

Élargir conceptuellement le modèle de l'alignement curriculaire pour comprendre la cohérence des pratiques évaluatives sommatives notées des enseignants : enjeux et perspectives

Raphaël Pasquini

Volume 42, Number 1, 2019

URI: <https://id.erudit.org/iderudit/1066598ar>

DOI: <https://doi.org/10.7202/1066598ar>

[See table of contents](#)

Publisher(s)

ADMEE-Canada - Université Laval

ISSN

0823-3993 (print)

2368-2000 (digital)

[Explore this journal](#)

Cite this article

Pasquini, R. (2019). Élargir conceptuellement le modèle de l'alignement curriculaire pour comprendre la cohérence des pratiques évaluatives sommatives notées des enseignants : enjeux et perspectives. *Mesure et évaluation en éducation*, 42(1), 63–92. <https://doi.org/10.7202/1066598ar>

Article abstract

Curriculum alignment refers to the links of coherence that exist in any teaching-learning process between curriculum objectives, learning tasks and assessment processes. This model makes it possible to understand the coherence of any assessment process. By mobilizing data from collaborative research with eight high school mathematics and French teachers, however, we will show that its meaning is limited when it comes to understanding this coherence captured in graded summative assessment practices, and that the model therefore needs to be conceptually expanded. To do this, we will use an example of a summative test modeled along these lines. Our results highlight the relevance of analyzing summative assessment practices using the expanded model, while considering the role that context plays in some of its specific dimensions.

Élargir conceptuellement le modèle de l'alignement curriculaire pour comprendre la cohérence des pratiques évaluatives sommatives notées des enseignants : enjeux et perspectives

Raphaël Pasquini

Haute école pédagogique du canton de Vaud

MOTS CLÉS : évaluation sommative, alignement curriculaire, pratiques évaluatives, notation

L'alignement curriculaire désigne les liens de cohérence existant dans tout processus d'enseignement-apprentissage entre les objectifs du curriculum, les tâches d'apprentissage et les démarches d'évaluation. Ce modèle permet notamment de comprendre la cohérence de toute démarche évaluative. En mobilisant des données issues d'une recherche collaborative menée avec huit enseignants de mathématiques et de français du secondaire, nous montrerons toutefois que son acception est limitée lorsqu'il s'agit de comprendre cette cohérence saisie dans des pratiques évaluatives sommatives notées et que, dès lors, le modèle demande à être conceptuellement élargi. Pour cela, nous nous appuyerons sur un exemple d'épreuve sommative modélisée dans ce sens. Nos résultats soulignent la pertinence d'analyser les pratiques évaluatives sommatives à l'aide du modèle élargi, tout en considérant le rôle que joue le contexte sur certaines de ses dimensions spécifiques.

KEY WORDS: summative assessment, curricular alignment, assessment practices, grading

Curriculum alignment refers to the links of coherence that exist in any teaching-learning process between curriculum objectives, learning tasks and assessment processes. This model makes it possible to understand the coherence of any assessment process. By mobilizing data from collaborative research with eight high school mathematics and French teachers, however, we will show that its meaning is limited when it comes to understanding this coherence captured in graded summative assessment practices, and

that the model therefore needs to be conceptually expanded. To do this, we will use an example of a summative test modeled along these lines. Our results highlight the relevance of analyzing summative assessment practices using the expanded model, while considering the role that context plays in some of its specific dimensions.

PALAVRAS-CHAVE: avaliação sumativa, alinhamento curricular, práticas avaliativas, notação

O alinhamento curricular designa os vínculos de coerência existentes em qualquer processo de ensino-aprendizagem entre objetivos curriculares, tarefas de aprendizagem e as abordagens de avaliação. Este modelo permite nomeadamente compreender a coerência de qualquer processo avaliativo. Ao mobilizar dados de uma investigação colaborativa realizada com oito professores de matemática e de francês do ensino secundário, mostraremos, porém, que o seu significado é limitado quando se trata de compreender essa coerência identificada nas práticas avaliativas sumativas e que, portanto, o modelo precisa ser conceitualmente ampliado. Neste sentido, apoiar-nos-emos num exemplo de um teste sumativo concebido em função destes pressupostos. Os nossos resultados destacam a pertinência de analisar práticas avaliativas sumativas usando o modelo alargado, considerando o papel que o contexto desempenha em algumas de suas dimensões específicas.

Problématique

Un nombre conséquent de recherches ont montré que l'évaluation formative est une démarche d'évaluation indispensable au soutien à l'apprentissage des élèves (p. ex., Allal et Mottier Lopez, 2005 ; Black et Wiliam, 1998 ; Mottier Lopez, 2015). L'évaluation sommative, quant à elle, est définie comme une démarche « qui a pour objet un inventaire. Son usage social est de vérifier (probation), sa fonction sociale de certifier. Elle a comme fonctions annexes classer, situer ou encore informer » (Hadji, 1989, cité par Mottier Lopez, 2015, p. 23). Malgré le fait qu'elle soit porteuse d'enjeux pédagogiques forts pour les élèves et que des travaux la pointent du doigt dans les questions d'échec scolaire (Antibi, 2003 ; Crahay, 2007 ; Guskey, 2011), elle fait paradoxalement l'objet de moins d'attention dans la recherche francophone, d'où la nécessité de lui accorder de l'importance.

En effet, que sait-on des pratiques d'évaluation sommative des enseignants ? Quels phénomènes mettent-elles en évidence ? Quels problèmes y observe-t-on ? Quelles sont leurs spécificités ? Des travaux dans le champ de l'évaluation abordent ces questions au travers d'enquêtes. Leurs résultats sont interpellants. Il en ressort notamment, toutes disciplines et années d'études confondues, que la cohérence des pratiques peine à être caractérisée, car il existe souvent de forts écarts entre les représentations que se font les enseignants des pratiques d'évaluation sommative cohérentes et la description de leurs propres pratiques (Braxmeyer, Guillaume et Levy, 2004 ; McKinney, Chappell, Berry et Hickmann, 2009 ; Rieg, 2007).

Peu d'études à ce jour ont abordé cette question du point de vue des pratiques effectives. Dans la communauté scientifique anglo-saxonne, Moss (2013) affirme : « Il existe un besoin pour des modèles de recherche qui vont au-delà de l'autoévaluation de l'enseignant, des études d'ensemble et des inventaires » (p. 252, trad. libre). C'est, de son point de vue, une condition pour aborder conceptuellement et pragmatiquement cette question de cohérence des pratiques. Un constat similaire est réalisé dans la communauté francophone (CNESCO, 2014 ; Sayac, 2017).

Malgré ces manques, certaines recherches ancrées dans le terrain mettent tout de même en évidence des résultats intéressants (p. ex., Bate-man, Taylor, Janik et Logan, 2009 ; Harlen, 2005, 2012 ; Royce Sadler, 2009). Elles relèvent par exemple la complexité à penser cette cohérence des pratiques évaluatives sommatives, et ce, même si les enseignants, en équipe et en amont, s'accordent sur les objectifs d'apprentissage à poursuivre et à évaluer ou encore sur le déroulement des séquences à mener avec les élèves. La cohérence est donc difficile à cerner et, *a fortiori*, à conceptualiser, tout en étant centrale (Gagné, Dumont, Brunet et Boucher, 2013 ; Martone et Sireci, 2009). Mottier Lopez et Laveault (2008) confirment cette dernière idée : « Cette recherche de cohérence et d'articulation des trois domaines – enseignement, apprentissage, évaluation – est tout à fait caractéristique des évolutions les plus récentes de l'évaluation en éducation » (p. 9).

Cet article s'empare de cette problématique. Il vise à examiner dans quelle mesure un modèle théorique peut permettre de mieux comprendre la cohérence des pratiques d'évaluation sommative notée et, dialectiquement, à quelles conditions ces pratiques peuvent interagir avec un modèle pour gagner en cohérence. Dans notre présente recherche¹, il nous apparaît que le modèle de l'alignement curriculaire est pertinent pour poursuivre ce but.

Dans le cadre conceptuel, nous exposons tout d'abord ce modèle en mettant en évidence ses caractéristiques principales et quelques-unes de ses limites théoriques pour penser les pratiques évaluatives sommatives notées. Dans un deuxième temps, nous présentons certains éléments méthodologiques de notre recherche, notamment en ce qui concerne notre méthode collaborative. Puis, en partant d'une épreuve sommative réalisée par un enseignant, nous documentons, sous forme d'étude de cas, comment ses pratiques ont évolué du point de vue de leur cohérence et, parallèlement, en quoi un élargissement conceptuel du modèle est devenu nécessaire pour appréhender cette cohérence dans l'épreuve. Nous poursuivons en exposant nos résultats principaux, puis présentons quelques éléments de discussion encore en suspens relativement à notre question de recherche, avant de conclure à l'aide de quelques pistes de réflexion.

