Document generated on 05/04/2024 5:13 a.m.

Mesure et évaluation en éducation

Conceptually expanding the curricular alignment model to understand the coherence of the graded summative assessment practices of teachers: issues and perspectives

Raphaël Pasquini

Volume 42, Number spécial, 2019

Translation Issue

URI: https://id.erudit.org/iderudit/1084128ar
DOI: https://doi.org/10.7202/1084128ar

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). Conceptually expanding the curricular alignment model to understand the coherence of the graded summative assessment practices of teachers: issues and perspectives. *Mesure et évaluation en éducation*, *42*(spécial), 39–68. https://doi.org/10.7202/1084128ar

Article abstract

Curricular alignment refers to the links of coherence, found in any teaching-learning process, between curriculum objectives, learning tasks and assessment approaches. This model makes it possible to understand the coherence of any assessment approach. By mobilizing data from a collaborative study carried out on eight secondary school teachers of mathematics and French, we will show, however, that its meaning is rather limited when it comes to understanding coherence in graded summative assessment practices and that, consequently, the model needs to be expanded conceptually. To this end, we will draw on an example of a summative test modelled in this way. Our findings demonstrate the relevance of analyzing summative assessment practices with the help of the expanded model, while considering the role that context plays in certain of its aspects.

Tous droits réservés © ADMEE-Canada - Université Laval, 2020

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/

Conceptually expanding the curricular alignment model to understand the coherence of the graded summative assessment practices of teachers: issues and perspectives*

Raphaël Pasquini
University of Teacher Education state of Vaud

Key words: summative assessment, curricular alignment, assessment practices, grading

Curricular alignment refers to the links of coherence, found in any teaching-learning process, between curriculum objectives, learning tasks and assessment approaches. This model makes it possible to understand the coherence of any assessment approach. By mobilizing data from a collaborative study carried out on eight secondary school teachers of mathematics and French, we will show, however, that its meaning is rather limited when it comes to understanding coherence in graded summative assessment practices and that, consequently, the model needs to be expanded conceptually. To this end, we will draw on an example of a summative test modelled in this way. Our findings demonstrate the relevance of analyzing summative assessment practices with the help of the expanded model, while considering the role that context plays in certain of its aspects.

* French version: *Élargir conceptuellement le modèle de l'alignement curriculaire pour comprendre la cohérence des pratiques évaluatives sommatives notées des enseignants : enjeux et perspectives* – vol. 42, n°1, 63-92

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 appuierons 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.

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 concetualmente 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.

Author's note: Correspondence relating to this article may be sent to [raphael.pasquini@ hepl.ch].

The Problem

A considerable body of research has demonstrated that formative assessment is an indispensable approach to assessment in supporting students' learning (e.g., Allal & Mottier Lopez, 2005; Black & Wiliam, 1998; Mottier Lopez, 2015). Summative assessment, by contrast, is defined as an approach "whose purpose is to carry out an inventory. Its social use is to verify (test), while its social function is to certify. Its ancillary functions include classifying, contextualizing and even informing" (Hadji, 1989, cited by Mottier Lopez, 2015, p. 23; Translator's translation). Despite the fact that it raises important pedagogical issues for students and that research has pointed to its role in academic failure (Antibi, 2003; Crahay, 2007; Guskey, 2011), it has, paradoxically, received less attention in French-language research, whence the need to investigate it.

What do we know about teachers' summative assessment practices? What phenomena do they highlight? What issues do they raise? What are their unique features? Research in the field of assessment addresses these questions through surveys. Their findings are quite striking. In particular, they show that, irrespective of the discipline and grade, it is difficult to characterize the coherence of practices, as there are often major discrepancies between teachers' representations of coherent summative assessment practices and the descriptions of their own practices (Braxmeyer, Guillaume & Levy, 2004; McKinney, Chappell, Berry & Hickmann, 2009; Rieg, 2007).

Few studies to date have addressed this issue from the standpoint of effective practices. In the Anglo-Saxon research community, Moss (2013) states: " there is a need for research designs that go beyond teachers' self-report, surveys, and inventories" (p. 252). She argues that this is a prerequisite for conceptually and pragmatically addressing the issue of coherence of practice. A similar observation was made in the Francophone community (CNESCO, 2014; Sayac, 2017).

Despite these shortcomings, certain field-based studies reveal some interesting results (e.g., Bateman, Taylor, Janik & Logan, 2009; Harlen, 2005, 2012; Sadler, 2009). They note, for example, the complexities involved in thinking about the coherence of summative assessment practices, even if teachers agree in advance, and as a team, on the learning objectives to be pursued and assessed, or on the activities to be carried out with the students. Coherence is therefore difficult to identify and, *a fortiori*, to conceptualize, though it is crucial (Gagné, Dumont, Brunet & Boucher, 2013; Martone & Sireci, 2009). Mottier Lopez and Laveault (2008) confirm this last idea: "The search for coherence and articulation of the three spheres - teaching, learning, assessment - is entirely characteristic of the most recent developments in educational assessment" (p. 9; Translator's translation).

