Document generated on 07/18/2025 9:34 p.m.

Mesure et évaluation en éducation

Évaluation formative et autorégulation des apprentissages des élèves : propositions méthodologiques pour l'analyse de processus *in situ*

Céline Lepareur and Michel Grangeat

Volume 41, Number 2, 2018

URI: https://id.erudit.org/iderudit/1059175ar
DOI: https://doi.org/10.7202/1059175ar

See table of contents

Publisher(s)

ADMEE-Canada - Université Laval

ISSN

0823-3993 (print)
2368-2000 (digital)

Explore this journal

Cite this article

Lepareur, C. & Grangeat, M. (2018). Évaluation formative et autorégulation des apprentissages des élèves : propositions méthodologiques pour l'analyse de processus *in situ*. *Mesure et évaluation en éducation*, *41*(2), 131–162. https://doi.org/10.7202/1059175ar

Article abstract

The research concerns the study of the effects of different practices of formative assessment on student self-regulation, in the context of inquiry-based science teaching. An indicator grid is constructed to analyze in situ processes, and two teaching situations are compared. The first corresponds to formative assessment such as teachers implement it in their daily practice. The second concerns the assessment practices implemented by the same teachers the following year, after a reflective workshop with the researchers. After confirming through a statistical analysis that teachers implement different modalities of formative assessment from one year to the next, we describe the evaluative practices of teachers and their effects on students. Our results show a better balance in the use of different formative assessment modalities by the teachers in the second situation and a more efficient self-regulation by the students in the proposed task.

Tous droits réservés © ADMEE-Canada - Université Laval, 2018

This document is protected by copyright law. Use of the services of Érudit (including reproduction) is subject to its terms and conditions, which can be viewed online.

https://apropos.erudit.org/en/users/policy-on-use/

This article is disseminated and preserved by Érudit.

Érudit is a non-profit inter-university consortium of the Université de Montréal, Université Laval, and the Université du Québec à Montréal. Its mission is to promote and disseminate research.

https://www.erudit.org/en/

Évaluation formative et autorégulation des apprentissages des élèves : propositions méthodologiques pour l'analyse de processus *in situ*

Céline Lepareur
Michel Grangeat
Université Grenoble Alpes
Laboratoire de recherche sur les apprentissages en contexte (LaRAC)

Mots clés : évaluation formative, autorégulation, enseignement scientifique fondé sur l'investigation, analyse vidéo

La recherche concerne l'étude des effets de différentes modalités d'évaluation formative sur l'autorégulation des apprentissages des élèves, dans le cadre de l'enseignement scientifique fondé sur l'investigation. Des indicateurs nous permettant d'analyser les processus in situ *sont construits et deux situations d'enseignement sont comparées. La première porte sur l'évaluation formative que les enseignants mettent en œuvre dans leur pratique quotidienne. La seconde concerne l'évaluation formative mise en œuvre par les mêmes enseignants l'année suivante, à la suite d'un travail réflexif avec des chercheurs. Après avoir confirmé, par le biais d'une analyse statistique, que les enseignants mettent en œuvre différentes modalités d'évaluation formative d'une année à l'autre, nous décrivons les pratiques évaluatives des enseignants et leurs effets sur les élèves. Nous observons un meilleur équilibre dans l'usage des différentes modalités d'évaluation formative par les enseignants dans la seconde situation ainsi qu'une autorégulation plus efficace des comportements des élèves dans la tâche proposée.*

Key words: formative assessment, self-regulation, inquiry-based science education, video analysis

The research concerns the study of the effects of different practices of formative assessment on student self-regulation, in the context of inquiry-based science teaching. An indicator grid is constructed to analyze in situ *processes, and two teaching situations are compared. The first corresponds to formative assessment*

such as teachers implement it in their daily practice. The second concerns the assessment practices implemented by the same teachers the following year, after a reflective workshop with the researchers. After confirming through a statistical analysis that teachers implement different modalities of formative assessment from one year to the next, we describe the evaluative practices of teachers and their effects on students. Our results show a better balance in the use of different formative assessment modalities by the teachers in the second situation and a more efficient self-regulation by the students in the proposed task.

Palavras-chave: avaliação formativa, autorregulação, ensino científica baseada em investigação, análise de vídeo

A investigação diz respeito ao estudo dos efeitos de diferentes modalidades de avaliação formativa sobre a autorregulação das aprendizagens dos alunos, no âmbito do ensino científico baseado na investigação. São construídos indicadores que nos permitem analisar processos in situ e são comparadas duas situações de ensino. A primeira tem a ver com avaliação formativa que os professores implementam na sua prática diária. A segunda diz respeito à avaliação formativa implementada pelos mesmos professores no ano seguinte após um trabalho reflexivo com investigadores. Após a confirmação, através do enviesamento da análise estatística, que os professores implementam diferentes modalidades de avaliação formativa de um ano para o outro, descrevemos as práticas de avaliação dos professores e os seus efeitos nos alunos. Observamos um melhor equilíbrio no uso de diferentes modalidades de avaliação formativa pelos professores na segunda situação, bem como uma autorregulação mais eficaz dos comportamentos dos alunos na tarefa proposta.

Note des auteurs : Cette recherche fait partie du projet européen ASSIST-ME, qui a reçu un financement de la part de la communauté européenne (nº de projet 321428).

La correspondance liée à cet article peut être envoyée à [celine.lepareur@hepl.ch] et à [michel.grangeat@univ-grenoble-alpes.fr].

Introduction

La recherche porte sur l'identification de différentes modalités d'évaluation formative mises en œuvre par des enseignants de sciences du second degré (élèves de 11 à 15 ans) en France et de leurs effets sur l'autorégulation des apprentissages. Inscrite dans le projet européen ASSIST-ME[1], la recherche concerne l'enseignement scientifique fondé sur l'investigation (ESFI)[2]. L'ESFI fait l'objet d'une réflexion internationale depuis le début des années 1990. L'introduction de cette démarche répond à la volonté d'accroître l'efficacité de l'enseignement des sciences en favorisant le développement d'une culture scientifique chez les élèves, et en donnant une image plus riche et diversifiée des démarches scientifiques (Boilevin, 2013). En France, les programmes mentionnent explicitement la démarche d'investigation au primaire depuis 2002 et au secondaire depuis 2005. L'objectif d'un renouvellement de ces pratiques d'enseignement est de rendre l'apprentissage plus actif ainsi que de permettre l'élaboration d'une véritable réflexion, facilitée par la collaboration entre les élèves, et la modification des conceptions par rapport au statut de l'erreur (Calmettes et Matheron, 2015 ; Grangeat, 2011).

Cette recherche vise à répondre à la question soulevée par Grangeat (2013) qui, en conclusion d'une étude européenne sur l'ESFI, montre que ce mode d'enseignement n'est pas toujours efficace pour tous les élèves et que « ces effets sont limités par le type d'évaluation proposé par les enseignants » (p. 248). Parallèlement à l'enjeu empirique de cette étude, l'aspect méthodologique apparaît crucial pour l'exploration de la question traitée. En effet, le développement de méthodes d'analyse qui permettent de saisir les pratiques d'évaluation informelles en situation écologique de classe et les processus d'autorégulation des élèves se révèle complexe, en raison du caractère situé et intégré de ces pratiques et processus dans les interactions entre les acteurs de la situation d'enseignement-apprentissage. Cette étude poursuit donc l'ambition de développer une classification générique des pratiques d'évaluation formative et d'autorégulation des apprentissages en vue d'appréhender les processus en jeu dans les situations de démarche d'investigation.

La première partie traite des ancrages théoriques de la recherche. À partir de deux cadres conceptuels, des indicateurs d'analyse sont dégagés. La seconde partie présente le contexte de l'étude, la procédure de recueil des données et les indicateurs sélectionnés. Les résultats obtenus sont présentés et interprétés dans une troisième partie. Une discussion autour des apports de l'étude mais aussi des limites clôture cet article.

Évaluation formative et autorégulation des apprentissages : quelle(s) articulation(s) ?

Cette première partie présente une synthèse des cadres théoriques mobilisés autour de l'évaluation formative, d'une part, et de l'autorégulation des apprentissages, d'autre part, en vue de les articuler dans un modèle qui les appréhenderait conjointement.

