

L'évaluation continue pour apprendre : enjeux de la pluralité des feedbacks entre pairs dans un cours universitaire

Lucie Mottier Lopez, Céline Girardet and Taha Naji

Volume 44, Number 2, 2021

URI: <https://id.erudit.org/iderudit/1090461ar>

DOI: <https://doi.org/10.7202/1090461ar>

[See table of contents](#)

Publisher(s)

ADMEE-Canada

ISSN

0823-3993 (print)

2368-2000 (digital)

[Explore this journal](#)

Cite this article

Mottier Lopez, L., Girardet, C. & Naji, T. (2021). L'évaluation continue pour apprendre : enjeux de la pluralité des feedbacks entre pairs dans un cours universitaire. *Mesure et évaluation en éducation*, 44(2), 1–34.
<https://doi.org/10.7202/1090461ar>

Article abstract

In the context of continuous assessment for learning in a university course in French-speaking Switzerland, this article studies students' perception when they receive multiple feedbacks from their peers about a written academic assignment carried out in small collaborative groups. How do students, who are both assessees and assessors, perceive these multiple feedbacks, especially when they are led to compare them to regulate their own initial assignment and their assessment skills? This article analyses in detail the processes and feelings at play when students notice similarities and differences between the feedbacks they received and produced. In light of the results, conceptual considerations are proposed around the notions of feedback, internal feedback, comprehensive feedback and metafeedback, seen as contributing to the regulation and learning processes involved.

L'évaluation continue pour apprendre : enjeux de la pluralité des feedbacks entre pairs dans un cours universitaire

Lucie Mottier Lopez

Céline Girardet

Taha Naji

Université de Genève

MOTS CLÉS : évaluation pour apprendre, évaluation continue, feedbacks écrits entre pairs, perception, régulation des apprentissages, enseignement supérieur

Dans le contexte d'une évaluation continue pour apprendre, expérimentée dans un cours universitaire en Suisse romande, l'article étudie la perception des étudiants quand ils reçoivent une pluralité de feedbacks des pairs à propos d'un travail réalisé en petits groupes collaboratifs. Comment les étudiants, qui sont à la fois évalués et évaluateurs, perçoivent-ils ces feedbacks multiples, notamment quand ils sont amenés à les comparer dans la perspective de réguler leur propre travail universitaire et leurs compétences évaluatives ? L'article analyse finement les processus et les ressentis en jeu quand les étudiants constatent des similarités et des différences entre les feedbacks reçus et produits. Au regard des résultats obtenus, une réflexion conceptuelle est proposée autour des notions de feedback, de feedback interne, de feedback global et de métafeedback, qui contribuent aux processus de régulation et d'apprentissage générés.

KEY WORDS: assessment for learning, continuous assessment, written peer feedback, perception, regulation of learning, higher education

In the context of continuous assessment for learning in a university course in French-speaking Switzerland, this article studies students' perception when they receive multiple feedbacks from their peers about a written academic assignment carried out in small collaborative groups. How do students, who are both assessees and assessors, perceive these multiple feedbacks, especially when they are led to compare them to regulate their own initial assignment and their assessment skills? This article analyses in detail the processes and feelings at play when students notice similarities and differences between the feedbacks they received and produced. In light of the results, conceptual considerations are proposed around the notions of feedback, internal feedback, comprehensive feedback and metafeedback, seen as contributing to the regulation and learning processes involved.

PALAVRAS-CHAVE: avaliação para a aprendizagem, avaliação contínua, feedback escrito entre pares, percepção, regulação das aprendizagens, ensino superior

No contexto de uma avaliação contínua para a aprendizagem, experimentada num curso universitário na Suíça francófona, o artigo estuda a percepção dos estudantes quando recebem uma pluralidade de feedbacks de colegas sobre o trabalho realizado em pequenos grupos colaborativos. Como é que os estudantes, que são ao mesmo tempo avaliados e avaliadores, percebem os feedbacks múltiplos, principalmente quando são levados a compará-los no âmbito da regulação do seu próprio trabalho académico e das suas competências avaliativas? O artigo analisa detalhadamente os processos e sentimentos em jogo quando os estudantes constatarem as semelhanças e as diferenças entre os feedbacks recebidos e produzidos. Perante os resultados obtidos, propõe-se uma reflexão conceitual em torno das noções de feedback, feedback interno, feedback global e metafeedback, que contribuem para os processos de regulação e aprendizagem desencadeados.

Introduction

La recherche présentée dans cet article s'intéresse à un dispositif d'évaluation continue dans un cours de première année en sciences de l'éducation dans une université de Suisse romande. Ce dispositif est fondé sur la production et la réception de multiples feedbacks écrits formatifs et anonymes entre étudiants à propos d'un travail initial réalisé en petits groupes collaboratifs.

Fondés sur des critères explicitement travaillés pendant le cours, ces feedbacks qualitatifs ne conduisent à formuler aucun score ni aucune note au travail évalué. Leur fonction, annoncée aux étudiants, est de contribuer à l'amélioration de leur travail initial dans un rapport de réciprocité, c'est-à-dire de produire des feedbacks formatifs et bienveillants aux pairs et, réciproquement, en recevoir sur leur propre travail universitaire à des fins de régulation.

La conception de cette évaluation est celle d'une «évaluation pour apprendre» (Black & Wiliam, 1998) articulant des démarches d'évaluation formative et certificative. Cette évaluation met l'accent sur les relations contingentes et dialogiques entre enseignement et apprentissage; elle vise à impliquer les apprenants tant dans la production des évaluations que dans la façon de les recevoir à des fins d'autoévaluation, de jugement critique et d'autorégulation (Allal & Laveault, 2009; Mottier Lopez, 2016, 2020; Nicol, 2020). De nombreux travaux existent sur les feedbacks entre pairs dans l'enseignement supérieur (p. ex., Carless & Boud, 2018; Falchikov & Boud, 1989; McConlogue, 2015; Nicol et al., 2014). Dans un environnement d'apprentissage collaboratif, la production et la réception de feedbacks formatifs entre apprenants sont susceptibles de représenter un dispositif puissant pour soutenir la régulation de leurs apprentissages (p. ex., Nicol, 2020; Strijbos et al., 2009).

Dans notre recherche, l'étudiant est à la fois dans la posture de l'évalué (c.-à-d. qu'il reçoit des feedbacks écrits individuels de pairs anonymes sur un travail initial) et de l'évaluateur (c.-à-d. qu'il rédige un feedback

anonyme à un pair, puis prend connaissance d'autres feedbacks produits sur le travail universitaire qu'il a évalué). Des critères prédéfinis lui offrent un cadre de référence pour qu'il rédige des feedbacks formatifs et bienveillants. L'étudiant reçoit alors plusieurs feedbacks individuels de pairs qui ne se sont pas coordonnés préalablement entre eux. Les feedbacks écrits peuvent être différents les uns des autres ; ils peuvent ne pas être cohérents entre eux ou, au contraire, apporter des regards complémentaires sur le travail concerné.

Les travaux de Nicol (2013, 2020) montrent que le fait de disposer de plusieurs feedbacks de pairs incite l'étudiant qui les reçoit à les comparer et à se positionner pour décider des actions de régulation à entreprendre afin d'améliorer son travail et, plus généralement, de développer de nouveaux apprentissages, y compris métacognitifs. Ces travaux montrent que les comparaisons appropriées entre feedbacks représentent de réelles ressources pour la régulation des performances des étudiants ainsi que pour le développement de leurs capacités d'autorégulation, sans nécessiter obligatoirement un étayage de l'enseignant (Nicol, 2020).

La recherche présentée dans cet article vise ainsi à répondre à la question générale de recherche suivante : *Dans le cadre d'un dispositif d'évaluation pour apprendre, comment les étudiants, qui sont à la fois évalués et évaluateurs, perçoivent-ils des sources multiples de feedbacks des pairs, notamment quand ils sont amenés à les comparer dans la perspective de réguler leur propre travail universitaire ?*

Notre hypothèse de travail est que les différents feedbacks écrits, s'ils sont comparés et confrontés entre eux par l'étudiant, engendrent des perceptions différentes selon le rôle tenu (évalué ou évaluateur) à propos des apprentissages associés au travail universitaire et à propos des compétences évaluatives sollicitées pour produire et recevoir des feedbacks formatifs et bienveillants.

Après cette première partie introductive, la deuxième partie de l'article expose un état de la question au sujet, premièrement, de l'évaluation pour apprendre, plus particulièrement quand celle-ci ambitionne une évaluation continue visant à créer des liens entre différentes tâches d'apprentissage et d'évaluation sur une temporalité longue. L'état de la question porte ensuite sur les pratiques de feedback entre pairs, notamment quand une pluralité de feedbacks formatifs est proposée à des fins de comparaison et de confrontation des jugements évaluatifs des pairs sur leur propre travail

universitaire. Enfin, deux configurations de comparaison de feedbacks sont définies et circonscrivent le cadre conceptuel de notre recherche. La troisième partie de l'article présente le contexte et les grandes lignes de l'évaluation continue pour apprendre (ECPA), qui intègre l'évaluation entre pairs qui nous intéresse. La quatrième partie expose le cadre méthodologique et les analyses entreprises. Les résultats de la recherche sont présentés dans la cinquième partie. Une discussion conclusive reprend les principaux constats et énonce prudemment quelques avancées conceptuelles au regard des résultats obtenus.

État de la question et cadre conceptuel

Une évaluation continue pour apprendre durablement

L'évaluation dont la fonction est d'aider les élèves à apprendre est traditionnellement associée à l'évaluation formative depuis les propositions de Bloom et al. (1971) énoncées dans la pédagogie de maîtrise. Depuis la fin des années 1990, d'autres termes ont été proposés, notamment dans le contexte anglo-saxon, afin d'insister sur le rôle de soutien à l'apprentissage que peut jouer l'évaluation quand elle est fortement articulée aux processus d'enseignement et de régulation des apprentissages des élèves. Le Tableau 1 présente plusieurs de ces propositions¹.