Cadre conceptuel

L'alignement curriculaire pour modéliser la cohérence en évaluation sommative

La littérature qui traite du modèle de l'alignement curriculaire postule que toute pratique d'évaluation doit être en forte cohérence avec les objectifs prescrits, l'enseignement offert, les tâches d'apprentissage proposées aux élèves et les activités effectivement réalisées par ces derniers (p. ex., Anderson, 2002; Biggs, 1999, 2003; Gauthier, Mellouki, Bissonnette et Richard, 2005). L'enjeu pour l'évaluation est ici de porter sur les apprentissages prévus et effectifs. Dans cette perspective, l'élaboration d'épreuves évaluatives sommatives est conçue comme un processus complexe, dynamique, contextuel et centré sur les apprentissages. La notion d'*alignement* relève d'une conceptualisation de la cohérence à l'aune d'une taxonomie d'habiletés cognitives appariées à des contenus, sur laquelle nous revenons plus loin. La dimension *curriculaire* renvoie au fait que cette cohérence est pensée depuis un curriculum.

En 2005, Gauthier, Mellouki, Bissonnette et Richard publient un état de la recherche sur les écoles efficaces et la réussite scolaire des élèves à risque en Amérique du Nord. Ils montrent que «l'alignement curriculaire apparaît grandement susceptible d'améliorer la qualité de l'enseignement et l'efficacité des écoles» (p. 28), même si sa mise en œuvre ne garantit pas à elle seule de meilleurs apprentissages. En reprenant notamment Guskey (2003), Gauthier et ses collègues définissent l'alignement curriculaire comme l'élément «visant à assurer une correspondance élevée entre le curriculum, l'enseignement et l'évaluation» (p. 24). Ils précisent par ailleurs que le modèle concerne aussi «les opérations intellectuelles et les différentes catégories de connaissances qui y sont reliées» (p. 25). Ces chercheurs décrivent trois étapes pour sa mise en œuvre.

Étape 1

La première étape consiste à définir ce qui doit être maîtrisé par les élèves en matière de performances : l'*output*. La question clé est : «Comment peut-on observer et juger le degré de maîtrise de telle connaissance ou compétence en termes de comportements observables ou de manifestations?» (Gauthier et al., 2005, p. 24). La formulation des objectifs d'apprentissage et des critères d'évaluation est ici centrale. Il s'agit de concrétiser le curriculum afin de faciliter l'enseignement et l'évaluation de son contenu.

Étape 2

La deuxième étape revient à planifier les séquences d'enseignement et d'apprentissage en rapport avec les objectifs du curriculum retenus, de manière à amener tous les élèves à progresser. Elle se concrétise dans la question suivante : « Comment permettre aux élèves d'apprendre ce qui doit être maîtrisé ? » (Gauthier et al., 2005, p. 25).

Étape 3

Enfin, la troisième étape vise à évaluer les apprentissages effectués par les élèves. Elle porte sur ce « qui fait objet d'apprentissage en salle de classe » (Gauthier et al., 2005, p. 25). Il est alors important que les tâches évaluatives soient reconnaissables et congruentes avec ce qui a été enseigné et avec le curriculum.

Relativement à notre problématique, cette opérationnalisation du modèle est remise en question, selon nous, sur trois plans. Tout d'abord, nous faisons l'hypothèse que son acception étapiste ne rend pas compte de la complexité des pratiques effectives lorsque l'enseignant tente de mettre en lien le curriculum, les tâches d'apprentissage et les démarches d'évaluation. En effet, nous savons que, lorsque les enseignants élaborent leurs démarches d'évaluation sommative, le curriculum est pour eux une référence parmi d'autres (Mottier Lopez, Tessaro, Dechamboux et Morales Villabona, 2012). Dès lors, ces pratiques ne peuvent être pensées sans un travail préalable sur les contenus du curriculum lorsque ce dernier en est la référence.

La deuxième interrogation concerne la notion de congruence et le rapport qu'elle entretient avec celle de cohérence, qui est principalement théorisée dans la littérature sur le modèle (p. ex., Anderson, 2002 ; Hamerness, 2006).

Enfin, la troisième source de questionnement – celle que nous abordons principalement dans cet article – relève les limites de la théorisation de l'évaluation sommative au sein même du modèle. En effet, le raisonnement de Gauthier et de ses collègues s'arrête sur le plan des tâches évaluatives. Il n'aborde pas la manière dont le poids est alloué aux apprentissages évalués au travers de points ou de critères, ni les questions de construction de la note. Cela signifie que les deux pratiques clés que sont la pondération et la notation ne sont pas suffisamment conceptualisées, alors qu'elles s'articulent (Anderson, 2002 ; Biggs, 2003) et sont indisso-

ciables de l'élaboration de l'évaluation sommative en référence à un curriculum (Marzano, 2002; Brookhart, 2005). Ce manque est d'autant plus regrettable, car des recherches ont permis de montrer que les pratiques de notation et les pratiques d'évaluation sommative se différencient et se pensent spécifiquement. Modéliser les pratiques de notation est dès lors très complexe (McMillan et Nash, 2000; Randall et Engelhard, 2010; Walvoord et Johnson Anderson, 2009). En effet, ces travaux documentent notamment que certains facteurs exercent des influences majeures sur les pratiques de notation (p. ex., les normes prescrites au travers des outils comme les échelles, la question des seuils de suffisance, la prise en compte de l'effort fourni par l'élève ainsi que des performances supplémentaires démontrées, la distribution des notes ou encore la question du zéro comme signe d'absence totale d'apprentissage). En d'autres termes, un enseignant peut être cohérent pour construire des tâches évaluatives ou des critères, tout en étant très incohérent lorsqu'il note. Il s'agit donc d'en tenir compte pour penser la cohérence de l'évaluation sommative dans sa globalité.

Des incomplétudes du même ordre s'observent dans la théorisation du modèle selon Anderson (2002), qui fait référence. L'alignement curriculaire y est défini comme : «le niveau de cohérence entre les objectifs du curriculum et l'évaluation des apprentissages, entre les objectifs du curriculum et les tâches d'apprentissage, et, enfin, entre l'évaluation des apprentissages et les tâches d'apprentissage» (p. 257, trad. libre). La validité du contenu, la façon dont il est enseigné ainsi que les occasions d'apprentissage de ce contenu offertes aux élèves au travers de tâches d'apprentissage font également partie du modèle. L'auteur schématise l'alignement curriculaire sous forme de triangle dont chaque sommet en est une composante (voir Figure 1).

Les limites théoriques relatives à l'évaluation sommative s'observent dans la dimension générique de l'évaluation des apprentissages, puis plus distinctement dans les questions formulées par le chercheur pour appréhender le modèle. Par exemple, pour la relation A, qui explicite le lien de cohérence entre l'évaluation des apprentissages et les objectifs d'apprentissage issus du curriculum, Anderson propose : « Dans quelle mesure l'épreuve mesure-telle les objectifs importants issus du curriculum ? » (p. 255, trad. libre), sans développer cette question de mesure. Quant à la relation C, qui met en évidence la cohérence qui devrait exister entre l'évaluation et les tâches d'apprentissage, l'auteur suggère : « Est-ce que ce qui

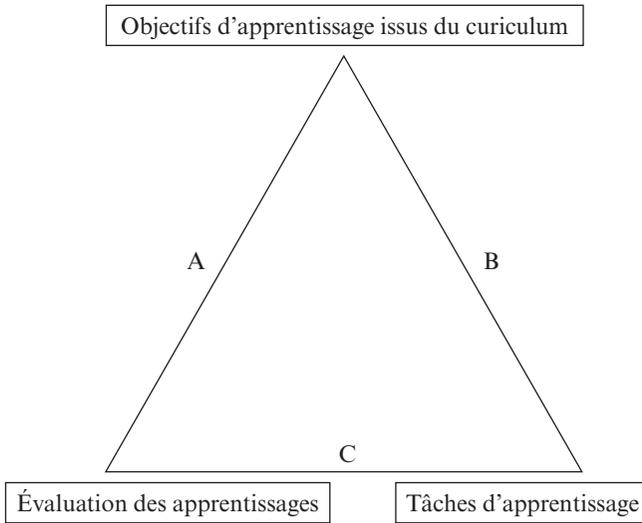


Figure 1. L'alignement curriculaire selon Anderson (2002, p. 256)

est enseigné est évalué?» (p. 255, trad. libre), en omettant de développer les modalités fines de l'élaboration de l'épreuve sommative. Nous pouvons observer que les questions centrales de pondération et de notation ne sont effectivement pas abordées.

La question de la cohérence curriculaire

Ces limites demandent d'approfondir cette notion de cohérence au cœur du modèle. Les chercheurs consultés parlent d'une cohérence d'alignement curriculaire construite à l'aide d'une taxonomie (Anderson, 2002) ou de « *constructive alignment* » (Biggs, 2003).