Évaluation formative : choix d'un modèle

Dans son acception générale, l'évaluation constitue un levier pour permettre aux enseignants d'identifier le niveau de compréhension de leurs élèves et le degré d'atteinte des objectifs visés par un enseignement (Wiliam, 2011). L'évaluation dite formative permet, elle, d'aider à la fois les enseignants à réguler leur enseignement, mais aussi les élèves, dans la mesure où elle est susceptible de soutenir et de développer l'autorégulation des apprentissages (Cartier et Mottier Lopez, 2017). Caractérisée en tant qu'évaluation *pour* les apprentissages dans les travaux anglo-saxons (Earl, 2003 ; Swaffield, 2011) ou en tant qu'« évaluation-soutien d'apprentissage » dans les travaux francophones (Allal et Laveault, 2009), la fonction principale de l'évaluation formative est de soutenir la régulation des enseignements et des apprentissages en cours de réalisation (Mottier Lopez, 2015). Ainsi appréhendée en tant que processus, elle « englobe l'ensemble des pratiques de classe quotidiennes à travers lesquelles les enseignants, les pairs et les apprenants s'informent, reconnaissent le processus d'apprentissage de l'élève et y répondent, d'une manière qui vise à améliorer son apprentissage ainsi que sa capacité à apprendre » (Cowie, Moreland et Otrel-Cass, 2013, p. 9, trad. libre). Cette définition met en avant l'utilisation conjointe, par l'enseignant et les élèves, d'informations pour réduire l'écart existant entre un niveau de performance atteint par l'élève et un niveau de performance désiré (Sadler, 1998 ; Shepard, 2009 ; Wiliam, 2010).

Dans la littérature anglophone, cette notion d'écart constitue un aspect central de l'évaluation, généralement développé en relation avec le feedback.

Le feedback

La méta-analyse réalisée par Hattie et Timperley (2007) sur l'analyse conceptuelle du feedback et de ses effets montre comment celui-ci peut être utilisé à la fois pour améliorer la réussite des élèves et l'efficacité des enseignants dans la classe. Les auteurs définissent le feedback comme étant l'information fournie par une source (enseignant, parent, pair, livre, etc.) concernant les aspects d'une performance ou d'une compréhension. Le feedback produit de l'information reliée à une tâche ou au processus d'apprentissage, ce qui permet de combler l'écart entre ce qui est compris et ce qui doit être compris (Sadler, 1998). L'objectif du feedback étant de permettre de réduire cet écart, l'environnement d'apprentissage serait efficace lorsque l'enseignant et les élèves tentent de répondre ensemble à ces questions : Quels sont les buts ? Comment faire pour atteindre ces buts ? Que faire ensuite ? (Hattie et Timperley, 2007).

Cinq stratégies de l'évaluation formative

La caractérisation des pratiques d'évaluation formative constitue un enjeu de recherche important lorsque les études visent à étudier leurs effets sur les apprentissages des élèves. Eu égard au nombre important de modèles existants, la classification des cinq stratégies d'évaluation formative développée par Wiliam (2010)[3] est retenue ici. Ce choix résulte du fait que le modèle synthétise les principaux éléments constitutifs d'une pratique d'évaluation formative les plus fréquemment développés dans la littérature anglophone, mais aussi du large spectre de contextes d'application possibles de ces stratégies. En effet, les enseignants ont besoin d'une certaine flexibilité pour être en mesure de modifier leurs pratiques évaluatives en fonction de leur propre contexte de classe. Les auteurs conceptualisent donc l'évaluation formative en cinq stratégies génériques, à partir des trois questions définies par Hattie et Timperley en relation avec le feedback. La catégorisation s'effectue en considérant à la fois le rôle de l'enseignant, des apprenants et de leurs pairs dans le processus évaluatif (voir Tableau 1).

Tableau 1
Stratégies de l'évaluation formative d'après Wiliam (2010)
Les nombres entre parenthèses indiquent à quelle stratégie chacun des aspects réfère.

Acteurs	**Où les élèves vont-ils?**	**Où en sont-ils dans leurs apprentissages?**	**Comment et par quelles étapes peuvent-ils atteindre les buts?**
Enseignant	Clarifier, partager et faire comprendre les intentions d'apprentissage et les critères de réussite (1)	Organiser des discussions, activités et tâches qui produisent des données d'apprentissage (2)	Donner du feedback qui fait progresser les élèves (3)
Apprenants		Inciter les élèves à être responsables de leurs apprentissages (4)	
Pairs		Inciter les élèves à être des ressources pour leurs pairs (5)	

Ce modèle vise à fournir un point de départ aux enseignants sur la pratique de l'évaluation formative, tout en offrant la possibilité de tirer parti des opportunités locales. La mise en œuvre de ces stratégies peut ainsi varier d'une classe à l'autre, d'une discipline à l'autre et d'un enseignant à l'autre, en fonction du contenu enseigné et du style d'enseignement, et ce, afin de prendre en compte les besoins différenciés des élèves. Ce modèle positionne ainsi l'enseignant et les élèves dans un processus interactif dans lequel ils s'engagent ensemble pour soutenir et piloter l'apprentissage.

L'évaluation formative dans l'ESFI

Dans le cadre de l'enseignement des sciences, des études montrent que les pratiques d'évaluation formative varient selon plusieurs dimensions : leur aspect formel ou informel, leur degré de planification, le type de données collectées, la nature et le temps lors duquel est dispensé le feedback de l'enseignant et ses actions formatives (Ruiz-Primo, 2011). Harlen (2013), qui s'est particulièrement intéressée aux liens qu'entretiennent l'évaluation formative et l'ESFI, propose un modèle permettant d'appréhender conjointement ces pratiques. Ce modèle spiralaire, où l'élève est à la fois acteur de son apprentissage et objet des actions de l'enseignant, met en avant le besoin pour l'élève de comprendre les buts et les critères de réussite de l'apprentissage, ainsi que l'importance de ces informations pour réguler l'enseignement et l'apprentissage. L'ESFI, lors duquel « les élèves

appliquent des compétences utilisées par les scientifiques telles que le questionnement, la collecte de données, le raisonnement et l'examen des preuves à la lumière des connaissances disponibles » (Harlen, 2013, p. 14), ainsi renforcé par l'évaluation formative, permettrait un équilibre entre la construction des connaissances et l'appropriation des savoirs par les élèves.

En résumé, Harlen (2013) comme Wiliam (2010) insistent sur la double visée du feedback ; ce dernier permet d'informer, d'un côté, les enseignants afin d'ajuster les objectifs proposés aux élèves et, de l'autre, les élèves pour les aider à prendre les dispositions nécessaires en vue d'améliorer leur compréhension ou leurs compétences. Nous considérons que ces modèles se rejoignent dans une vision de l'évaluation formative qui vise à capitaliser des moments de rencontre entre l'enseignant et les élèves afin de réguler les processus d'enseignement-apprentissage. Une intégration des pratiques d'évaluation formative à l'ESFI apparaîtrait donc favorable à la prise en charge par les élèves de leurs apprentissages. Pour dégager de tels effets, il reste à s'interroger sur ces processus en question : Comment se manifestent-ils et peut-on les mesurer?

Cadre conceptuel pour l'étude de l'autorégulation des apprentissages en ESFI

L'autorégulation des apprentissages : définition

Le concept d'apprentissage autorégulé (*self-regulated learning* ou SRL) peut se définir en tant que

> « processus actif et constructif avec lequel les apprenants se fixent des objectifs pour leur apprentissage et tentent de surveiller, réguler et contrôler leur cognition, motivation et comportement, guidés et limités par leurs objectifs et les caractéristiques contextuelles de l'environnement » (Pintrich, 2000, p. 453, trad. libre).

Nous savons qu'il ne suffit pas de disposer d'un répertoire de stratégies efficaces pour apprendre, car le maintien de l'effort et de l'attention pour réaliser une tâche se trouve en permanence menacé par les difficultés éprouvées ou par des activités plus attrayantes (Cosnefroy, 2011). Puisque l'autorégulation des apprentissages correspond à une conduite dirigée vers l'atteinte d'un but, un diagnostic de la situation qui met en jeu à la fois une analyse, par l'élève, de la difficulté de la tâche à effectuer et une estimation de ses propres ressources pour l'accomplir semble alors nécessaire.

Les boucles de feedback

Dans le champ de la psychologie sociale et développementale, la théorie des boucles de feedback développée par Carver et Scheier (1999) offre un éclairage sur le caractère dynamique et circulaire de l'autorégulation en se concentrant sur le rôle du feedback. L'aspect que nous retenons concerne les évaluations que réalisent les individus de leur état actuel (où ils en sont à un moment donné) en fonction du feedback perçu. Carver et Scheier (1999) considèrent que les individus, avant de se mettre en action, évaluent où ils en sont par rapport au(x) but(s) à atteindre. C'est à partir de cette évaluation que les individus mettent en place des actions qui produiront des effets sur leur environnement social et/ou matériel. L'environnement fournit en retour des informations sur l'écart au but et sur l'efficacité des actions mises en place. Chaque information reçue donne suite à une nouvelle évaluation de l'écart au but, puis à de nouvelles actions, jusqu'à l'atteinte (ou l'abandon) du but. Ainsi, l'évaluation que réalise l'individu pour effectuer une tâche donnée et le feedback perçu de son action sont indispensables à son autorégulation. Ces auteurs précisent que l'évaluation des chances de succès et la confiance dans les actions mises en place pour réduire l'écart sont déterminantes dans la persistance à l'exécution d'une tâche et dans l'effort fourni. En l'absence de confiance dans les actions accomplies, l'individu se déconcentre, voire se désengage de la tâche.