Le but commun visé par ces propositions conceptuelles est de *reconfigurer* l'évaluation, tant formative que certificative, afin qu'elle soit conçue et utilisée comme un levier d'action positif pour soutenir des apprentissages de qualité chez les élèves. À quelques nuances près, les auteurs s'accordent sur un ensemble de caractéristiques considérées comme cruciales pour atteindre ce but ambitieux, lesquelles ont orienté l'ECPA expérimentée dans notre recherche :

- Concevoir des tâches *d'apprentissage* qui intègrent des évaluations associées à des possibilités de régulation et d'autorégulation des étudiants ;
- Concevoir des tâches *évaluatives* complexes qui encouragent des collaborations entre étudiants, par exemple par le moyen de travaux de groupe et/ou par la réalisation de projets ;

1. Les dimensions énoncées dans le Tableau 1 ont été extraites des références citées sans viser l'exhaustivité. Dans des textes ultérieurs aux sources citées dans le tableau, d'autres prolongements ont été proposés. Nous ne les repreneons pas ici.

Tableau 1

Propositions conceptuelles visant à reconfigurer l'évaluation des apprentissages

Terme en anglais	Auteurs à l'origine	Traduction en français	Dimensions principales mises en évidence
<i>Assessment for learning</i>	<i>Assessment Reform Group</i> (Black & Wiliam, 1998)	Évaluation pour apprendre Évaluation soutien d'apprentissage (Allal & Laveault, 2009)	Englobe toutes formes d'évaluations (formelles et informelles, formatives et sommatives) susceptibles d'encourager les élèves à s'engager dans des apprentissages à des fins de régulation et d'autorégulation ; fortement intégrées aux activités d'enseignement et d'apprentissage
<i>Authentic assessment</i>	Wiggins (1998)	Évaluation authentique	Les contenus, la structure, les caractéristiques et la validité des évaluations s'associent à des situations complexes du monde « réel », sans distinction entre évaluations formatives et certificatives
<i>Sustainable assessment</i>	Boud (2000)	Évaluation durable	Les évaluations, quelles que soient leurs formes et leurs fonctions, répondent aux besoins d'apprentissage dans des temporalités courtes et longues (apprentissage tout au long de la vie)
<i>Assessment as learning</i>	Earl (2003)	Évaluation comme apprentissage	Les évaluations sont vues comme susceptibles de représenter des occasions d'apprentissage, particulièrement quand les élèves sont impliqués dans des autoévaluations, des évaluations entre pairs et des évaluations collaboratives formatives. L'évaluation devient objet d'apprentissage

- Clarifier avec les étudiants les contrats évaluatifs en jeu, c'est-à-dire les attentes, les objectifs et les critères d'évaluation, y compris pour renforcer la confiance mutuelle entre apprenants et avec les enseignants au regard des différents enjeux de l'évaluation ;
- Impliquer les étudiants dans des démarches d'autoévaluation et d'évaluation entre pairs qui, en plus de se mettre au service de leurs

- apprentissages universitaires, leur permettent de développer des compétences évaluatives, de jugement critique et d'autorégulation ;
- Accorder de l'importance à la voix des étudiants pour les considérer comme des partenaires à part entière dans l'évaluation ;
 - Offrir des opportunités nombreuses de rétroaction (feedback) avec les enseignants et entre pairs dans la perspective de soutenir les apprentissages dans les tâches concernées, mais également pour leurs apprentissages futurs.

Ailleurs (Mottier Lopez, 2021 ; Mottier Lopez & Girardet, 2022), nous avons conceptualisé cette ECPA sur un temps long, au fil de différentes tâches qui intègrent des évaluations formatives, formelles et informelles, fondées sur une diversité de modalités : écrites ou orales, individuelles ou en groupes, en présentiel ou à distance, à l'aide d'outils numériques.

L'ECPA se distingue du contrôle continu traditionnel consistant à effectuer une succession de bilans microsomatifs plus ou moins détachés les uns des autres, dont l'enseignant fait une somme (ou une « moyenne ») en fin de période de formation. Elle vise plutôt le développement de compétences universitaires et transversales dépassant le strict cadre du cours concerné en faveur d'apprentissages tout au long de la vie (Boud & Soler, 2015). Entre autres, elle permet de soutenir l'apprentissage de compétences évaluatives et autoévaluatives par le moyen de feedbacks entre pairs dans le but, entre autres, de contribuer au développement du « pouvoir d'agir » et de l'autonomie des étudiants (Bain, 2010).

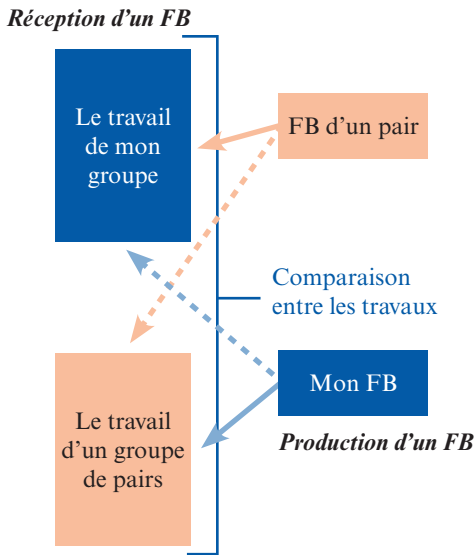
Une pluralité de feedbacks entre pairs à des fins de comparaison et de régulation

Dans les écrits spécialisés, la fonction du *feedback* est considérée comme cruciale pour produire un étayage ajusté aux caractéristiques et aux besoins de l'apprenant ainsi que pour soutenir sa progression d'apprentissage (p. ex., Black & Wiliam, 1998 ; Hattie & Timperley, 2007). Boud et Molloy (2013) définissent le feedback comme un « *mechanism through which students discover whether they are successful in their work and if they are on track to meet expectations* » (p. 1). Le feedback est conceptualisé comme un processus actif pour l'élève, dépassant le « simple » contenu informationnel transmis oralement ou par écrit (Boud & Molloy, 2013 ; Hattie & Timperley, 2007). Nicol (2020) parle de « feedback interne »

(*internal feedback*) pour désigner l'appropriation active faite par l'élève des informations produites par l'environnement (social, matériel) l'amenant à produire de nouvelles connaissances.

Dans un contexte universitaire, les feedbacks entre pairs représentent une modalité visant à engager activement les étudiants dans les processus évaluatifs et de régulation formative (p. ex., Nicol et al., 2014; Panadero et al., 2016). Ils contribuent non seulement au développement des compétences cognitives en jeu dans les tâches universitaires, mais également des compétences (auto)évaluatives et métacognitives (Girardet, 2020). Plus précisément, des travaux de recherche montrent que le fait de demander à l'étudiant non seulement de *recevoir* mais également de *produire* un feedback sur le travail des pairs enclenche une *comparaison* entre le travail des pairs et son propre travail (voir Figure 1), source de nouvelles connaissances (Nicol, 2013, 2020).

Figure 1
*Réception et production de feedbacks entre pairs :
comparaison entre les productions académiques*



Plusieurs études constatent que la modalité de produire un feedback aux pairs tend à être plus efficace que celle qui consiste à recevoir des feedbacks (McConlogue, 2015 ; Nicol et al., 2014). D'autres relèvent la complémentarité entre les deux modalités (Cao et al., 2019 ; Girardet, 2020). Nicol (2020) observe que les élèves tendent à ne pas réguler de la même façon : évaluer le travail d'un pair les conduit à apporter de nouvelles perspectives à leur propre travail ; recevoir un feedback entraîne plutôt une correction d'éléments manquants.

Ce même auteur souligne l'intérêt de multiplier les sources de feedbacks entre pairs, à la fois en tant que récepteurs et producteurs de feedbacks, dans une conception d'évaluation authentique par rapport aux situations du monde professionnel :

In professional and workplace settings and in life beyond university, feedback rarely comes from a single source. Rather, faced with multiple sources of feedback, the task is usually to evaluate, weigh up, reconcile and respond to different and sometimes quite contradictory feedback perspectives. Also, in professional life, graduates are not just consumers of feedback but they are also producers; they will be invariably required to evaluate and comment on the work of others from a range of perspectives (Nicol, 2013, s. p.).

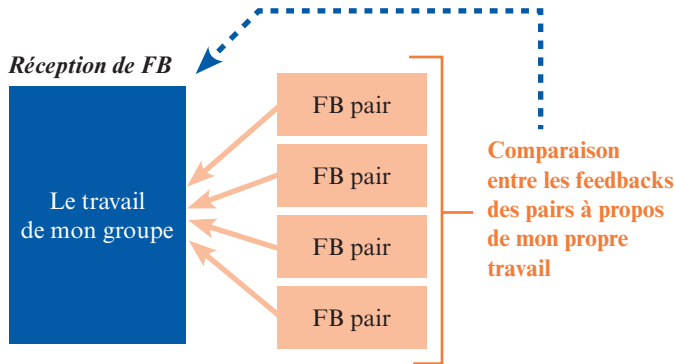
L'étude expérimentale de Cho et MacArthur (2010) va également dans ce sens ; elle montre que recevoir de multiples feedbacks de pairs amènerait à une meilleure régulation du travail initial que de recevoir un seul feedback de pair ou d'une personne experte.

L'objectivation de deux configurations de comparaison de feedbacks entre pairs dans une ECPA

Outre la comparaison entre productions académiques (voir Figure 1), d'autres formes de comparaison peuvent être intéressantes à explorer, notamment celle consistant à comparer des feedbacks de pairs reçus et produits. Dans cette perspective, nous définissons ci-dessous deux configurations de comparaison de feedbacks écrits entre pairs à des fins d'évaluation continue pour apprendre durablement.

Dans la configuration exposée dans la Figure 2, la comparaison intervient au niveau de la réception de plusieurs feedbacks de pairs sur son propre travail universitaire.