The present article focuses on this issue. It aims (i) to examine the extent to which a theoretical model can provide a better understanding of the coherence of graded summative assessment practices and (ii) simultaneously reveal the conditions in which these practices can interact with the model to gain coherence. In the present research, we argue that the model of curricular alignment is relevant in pursuing this goal.

In the conceptual framework, we begin by describing this model, highlighting its main characteristics and some of its theoretical limitations in thinking about graded summative assessment practices. We then present certain methodological aspects of our research, particularly with regard to its collaborative method. Then, starting with a summative test carried out by a teacher, we use a case study to illustrate how this teacher's practices have evolved in terms of their coherence and, at the same time, how a conceptual extension of the model became necessary in order to understand this coherence in the test. We then proceed to present our main results, and later introduce a few remaining discussion points related to our research question, before concluding with some avenues for further research.

Conceptual framework

Using curricular alignment to model coherence in summative assessment

The literature on the curricular alignment model postulates that any assessment practice must be highly consistent with the prescribed objectives, the instruction provided, the learning tasks assigned to students and the activities actually carried out by the latter (e.g., Anderson, 2002; Biggs, 1999, 2003; Gauthier, Mellouki, Bissonnette & Richard, 2005). Here, the challenge for assessment is to focus on planned and actual learning. In this perspective, the development of summative assessment tests is conceived as a complex, dynamic and contextual process focused on learning. The concept of alignment refers to a conceptualization of coherence based on a taxonomy of cognitive abilities related to content, to which we will return later. The curricular dimension refers to the fact that this coherence is based on a curriculum.

In 2005, Gauthier, Mellouki, Bissonnette and Richard published a review of the research on effective schools and academic achievement for at-risk students in North America. They demonstrated that "curricular alignment appears to be highly likely to improve the quality of teaching and the effectiveness of schools" (p. 28, Translator's translation), even if its implementation alone does not guarantee improved learning. Drawing on Guskey (2003), Gauthier and colleagues defined curricular alignment as "ensuring a high degree of correspondence between curriculum, instruction and assessment" (p. 24). They further stated that the model also involves "intellectual operations and the various categories of knowledge that are related to them" (p. 25, Translator's translation). These researchers described three steps for implementing the model.

Step 1

The first step is to identify what students need to achieve in terms of performance, that is, their output. The key question is: "How can one observe and judge the degree of mastery of a given skill or knowledge in terms of observable behaviour or manifestations?" (Gauthier et al., 2005, p. 24, Translator's translation). The formulation of learning objectives and assessment criteria is central here. It involves making the curriculum concrete in order to facilitate the teaching and assessment of its content.

Step 2

The second step is to plan the teaching and learning sequences related to the selected curriculum objectives, so that all students make progress. It is expressed through the following question: "How can we help students learn the content that needs to be mastered?" (Gauthier et al., 2005, p. 25, Translator's translation).

Step 3

The third and final step is to assess what the students have learned. It focuses on what they have learned in the classroom (Gauthier et al., 2005, p. 25). It is therefore important that the assessment tasks are recognizable and consistent with what has been taught and with the curriculum.

With regard to the issue being examined in the present article, this operationalization of the model may be challenged on three levels. First, we hypothesize that its step-by-step approach does not take into account the complexity of effective practices when a teacher attempts to link the curriculum, the learning tasks and the assessment process. Indeed, we know that when teachers develop their summative assessment approaches, the curriculum is one reference among others (Mottier Lopez, Tessaro, Dechamboux & Morales Villabona, 2012). Consequently, these practices cannot be considered without first working on the curriculum content when the latter is the point of reference.

The second concern involves the concept of congruence and its relationship to that of coherence, which is theorized primarily in the literature on the model (e.g., Anderson, 2002; Hammerness, 2006).

The third concern - the main one addressed in the present article – pertains to the limits of theorizing summative assessment in the model itself. Indeed, the analysis of Gauthier and colleagues stops at consideration of assessment tasks. It does not address the way in which weight is assigned to the learning assessed through points or criteria, nor issues of grade construction. This means that the two key practices of scoring and grading are not sufficiently conceptualized, even though they are interrelated (Anderson, 2002; Biggs, 2003) and inseparable from the development of summative assessment with reference to a curriculum (Marzano, 2002; Brookhart, 2005). This shortcoming is all the more regrettable because research has shown that grading practices and summative assessment

practices are differentiated and conceived in a specific way. Modeling grading practices is therefore very complex (McMillan & Nash, 2000; Randall & Engelhard, 2010; Walvoord & Johnson Anderson, 2009). These studies demonstrate that certain factors have a major influence on grading practices (e.g., the standards prescribed by tools such as scales, the question of adequacy or sufficiency (pass) thresholds, consideration of the effort - and even the additional performance - demonstrated by the student, the distribution of grades and the issue of assigning zero as indicative of a total absence of learning. In other words, a teacher can be coherent in constructing assessment tasks or criteria, while being very incoherent when grading. This must be taken into account when considering the coherence of the summative assessment as a whole.