Nous transposons ce modèle aux situations d'apprentissage en classe, car nous pensons qu'il est compatible avec celui de l'évaluation formative. En effet, l'étude de l'autorégulation en contexte scolaire rejoint, à travers cette modélisation, la notion d'écart entre un niveau de performance atteint par l'élève et un niveau de performance désiré, qu'il s'agit de combler avec le soutien du feedback. Un croisement des modèles de l'évaluation formative et de l'autorégulation apparaît alors possible, chacun étant susceptible de nourrir l'autre dans la recherche d'une réduction de l'écart.

La conception d'un modèle intégré

L'évaluation formative est considérée comme un levier pour soutenir l'autorégulation des apprentissages des élèves et pour réguler efficacement l'enseignement (Allal et Mottier Lopez, 2007 ; Clark, 2012 ; Nicol et Macfarlane-Dick, 2006). Toutefois, différentes pratiques la caractérisent et certaines d'entre elles seraient plus ou moins favorables à l'autorégulation. Notre recherche vise ainsi à identifier les pratiques les plus porteuses

au soutien de l'autorégulation dans le cadre de l'ESFI. Nous présentons dans la figure 1 une mise en perspective de ces cadres en vue de les opérationnaliser.

Processus	Questions de l'enseignant à des fins d'évaluation formative		Questions de l'apprenant à des fins d'autorégulation
1. Collecte de preuves	Où sont les élèves dans leurs apprentissages?	⟺	Où est-ce que j'en suis? Quel est mon état actuel?
2. Jugement sur les buts et les critères de réussite	Où les élèves doivent-ils aller? Quelles sont les attentes?	⟺	Quel est mon but? Quel est l'objectif, le standard à atteindre?
3. Choix sur les prochaines étapes	Comment et par quelles étapes peuvent-ils y aller?	⟺	Comment puis-je atteindre mon but? Quelles stratégies mettre en œuvre pour l'atteindre?

Figure 1. ***Modèle de l'évaluation formative à visée d'autorégulation des apprentissages***

Ce modèle met en correspondance des questions qui guident l'action de l'enseignant avec celles des apprenants, questions alors indispensables à l'évaluation formative. Comme nous l'avons présentée plus haut, l'idée qu'une conception partagée des principes de l'évaluation formative et de l'ESFI produise des effets positifs sur la régulation des enseignements et des apprentissages semble cohérente. La question reste de savoir comment mettre en lumière ce lien, car, jusqu'alors, les effets de ces pratiques sur l'autorégulation des apprentissages n'ont pas encore été formellement démontrés.

Méthodologie

La question de l'identification et de la mesure des processus d'évaluation et d'autorégulation apparaît cruciale. En effet, nous nous intéressons à l'activité d'évaluation, la plupart du temps informelle, qui vise à se coordonner avec l'activité de l'élève (Mottier Lopez, 2015). Aussi, puisque l'autorégulation est par définition un processus interne à l'individu, son étude nécessite une réflexion sur les outils qui permettent de la mesurer.

Pour répondre à ces préoccupations, nous avons conduit notre recherche en contexte écologique de classe et avons choisi d'opérationnaliser les stratégies d'évaluation formative à travers deux modalités : 1) les stratégies d'évaluation dites « intuitives », c'est-à-dire celles mises en œuvre par des enseignants de sciences et de mathématiques dans le cadre de leur pratique de classe habituelle et 2) les stratégies d'évaluation formative dites « élaborées », en ce qu'elles répondent aux objectifs d'efficacité relevés dans la littérature pour soutenir l'autorégulation. De manière à dégager les effets de ces deux modalités, nous avons choisi d'isoler les variables « contexte » (c.-à-d. l'établissement scolaire), « enseignant » et « contenu d'enseignement ». En d'autres termes, nous comparons différentes pratiques d'évaluation formative mises en œuvre par le même enseignant, dans le même établissement, à la même année d'études et sur un même contenu. Ce choix implique, en contrepartie, une analyse des processus d'autorégulation sur des élèves différents.

2.1. Contexte de l'étude

Cette recherche s'inscrit dans un lieu d'éducation associé (LéA), dispositif de formation-recherche mis en place par l'Institut français de l'éducation, qui vise une collaboration entre les acteurs de terrain et de la recherche. Ce LéA[4] porte sur l'approche par compétences dans les démarches d'investigation et représente la partie expérimentale du projet européen ASSIST-ME, qui étudie la mise en œuvre de l'évaluation formative dans l'ESFI. Accompagnés d'enseignants chercheurs et de formateurs en didactique des sciences, les enseignants participent à la conception et au test de séances fondées sur l'investigation et intégrant des outils d'évaluation formative. Les situations de formation sont construites autour d'une question professionnelle en lien avec l'évaluation des compétences des élèves, que les participants doivent résoudre collectivement.

La durée de ce LéA était de trois ans (2012-2015). Notre recherche concerne les deux premières années du projet, la dernière étant destinée au test et à la diffusion des séquences. Chaque année, trois journées de regroupement en présentiel étaient planifiées à intervalle régulier. Les séances de travail au sein des équipes disciplinaires visaient la modification et l'enrichissement de séances préexistantes, sélectionnées par les enseignants, en vue d'intégrer les objectifs développés en formation. Des aménagements didactiques ont été apportés pour une meilleure prise en compte des difficultés des élèves, mais aussi pour augmenter les moyens de prise

d'information, d'une part, par l'enseignant sur la progression de ses élèves dans la tâche et dans leurs apprentissages (temps de débats, temps de travail individuel, etc.) et, d'autre part, par les élèves eux-mêmes pour évaluer leur propre progression (autoévaluation, débats collectifs, etc.).

Les modifications apportées par les enseignants sur le contenu de leurs séances sont le fruit d'un travail collaboratif mené à partir de leurs expériences, des difficultés éprouvées lors de la première mise en œuvre et des productions d'élèves. La séance travaillée par l'équipe des sciences de la vie et de la Terre (SVT) est celle du moteur des éruptions volcaniques. Il s'agit d'un travail sur la modélisation de l'éruption volcanique et de son interprétation. La séance de mathématiques est celle du nombre de diagonales d'un polygone. Il s'agit d'un problème ouvert pour lequel les élèves sont amenés à réinvestir les concepts de diagonale et de polygone et à développer une démarche de recherche. La séance de sciences physiques et chimiques (SPC) porte sur le changement d'état de la matière et intègre une expérience destinée à identifier la température d'ébullition de l'eau. Pour chacune des séances, les enseignants devaient concevoir ou améliorer un outil formel d'évaluation mis à disposition des élèves. Les séances de mathématiques duraient 55 minutes, tandis que celles des sciences expérimentales duraient 1 h 20[5].

2.2. Groupe de participants

Les établissements impliqués dans notre recherche sont deux collèges classés en Réseau d'éducation prioritaire (REP) qui mènent une réflexion depuis 2011 de « classes compétences sans note » pour les niveaux de 6e et 5e années du collège secondaire[6]. Le groupe de participants se compose d'un enseignant de SVT (P1), de deux enseignants de mathématiques (P2 et P3) et de deux enseignants de SPC (P4 et P5). Ces enseignants possèdent au minimum sept années d'expérience d'enseignement.

En regard de la précision nécessaire pour identifier les processus d'autorégulation en jeu, nous ne pouvions analyser les conduites de l'ensemble des élèves d'une classe. Un groupe d'élèves a donc été sélectionné dans chacune des classes observées. Le choix du groupe, délégué à l'enseignant, reposait sur le niveau scolaire plutôt moyen des élèves et sur une expression orale facile. Les groupes, dont le nombre était contraint par la configuration de la classe, se composaient de deux élèves en sciences expérimentales et de quatre en mathématiques. Aussi, l'année d'études des élèves

varie en fonction des séances : des élèves de 6e (11-12 ans) pour les séances de mathématiques, de 5e (12-13 ans) pour celles de SPC et de 4e (13-14 ans) pour celles de SVT. Les classes se composent en moyenne de 28 élèves.

2.3. Opérationnalisation et procédure de codage

L'étude est réalisée à partir d'un enregistrement vidéo et audio de 10 séances de classe. Dans une posture exploratoire de construction d'indicateurs, ces enregistrements apparaissent particulièrement adaptés pour réaliser des allers-retours entre les données et les cadres conceptuels mobilisés. Par ailleurs, la vidéo nous permet d'analyser conjointement les comportements, les verbalisations et l'ensemble des éléments contextuels essentiels à la compréhension des situations. Nous avons disposé une première caméra au fond de la classe, orientée vers l'enseignant. Une seconde caméra était dirigée vers le groupe d'élèves sélectionné, placée suffisamment haut et orientée vers le bas de manière à capter leurs comportements et leurs productions.

L'ensemble des interactions verbales des groupes d'élèves et des enseignants a été transcrit. Les comportements ont également été reportés lorsqu'ils étaient pertinents pour comprendre la situation. Les transcriptions ont été réalisées à l'aide du logiciel d'analyse vidéo Transana. Nous avons procédé à un découpage de la transcription à l'aide de balises temporelles, puis à l'affectation de « mots clés » correspondant aux indicateurs de recherche. À l'issue de ce codage, le logiciel fournit un rapport sur la proportion de chaque mot clé (durée de chaque comportement codé). Les données sont ensuite exportées vers d'autres logiciels en vue de réaliser des analyses statistiques.