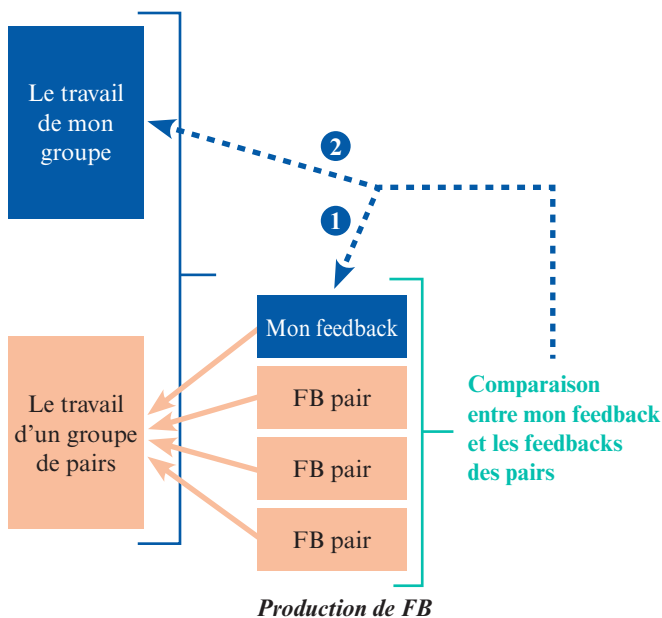
Figure 2
*Réception de multiples sources de feedbacks de pairs :
 comparaison entre les feedbacks*



Recevoir différents feedbacks sur son travail devrait permettre à l'étudiant de les comparer ainsi que de prendre conscience de la diversité possible des jugements évaluatifs, des similarités et des différences possiblement contradictoires entre feedbacks, y compris face à différentes options de régulation du travail sur lequel portent les feedbacks des pairs. Les recherches citées par Nicol et al. (2014) montrent que cette diversité de feedbacks représente un enrichissement potentiellement favorable à une régulation plus en profondeur de son propre travail. Cette diversité permet une décentration, ce qui engage une prise de recul soutenant l'autoévaluation et la régulation du travail (Girardet, 2020 ; Nicol, 2020). Ces mêmes auteurs observent que la diversité peut également représenter une source de déstabilisation. La validité et la fiabilité des jugements évaluatifs des pairs peuvent être mises en doute, ce qui pourrait refléter par exemple un sentiment d'insécurité lié à la posture d'un évaluateur vu comme un non-expert (Girardet, 2020).

Une configuration complémentaire consiste à faire intervenir la comparaison au niveau de la production de feedbacks sur le travail de pairs, notamment entre son propre feedback et le feedback des autres étudiants portant sur le même travail universitaire évalué. Cette configuration est présentée dans la Figure 3. Elle apparaît moins étudiée dans la littérature.

Figure 3
*Production de multiples sources de feedbacks de pairs :
 comparaison entre son propre feedback et celui des autres*



Cette configuration incite les étudiants qui ont produit un feedback à prendre connaissance des autres feedbacks des pairs sur le travail évalué, toujours à des fins de comparaison des jugements évaluatifs et des pistes de régulation suggérées dans les feedbacks. D'une part, nous postulons que cette comparaison est susceptible de produire un retour critique sur son propre feedback (voir ❶ dans la Figure 3) et, plus généralement, de donner des informations sur les compétences évaluatives impliquées dans la production de feedbacks aux pairs, y compris utiles pour une autoévaluation de sa propre production universitaire (voir Figure 1). D'autre part, elle peut engendrer une rétroaction sur son propre travail universitaire à des fins de régulation potentielle (voir ❷ dans la Figure 3).

Tout comme dans la configuration précédente (voir Figure 2), si des effets positifs sont attendus des comparaisons entre feedbacks, des effets possiblement contreproductifs pourraient également émerger. Les travaux en psychologie sociale, par exemple, soulignent les tensions qui existent entre le support informationnel constructif que la comparaison à autrui

peut représenter et le sentiment de menace sur les compétences si un climat de compétition domine entre les individus impliqués (p. ex., Butera et al., 2011).

D'autres recherches soulignent également les limites des feedbacks strictement écrits. Elles insistent sur l'importance de les inscrire dans des contextes de dialogue et de coconstruction de sens (p. ex., Schillings et al., 2021 ; Yang & Carless, 2013), y compris pour négocier les attentes, les droits et les devoirs associés à leurs usages dans une culture de classe visant à soutenir la régulation des apprentissages (Mottier Lopez, 2016).

Notre recherche explore, dans une approche qualitative, ces deux configurations de comparaison de feedbacks entre pairs. Le but scientifique est de mieux comprendre comment les étudiants perçoivent et exploitent des évaluations entre pairs qui s'appuient sur des sources multiples de feedbacks dans une réciprocité de rôles (évalué/récepteur de feedbacks et évaluateur/producteur de feedback), ce qui amène à des comparaisons susceptibles de contribuer à la régulation de leurs travaux universitaires et de soutenir le développement de leurs compétences évaluatives.

Contexte et méthodologie

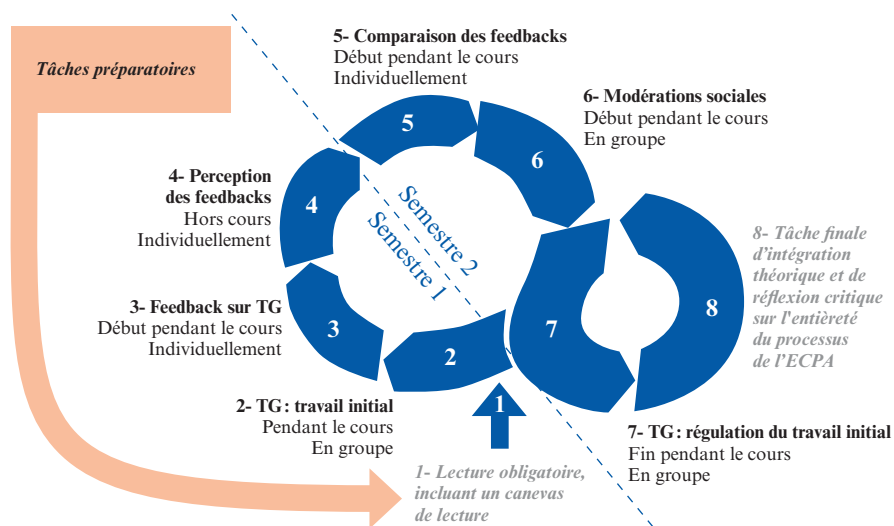
Présentation de l'ECPA

L'évaluation continue pour apprendre (ECPA) a été expérimentée durant l'année universitaire 2020-2021 dans un cours en sciences de l'éducation de l'université d'attache, en première année de bachelor². Cette ECPA est fondée sur une succession de tâches articulées aux contenus d'enseignement et réparties sur une année. Les tâches interreliées entre elles ont été conçues à la fois pour soutenir la régulation des apprentissages des étudiants et pour produire des données empiriques pour la recherche (Mottier Lopez & Girardet, 2022). La Figure 4 schématise ces tâches.

Les différentes tâches sont décrites ci-dessous. L'accent est mis sur les tâches 2, 3 et 4, qui sont au cœur du questionnement présenté dans cet article.

2. Cette évaluation s'inscrit dans une modalité de «contrôle continu» qui débouche sur une validation certificative du cours.

Figure 4
Schématisation de l'ECPA



Note. TG = travail de groupe.

Les tâches préparatoires A (quiz sur les fonctions de l'évaluation) et B (questionnaire de perceptions sur deux feedbacks écrits afin d'engager un questionnement sur ce qu'est un feedback formatif et bienveillant) ainsi que la tâche 1 (lecture obligatoire sur l'autoévaluation au sens large avec un canevas de lecture) ont été réalisées avant le travail universitaire initial faisant l'objet des feedbacks anonymes entre pairs.

Tâche 2: Travail de groupe

Ce travail initial a été réalisé en 32 groupes de 4-5 étudiants dans des salles de visioconférence Zoom sur un fichier Google Document. La consigne était d'imaginer des dispositifs d'évaluation pour favoriser l'apprentissage et l'autonomie d'un étudiant fictif, Victor, qui avait éprouvé des difficultés à réaliser la tâche 1 (lecture obligatoire incluant un canevas de lecture). Les étudiants avaient à disposition le canevas de lecture fictif de Victor, les consignes ainsi que des critères d'évaluation de leur travail de groupe.

Tâche 3: Feedback sur le travail de groupe

À la suite d'enseignements ciblés sur l'évaluation et sur les feedbacks formatifs et bienveillants³, chaque étudiant a dû individuellement rédiger deux feedbacks sur deux travaux de groupe des pairs en se basant sur les critères d'évaluation de la tâche 2. Chaque feedback devait compter un minimum de 1000 mots.

Tâche 4: Perceptions des feedbacks

Une fois tous les feedbacks réalisés, chaque étudiant a notamment reçu quatre feedbacks écrits par des pairs sur son travail de groupe initial et trois feedbacks de pairs sur le travail d'un autre groupe pour lequel lui-même a rédigé un feedback. Pour les sept feedbacks reçus, chaque étudiant a répondu – individuellement – à deux rubriques réflexives présentées ci-après.