Similar shortcomings can be observed in the theorization of Anderson's (2002) model, which remains a reference. Curricular alignment is defined as "the degree of coherence between curriculum objectives and learning assessment, between curriculum objectives and learning tasks, and between learning assessment and learning tasks" (p. 257, Translator's translation of author's free translation). The content validity, the way it is taught, and the opportunities for students to learn it through learning tasks are also part of the model. The author depicts curricular alignment as a triangle with each vertex as a component (see Figure 1).

The theoretical limitations of summative assessment can be seen in the generic dimension of learning assessment, and more distinctly in the questions the researcher formulates to understand the model. For example, for relationship A, which makes explicit the coherence between learning assessment and the learning objectives derived from the curriculum, Anderson asks: "To what extent does the test measure the important curricular objectives?" (p. 255), without elaborating on this measurement issue. As for relationship C, which highlights the coherence that should exist between learning assessment and learning tasks, the author asks: "Is what we are teaching being tested?" (p. 255), but fails to develop the fine details of the summative test development. We can see that the central issues of scoring and grading are not addressed.

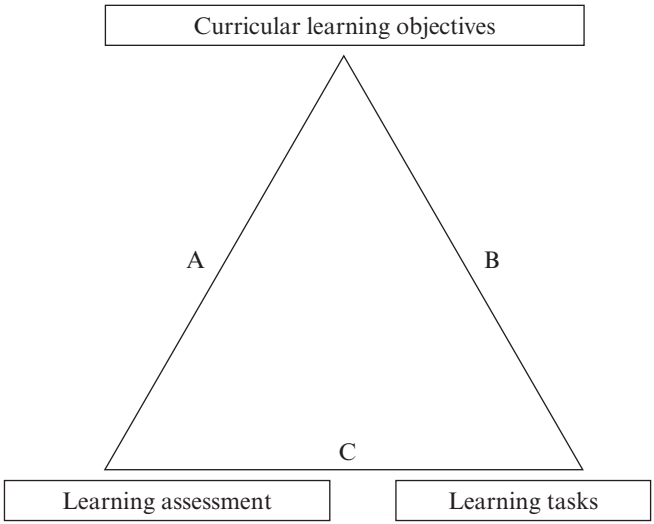


Figure 1. Curricular alignment according to Anderson (2002, p. 256)

The issue of curricular coherence

These limitations call for a more thorough examination of the concept of coherence that lies at the heart of the model. The researchers surveyed speak of a coherence of curricular alignment constructed with a taxonomy (Anderson, 2002) or of "constructive alignment" (Biggs, 2003).

As part of our research, we used Anderson and Krathwohl's (2001) taxonomy, as it is the one most frequently mentioned in the works reviewed. It conceptualizes learning as the capacity to master objectives, formulated as cognitive abilities related to content. We are therefore talking about objectives, not skills. From this perspective, two key elements emerge: the complexity of the cognitive ability and, simultaneously, the characteristics of the subject to which it pertains (McGrath & Noble, 2008). The learning objective is thus defined as an ability that includes both a general and a specific dimension.

The first dimension refers to the student's mental activity across the taxonomy's six cognitive abilities (remembering, understanding, applying, analyzing, evaluating and creating). The second refers to the content on which the learning is focusing: relative clauses in French, isometrics in mathematics, etc. The complexity of the objective therefore depends

on these two dimensions (e.g., memorizing a mathematical vocabulary pertaining to trigonometry), which should, ideally, make it possible to establish coherent relationships between each component of the model.

Coherence until grading

In summative assessment, this conceptualization of curricular coherence remains relevant to understanding the relationships that should exist between all the components of the model, until scoring and grading (Airasian & Miranda, 2002). Consider the following: scoring is the act of representing the importance of learning - its weight - through tools such as criteria or points (Marzano, 2002); grading as it relates to learning is "a complex, context-dependent process that plays multiple roles" (Walvoord and Johnson Anderson, 2009, p. 2).). In the light of these two considerations, we will analyze how this taxonomic coherence can be reflected conceptually and in terms of effective practices up until the point at which teachers assign points/grades or develop criteria, but also when they establishe grade thresholds (in our case, from 1 to 6, including half grades, with grade 4 signifying the sufficiency or pass threshold).