2.4. Indicateurs d'analyse

La catégorie d'indicateurs de l'évaluation formative s'appuie sur les cinq stratégies de Wiliam (2010) (voir Annexe 1) :

- Explicitation du but et des sous-buts (S1) ;
- Évaluation de l'état actuel des élèves (S2) ;
- Feedback de progression (S3) ;
- Responsabilisation des élèves (S4) ;
- Incitation aux échanges entre pairs (S5).

Nous déclinons la stratégie 2 en fonction de la nature des indices collectés : leurs connaissances (p. ex., *qu'est-ce qu'une diagonale ?*), leur compréhension (*qu'as-tu compris de la question ?*), leur avancée dans la tâche (*où en êtes-vous sur la partie hypothèse ?*) et les critères définis dans l'outil formel d'évaluation (*où est-ce que tu te situes par rapport à cette compétence ?*). Nous distinguons également si l'incitation aux échanges se réalise au sein des groupes ou de la classe entière (S5).

Pour les indicateurs d'autorégulation, deux composantes sont distinguées (voir Annexe 2). La composante des indicateurs cognitifs, issue du modèle de Carver et Scheier (1998), comprend la perception des buts poursuivis par les élèves (p. ex., *on va faire un schéma*), l'évaluation de leur état actuel et les stratégies de résolution mises en œuvre pour atteindre leurs buts (*je pense qu'il faut colorier les diagonales pour les compter*). Comme pour l'évaluation formative, les indicateurs de l'évaluation de leur état actuel sont déclinés selon qu'ils se réfèrent à leurs connaissances (*c'est quoi, effusif ?*), à leur compréhension (*j'ai pas compris la question*), à leur avancée dans la tâche (*on a fini la partie hypothèse ?*) ou aux critères décrits dans le support formel d'évaluation (*je suis expert sur cette compétence*). Enfin, nous distinguons les stratégies liées à la résolution du problème de celles situées à un niveau *méta*, qui relèvent de stratégies plus globales d'apprentissage (*je pense qu'il faut lire plusieurs fois la règle pour apprendre*).

La composante des indicateurs motivationnels et affectifs se décline en trois sous-catégories d'indicateurs. Ces indicateurs, alors issus de l'analyse du corpus, n'étaient pas construits a priori :

1) Les attitudes correspondent au codage des temps d'attente de feedback de l'enseignant, d'inactivité et de déconcentration (l'activité des élèves n'est pas liée à la tâche). Ces indicateurs traduisent le niveau d'engagement ou de désengagement des élèves dans la tâche ;
2) Les états émotionnels que les élèves peuvent être amenés à extérioriser lors de l'exécution d'une tâche. Ils sont distingués selon leur orientation positive (enthousiasme) ou négative (anxiété) ;
3) Les perceptions exprimées sur la tâche et leurs compétences. Ces indicateurs réfèrent à l'avis qu'ils émettent sur le niveau de difficulté/facilité de la tâche et sur l'enjeu perçu.

Résultats

Une analyse factorielle des correspondances multiples (AFCM) permet de vérifier que les deux modalités étudiées (intuitives en A1/élaborées en A2) sont différentes. Elles seront caractérisées plus finement ensuite, à partir des résultats du codage des indicateurs d'évaluation formative et d'autorégulation des apprentissages. La mise en correspondance de ces résultats permettra de dégager les stratégies d'évaluation les plus favorables au soutien de l'autorégulation des élèves.

Exploration statistique des données

Nous cherchons, dans un premier temps, à vérifier que les enseignants mettent bien en œuvre des modalités d'évaluation différentes entre les deux années d'observation. Pour cela, nous procédons à une AFCM à partir des données codées sur le logiciel Transana.

Une première étape de cette analyse consiste à identifier la participation de chaque variable à la définition des deux axes produits par l'analyse dans le plan factoriel. La figure 2 illustre ces deux axes. L'axe horizontal se définit d'un côté (à droite) par la responsabilisation des élèves (S4) et par la valorisation des échanges au sein de la classe (S5) et, à son autre extrémité (à gauche), par l'explicitation des buts (S1). L'axe vertical se définit par le feedback de progression (S3) et par l'évaluation de l'avancée des élèves dans la tâche (S2). Les modalités relatives à l'évaluation des connaissances des élèves et de leur compréhension (S2) participent aux deux axes et sont, de fait, peu spécifiques de chacun des axes. Aussi, les modalités d'explicitation des critères de réussite (S1) et de valorisation des échanges au sein des groupes (S5) ne participent aucunement à la définition d'un des axes. Nous ne les faisons donc pas figurer dans la figure.

L'AFCM peut s'interpréter géométriquement, à partir d'un nuage de points, tel que l'illustre la figure 3. Chaque point, identifié par un nombre, représente la position d'un enseignant dans le plan. Ainsi les nombres 1, 2, 3, 4 et 5 correspondent à la position des enseignants P1, P2, P3, P4 et P5 respectivement. Chacun de ces points est projeté dans le plan de telle manière que sa position permet de situer leurs pratiques évaluatives lors de la première séance. Ces mêmes nombres, suivis d'un « 2 », correspondent à la position de ces mêmes enseignants lors de la seconde séance. Ce nuage de points permet ainsi d'identifier la proximité ou la distance des pratiques évaluatives des enseignants par rapport aux autres et leur dominante.

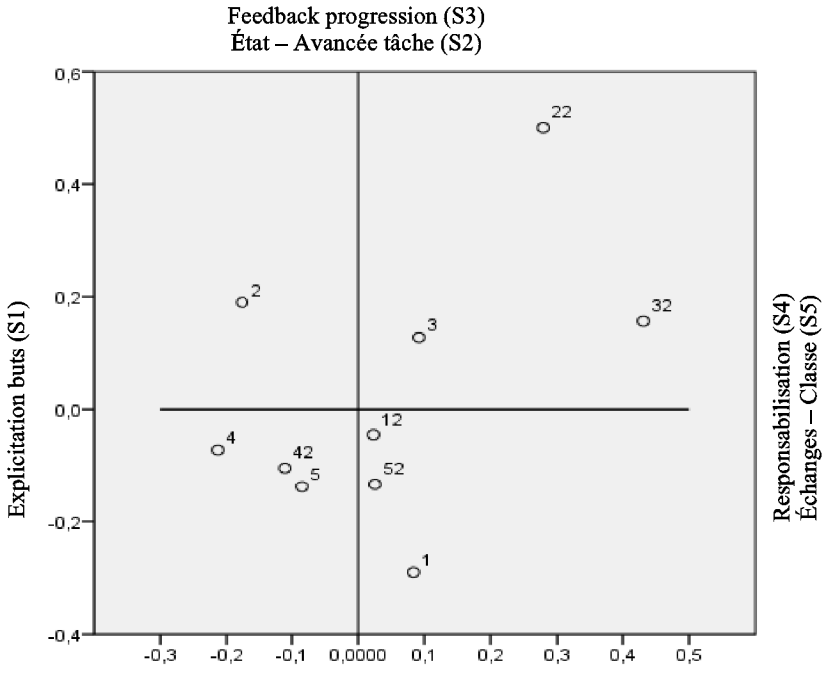

Figure 2. ***Projection des séances d'enseignement dans le plan orthogonal produit par l'AFCM***

La figure 2 montre que les enseignants ne mobilisent pas les mêmes modalités d'évaluation formative entre leurs deux séances. Lors de la première séance, les enseignants de mathématiques (P2 et P3) se situent vers le haut de l'axe vertical, c'est-à-dire en direction de plus de feedback de progression et d'évaluation de l'état des élèves en matière d'avancée dans la tâche, mais ils se distinguent entre eux sur l'axe horizontal, soit vers une plus grande explicitation des buts chez l'enseignant 2 et vers une plus grande responsabilisation et incitation aux échanges au sein de la classe chez l'enseignant 3. Les enseignants de sciences expérimentales (P1, P4 et P5) se situent, quant à eux, vers le bas de l'axe vertical, donc vers moins de feedback de progression et d'évaluation de l'avancée dans la tâche. En revanche, ils se distinguent sur l'axe horizontal, dans la mesure où les pratiques de l'enseignant de SVT (P1) se dirigent vers une plus grande responsabilisation et valorisation des échanges au sein de la classe, alors que celles des enseignants de SPC (P4 et P5) vont vers une plus grande explicitation des buts.

Lors de la seconde séance, les enseignants P2, P3, P4 et P5 se déplacent vers la droite de l'axe horizontal, c'est-à-dire vers une plus grande responsabilisation et valorisation des échanges au sein de la classe, alors que l'enseignant P1 se déplace sur l'axe vertical en direction du centre du plan. Nous observons que l'écart de déplacement chez les enseignants de SPC (P4 et P5) est relativement faible par rapport à celui des enseignants de mathématiques (P2 et P3). Ce résultat révèle une moindre modification des modalités d'évaluation mises en œuvre chez ces deux enseignants de SPC entre les deux années d'observation.

