permis d'aligner nos styles et de garantir une plus grande fluidité et cohérence dans l'ensemble du travail.
Rôles clairement définis dans le groupe et tenus de façon satisfaisante	5	Comme mentionné précédemment, la répartition des rôles était claire dès le départ et a été respectée tout au long du projet.
Ambiance de travail conviviale, bonne entente entre les membres du groupe	5	L'ambiance de travail au sein de notre binôme a été très positive. Étant déjà amies en dehors du cadre universitaire, nous connaissions nos modes de fonctionnement respectifs ce qui a grandement facilité la collaboration. Cette relation de confiance et de bienveillance a permis d'instaurer un climat serein et propice à une coopération efficace et harmonieuse tout au long du projet.
Écoute et prise en compte des avis de chacune, soutien mutuel en cas de besoin	5	Les avis de chacune ont été pris en compte tout au long du projet. Chacune a pu exprimer ses idées et ses attentes, ce qui nous a permis de trouver rapidement un équilibre dans notre manière de travailler.

À l'exception de la mise en œuvre du dispositif de recherche, qui a été réalisée dans les classes respectives de chacune des étudiantes (100 % pour Louise en sciences et 100 % pour Léa en anglais), l'ensemble des étapes du projet a été mené conjointement. Les contributions de chacune sont donc estimées à **50 %** pour toutes les autres composantes du travail. Ce choix reflète notre volonté de construire le mémoire dans une démarche véritablement collaborative, fondée sur des échanges constants, des relectures croisées et des prises de décision communes, afin de garantir la cohérence de notre recherche et de limiter les biais d'analyse.