Dans le cadre de notre recherche, nous avons travaillé avec la taxonomie d'Anderson et de Krathwohl (2001), car elle est majoritairement exploitée dans les travaux recensés. Elle conceptualise l'apprentissage en tant que capacité à maîtriser des objectifs formulés en habiletés cognitives appariées à des contenus. Nous parlons donc bien d'objectifs, et non de compétences. Dans cette perspective, deux éléments clés émergent : la complexité de l'habileté cognitive en jeu et, conjointement, les caractéristiques de l'objet disciplinaire sur lesquelles elle porte (McGrath et Noble, 2008). L'objectif d'apprentissage est donc défini comme une capacité comprenant une dimension générale et spécifique.

La première dimension renvoie à l'activité mentale de l'élève au travers des six niveaux d'habiletés cognitives de la taxonomie (se souvenir, comprendre, appliquer, analyser, évaluer ou créer). La seconde se rapporte au contenu sur lequel porte l'apprentissage en question : la relative en français, les isométries en mathématiques, etc. La complexité de l'objectif dépend donc de ces deux dimensions (p. ex., mémoriser un vocabulaire mathématique relatif à la trigonométrie), qui devraient, idéalement, permettre d'établir les relations de cohérence entre chaque composante du modèle.

Une cohérence jusqu'à la notation

En évaluation sommative, cette conceptualisation de la cohérence curriculaire reste pertinente pour comprendre les relations qui devraient exister entre toutes les composantes du modèle jusqu'à la pondération et la notation (Airasian et Miranda, 2002). En effet, considérant que la pondération est le fait de signifier l'importance d'un apprentissage – son poids – au travers d'outils comme les critères ou les points (Marzano, 2002) et que la notation en lien avec l'apprentissage est « un processus complexe dépendant du contexte, qui joue de multiples rôles » (Walvoord et Johnson Anderson, 2009, p. 2, trad. libre), nous analyserons comment cette cohérence taxonomique peut se traduire conceptuellement et dans les pratiques effectives jusqu'à la phase d'attribution des points ou l'élaboration des critères par l'enseignant, mais également lorsque ce dernier établit les seuils des notes (dans notre cas, de 1 à 6, avec demi-notes, la note 4 signifiant le seuil de suffisance).

Un premier pas vers un modèle de l'alignement curriculaire élargi

Biggs (1999, 2003) est à notre connaissance un des chercheurs à donner de l'importance à une forme de conceptualisation de l'évaluation sommative dans le modèle de l'alignement curriculaire incluant la pondération et la notation. Concernant les relations que l'évaluation sommative entretient avec, d'une part, les objectifs du curriculum et, d'autre part, les tâches d'apprentissage, ses travaux insistent sur trois caractéristiques : 1) l'évaluation doit porter sur des objectifs complexes et donner à voir la qualité de l'apprentissage ; 2) l'utilisation de critères en rapport avec les contenus et les objectifs est centrale et présentée comme « la clé pour une notation facile et réussie » (p. 6, trad. libre) ; 3) la notation doit rendre compte de l'apprentissage par le biais des critères et des tâches.

Ces éléments constituent un premier pas significatif dans l'élargissement conceptuel du modèle, et c'est sur leurs bases que nous avons abordé notre travail avec les enseignants impliqués dans notre recherche. Toutefois, deux limites perdurent dans ces travaux et viennent encore remettre en question le modèle.

La première consiste à ne pas considérer les spécificités des disciplines, alors que les objets de savoir ont une influence sur les pratiques d'évaluation (p. ex., Meier, Rich et Cady, 2006; Sayac, 2017). La seconde est de ne se référer qu'à une seule échelle de notation (à six positions, avec des notes de A à F et une seule note insuffisante – le F), alors qu'il est avéré que cet outil infléchit aussi fortement les pratiques évaluatives (Hadji, 2017). Il s'agit donc d'envisager un modèle compréhensif transposable à d'autres environnements scolaires et disciplinaires, et adapté à diverses échelles de notation.

Ainsi, d'un côté, la pratique semble avoir besoin d'un modèle théorique pour conceptualiser la cohérence en évaluation sommative. De l'autre, les travaux sur la pondération (p. ex., Jonsson, 2014; Marzano, 2002; Royce Sadler, 2009) et sur la notation (p. ex., Brookhart, 2017; Walvoord et Johnson Anderson, 2009; Winger, 2009) demandent à approfondir la conceptualisation de la cohérence en évaluation sommative en fonction du contexte, des curricula et des outils de notation.

Méthodologie

La question à laquelle nous tentons de répondre est la suivante : Du point de vue des pratiques effectives et des apports de la recherche, comment élargir conceptuellement le modèle de l'alignement curriculaire pour comprendre la cohérence en évaluation sommative notée ?

Pour cela, nous commençons par présenter notre dispositif, puis notre méthode de recherche collaborative. Nous poursuivons en explicitant nos données et les méthodes d'analyse choisies.

Les modalités et la temporalité du dispositif de recherche

Nous avons travaillé avec quatre enseignants de mathématiques et quatre de français du secondaire, tous expérimentés et volontaires, dans le canton de Vaud, en Suisse romande. Nous avons élaboré notre dispositif de façon à ce qu'il nous permette d'accéder à des épreuves somma-

tives qu'ils ont construites. Trois autres types de données ont été récoltés : 1) les enregistrements des échanges lors des discussions de ces épreuves pendant les rencontres entre les enseignants et le chercheur ; 2) dans l'idée de comprendre leurs pratiques évaluatives en profondeur, deux entretiens semi-dirigés avant et après le dispositif ; 3) trois écrits exemplaires de pratiques (Desgagné, 2005 ; Pasquini, 2013, 2016), dont un en amont, un en cours et un en fin de processus. La nature et l'agencement de ces différents temps ainsi que les moments de collecte de données sont présentés dans le tableau 1.

Une recherche collaborative

En partant des questions et problèmes évoqués par les enseignants, nous avons conçu un dispositif de recherche leur offrant la possibilité de s'appropriier des connaissances théoriques pour analyser, concevoir et discuter leurs évaluations sommatives. Pour cela, nous leur avons présenté le modèle de l'alignement curriculaire élargi lors de la première séance de travail (Temps 3), dans une version simplifiée élaborée uniquement avec les apports de la recherche, comme nous l'expliquons dans le cadre conceptuel. Nous avons privilégié une méthode de recherche collaborative, définie par Desgagné (1997) de la manière suivante : « À sa base, nous dirons qu'une recherche collaborative suppose la contribution des praticiens enseignants à la démarche d'investigation d'un objet de recherche, démarche le plus souvent encadrée par un ou plus d'un chercheur universitaire » (p. 372). Van Nieuwenhoven et Colognesi (2015) précisent les deux buts incontournables de cette méthode : « faire émerger une production de connaissances et viser un développement professionnel de chaque catégorie d'acteurs concernés par le processus de recherche » (p. 105). La figure 2 rend compte du schéma de notre recherche collaborative fidèlement à ces définitions et aux travaux de Desgagné, Bednarz, Lebus, Poirier et Couverture (2001). Nous en détaillons le contenu plus bas.

Une étape de négociation entre les enseignants et le chercheur a participé à circonscrire un objet de formation : travailler la cohérence des pratiques évaluatives sommatives notées des participants. L'objet de recherche s'est en conséquence imposé. L'enjeu a été de récolter des données pour documenter et comprendre comment la cohérence des pratiques évaluatives sommatives réagissait à la confrontation avec le modèle théorique de l'alignement curriculaire élargi à l'aide des apports de la recherche –

Tableau 1
Dispositif de recherche

Temps	0	1	2	3	4	5	6	7	8
Séances	Constitution du groupe Recensement des questions	Entretien individuel 1	Récit de pratique individuel 1	Rencontre 1 <i>Commune</i>	Récit de pratique individuel 2	Rencontre 2 <i>Différenciée mathématiques et français</i>	Rencontre 3 <i>Commune</i>	Récit de pratique individuel 3	Entretien individuel 2
Données récoltées		Entretien semi-dirigé	Écrit sur les pratiques	Première évaluation sommative créée par les enseignants Présentation du modèle théorique	Écrit sur les pratiques	Discussion des évaluations sommatives réalisées Seconde évaluation sommative créée par les enseignants	Discussion des évaluations sommatives réalisées	Écrit sur les pratiques	Entretien semi-dirigé

Note. 1 = mathématiques ; 2 = français.