L'analyse confirme que les enseignants ne mobilisent pas les mêmes modalités d'évaluation formative entre les deux séances, validant ainsi nos deux conditions d'étude (intuitives/élaborées), et que certains enseignants mettent davantage l'accent sur certaines stratégies que sur d'autres. Notons également que les pratiques évaluatives observées la seconde année semblent évoluer dans le même sens, à savoir vers une plus grande responsabilisation des élèves et vers une incitation aux échanges au sein de la classe.

De précédentes publications ont porté sur l'analyse détaillée des séances de classe et sur leur inscription dans le cadre de démarches d'investigation (Lepareur, Gandit et Grangeat, 2017 ; Lepareur et Grangeat, 2017). L'enjeu de cet article est de présenter une analyse transversale aux disciplines et concernant l'ensemble des participants de l'étude. Il s'agit notamment de décrire les modalités d'évaluation formative identifiées chez les cinq participants de l'étude et à deux années consécutives en vue de dégager des modalités de pratiques favorables au soutien de l'autorégulation des apprentissages.

Analyse transversale des modalités d'évaluation formative

Les données relatives aux 10 séances d'enseignement sont regroupées en vue de dresser un panorama des processus en jeu. Nous procédons à une analyse des durées d'encodage des indicateurs répertoriés sous la forme de mots clés. Dans la mesure où la durée d'encodage de chaque mot clé est extrêmement variable (de plusieurs minutes à quelques secondes), se limiter à comparer le nombre d'occurrences serait peu informatif sur le poids véritablement occupé par chaque indicateur. Le tableau 2 présente la durée totale obtenue pour chaque indicateur d'évaluation formative codé, en distinguant, d'un côté, les modalités d'évaluation intuitives mises en œuvre

lors de la première année d'observation (A1) et, de l'autre, les modalités élaborées mises en œuvre lors de la seconde année (A2). Nous détaillons consécutivement ces deux modalités afin de dégager leurs caractéristiques.

Tableau 2
Durée totale des modalités d'évaluation formative mises en œuvre lors des deux séances

		Code de l'enseignant									
		P1		**P2**		**P3**		**P4**		**P5**	
	Indicateurs	A1	A2	A1	A2	A1	A2	A1	A2	A1	A2
S1	Explicitation buts	6:16	2:04	4:57	3:27	1:00	4:19	16:02	8:36	4:13	3:44
	Critères de réussite	0:29	1:19	0:00	1:40	3:03	0:20	0:44	2:26	0:19	0:00
S2	État – Connaissances	7:54	6:43	4:13	0:54	6:44	8:45	7:57	5:18	11:42	13:11
	État – Compréhension	4:01	1:24	6:41	8:19	4:38	8:28	5:39	9:33	5:17	9:48
	État – Avancée tâche	0:00	0:49	5:32	9:51	2:32	2:55	4:31	2:45	1:48	2:51
	État – Autoévaluation	0:00	0:29	0:00	0:15	0:17	0:42	0:25	0:34	0:00	0:00
S3	Feedback progrès	0:13	0:00	1:08	2:13	0:50	0:49	0:00	0:00	0:09	0:03
S4	Responsabilisation	0:18	2:25	0:06	4:10	1:25	6:31	0:44	1:48	0:45	0:55
S5	Échanges – Classe	0:00	0:00	1:11	2:38	1:52	15:38	0:00	1:10	0:31	4:32
	Échanges – Groupes	0:00	0:22	1:19	0:00	0:32	0:02	0:00	0:00	0:00	0:13

Les pratiques d'évaluation intuitives (A1)

Les modalités d'évaluation observées lors de la première année se caractérisent par une explicitation importante des buts, excepté pour l'enseignant 3, qui semble compenser cette modalité avec l'explicitation des critères de réussite. Quatre enseignants (P1, P3, P4 et P5) consacrent du temps à l'explicitation des critères, tandis que tous les enseignants prennent beaucoup d'informations sur l'état actuel des élèves, tant en ce qui concerne la compréhension, les connaissances et, dans une moindre mesure, leur avancée dans la tâche. Ainsi, les stratégies 1 et 2 apparaissent les plus caractéristiques des pratiques évaluatives du groupe d'enseignants.

Les pratiques se distinguent en revanche pour les stratégies 3, 4 et 5. D'abord, la stratégie 3, relative au feedback, n'est pas ou peu représentée pour les disciplines expérimentales (P1, P4 et P5). Nous pouvons suggérer

que la tâche mathématique, problème ouvert à résoudre par groupe, nécessite plus de feedback de la part de l'enseignant. Les séances de sciences expérimentales observées intègrent, elles, un support destiné à chaque élève présentant un ensemble de questions hiérarchisées qui guident l'investigation. Cet indicateur semble donc plus ou moins présent en fonction des tâches à réaliser. Notons également que la stratégie 4 (responsabilisation des élèves) est peu présente pour l'ensemble des enseignants ; ces derniers intègrent donc peu les propositions des élèves. Les modalités de la stratégie 5 sont aussi plus représentées dans les séances de mathématiques, ce qui indique que ces enseignants incitent davantage les échanges argumentés au sein de la classe.

Les pratiques d'évaluation élaborées (A2)

Les modalités d'évaluation observées lors de la seconde année se définissent par une moindre explicitation des buts (S1) chez quatre enseignants (P1, P2, P4 et P5). Les aménagements didactiques et l'anticipation des difficultés d'élèves ont possiblement conduit les enseignants à une meilleure clarification des buts et standards à atteindre, ce qui nécessite un temps moindre à les définir. En revanche, il est surprenant que l'indicateur d'explicitation des critères de réussite, lié au tableau de progression, ne soit pas identifié chez tous les enseignants. Il est en effet plus présent chez trois enseignants (P1, P2 et P4), mais absent pour les deux autres (P3 et P5). Dans le cadre du LéA, la question de l'explicitation des critères était pourtant un point de réflexion central. Nous pouvions donc nous attendre à ce que le travail réalisé sur le support d'évaluation conduise les cinq enseignants à expliciter ces critères et à les intégrer formellement à leur séance, en incitant les élèves à s'y référer pour guider leur avancée.

L'évaluation de l'état actuel des élèves (S2) reste la stratégie la plus mobilisée par les enseignants, et ses modalités conservent approximativement les mêmes proportions. De la même façon, la stratégie relative au feedback (S3) semble, ici encore, caractéristique des séances de mathématiques. En revanche, les modalités d'évaluation élaborées se distinguent des modalités intuitives par une plus forte responsabilisation des élèves (S4). En outre, tous semblent intégrer davantage les idées des élèves et favoriser la recherche autonome des ressources nécessaires à leur avancée. Les enseignants notent les propositions des élèves au tableau et en débattent collectivement. Ils promeuvent ainsi davantage les échanges entre élèves (S5), l'enseignant 1 valorisant une communication à l'intérieur des groupes, les quatre autres favorisant les échanges au sein de la classe.

En définitive, nous retiendrons que les modalités d'évaluation formative élaborées se définissent par une moindre explicitation des buts en faveur d'une plus grande responsabilisation des élèves et d'une valorisation des échanges entre pairs. Ces modalités impliquent une plus grande prise en compte des conceptions des élèves et de leurs initiatives pour chercher les ressources, à la fois sociales et matérielles, dont ils ont besoin. En ce sens, elles sont censées être favorables à l'autorégulation des élèves et en adéquation avec les objectifs poursuivis par les démarches d'investigation. L'analyse des processus d'autorégulation nous permettra de le vérifier.

Analyse des processus d'autorégulation des élèves

Nous commençons par développer les résultats pour la composante cognitive de l'autorégulation, puis développerons la composante motivationnelle et affective.

La composante cognitive de l'autorégulation

Excepté pour les élèves de P1, le tableau 3 montre une baisse du processus lié à l'évaluation des connaissances en jeu lors de la seconde année d'observation (A2). Ce résultat est possiblement lié à une meilleure prise en compte des connaissances initiales des élèves par les enseignants. Nous notons que ce sont les comportements d'autorégulation liés à l'évaluation de leur compréhension de la tâche qui composent la majorité des indicateurs identifiés en A2. Nous observons également une augmentation des stratégies de résolution mises en œuvre chez l'ensemble des groupes d'élèves. Ces stratégies sont considérées comme déterminantes en ce qu'elles révèlent la mise en place de processus d'autorégulation directement liés à la réduction de l'écart entre le but perçu par les élèves et leur état actuel. En effet, des régulations centrées uniquement sur les buts ou leur état peuvent témoigner d'une stagnation sur un de ces processus, sans qu'ils s'accompagnent d'une avancée dans la résolution de la tâche. L'accroissement du temps passé par les élèves à produire ces stratégies en A2 est donc révélateur d'une autorégulation plus efficace. Nous observons des stratégies d'apprentissage chez les deux groupes d'élèves de P3 (A1 et A2). Si celles-ci apparaissent uniquement chez cet enseignant, un lien avec sa pratique d'enseignement peut être supposé ; celui-ci promeut possiblement une réflexion sur les stratégies d'apprentissage au sens large, lesquelles se dégagent ensuite dans les interactions entre les élèves.

