

Pour citer cet article : Girardet, C. (2020). Lorsque les étudiant·es produisent et reçoivent des feedbacks entre pairs à l'université : étude des perceptions d'un dispositif d'évaluation continue facilité par le numérique. *La Revue L'EeE*, 2. <https://doi.org/10.48325/rleee.002.04>

LORSQUE LES ÉTUDIANT·ES PRODUISENT ET REÇOIVENT DES FEEDBACKS ENTRE PAIRS À L'UNIVERSITÉ

Étude des perceptions d'un dispositif d'évaluation continue facilité par le numérique

Céline GIRARDET

Version de la publication : octobre 2020

Évaluation ouverte et collaborative

Rétroacteurs : Amaury DAELE & Vincent GRENON

Résumé

Cet article étudie les perceptions d'étudiant·es universitaires de tâches de production et de réception de feedbacks entre pairs articulées dans un dispositif d'évaluation continue (faisant se rejoindre évaluations formative et certificative) conçu dans une vision d'évaluation soutient d'apprentissage (AfL) et d'évaluation comme moyen pour apprendre (AaL). L'analyse inductive des perceptions des étudiant·es montre des apprentissages en profondeur liés à la compréhension des concepts et au développement de diverses compétences « transversales » (compétences réflexives, compétences communicationnelles, bienveillance, empathie, compétences collaboratives, compétences évaluatives). Les résultats sont discutés à la lumière de la recherche sur les feedbacks entre pairs.

Mots-clés : enseignement supérieur, évaluation entre pairs, feedback, perceptions, apprentissage

Abstract

This article examines university students' perceptions of giving and receiving peer feedback as part of a continuous evaluation scenario (linking formative and summative assessment) designed in the perspective of assessment for learning (AfL) and assessment as learning (AaL). The inductive analysis of students' perceptions shows deep learning of concepts and the development of various "transversal" skills (reflexive skills, communication skills, caring, empathy, collaboration skills, evaluative competences). Results are discussed in light of existing research on peer feedback.

Keywords: higher education, peer assessment, feedback, perceptions, learning

1. Introduction

Cette contribution interroge les perceptions¹ d'étudiant·es d'une tâche de production de **feedbacks** entre pairs (évaluation par les pairs) et de leurs perceptions de la réception et mobilisation de ces feedbacks à des fins de **régulation des apprentissages**. Ces perceptions sont étudiées dans le contexte d'un cours universitaire en sciences de l'éducation, et plus particulièrement de son évaluation continue, dont la réalisation a été soutenue par une plateforme numérique.

Un des constats décrits dans Mottier Lopez (2015a) est qu'« on ne peut plus penser l'évaluation formative et la régulation de l'apprentissage de l'élève sans impliquer ce dernier dans l'évaluation et, plus généralement, sans construire une compréhension partagée de la signification de l'évaluation avec l'ensemble des partenaires concernés » (p.28). Dans le contexte universitaire, l'évaluation est typiquement imposée aux étudiant·es. Impliquer les étudiant·es dans les processus évaluatifs peut redonner du pouvoir à ces dernier·ères (*empowerment*) en les exerçant à l'évaluation dans le but de les préparer à leur vie future (*sustainable assessment*² ; Boud, 2000 ; Boud & Soler, 2015). Dans le cadre de la formation à l'enseignement, cette évaluation a pour but, entre autres, de préparer les étudiant·es à leur vie professionnelle de futur·es enseignant·es qui seront amené·es à évaluer les apprentissages d'une multitude d'élèves. La conception de l'évaluation-soutien d'apprentissage (*assessment for learning*, AfL) a été étendue à l'évaluation comme moyen pour apprendre (*assessment as learning*, AaL), suggérant une reconfiguration de l'évaluation « traditionnelle » en la posant comme un processus par lequel l'implication de l'apprenant·e dans l'évaluation participe à ses apprentissages (Dann, 2014). Une façon d'impliquer les apprenant·es dans l'évaluation est l'évaluation par les pairs, dont les bénéfices pour l'apprentissage autonome (e.g., Shen et al., 2020), la motivation et l'engagement dans l'apprentissage (e.g., Topping, 2003) ont été largement démontrés.

Dans le cas de cette étude, nous nous intéressons à l'évaluation par les pairs par le moyen de feedbacks, dont l'importance et la force (autant positive que négative) pour l'apprentissage a été établie (e.g., Black & William, 1998 ; Hattie & Timperley, 2007 ; Sadler, 1998 ; Topping, 2003). La pratique de feedbacks par les pairs a l'avantage de permettre un grand nombre de feedbacks immédiats et individualisés (Topping, 2009), ce qui est particulièrement utile dans le contexte universitaire, dans lequel il peut être difficile pour l'enseignant·e de produire des feedbacks individuels lorsqu'elle ou il donne plusieurs cours par semestre avec des nombres d'étudiant·es pouvant atteindre plusieurs centaines. Mettre en place des dispositifs d'évaluation entre pairs est donc un bon moyen de faire en sorte que les étudiant·es aient accès à des feedbacks individualisés en cours de semestre. Dans ses nombreux travaux sur les feedbacks en enseignement supérieur, Sadler (2010) propose un modèle de l'enseignement à l'université basé sur l'utilisation intensive et ciblée de l'évaluation entre pairs en tant que stratégie pédagogique, et ce, non seulement pour que les étudiant·es reçoivent de nombreux feedbacks formatifs, mais également à des fins d'enseignement des contenus du cours. Selon cet auteur, trois éléments doivent être travaillés avec les étudiant·es afin qu'elles et ils puissent réellement bénéficier de l'évaluation entre pairs : la compréhension de la tâche, la compréhension de ce qu'est une

¹ Par perception, nous entendons la compréhension construite par une personne au travers d'expériences sensorielles et cognitives d'objets, de personnes, d'environnements, de situations. Parmi les éléments constitutifs de la perception, Lewis (2001) distingue (a) la présence d'une personne perceptrice, (b) la présence d'un objet (personne, situation, événement, relations, ou objet) qui est perçu, (c) un contexte dans lequel l'objet est perçu, et (d) un processus de perception qui débute avec l'expérience sensorielle de multiples stimuli et se termine par la formation de percepts (entités cognitives se rapportant à la perception).

² *Sustainable assessment* est défini comme une évaluation « that meets the needs of the present and prepares students to meet their own future learning needs » (Boud, 2000, p.151)



évaluation de qualité, et la compréhension des critères d'évaluation. Cette focale sur le développement des compétences évaluatives des étudiant·es est essentielle dans la perspective d'une évaluation « durable » qui répond aux besoins présents de l'évaluation tout en préparant les étudiant·es à répondre à leurs besoins d'apprentissage futurs (Boud, 2000 ; Boud & Falchikov, 2007 ; Tai et al., 2017).

Alors que les effets des feedbacks ont été largement étudiés, peu d'études à notre connaissance se sont intéressées à la réceptivité des étudiant·es aux feedbacks, et encore moins à leur perception de tâches de production de feedbacks. Au-delà du contenu du feedback, nous pensons que la manière dont celui-ci est perçu participe de l'apprentissage de l'étudiant·e et interagit avec les actions de régulation entreprises suite au feedback. De même, nous considérons que le fait de produire un feedback convoque un ensemble de processus participant également à l'apprentissage des savoirs en jeu. Ce sont ces processus que cette étude vise à mettre en lumière, en donnant à voir comment les étudiant·es ont perçu les tâches de production et de réception de feedbacks de leurs pairs, ainsi que de quelle manière ces tâches leur ont permis, toujours selon leurs perceptions, d'apprendre les contenus du cours ou des compétences plus transversales.

Ainsi, un des buts de cette étude est d'étendre la compréhension des perceptions de l'évaluation par les pairs d'étudiant·es, dans le but de mieux saisir comment promouvoir des dispositifs incluant des pratiques d'évaluation par les pairs qui encouragent une évaluation-soutien d'apprentissage (AfL) et une évaluation comme moyen pour apprendre (AaL) dans des cours universitaires. L'étude vise également à interroger le rapport de l'évaluation au numérique, dans une perspective « d'évaluation bienveillante et exigeante par le numérique³ », au regard plus spécialement de la caractéristique de l'anonymat des feedbacks produits et reçus par les étudiant·es.

Dans le texte qui va suivre, l'évaluation se situe à deux niveaux :

- La thématique du cours dans lequel le dispositif a été mis en place est l'évaluation. Les *contenus* travaillés par les étudiant·es portent donc sur l'évaluation.
- L'*activité* dont fait état cet article est l'évaluation par les pairs.

En préambule, il est encore utile de préciser que le format de cette contribution ne suit pas la structure « classique » d'un article scientifique, mais se calque sur l'approche méthodologique adoptée. En effet, cette étude a privilégié une approche inductive qui se reflète dans la structure de l'article. Ainsi, nous plongerons dans l'empirie directement après la description du dispositif. Une fois les résultats présentés, des liens seront faits avec la littérature existante sur l'évaluation par les pairs dans la discussion. La discussion explorera une littérature diverse issues de différents courants, sélectionnée en fonction des résultats émergents.

2. Description du dispositif d'évaluation continue

Nous adoptons ici la perspective située (Mottier Lopez, 2008), dans laquelle « on ne peut pas définir les apprentissages évalués indépendamment des tâches et plus généralement des situations et contextes qui, pour ces derniers, contribuent également à définir les attentes évaluatives » (Mottier Lopez, 2015b, p.62). Pour cette raison, il est nécessaire de

³ Préoccupation débattue dans un séminaire organisé à Metz en février 2018. <http://www.ac-nancy-metz.fr/seminaire-evaluation-bienveillante-et-exigeante-par-le-numerique--111302.kjsp>



commencer par détailler le dispositif et les tâches d'évaluation continue sur lesquelles s'appuie la présente étude. Une description volontairement « épaisse » du dispositif est disponible en annexe ([Annexe A](#)) afin d'en permettre sa reproduction. Une compréhension du dispositif dans son ensemble est nécessaire pour la compréhension de l'étude, mais il convient de préciser que les résultats se basent principalement sur la perception de deux tâches d'un scénario pédagogique⁴ : la production et la réception de feedbacks entre pairs.

Le dispositif dont il est question dans cet article est une **évaluation continue pour apprendre** qui a été expérimentée dans un cours de Bachelor à l'université de Genève intitulé « Évaluation et régulation des apprentissages » au semestre d'automne 2018. Le cours était donné par la professeure Lucie Mottier Lopez secondée par Céline Girardet (assistante post-doctorante) et comptait 73 étudiant·es inscrit·es.

La visée du dispositif était d'amener les étudiant·es à co-élaborer des savoirs sur des thèmes relatifs à l'évaluation formative et certificative au travers de trois travaux pratiques qui faisaient vivre aux étudiant·es les types d'évaluation étudiés. Il s'agissait donc d'amener les étudiant·es à étudier l'évaluation en la pratiquant, dans un cours qui vise à fonctionner comme une communauté d'apprentissage⁵. Nous nous situons donc dans des démarches d'AfL/AaL qui proposent des tâches évaluatives dites « authentiques » (Wiggins, 1990).

Le dispositif pensé intègre les fonctions formative et certificative de l'évaluation, en partant du postulat que ces différentes fonctions peuvent être complémentaires (Mottier Lopez, 2015a). L'hypothèse est qu'une évaluation certificative construite dans la continuité du semestre soutient l'engagement des élèves dans les tâches évaluatives formatives qui sont proposées au fil des cours.

L'ensemble des tâches et activités à réaliser étaient transmises puis centralisées sur Moodle, en tant que la plateforme numérique qui « véhicule ainsi une fonction réflexive à deux niveaux : c'est un miroir du vécu et une occasion d'analyse de sa propre activité » (Mailles-Viard Metz, 2015, para.14), facilitant ainsi des démarches d'autoévaluation et d'évaluation entre pairs. Plus précisément, les buts étaient :

- de permettre une gestion plus facile de toutes les pièces d'un dispositif complexe au niveau des tâches à réaliser (en présentiel, hors des cours), de leur articulation et de leur temporalité ;
- pour les étudiant·es, de télécharger les documents à remplir pendant les TP puis de les redéposer à la fin de chaque séance de façon autonome ;
- de distribuer hors des cours différents travaux aux étudiant·es selon une répartition organisée pour la réalisation de feedbacks individualisés ;
- de faciliter l'anonymat des étudiant·es dans les tâches d'évaluation formative et de rétroactions entre pairs ;

⁴ Nous adoptons la proposition de Daele et al. (2003), selon laquelle « le scénario pédagogique est la partie d'un dispositif de formation qui décrit le déroulement des activités d'enseignement et d'apprentissage. Le dispositif met à la disposition du scénario des moyens logistiques et des ressources (techniques, humaines, administratives, etc.) pour être mis en œuvre » (p.6). De plus, le dispositif est pensé au sein du contexte institutionnel et organisationnel dans lequel il prend place (Daele et al., 2003).

⁵ Nous reprenons ici une définition de la communauté d'apprentissage de Bielaczyc et Collins (1999) : « une communauté où se manifeste une culture d'apprentissage dans laquelle tous sont impliqués dans un effort collectif de compréhension » (p.271), traduite par Laferrière (2005, p.9). Nous utilisons ce concept pour présenter les visées du dispositif, mais nous n'y reviendrons pas dans la suite de ce texte, car, bien que très intéressant, il n'entre pas directement dans nos questions de recherche.



- de centraliser les consignes associées à chaque pièce, afin que les étudiant·es puissent facilement s'y référer ;
- de déposer et de centraliser les rendus des étudiant·es qui étaient en grand nombre ;
- plus généralement, d'offrir un espace gardant les multiples traces des productions des étudiant·es, susceptible d'être facilement mises en relation entre elles et avec les ressources de l'enseignement (supports des cours, lectures obligatoires et d'approfondissement, notamment).

Le déroulement détaillé du dispositif est décrit dans l'[Annexe A](#), en incluant une schématisation des tâches de l'évaluation continue et des informations sur leur organisation concrète sur la plateforme Moodle. Afin d'avoir un meilleur aperçu des tâches demandées aux étudiant·es, des exemples de tâches réalisées sont également disponibles en annexes de l'article.

Tâche 1 : TP1. Les étudiant·es ont travaillé en groupes de trois à partir d'un même cas concret de productions d'élèves. La consigne était d'élaborer, à partir du cas concret, des démarches d'évaluation formatives soit du point de vue des régulations de l'enseignement et de l'apprentissage par l'enseignant·e (thème A), soit du point de vue de l'implication de l'élève dans l'évaluation (thème B), et ce en mobilisant des catégories conceptuelles présentées dans le cours. Un exemple de TP1 est disponible en annexe ([Annexe B](#)).