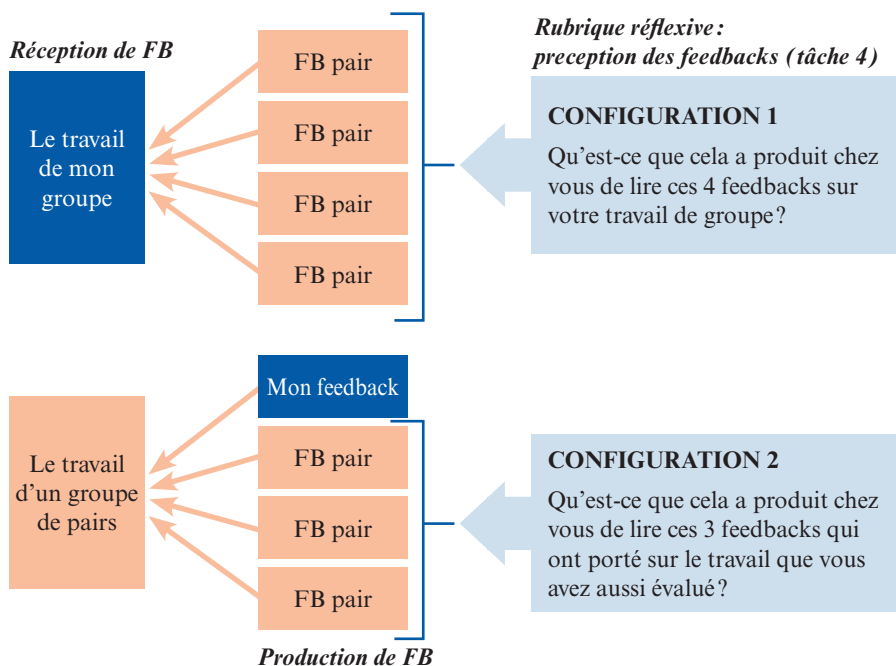
Les tâches 5 et 6 ont incité les étudiants à poursuivre la comparaison de l'intégralité des feedbacks des pairs portant sur leur travail initial (incluant les quatre feedbacks de la tâche 4) dans le but de se positionner explicitement, d'abord individuellement (tâche 5), puis sous forme de modération sociale (tâche 6). Pour celle-ci, les étudiants des mêmes groupes que lors du travail de groupe avaient pour consigne de construire collectivement des consensus à propos des pistes de régulation à mobiliser pour améliorer leur travail qui était à rendre lors de la tâche 7. La tâche 8 consistait en un retour sur l'entièreté de l'ECPA expérimentée.

Rubriques réflexives de la tâche 4

Les résultats présentés dans cet article concernent les deux rubriques réflexives de la tâche 4, qui implique les deux configurations de comparaison de feedbacks définies dans le cadre conceptuel. La Figure 5 présente l'énoncé de la rubrique réflexive associée à chacune des configurations.

3. Un des objectifs du cours (Mottier Lopez, 2020) communiqués aux étudiants était de reconnaître les caractéristiques d'un feedback susceptible de soutenir la régulation des apprentissages. Les contenus enseignés ont porté sur la définition du feedback, sur les visées d'un feedback formatif, sur les effets et limites sur les apprenants ainsi que sur les recommandations issues de la littérature (notamment, Allal & Mottier Lopez, 2005; Black & Wiliam, 1998; Hattie & Timperley, 2007). Les recommandations 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 ; (d) être prudent dans ses interprétations ; (e) considérer que la bienveillance inclut l'exigence (Mottier Lopez, 2020). La tâche préparatoire B a été exploitée afin de dégager des pistes pour produire un feedback formatif et bienveillant selon eux.

Figure 5
*Les deux configurations de comparaison de feedbacks
 et leur rubrique réflexive de la tâche 4*



Dans la configuration 1, l'étudiant répond en tant que coauteur du travail de son groupe (tâche 2) ayant fait l'objet des quatre feedbacks rédigés par des pairs (tâche 3). Il est dans la posture de réception des feedbacks des pairs. Dans la configuration 2, l'étudiant répond en tant qu'évaluateur du travail d'un groupe de pairs à propos des feedbacks des pairs sur le travail qu'il a lui-même évalué. Il est ici dans la posture de production de feedbacks.

Corpus et analyses

Au total, 123 étudiants (88 femmes et 35 hommes) ont rempli les deux rubriques réflexives de la tâche 4. La commission d'éthique de l'université en question n'a pas exigé que ce projet passe par une procédure d'approbation éthique. Nous avons sollicité auprès des étudiantes leur consentement écrit pour analyser leurs productions anonymisées à des fins de recherche,

sans possibilité de revenir à leur nom, et seulement une fois le cours terminé. Ces conditions leur ont été très explicitement présentées. L'analyse a donc porté sur 246 rubriques, chacune devant comporter minimalement 100 mots. La plupart des étudiants ont dépassé ce minimum, allant jusqu'à produire le double, voire parfois le quadruple.

Une analyse de contenu (Bardin, 1998) des écrits réflexifs des étudiants a été effectuée à l'aide du logiciel NVivo. La codification choisie est justifiée par notre question de recherche et par notre cadre conceptuel. Ainsi, les deux configurations ont fait l'objet d'un codage séparé.

Tout d'abord, tous les extraits dans lesquels les étudiants ont relaté des éléments suggérant des *comparaisons* entre les feedbacks ont été identifiés. Ensuite, les extraits ont été codés selon si les étudiants relevaient des *similarités* et/ou des *différences* entre les feedbacks. Pour chaque cas de figure, nous avons observé que les étudiants exprimaient principalement des *jugements*, des *ressentis* et des *intentions de régulation*, ce qui nous a amenés à élaborer une codification analytique de ces trois catégories émergentes (voir Tableau 2).

Tableau 2
Trois niveaux de codage

Identification d'éléments de comparaison					
Similarités entre feedbacks			Différences entre feedbacks		
Jugements	Ressentis	Intentions de régulation	Jugements	Ressentis	Intentions de régulation

Une fois cette codification stabilisée par les trois auteurs, un des auteurs a codé systématiquement l'ensemble du corpus. Celui-ci a fait l'objet de va-et-vient entre les codages effectués et des discussions collectives avec les deux autres auteurs pour assurer la validité de l'analyse et pour effectuer des rectifications, au besoin.

Résultats

Pour rappel, notre question de recherche est la suivante: *Dans le cadre du dispositif d'évaluation présenté, comment les étudiants, qui sont à la fois évalués et évaluateurs, perçoivent-ils les feedbacks multiples des pairs, notamment quand ils sont amenés à les comparer dans la perspective de réguler leur propre travail universitaire?*

Les résultats ci-dessous sont exposés pour chacune des deux configurations constitutives de la tâche 4 de l'ECPA. Ils sont structurés en fonction des comparaisons faites par les étudiants. Nicol (2020) insiste sur l'importance de rendre accessibles les référents de l'évaluation à des fins de comparaison. Dans notre cas, ce sont les critères d'évaluation qui fonctionnent comme références :

- les critères d'évaluation portant sur le travail de groupe associés aux apprentissages (ancrage dans le cas de Victor, adéquation des dispositifs avec les objectifs visés, et mobilisation correcte des concepts liés à l'évaluation et à la régulation des apprentissages) ; et
- les critères portant sur la production de feedbacks aux pairs relevant des compétences à produire et à recevoir des feedbacks a priori formatifs et bienveillants (adéquation avec les critères d'évaluation du travail de groupe, justesse des contenus rédigés, spécificité vis-à-vis du travail des pairs, présence de pistes d'amélioration et bienveillance).

Ces deux cadres de référence critériée organisent les comparaisons présentées ci-dessous, qui incluent des extraits de corpus à titre d'illustrations.

Configuration 1: « Moi » récepteur de feedbacks multiples sur le travail de mon groupe

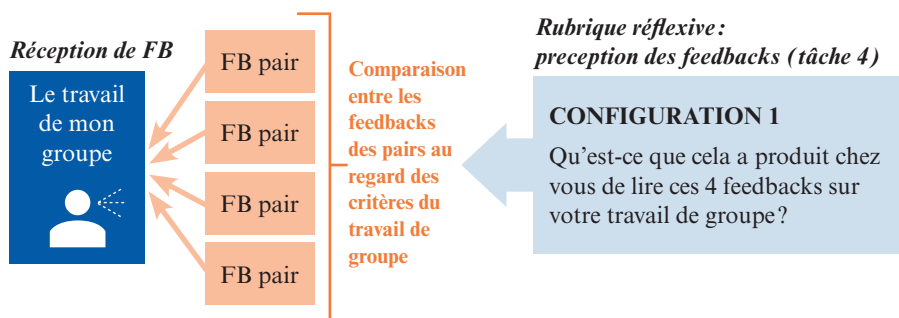
Comparaison au regard des critères portant sur le contenu du travail de groupe⁴

La Figure 6 rappelle les caractéristiques de cette première configuration de la tâche 4 de l'ECPA dans laquelle l'étudiant s'exprime en tant que (co)auteur de son travail de groupe.

4. Pour rappel, ces critères étaient : ancrage dans le cas de Victor, adéquation des dispositifs avec les objectifs visés, et mobilisation correcte des concepts liés à l'évaluation et à la régulation des apprentissages.

Figure 6

**Configuration 1 : comparaison entre les feedbacks des pairs
à propos de son travail de groupe**



Lorsque les étudiants comparent les contenus des quatre feedbacks reçus sur leur propre travail de groupe au regard des critères d'évaluation de ce travail, il apparaît que les éléments qu'ils pointent sont différents selon qu'ils constatent des similarités ou des différences entre les feedbacks des pairs. Pour des questions de lisibilité, ces deux cas de figure sont présentés ci-dessous de façon distincte, même s'ils sont susceptibles de s'entremêler dans les écrits réflexifs.

– ***Quand les étudiants constatent des similarités***

La constatation par les étudiants de similarités entre les contenus des quatre feedbacks reçus à propos de leur travail de groupe semble avoir un impact sur les *ressentis* exprimés par les étudiants, particulièrement lorsque les feedbacks sont positifs et qu'ils valident le travail effectué. Les commentaires portant sur les critères de production du travail de groupe qui sont positifs et similaires entre plusieurs feedbacks ont été perçus comme «réconfortants» (NIC⁵), ont procuré «un sentiment de satisfaction» (SEV) ou encore «permet[traient] de garder une motivation et de vouloir améliorer le travail» (LAM). Globalement, les étudiants se disent réconfortés et rassurés sur la qualité de leur travail de groupe lorsqu'ils prennent connaissance de feedbacks positifs et convergents portant sur les critères de production de celui-ci.