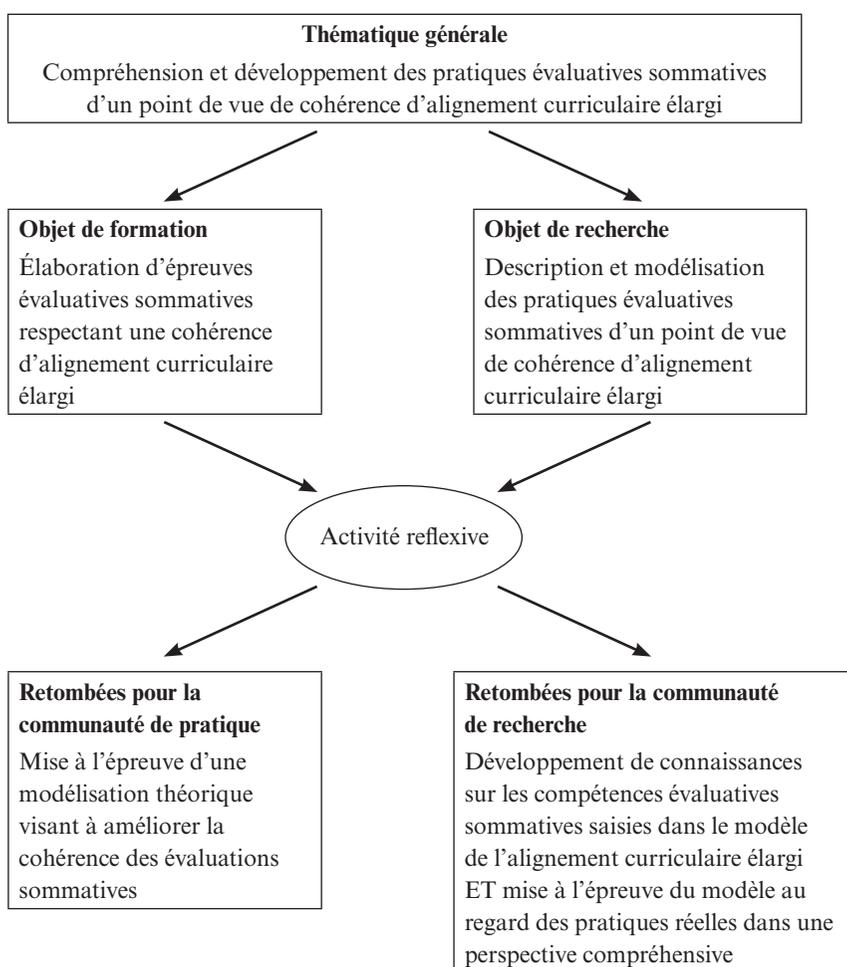


Figure 2. Modélisation de notre recherche collaborative selon Desgagné et ses collègues (2001)

notamment au regard des caractéristiques disciplinaires contrastées (français et mathématiques) et des contextes des enseignants (types d'élèves, prescriptions locales, logiques d'équipes) –, tout en observant comment ce même modèle réagissait à ces pratiques en contexte.

Dans ce sens, l'activité réflexive a traversé tout notre dispositif. En effet, dès le départ, les enseignants ont été amenés à thématiser, au travers d'un entretien et d'un écrit, les problèmes et/ou questions qu'ils souhaitent traiter. Il en a été de même dans l'analyse des évaluations en vue de

les modéliser ; dans l'élaboration d'évaluations en référence à l'enseignement et à l'aune de l'alignement curriculaire élargi ; dans la confrontation des évaluations avec les éléments de discussion proposés par le chercheur ou les pairs ; ou encore dans les écrits faisant référence à la pratique évaluative.

Ainsi élaboré, notre dispositif de recherche collaborative a notamment tenté de mettre en évidence le rôle de l'alignement curriculaire élargi sur le développement de pratiques évaluatives sommatives cohérentes référées à des contextes de classe (Balach et Szymanski, 2003) et de soumettre un modèle théorique à la pratique.

Les données considérées

Dans cet article, nous avons prioritairement pris en compte les épreuves sommatives réalisées par les enseignants et nous nous appuyons sur l'une d'entre elles pour illustrer notre propos. D'un point de vue méthodologique, nous avons respecté la manière avec laquelle les enseignants les ont créées (travail individuel, en duo fidèlement à leur collaboration dans leur établissement, choix des objets et objectifs à évaluer), de manière à ce qu'elles soient en lien avec les apprentissages des élèves. Toutefois, pour étayer notre propos relativement aux phénomènes de cohérence dans l'évaluation sommative, nous aurons également recours à des extraits issus d'écrits et d'entretiens.

Un aller-retour entre pratiques évaluatives sommatives et modèle théorique

Notre méthodologie a suivi d'abord un mouvement inductif, mettant au centre les « questions vives de la pratique évaluative » (Pasquini, 2017) des enseignants récoltées au début du processus lors du Temps 0. Puis, à dessein et assez rapidement, nous avons présenté le modèle théorique pour voir dans quelle mesure il permettait aux enseignants de comprendre les phénomènes d'incohérence dans leurs pratiques, et ainsi de trouver des pistes d'action à leurs questions et problèmes. C'est alors un mouvement déductif qui a pris le relais puisque c'est à l'aune de ce modèle que les données ont été interprétées à l'aide de catégories conceptualisantes² (Paillé et Mucchielli, 2012), dans une démarche de triangulation méthodologique (Denzin, 1978 ; Silverman, 2009). Le recours à un modèle – bien que non suffisant puisque tout modèle théorique doit être éprouvé au regard de données reflétant des éléments de la pratique – s'est avéré nécessaire dans

notre démarche qualitative (Anadon et Guillemette, 2007). En effet, le but était d'ouvrir des possibles sur une potentielle reconfiguration du modèle théorique de départ, tout comme d'être flexible sur les phénomènes de cohérence qu'éclaireraient nos données. Dès lors, l'un pouvait influencer l'autre, et vice versa.

À la lumière de ces éléments, exploitons maintenant une épreuve réalisée par un enseignant engagé dans notre recherche, sous forme d'étude de cas. Nous y pointons un certain nombre de dimensions de la pratique évaluative dont la cohérence a évolué et qui, dialectiquement, ont influencé conceptuellement notre modèle élargi.

Étude de cas

L'exemple présenté est une épreuve de mathématiques. Nous l'avons choisie parmi une douzaine d'autres produites dans notre recherche, car elle donne à voir des tentatives d'optimisation de cohérence d'alignement curriculaire élargi qui mettent en jeu toutes les composantes du modèle, jusqu'à la notation. Réalisée par Paul³ lors de la deuxième rencontre (Temps 5, entre spécialistes de mathématiques) et discutée en groupe pendant la dernière rencontre (Temps 6), cette évaluation porte sur les fonctions et l'algèbre, plus spécifiquement sur le calcul littéral. Constituée de neuf exercices, elle se destine à des élèves de 10^e année (l'avant-dernière de la scolarité obligatoire, moyenne d'âge de 15 ans), de Voie générale (la moins exigeante du système sur les deux existantes), de niveau 2 au sein de cette voie (niveau d'exigences supérieures, alors que le niveau 1 concerne les exigences de base). La figure 3 présente son contenu. L'objectif du Plan d'études romand⁴ évalué est « Résoudre des problèmes numériques et algébriques ». Paul sélectionne quatre de ses composantes, puis six éléments de progression des apprentissages de l'objectif (P1 à P6) mis en correspondance avec les neuf exercices de l'épreuve.

Sur un autre document, Paul explicite ses pratiques de pondération et de notation. Une première étape a consisté pour lui à mettre en rapport chaque exercice (Ex) avec la ou les progressions des apprentissages (P1 à P6) qui correspondent (voir Tableau 2).

Nous observons que Paul décide d'évaluer la P1 dans chaque exercice. En revanche, chacune des autres progressions est évaluée séparément dans les différents exercices (p. ex., la P4 est évaluée dans les exercices 4 et 5).

Test significatif 2

Objectif du plan d'étude évalué: MSN 33 – Résoudre des problèmes numériques et algébriques

...en reconnaissant les caractéristiques mathématiques d'une situation et en la traduisant en écriture numérique

...en utilisant des propriétés des opérations

...en mobilisant l'algèbre comme outil de calcul

...en construisant, en exerçant et en utilisant des procédures de calcul.

Progressions des apprentissages évalués:

P1 – Connaissance et utilisation des règles et conventions usuelles d'écriture algébrique (dans tout le test et exercice 1)

P2 – Détermination de la valeur numérique d'une expression littérale en substituant des nombres aux variables (exercices 8 et 9)

P3 – Élaboration d'expressions littérales à partir de problèmes, figures géométriques ou expressions verbales (exercices 7, 8, 9)

P4 – Interprétation d'expressions littérales et identification de celles qui sont équivalentes (exercices 4 et 5)

P5 – Connaissance de la terminologie, écriture réduite et ordonnée de monômes et polynômes (exercices 1, 2, 3, 6, 7)

P6 – Opérations sur les monômes et polynômes (+/-/x) (exercices 2, 3, 6, 7).