Tâche 2 : Feedbacks entre pairs. Les enseignantes ont attribué à chaque étudiant·e trois TP1 : un sur le même thème que leur propre TP1 et deux sur l'autre thème. Pour chaque travail, l'étudiant·e a dû rédiger un feedback formatif en fonction de plusieurs critères (entre autres, le critère de pertinence de la démarche décrite par rapport au cas concret et le critère de justesse des concepts mobilisés par les pairs). Les commentaires devaient être bienveillants⁶ et avoir une fonction formative (le cours de la semaine 7 a été consacré aux principes du commentaire formatif⁷). Un exemple de feedback est présenté en annexe ([Annexe C](#)).

Tâche 3 : TP2. Hors du cours, avant la seconde session de travaux pratiques, les étudiant·es ont dû prendre connaissance des feedbacks des pairs (environ 10), toujours anonymes, sur leur TP1. Lors du TP2, les étudiant·es ont ajusté en groupe (mêmes groupes que lors du TP1) leur production initiale au regard des feedbacks reçus. Les buts étaient d'améliorer leur production, de la compléter, et surtout de donner à voir une progression de leur réflexion depuis le TP1. Pour cela, les enseignantes ont demandé aux étudiant·es d'activer le suivi des modifications de Word et de justifier les modifications apportées. Un exemple de travail réalisé dans le TP2 est disponible dans les annexes ([Annexe D](#)).

Tâche 4 : TP3. Le TP3 était structuré en deux tâches à réaliser individuellement. La tâche 4a consistait à formuler des critères pour une évaluation formative et pour une évaluation

⁶ Durant le cours, les recommandations présentées aux étudiant·es pour formuler des commentaires bienveillants étaient : (a) pas d'expression d'exaspération, d'ironie, de mépris, ... ; (b) pas de jugement sur la personne ; (c) s'appuyer sur des faits objectifs, observables ; et (d) être prudent·e dans ses interprétations (Mottier Lopez, 2018).

⁷ L'objectif de ce cours communiqué aux étudiant·es était de connaître les caractéristiques d'un feedback formatif. Le cours (Mottier Lopez, 2018) portait notamment sur : la définition du feedback au sens large, les visées du feedback formatif, les effets et limites des feedbacks sur les apprenant·es, les recommandations issues de la littérature pour produire un feedback formatif (clair, spécifique, lié à des critères, ne portant pas sur la personne, contenant des pistes de régulation et d'amélioration, relevant des aspects positifs, etc.). Le cours était basé, entre autres, sur les travaux d'Allal & Mottier Lopez, 2005 ; Black & William, 1998 ; Hattie & Timperley, 2007). De plus, la méta-analyse de Hattie & Timperley a été mise à disposition des étudiant·es.



certificative, prenant appui sur la même situation concrète que pour les TP1 et 2. La tâche 4b demandait une réflexion critique sur l'expérience vécue dans le dispositif du cours, plus particulièrement sur la rédaction et la réception des feedbacks. Un exemple de tâche 4b du TP3 est présentée en annexe ([Annexe E](#)).

3. Méthodologie

3.1 Participant·es

Les participant·es à la présente étude étaient les 73 étudiant·es (57 femmes et 16 hommes) inscrit·es au cours de Bachelor « Évaluation et régulation des apprentissages » au semestre d'automne 2018. Soixante-deux étudiant·es étaient inscrit·es en Bachelor en Sciences de l'éducation, et 11 en complément de formation MESP (maîtrise en enseignement spécialisé).

3.2 Récolte des données

Notre étude s'appuie sur la tâche réflexive du TP3 afin d'interroger les perceptions des étudiant·es de l'exercice d'évaluation par les pairs en double aveugle, du point de vue :

- de l'écriture des feedbacks : Question 1 - « Qu'est-ce que cela vous a apporté de devoir rédiger des feedbacks sur les TP des pairs ? Quelles ont été vos difficultés pour rédiger ces feedbacks aux pairs ? Justifiez. »
- de la réception des feedbacks : Question 2 - « Vous avez donc reçu des feedbacks de la part d'autres étudiant·es pour réaliser le TP2. Qu'est-ce que ces feedbacks vous ont apporté ? Quelles sont les limites de ces feedbacks selon vous ? Justifiez. »

Cette tâche réflexive a été conçue dans l'optique de répondre à trois visées :

1. Une visée de formation à l'enseignement, en encourageant la réflexivité des étudiant·es.
2. Une visée d'amélioration du dispositif d'évaluation continue proposé aux étudiant·es grâce à leurs perceptions.
3. Une visée de recherche sur l'évaluation entre pairs.

Les étudiant·es ont indiqué leur consentement pour l'utilisation de leurs réflexions critiques à des fins de recherche et d'enseignement⁸.

Des indicateurs quantitatifs recueillis au moyen des questionnaires anonymes standards d'Évaluation des Enseignements par les Étudiant·es (EEE) de l'Université de Genève⁹ seront parfois mobilisés afin d'étayer nos propos. Le recueil et l'analyse de ces questionnaires n'étant pas réalisés par les enseignantes-chercheuses mais par un service institutionnel

⁸ Étant donné que les réflexions critiques faisaient partie de l'évaluation continue du cours, nous pouvons nous demander si les étudiant·es ont pu faire preuve d'une certaine complaisance, malgré l'engagement des enseignantes à analyser ces travaux uniquement après les avoir anonymisés. Ceci peut constituer une limite à cette étude.

⁹ Pour des informations sur cette procédure institutionnelle, voir <https://www.unige.ch/dife/enseigner-apprendre/soutien-enseignement/evaluer-son-enseignement/formation-de-base/>



spécialisé, nous ne décrivons pas ici d'éléments méthodologiques et nous ne mobiliserons ces indicateurs qu'en soutien à la discussion.

3.3 Analyse des données

L'analyse des écrits réflexifs a été réalisée avec l'aide du logiciel NVivo et a consisté en un codage thématique dans une approche générale d'analyse inductive (Blais & Martineau, 2006). Les réflexions critiques des étudiant·es ont fait l'objet d'un codage séparé pour chaque question (écriture et réception des feedbacks). Le codage a été fin et détaillé afin de « coller » le plus possible aux productions des étudiant·es. Ensuite, pour chacune des deux questions, les codes ont été regroupés dans des catégories un peu plus larges, tout en gardant le codage fin pour ne pas perdre de nuances dans le discours des étudiant·es. Le codage thématique a donc résulté en deux grilles de codes (une par question) constituées – dans le langage NVivo – de nœuds-enfants (sous-codes) et de nœuds-parents (codes), les nœuds-parents ayant émergé d'une catégorisation des nœuds-enfants. Pour la question sur la production de feedbacks, 421 segments de texte ont été codés en 35 nœuds-enfants, regroupés ensuite en 12 nœuds-parents. Pour la question sur la réception de feedbacks, 504 segments ont été codés en 42 nœuds-enfants, regroupés ensuite en 16 nœuds-parents. Les nœuds-parents retenus pour cet article étaient ceux comportant le plus de segments ainsi que d'autres qu'il semblait pertinent d'inclure malgré un nombre de segments plus limités. Ainsi, bien que l'analyse ait été guidée par les objectifs de la recherche, les catégories présentées dans les résultats proviennent des données brutes et non de catégories préexistantes dans la littérature ou établies en amont par la chercheuse. Une fois les résultats obtenus, les catégories ont été confrontées à la littérature existante. Cette dernière étape de l'analyse est présentée dans la discussion.

Le choix de l'analyse inductive découle de la première visée de la tâche réflexive : la formation à l'enseignement. Il nous est paru important de donner du sens aux données brutes en partant des mots des étudiant·es pour en dégager du sens, afin que nos résultats puissent être réutilisés dans une visée de formation. En effet, comme l'indiquent Blais et Martineau (2006), l'utilisation de l'analyse inductive permet, entre autres, de « condenser des données brutes, variées et nombreuses, dans un format résumé » (p.4). Résumé qui, dans notre cas, pourra constituer un outil de formation.

Afin d'assurer la rigueur de l'analyse thématique, la chercheuse a donné à voir les catégories à l'enseignante principale du cours, afin de vérifier si celles-ci faisaient sens pour elle également. Cette démarche correspond au critère de confirmation externe de Mucchielli (1996)¹⁰. De plus, une première synthèse de l'analyse des tâches réflexives (anonymisées une fois les points attribués) a été présentée aux étudiant·es lors du bilan final du cours. Celles et ceux-ci ont pu donner leur avis sur cette première synthèse, ce qui a permis une vérification de la pertinence des catégories au regard de leurs expériences personnelles. Pour aller plus loin, la prépublication de l'article a été envoyée aux participant·es dont l'avis quant à l'adéquation des résultats avec leurs perceptions a été sollicité (vérification auprès des participant·es ; Blais & Martineau, 2006). Nous n'avons reçu aucune demande d'ajustement suite à cette démarche.

¹⁰ Une limite réside dans le fait que nous n'avons pas calculé d'accord interjuges, ni de « codage parallèle en aveugle » (Blais et Martineau, 2006). Une validité externe a cependant été effectuée par des discussions sur les catégories de codage entre l'autrice de ce texte et l'enseignante du cours.



4. Résultats

Les résultats font état de l'analyse des perceptions de la rédaction et de la réception des feedbacks des pairs par les étudiant·es. Nous avons choisi de présenter ces résultats par thème, organisés en quatre axes, une catégorisation issue de notre analyse des données empiriques qui regroupe nos principaux constats :

- Apprentissages des concepts étudiés en le cours
- Développement de compétences « transversales »
- Expérimentation d'une posture d'évaluateur·rice et bénéfiques pour la vie de futur·e enseignant·e
- Rapport à l'anonymat des productions des pairs facilité par le numérique

En préambule, il est important de mentionner que, globalement, les étudiant·es ont perçu ces activités comme enrichissantes, mais difficiles. Les résultats ci-dessous montrent que le fait de surmonter ces difficultés semble avoir favorisé l'apprentissage des étudiant·es.

4.1 Apprentissages des concepts étudiés en cours

Les principaux apports perçus par les étudiant·es, tant dans la production que la réception des feedbacks, ont été l'appropriation des concepts vus en cours ainsi que l'amélioration de leur travail de groupe pour le TP2. Les processus d'apprentissage exprimés par les étudiant·es semblent différer d'une activité à l'autre, c'est pourquoi les deux activités seront traitées séparément.

4.1.1 Apprentissage des concepts par l'activité de production de feedbacks

Pour rappel, chaque étudiant·e a dû effectuer trois feedbacks sur les TPs produits par d'autres groupes que le leur. Le principal apport de cet exercice selon les étudiant·es, mentionné par 46 étudiant·es¹¹, a été que l'exercice leur a permis une autoévaluation de leur propre production lors du TP1. Vu qu'ils devaient ensuite (lors du TP2) améliorer leur texte, les étudiant·es ont mentionné que ce travail de rédaction de feedback leur a permis « de pouvoir juger la qualité de [leur] propre travail par rapport à un autre » (DEV¹²). Cela leur a permis de « voir d'autres idées pour la résolution d'un même problème » (DEV).

En plus de découvrir de nouvelles idées pour le TP2, cette autoévaluation leur a également permis de vérifier leur connaissance des différents concepts étudiés, comme mentionné par cet étudiant :

Les feedbacks rédigés aux pairs m'ont également permis de visualiser les concepts dans leur contexte. Il y avait énormément de choses intéressantes à lire et l'avis de chaque groupe sur le TP, m'ouvrait parfois une nouvelle compréhension des concepts. (NAB)

¹¹ Tout au long de la section « résultats », nous utilisons des nombres pour donner du sens aux données analysées. Nous le faisons à des fins descriptives, et non à des fins statistiques. Nous refusons en effet la prescription tacite selon laquelle les nombres et les quantités doivent être bannis des recherches qualitatives. Au contraire, nous considérons, comme le souligne Sandelowski (2001), que le fait de compter et de quantifier fait partie intégrante de tout processus d'analyse, y compris qualitative.

¹² Chaque étudiant·e s'est vu·e attribuer un code de trois lettres systématiquement formé d'une combinaison de lettres figurant dans ses nom(s) et prénom(s).



En effet, lire des pistes auxquelles ils et elles n'avaient pas pensé les a amené·es à se « questionner sur la pertinence de ce qu'ils [les pairs] proposaient » et ainsi d'apprendre « de nouveaux éléments sur l'évaluation formative et de nouveaux raisonnements » (SEA).

La meilleure compréhension des concepts grâce à cette activité de rédaction de feedbacks (mentionnée par 37 étudiant·es) semble avoir été amenée de deux façons. Le premier processus, mentionné ci-dessus, semble se référer à la comparaison entre sa propre compréhension des concepts et la manière de les mobiliser par les pairs. Le deuxième moyen d'appropriation des concepts du cours a été verbalisé par les étudiant·es en termes de sentiment de responsabilité. Cela a débouché sur une exigence envers soi-même, de besoin de se sentir « expert·e » des contenus, ou plus souvent, d'un sentiment d'insécurité (N=43) lié notamment à une perception de manque de compétences nécessaires (N=26) ou à un manque de légitimité (N=13) pour la réalisation du feedback. Ce sentiment d'insécurité, cité comme principale difficulté rencontrée par les étudiant·es dans la rédaction de leurs feedbacks, les a amené·es à « bien revoir [leur] compréhension des concepts pour [s]'assurer de ne pas faire un feedback incorrect » (EVF).

Le fait de devoir corriger des travaux demande au correcteur une certaine maîtrise des concepts pour pouvoir corriger de manière claire et compréhensible pour les pairs. J'ai ainsi repris mon cours à plusieurs reprises pour comprendre dans les moindres détails les différents concepts et ainsi diminuer certaines lacunes. J'ai ainsi pu par exemple, faire la différence entre une grille de vérification et une liste d'appréciation. (ALC)

Ce travail m'a permis une implication beaucoup plus approfondie dans le cours lors de la rédaction du feedback. En effet, je voulais pouvoir faire le meilleur retour possible et pour cela il m'a fallu reprendre le cours en entier et l'approfondir. Mes difficultés ont été le fait que je ne suis pas une experte et cela m'a donc demandé un travail au préalable. (NAL)

Ces extraits tendent à montrer une représentation de l'évaluateur·rice comme nécessairement expert·e, qu'il serait intéressant d'étudier davantage dans nos recherches futures. Le fait même de donner à un individu le rôle de l'évaluateur·rice semble amener à la représentation d'une certaine obligation d'expertise sur le contenu évalué¹³. De manière intéressante, mettre les étudiant·es dans une posture d'expert·e qu'elles et ils ne sont pas encore semble avoir généré une tension entre leur rôle d'apprenant·es et leur rôle projeté d'expert·e (les faisant alors ressentir une certaine insécurité). Cette tension semble avoir participé à les faire prendre conscience de ce qu'elles et ils ne savent pas encore, constituant ainsi un moteur d'apprentissage.