Tableau 3
Durée totale des processus cognitifs d'autorégulation des élèves choisis par chaque enseignant

		Code de l'enseignant									
		P1		**P2**		**P3**		**P4**		**P5**	
	Indicateurs	A1	A2	A1	A2	A1	A2	A1	A2	A1	A2
Processus cognitifs d'autorégulation	Buts	1:08	3:28	1:07	0:01	1:32	0:32	0:50	1:57	1:46	1:38
	Connaissances	0:49	13:22	5:48	1:26	11:40	5:23	4:26	6:33	5:16	2:26
	Compréhension	1:08	0:00	1:54	9:41	7:53	16:46	5:28	12:28	7:47	9:52
	Avancée tâche	0:52	4:13	4:12	2:13	2:23	2:16	3:56	4:29	1:29	4:21
	Autoévaluation	0:00	0:00	0:00	0:29	1:05	0:57	0:10	1:56	0:00	0:00
	Stratégie résolution	0:20	7:09	4:48	6:12	3:43	7:54	1:37	2:05	2:34	4:06
	Stratégie apprentissage	0:00	0:00	0:00	0:00	1:28	0:10	0:00	0:00	0:00	0:00

En résumé, une augmentation globale des processus cognitifs identifiés chez les élèves est observée pour toutes les séances intégrant une évaluation formative élaborée (A2). Notons que le nombre de processus identifiés dans les classes de mathématiques (P2 et P3) est remarquable au regard de la durée des séances, celles-ci étant plus courtes que celles de sciences. En effet, certaines tâches, en l'occurrence ici des problèmes ouverts en mathématiques, semblent engendrer un nombre plus important de processus d'autorégulation chez les élèves observés. Avant de tirer des conclusions sur l'efficacité supposée de certaines modalités, nous exposons les résultats obtenus pour les indicateurs motivationnels et affectifs.

La composante motivationnelle et affective de l'autorégulation

La rareté d'apparition et l'hétérogénéité des indicateurs motivationnels et affectifs ne nous permettent pas de les analyser adéquatement en fonction de la durée[7]. Une analyse du nombre d'occurrences des indicateurs est donc d'abord proposée et présentée dans le tableau 4.

Les comportements de désengagement et les émotions négatives sont moindres lors de la seconde année (A2) pour l'ensemble des groupes d'élèves, ce qui va dans le sens d'une meilleure régulation de la composante motivationnelle lorsqu'ils sont placés dans un contexte d'évaluation élaborée.

Tableau 4
Nombre d'occurrences des indicateurs de régulation motivationnelle et affective

		Code de l'enseignant									
		P1		**P2**		**P3**		**P4**		**P5**	
	Indicateurs	A1	A2	A1	A2	A1	A2	A1	A2	A1	A2
Motivation et émotion	Anxiété	1	1	0	0	1	0	0	0	0	0
	Enjeu/note	5	0	0	0	2	0	0	1	0	0
	Découragement	1	0	1	1	4	3	2	0	0	0
	Perception facilité	0	0	0	0	0	0	0	0	2	0
	Perception difficulté	5	1	3	1	4	4	0	1	1	1
	Perception soi négative	0	0	1	0	2	0	0	0	0	0
	Intérêt	0	0	0	1	0	0	0	0	0	0

Deux groupes d'élèves se distinguent particulièrement des autres. D'abord, le groupe P1 lors de la première année (A1) témoigne d'un nombre important d'affects négatifs. Les élèves de ce groupe manifestent une certaine pression évaluative au regard de leur questionnement sur le caractère noté ou non de l'activité. S'ajoutent à cela des verbalisations liées à leur perception de difficulté pour accomplir la tâche, au découragement et à l'anxiété. Les stratégies évaluatives de cet enseignant en A1 semblent donc néfastes à l'autorégulation des élèves. Un nombre important d'émotions négatives est également relevé chez les élèves de P3 en A1. Les indicateurs de découragement, de perception de difficulté, mais aussi de perception négative de soi sont identifiés. Pour creuser ces constats, nous présentons, dans le tableau 4, la durée des indicateurs codés concernant l'engagement des élèves dans la tâche. Cette variable est mesurée en comptabilisant les temps de déconcentration, d'attente de feedback et d'inactivité.

Tableau 5
Durée totale des indicateurs d'engagement dans la tâche

		Code de l'enseignant									
		P1		**P2**		**P3**		**P4**		**P5**	
	Indicateurs	A1	A2	A1	A2	A1	A2	A1	A2	A1	A2
Désengagement	Inactivité	0:00	5:55	0:00	0:00	3:12	0:00	2:39	0:00	1:00	0:00
	Attente feedback	0:00	0:00	0:05	0:40	7:05	0:00	0:00	0:18	0:00	1:34
	Déconcentration	0:00	0:00	3:19	0:45	7:40	0:10	4:47	0:00	6:25	0:00

Les temps de déconcentration diminuent, voire disparaissent chez les groupes d'élèves en A2. Nous percevons également que les temps d'inactivité (les élèves ne sont pas engagés dans la tâche) ont tendance à se transformer en temps d'attente de feedback (les élèves lèvent la main, mais continuent d'échanger). La différence la plus criante se situe entre les deux groupes de P3. Celui en A2 semble réguler ses comportements et son avancée de façon bien plus efficace que celui en A1 concernant ces variables.

Discussion

Les résultats sont discutés en deux temps : les apports empiriques obtenus, puis les apports et limites méthodologiques.

Apports empiriques

Les résultats obtenus montrent que les élèves placés dans un contexte d'évaluation formative élaborée, caractérisée par les pratiques d'enseignement mises en œuvre la seconde année, font preuve d'une autorégulation plus efficace de leurs comportements. Cette efficacité est d'abord appréciée par un temps plus conséquent passé à produire des stratégies de résolution qui visent à réduire l'écart entre le but perçu et leur état actuel. Les comportements des élèves observés la seconde année témoignent également d'une meilleure régulation motivationnelle et affective pour conduire leur activité.

La mise en œuvre d'une évaluation élaborée se traduit, chez les enseignants participants, par une mobilisation plus diversifiée des modalités d'évaluation formative. Plus précisément, les stratégies 4 et 5, liées respectivement à la responsabilisation des élèves et à l'incitation aux échanges, prennent une place plus importante dans leurs pratiques évaluatives. À l'instar de Coppé (2016), nous pensons que la viabilité de l'évaluation formative, à travers le débat argumentatif, implique un contrat didactique qui laisse aux élèves la responsabilisation et la validation de leurs réponses. Ces modalités révèlent également une évolution relativement homogène de leurs pratiques, en particulier pour les stratégies 1 et 2. Par exemple, la stratégie relative à l'explicitation des buts se réduit considérablement chez les enseignants qui la mobilisaient beaucoup lors de la première année (P1 et P4), tandis qu'elle augmente pour celui qui la mobilisait peu (P3). L'évolution de la stratégie 2, relative à l'évaluation de l'état actuel des élèves, montre que les enseignants différencient le type d'information recueillie en fonction de ce qu'ils voient du travail des élèves et de l'anticipation de leurs difficultés. L'efficacité d'une évaluation formative élaborée serait liée à la répartition équilibrée des différentes modalités et à leur articulation cohérente à l'activité des élèves.

Parallèlement, nous observons que la tâche mathématique génère un nombre plus important de stratégies de résolution que la tâche de SPC, en considération de l'écart de durée des séances. Elle produit aussi un plus grand nombre d'affects négatifs que les autres tâches proposées aux élèves. Ces résultats renforcent l'importance d'une adaptation des modalités d'évaluation en fonction des tâches et objectifs afférents. Les séances de mathématiques visent ici une responsabilisation des élèves en lien avec l'élaboration concertée de procédures par groupe pour résoudre le problème donné. Les savoirs notionnels font seulement, dans ce cas, l'objet d'un réinvestissement, et non d'un apprentissage en soi. Pour les séances de sciences expérimentales étudiées, les processus d'évaluation et d'autorégulation en jeu diffèrent, car le protocole expérimental est donné. Les enseignants visent alors l'apprentissage de la démarche scientifique par les élèves, laquelle doit permettre d'aboutir à la construction de nouveaux éléments de savoirs. Cette approche de l'évaluation formative, qui implique de considérer les savoirs enseignés et les processus d'apprentissage des élèves en conséquence, rejoint les propositions déjà énoncées par Allal, Bain et Perrenoud (1993) pour penser l'articulation entre la didactique et l'évaluation formative.