Figure 3. Objectifs évalués de la seconde épreuve de Paul en calcul littéral

À la fin, Paul accorde à chaque progression la même pondération, indépendamment du nombre de fois qu'elle est évaluée dans les exercices, au travers d'une appréciation pondérée en points (où insuffisant = 1 ; suffisant = 2 ; complètement acquis = 3). Cette lecture horizontale demande donc

Tableau 2

Mise en lien entre exercices et progressions des apprentissages

	Ex 1	Ex 2	Ex 3	Ex 4	Ex 5	Ex 6	Ex 7	Ex 8	Ex 9	Total 1/2/3
P1	√	√	√	√	√	√	√	√	√	
P2								√	√	
P3							√	√	√	
P4				√	√					
P5	√	√	√			√	√			
P6		√	√			√	√			

à Paul d'observer jusqu'à neuf reprises si une progression est insuffisante, suffisante, voire complètement acquise, puis de juger quelle appréciation finale il attribue à chacune d'entre elles. Les figures 4 et 5 exemplifient sa manière de définir un niveau de suffisance attendu relativement à chaque progression au travers des exercices. À titre de comparaison, Paul a également mentionné le nombre de points qu'il donnerait à chaque exercice, comme il le fait d'habitude.

Exercice 2 (pts)**2.5 pts**

Réduis ces expressions littérales :

a) $4 \cdot y \cdot y = 4y^2$

b) $(-6) \cdot (x \cdot y) = -6xy$

c) $(5y)^2 = 25y^2$

d) $(3n \cdot 6) \cdot 2 = 36n$

e) $-4z^2 \cdot z^3 \cdot z = -4z^6$

P6 Seuil de suffisance : 3 justes**P5** Seuil de suffisance : 0 faute d'écriture

Figure 4. Seconde épreuve de Paul en calcul littéral avec pose des seuils pour les P1, P5 et P6 dans l'exercice 2

Exercice 4 (pts)**2 pts**

Relie les expressions équivalentes :

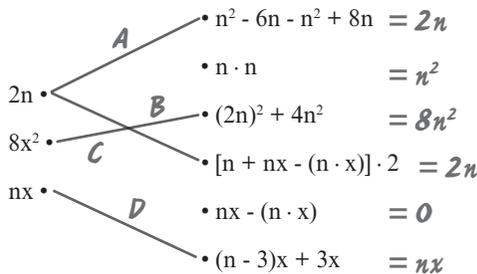
**P4** Seuil de suffisance : identifier 2 des 4 liens. Max 2 liens faux.

Figure 5. Seconde épreuve de Paul en calcul littéral avec pose des seuils pour la P4 dans l'exercice 4

Dans ces nouvelles pratiques, il est dès lors possible pour l'élève d'obtenir 18 points (6P x 3) au maximum dans l'évaluation, ce qui représente la note maximale de 6 (voir Tableau 3). Si toutes les progressions sont suffisantes, il obtient 12 points (6P x 2), ce qui correspond à une note de 4. En dessous, l'élève est en échec. Sur cette base, Paul crée l'échelle suivante :

Tableau 3
Échelle de notation de Paul

Note	1	1,5	2	2,5	3	3,5	4	4,5	5	5,5	6
Points		6	7	8	9	10	12	13	15	16	18

Résultats

Fidèlement à notre question de recherche et à notre méthodologie, nous exposons nos résultats en deux temps : ceux qui concernent les changements dans la cohérence des pratiques évaluatives sommatives de Paul, puis ceux qui concernent l'évolution du modèle théorique.

Les résultats sur la cohérence des pratiques

L'analyse de nos diverses données montre que les pratiques évaluatives sommatives de Paul convergent vers plus de cohérence curriculaire : le rapport entre les exercices de complexité homogène et les objectifs est lisible du point de vue des objets de savoir et des habiletés cognitives ; les exercices sont pensés en référence aux apprentissages menés en classe ; les liens entre les progressions des apprentissages considérées comme des critères et les exercices sont explicites.

Les changements en matière de pratiques de pondération et de notation, bien qu'également présents, méritent quant à eux d'être détaillés, car ils contrastent avec les pratiques traditionnelles. En effet, alors qu'au début de la recherche, Paul reste vague sur ces points, il réalise, à l'aide du modèle élargi, un travail approfondi autour d'une recherche de cohérence systémique entre les objectifs, les tâches ainsi que sur des outils de pondération centrés sur les apprentissages, donnant à voir ce que l'élève réussit ou non. Dans cette épreuve, nous observons qu'au lieu de compter des points alloués à des exercices, l'enseignant juge de l'atteinte de progressions d'apprentissage du Plan d'études romand au travers des exercices, ce qui représente un changement de perspective important. Toutefois, nous remarquons que de fortes tensions sont en jeu chez lui :

«Signaler quand ce n'est pas réussi, et puis qu'est-ce qui n'est pas réussi. C'est ce que je pense, ce que j'aimerais faire. Alors... dans ma dernière évaluation, j'avais travaillé sur... plus des critères de réussite, avec des appréciations, et puis je trouvais ça pas mal. Mais je ne voyais pas comment passer de ces appréciations à la note. Enfin, je voyais difficilement. On peut toujours trouver des artifices, des trucs, mais ça revient toujours à choisir une échelle, à choisir un 4, un seuil de réussite, et puis après, bien voilà, de faire ça proportionnellement ou pas... c'est un peu toujours la même chose.»

C'est donc entre les critères – ici, pour Paul, les progressions des apprentissages – et la notation que perdurent des obstacles à une cohérence curriculaire. Comme le dit l'intéressé, il assume des compromis en utilisant finalement des barèmes standardisés : «C'est ma base de travail. Je ne pense pas que c'est bien, mais il faut bien que je travaille avec quelque chose.» À ce sujet, Paul revient donc à des pratiques traditionnelles. Conscient que cette procédure entraîne une rupture d'alignement curriculaire élargi puisque le barème n'est pas pensé en lien avec les apprentissages évalués, il s'en accommode tout de même, tant une solution alternative lui paraît difficile à mettre en œuvre. D'un côté, Paul est donc en mesure d'explicitier le chemin parcouru vers une notation plus en lien avec l'apprentissage : «Concernant l'échelle, [...] j'ai imaginé évaluer les élèves en prenant les progressions comme des critères, les unes après les autres et sur un pied d'égalité.» Mais, d'un autre côté, sa crainte d'une trop grande subjectivité semble prendre le dessus :

«Mais comment justifier que le choix de tel ou tel critère est plus équitable qu'un autre ? Sur le fond, je suis convaincu par ce type d'évaluation, mais comment transformer nos critères en note, de la façon la plus *juste* possible ? Et le problème reste le même : quelle échelle choisir ?»

Une forme d'insécurité domine. Nous y reviendrons dans notre discussion.

Il importe maintenant de se demander comment conjuguer ces observations de la pratique avec les apports de la recherche en évaluation sommative, de manière à conceptualiser un modèle permettant de mieux appréhender cette question de cohérence en évaluation jusqu'à la notation.

Les résultats sur l'évolution du modèle

Pour réaliser ce travail de conceptualisation, nous nous sommes appuyé sur les phénomènes relevés dans les pratiques des participants, comme dans notre exemple, ainsi que sur deux caractéristiques du modèle de départ d'Anderson (2002) : 1) une perspective dynamique pour penser les relations entre ses différentes composantes (objectifs du curriculum,

tâches d'apprentissage, évaluation); 2) l'inscription première du modèle dans les processus d'enseignement-apprentissage. Les rapports de cohérence taxonomique qu'entretiennent idéalement les différentes composantes traduisent donc les résultats de recherche et les pratiques effectives.

Dans cette perspective, nous avons été attentif aux liens de cohérence entre les critères d'une épreuve et les objectifs évalués (ou annoncés comme tels), à leur rapport respectif avec les tâches évaluatives et avec les outils de notation. Dans le même sens, nous avons observé dans quelle mesure les objets de savoir et les habiletés cognitives décrivant les objectifs d'apprentissage évalués caractérisent les seuils de réussite dans l'échelle de notation construite, par le biais des tâches évaluatives et/ou des critères d'évaluation qui devraient, idéalement, les rendre explicites.