Alors que le sentiment d'insécurité a pu être un tremplin pour la plupart des étudiant·es pour approfondir leur compréhension des concepts, celles et ceux-ci ont tout de même relevé qu'il s'agissait d'une difficulté majeure pour la réalisation de l'exercice, qui a pu les faire se sentir « mal à l'aise de prendre la place de 'l'expert' » (MAB), avoir « peur de faire faux » (CER), se trouver dans une situation « relativement stressante car [elle·il] ne voulait [pas] se tromper » (ALC), ou encore dans une situation « délicate » (ALC, ARS), « difficile » (SAK ; VIM) ou même « malaisante » (MAJ). Cette difficulté est bien résumée par les propos de l'étudiant ci-dessous :

Il était difficile pour moi de donner mon avis car je ne maîtrisais pas totalement tous les concepts vus en cours... Comment dire à un groupe qu'ils ont fait une erreur si nous ne sommes pas sûrs de ce que nous avançons ? En fait, on doit adopter une position d'expert sans pour autant l'être vraiment. (CHF)

¹³ Nous précisons qu'il n'a jamais été dit aux étudiant·es qu'elles et ils devaient endosser une telle posture d'expertise.



Le sentiment d'insécurité, principale difficulté mentionnée (N=43), semble avoir produit dans la majorité des cas une meilleure compréhension des concepts, un des principaux apports mentionnés par les étudiant·es en ce qui concerne la production de feedbacks (N=37). Ainsi, il est intéressant de constater que l'apprentissage des concepts semble, dans cette situation, passer par le fait de surmonter une difficulté et un sentiment de malaise engendré par la tâche de devoir « évaluer des pairs alors qu'[elles et ils sont] au même niveau » (MAJ). Il est important de préciser qu'afin de surmonter cette difficulté, les étudiant·es rapportent un retour aux contenus enseignés et autres ressources disponibles sur Moodle. De ce fait, nous postulons que l'activité de production de feedback a également entraîné les étudiant·es à aller rechercher les informations appropriées.

4.1.2 Apprentissage des concepts dans l'activité de mobilisation des feedbacks reçus

Avant le TP2, chaque étudiant·e a reçu entre 8 et 10 feedbacks de leurs pairs sur sa production du TP1. Le principal apport de la lecture de ces feedbacks, mentionné par 54 étudiant·es, a été de leur permettre d'améliorer leur production de groupe lors du TP2, en réalisant les forces et faiblesses de leur travail (N=28), en pointant des besoins de clarification (N=13), en permettant un questionnement et une remise en question sur leur travail (N=17) et en apportant de nouvelles idées (N=7).

La lecture des feedbacks a également permis une meilleure appropriation des concepts étudiés (N=40), qui semble cette fois s'exprimer au travers d'une affirmation des positions (N=19), comme le décrit cette étudiante :

La lecture des feedbacks m'a beaucoup apporté. Cela a signifié un continuum dans une appropriation de plus en plus approfondie et éclairée du contenu du cours. En effet, j'étais capable de me dire 'lui il a raison, lui il n'a rien compris' (sans être méchante). (AIF)

Cette affirmation des positions semble avoir été facilitée par le fait que les feedbacks aient été produits par des pairs et non par les enseignantes :

Lorsque ce sont les enseignants qui écrivent les feedbacks, j'ai tendance à appliquer ce qu'ils me disent et ne pas oser défendre ma position. Je dois reconnaître que je mets un peu ce que je pense de côté. Le fait que les feedbacks aient été écrits par mes collègues a eu pour conséquence de me permettre d'oser plus facilement prendre position et d'argumenter lorsque je n'étais pas d'accord avec un commentaire, chose que je n'aurais sûrement pas fait avec le feedback d'un enseignant. (LAA)

Il semblerait que la combinaison des deux pièces complémentaires du dispositif – la production et la réception des feedbacks – ait facilité l'appropriation des concepts étudiés en cours, ce qui a permis aux étudiant·es de se sentir assez « expert·es » des contenus pour pouvoir argumenter et affirmer leurs positions face à d'autres étudiant·es.

Cette affirmation des positions semble ne pas avoir eu lieu dans tous les cas. Le sentiment d'insécurité lié à la posture d'étudiant·e non-expert·e, qui semble avoir suscité un dépassement et une appropriation des concepts lors de l'écriture des feedbacks, semble avoir représenté un obstacle majeur lorsque les étudiant·es ont dû mobiliser les feedbacks de leurs pairs pour le TP2. En effet, l'obstacle principalement mentionné par les étudiant·es pour la réception des feedbacks (N=33) est la perception d'un manque de fiabilité des feedbacks reçus.



Comme je recevais les feedbacks de la part de mes collègues, je ne savais pas à quel degré, je pourrais leur faire confiance. (NAN)

Une autre limite serait le fait que nous ne sommes pas sûr que l'étudiant ayant fait le feedback l'ait fait consciencieusement et donc nous ne sommes pas totalement sûr que ce qu'il avance est correct. Peut-être n'a-t-il pas compris une notion et nous envoie un feedback faux. (DEV)

Le sentiment d'insécurité et de manque de fiabilité liés à la posture d'évaluateur·rice non-expert·e semble donc avoir représenté un obstacle à l'apprentissage au moment de la réception des feedbacks des pairs. À noter que les étudiant·es n'ont eu que peu de temps avant le TP2 pour prendre connaissance des travaux des pairs. On pourrait faire l'hypothèse qu'avec davantage de temps, cette limite aurait également suscité une appropriation des concepts, car les étudiant·es auraient eu plus de temps pour « vérifier » la qualité des feedbacks reçus, une vérification qui a été tout de même mentionnée par quelques étudiantes (N=3). Néanmoins, nous postulons également que la mobilisation de la compétence à aller rechercher des informations a été mise en œuvre au niveau de la réception des feedbacks, et l'a encore davantage été lors du TP2, lorsque les étudiant·es ont dû réviser leur production initiale. Nous concluons cette section avec un extrait qui résume un cas « idéal » d'apprentissage lors de la réception de feedbacks des pairs :

Je reconnais avoir été dans un premier temps un peu sceptique lorsque Madame Mottier Lopez nous a annoncé que les feedbacks qu'on allait recevoir allaient être produits par nos collègues et que ça ne serait par conséquent pas les enseignantes qui allaient les écrire. Étant donné que l'intégralité des feedbacks que j'avais pu recevoir dans ma scolarité avait été produits par mes enseignants, je pense que j'ai eu peur que les feedbacks de mes collègues ne soient pas corrects. Lorsque j'ai dû lire les 9 feedbacks faits par mes collègues pour préparer le TP2, mes craintes se sont très rapidement envolées et j'ai pu voir l'intérêt de l'exercice qui m'avait été proposé. Alors que j'avais l'habitude de lire les feedbacks de mes enseignants sans vraiment me poser de question, l'exercice a été dans ce cas très différent. Je me connais suffisamment pour pouvoir affirmer que si les feedbacks avaient été réalisés par les enseignantes, je me serais contentée de les lire et d'appliquer sans me poser de questions les conseils donnés par ces dernières. Étant donné que les auteurs des feedbacks que j'ai reçus étaient dans ce cas mes collègues, je les ai lus avec un esprit beaucoup plus critique. Le fait de savoir que tout n'était peut-être pas correct dans les feedbacks reçus m'a permis de faire un travail de recherches, de vérifications qui m'a énormément aidé à comprendre les concepts vus en classe. (LAA)

4.2 Développement de compétences « transversales »¹⁴

Au-delà de la compréhension des concepts étudiés en cours, les étudiant·es ont relevé d'autres apprentissages, liés à la thématique de l'évaluation mais pas directement aux contenus des TP1 et 2.

4.2.1 Compétences de décentration et développement de l'esprit critique

Lors de la lecture des TP1 des pairs, 13 étudiant·es ont explicitement mentionné une prise de conscience de l'hétérogénéité des pistes de régulation possibles pour la même situation. La difficulté de produire un feedback qui soit aligné à la production des pairs,

¹⁴ Dans le texte qui suit, nous utilisons le terme « compétence » pour chapeauter plusieurs éléments. Ce choix s'ancre dans notre volonté de mettre en avant le caractère entraînable de ces éléments, auxquels il est possible de former les étudiant·es.



relevée par 18 étudiant·es, a donc été reliée à un entraînement de la compétence de décentration :

Cette expérience nous a donc mis à l'épreuve sur le développement de plusieurs compétences. Dont, je pense, surtout sur la piste pour le développement de nos capacités de décentration. Une capacité qui est selon moi, essentielle pour quiconque souhaite travailler dans l'enseignement. (COP)

Ce besoin de prendre du recul par rapport à sa propre production a été exprimé en termes de « mettre une distance avec [leur] propre travail » (KAL), de « se détacher du point de vue qu'[elles et ils] avai[en]t sur les duos » (COP), de « se décentrer de ce qu'[ils et elles] avai[ent] produit avec [leur] groupe » (MEC), ou encore de « ne pas avoir une vision teintée par son propre TP » (MAC). Ces deux étudiantes ont exprimé clairement cette difficulté :

Ce qui a été difficile ici c'est d'être à la fois producteur et évaluateur du même travail. Si on reste accroché à son point de vue ou à son processus de régulation on peut louper l'opportunité d'analyser la pertinence de celle des autres. (KAK)

Les difficultés rencontrées sont d'une part, le fait qu'il est difficile de ne pas avoir une vision teintée par ton propre TP. En effet, il est difficile d'analyser le TP sans se dire 'moi je n'ai pas fait ça donc j'estime que ce n'est pas la bonne solution'. (MAC)

Cette difficulté a été surtout relevée lorsque les étudiant·es devaient produire un feedback sur un TP d'un groupe qui avaient le même thème que le leur. « En effet, il est difficile, ayant travaillé sur le même sujet, de laisser notre vision de celui-ci de côté afin de leur rendre un feedback entièrement objectif » (MIF).

À cette compétence de décentration se combine le développement de l'esprit critique, dont 6 étudiant·es ont fait mention :

Les apports de la rédaction de feedback m'a apporté, tout d'abord, un développement de mon esprit critique. Auparavant, je n'avais réellement un grand esprit critique et était plutôt indulgente lorsque l'on me demandait mon avis. Ici, j'ai dû être plus ferme et honnête avec ces feedbacks. (DEV)

La thématique de décentration est également apparue lors de la réception des feedbacks des pairs (mentionnée par 46 étudiant·es). L'hétérogénéité et la diversité des feedbacks reçus a également produit, chez quelques étudiant·es, la réalisation de l'hétérogénéité des commentaires : lire plusieurs feedbacks leur a « permis de voir qu'un travail peut être vu différemment en fonction des personnes » (CAR), comme en témoigne les extraits suivants :

Grâce au feed-back j'ai également eu l'impression de grandir sur le plan « professionnel ». En effet, chaque feed-back se basent sur des critères différents et ne corrigent ou ne se focalisaient pas forcément les mêmes points. Cela permet d'avoir une grande palette de possibilités pour améliorer son travail. (ALC)

Les feedbacks que j'ai reçu de la part de mes camarades m'ont permis de prendre du recul sur mes propres feedbacks. Se dire qu'il y a une multitude de feedbacks qui peuvent sortir d'un seul exercice dépendant de l'enseignant est très intéressant. Il existe des facteurs qui influencent le jugement évaluatif de l'enseignant et de voir que certaines personnes ont des feedbacks différents que les nôtres permet une vue plus large de ce qui est correct à faire. (KOA)

Recevoir un nombre conséquent de feedbacks a permis aux étudiant·es d'être « exposé à différentes manières de penser et par extension à différentes focalisations sur les difficultés



que présentent les élèves ciblés dans le TP1 » (MVA), de « voir de nouvelles perspectives » (MME). Cette décentration semble être jugée importante pour les étudiant·es du cours :

Je trouve toujours important de pouvoir avoir ce regard, car lorsque nous sommes trop impliqués dans une tâche nos œillères se ferment rapidement et nous sommes focalisés sur certains aspects et en mettant un nombre conséquent de côté. (ANN)

Ce qui m'a le plus apporté était de lire les feedbacks des autres groupes. J'ai pu sortir de mon avis personnel, et de ma propre vision de l'exercice et des attentes pour avoir un autre point de vue. C'est une forme de collaboration importante à mes yeux. Je pense que ce type d'action devrait se faire plus régulièrement au sein de l'institution scolaire genevoise. (KAB)

La décentration semble avoir été produite par la diversité des feedbacks reçus, diversité qui a été relevée comme une des difficultés de l'exercice. En effet, les étudiant·es se sont trouvé·es en difficulté face à des feedbacks parfois perçus comme contradictoires (N=21). En conséquence, les étudiant·es ont relevé avoir dû trier les feedbacks reçus dans leurs réflexions critiques :

Le seul élément qui m'a posé un peu soucis a été le fait de devoir faire le tri parmi tous ces feedbacks. J'ai en effet pris les critiques qui me paraissaient les plus pertinentes. Cela a donc contribué à en mettre certaines de côté qui j'en suis sûre auraient pu paraître pertinente/valable du point de vue de quelqu'un d'autres. (ANN)

Ce tri, que les étudiant·es ont commencé à faire au moment de la lecture des feedbacks, mais qui semble surtout s'être produit au moment de la révision collaborative, ne faisait pas partie des intentions des enseignantes au moment de la conception du dispositif. Cependant, il semblerait que le fait de devoir trier les feedbacks ait constitué une activité supplémentaire indispensable pour la réalisation du TP2. Trier les commentaires des feedbacks reçus a demandé des compétences évaluatives de la part des étudiant·es. À nouveau, nous voyons donc qu'une difficulté peut représenter une opportunité d'apprentissage.

4.2.2 Compétences à communiquer des appréciations

La difficulté de formuler des appréciations dans le cadre des feedbacks (choix des mots) a été mentionnée par 45 étudiant·es, que ce soit en termes de bienveillance (N=25), d'alignement à la production des pairs (N=18), de formulation de commentaires qui soient formatifs (N=16), ou de clarté et explicitation des propos (N=10). Ces difficultés à surmonter ont permis le développement de diverses compétences relatives à la communication.

Dans l'activité de rédaction des feedbacks, les étudiant·es ont pu entraîner leurs compétences communicationnelles, vu qu'elles et ils ont dû s'appliquer à formuler des appréciations claires, précises, formatives et bienveillantes.

Je l'ai pris comme une nouvelle occasion d'entraîner l'aspect formatif des commentaires, la bienveillance. Pour ce faire, je me suis constamment posé les mêmes questions : « Comment être clair et précis sans donner l'impression que l'on sait tout, comment proposer une démarche de manière concise, faire valoir quels arguments afin de justifier l'emploi de la démarche proposée plutôt qu'une autre. (THW)

Il faut rester bienveillant, trouver les mots, faire attention de suggérer des pistes d'améliorations sans pour autant donner les réponses. Il faut ouvrir à la discussion, poser des questions, mais surtout se demander si moi je recevais ce feedback est-ce que ça me permettrait de progresser, de m'améliorer ? Est-ce que les mots me blessent ? (MAJ)



Les étudiant·es ont également dû développer leur capacité à justifier leurs commentaires et appréciations, comme le décrit l'étudiante ci-dessous.