5. À des fins d'anonymisation, nous indiquons une combinaison de trois lettres issues des noms et prénoms de l'étudiant ou de l'étudiante dont nous citons un extrait de l'écrit réflexif.

Lorsque des pistes d'amélioration reviennent dans plusieurs feedbacks, les étudiants ont tendance à les considérer comme valides. Ils expriment, dans la majorité des cas, une intention de régulation de leur travail de groupe en prenant en compte les commentaires de leurs pairs :

Beaucoup de remarques se rejoignent, ce qui veut dire que notre travail n'était pas complet. Il nous faut redéfinir certaines notions afin de montrer que nous les avons saisies. Il nous faut également expliciter plus clairement et donner plus de détails concernant les dispositifs proposés. (INR)

La majorité des personnes ont soulevé des éléments très pertinents, et ceux-ci m'ont permis de revenir sur mon travail et de constater que, pour la plupart, les aspects « critiqués » méritaient une amélioration. (ISV)

Globalement, les feedbacks des pairs contenant des pistes d'amélioration convergentes apparaissent aux yeux des étudiants comme motivants pour apporter des régulations à leur travail de groupe.

– *Quand les étudiants constatent des différences*

La constatation de différences entre les contenus des feedbacks reçus à propos de leur travail de groupe enclenche chez de nombreux étudiants une prise de conscience de la diversité possible des jugements évaluatifs sur un même travail :

Je me suis rendu compte qu'un feedback était quelque chose de très personnel. Même si nous avons des critères pour nous guider, nous n'avons pas les mêmes points de vue [...]. C'est pourquoi j'ai reçu quatre feedbacks complètement différents sur un même travail. (MAU)

Pour certains, cette prise de conscience semble avoir été appréciée comme une « expérience stimulante, motivante et enrichissante » (NIA), car il était « intéressant de voir que chaque personne a une façon et un regard spécial [...] porté sur [son] dispositif » (NIA), ce qui peut avoir « enrichi [s]es idées pour l'améliorer » (OCA).

D'autres expriment au contraire une déstabilisation face à cette prise de conscience, ce qui les amène à se questionner sur la validité des feedbacks des pairs :

Les quatre feedbacks amenaient des critiques et des conseils différents. Cela m'interroge et me laisse perplexe. En effet, je ne sais pas quels feedbacks je dois réellement prendre en compte et auxquels je peux me fier. De plus, je les trouve tous pertinents et cela me conforte dans mes questionnements. (SYB)

Il y a eu un feedback qui m'a un peu perturbé, car il allait contre tous les autres. Cela fait qu'on ne sait pas si on doit se fier à la majorité ou à l'autre, ce qui va rendre la correction un peu plus ardue. (BEF)

Les étudiants se disent souvent perturbés par la diversité des feedbacks reçus. En revanche, pour certains, cette diversité les conduit à mener une analyse critique des feedbacks des pairs :

Le fait de recevoir différents feedbacks pour un même travail permet d'avoir différents points de vue sur ce qu'on a produit. J'ai remarqué que chaque personne ne relevait pas les mêmes points, les mêmes aspects. En mélangeant tous ces feedbacks, on peut avoir un bon commentaire sur l'ensemble de notre travail. (SAF)

Quelques étudiants déclarent s'être créé un *feedback global* à partir des points de vue et des jugements différents des pairs.

Comparaison au regard des critères portant sur la production de feedbacks⁶

Toujours dans la configuration 1 (« moi » récepteur de feedbacks multiples sur le travail de mon groupe), il apparaît que les étudiants n'ont pas seulement comparé les feedbacks par rapport à ce qu'ils disaient de leur travail de groupe, mais également au regard des critères de qualité des feedbacks (c.-à-d. à propos de leur caractère formatif et bienveillant). Si les étudiants s'expriment toujours en tant que « moi récepteur de feedbacks », certains comparent également les feedbacks reçus en revêtant les lunettes du « moi producteur de feedback ». La Figure 7 illustre ce cas de figure.

Il ressort de notre analyse des écrits réflexifs que les étudiants ne font état que de *différences* entre les feedbacks des pairs dans ce cas de figure. Parfois, les étudiants s'expriment plutôt en tant que (*co*)auteurs de leur travail de groupe. Ils expriment alors essentiellement leurs *ressentis* à la lecture des feedbacks reçus.

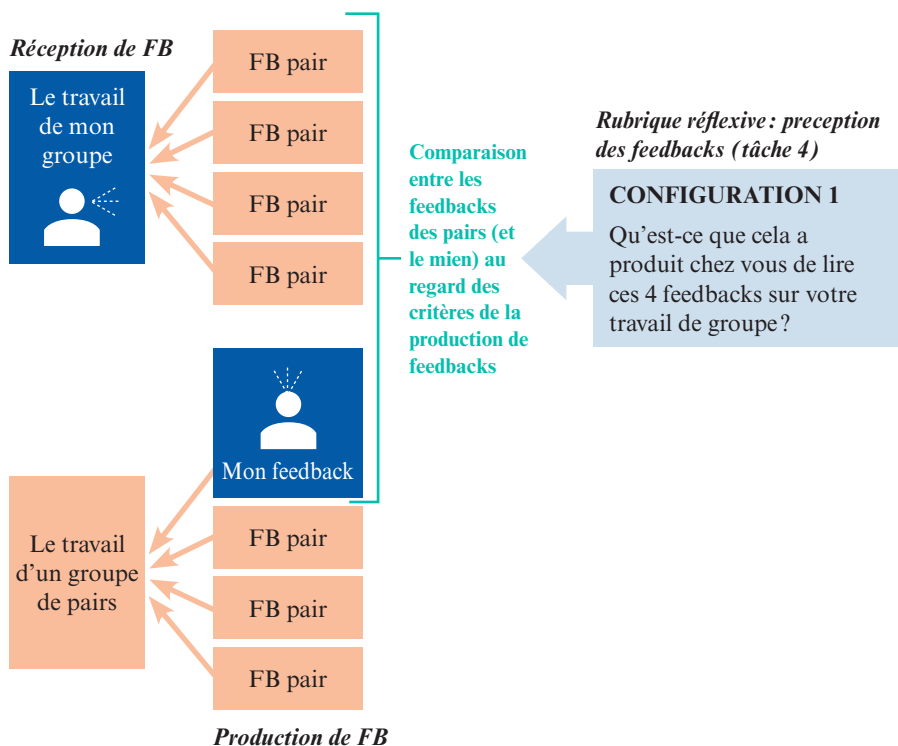
Un étudiant explique que « certains [feedbacks] [l]'ont certes un peu énervé et c'est à ce moment qu'[il a] compris toute l'importance de la bienveillance » (PAS). Le ressenti et la perception du jugement évaluatif apparaissent intimement liés :

La conclusion pour moi de tous ces feedbacks, c'est que je préfère largement un feedback bienveillant qu'un autre ! Ce qui m'a le plus plu, ce sont les feedbacks où l'auteur est impliqué dans son texte en donnant des exemples personnels ou

6. Pour rappel, ces critères étaient : adéquation avec les critères d'évaluation du travail de groupe, justesse des contenus rédigés, spécificité vis-à-vis du travail des pairs, présence de pistes d'amélioration et bienveillance.

Figure 7

Configuration 1: comparaison entre les feedbacks des pairs (et le sien) au regard des critères d'évaluation portant sur la production de feedbacks



en essayant de comprendre ce qui a pu se passer sur une question mal comprise et en se mettant au même niveau de celui qui a fait l'exercice en l'encourageant. En revanche, ce qui m'a déplu, ça a été un feedback à mon sens « hyper-complet » avec beaucoup de choses à revoir et des termes parfois assez secs. Ça a été le pire, car j'ai eu l'impression que tout ce que disait l'auteur était vrai et il y avait énormément de choses à revoir ! (ELV)

La perception de la bienveillance, notion travaillée dans le cadre du cours, est fortement soulignée dans les écrits réflexifs. Les éléments du feedback qui procurent un ressenti de bienveillance semblent propres à chaque étudiant. Par exemple, pour ELV, ce sont les marques d'empathie

ou encore l'implication de l'auteur du feedback. Pour MAV, la bienveillance se manifeste dans la forme des phrases, dans la structure et dans la mise en page du texte :

J'ai remarqué que la formulation des phrases ou la mise en page avaient une grande influence sur mon ressenti. Les phrases trop longues, trop complexes ne sont pas motivantes et je n'arrivais pas à ne pas me sentir offensée si les dites phrases étaient des reproches. Des commentaires négatifs expliqués en plusieurs phrases simples et limpides étaient bien plus digestes et semblaient moins agressifs. (MAV)

Le constat de différences entre les feedbacks conduit certains étudiants à exprimer des effets sur leur motivation à améliorer leur travail de groupe :

Ce que je trouve très intéressant est que mon degré de motivation et d'envie concernant l'amélioration de mon travail changeait en fonction des différents feedbacks. Si le feedback était trop positif, je n'avais pas forcément envie de m'améliorer, et c'est en lisant le feedback qui m'a paru le plus exigeant, soulevant le plus de points à améliorer, que j'ai eu le plus de motivation et l'envie de m'améliorer. (JUN)

J'ai pu aussi lire un feedback qui ne m'a pas du tout plu. En effet, celui-ci contenait quasiment des aspects négatifs, ce qui m'a considérablement démotivée. Ce feedback manquait beaucoup d'équilibre, car je sentais que des reproches. (ISV)

La motivation à engager des régulations du travail de groupe apparaît liée au ressenti produit par les commentaires négatifs ou positifs perçus.