A first step toward an expanded curricular alignment model

Biggs (1999, 2003) is one of the researchers highlighting the importance of a form of conceptualization of summative assessment in the curricular alignment model that includes scoring and grading. His work emphasizes three elements regarding the relationship between summative assessment and curricular objectives on the one hand, and between summative assessment and learning tasks on the other: 1) the assessment must address complex objectives and show the quality of the learning; 2) the application of criteria related to the content and objectives is central and presented as "the key to easy and successful grading" (p. 6); 3) grading must reflect learning through the use of criteria and tasks.

These elements constitute a significant first step in expanding the model conceptually, and we have drawn on them in our work with the teachers involved in our research. However, this research still has two limitations that continue to challenge the model.

The first is the failure to consider disciplinary specificities, even though the subjects influence assessment practices (e.g., Meier, Rich, & Cady, 2006; Sayac, 2017). The second is to refer only to a single grading scale (a six-position scale, with grades from A to F and only one failing grade:

F), whereas this tool has also been shown to strongly influence assessment practices (Hadji, 2017). Thus, we need to consider a comprehensive model that can be transposed to other academic and disciplinary environments, and adapted to various grading scales.

Thus, on the one hand, practice seems to need a theoretical model to conceptualize coherence in summative assessment. On the other hand, the research on scoring (e.g., Jonsson, 2014; Marzano, 2002; Sadler, 2009) and grading (e.g., Brookhart, 2017; Walvoord & Johnson Anderson, 2009; Winger, 2009) calls for further conceptualization of coherence in summative assessment related to context, curricula, and grading tools.

Methodology

The question we wish to answer is the following: In terms of effective practices and research findings, how can we conceptually expand the curricular alignment model to understand coherence in graded summative assessment?

To this end, we begin by presenting our research device and our collaborative research method. We then clarify our data and the selected analytical methods.

The modalities and time frame of our research process

We worked with eight volunteers, all of whom were experienced teachers at the secondary school level in the state of Vaud in French-speaking Switzerland; four were mathematics teachers and four were French teachers. We designed our research scheme so that it would enable us to access the summative tests they had created.

Three other types of data were collected:

1) recordings of conversations at meetings during which the teachers and the researcher discussed these tests; 2) two semi-structured interviews before and after the presentation of the research scheme, in order to gain an in-depth understanding of their assessment practices; 3) three exemplary texts on practices (Desgagné, 2005; Pasquini, 2013, 2016), one of which was written beforehand, one during the process and one at the end of the process. The nature and arrangement of these different time periods as well as the times of data collection are presented in Table 1.

Collaborative research

Drawing on the questions and problems raised by the teachers, we designed a research scheme that gave them the opportunity to apply the theoretical knowledge in analyzing, designing and discussing their summative assessments. To this end, during the first work session we introduced them to the expanded curricular alignment model (Phase 3), using a simplified version developed solely with the research input, as explained in the conceptual framework. We opted for a collaborative research method, defined by Desgagné (1997) as follows: "At its core, we would say that collaborative research involves the contribution of teachers to the process of investigating a research question, a process that is usually supervised by one or more researchers" (p. 372, Translator's translation). Van Nieuwenhoven and Colognesi (2015) specify the two unavoidable goals of this method: "to foster knowledge creation and aim for the professional development of each category of actors involved in the research process" (p. 105, Translator's translation). Figure 2 shows the schema of our collaborative research, which follows these definitions and the work of Desgagné, Bednarz, Lebuis, Poirier and Couture (2001). We describe the content of this diagram in greater detail further on.

A stage of negotiation between the teachers and the researcher helped to delineate a training goal: improving the coherence of the participants' graded summative assessment practices. The object of the research then became clear. The challenge was to collect data to document and understand how the coherence of summative assessment practices was affected when confronted with the theoretical model of expanded curricular alignment using the research contributions; here, we paid particular attention to the contrasting subjects (French and mathematics) and teacher/teaching contexts (types of students, local requirements/policies, team approaches) - while also observing how the model reacted to these practices in context.

Table 1
Research scheme

Time	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
Session	Group formation Recording of questions	One on one interview 1	Individual writing tale 1	Meeting 1 *All together*	Individual writing tale 2	Meeting 2 *Mathematics and French separately*	Meeting 3 *All together*	Individual writing tale 3	One-on-one interview 2
Data collected		Semi-structured interview	Writing tale	First summative assessment created by teachers **Presentation of theoretical model**	Writing tale	Discussion about summative assessments completed Second summative assessment created by teachers	Discussion about summative assessments completed	Writing tale	Semi-structured interview

Note. 1 = Mathematics; 2 = French.