On doit éviter de s'attarder énormément sur l'énumération des points faibles de la production sans vraiment justifier cette approche. J'ai pu donc comprendre que lorsque nos commentaires ou notre avis sur quelques éléments de la production sont négatifs, il faut justifier de manière précise et ciblée en donnant des exemples spécifiques. Dans ce cas-là, il faudrait aborder les éléments faibles de façon à donner des commentaires qui peuvent renseigner l'apprenant sur les points à améliorer et les efforts qu'il doit fournir. (WES)

Ces compétences sont liées à d'autres compétences qui sont ressorties des réflexions critiques des étudiant·es, qui sont l'entraînement de la bienveillance et de l'empathie.

4.2.3 Compétences de bienveillance et d'empathie¹⁵

Le fait de devoir surmonter la difficulté de produire un feedback bienveillant a également permis aux étudiant·es d'entraîner leurs compétences de bienveillance et d'empathie.

La difficulté à produire des feedbacks bienveillants a été relevée par 25 étudiant·es.

Trouver les bons mots, les bons arguments, le bon message dans la formulation des propos et du commentaire (positif ou négatif) sans juger la personne est très, très difficile. (ARS)

Jusqu'à ce cours, je n'avais jamais réellement réfléchi sur le feed-back d'un enseignant. Je vois maintenant autrement et consciente de ses bien faits et de ses répercussions négatives qu'il peut avoir s'il n'est pas rédigé de manière correcte et bienveillante. (CAR)

Les étudiant·es, au-delà de relever la difficulté de l'exercice, ont affirmé avoir fait de leur mieux afin de produire des feedbacks bienveillants pour leur pairs. Pour cela, elles et ils ont dû se mettre à la place des étudiant·es qui allaient recevoir ce feedback. L'entraînement de l'empathie a d'ailleurs été mentionné par 5 étudiant·es.

J'ai aussi réalisé la complexité de choisir des mots qui ne blessent pas et qui soient suffisamment explicites pour qu'eux-mêmes me comprennent. J'ai donc toujours dû faire l'effort de me mettre à leur place pour vérifier que les commentaires étaient bienveillants et avaient une fonction formative. (NAC)

Malgré le soin et les efforts entrepris par les étudiant·es pour rédiger des commentaires bienveillants, une des principales difficultés abordées dans la réception des feedbacks a été le manque perçu de bienveillance (N=18).

Ceux-ci étaient censés répondre au critère de bienveillance. Pourtant cela n'a pas été le cas pour l'un d'entre eux qui était miné de remarques désobligeantes. Ces critiques fortes et crues m'ont fait remettre en question l'ensemble de mon travail et m'ont fortement découragée. (AMF)

¹⁵ Le fait d'associer l'empathie à une compétence peut être débattu, comme le donnent à voir les extraits d'échanges issus de l'évaluation ouverte et collaborative présentés en annexe de l'article. En effet, l'empathie peut être vue comme une disposition relativement stable dans le temps. Comme le décrivent Girard et al. (2014), « deux composantes peuvent être distinguées à l'intérieur du construit d'empathie, soit les composantes affectives et cognitives. La composante affective réfère à l'expérience vicariante d'émotions qui sont concordantes avec celles vécues par autrui (Wied et al., 2005). La composante cognitive fait allusion à la compréhension des sentiments d'une autre personne, soit par une simple association de sens ou par un processus plus complexe consistant à prendre la perspective d'autrui (Wied et al., 2005) » (p.465). Pour ce texte, nous gardons l'empathie sous le chapeau des compétences ; cependant, nous sommes consciente qu'il s'agira de mieux la définir.



L'un des feedbacks m'a vraiment déstabilisé au départ car il remettait vraiment tout notre travail en question et n'y a trouvé que des points négatifs. Le caractère non bienveillant d'un feedback qui ne relève aucun point positif, loin de m'inciter à revoir les choses reprochées, me rebute. Je n'ai pas analysé suffisamment ce feedback alors qu'il pouvait comporter des remarques pertinentes. (KAK)

Les réflexions des étudiant·es sur la réception des feedbacks donnent clairement à voir l'impact des mots sur la réception du feedback. Certain·es ont d'ailleurs soulevé ce point comme une prise de conscience :

Mais malgré cela, même les feedbacks que nous n'avons pas pris en compte pour la correction de notre travail, m'ont permis, personnellement, de voir ce qu'il ne fallait pas faire et de ressentir ce que les apprenants ressentent parfois en lisant des commentaires mal rédigés et qui par conséquent sont perçus comme désagréables ou pas aidant. (NEA)

Un des impacts du manque de bienveillance perçu dans les feedbacks a été, lors de la lecture, de consciemment les écarter et de ne pas les prendre en considération pour la réalisation du TP2. Ce tri en fonction de la bienveillance a été mentionné par 5 étudiant·es.

Quand j'ai trouvé des formulations un peu maladroites ou jugeantes, du coup je gardais moins en tête les commentaires parce que je n'avais pas envie d'être démotivée ou de me dire qu'il fallait tout changer dans notre production. (EVF)

La bienveillance a donc été pour certain·es un critère de sélection des feedbacks à exploiter pour le TP2. Une étudiante raconte même que pour elle et les membres de son groupe, la bienveillance faisait partie des critères permettant de juger de la fiabilité des feedbacks reçus :

Comme nous recevions différents feedbacks, tous n'allaient pas dans la même direction. Il a donc fallu se mettre d'accord sur une sélection des feedbacks que l'on utiliserait ; que l'on privilégierait dans l'optique d'une amélioration de notre travail. Nous avons choisi d'utiliser ceux qui nous avaient été rédigés par un étudiant qui semblait avoir une bonne maîtrise des concepts et de la notion de feedbacks. Il nous a paru plus pertinent d'utiliser les commentaires qui se voulaient bienveillants et qui proposaient des pistes précises pour améliorer notre travail, sans toutefois nous donner directement les réponses. Cela a été pour nous une source de fiabilité. (PAM)

Ainsi, nous pouvons voir que la perception de la bienveillance a eu une grande importance pour les étudiant·es du cours, que ce soit lors de la rédaction des feedbacks ou lors de la réception de ceux-ci. D'ailleurs, les feedbacks perçus comme bienveillants ont apporté un sentiment de satisfaction / de valorisation (N=10), de sécurité face à sa production (N=15), et de motivation (N=6).

4.2.4 Compétences de collaboration

Bien que la question posée aux étudiant·es concernait la réception des feedbacks des pairs et pas directement leur remobilisation pour réviser leur production lors du TP2, certain·es (N=7) ont souhaité mettre en avant les difficultés et/ou apports de la collaboration avec leurs pairs. Les propos présentent cette collaboration comme « une difficulté à la fois enrichissante et encombrante » (JOD), représentant à nouveau une difficulté source d'apprentissage.

Nous retrouvons dans les réflexions des étudiant·es des liens entre leur perception de la collaboration et les dimensions identifiées dans les sections précédentes :



- Le fait de devoir collaborer avec des pairs peut créer des émotions qu'il faut surmonter. Par exemple, KOA exprime comme difficulté le fait qu'« il est important de garder son calme en face d'une personne qui n'est pas en accord avec ce que vous pensez et n'est pas ouverte à la discussion ».
- La collaboration a poussé les étudiant·es à affirmer leurs positions et à débattre, ce qui mobilise « la capacité des membre du groupe à s'exprimer quitte à parfois être contre son groupe » et nécessite d'« accepter de rentrer dans une logique de conflit sociocognitif en échangeant tout en gardant son calme » (NEA).
- Ces débats semblent avoir pu créer à nouveau un certain sentiment d'insécurité. Par exemple, SIH explique s'être « rendu compte que face aux autre membre de [s]on groupe, [s]es arguments n'étaient pas assez assurés, [s]a connaissance du cours pas si claire ».
- À nouveau, la collaboration entre pairs semble avoir nécessité des compétences de décentration :

Je pense que, par la collaboration et la discussion, j'ai été poussée à me décentrer et prendre en compte les avis de mes camarades. Si, parfois nous sommes rapidement arrivées à un consensus, d'autres fois, les affrontements d'idées au sein du groupe ont nécessité un retour au cours pour appuyer ses avis. Quand l'une ne s'accordait pas avec l'autre, nous montrions dans les powerpoints du cours ou dans nos notes des passages qui appuyaient ce que nous avançons. Ces conflits sociocognitifs nous ont permis, selon moi, d'améliorer la qualité de notre travail mais aussi d'avancer dans nos compréhensions personnelles du cours. (AMF)

- Finalement, ce même extrait montre que les désaccords ont amené les étudiant·es à revenir aux ressources du cours afin de pouvoir arriver à un consensus. À noter que certain·es ont également fait appel aux enseignantes pour pouvoir résoudre un désaccord durant le TP2 en présentiel.

4.3 Expérimentation d'une posture d'évaluateur·rice et bénéfices perçus pour la vie professionnelle future

L'activité de rédaction de feedback a permis aux étudiant·es de se mettre dans la peau d'un·e enseignant·e, ce qui était une première pour la plupart d'entre eux. Vingt-et-un étudiant·es ont mentionné que l'activité leur a permis d'adopter une posture d'évaluateur·rice et discutent leur rapport avec cette posture en faisant des liens avec les concepts vus en cours (par exemple, certain·es ont réalisé qu'ils ou elles étaient influencé·es par l'ordre des TPs à commenter), ou en explicitant leur réalisation de la complexité de cette posture.

Auparavant je n'avais jamais produit de feedbacks et cette démarche m'a fait prendre conscience de la complexité de rédiger des feedbacks le plus correct et formateur possible (être précis, ne pas faire intervenir de jugement [de valeur], choisir les bons et justes mots, encourager, etc). Personnellement je trouve qu'écrire 3 feedbacks n'est pas trop en termes de quantité mais est au contraire très utile car étant une personne qui se destine à l'enseignement primaire, dans mon futur professionnel, je n'échapperai pas à la production de feedback. (SEA)

Tout comme cette étudiante, qui semble voir la production de feedbacks (et peut-être même l'évaluation) comme une fatalité à laquelle elle ne pourra pas « échapper », 11 autres personnes ont relevé l'utilité de l'exercice de production de feedback pour leur futur en tant que professionnel·les de l'enseignement.



Rédiger des feedbacks sur les TPs des pairs a ensuite été l'occasion de se confronter concrètement à un travail qu'on sera amené à faire régulièrement dans l'exercice de notre profession, à savoir la difficulté dans le choix des mots et des formules qui permettent à l'élève de lui faire un retour sur son travail qui lui donne envie de progresser. Cela n'est pas un travail évident. (PAM)

Cet exercice formatif représentait bien, selon moi, la difficulté du métier d'enseignant. D'autre part je me suis rendue compte qu'il s'agit d'un outil important qui permet à l'élève de progresser dans ses apprentissages. D'où sa difficulté à le rédiger pour qu'il ait l'effet voulu. (CAR)

L'utilité pour le futur professionnel a également été mentionnée par sept étudiant·es pour la réception des feedbacks des pairs :

Cela a été une source de régulation pour moi en tant que future enseignante. Je sais maintenant, par expérience, en quoi cela peut être déstabilisant de recevoir un retour très négatif et critique. J'ai pu voir que cela peut être fortement décourageant et n'aide en aucun cas à progresser. Cet exercice me sera utile dans mon futur professionnel. En effet, je porterai davantage d'attention à la formulation de mes phrases en faisant en sorte de rester dans une démarche critique formative et non pas blessante pour celui qui recevra ce que je rédige. (AMF)

En nous basant sur ces résultats, nous postulons donc un développement de certaines compétences évaluatives grâce au dispositif, telles que la reconnaissance des biais évaluatif, l'exercice du jugement évaluatif, l'utilisation de critères d'évaluation, ou encore la formulation d'appréciations sur le travail d'autrui, ce qui est à la fois en lien avec des apprentissages visés dans le cours et des apprentissages plus transversaux dont les étudiant·es perçoivent les bénéfices pour leur vie future.

4.4 Rapport des étudiant·es au numérique

4.4.1 Bénéfices et difficultés face au numérique

Ce point n'a été que peu relevé par les étudiant·es dans les tâches réflexives. Cela fait sens car cette tâche portait sur des aspects spécifiques du dispositif – l'écriture sur un document Word et la lecture de feedbacks – qui ne demandaient pas des compétences numériques farmineuses. Les quelques éléments ci-dessous sont donc essentiellement issus des constats des enseignantes sur les bénéfices et difficultés des étudiant·es face au numérique sur l'ensemble du dispositif.

Le bénéfice principal de la plateforme Moodle pour les étudiant·es (et pour les enseignantes !) est qu'elle permet de centraliser l'ensemble du travail et des ressources. Les résultats présentés précédemment laissent entrevoir que l'apprentissage des étudiant·es s'est souvent produit au travers de va-et-vient entre les tâches et les ressources centralisées sur Moodle (textes scientifiques, supports de cours, consignes, etc.), comme le montre cet extrait : « Par exemple lorsqu'un groupe parlait de régulation pro ou rétroactive et que je n'étais pas sûre de leur définition, je pouvais moi-même aller dans les documents vérifier et mieux intégrer moi-même ces notions » (MEM), ou encore celui-ci : « Pour réaliser au mieux les TPs, je devais rechercher les informations dans le cours et de ce fait, comme j'allais à la recherche de ces informations, les concepts rentraient mieux en mémoire et je trouve cela bénéfique pour l'apprentissage » (CAD). Nous postulons que cette centralisation des ressources et activités facilite la réflexion sur son activité (Mailles-Viard Metz, 2015) en permettant aux étudiant·es de visualiser la progression du cours et leur



propre progression dans le même espace. Nous n'avons cependant pas recueilli de données nous permettant d'investiguer ce postulat.

Une difficulté des étudiant·es face au numérique a été constatée par les enseignantes lors du déroulement du TP2 en classe. Elle concerne l'utilisation des fonctions de révision (suivi de modifications et commentaires) sur Word, qui, à notre surprise, semblaient inconnues d'un certain nombre d'étudiant·es. Nous avons dû passer dans de nombreux groupes (selon nos souvenirs, plus de la moitié) pour leur expliquer l'utilisation de ces fonctions. Le fait que certain·es étudiant·es (N=7) aient préféré réaliser le TP3 en format papier plutôt que numérique nous permet également d'avancer prudemment que le rapport au numérique pourrait peut-être encore représenter une difficulté pour certain·es. Sinon, dans l'ensemble, tout s'est passé très naturellement au niveau du numérique. La section suivante donne à voir des résultats sur le rapport à l'anonymat dans les tâches de production et de réceptions des feedbacks, une anonymisation qui a été grandement facilitée par le numérique.