De façon intéressante, dans la configuration 1, les étudiants s'expriment également comme producteurs de feedback aux pairs (voir Figure 7, double posture d'évalué et d'évaluateur) :

Le premier feedback que j'ai découvert était extrêmement positif et élogieux, presque trop, ne mettant en valeur que le positif. Puis, les deux suivants ont été plus incisifs, pointus, critiques, mais en même temps extrêmement bien rédigés, constructifs et propices à me permettre de vraiment cerner (confirmer plutôt) mes faiblesses, que j'avais déjà en tête concernant mon [travail de groupe]. C'était très pertinent et, en les découvrant, j'ai vraiment compris que je n'avais pas rédigé mes feedbacks avec assez de précision, de technique, et de références au cours. (SAR)

Que ce soit par la comparaison des ressentis produits par les feedbacks reçus ou par la comparaison entre les feedbacks des pairs et leur propre feedback, certains étudiants expriment des intentions de régulation de leurs *compétences à rédiger un feedback* :

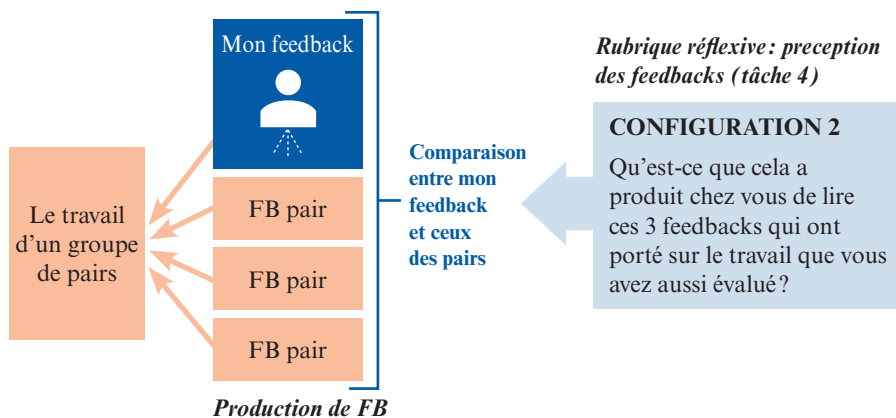
Certains d'entre eux ont très bien saisi le but des feedbacks, rédigent de façon claire, avec des objectifs précis. Ils relèvent également de bienveillance et proposent des pistes d'amélioration bien explicitées. Cela me donne vraiment envie de revoir mon feedback avec leurs commentaires en parallèle. (LAG)

Configuration 2: « Moi » producteur d'un feedback sur le travail d'un groupe de pairs

La Figure 8 rappelle les détails de cette configuration.

Figure 8

Comparaison entre son feedback et ceux des pairs sur le même travail



Tout comme pour la configuration 1, la configuration 2 amène les étudiants à comparer les feedbacks avec comme cadre de référence les critères relatifs au travail de groupe et les critères relatifs à la production de feedbacks.

Comparaison au regard des critères portant sur le contenu du travail de groupe⁷

Lorsque les étudiants comparent les feedbacks des pairs sur le travail de groupe qu'ils ont eux-mêmes évalué, une grande partie de leurs réflexions portent sur des *enjeux de contenus* liés au travail de groupe.

7. Pour rappel, ces critères étaient : ancrage dans le cas de Victor, adéquation des dispositifs avec les objectifs visés, et mobilisation correcte des concepts liés à l'évaluation et à la régulation des apprentissages.

– *Quand les étudiants constatent des similarités*

Quand les étudiants constatent des similarités entre les contenus des feedbacks produits, ils expriment principalement des *ressentis*. Ils se disent rassurés sur leur propre feedback :

La majeure partie du temps, les retours de mes camarades allaient dans le même sens que le mien et les éléments que j'avais jugé positifs ou négatifs se retrouvaient dans leurs feedbacks. Le fait d'avoir un avis similaire que ceux de mes camarades me donne plus confiance en la pertinence de mon analyse et surtout me fait prendre conscience de ma légitimité dans ce travail. (RGR)

Je trouve qu'il est intéressant de voir les avis et remarques de ses camarades concernant un travail [l'évaluation] que j'ai aussi dû faire. Je me rends compte que j'ai amené des remarques similaires à celles de mes camarades, ce qui est plutôt rassurant. En effet, je craignais que mes remarques ne soient pas pertinentes ou à côté de la plaque, mais je me rends compte que pas forcément. (INR)

Certains parlent de sentiment de soulagement. Des étudiants affirment avoir gagné en confiance en découvrant que certaines des remarques qu'ils avaient faites concordaient avec celles de leurs pairs ayant évalué le même travail de groupe. Ce constat de similarité entre les feedbacks permet aux étudiants de dissiper leurs doutes à propos de la validité de leur feedback. Globalement, il en ressort que, dans ce cas de figure, aucune intention de régulation n'est exprimée par les étudiants.

– *Quand les étudiants constatent des différences*

Dans la configuration 2, quand les étudiants constatent des différences entre les contenus des feedbacks produits, c'est pour exprimer des *ressentis* principalement *négatifs*. Les étudiants expriment des doutes, voire de l'inquiétude par rapport à leurs compétences évaluatives :

Cela est déstabilisant de voir que son avis est minoritaire sur certains points. Il peut même en résulter un certain «stress» d'être complètement «à côté de la plaque» (spécialement quand la majorité supporte cet avis différent du vôtre). (CMU)

La plupart du temps, lorsque je lisais quelque chose de différent par rapport à ma conception de la rédaction du groupe, je trouvais intéressant, toujours très bien justifié et expliqué ce qui pouvait me faire douter par rapport à mon travail. (MAP)

Les différences observées entre les feedbacks produits amènent à nouveau certains étudiants à prendre conscience de la diversité des jugements évaluatifs. Cette prise de conscience leur semble utile pour la vie future :

Cela m'a fait bizarre de voir qu'il pouvait y avoir autant d'avis différents sur un simple texte. Cela démontre que nous sommes tous subjectifs et que, même avec des critères d'évaluation, nous n'évaluons pas tous de la même manière. C'est quelque chose que je garderai pour plus tard. (BEF)

Cela permet aussi d'identifier d'autres lectures du travail du groupe – et parfois même une autre compréhension ou positionnement. Cela souligne le fait que le feedback peut être tellement différent selon le correcteur ! Ce qui, je trouve, est très important d'en prendre conscience dès maintenant, car nous serons confrontés à cela pendant toute notre carrière professionnelle d'enseignants. (CAF)

La décentration grâce à la prise de connaissance des autres feedbacks et de leurs différences amène les étudiants à porter un jugement évaluatif sur leur propre feedback et sur le feedback d'autrui :

Le fait de me comparer aux autres m'a procuré un sentiment d'infériorité, c'est-à-dire que j'avais l'impression que le mien ne valait pas les autres et, du coup, que j'aurais pu faire mieux. (MIM)

Cela m'a permis de voir ce que d'autres personnes avaient pu trouver comme pistes d'amélioration et comment ils avaient relevé les points corrects. J'ai également trouvé très intéressant de lire des feedbacks sur un autre travail de groupe que le mien, car cela permet de prendre du recul et de les juger sans la dimension personnelle. (JES)

Ils expriment aussi des intentions de régulation de leur feedback :

Le fait de voir des étudiants mieux gérer les concepts du cours que nous, de voir qu'ils s'en emparent bien pour analyser les productions de nos pairs, c'est très *challenging*. Ça donne envie de faire mieux, de savoir davantage, de maîtriser aussi. Je sens que j'étais encore trop dans le flou lors de la rédaction de mon propre feedback et ça me donne très envie de mieux faire, en m'appuyant sur des façons de faire que je vois chez mes collègues. (SAR)

Le constat de différences conduit parfois les étudiants à vérifier l'argumentation de leurs pairs en retournant aux ressources à disposition et, ce faisant, à approfondir les concepts travaillés au cours :

Lire les productions de mes pairs m'a permis d'aiguiser ma maîtrise des concepts, car je suis allé plusieurs fois vérifier les *slides* et mes prises de notes pour juger l'argumentation de mes camarades. (NMA)

Comparaison au regard des critères portant sur la production de feedbacks⁸

Dans ce dernier cas de comparaison entre feedbacks (le leur et celui des autres), seules des différences ont été mises en évidence dans les écrits réflexifs des étudiants. Les étudiants expriment majoritairement de l'empathie envers le groupe dont le travail a été évalué :

J'ai ressenti de l'empathie pour le groupe qui reçoit les feedbacks par rapport à certains feedbacks qui me semblaient très cassants et critiques. En effet, ces derniers peuvent avoir un fort impact émotionnel, blesser et démotiver, alors que ce n'est pas le but de l'exercice ni d'une évaluation en général. (MAA)

Dans cet extrait, l'étudiante prend conscience de l'impact émotionnel des feedbacks qui lui semblent manquer de bienveillance. Il est intéressant d'observer que les étudiants mobilisent explicitement les critères d'évaluation pour porter un jugement sur la qualité des feedbacks des pairs :

Ces feedbacks [...] restaient, dans leur globalité, bienveillants, faciles de compréhension et pertinents par rapport au travail réalisé. Ils sont formatifs selon moi puisqu'ils argumentent bien, donnent des pistes d'amélioration. J'ai été donc satisfaite en lisant ces feedbacks, car ils respectaient les critères demandés et sont formatifs. (CVI)

Certains des feedbacks sont notamment de qualité grâce aux améliorations qu'ils proposent [...], grâce à leur précision dans les commentaires [...], grâce à l'interprétation des réponses qui sont faites [...], grâce à leur structure claire. (DAM)

D'une façon générale, lorsque les étudiants constatent des différences dans la configuration 2, que ce soit à partir des critères d'évaluation portant sur le travail de groupe ou sur la production de feedbacks, les intentions de régulation qu'ils expriment concernent les feedbacks qu'ils auront à rédiger *dans le futur* :

Cela m'a permis de relever les points forts des différents feedbacks que je pourrais ainsi mettre en pratique dans mes futurs travaux. (DAM)

Cela va permettre d'améliorer mes prochains feedbacks puisque j'ai là sous les yeux des travaux avec des qualités et des défauts. Il ne me reste qu'à trier et ne garder que ce qui fonctionne bien. (NIC)

8 Pour rappel, ces critères étaient : adéquation avec les critères d'évaluation du travail de groupe, justesse des contenus rédigés, spécificité vis-à-vis du travail des pairs, présence de pistes d'amélioration et bienveillance.