Ainsi pensé, le modèle élargi devrait permettre de comprendre les phénomènes de cohérence ou d'incohérence (ou de désalignement curriculaire) présents dans les pratiques évaluatives sommatives et perceptibles dans les épreuves, puis de renvoyer à un raisonnement saisi dans une perspective systémique et contextuelle. Aussi, nous l'avons schématisé de la manière suivante :

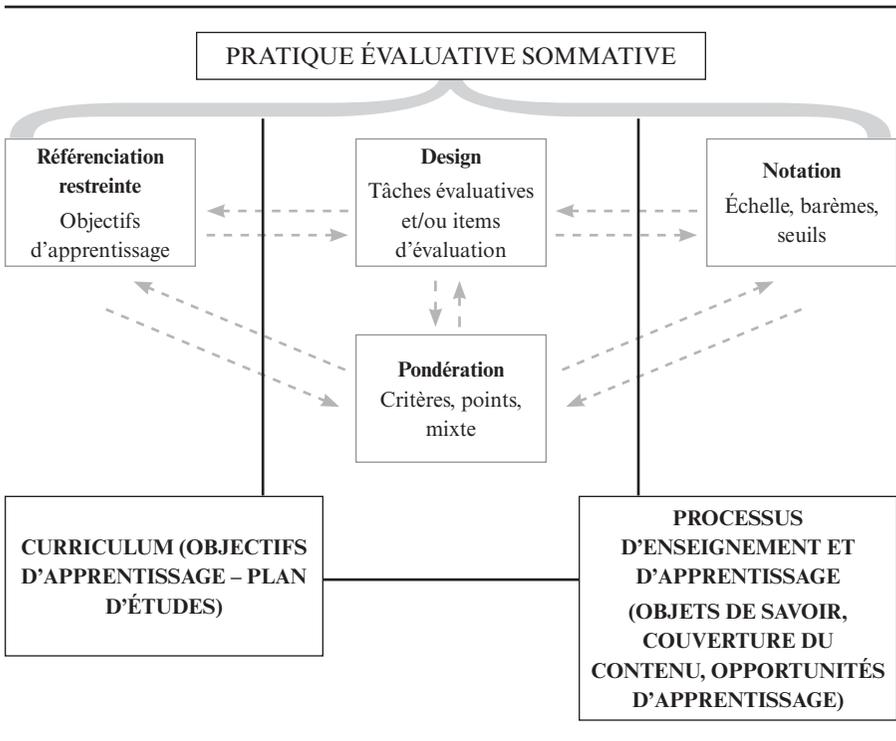


Figure 6. L'alignement curriculaire élargi (Pasquini, 2018)

Le modèle ainsi conceptualisé s'inscrit au cœur du modèle premier où interagissent en cohérence les processus d'enseignement-apprentissage, les objectifs prescrits dans le curriculum et les pratiques d'évaluation, que nous centrons volontairement sur leur fonction de certification, soit l'évaluation sommative (traits foncés continus dans la Figure 6). Cette dernière déploie ensuite ses quatre composantes (acolade en clair) en les mettant en liens réciproques (doubles lignes pointillées en clair). Ainsi, la référenciation restreinte, inspirée de Vial (2012), décrit « ce qui fait foi » chez les enseignants quand ces derniers opèrent des choix dans les objectifs d'apprentissage à évaluer d'un point de vue sommatif. Le « design » (Wiggins et McTighe, 2005) explicite les processus de construction des tâches évaluatives, tant en ce qui a trait aux objets de savoir qui y circulent et aux habiletés cognitives que les élèves doivent mobiliser qu'à la nature des consignes. La pondération renvoie au poids alloué aux apprentissages et aux outils exploités pour ce faire (des critères, des points ou des critères appariés à des points comme modalité mixte ; Andrade, 2005 ; Jonsson, 2014). Enfin, la notation cible les démarches de construction de la note chiffrée en lien avec l'apprentissage à l'aide d'outils comme les échelles (Brookhart, 2017 ; Marzano et Heflebower, 2011). Toutes ces composantes devraient entretenir des liens de cohérence d'un point de vue taxonomique et de contenu.

Ainsi, relativement à notre exemple, nous voyons que les pratiques évaluatives sommatives de Paul mettent en cohérence les objectifs, les tâches et les critères, et proposent une tentative de notation liée aux apprentissages évalués, même si cette dernière reste difficile à mettre en œuvre. Nous pouvons dès lors remettre en question en quoi le poids alloué à tel ou tel apprentissage par le biais d'exercices correspond à la complexité de ces derniers, relativement aux objets mathématiques et aux objectifs qu'ils évaluent. De manière complémentaire et non exhaustive, sachant que des tâches peu référées aux objets de savoir traduisent « un alignement pauvre » (Biggs et Tang, 2011, p. 163, trad. libre) qui ne permettra vraisemblablement pas de pondérer et de noter en référence à ces mêmes objets, nous pouvons interroger la pertinence chez Paul de convoquer, *in fine*, une échelle standardisée pour construire les notes.

À la lumière de ces quelques réflexions articulant éléments de la pratique et apports de la recherche, nous observons que, dans cette nouvelle acception systémique, la cohérence curriculaire de l'évaluation sommative est le fruit d'une triangulation entre plusieurs composantes du modèle, jusqu'à la notation. Dans ce sens, l'alignement curriculaire élargi désigne,

en évaluation sommative, le fort lien de cohérence systémique existant entre les objectifs évalués (référenciation restreinte), les tâches évaluatives (*design*), les points et/ou les critères attribués (pondération) et l'échelle (notation) (Pasquini, 2018).

Discussion

Relativement à notre question de recherche, notre discussion se déploie en trois temps. Tout d'abord, nous étayons une tension entre les pratiques évaluatives de Paul et le modèle théorique élargi, pour donner encore davantage à voir certains obstacles potentiels lorsqu'il s'agit de mettre en œuvre une cohérence dans les pratiques évaluatives jusqu'à la notation. Puis, nous faisons brièvement référence à d'autres de nos résultats afin de mettre en perspective certaines questions en suspens. Enfin, nous explicitons une réflexion sur le modèle ainsi élargi.

Une tension encore très présente

En fin de dispositif, Paul affirme, au sujet de l'alignement curriculaire :

« Ça m'a aidé à plus prendre conscience que la notation, le fait de choisir une échelle, eh bien, ça pose un... c'est un nœud. Ça peut former une cassure entre tout le travail que je me suis donné la peine de faire avant pour transposer les savoirs, pour faire un choix d'objectifs cohérents. Tout ce travail peut simplement être anéanti (rire)... par le choix d'une mauvaise échelle ».

Il assume donc qu'il est amené à rompre cette cohérence en matière de notation par l'impossibilité de la penser différemment, mais il peut vivre avec puisqu'il a bien amélioré sa pratique d'un point de vue général, selon lui : « Je pense qu'en améliorant la qualité de mes évaluations, donc dans leur conception, il ne devrait pas y avoir ce problème-là [de la notation]. »

Il persiste donc chez Paul une difficulté à appréhender le modèle précisément jusqu'à la notation : « Pour moi, l'alignement curriculaire, il s'arrête à la conception de l'évaluation. Après, quand on met la note, quand on choisit l'échelle, ce n'est plus dans l'alignement curriculaire. » Nous faisons l'hypothèse que cette compréhension du modèle en toute fin de recherche rend compte des compromis que les enseignants font en permanence quand ils évaluent et qui sont de nature profondément dynamique (Carless, 2011). Ce qui n'invalide en rien la pertinence du modèle, comme nous l'argumentons dans notre conclusion.

Une mise en perspective avec d'autres résultats

Les résultats que nous exemplifions avec la situation de Paul s'inscrivent dans des résultats plus transversaux qui viennent encore enrichir cette nouvelle conceptualisation élargie et systémique du modèle, tout en y mettant en évidence des points de tension (Mottier Lopez et Pasquini, 2017).

Par exemple, la confrontation du modèle avec les pratiques habituelles des enseignants a un impact sur plusieurs de leurs dimensions : une appropriation plus approfondie du curriculum permettant de mieux hiérarchiser et traduire les objectifs à enseigner et évaluer ; une plus grande homogénéité du niveau de complexité des tâches évaluatives (Guskey, 2006) ; une pondération en points, en critères ou en modalité mixte référée aux objectifs ; ou encore une remise en question de certains processus d'enseignement, ce qui validerait notre nouvelle acception systémique du modèle.

Notons toutefois que Paul n'est pas le seul à avoir des difficultés à garder une cohérence curriculaire jusqu'à la notation. Les trois autres enseignants de mathématiques rencontrent également des obstacles lorsqu'ils doivent construire la note à l'aide de critères et assument, finalement, de recourir à des barèmes qui ne tiennent pas compte des apprentissages évalués. Il en va de même chez deux enseignants de français dans des épreuves de structure de la langue ou de compréhension écrite.

Une interrogation sur le modèle élargi

Ces résultats approfondissent selon nous les apports de Biggs (2003) relativement au fort potentiel de désalignement que possèdent des outils comme les barèmes standardisés externes aux apprentissages réellement menés ou encore les échelles de notation amples. Nous faisons dès lors l'hypothèse que modéliser ainsi la cohérence curriculaire en évaluation sommative permet de mieux comprendre les causes de ces désalignements. En effet, l'impossibilité d'une majorité des participants à construire une note en lien avec l'apprentissage – et ce, même si la cohérence d'alignement curriculaire élargi est perceptible jusqu'à la pondération – souligne l'impact des outils de notation sur les pratiques évaluatives sommatives.