4.4.2 Rapport à l'anonymat des productions des pairs

Une des difficultés relevées les étudiant·es tant dans la réception des feedbacks (N=15) que dans la production des feedbacks (N=9) était les propos des pairs parfois difficilement compréhensibles.

Pour pallier les problèmes de compréhension des feedbacks, certain·es étudiant·es (N=8) ont décrit une certaine frustration quant à l'anonymat des TP et des feedbacks. En effet, elles·ils auraient voulu pouvoir demander des clarifications aux auteur·rices. Cela était surtout le cas lors de la réception des feedbacks des pairs, comme le montre ces deux exemples :

Un retour par écrit peut être difficile si la personne qui lit le retour ne comprend pas la phrase ou le raisonnement du correcteur. Ici, une régulation des apprentissages s'avère alors beaucoup plus difficile (comment réguler quand on ne comprend pas ce qui est faux ou quelle est l'argumentation, la pertinence d'une remarque ?). C'est pour cela que je pense que les feedbacks écrits peuvent s'avérer plus délicats surtout si la possibilité d'agrémenter ces retours d'une médiation ou d'une discussion orale n'est pas possible ou pas envisagée. (ALC)

L'anonymat a cependant représenté pour moi une limite lorsque j'ai endossé la posture de réceptrice de feedbacks. L'anonymat des auteurs des feedbacks que j'ai reçu a été une limite en raison du manque de clarté de certains feedbacks que j'ai reçu. Il m'est arrivé à plusieurs reprises de ne pas comprendre ce qui se cachait derrière le commentaire d'un de mes collègues. Le fait d'avoir tout anonymisé a fait que je me suis retrouvée dans l'incapacité d'aller demander des clarifications aux auteurs des feedbacks et cela a eu pour conséquence que j'ai dû à certains moments laisser de côté des commentaires qui auraient peut-être pu être pertinents si j'avais eu la possibilité de demander des explications à ses auteurs. (LAA)

En ce qui concerne la production des feedbacks, les avis sont plus nuancés et certain·es relèvent des avantages de l'anonymat :

L'absence de mutualité dans les feedbacks a été également un apport : l'anonymat aide forcément à l'impartialité et cela force à construire des feedbacks objectifs et consciencieux. (LED)

Malgré quelques réflexions positives sur l'anonymat, la frustration de ne pas pouvoir communiquer avec les auteur·rices des TP a été exprimée par 4 étudiant·es en référence



à l'activité de production des feedbacks et semble être liée à la même difficulté, exprimée par 9 personnes : comprendre le travail des pairs.

Le fait de ne pas connaître l'identité des auteurs des travaux que j'avais à analyser m'a empêché d'aller leur demander des explications sur ce qu'ils avaient écrit. J'ai donc parfois dû laisser de côté certains éléments sur lesquels il aurait peut-être pu être intéressant de s'attarder par manque de clarté. (LAA)

Dans la lecture de certains TP, certaines parties n'étaient pas claires et je ne comprenais pas toujours ce que les auteurs essayaient d'expliquer. Étant donné que je ne savais pas qui étaient les auteurs de ces TP, je n'avais pas de possibilité de leur demander de précision. C'était donc compliqué pour moi d'établir un feedback pertinent et de leur proposer d'éventuels conseils. (SEA)

5. Discussion

Cette discussion a pour but de mettre en relation les résultats de la présente étude avec la littérature sur l'apprentissage dans des activités de feedbacks entre pairs et de modéliser les apprentissages réalisés au moyen des deux tâches « phares » du dispositif : la production de feedback et la réception de feedback à des fins de régulation. Le choix de la littérature s'est fait sur la base des résultats émergents. Ainsi, nous explorons ci-dessous (section 5.1) les apprentissages réalisés lors d'activités d'évaluation entre pairs. Nous revenons également sur la dimension numérique dans la section 5.2. Nous concluons la discussion par une section (5.3) sur des considérations au sujet du dispositif.

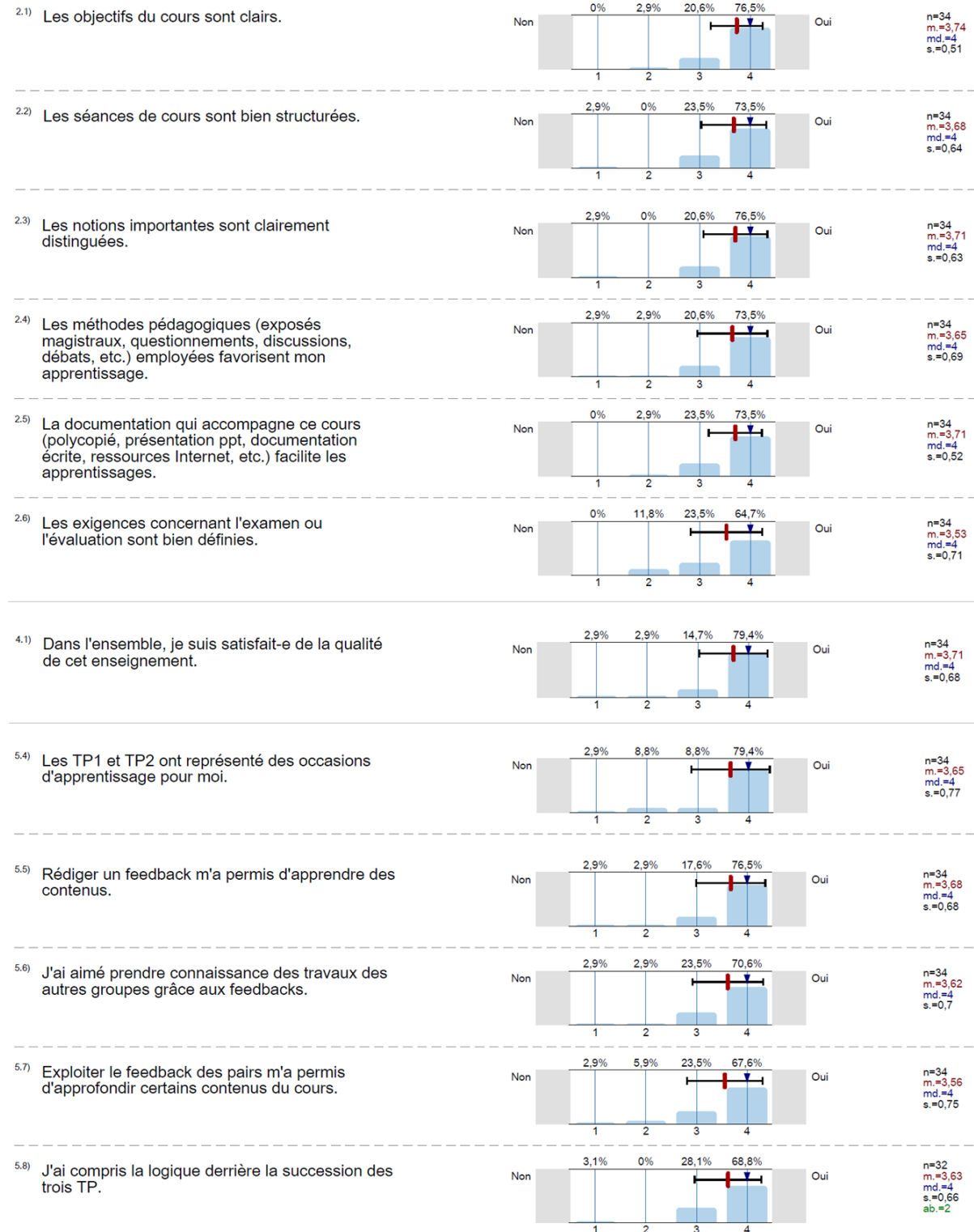
5.1 L'évaluation entre pairs comme moteur pour l'apprentissage

La puissance du scénario d'évaluation entre pairs (et du dispositif dans lequel il s'inscrit) pour soutenir les apprentissages des étudiant·es est mise en lumière par les résultats de notre étude. Dans les réflexions critiques, cela s'est notamment traduit en termes de prises de conscience, qui représentent des indices forts des apprentissages réalisés au travers du dispositif du cours. Ces prises de conscience ont porté sur :

- la difficulté de produire un feedback formatif, spécifique, bienveillant, et fondé sur des critères d'évaluation ;
- l'impact émotionnel de feedbacks perçus comme manquant de bienveillance ;
- l'hétérogénéité des jugements inter-évaluateur·rices formulés sur la même production ;
- la diversité des démarches d'évaluation formative et de régulation proposées par les étudiant·es pour le même duo d'élèves ;
- et plus largement, la complexité de l'évaluation en classe et l'importance de former les futur·es enseignant·es à l'évaluation.

L'efficacité du dispositif pour les apprentissages est corroborée par l'Évaluation de l'Enseignement par les Étudiant·es (EEE) consistant en un questionnaire institutionnel anonyme que ces dernier·es sont invité·es à remplir en fin de semestre. Dans les EEE, les étudiant·es (N=34) ont indiqué une appréciation globale du cours avec une moyenne de 3.71 sur 4 (sd = 0.68). La figure 1 montre des indicateurs concernant les questions standards et personnalisées de ce questionnaire :



Figure 1*Indicateurs de l'Évaluation de l'Enseignement par les Étudiants (EEE)*

Note. n. = nombre ; m. = moyenne ; md. = médiane ; s. = écart-type ; ab. = abstention.

La littérature montre que l'évaluation par les pairs implique des processus de réflexions, comparaisons, contrastes, et communication, qui peuvent amener l'évaluateur·rice à consolider et approfondir sa propre compréhension. Dans sa revue de littérature sur

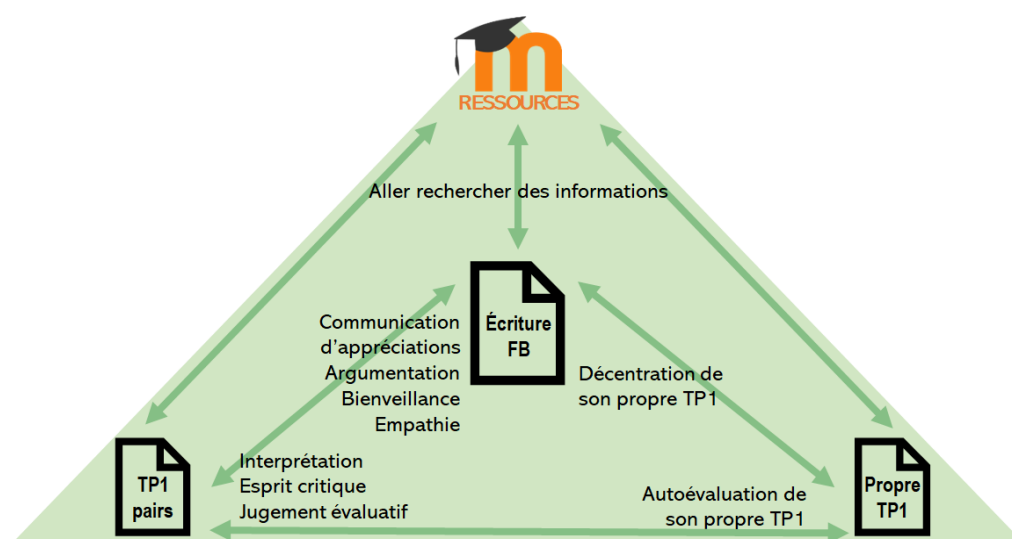


l'évaluation par les pairs entre étudiant·es universitaires, Topping (1998) nomme ce processus « *learning by assessing* ». Il a été montré que les feedbacks entre pairs contribuent à l'acquisition des contenus du cours ainsi que d'autres compétences, plus transversales. Les sections suivantes tentent de schématiser les apprentissages réalisés au travers de l'activité de production de feedbacks (section 5.1.1) et de réception de feedbacks (section 5.1.2). La littérature sur les apprentissages lors de l'implication des étudiant·es dans l'évaluation est abondante. Falchikov (2005) réalise une synthèse des bénéfices identifiés dans la littérature de l'implication des apprenant·es dans l'évaluation. Sans entrer dans les détails, l'autrice met en avant l'amélioration de la performance, le développement de compétences cognitives et métacognitives (dont la réflexivité et l'esprit critique), et le développement d'autres compétences comme la compétence à écouter, à présenter, à écrire, à apprendre tout au long de la vie. Ressortent également le développement personnel (augmentation de l'autonomie, de la responsabilité, du sentiment de compétence), le développement des compétences sociales (diplomatie, coopération), et le développement des dispositions affectives (diminution de l'anxiété et du stress, augmentation de la confiance en soi et de la motivation). Dans les sections suivantes, nous nous concentrerons sur la littérature portant plus spécifiquement sur les feedbacks entre pairs, voire sur l'évaluation entre pairs plus généralement.

5.1.1 Les apprentissages au travers de l'activité de production de feedbacks sur le TP1 des pairs

Dans notre étude, l'apprentissage des contenus du cours s'est réalisé au travers de va-et-vient entre la production initiale des pairs à évaluer, la tâche d'écriture du feedback, les ressources du cours et sa propre production initiale. L'apprentissage des contenus du cours s'est produit au travers de processus dynamiques entre les tâches et ressources, et ont convoqué ou impliqué d'autres apprentissages plus transversaux. La figure 2 tente de modéliser ces processus ayant contribué à l'apprentissage des contenus du cours.



Figure 2Apprentissages des contenus du cours lors de la production de feedback (FB) ¹⁶

Afin de produire leur feedbacks, les étudiant·es ont dû endosser la posture d'un·e évaluateur·rice. Cette tâche a en effet impliqué les étudiant·es dans les démarches d'évaluation (Allal, 2008) que sont la récolte d'informations en rapport avec l'objet et les critères d'évaluation définis en amont, l'interprétation de ces informations (nécessitant un certain esprit critique et de l'empathie), le jugement évaluatif (prises de décision) et la communication d'appréciations à autrui. Pour communiquer leurs appréciations, les étudiant·es ont dû produire des commentaires formatifs en faisant preuve de bienveillance et en argumentant leurs points de vue. Pour réaliser ces démarches d'évaluation, elles et ils ont eu besoin d'aller rechercher des informations dans les ressources centralisées sur la plateforme Moodle : critères pour évaluer les TP1, ressources sur l'écriture de feedbacks formatifs, et ressources sur l'évaluation formative et les démarches de régulation étudiées en cours. Les étudiant·es rapportent aussi avoir dû faire preuve de décentration par rapport à leur propre TP1 pour réussir à produire un feedback qui soit aligné au TP1 des pairs. L'activité de production de feedbacks les a également amené·es à revenir à leur propre TP1 et à l'autoévaluer. Bien que tout se soit déroulé de manière dynamique, avec des va-et-vient constants entre les différents éléments de la figure (symbolisés par les flèches bidirectionnelles), nous avons fait le choix de positionner les compétences à des endroits précis dans la modélisation, signifiant la proximité de la compétence avec l'activité réalisée. Nous reprenons ci-dessous ces compétences en les liant aux résultats trouvés dans la littérature.