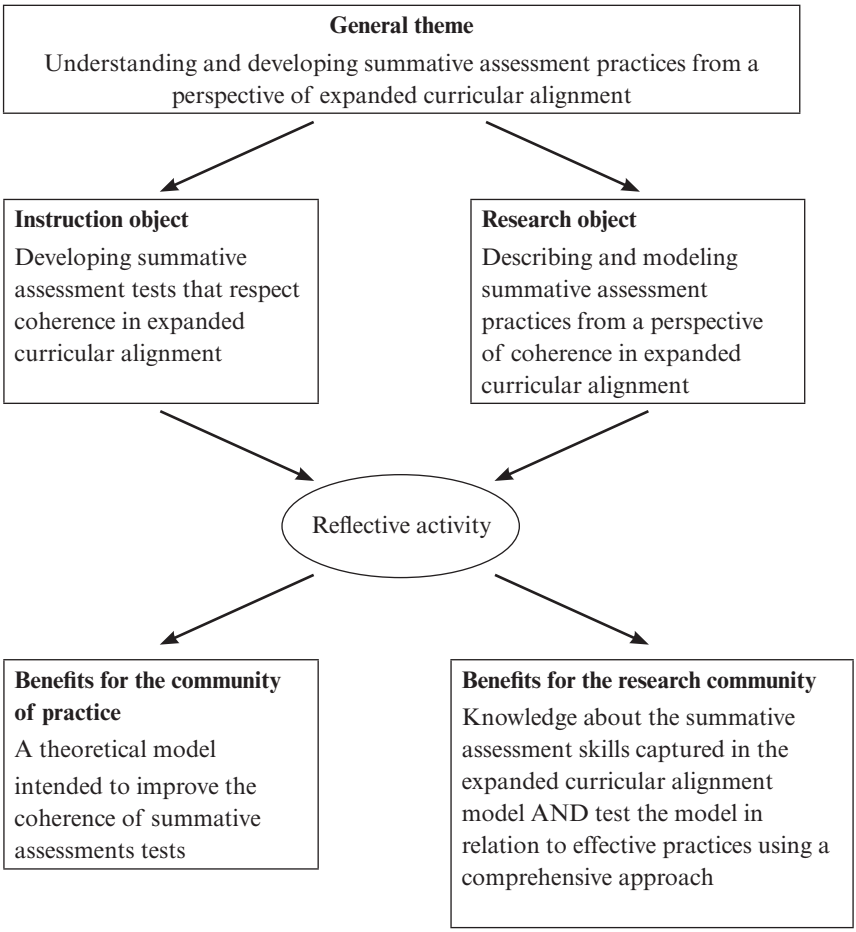


Figure 2. Our collaborative research model according to Desgagné et al. (2001)

In this way, reflective activity cut across our entire scheme. From the outset, the teachers were encouraged to thematize the issues and/or questions they wished to address, and to do this using an interview and written material. This was also the case in the analysis of the assessments for the purpose of modelling them; in the development of assessments, with reference to teaching, and in the light of broader curricular alignment; in

the comparison between the assessments and the discussion points proposed by the researcher or peers; and in the written material referring to the assessment practice.

Created in this manner, our collaborative research tried to highlight the role of expanded curricular alignment in the development of coherent summative assessment practices relating to classroom contexts (Balach & Szymanski, 2003), and to put a theoretical model into practice.

The data examined

In the present article, we have given priority to the summative tests created by the teachers, and we use one of them to illustrate our point. From a methodological perspective, we have complied with the way in which the teachers created the tests (working individually; working in pairs and consistent with their collaborative undertaking at the schools; selecting the subjects and the goals to assess), so that they were linked to the students' learning. However, to support our discussion on the phenomena of coherence in the summative assessment, we also used excerpts from teachers' tales and interviews.

A two-way process between summative assessment practices and the theoretical model

Our methodology first involved an inductive approach, focusing on the teachers' "strong questions regarding their assessment practice" (Pasquini, 2017, Translator's translation), collected at the beginning of the process during Time 0. Then, promptly and by design, we introduced the theoretical model to see the extent to which it could help the teachers understand the in-coherence in their practices, and consequently discover ways to deal with their questions and problems. A deductive approach then took over, since it was in the light of this model, and as part of a methodological triangulation process (Denzin, 1978; Silverman, 2009), that the data were interpreted with the use of conceptualizing categories (Paillé & Mucchielli, 2012).

While relying on a model is not satisfactory since any theoretical model must be tried and tested using data that reflect actual practice, it became necessary in our qualitative process (Anadon & Guillemette, 2007). Indeed, the aim was to open up possibilities for a potential reconfiguration

of the initial theoretical model, as well as to remain flexible with regard to the coherence phenomena that our data would clarify. Consequently, one could influence the other, and vice versa.