Malgré cette meilleure répartition des modalités d'évaluation, nous pensons qu'elles gagneraient à se répartir entre elles de façon encore plus équilibrée. Certaines modalités sont en effet sous-représentées ; c'est notamment le cas de l'évaluation de l'état actuel des élèves par rapport à l'outil d'autoévaluation, du feedback ou encore de la valorisation des échanges au sein des groupes. En nous référant aux principes développés par Nicol et Macfarlane-Dick (2006) pour soutenir l'autorégulation des apprenants, ces trois modalités pourraient être développées de concert. Pour ces auteurs, le feedback doit fournir de l'information qui aide l'élève, d'une part, à comprendre sa performance et, de l'autre, à s'autocorriger au regard de la perception de l'écart entre ses intentions et ses résultats. Ainsi, il doit être relié aux buts, aux standards ou aux critères de réussite. Les enseignants pourraient ainsi s'appuyer davantage sur des outils formels d'évaluation (p. ex., tableau de progression, livret personnel de compétences, livret scolaire) pour favoriser le développement d'une réflexion sur les critères et leurs productions. La valorisation des échanges sur l'apprentissage au sein des groupes pourrait alors intervenir dans le cadre d'activités de coévaluation ou de coconstruction des critères.

Apports et limites méthodologiques

Nos propositions méthodologiques reposent en partie sur la construction d'indicateurs d'analyse spécifiques aux variables étudiées, identifiables en situation écologique de classe et quantifiables. Nous discutons maintenant le dispositif de recherche et ces indicateurs.

Le dispositif de recherche

Notre dispositif de recherche nous a permis d'étudier des situations d'enseignement comparables : stables en ce qui concerne les enseignants, les séances et les contextes, mais variables quant aux stratégies évaluatives mises en œuvre. Ce dispositif impliquait un travail didactique sur chacune des séances d'ESFI sélectionnées. Les aménagements didactiques apportés, conjoints à l'action de formation, nous placent devant la difficulté de démêler les effets propres de l'évaluation formative de ceux de la modification des tâches. Un moyen de contourner cette limite aurait été d'axer la formation-recherche uniquement sur les modalités d'évaluation formative. Or, nous considérons qu'il serait contre-productif de mettre en place un dispositif de formation qui ne s'accompagne pas d'une articulation cohérente aux contenus d'enseignement (Timperley, 2011). Nous avons donc choisi d'analyser des situations d'enseignement différentes

pour ces deux dimensions (évaluation et tâche), tout en étant conscients des difficultés interprétatives qu'implique ce choix. Notre dispositif de recherche ne nous permet pas non plus d'accéder aux pratiques habituelles de l'enseignant, celles mises en place tout au long de l'année, qui sont pourtant susceptibles d'affecter directement les variables motivationnelles et affectives que nous mesurons.

Les indicateurs d'évaluation formative

Un des objectifs de cette recherche était l'élaboration d'une classification générique des pratiques d'évaluation formative. Trois des catégories d'indicateurs (S1, S2 et S5) ont trouvé des déclinaisons permettant de préciser les pratiques évaluatives des enseignants. D'abord, nous avons distingué l'explicitation des buts de celle des critères de réussite. Nous avons ainsi relevé que la majorité des enseignants explicitent fréquemment les buts de l'activité, mais que cela ne s'accompagne pas nécessairement d'une explicitation des critères de réussite et des enjeux. Cette distinction qualitative semble donc déterminante pour la compréhension de leurs pratiques évaluatives. Manifestement, tous les enseignants intègrent un outil formel d'évaluation à leur activité, mais les visées de cet outil pour l'apprentissage des élèves ne font pas toujours l'objet d'une présentation explicite. De la même façon, la distinction des différentes sources d'information recueillie par les enseignants sur l'état actuel des élèves (connaissances, avancée, compréhension, positionnement par rapport aux critères d'autoévaluation) apparaît décisive pour définir la cohérence entre les actions de l'enseignant et celles des élèves dans le processus évaluatif. Par exemple, en fonction des tâches et de l'activité des élèves pour les résoudre, l'évaluation des connaissances préalables des élèves est plus ou moins nécessaire pour réguler l'apprentissage.

Deux stratégies n'ont pas été déclinées. D'abord, la stratégie 3, relative au feedback, est plus mobilisée par les enseignants de mathématiques. Les étapes étant définies dans les séances de sciences expérimentales, nous pouvons penser que le feedback prospectif, qui porte sur ce qu'il reste à faire et comment le faire, a potentiellement moins d'opportunités d'émerger. La stratégie 4 liée à la responsabilisation n'est également pas déclinée. Au regard de l'exploitation des données, nous suggérons, a posteriori, qu'elle puisse être précisée. Nous pourrions envisager une distinction entre les actions visant à intégrer les idées des élèves (p. ex., lister les propositions au tableau), à suivre leurs initiatives (utiliser leurs propositions pour en débattre, pour mettre en place un protocole de recherche) et à favoriser leur accès aux ressources (manuels, cours, grille de critères, etc.).

En somme, la nécessité d'un grain d'opérationnalisation plus fin émerge. Des améliorations, liées à une plus grande spécification des indicateurs, pourraient être apportées afin d'accroître la qualité informative de chacune des catégories. Néanmoins, notre volonté était d'aboutir à une classification générique des pratiques d'évaluation formative en vue de nous inscrire dans une approche transversale pour comprendre les phénomènes en jeu, à la fois au sein d'une même discipline, mais aussi entre les disciplines. Cette catégorisation nous permet de tester ces pratiques dans différents contextes d'enseignement-apprentissage.

Les indicateurs d'autorégulation des apprentissages

Les indicateurs relatifs à l'évaluation de l'état actuel par les élèves, déclinés en fonction de ce sur quoi porte l'évaluation, apparaissent pertinents pour comprendre leur activité. L'apparition répétée d'un de ces indicateurs ainsi que les allers-retours entre les deux renseignent en effet sur la visée du processus d'autorégulation, mais aussi sur ses obstacles. En revanche, la qualité informative de l'indicateur « stratégies de résolution » est plus nuancée. En effet, en fonction des tâches, les stratégies de résolution ne sont pas de même nature et pourraient être appréhendées distinctement. Les stratégies intermédiaires (p. ex., colorier les diagonales pour les compter) devraient être distinguées de stratégies plus générales de résolution de la tâche (proposer une conjecture). Enfin, l'indicateur relatif aux stratégies d'apprentissage fournit des données intéressantes sur des compétences d'autorégulation plus générales, située à un niveau *méta* sur les conditions de l'apprentissage perçues par les élèves et détachées de l'activité en question.

Les indicateurs motivationnels et affectifs ont pour leur part été construits à partir de ce corpus de données et peuvent donc être enrichis en les étudiant dans d'autres disciplines et contextes et sur d'autres élèves. Nous savons que l'autorégulation implique de se fixer des buts qui sont correctement calibrés, et que leur définition dépend de variables cognitives, métacognitives et motivationnelles (Cosnefroy, 2011). Cette recherche pourrait être approfondie en analysant les facteurs qui influencent l'engagement dans le processus d'autorégulation. Nous pourrions compléter notre batterie d'observations en analysant, en amont de l'activité, des variables susceptibles d'affecter la quantité d'efforts fournis, les perceptions de compétences ou encore l'intérêt suscité par la tâche. Des mesures complémentaires réalisées par questionnaire ou entretien, avant le com-

mencement de l'activité, permettraient d'aboutir à un profil d'apprenant, que nous pourrions ensuite étudier en rapport avec les processus observés lors de l'exécution d'une tâche.

Conclusion

Cette recherche visait à analyser les effets produits par la mise en œuvre de différentes modalités d'évaluation formative sur l'autorégulation des apprentissages dans le cadre de l'ESFI. Des effets différenciés en fonction des situations d'enseignement ont été dégagés. La première situation correspondait à la mise en œuvre d'une évaluation formative dite intuitive, en référence aux pratiques quotidiennes des enseignants de sciences et de mathématiques, lesquelles ont été étudiées lors d'une première observation au sein des classes. Nos résultats ont montré un déséquilibre dans l'usage des différentes modalités qui caractérisent l'évaluation formative. Les enseignants observés ont produit peu de stratégies visant à informer l'élève des prochaines étapes du processus d'apprentissage. En retour, les processus d'autorégulation identifiés chez quelques groupes d'élèves se sont révélés peu efficaces.

La seconde situation d'enseignement correspondait à la mise en œuvre de pratiques évaluatives dites élaborées. Celles-ci ont été mises en œuvre par les mêmes enseignants l'année suivante, après un travail réflexif mené sur leur séance. Une distribution plus équilibrée des modalités d'évaluation mobilisées par les enseignants a été observée, notamment en direction d'une plus grande responsabilisation des élèves pour conduire leur activité ainsi que d'une plus forte valorisation des pairs en tant que ressource à part entière. Les élèves, ainsi placés dans ce contexte, ont passé plus de temps à produire des stratégies de résolution pour réduire l'écart entre leur état initial et le but visé. Les élèves observés tendent également à faire preuve d'une meilleure régulation motivationnelle et affective envers la tâche proposée.

Nos résultats mettent au jour un ensemble de pistes de réflexion sur la façon dont l'évaluation formative peut être combinée à l'activité de démarche d'investigation en classe de sciences. Cette recherche vise à être approfondie dans d'autres contextes d'enseignement et sur une plus large temporalité en vue d'affiner et de compléter les indicateurs de recherche élaborés.