Du côté de la lecture des TP1 des pairs : interprétation, esprit critique, jugement évaluatif

Notre étude montre que l'évaluation des productions des pairs a amené les étudiant·es à mobiliser des compétences évaluatives et réflexives, telles que l'interprétation, l'esprit critique et le jugement évaluatif. Le développement des compétences réflexives et critiques est également un constat de Yu (2019) qui a étudié les perceptions d'étudiant·es universitaires de leur apprentissage au moyen d'une activité de production de feedback aux pairs sur une production écrite académique (*postgraduate dissertation*). De manière plus globale, les résultats de la revue de littérature de Topping (1998) sur l'évaluation entre

¹⁶ Notons que la figure se base sur la production d'un seul feedback. Étant donné que les étudiant·es devaient en produire trois, les trois feedbacks à produire et les trois différents TP1 des pairs s'ajoutent à ces interactions et participent aux apprentissages.



pairs montrent un développement des compétences cognitives et métacognitives du côté de l'évaluateur·rice dans une situation d'évaluation par les pairs. La littérature qu'il convoque suggère qu'évaluer les productions de pairs est une activité coûteuse en temps et en implication, convoquant, entre autres, capacités de réflexion, esprit de synthèse, et stratégies diagnostiques.

Du côté de l'écriture du feedback : communication d'appréciations, argumentation, bienveillance, empathie, décentration

Pour écrire leur feedback aux pairs, les étudiant·es de notre étude ont dû faire preuve de compétences évaluatives et communicationnelles : la communication d'appréciation, l'argumentation, la bienveillance et l'empathie. Apprendre à justifier ses positions, ainsi que le développement des compétences communicationnelles sont également des compétences « transversales » que Topping (1998) a identifié dans sa revue de littérature sur l'évaluation par les pairs à l'université (sans différencier, par contre, entre production et réception). L'étude de Yu (2019), elle, a révélé une amélioration des compétences d'écriture académique. Si l'amélioration des compétences d'écriture académique n'a pas été relevée par les étudiant·es de cette étude (qui, contrairement aux étudiant·es de l'étude de Yu, ne devaient pas évaluer des thèses), elles et ils ont par contre perçu un développement de leur compétence d'écriture d'appréciations et de commentaires formatifs, ainsi que de leur capacité à argumenter leurs propos. Finalement, écrire un feedback semble avoir convoqué des compétences liées à la bienveillance. L'entraînement de la bienveillance n'apparaît que très peu dans la littérature. Cependant, Topping (1998), en citant Riley (1995), rapporte que l'évaluation par les pairs peut permettre un développement de la compétence de diplomatie, qui peut se rapprocher de la bienveillance. Néanmoins, il se pourrait que l'insistance de « nos » étudiant·es sur la dimension de la bienveillance soit associée à notre enseignement, du fait que la bienveillance était l'un des critères demandés pour la réalisation de leurs feedbacks¹⁷. Nous n'avons trouvé mention d'un développement de l'empathie que dans l'étude d'Hanrahan et Isaacs (2001), dans laquelle les étudiant·es ont rapporté avoir développé leur empathie envers le corps enseignant grâce à l'évaluation par les pairs, ce qui semble aussi être le cas des étudiant·es de cette étude mais qui ne correspond pas à l'empathie dont il est question ici. Dans notre étude, les étudiant·es rapportent plutôt avoir dû faire preuve d'empathie envers leurs pairs, en se mettant à leur place pour imaginer comment celles et ceux-ci allaient recevoir leurs commentaires. L'empathie a donc été placée du côté de l'écriture des feedbacks aux pairs, car elle semble avoir été éprouvée en lien avec la formulation des appréciations et commentaires.

Toujours du côté de l'écriture du feedback, mais cette fois au regard de leur propre production initiale, on trouve la compétence de décentration, ressortie comme très importante dans notre étude, mais qui n'a pas été relevée explicitement dans la littérature à notre connaissance, bien que pouvant faire partie des compétences réflexives. On pourrait se demander si cette compétence a été développée spécifiquement dans les activités demandées dans le cadre du cours dont fait objet la présente étude. En effet, l'activité proposée permettait aux étudiant·es d'être créatif·ves et de proposer une large palette de démarches de régulation pour la même situation, et les étudiant·es devaient donc évaluer des démarches différentes de celles qu'elles et ils avaient produites. Nous pouvons donc imaginer que c'est l'hétérogénéité des propositions de régulation qui a

¹⁷ À noter aussi que la notion de bienveillance ne se traduit pas de manière fidèle en anglais, et que sa traduction directe, « benevolence », n'est que rarement mobilisée en éducation et se rapporte davantage, dans le langage courant, à la notion de « bienfaisance » en français. Il n'est donc pas étonnant de ne pas trouver de résultats dans la littérature prenant en considération ce concept.



demandé de la décentration aux étudiant·es, au moment de produire un feedback sur les propositions d'autrui.

Du côté de sa propre production initiale : autoévaluation

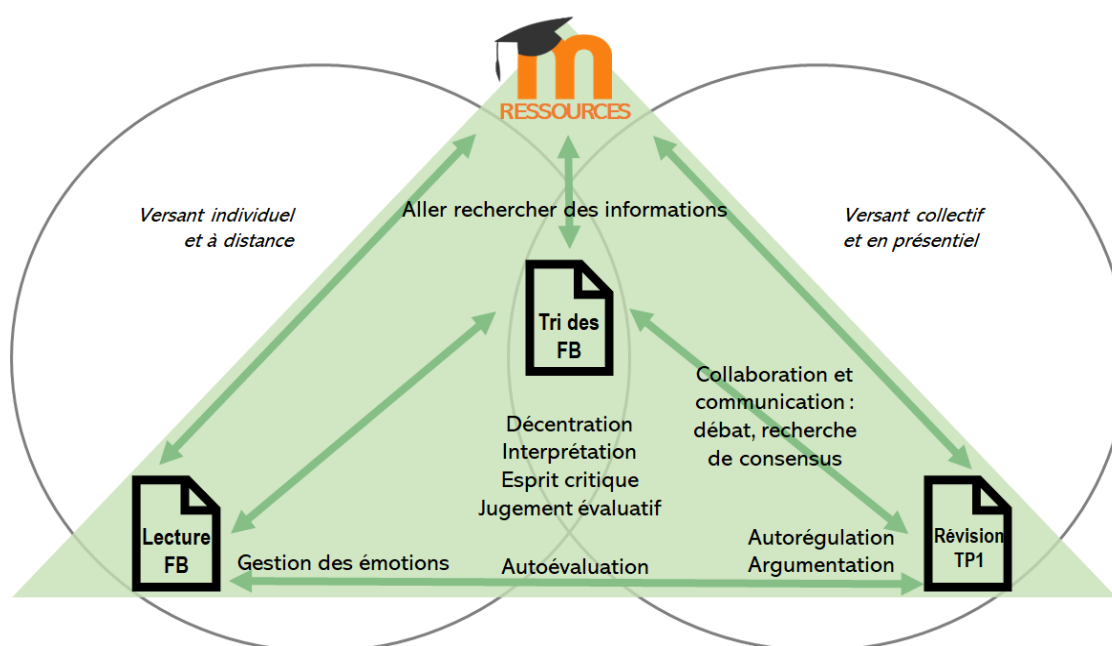
Les étudiant·es de l'étude de Hanrahan et Isaacs (2001) – qui portait sur les perceptions de 233 étudiant·es universitaires d'un dispositif d'autoévaluation et d'évaluation par les pairs sur une production écrite de type dissertation – ont également relevé les avantages de découvrir les productions d'autres étudiant·es (afin de produire un feedback sur ces productions) en termes d'apprentissage. En effet, cela permet de comparer sa propre production avec celles d'autres étudiant·es, et ainsi d'engager une autoévaluation de sa production. Cao et collègues (2019) font le même constat au travers de leur étude qualitative investiguant les perceptions des apprentissages de 15 étudiant·es lors de la production et de la réception de feedbacks entre pairs : certain·es ont même rapporté être allé réviser leurs propres productions suite à la réception de feedback. La présente étude est donc un exemple supplémentaire montrant que la production de feedbacks dans un scénario d'évaluation entre pairs est susceptible d'être liée à l'autoévaluation et encourage l'autorégulation des apprenant·es, bien que rien ne montre, dans les propos des apprenant·es, que celles et ceux-ci aient apporté à ce stade des révisions à leur production initiale au regard de cette autoévaluation.

5.1.2 Les apprentissages au travers de l'activité de réception de feedbacks dans le but d'améliorer sa production initiale

Concernant maintenant la réception de feedbacks à des fins de régulation, l'apprentissage des contenus du cours semble s'être réalisé au travers de va-et-vient entre les feedbacks reçus, les ressources du cours et sa propre production à réviser. La figure 3 tente de modéliser ces apprentissages en distinguant les deux versants de la tâche : le versant individuel et à distance, comprenant la lecture des feedbacks et une première sélection des commentaires, et le versant collectif et en présentiel, reprenant le tri des feedbacks en ajoutant la révision collaborative de la production initiale lors du TP2.

Figure 3

Apprentissages lors de la réception de feedback (FB)



Afin de pouvoir réviser leur TP1, les étudiant·es ont tout d'abord lu les feedbacks de leurs pairs. Ceci a créé de nombreuses émotions, positives comme négatives, ce qui suggère que la lecture de feedbacks implique des compétences de gestion des émotions. Une fois les feedbacks lus, les étudiant·es ont dû trier les commentaires et propositions des pairs afin de choisir lesquelles mobiliser dans leur révision. Pour réaliser ce tri, elles·ils ont dû à nouveau s'impliquer dans des démarches d'évaluation : se décentrer, interpréter, et exercer leur jugement évaluatif et leur esprit critique vis-à-vis des feedbacks, ce qui a nécessité d'aller à nouveau rechercher des informations dans les ressources du cours sur l'évaluation formative et les démarches de régulation. Une fois le tri effectué, les étudiant·es ont pu réviser leur TP1. Cette révision a impliqué des compétences d'autoévaluation, d'autorégulation, et d'argumentation (les étudiant·es devant justifier leurs révisions au regard des feedbacks reçus et de la théorie). Finalement, la révision du TP1 a impliqué des compétences de collaboration et de communication interpersonnelle, étant donné que les membres du groupe devaient se mettre d'accord sur les ajustements et justifications à apporter.

Du côté du tri des feedbacks : décentration, interprétation, esprit critique, jugement évaluatif

Un des éléments apparaissant dans la littérature est le fait que l'évaluation par les pairs permet de multiplier les feedbacks reçus (Hanrahan & Isaacs, 2001 ; Topping, 2009), offrant davantage de possibilités aux étudiant·es pour approfondir les concepts. Notre étude va plus loin en ce qui concerne la multiplicité des feedbacks. En effet, il ressort que la multiplicité et la diversité des feedbacks (hétérogènes et parfois même contradictoires) ont été perçues comme une difficulté par les étudiant·es et les ont « forcées » à engager leur jugement évaluatif et leur esprit critique afin de sélectionner des commentaires selon une articulation de critères :

- Les commentaires se recoupent (ou non) → esprit critique et analytique.
- Les propositions sont pertinentes (ou non) → jugement évaluatif, mobilisation des connaissances et besoin d'un retour aux ressources du cours.
- Les propos sont bienveillants (ou non) → stratégies de sélection selon les réponses émotionnelles (par exemple, évitement ou rejet du feedback si sa lecture engendre des émotions négatives).

Les étudiant·es ont donc dû trier les commentaires reçus et choisir lesquels mobiliser et lesquels rejeter. La revue de littérature de Topping (1998) relève d'ailleurs que l'évaluation par les pairs permet d'apprendre à rejeter des suggestions. Nous pouvons penser que ce processus de sélection a été une étape supplémentaire dans la compréhension des concepts, puisqu'il a encouragé les récepteur·rices à évaluer les feedbacks reçus en revenant aux ressources du cours afin de faire des choix dans les éléments à retenir pour améliorer leurs productions de groupe.

Du côté de la révision collaborative du TP1 : autorégulation, argumentation, collaboration et communication

Dans notre étude, les étudiant·es ont pu, grâce aux feedbacks de leurs pairs, réguler leurs productions initiales. Topping (1998) montre qu'avec un travail sur les critères d'évaluation, l'évaluation par les pairs peut amener à un travail de meilleure qualité du côté des « évalué·es » (l'auteur distingue les apprentissages pour « *the assessor* » versus « *the assessee* »). Il décrit que l'amélioration des performances et de la qualité des travaux résulte d'une meilleure capacité des étudiant·es à se situer par rapport aux attendus en comparant leur travail avec ceux des pairs. Nous pouvons donc penser que la régulation de la production initiale est favorisée non seulement au travers de la réception et mobilisation



des feedbacks, mais bien par l'ensemble du processus, incluant la production de feedbacks. Les étudiant·es ont également mentionné avoir dû faire preuve de compétences en collaboration et communication interpersonnelle lors de cette « pièce » du dispositif. Ce résultat corrobore ceux de Donia et collègues (2018) qui ont conduit une étude longitudinale sur les effets des feedbacks entre pairs dans le développement des compétences collaboratives (*teamwork skills*). Ces auteur·rices montrent non seulement que des activités de feedbacks entre pairs permettent de développer les compétences en collaboration des étudiant·es, mais également que ces compétences sont « transférées » de l'université au lieu de travail. Topping (1998) met aussi en avant un développement des *teamwork skills* au travers de l'évaluation par les pairs, telles qu'apprendre à accepter la critique, à justifier ses positions, ou encore à rejeter des suggestions.

5.1.3 Une complémentarité pour l'apprentissage : tâches de production de feedbacks et réception de feedbacks à des fins de régulation de sa production initiale, intégré dans un dispositif d'évaluation continue

Apprendre à donner et à recevoir des feedbacks créerait des opportunités pour les étudiant·es de développer différents types de compétences que lors de la seule réception de feedbacks (Rowe, 2017). Nous postulons ainsi que les deux tâches – production et réception de feedbacks – sont complémentaires pour le développement des compétences transversales et de l'apprentissage des contenus du cours. Les résultats de cette étude renforcent le constat existant que l'impact de l'évaluation entre pairs ne se manifeste pas seulement dans l'apprentissage des concepts, mais également dans le développement de compétences transversales qui peuvent varier en fonction de la nature et des objectifs de la tâche demandée.