Given these considerations, we now turn to a case study: a test created by a teacher involved in our research. We point to several aspects of an assessment practice whose coherence has evolved and which, in the process, has influenced our expanded model conceptually.

The case study

The example provided comes from a mathematics test. We chose it from among a dozen other examples in our research because it reveals attempts to optimize expanded curricular alignment coherence that involve all model components, until grading.This assessment, carried out by Paul[2] during the second meeting (Time 5, involving mathematics teachers) and discussed by the group during the final meeting (Time 6), pertained to functions and algebra, and more specifically to literal calculation. It consists of nine exercises and is designed for students in grade 10 (the second-to-last grade of compulsory schooling, average age 15), in the *Voie générale*, ("general stream" or "general section", the least demanding of the two existing streams in the system), and in level 2 of this stream (a higher level of requirements, whereas level 1 only has basic performance requirements). Figure 3 shows its content. The learning objective being assessed from the Swiss French regional curriculum (Plan d'études romand[4]) is "to solve numerical and algebraic problems" (Translator's translation). Paul selects four of its constituent elements, then six learning progression elements of the objective (P1 to P6) corresponding to the nine exercises in the test.

In another document, Paul explains his scoring and grading practices. As a first step, he linked each exercise (Ex) to the corresponding learning progression(s) (P1 to P6) (see Table 2).

We observe that Paul decided to assess P1 in each exercise. By contrast, each of the other progressions is assessed separately in the various exercises (e.g., P4 is assessed in exercises 4 and 5).

2. See notes at end of article

Core Test 2

Curriculum objectives assessed: MSN 33 – Solve numeric and algebraic problems

…by recognizing a situation's mathematical characteristics and translating it into numeric writing

…by using operational properties

…by using algebra as a calculation tool

…by developing, practising and using calculation procedures

Learning progressions assessed:

P1 – Knowledge and use of standard algebraic writing rules and conventions (throughout the test and in Exercise 1)

P2 – Determination of the numeric value of a literal expression by substituting numbers in place of variables (Exercises 8 and 9)

P3 – Creation of literal expressions from problems, geometric figures or verbal expressions (Exercises 7, 8 and 9)

P4 – Interpretation of literal expressions and identification of those that are equivalent (Exercises 4 and 5)

P5 – Knowledge of the terminology and organized shorthand of monomials and polynomials (Exercises 1, 2, 3, 6 and 7)

P6 – Monomial and polynomial operations (+/-/x) (Exercises 2, 3, 6 and 7)

Figure 3. Objectives assessed in Paul's second literal calculation test (translation)

In the end, Paul gives each progression the same scoring, regardless of the number of times it was assessed in the exercises, using a points system (where insufficient = 1; sufficient = 2; completely achieved =3). This horizontal reading therefore requires that Paul observe up to nine times whether a progression is inadequate, adequate or completely achieved, and then to determine

Table 2
Linking exercises and learning progressions

	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		√	√			√	√			

what final assessment to give each of them. Figures 4 and 5 illustrate how he defines the expected level of sufficiency (pass level) for each progression in the exercises. For purposes of comparison, Paul also mentioned, as was his usual procedure, the number of points he would assign to each exercise.

Exercise 2 (pts) 2.5 pts

Simplify the following literal expressions:

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

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

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

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

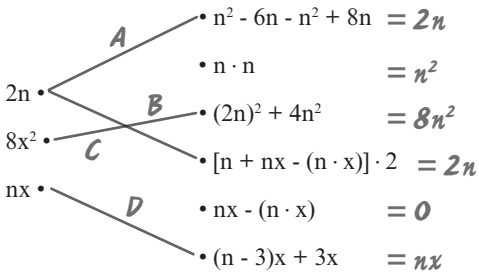
e) $-4z^2 \cdot z^3 \cdot z =$ $-4z^6$

P6 Pass threshold: 3 correct
P5 Pass threshold: 0 writing errors

Figure 4. Paul's second literal calculation test with thresholds indicated for P1, P5 and P6 in Exercise 2

Exercise 4 (pts) 2 pts

Connect the equivalent expressions:



P4 Pass threshold: Identify 2 of the 4 links. Max 2 incorrect links.

Figure 5. Paul's second literal calculation test with threshold indicated for P4 in Exercise

Figures 4 and 5 show how he defines the expected pass level for each progression through the exercises. For comparison, Paul also indicated the number of points he would give to each exercise, as he usually does.

In these new practices, it is therefore possible for the student to obtain a maximum of 18 points (6P x 3) in the test, which represents the maximum grade of 6 (see Table 3). If all the progressions are sufficient, he gets 12 points (6P x 2), which corresponds to a grade of 4. Below that, the student fails. On this basis, Paul created the following scale.