Réception : 5 avril 2017

Version finale : 1 mai 2018

Acceptation : 24 août 2018

NOTES

1. Assess Inquiry in Science, Technology and Mathematics Education (2012-2016). Voir https://assistme.ku.dk.
2. L'ESFI désigne en France les démarches d'investigation et inclut les tâches complexes, les problèmes ouverts, les narrations de recherche, etc.
3. Voir aussi Wiliam et Thompson (2007).
4. http://ife.ens-lyon.fr/lea/le-reseau/anciens-lea/evacodice
5. Une description détaillée des séances est disponible dans Lepareur, 2016.
6. Les 6e et 5e années du collège secondaire en France équivalent à la 6e année du primaire et à la 1re année du secondaire au Québec.
7. Il s'agit la plupart du temps de quelques mots («c'est noté?», «c'est dur», «ça me stresse»).

RÉFÉRENCES

Allal, L., Bain, D. et Perrenoud, P. (1993). *Évaluation formative et didactique du français.* Lausanne : Delachaux et Niestlé.

Allal, L. et Laveault, D. (2009). Assessment for Learning : évaluation-soutien d'apprentissage. *Mesure et évaluation en éducation*, *32*(2), 99-106. doi: 10.7202/1024956ar

Allal, L. et Mottier Lopez, L. (dir.). (2007). *Régulation des apprentissages en situation scolaire et en formation.* Bruxelles : De Boeck.

Boilevin, J.-M. (2013). *Rénovation de l'enseignement des sciences physiques et formation des enseignants : regards didactiques.* Bruxelles : De Boeck.

Calmettes, B. et Matheron, Y. (2015). Les démarches d'investigation : utopie, mythe ou réalité ? *Recherches en éducation*, *21*, 3-11.

Cartier, S. et Mottier Lopez, L. (2017). *Soutien à l'apprentissage autorégulé en contexte scolaire : perspectives francophones*. Québec : Presses universitaires du Québec.

Carver, C. S., & Scheier, M. F. (1999). Themes and issues in the self-regulation of behavior. In R. S.Wyer, Jr. (Ed.), *Advances in social cognition* (vol. 12, pp. 1-105). Mahwah, NJ: Lawrence Erlbaum Associates.

Clark, I. (2012). Formative assessment: Assessment is for self-regulated learning. *Educational Psychology Review*, *24*, 205-249. doi: 10.1007/s10648-011-9191-6

Coppé, S. (Janvier, 2016). Des dispositifs d'évaluation formative en mathématiques aux apprentissages des élèves. *Actes du 28e colloque de l'ADMEE-Europe*, Lisbonne, Portugal.

Cosnefroy, L. (2011). *L'apprentissage autorégulé : entre cognition et motivation.* Grenoble : Presses universitaires de Grenoble.

Cowie, B., Moreland, J., & Otrel-Cass, K. (2013). *Expanding notions of assessment for learning : Inside science and technology primary classrooms*. Rotterdam: Sense Publishers.

Earl, L. M. (2003). *Assessment as learning: Using classroom assessment to maximize student learning*. Thousand Oaks, CA: Corwin Press.

Grangeat, M. (dir.). (2011). *Les démarches d'investigation dans l'enseignement scientifique : pratiques de classe, travail collectif enseignant, acquisitions des élèves*. Lyon : ENS.

Grangeat, M. (2013). Renouveler le questionnement sur le travail collectif enseignant dans les démarches d'investigation. Dans M. Grangeat (dir.), *Les enseignants de sciences face aux démarches d'investigation* (pp. 245-255). Grenoble : Presses universitaires de Grenoble.

Harlen, W. (2013). *Évaluation et pédagogie d'investigation dans l'enseignement scientifique : de la politique à la pratique*. Trieste (Italie) : IAP Le Réseau mondial des Académies des sciences (Science Education Program). Repéré à www.interacademies.org/File.aspx?id=22668&v=42369a72

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research, 77*(1), 81-112. doi: 10.3102/003465430298487

Lepareur, C. (2016). *L'évaluation dans les enseignements scientifiques fondés sur l'investigation : effets de différentes modalités d'évaluation formative sur l'autorégulation des apprentissages* (Thèse de doctorat, Université Grenoble Alpes, Grenoble). Repéré à https://tel.archives-ouvertes.fr/tel-01488023

Lepareur, C., Gandit, M. et Grangeat, M. (2017). Évaluation formative et démarche d'investigation en mathématiques : une étude de cas. *Éducation et didactique*, *11*(3), 101-120. doi: 10.4000/educationdidactique.2857

Lepareur, C. et Grangeat, M. (2017). L'évaluation formative : soutien à l'autorégulation des apprentissages dans les enseignements scientifiques ? Dans S. Cartier et L. Mottier Lopez (dir.), *Soutien à l'apprentissage autorégulé en contexte scolaire* (pp. 183-209). Québec : Presses universitaires du Québec.

Mottier Lopez, L. (2015). *Évaluations formative et certificative des apprentissages*. Bruxelles : De Boeck.

Nicol, D., & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. *Studies in Higher Education*, *31*(2) 199-218. Retrieved from www.reap.ac.uk/reap/public/Papers/DN_SHE_Final.pdf

Pintrich, P. (2000). The role of goal orientation in self-regulated learning. In M. Boeckaerts, P. Pintrich & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 451-502). San Diego, CA: Academic Press.

Ruiz-Primo, M. A. (2011). Informal formative assessment: The role of instructional dialogues in assessing students' learning. *Studies in Educational Evaluation*, *37*, 15-24. Retrieved from https://sde.idaho.gov/assessment/isat-cas/training/digital-library/Ruiz-Primo-Informal-formative-assessment.pdf

Sadler, D. R. (1998). Formative assessment: Revisiting the territory. *Assessment in Education*, *5*(1), 77-84. doi: 10.1080/0969595980050104

Shepard, L. A. (2009). Commentary: Evaluating the validity of formative and interim assessment. *Educational Measurement: Issues and Practice*, *28*(3), 32-37. doi: org/10.1111/j.1745-3992.2009.00152.x

Swaffield, S. (2011). Getting to the heart of authentic Assessment for Learning. *Assessment in Education: Principles, Policy & Practice*, *18*(4), 433?449. doi: 10.1080/0969594X.2011.582838

Timperley, H. (2011). Le développement professionnel des enseignants et ses effets positifs sur les apprentissages des élèves. *Revue française de pédagogie*, *174*, 31-40. doi: 10.4000/rfp.2910

Transana [Computer software]. Madison, WI: Transana.

Wiliam, D. (2010). An integrative summary of the research literature and implications for a new theory of formative assessment. In H. A. Andrade & G. J. Cizek (Eds.), *Handbook for formative assessment* (pp. 18-40). New York: Routledge.

Wiliam, D. (2011). What is assessment for learning? *Studies in Educational Evaluation*, *37*(1), 3-14. doi: 10.1016/j.stueduc.2011.03.001

Wiliam, D., & Thompson, M. (2007). Integrating assessment with instruction: What will it take to make it work? In C. A. Dwyer (Ed.), *The future of assessment: Shaping teaching and learning* (pp. 53-82). Mahwah, NJ: Lawrence Erlbaum Associates.

Annexe 1. Grille d'indicateurs de pratiques d'évaluation formative

Indicateurs d'évaluation formative

Catégories	Mots clés	Description
S1	Buts	Expliciter les buts et les sous-buts intermédiaires de l'activité
	Critères	Expliciter les critères de réussite et les compétences évaluées
S2	Avancée tâche	Prendre des indices sur l'état d'avancée des élèves dans la tâche
	Compréhension	Prendre des indices sur la compréhension de ce qui est demandé
	Connaissances	Prendre des indices sur les connaissances préalables
	Autoévaluation	Décrire les critères de réussite et s'y référer pour aider les élèves à se situer
S3	Feedback progression	Expliciter la progression des élèves dans les tâches et leurs connaissances, pointer ce qu'il reste à faire et comment le faire
S4	Responsabilisation	Intégrer les idées des élèves, suivre leurs initiatives et autoriser une certaine autonomie dans leur accès aux ressources
S5	Échanges groupe	Inciter les échanges et l'argumentation à l'intérieur des groupes, mise en avant des pairs comme ressource
	Échanges classe entière	Prendre en compte les propositions des élèves au sein de la classe et leur proposer d'interagir entre eux sur ces propositions

Annexe 2. Grille d'indicateurs d'autorégulation des apprentissages

Indicateurs d'autorégulation

Composantes	Catégories		Mots clés
Indicateurs cognitifs	Perception des buts		Buts
	Évaluation de l'état actuel		Avancée tâche Connaissances Compréhension Autoévaluation
	Stratégies		Stratégie résolution Stratégie apprentissage
Indicateurs motivationnels et affectifs	Attitudes		Inactivité Attente feedback Déconcentration Découragement Désintérêt
	Émotions		Enthousiasme Anxiété
	Perceptions	Tâche	Perception difficulté Perception facilité Enjeu/note
		Compétences	Perception négative Perception positive