Dans l'étude de Cao et al. (2019), les étudiant·es n'ont pas attribué les bénéfices perçus aux tâches de manière homogène. Sur les 15 participant·es à l'étude, 4 ont perçus des bénéfices uniquement dans la réception de feedbacks, 3 uniquement dans la production de feedbacks, 7 à la fois dans la production et la réception de feedbacks, et 1 n'a pas perçu de bénéfices. Nos résultats n'ont pas donné à voir les réponses des étudiant·es à la troisième question de la réflexion critique (à partir de votre expérience personnelle, quelles sont les « parties » des TP qui vous ont été les plus utiles pour comprendre les concepts du cours ?). Nous pensons utile d'indiquer ici que les réponses des étudiant·es ont été diverses et variées. En mettant en relation ce résultats avec ceux de l'étude de Cao et al. (2019), nous pouvons avancer que la complémentarité des tâches de production et de réception de feedbacks trouve également sa force dans sa capacité à satisfaire les stratégies d'apprentissages hétérogènes des étudiant·es.

Nous souhaitons ici mettre l'accent sur le développement des compétences évaluatives, qui ont un statut particulier dans la présente étude¹⁸ : elles représentent à la fois des apprentissages relatifs aux contenus du cours (le cours portant sur l'évaluation), et des compétences transversales telles qu'émettre un jugement évaluatif, faire preuve de discernement, d'esprit critique, de décentration. Mettre l'accent sur les compétences évaluatives serait en effet nécessaire au bon déroulement des activités impliquant des feedbacks entre pairs. Déjà en 1989, Sadler, en parlant de « *evaluative knowledge* » (p.135)

¹⁸ Nous faisons l'hypothèse que le développement des compétences évaluatives, rapporté dans des études portant sur l'évaluation entre pairs sur d'autres objets de savoirs (e.g., Hanrahan & Isaacs, 2001 ; Tai et al., 2017), est encore renforcé dans notre étude, par le but explicite du cours de former les étudiant·es à l'évaluation. En effet, la réflexion des étudiant·es était explicitement orientée non seulement sur les contenus relatifs à l'évaluation, mais aussi sur les mécanismes évaluatifs à l'œuvre lors de la réalisation des tâches.



ou de « *evaluative expertise* » (p.138), émet la proposition que pour qu'un feedback reçu soit efficace, les étudiant·es doivent comprendre les critères d'évaluation afin de pouvoir se rapprocher des attendus. En ce sens, un constat récurrent dans la littérature est que l'évaluation entre pairs nécessite un travail avec les apprenant·es sur les critères et procédures d'évaluation pour avoir des effets plus significatifs sur l'apprentissage (Cao et al., 2019 ; Gielen & De Wever, 2012 ; Liu & Carless, 2006 ; Sadler, 1989, 2010 ; Tsivitanidou et al., 2011 ; Yu & Lee, 2016). Lorsque c'est le cas, il a été mis en évidence que l'évaluation par les pairs permet aux apprenant·es de développer leurs compétences évaluatives comme aller rechercher des informations (Yu, 2019), comprendre et définir des critères d'évaluation, porter un jugement sur le travail d'autrui, ou encore formuler un feedback (Tsivitanidou et al., 2011), ainsi que leur confiance en elles-eux en tant qu'évaluateur·rices (Lindschouw, 2016). Dans un dispositif tel que celui expérimenté dans cette étude, les critères d'évaluation sont travaillés et retravaillés tout au long du semestre (apports théoriques dans le cours, puis lors de la production initiale, encore lors de la production des feedbacks, et finalement dans la réception et mobilisation des feedbacks pour réguler sa production). Ce travail itératif sur les critères a permis aux étudiant·es d'affiner leur compréhension des attendus, comme le montre notre évaluation de leurs TP2. De plus, en développant ces compétences, les étudiant·es apprennent à s'appuyer sur leur jugement évaluatif afin de devenir des étudiant·es (et praticien·nes) autonomes (Tai et al., 2017).

Le développement d'une meilleure compréhension de l'évaluation est également une des dimensions qui ressort de l'analyse du discours des étudiant·es dans l'étude de Hanrahan et Isaacs (2001), en lien avec le développement de l'empathie envers le corps enseignant évaluateur. Dans sa revue de la littérature, Topping (1998) décrit ces apprentissages en termes de « *bénéfices systémiques* », en expliquant que l'évaluation par les pairs peut donner aux étudiant·es une meilleure idée des processus d'évaluation institutionnels, par exemple en leur permettant de se rendre compte des difficultés inévitables et des biais évaluatifs, les rendant plus tolérants avec les évaluations qu'elles et ils recevront dans le futur. Ce constat, qui corrobore le fait que l'évaluation par les pairs permet de développer le jugement évaluatif des étudiant·es (Tai et al., 2017) nous amène vers les conceptions de l'évaluation comme moyen pour apprendre (*assessment as learning* ; e.g., Dann, 2014) et de l'évaluation « durable » (*sustainable assessment* ; Boud, 2000), en renforçant l'idée que faire pratiquer l'évaluation aux étudiant·es leur permet de développer des compétences évaluatives (ou jugement évaluatif) qui leur seront utiles dans leur vie estudiantine et professionnelle.

Nous postulons donc que l'interaction entre les différentes pièces du dispositif, convoquant diverses compétences, a facilité la régulation des apprentissages relatifs aux contenus du cours chez les étudiant·es de notre étude.

Un autre élément qui a son importance est le fait que toutes ces « pièces » du dispositif comptaient dans l'évaluation continue du cours. En effet, sachant que les feedbacks devaient être mobilisés par leurs pairs et que les feedbacks étaient évalués par les enseignantes, les étudiant·es ont pris la tâche de production de feedbacks très à cœur. De même, elles et ils ont analysés les feedbacks reçus en profondeur afin d'améliorer leur production dans le TP2, dont la révision était également évaluée par les enseignantes. En ce sens, nous prenons appui sur l'étude de Tsivitanidou et al. (2011), qui, dans une activité de feedbacks entre pairs non-guidée et libre, ont constaté que la majorité des étudiant·es (du secondaire) n'avaient pas révisé leurs productions suite aux feedbacks de leurs pairs. Ces auteur·rices émettent l'hypothèse d'un manque de confiance en les compétences évaluatives des pairs (perception partagée par les étudiant·es de notre étude, nous y reviendrons). Nous pensons aussi que cela est dû au caractère non obligatoire de leurs



tâches. Dans notre étude, le fait que la mobilisation des feedbacks des pairs ait été obligatoire et évaluée par les enseignantes a pu représenter un point primordial dans le « succès » du dispositif. Ceci pose la question de l'utilité d'une certaine pression externe (et donc un certain niveau de stress) engendrée par l'évaluation certificative pour engager les apprenant·es dans l'apprentissage... ce qui nous permet de faire le lien avec la section suivante.

5.2 Quid de la dimension numérique ?

Un des objectifs de cette étude était d'interroger l'articulation de l'évaluation au numérique. Cependant, bien que toutes les tâches et activités aient été effectuées à partir de et déposées sur une plateforme numérique, la dimension numérique n'est pas apparue dans les réflexions critiques des étudiant·es. Comme indiqué précédemment, ceci fait sens étant donné que les questions de la tâche réflexive portaient sur des activités précises (production et réception de feedbacks) qui ne nécessitaient en soi pas de particularités intéressantes au niveau numérique. L'apport principal de la plateforme numérique utilisée a été de permettre une centralisation des activités et ressources pour faciliter un va-et-vient entre les tâches et les contenus théoriques. Un des points que nous pouvons relever est le fait que le numérique a facilité une anonymisation complète des travaux des étudiant·es afin d'assurer une évaluation par les pairs « en double aveugle ». La dimension de l'anonymat de l'évaluation par les pairs a été abordée par certain·es étudiant·es de notre étude et a été étudiée dans la littérature.

L'évaluation par les pairs est une activité fondamentalement sociale et peut ainsi toucher la dimension relationnelle entre les étudiant·es. Rotsaert et al. (2017) rappellent, en s'appuyant sur d'autres auteur·rices (e.g., Chen & Tsai, 2002 ; Harris & Brown, 2013 ; van Gennip et al., 2010 ; Vanderhoven et al., 2015) que l'évaluation par les pairs sans anonymat peut avoir des effets émotionnels négatifs, comme une incertitude psychologique, une peur des récriminations en cas de feedback négatif, et un manque de confiance dans ses compétences évaluatives ainsi que celles des pairs. Des recherches ont alors montré que le fait de rendre les évaluations par les pairs anonymes peut aider à réduire ces impacts négatifs (Yu & Liu, 2009, cités par Rotsaert et al., 2017). Les perceptions des activités d'évaluation par les pairs seraient en effet plus positives en cas d'anonymat (Vanderhoven et al., 2015, cités par Rotsaert et al., 2017) et les étudiant·es semblent accorder une grande importance à ce dernier (Topping, 1998).

Cependant, l'anonymat ne reflète pas - et empêche - les dialogues entre des pairs qui connaissent leurs identités réciproques, qui sont des situations d'évaluation en face-à-face rencontrées dans la vie de tous les jours. Cette limite a été relevée par certain·es étudiant·es de cette étude, autant dans la production que dans la réception des feedbacks, accompagnée par une frustration de ne pas pouvoir lever des incompréhensions à cause de l'anonymat.

Selon Panadero (2016), l'anonymat dans des situations d'évaluation par les pairs peut représenter un obstacle à la création d'un environnement de feedback riche et interactif. De plus, sa revue de la littérature montre que plus il y a d'interactions entre l'évaluateur·rice et l'évalué·e dans des situation d'évaluation par les pairs, plus les résultats en termes de variables intra et interpersonnelles sont optimaux (Panadero, 2016). Un feedback dialogique permettrait alors de partager des informations, de négocier des significations, et de clarifier les attentes entre l'évaluateur·rice et l'évalué·e (Carless et al., 2011, cités dans Rotsart et al., 2017). Ces auteur·rices ont expérimenté le passage graduel entre des



feedbacks par les pairs anonymes et dialogiques, et concluent que, si l'enseignant·e dispose d'assez de temps, l'anonymat peut être utilisé comme un tremplin pour rassurer les étudiant·es alors qu'ils entraînent leurs compétences évaluatives. Ensuite, une évaluation par les pairs dialogique en face-à-face peut être instaurée et présenter des bénéfices tout en étant de bonne qualité. L'usage du numérique offre des possibilités de réalisation intéressante pour réaliser ce type de dispositif dans des cours à grands effectifs.¹⁹

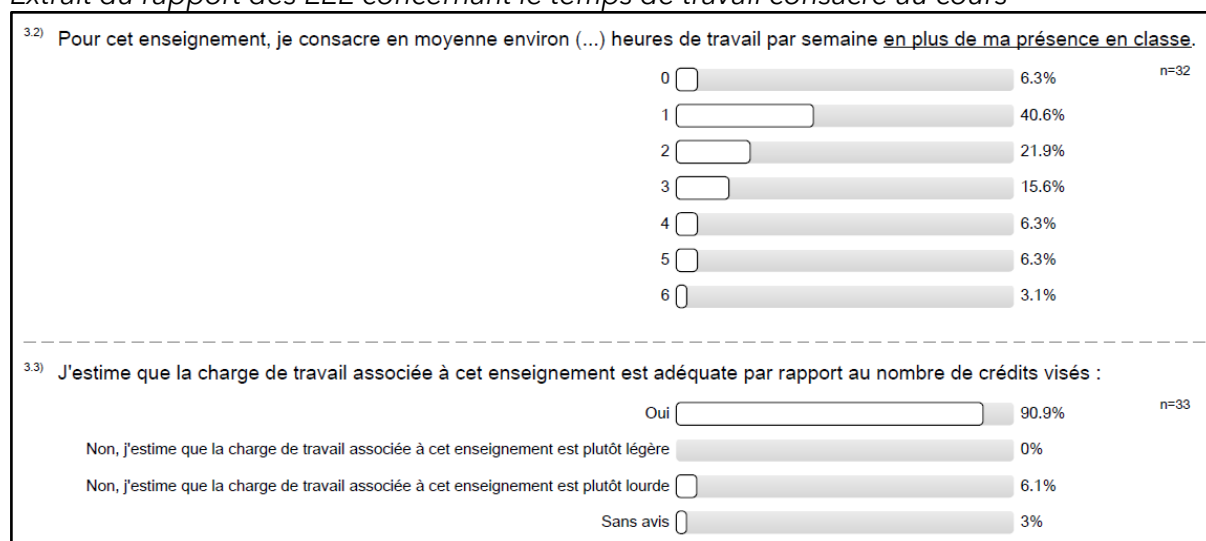
Notre étude aurait bénéficié de davantage de données permettant d'interroger le rapport au numérique des étudiant·es dans le dispositif d'évaluation continue proposé.

5.3 Considérations au sujet du dispositif

Une critique que l'on entend souvent à l'égard des activités d'évaluations entre pairs est qu'elles sont chronophages. Du côté des enseignantes, la mise en place d'un tel dispositif a pris un temps considérable (scénarisation, organisation, répartition des feedbacks, etc.). Du côté des étudiant·es, il est parfois difficile d'estimer si la charge de travail demandée correspond aux crédits ECTS alloués pour le cours. Afin de vérifier cela, des indicateurs sont disponibles dans le rapport institutionnel des Évaluations des Enseignements par les Étudiant·es (EEE). Voici ces indicateurs pour le cours concerné :

Figure 4

Extrait du rapport des EEE concernant le temps de travail consacré au cours



Ces indicateurs nous permettent de dire que la charge de travail semble avoir été estimée comme adéquate par les étudiant·es.

Une autre élément de considération lors de la mise en place d'un tel dispositif est la quantité d'éléments (principes, précautions) que les enseignant·es doivent prendre en

¹⁹ Le dispositif pour ce même cours a été réitéré l'année académique 2019-2020 avec plusieurs changements, dont la mise en place de modérations sociales au sein de forums en ligne regroupant quatre étudiant·es, deux auteur·rices et deux rétroacteur·rices. Dans cette étape, il y avait une levée de l'anonymat partielle. Tous les noms étaient connus. Les étudiant·es pouvaient également voir leurs rôles respectifs. Cependant, les quatre feedbacks portaient des numéros et les rétroacteur·rices pouvaient choisir d'indiquer quel feedback était le leur, ou choisir de ne pas le dire. Ainsi, chacun·e a pu se sentir libre de révéler ou non son travail. Des travaux de recherches seront envisagés sur ce dispositif également, mais nos résultats préliminaires laissent penser que le dispositif 2019-2020 a été perçu tout autant positivement que celui de 2018-2019, voire davantage.



compte afin d'œuvrer à sa réussite. Parmi ces éléments, on retrouve : initiation préalable aux concepts du cours ; formation des apprenant·es à la production de feedback ; consignes précises ; mise à disposition des critères d'évaluation ; articulation de l'évaluation formative à l'évaluation sommative assurant un engagement des étudiant·es dans les tâches. Dans des études ultérieures, il serait intéressant d'aborder comment le dispositif dans son ensemble (scénario, consignes, façon dont les contenus, consignes et articulations des tâches sont présentées aux étudiant·es, ou encore le style d'enseignement et les attitudes des enseignantes) contribue aux apprentissages des étudiant·es.²⁰

6. Brève conclusion

Cette étude avait pour but d'investiguer les perceptions d'étudiant·es universitaires d'un dispositif d'évaluation continue impliquant des tâches de production et de réception de feedbacks entre pairs. Un constat est la diversité des apprentissages réalisés au moyen de ces tâches, tant du point de vue de l'apprentissage des contenus étudiés en cours que du développement de compétences transversales, ce qui constitue un exemple d'une évaluation comme moyen pour apprendre (*assessment as learning*). Le fait de surpasser des difficultés semble avoir « boosté » les apprentissages de façon particulièrement forte. Nous pensons que tout cela n'aurait pas été possible sans l'articulation de l'évaluation formative à l'évaluation certificative. En effet, le fait de se savoir évalué·es semble avoir joué un rôle dans l'investissement des étudiant·es dans les tâches, qui, à leur tour, ont produit de nombreux apprentissages.