Discussion

La première partie de cette discussion conclusive présente les principaux constats qui ressortent de notre analyse des écrits réflexifs des étudiants. La deuxième propose une élaboration conceptuelle au regard de ces constats. La dernière, sous forme de brève conclusion, revient à l'importance d'inscrire ces constats dans le contexte d'une évaluation située, ici celle de l'ECPA dans un contexte d'enseignement universitaire.

Principaux constats empiriques

Le Tableau 3 synthétise nos résultats afin de dégager ce qui ressort des perceptions des étudiants quand ils constatent des similarités et des différences entre les feedbacks comparés.

Trois ensembles de constats principaux se dégagent.

Le premier montre que l'expérience des deux postures (moi producteur/moi récepteur) offre aux étudiants des opportunités de développer à la fois leur maîtrise des contenus du travail de groupe (concepts liés au cours) et leurs compétences évaluatives à produire un feedback, et ce, dans les deux configurations. Des travaux de la littérature ont mis en évidence une complémentarité entre des tâches de production et de réception de feedbacks pour la régulation des apprentissages des étudiants (Cao et al., 2019 ; Girardet, 2020).

À notre sens, notre étude va plus loin. Elle explore deux configurations différentes dans lesquelles les étudiants se retrouvent lors d'une situation de réception de feedbacks et non étudiées à notre connaissance : la lecture de feedbacks sur leur propre travail et la lecture de feedbacks sur un travail qu'ils ont eux-mêmes évalué. Nos résultats montrent qu'amener les étudiants à exercer leur réflexivité sur les feedbacks des pairs dans ces deux configurations apporte une complémentarité intéressante pour ce qui est de la régulation de leurs compétences associées à l'enseignement et de leurs compétences à évaluer/s'autoévaluer. Nous postulons que ces configurations permettent de multiplier et de diversifier les opportunités de régulation favorables à des profils d'apprenants hétérogènes.

Le deuxième ensemble de constats, nouveau à notre connaissance, est que les comparaisons entre les feedbacks soutiennent différemment les intentions de régulation et leurs objets selon que les étudiants observent des similarités ou des différences. Les constats de similarités tendent à susciter une sécurité émotionnelle qui rassure les étudiants sur la qualité

Tableau 3
Synthèse des résultats

	Similarités	Différences
«Moi» récepteur de feedbacks		
Comparaison/ critères du TG	Convergence de commentaires positifs → soulagement sur la qualité de son TG	Prise de conscience de la diversité → intérêt et/ou déstabilisation sur les pistes à retenir
	Convergence des pistes d'amélioration → validité des feedbacks → intentions de régulation du TG	Jugement évaluatif → analyse critique des feedbacks reçus → feedback global
Comparaison/ critères de la production de feedbacks		Ressentis → jugement évaluatif sur les feedbacks reçus → (dé)motivation à réguler son TG
		Glissement vers le « moi » producteur de feedback → intentions de régulation de ses compétences évaluatives
«Moi» producteur de feedback		
Comparaison/ critères du TG	Soulagement sur ses compétences évaluatives	Doutes sur ses compétences évaluatives
		Prise de conscience de la diversité → intérêt Décentration → jugement évaluatif sur les feedbacks d'autrui et/ou sur son propre feedback Vérification des feedbacks (retour au cours) → régulation de leur compréhension des concepts
Comparaison/ critères de la production de feedbacks		Empathie envers les récepteurs Jugement évaluatif sur le respect des critères → intentions de régulation de ses compétences évaluatives

Note. TG = travail de groupe.

de leur travail de groupe dans le cas de commentaires positifs convergents (configuration 1) et sur la qualité de leur feedback aux pairs (configuration 2). Constaté des différences entre les feedbacks représente un levier de régulation plus puissant que les similarités, que ce soit au niveau des contenus du travail de groupe ou au niveau de la production de feedbacks.

Les étudiants disent alors prendre conscience de la diversité des points de vue évaluatifs des pairs. Déstabilisés, ils sont amenés à évaluer et à analyser les feedbacks. Ce résultat souligne l'importance de créer des occasions pour les étudiants de recevoir des feedbacks multiples sur leurs travaux, comme le relèvent notamment Cho et MacArthur (2010) ainsi que Nicol (2013, 2020), afin d'intensifier les opportunités de divergences favorables à la régulation des apprentissages. Globalement, la convergence des feedbacks tend à amener les étudiants à conclure à leur validité. Le consensus ayant tendance à être perçu comme valide (Cho & MacArthur, 2010), il n'encourage pas les étudiants à prendre une distance critique.

Ces constats mettent en évidence les ressentis exprimés par les étudiants. Nicol (2020) considère que la lacune principale des travaux qu'il mène sur les comparaisons faites par les étudiants concerne la dimension affective des feedbacks internes, peu documentée, bien que connaissant le rôle joué notamment par les émotions. D'ailleurs, les affects liés aux feedbacks font partie des éléments mis en évidence par la littérature comme étant importants à prendre en compte dans le développement des compétences des étudiants relatives aux feedbacks (*student feedback literacy*; Carless & Boud, 2018; Carless & Winstone, 2020).

Notre troisième ensemble de constats concerne précisément l'omniprésence de ces ressentis qui apparaissent intimement liés aux jugements évaluatifs portés par les étudiants sur les feedbacks. Dans la configuration 1, les étudiants tendent à exprimer un jugement évaluatif sur les feedbacks reçus en comparant leurs réactions émotionnelles à leur lecture. Dans la configuration 2, ils appréhendent les ressentis en se projetant par empathie dans le rôle du pair évalué, ce qui les conduit à porter un jugement évaluatif sur leur propre feedback et sur ceux des pairs, mais en étant cette fois-ci « préservés » de la dimension personnelle liée à la posture d'évalué.

Dans les deux configurations, la comparaison des feedbacks semble avoir contribué à construire toujours plus de sens aux critères concernant la production de feedbacks formatifs et bienveillants, et ce, en interaction étroite avec la dimension affective inhérente à la réception de feedbacks. Les réflexions des étudiants montrent une prise de recul sur les ressentis qu'un feedback peut enclencher sur eux-mêmes et sur autrui. De façon concordante avec la littérature, ces ressentis semblent jouer un rôle sur leur motivation (ou non) à réguler leur travail (Girardet, 2021; Rowe et al., 2014). Un certain inconfort émotionnel pourrait favoriser la régulation des apprentissages

(Molloy et al., 2013). En effet, les ressentis de doute et de déstabilisation semblent inciter certains étudiants à procéder à une évaluation critique des feedbacks pour «résoudre» la déstabilisation ressentie (Girardet, 2021).

Élaboration conceptuelle autour du feedback interne et mises en relation

Selon Nicol (2020), «*internal feedback is the new knowledge that students generate when they compare their current knowledge and competence against some reference information*» (p. 2). Les constats empiriques présentés dans cet article nous incitent à énoncer deux notions complémentaires : le *feedback global* et le *métafeedback*, liés non seulement aux critères d'évaluation prédéfinis dans l'ECPA (concernant le travail de groupe et la rédaction d'un feedback), mais également aux ressentis émergeant des étudiants. Des «effets de rebond» entre les configurations de production et de réception de feedbacks multiples ont été observés.

– Feedback global

Reprenons nos deux configurations. Dans la première, le fait de recevoir des feedbacks multiples et différents sur leur travail de groupe conduit les étudiants, par le jeu de la comparaison, à un feedback interne, puis à élaborer un *feedback global* qui intègre en partie les points de vue et évaluations des pairs. Pour ce faire, les étudiants procèdent à un jugement évaluatif et à une analyse critique des feedbacks reçus afin de décider ce qu'il faut retenir de chaque feedback dans une perspective de régulation potentielle. Ce feedback global témoigne d'une appropriation des feedbacks individuels des pairs, tout en incluant les critères de production des feedbacks et des travaux de groupe révélateurs des valeurs associées à l'ECPA et des contenus à s'approprier.

Nos analyses montrent le rôle crucial des ressentis dans cette appropriation, rôle mis à distance par le moyen des critères d'évaluation. Cette mise à distance critique est favorisée par le dispositif de l'ECPA, qui fait alterner les rôles d'évaluateur et d'évalué aux étudiants, et par le fait d'avoir travaillé les notions de feedback formatif et de bienveillance dans le cadre du cours. Notre hypothèse est que cette première configuration offre des conditions de développement de *métaconnaissances sur soi* en tant que récepteur de feedbacks, notamment par la prise de conscience et par la possibilité de verbaliser les caractéristiques des feedbacks qui «conviennent le mieux» et des émotions qu'ils génèrent.

Un premier effet de rebond s'observe alors de la configuration 1 vers la configuration 2, qui sollicite les compétences évaluatives des étudiants en tant que « producteur de feedback », autrement dit d'un acteur amené à produire lui-même une évaluation sur le travail d'autrui. Cet effet de rebond conduit certains étudiants à évaluer leur propre feedback de façon critique dans la perspective de réguler des feedbacks futurs qu'ils pourraient être amenés à faire. Cette autoévaluation de leur feedback se fonde non seulement sur les critères prédéfinis, mais également au regard de la comparaison avec les feedbacks des pairs et des émotions qu'ils ressentent. Ce faisant, nous pouvons postuler le développement de *métaconnaissances sur la tâche de production de feedbacks*.