Il est par ailleurs important de rappeler le rôle que joue également le contexte sur les pratiques évaluatives. La culture de l'établissement, son leadership, les habitudes des enseignants et leurs modalités de travail en équipe, mais aussi les prescriptions et règlements sont des facteurs déterminants qui infléchissent grandement la mise en œuvre de pratiques

évaluatives différentes, voire innovantes (Timperley et Parr, 2009). Une des limites de notre recherche est de n'avoir pas assez thématisé en quoi les interactions entre de tels environnements et des pratiques en changements pouvaient avoir un impact sur le modèle. Nous y revenons en fin de texte.

Conclusion

Dans l'ensemble et grâce à l'exemple proposé, nos résultats soulignent l'intérêt à exploiter le modèle élargi comme outil compréhensif des pratiques évaluatives sommatives, ce qui permettrait de mieux appréhender leur complexité. Dans ce sens, notre recherche montre que penser la cohérence des pratiques évaluatives sommatives est pertinent. En effet, conceptualisée ainsi, la quête de cohérence s'apparente à une recherche de validité, notion centrale en évaluation. Bonner (2013) conforte ce postulat lorsqu'elle explicite les caractéristiques que devrait posséder toute démarche d'évaluation valide : une forte cohérence d'alignement curriculaire entre les processus d'enseignement, d'apprentissage et d'évaluation ; la résolution maximale des biais durant la démarche évaluative ; le fait de juger de la pertinence des résultats obtenus ; et la transparence des démarches perceptible pour tous les acteurs concernés.

Il reste toutefois important d'examiner les rapports qui peuvent exister entre les pratiques effectives et un modèle théorique : à quelles conditions ce dernier peut-il rester compréhensif ? Dans quelle mesure son recours n'induit-il pas, malgré tout et au-delà des intentions de recherche, une dimension prescriptive de la pratique, comme pourraient le suggérer les compromis de Paul et de certains de ses collègues vis-à-vis de leur incapacité à noter sur la base de critères ? Et quel rôle joue le dispositif de recherche ou de formation à ce sujet ?

Par ailleurs et comme nous l'avons rapidement évoqué, la cohérence systémique qu'appelle notre modèle est éminemment contextuelle. Sa conceptualisation ne peut dès lors être généralisable, car elle dépend en grande partie de la nature du curriculum, des caractéristiques disciplinaires et des outils régissant la notation. À ce sujet, nos résultats auront conforté l'hypothèse selon laquelle le problème est moins l'existence de la note que son mode de construction et d'exploitation (Brookhart, 2017 ; Walvoord et Johnson Anderson, 2009) et qu'à certaines conditions, la note peut être constructive et à haute valeur informative (Hadji, 2016). Or, jusqu'où ce modèle théorique élargi peut-il cohabiter avec les pratiques évaluatives

dans divers environnements scolaires? À quelles conditions peut-il interagir avec les spécificités des divers curricula, avec les référentiels multiples qu'utilisent quotidiennement les enseignants et avec les outils proposés – voire imposés – à ces derniers pour noter?

Il serait intéressant de mener une recherche similaire avec des enseignants qui sont amenés à utiliser des notes étendues (p. ex., de 1 à 20) ou soumis à des échelles de notation lettrées dont les logiques ordinales sont spécifiques. De manière complémentaire, comment cette cohérence pourrait-elle être intelligible dans une approche par compétences dont l'épistémologie est très différente de la nôtre? Et, à partir du constat du rôle prépondérant que joue la discipline dans les démarches d'évaluation sommative, une analyse plus approfondie pourrait étudier dans quelle mesure certains objets de savoir permettent de construire une cohérence curriculaire jusqu'à la notation (p. ex., en production écrite), alors que d'autres semblent être des obstacles importants à de telles pratiques, comme dans le cas de Paul.

Comme le souligne Moss (2013), des recherches futures considérant les pratiques évaluatives sommatives effectives dans différents contextes demandent à être conduites urgemment. L'enjeu est double : produire des savoirs sur la cohérence de ces pratiques jusqu'aux processus de notation, et enrichir les connaissances et compétences des enseignants en évaluation sommative, bien trop insuffisantes à ce jour.

Réception : 11 décembre 2018

Version finale : 5 septembre 2019

Acceptation : 6 septembre 2019

NOTES

1. Cet article est issu d'une recherche doctorale menée à terme.
2. Dans cette perspective inductive-déductive, nous avons construit cinq catégories : rapport au référentiel, expression de l'objet, élaboration des tâches évaluatives, pratiques de pondération et pratiques de notation.
3. Prénom d'emprunt.
4. www.plandetudes.ch/per

RÉFÉRENCES

- Airasian, W. & Miranda, H. (2002). The role of assessment in the revised taxonomy. *Theory into Practice*, 41(4), 249-254. doi: 10.1207/s15430421tip4104_8
- Allal, L. & Mottier Lopez, L. (2005). Formative assessment of learning: A review of publications in French. In *Formative assessment: Improving learning in secondary classrooms* (pp. 241-264; version française disponible). Paris: OECD-CERI.
- Anadon, M. et Guillemette, F. (2007). La recherche qualitative est-elle nécessairement inductive? *Recherches qualitatives*, 5, 26-37. Repéré à www.recherche-qualitative.qc.ca/documents/files/revue/hors_serie/hors_serie_v5/anadon.pdf
- Anderson, L. W. (2002). Curricular alignment: A re-examination. *Theory into Practice*, 41(4), 255-260. Retrieved from <http://iowaascd.org/files/5313/4045/2315/BloomAI.pdf>
- Anderson, L. W. & Krathwohl, D. R. (Eds.) (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives*. New York, NY: Longman.
- Andrade, H. G. (2005). Teaching with rubrics: The good, the bad, and the ugly. *College Teaching*, 53(1), 27-31. doi: 10.3200/CTCH.53.1.27-31
- Antibi, A. (2003). *La constante macabre ou comment a-t-on découragé des générations d'élèves*. Toulouse: Math'Adore.
- Balach, C. A. & Szymanski, G. J. (2003, April). *The growth of a professional learning community through collaborative action research*. Paper presented at the Annual Meeting of the American Educational Research Association, Chicago, IL. Retrieved from <http://files.eric.ed.gov/fulltext/ED477527.pdf>
- Bateman, D., Taylor, S., Janik, E., & Logan, A. (2009). Curriculum coherence and student success. *Pédagogie collégiale*, 22(5), 8-18. Retrieved from <http://aqpc.qc.ca/sites/default/files/revue/bateman.pdf>
- Biggs, J. (1999). What the student does: Teaching for enhanced learning. *Higher Education Research & Development*, 18(1), 57-75. doi: 10.1080/0729436990180105
- Biggs, J. (2003). *Aligning teaching for constructive learning*. The Higher Education Academy. Retrieved from www.heacademy.ac.uk/system/files/resources/id477_aligning_teaching_for_constructing_learning.pdf
- Biggs, J. & Tang, C. (2011). *Teaching for quality learning at university*. New York, NY: McGraw-Hill.
- Black, P. & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy and Practice*, 5, 1, 7-74. doi: 10.1080/0969595980050102
- Bonner, S. M. (2013). Validity in classroom assessment: Purposes, properties, and principles. In J. McMillan (Ed.), *SAGE handbook of research on classroom assessment* (pp. 87-106). Thousand Oaks, CA: SAGE.
- Braxmeyer, N., Guillaume, J.-C. et Levy, J.-F. (2004). Les pratiques d'évaluation des enseignants au collège. *Les dossiers du ministère de l'Éducation nationale, de l'Enseignement supérieur et de la Recherche*, Direction de l'évaluation et de la prospective, 160, 5-117.
- Brookhart, S. M. (2005). Assessment theory for college classrooms. *New directions for teaching and learning*, 100, 5-14. doi: 10.1002/tl.165