Cette étude a identifié deux principaux processus « moteurs » de régulation des apprentissages, associés à :

- l'articulation entre évaluations formative et certificative ;
- l'articulation entre production de feedbacks aux pairs sur une production initiale, et réception de feedbacks à trier puis à mobiliser dans la révision de la production initiale.

Nous allons dans le sens de plusieurs études (Boud, 2000 ; Boud & Molloy, 2013 ; Liu & Carless, 2006 ; Rowe, 2017) préconisant un changement de la vision du feedback, non plus comme un acte qui est produit ponctuellement, mais comme un processus intégré dans les dispositifs d'enseignement et d'évaluation de l'apprentissage, et dans lequel les étudiant·es sont fortement impliqués.

À titre d'ouverture, nous proposons d'explorer davantage le rôle des émotions et sentiments dans les processus d'apprentissage dans des activités de feedback entre pairs, que cette étude laisse entrevoir. Ceci fera l'objet d'un second article dans un prochain numéro.

Références

Allal, L. (2008). Évaluation des apprentissages. In A. van Zanten (Ed.), *Dictionnaire de l'éducation* (pp. 311-314). Presses Universitaires de France.

²⁰ Avec mes remerciements à Amaury Daele et Vincent Grenon qui m'ont suggéré ces pistes durant l'évaluation ouverte et collaborative de l'article.



- Allal, L., & Mottier Lopez, L. (2005). L'évaluation formative de l'apprentissage : revue de publications en langue française. In *L'évaluation formative : pour un meilleur apprentissage dans les classes secondaires* (pp. 265-290). OCDE-CERI Publication.
- Bielaczyc, K., & Collins, A. (1999). Learning communities in classrooms: Advancing knowledge for a lifetime. *National Association of Secondary School Principals Bulletin*, 83(604), 4-10. <https://doi.org/10.1177/019263659908360402>
- Black, P., & Wiliam, D. (1998). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan*, 80(2), 139-148. <https://doi.org/10.1177/003172171009200119>
- Blais, M., & Martineau, S. (2006). L'analyse inductive Générale : description d'une démarche visant à donner un sens à des données brutes. *Recherches Qualitatives*, 26(2), 1-18.
- Boud, D. (2000). Sustainable assessment: Rethinking assessment for the learning society. *Studies in Continuing Education*, 22(2), 151-167. <https://doi.org/10.1080/713695728>
- Boud, D., & Falchikov, N. (Eds.) (2007). *Rethinking assessment in higher education: Learning for the longer term*. Routledge.
- Boud, D., & Molloy, E. (Eds.) (2013). *Feedback in Higher and Professional Education: Understanding It and Doing It Well*. Taylor & Francis Group. <https://ebookcentral.proquest.com/lib/graduateinstitute/detail.action?docID=1101395>
- Boud, D., & Soler, R. (2015). Sustainable assessment revisited. *Assessment & Evaluation in Higher Education*, 41(3), 400-413. <https://doi.org/10.1080/02602938.2015.1018133>
- Carless, D., Salter, D., Yang, M., & Lam, J. (2011). Developing sustainable feedback practices. *Studies in Higher Education*, 36(4), 395-407. <https://doi.org/10.1080/03075071003642449>
- Cao, Z., Yu, S., & Huang, J. (2019). A qualitative inquiry into undergraduates' learning from giving and receiving peer feedback in L2 writing: Insights from a case study. *Studies in Educational Evaluation*, 63, 102-112. <https://doi.org/10.1016/j.stueduc.2019.08.001>
- Daele, A., Brassard, C., Esnault, L., O'Donogue, M., Uyttebrouck, E., & Zeiliger, R. (2003). WP2 - Conception, mise en œuvre, analyse et évaluation de scénarios pédagogiques recourant à l'usage des Technologies de l'Information et de la Communication. Rapport recre@sup. <http://tecfa.unige.ch/proj/recreasup/rapport/WP2.pdf>
- Dann, R. (2014). Assessment as learning: blurring the boundaries of assessment and learning for theory, policy and practice. *Assessment in Education: Principles, Policy & Practice*, 21(2), 149-166. <https://doi.org/10.1080/0969594X.2014.898128>
- Donia, M. B. L., O'Neill, T. A., & Brutus, S. (2018). The longitudinal effects of peer feedback in the development and transfer of student teamwork skills. *Learning and Individual Differences*, 61, 87-98. <https://doi.org/10.1016/j.lindif.2017.11.012>
- Falchikov, N. (Ed.) (2005). *Improving Assessment Through Student Involvement: Practical Solutions for Aiding Learning in Higher and Further Education*. Taylor & Francis Group. <http://ebookcentral.proquest.com/lib/graduateinstitute/>
- Gielen, M., & De Wever, B. (2012). Peer assessment in a wiki: Product improvement, students' learning and perception regarding feedback. *Procedia - Social and Behavioral Sciences*, 69, 585-594. <https://doi.org/10.1016/j.sbspro.2012.11.450>
- Girard, É., Terradas, M., & Matte-Gagné, C. (2014). Empathie, comportements pro-sociaux et troubles du comportement. *Enfance*, 4(4), 459-480. <https://doi.org/10.4074/S0013754514004030>
- Hanrahan, S. J., & Isaacs, G. (2001). Assessing self- and peer-assessment: The students' views. *Higher Education Research & Development*, 20(1), 53-70. <https://doi.org/10.1080/07294360123776>
- Harris, L. R., & Brown, G. T. L. (2013). Opportunities and obstacles to consider when using peer- and self-assessment to improve student learning: case studies into teachers' implementation. *Teaching and Teacher Education*, 36, 101-111. <https://doi.org/10.1016/j.tate.2013.07.008>



- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81-112. <https://doi.org/10.3102/003465430298487>
- Laferrière, T. (2005). Les communautés d'apprenants en réseau au bénéfice de l'éducation. *Rencontres en Education*, 6, 5-21.
- Lewis, A. (2001). The issue of perception: some educational implications. *Educare*, 30(1), 272 - 288. http://www.andrewlewis.co.za/Lewis.Perception.Educare1_v30_n1_a15.pdf
- Lindschouw, J. (2016). Le feedback entre pairs, points de focalisation et attitudes lors des tâches d'écriture produites par les apprenants de FLE. *Pratiques : Linguistique, littérature, didactique*, 169-170. <https://doi.org/10.4000/pratiques.3030>
- Liu, N.-F., & Carless, D. (2006). Peer feedback: the learning element of peer assessment. *Teaching in Higher Education*, 11(3), 279-290. <https://doi.org/10.1080/13562510600680582>
- Mailles-Viard Metz, S. (2015). L'aide du numérique aux activités d'auto-évaluation. *Revue internationale de pédagogie de l'enseignement supérieur [en ligne]*, 31(3). <http://journals.openedition.org/ripes/991>
- Mottier Lopez, L. (2008). *Apprentissage situé. La microculture de classe en mathématiques*. Berne: Peter Lang. <https://archive-ouverte.unige.ch/unige:17430/ATTACHMENT01>
- Mottier Lopez, L. (2015a). *Évaluations formative et certificative des apprentissages : Enjeux pour l'enseignement*. De Boeck. <https://archive-ouverte.unige.ch/unige:106923/ATTACHMENT01>
- Mottier Lopez, L. (2015b). Tâches pour apprendre, tâches pour évaluer. *Formation et pratiques d'enseignement en questions*, 19, 55-65. <http://revuedeshep.ch/no-19/>
- Mottier Lopez, L. (2018, 31 octobre). *Le feedback formatif* [Support de cours non publié].
- Mucchielli, A. (1996). *Dictionnaire des méthodes qualitatives en sciences humaines et sociales*. Armand Colin.
- 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 (Eds.), *Handbook of Social and Human Conditions in Assessment* (pp. 247-266). Routledge. https://www.researchgate.net/profile/Ernesto_Panadero/publication/
- Riley, S. M. (1995). *Peer response in an ESL writing class: Student interaction and subsequent draft revision* [Thèse de doctorat], Florida State University.
- Rotsaert, T., Panadero, E., & Schellens, T. (2017). Anonymity as an instructional scaffold in peer assessment: its effects on peer feedback quality and evolution in students' perceptions about peer assessment skills. *European Journal of Psychology of Education*, 33(1), 75-99. <https://doi.org/10.1007/s10212-017-0339-8>
- Rowe, A. D. (2017). Feelings about feedback: The role of emotions in assessment for learning. In D. Carless, S. M. Bridges, C. Ka Yuk Chan & R. Glofcheski (Eds.). *Scaling up assessment for learning in higher education* (pp. 159-172). Springer. <https://doi.org/10.1007/978-981-10-3045-1>
- Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional Science*, 18, 119-144. <https://doi.org/10.1007/BF00117714>
- Sadler, D. R. (1998). Formative assessment: Revisiting the territory. *Assessment in Education: Principles, Policy & Practice*, 5(1), 77-84. <https://doi.org/10.1080/0969595980050104>
- Sadler, D. R. (2010). Beyond feedback: developing student capability in complex appraisal. *Assessment & Evaluation in Higher Education*, 35, 535-550. <https://doi.org/10.1080/02602930903541015>
- Sandelowski, M. (2001). Real qualitative researchers do not count: The use of numbers in qualitative research. *Research in Nursing & Health*, 24(3), 230-240. <https://doi.org/10.1002/nur.1025>



- Shen, B., Bai, B., & Xue, W. (2020). The effects of peer assessment on learner autonomy: An empirical study in a Chinese college English writing class. *Studies in Educational Evaluation*, 64, 1-10. <https://doi.org/10.1016/j.stueduc.2019.100821>
- Tai, J., Ajjawi, R., Boud, D., Dawson, P., & Panadero, E. (2017). Developing evaluative judgement: Enabling students to make decisions about the quality of work. *Higher Education*, 76(3), 467-481. <https://doi.org/10.1007/s10734-017-0220-3>
- Topping, K. (1998). Peer assessment between students in colleges and universities. *Review of Educational Research*, 68(3), 249-276. <https://doi.org/10.3102/00346543068003249>
- Topping, K. J. (2003). Self and peer assessment in school and university: Reliability, validity and utility. In M. Segers, F. Dochy, & E. Cascallar (Eds.), *Optimising new modes of assessment: In search of qualities and standards* (Vol.1, pp. 55-87). Springer.
- Topping, K. (2009). Peer Assessment. *Theory Into Practice*, 48(1), 20-27. <https://doi.org/10.1080/00405840802577569>
- Tsivitanidou, O., Zacharia, Z. C., & Hovardas, T. (2011). Investigating secondary school students' unmediated peer assessment skills. *Learning and Instruction*, 21, 506-519. <https://doi.org/10.1016/j.learninstruc.2010.08.00>
- Van Gennip, N. A. E., Segers, M. S. R., & Tillema, H. H. (2010). Peer assessment as a collaborative learning activity: the role of interpersonal variables and conceptions. *Learning and Instruction*, 20(4), 280-290. <https://doi.org/10.1016/j.learninstruc.2009.08.010>
- Vanderhoven, E., Raes, A., Montrieux, H., Rotsaert, T., & Schellens, T. (2015). What if pupils can assess their peers anonymously? A quasi-experimental study. *Computers & Education*, 81, 123-132. <https://doi.org/10.1016/j.compedu.2014.10.001>
- Wied, M., Goudena, P. P., & Matthys, W. (2005). Empathy in boys with disruptive behavior disorders. *Journal of Child Psychology and Psychiatry*, 46(8), 867-880. <https://doi.org/10.1111/j.1469-7610.2004.00389.x>
- Wiggins, G. (1990). The Case for Authentic Assessment. *Practical Assessment, Research & Evaluation*, 2(2), 1-3. <https://www.pareonline.net/getvn.asp?v=2&n=2>
- Yu, S. (2019). Learning from giving peer feedback on postgraduate theses: Voices from Master's students in the Macau EFL context. *Assessing Writing*, 40, 42-52. <https://doi.org/10.1016/j.asw.2019.03.004>
- Yu, S., & Lee, I. (2016). Peer feedback in second language writing (2005-2014). *Language Teaching*, 49(4), 461-493. <https://doi.org/10.1017/S0261444816000161>
- Yu, F.-Y., & Liu, Y.-H. (2009). Creating a psychologically safe online space for a student-generated questions learning activity via different identity revelation modes. *British Journal of Educational Technology*, 40(6), 1109-1123. <https://doi.org/10.1111/j.1467-8535.2008.00905.x>

Remerciements

Un merci tout particulier à Lucie Mottier Lopez, principale conceptrice du dispositif et enseignante du cours, sans qui cette étude n'aurait pas pu avoir lieu. Merci à elle également pour les précieuses conversations que nous avons eues à l'occasion de la conception et réalisation de la présente recherche ainsi que de l'écriture de ce texte. Je souhaite également remercier les rétroacteurs et rétroactrices qui m'ont donné des feedbacks sur la première version de ce texte lors du symposium REF 2019. Je tiens à remercier également les doctorant·es de l'équipe EReD pour leurs retours sur ce texte à l'occasion d'un exercice sur l'évaluation scientifique dans le cadre d'un séminaire doctoral que j'ai co-organisé avec la professeure Mottier Lopez. Finalement, je remercie chaleureusement Amaury Daele et Vincent Grenon pour leurs rétroactions durant l'évaluation ouverte et collaborative du présent article au sein de La Revue LEeE.



Céline Girardet

Maîtresse assistante dans l'équipe EReD (évaluation, régulation et différenciation des apprentissages dans les systèmes d'enseignement) à l'Université de Genève, elle est détentrice d'une thèse de doctorat en sciences de l'éducation portant sur l'évolution des croyances relatives à la gestion de classe des enseignant·es en formation professionnelle. Elle est également chargée d'enseignement en suppléance pour un cours de maîtrise portant sur la recherche collaborative en éducation, et elle est co-éditrice de La Revue LEeE. Ses intérêts actuels portent sur les apprentissages réalisés au travers de dispositifs d'évaluation entre pairs dans des contextes d'enseignement supérieur et d'édition scientifique.