– *Métafeedback*

Dans la configuration 2, lorsque les étudiants comparent le feedback qu'ils ont eux-mêmes rédigé avec ceux de leurs pairs, ils tendent à verbaliser un *métafeedback*, c'est-à-dire un feedback sur les feedbacks. Les critères d'évaluation portant sur la production de feedbacks jouent un rôle important en tant que cadre d'interprétation permettant aux étudiants d'objectiver le degré de pertinence qu'ils perçoivent aux feedbacks (le leur et ceux des pairs). Cette objectivation leur permet de faire le tri entre les feedbacks qu'ils jugent pertinents au regard des compétences de production de feedbacks. Ainsi, cette configuration est susceptible de contribuer au développement d'un regard critique plus fin, ce qui amène les étudiants à différencier la qualité des feedbacks et à déterminer leur valeur au regard des critères prédéfinis et également au regard des différences entre les feedbacks comparés (incluant le leur et débouchant sur son autoévaluation).

Tout comme dans la première configuration, les ressentis des étudiants deviennent des *référents* aux jugements évaluatifs qu'ils portent. Certains étudiants se projettent dans le rôle de l'évalué (en imaginant la façon dont les feedbacks peuvent être reçus par leurs pairs) tout en étant l'évaluateur. Une prise de conscience se produit en lien avec le critère de bienveillance du feedback, qui amène une décentration, médiatisée par le sentiment d'empathie. Dans cette configuration, nos constats empiriques nous permettent de postuler le développement de *métaconnaissances sur ce qu'un feedback peut produire en termes de ressentis* et sur l'impact potentiel de ces ressentis sur l'apprentissage.

Un second effet de rebond s'observe alors entre la configuration 2 vers la configuration 1. En comparant les contenus de leur feedback et ceux des autres à propos du travail des pairs (configuration 2), les étudiants développent une compréhension accrue des contenus du cours mobilisés dans les travaux de groupe. Potentiellement, cette meilleure compréhension, dont nous faisons l'hypothèse qu'elle est soutenue par les feedbacks internes englobant des feedbacks globaux et des métafeedbacks, représente une ressource pour la régulation de leur propre travail de groupe (configuration 1).

Brève conclusion

Au-delà de ses limites, qui nous interdisent une généralisation en raison de son caractère singulier, notre recherche met en évidence les apports potentiels des feedbacks multiples entre pairs dans une évaluation continue pour apprendre (ECPA) qui articule un ensemble de tâches interreliées offrant des possibilités de progression conceptuelle (compétences propres à des contenus et compétences transversales) et de régulation de travaux concrets à des fins de certification finale.

Cette ECPA est vue comme essentielle pour construire une culture d'évaluation partagée avec les étudiants qui donne du sens et qui finalise les pratiques de feedbacks entre pairs données à voir dans cet article (Mottier Lopez, 2021). Sur le plan pédagogique, un apport de notre recherche est de montrer l'importance non seulement d'apprendre à produire des feedbacks formatifs, mais également d'apprendre à les recevoir afin d'être en mesure de les exploiter de façon critique à des fins de régulation. Des études complémentaires portant sur les autres tâches de l'ECPA expérimentées viseront à toujours mieux comprendre les processus en jeu.

Réception : 7 juillet 2021

Version finale : 14 décembre 2021

Acceptation : 29 décembre 2021

RÉFÉRENCES

- Allal, L. & Laveault, D. (2009). Assessment for learning: évaluation-soutien d'apprentissage. *Mesure et évaluation en éducation*, 32(2), 99-106. <https://doi.org/10.7202/1024956ar>
- Allal, L. & Mottier Lopez, L. (2005). L'évaluation formative de l'apprentissage: revue de publications en langue française. Dans OCDE, *L'évaluation formative: pour un meilleur apprentissage dans les classes secondaires* (pp. 265-290). Éditions OCDE/CERI.
- Bain, J. (2010). Integrating student voice: Assessment for empowerment. *Practitioner Research in Higher Education*, 4(1), 14-29. <https://files.eric.ed.gov/fulltext/EJ1130598.pdf>
- Bardin, L. (1998). *L'analyse de contenu*. PUF.
- Black, P., & Wiliam, D. (1998). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan*, 80(2), 139-148. <http://electronicportfolios.org/afl/InsideBlackBox.pdf>
- Bloom, B. S., Hastings, J. T., & Madaus, G. F. (1971). *Handbook on formative and summative evaluation of student learning*. McGraw-Hill.
- 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., & Molloy, E. (2013). What is the problem with feedback? In D. Boud & E. Molloy (Eds.), *Feedback in higher and professional education: Understanding it and doing it well* (pp. 1-10). Routledge. <https://doi.org/10.4324/9780203074336>
- 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>
- Butera, F., Buchs, C. & Darnon, C. (dir.). (2011). *L'évaluation, une menace?* PUF. <https://doi.org/10.3917/puf.darno.2011.01>
- 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>
- Carless, D., & Boud, D. (2018). The development of student feedback literacy: Enabling uptake of feedback. *Assessment & Evaluation in Higher Education*, 43(8), 1315-1325. <https://doi.org/10.1080/02602938.2018.1463354>
- Carless, D., & Winstone, N. (2020). Teacher feedback literacy and its interplay with student feedback literacy. *Teaching in Higher Education*, 1-14. <https://doi.org/10.1080/13562517.2020.1782372>
- Cho, K., & MacArthur, C. (2010). Student revision with peer and expert reviewing. *Learning and Instruction*, 20(4), 328-338. <https://doi.org/10.1016/j.learninstruc.2009.08.006>
- Earl, L. M. (2003). *Assessment as Learning. Using assessment to maximize student learning*. Hawker Brownlow Education.
- Falchikov, N., & Boud, D. (1989). Student self-assessment in higher education: A meta-analysis. *Review of Educational Research*, 59(4), 395-430. <https://doi.org/10.3102/00346543059004395>
- 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 LEEe*, 2, 1-33. <https://doi.org/10.48325/rleee.002.04>
- Girardet, C. (2021). Lorsque les étudiant-es produisent et reçoivent des feedbacks entre pairs à l'université: les émotions pour apprendre. *La Revue LEEe*, 5, 1-35. <https://doi.org/10.48325/rleee.005.01>

- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81-112. <https://doi.org/10.3102/003465430298487>
- McConlogue, T. (2015). Making judgements: Investigating the process of composing and receiving peer feedback. *Studies in Higher Education*, 40(9), 1495-1506. <https://doi.org/10.1080/03075079.2013.868878>
- Molloy, E., Borrell-Carrió, F., & Epstein, R. (2013). The impact of emotions in feedback. In D. Boud & E. Molloy (Eds.), *Feedback in higher and professional education: Understanding it and doing it well* (pp. 50-71). Routledge. <https://doi.org/10.4324/9780203074336>
- Mottier Lopez, L. (2016). La microculture de classe: un cadre d'analyse et d'interprétation de la régulation située des apprentissages des élèves. Dans B. Noël & S. C. Cartier (dir.), *De la métacognition à l'apprentissage autorégulé* (pp. 67-78). De Boeck.
- Mottier Lopez, L. (2020, 11 novembre). *Les feedbacks en évaluation pour apprendre* [support de cours non publié].
- Mottier Lopez, L. (2021). Une évaluation continue pour apprendre durablement, une évaluation à visée inclusive. *Revue suisse de pédagogie spécialisée*, 4, 9-16.
- Mottier Lopez, L., & Girardet, C. (2022). L'évaluation pour mieux soutenir les apprentissages: un exemple dans l'enseignement supérieur. In B. Albero & J. Thievenaz (Eds.), *Enquêteur dans les métiers de l'humain: traité de méthodologie de la recherche en sciences de l'éducation et de la formation* (Tome III, pp. 283-297). Editions Raison et Passions.
- Nicol, D. (2013). Resituating feedback from the reactive to the proactive. In D. Boud & E. Molloy (Eds.), *Feedback in higher and professional education* [Kindle DX version]. Routledge.
- Nicol, D. (2020). The power of internal feedback: Exploiting natural comparison processes. *Assessment & Evaluation in Higher Education*, 23(1), 756-778. <https://doi.org/10.1080/02602938.2020.1823314>
- Nicol, D., Thomson, A., & Breslin, C. (2014). Rethinking feedback practices in higher education: A peer review perspective. *Assessment & Evaluation in Higher Education*, 39(1), 102-122. <https://doi.org/10.1080/02602938.2013.795518>
- Panadero, E., Jonsson, A., & Strijbos, J.-W. (2016). Scaffolding self-regulated learning through self-assessment and peer assessment: Guidelines for classroom implementation. In D. Laveault & L. Allal (Eds.), *Assessment for learning: Meeting the challenge of implementation* (pp. 311-326). Springer. https://doi.org/10.1007/978-3-319-39211-0_18
- Rowe, A. D., Fitness, J., & Wood, L. N. (2014). The role and functionality of emotions in feedback at university: A qualitative study. *The Australian Educational Researcher*, 41, 283-309. <https://doi.org/10.1007/s13384-013-0135-7>
- Schillings, M., Roebertsen, H., Savelberg, H., van Dijk, A., & Dolmans, D. (2021). Improving the understanding of written peer feedback through face-to-face peer dialogue: Students' perspective. *Higher Education Research & Development*, 40(5), 1100-1116. <https://doi.org/10.1080/07294360.2020.1798889>
- Strijbos, J. W., Ochoa, T. A., Sluijsmans, D. M. A., Segers, M. S. R., & Tillema, H. H. (2009). Fostering interactivity through formative peer assessment in (web-based) collaborative learning environments. In C. Mourlas, N. Tsianos & P. Germanakos (Eds.), *Cognitive and emotional processes in web-based education: Integrating human factors and personalization* (pp. 375-395). IGI Global. <https://doi.org/10.4018/978-1-60566-392-0.ch018>
- Wiggins, G. P. (1998). *Educative assessment: Designing assessments to inform and improve student performance*. Jossey-Bass.
- Yang, M., & Carless, D. (2013). The feedback triangle and the enhancement of dialogic feedback processes. *Teaching in Higher Education*, 18(3), 285-297. <https://doi.org/10.1080/13562517.2012.719154>