- Brookhart, S. M. (2017). *How to use grading to improve learning*. North Garden, VA: Virginia ASCD.
- Carless, D. (2011). *From testing to productive student learning: Implementing formative assessment in Confucian-heritage settings*. London: Routledge.
- Conseil national d'évaluation du système scolaire (CNESCO). (2014). *L'évaluation des élèves par les enseignants dans la classe et les établissements: réglementation et pratiques. Une comparaison internationale dans les pays de l'OCDE*. Paris: CNESCO.
- Crahay, M. (2007). *Peut-on lutter contre l'échec scolaire?* Bruxelles: De Boeck.
- Denzin, N. K. (1978). *The research act: A theoretical introduction to sociological methods*. New York, NY: Mc Graw-Hill.
- Desgagné, S. (1997). Le concept de recherche collaborative: l'idée d'un rapprochement entre chercheurs universitaires et praticiens enseignants. *Revue des sciences de l'éducation*, 23(2), 371-393. doi: 10.7202/031921ar
- Desgagné, S. (2005). *Récits exemplaires de pratique enseignante: analyse typologique*. Sainte-Foy, QC: Presses de l'Université du Québec.
- Desgagné, S., Bednarz, N., Lebuis, P., Poirier, L. et Couture, C. (2001). L'approche collaborative de recherche en éducation: un rapport nouveau à établir entre recherche et formation. *Revue des sciences de l'éducation*, 27(1), 33-64. doi: 10.7202/000305ar
- Gagné, P., Dumont, L., Brunet, S., & Boucher, G. (2013). Curriculum alignment: Establishing coherence. *Collected Essays on Learning and Teaching*, VI, 7-12. doi: 10.22329/celt.v6i0.3763
- Gauthier, C., Mellouki, M., Bissonnette, S. et Richard, M. (2005). *Écoles efficaces et réussite scolaire des élèves à risque: un état de la recherche* (Rapport de recherche). Québec, QC: CRIFPE.
- Guskey, T. R. (2003). How classroom assessments improve learning. *Educational Leadership*, 60(5), 7-11. Retrieved from www.ascd.org/publications/educational-leadership/feb03/vol60/num05/How-Classroom-Assessments-Improve-Learning.aspx
- Guskey, T. R. (2006). Making high school grades meaningful. *Phi Delta Kappan*, 87(9), 670-675. Retrieved from <https://pdfs.semanticscholar.org/620a/83c1e1911e49f2647bc6ec98ee5d46027e47.pdf>
- Guskey, T. R. (2011). Five obstacles to grading reform. *Effective Grading Practices*, 69(3), 16-21. Retrieved from www.ascd.org/publications/educational-leadership/nov11/vol69/num03/Five-Obstacles-to-Grading-Reform.aspx
- Hadji, C. (2016). Le système éducatif peut-il passer de l'évaluation normative à une évaluation constructive? Dans B. Boquet (dir.). *La fièvre de l'évaluation: quels symptômes, quels traitements?* (pp. 113-135). Villeneuve-d'Ascq: Presses universitaires du Septentrion.
- Hadji, C. (2017). Savoir mettre en œuvre une évaluation constructive. Dans A. Bentolila (dir.). *L'essentiel de la pédagogie* (pp. 229-250). Paris: Nathan.
- Hammerness, K. C. (2006). From coherence in theory to coherence in practice. *Teachers College Record*, 108(7), 1241-1265. doi: 10.1111/j.1467-9620.2006.00692.x
- Harlen, W. (2005). Teacher's summative practices and assessment for learning: Tensions and synergies. *The Curriculum Journal*, 16(2), 207-223. doi: 10.1080/09585170500136093
- Harlen, W. (2012). On the relationship between assessment or formative and summative purposes. In J. Gardner (Ed.), *Assessment and learning* (pp. 87-102). London: SAGE.

- Jonsson, A. (2014). Rubrics as a way of providing transparency in assessment. *Assessment & Evaluation in Higher Education*, 39(7), 840-852. doi: 10.1080/09585170500136093
- Martone, A. & Sireci, S. G. (2009). Evaluating alignment between curriculum, assessment, and instruction. *Review of Educational Research*, 79(4), 1332-1361. doi: 10.3102/0034654309341375
- Marzano, R. J. (2002). A comparison of selected methods of scoring classroom assessments. *Applied Measurement in Education*, 15(3), 249-267. doi: 10.1207/S15324818AME1503_2
- Marzano, R. J. & Heflebower, T. (2011). Grades that show what students know. *Effective Grading Practices*, 69(3), 34-39. Retrieved from www.math.arizona.edu/~vbohme/Grades%20That%20Show%20What%20Students%20Know.pdf
- McGrath, H. et Noble, T. (2008). *Huit façons d'enseigner, d'apprendre et d'évaluer*. Montréal, QC: Chenelière Éducation.
- McKinney, S. E., Chappell, S., Berry, R. Q., & Hickmann, B. T. (2009). An examination of the instructional practices of mathematics teachers in urban schools. *Preventing school failure: Alternative education for children and youth*, 53(4), 278-284. doi: 10.3200/PSFL.53.4.278-284
- McMillan, J. H. & Nash, S. (2000, April). *Teacher classroom assessment and grading practices decision making*. Paper presented at the NCME Annual Meeting, New Orleans, LA. Retrieved from <https://files.eric.ed.gov/fulltext/ED447195.pdf>
- Meier, S. L., Rich, B. S., & Cady, J. (2006). Teacher's use of rubrics to score non-traditional tasks: Factors related to discrepancies in scoring. *Assessment in Education*, 13(1), 69-95. doi: 10.1080/09695940600563512
- Moss, C. (2013). Research on classroom summative assessment. In J. McMillan (Ed.), *SAGE handbook of research on classroom assessment* (pp. 235-256). Thousand Oaks, CA: SAGE.
- Mottier Lopez, L. (2015). *Évaluations formative et certificative des apprentissages*. Bruxelles: De Boeck.
- Mottier Lopez, L. et Laveault, D. (2008). L'évaluation des apprentissages en contexte scolaire: développements, enjeux et controverses. *Mesure et évaluation en éducation*, 31(3), 5-34. doi: 10.7202/1024962ar
- Mottier Lopez, L. & Pasquini, R. (2017). Professionnal controversies between teachers about their summative assessment practices: A tool for building assessment capacity. *Assessment in Education: Principles, Policy & Practice*, 24(2), 228-249. doi: 10.1080/0969594X.2017.1293001
- Mottier Lopez, L., Tessaro, W., Dechamboux, L. et Morales Villabona, F. (2012). La modération sociale: un dispositif soutenant l'émergence de savoirs négociés sur l'évaluation certificative des apprentissages des élèves. *Questions vives*, 18, 159-175. Repéré à <https://journals.openedition.org/questionsvives/1235>
- Paillé, P. et Muchielli, A. (2012). *L'analyse qualitative en sciences humaines et sociales*. Paris: Armand Colin.
- Pasquini, R. (2013). *Quand les récits de pratique enseignante parlent d'apprentissage*. Québec, QC: Presses de l'Université Laval.
- Pasquini, R. (2016). Et si, pour mieux saisir la complexité de la pratique, nous passions par l'écriture? *Didactiques en pratique*, 2, 21-29. Repéré à <https://orfee.hepl.ch/handle/20.500.12162/2423>

- Pasquini, R. (2017). Évaluation : première question vive de la pratique. *L'Éducateur*, 4, 38-39. Repéré à <https://orfee.hepl.ch/handle/20.500.12162/1604>
- Pasquini, R. (2018). *Le modèle théorique de l'alignement curriculaire élargi pour étudier des pratiques évaluatives sommatives d'enseignants de mathématiques et de français du secondaire : enjeux conceptuels et pragmatiques* (Thèse de doctorat inédite). Université de Genève, Genève. Repéré à <https://archive-ouverte.unige.ch/unige:106442>
- Randall, J. & Engelhard, G. (2010). Examining the grading practices of teachers. *Teaching and Teacher Education*, 26(7), 1372-1380. doi: 10.1016/j.tate.2010.03.008
- Rieg, S. A. (2007). Classroom assessment strategies: What do students at-risk and teachers perceive as effective and useful? *Journal of Instructional Psychology*, 34(4), 214-225.
- Royce Sadler, D. (2009). Transforming holistic assessment and grading into a vehicle for complex learning. In G. Joughin (Ed.), *Assessment, learning and judgement in higher education* (pp. 1-15). London: Springer.
- Sayac, N. (2017). Étude des pratiques évaluatives en mathématiques de 25 professeurs des écoles français : une approche didactique à partir de l'analyse des tâches données en évaluation. *Mesure et évaluation en éducation*, 40(2), 1-31. doi: 10.7202/1043566ar
- Silverman, D. (2009). *Doing qualitative research: A practical handbook*. Thousand Oaks, CA: SAGE.
- Timperley, H. S. & Parr, J. M. (2009). Chain of influence from policy to practice in the New Zealand literacy strategy. *Research Papers in Education*, 24(2), 135-154. doi: 10.1080/02671520902867077
- Van Nieuwenhoven, C. et Colognesi, S. (2015). Une recherche collaborative sur l'accompagnement des futurs instituteurs : un levier de développement professionnel pour les maîtres de stage. *Évaluer : Journal international de recherche en éducation et formation*, 1(2), pp. 103-121. Repéré à http://usherbrooke.criipe.ca/wp-content/uploads/2016/05/vnh-colo-jiref_2015.pdf
- Vial, M. (2012). *Se repérer dans les modèles de l'évaluation*. Bruxelles : De Boeck.
- Walvoord, B. E. & Johnson Anderson, V. (2009). *Effective grading: A tool for learning and assessment*. San Francisco, CA: Jossey Bass.
- Wigging, G. & McTighe, J. M. (2005). *Understanding by design*. Alexandria, VA: ASCD.
- Winger, T. (2009). Grading what matters. *Educational Leadership*, 67(3), 73.