Table 3
Paul's grading scale

Grade	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

Results

In keeping with our research question and methodology, we will present our findings in two parts: those pertaining to changes in the coherence of Paul's summative assessment practices, and those pertaining to the evolving theoretical model.

The findings on the coherence of the practices

The analysis of our data shows that Paul's summative assessment practices converge towards greater curricular coherence: the relationship between the homogeneous complexity exercises and the objectives is understandable from the standpoint of the learning objects and cognitive abilities; the exercises are designed with reference to in-class learning; and the links between the learning progressions that are considered to be criteria and the exercises are explicit.

Changes in scoring and grading practices, which are also present, deserve to be discussed in detail, as they contrast with traditional practices. Whereas at the beginning of the research Paul was vague on these points, with the help of the expanded model he was able to carry out in-depth research on the systemic coherence between the objectives and the tasks, as well as on the scoring tools focusing on learning, thereby revealing what the student had or had not achieved. In this test, we observe that instead of counting points allocated to exercises, the teacher uses the exercises to assess the achievement of the learning progressions in the *Plan d'études romand* (i.e., the French-language study plan). This represents an important shift in perspective. However, we note that there were strong tensions at play in his practice:

> "One seeks to point it out when something is not successful; yet who is to say what's unsuccessful? This is what I think, and what I would like to do. So... in my last assessment, I worked on... additional success criteria, with assessments, and I thought this was pretty good. But I didn't see how to make the leap from these assessments to the grade. In the end, I found it difficult. You can always find tricks and devices, but it always comes down to choosing a scale, choosing a 4, a success or pass threshold, and then, later, well, there you have it, doing it proportionally or not... it's always comes down to the same thing".

Therefore, obstacles to curricular coherence persisted between the point at which criteria are determined - in this case, for Paul, the learning progressions - and the grading. As he says, he makes compromises when he uses standardized scales: "That's the foundation of my work. I don't think it's right, but I have to work with it." In this regard, Paul was therefore referring to conventional practices. Aware that this procedure leads to a misalignment, since the scale has not been designed in relation to the learning assessed, he accepted it all the same, since he found an alternative solution difficult to implement. On the one hand, Paul was able to explain the path he took towards a grading more in line with learning: "Concerning the scale, [...] I imagined assessing the students by taking the progressions as criteria, one after another and on an equal footing." On the other hand, his fear of too much subjectivity seemed to get the better of him:

"But how can we justify that the choice of this or that criteria is more equitable than another? In principle, I am convinced that this type of assessment is the right approach, but how can we transform our criteria into a grade in the fairest way possible? So, the problem remains the same: which scale to choose?"

One form of insecurity dominates. We will return to this in our discussion.

It is now time to ask ourselves how we can combine these observations of practice with the research findings in the area of summative assessment, so that we can conceptualize a model that will make it possible to better understand the coherence in assessment up to the point of grading.

The findings on the evolution of the model

To accomplish this task of conceptualization, we relied on the phenomena identified in the participants' practices, as in our example, as well as on two characteristics of Anderson's (2002) initial model: (1) a dynamic perspective for thinking about the relationships between its various components (curriculum objectives, learning tasks, assessment); and

(2) the initial inclusion of the model in the teaching-learning process. The taxonomic coherence relationships that ideally exist between the various components thus reflect the research findings and the effective practices.

In this connection, we paid attention to (a) the links of coherence between the test criteria and the objectives assessed (or presented as such), (b) their corresponding relationship with the assessment tasks and (c) the grading tools. Similarly, we observed the extent to which the learning object and the cognitive abilities describing the assessed learning objectives characterized the success (pass) thresholds in the grading scale constructed; we did this using the assessment tasks and/or the assessment criteria that should, ideally, make them explicit.

Conceived in this way, the expanded model should make it possible to understand the phenomena of coherence or incoherence (or curricular misalignment) present in summative assessment practices and observable in tests, and then to refer to a logic understood in a systemic and contextual perspective. We have therefore illustrated it as follows:

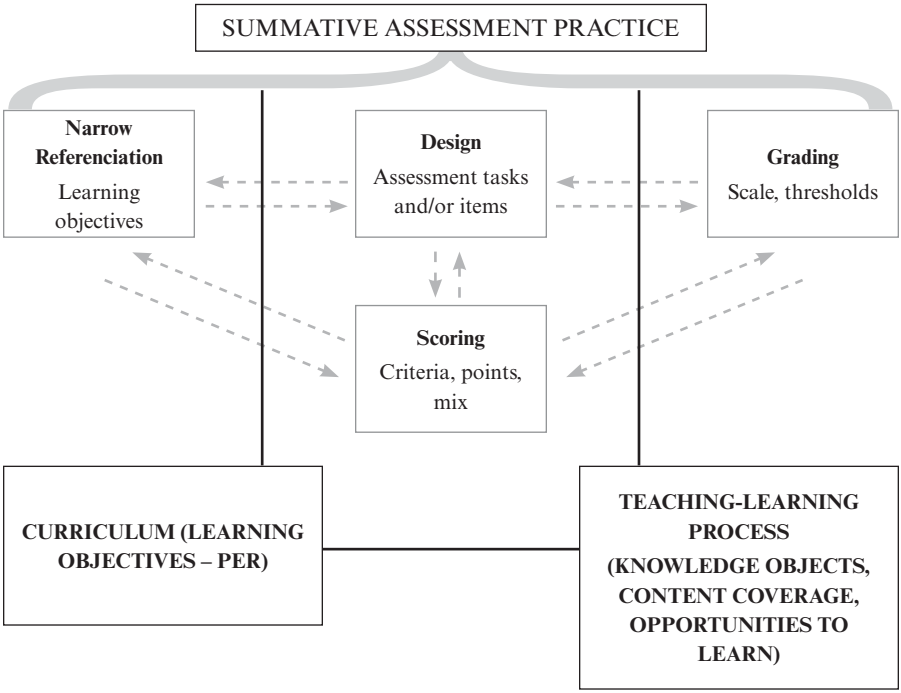


Figure 6. Expanded curricular alignment (Pasquini, 2018)

The model, conceptualized in this way, lies at the heart of the initial model, in which the teaching-learning processes, the objectives set out in the curriculum and the assessment practices interact coherently, and we are intentionally focusing on their certification function, i.e., the summative assessment (solid dark lines in Figure 6). The latter then displays its four components (light-coloured bracket) by linking them together (light-coloured double dotted lines). Thus, the narrow referenciation, inspired by Vial (2012), describes "what is believed" by teachers when choosing which learning objectives to assess from a summative perspective. The "design" (Wiggins & McTighe, 2005) explains the processes involved in constructing assessment tasks, both in terms of the knowledge objects that circulate in them and the cognitive abilities that students must mobilize, as well as the nature of the guidelines. Scoring refers to the weight assigned to learning and to the tools used for this purpose (criteria, points or a mixed method of criteria matched to points; Andrade, 2005; Jonsson, 2014). Finally, grading focuses on approaches to constructing the numerical grade as it relates to learning through tools such as scales (Brookhart, 2017; Marzano & Heflebower, 2011). All of these components should have coherent relationships from a taxonomic and content standpoint.

Thus, in terms of our example, we can see that Paul's summative assessment practices ensure coherence between the learning objectives, tasks and criteria, and propose an attempt at grading linked to the learning assessed, even if this remains difficult to implement. We can now question how the weight assigned to a given learning item in the exercises corresponds to the complexity of the latter, with regard to the mathematics subjects and the objectives they assess. In a complementary and non-exhaustive way, we can question the relevance of Paul's decision to use, *in fine*, a standardized scale to construct grades; we do this knowing that tasks with little reference to the learning objects result in a "poor alignment" (Biggs and Tang, 2011, p. 163), which is unlikely to allow for scoring and grading in reference to these same objects.

In light of these considerations, which bring together research findings and elements of practice, we observe that in the new systemic approach, the curricular coherence in the summative assessment is the result of a triangulation between several components of the model, until grading. In this sense, expanded curricular alignment refers, in summative assessment,

to the strong link of systemic coherence that exists between the objectives assessed (narrow referenciation), the assessment tasks (design), the points and/or criteria assigned (scoring), and the scale (grading) (Pasquini, 2018).

Discussion

Our discussion, as it relates to our research question, is organized into three parts. First, we substantiate the tension between Paul's assessment practices and the expanded theoretical model, thereby further highlighting some of the potential obstacles to implementing coherence in practices until grading. Then, we briefly refer to other findings in order to put certain unresolved issues into perspective. Lastly, we reflect upon the expanded model.

The tension is still there

At the end of the process, Paul states, regarding curricular alignment:

> "It helped me become more aware that grading, choosing a scale, is, well, a conundrum. It can undo all the work that I've gone to the trouble of doing previously to transpose the knowledge and choose coherent objectives. All of that work can just get wiped out (laughs)...by choosing the wrong scale."

Due to the impossibility of thinking differently about this coherence in grading, he thus assumed that he was forced to lose it; but he could live with this situation since he had significantly improved his practice in a general sense. In his words, "I think that by improving the design and hence the quality of my assessments, the problem [of grading] should not exist."

Thus, Paul still had difficulty accurately understanding the model up to the point of grading: "For me, curricular alignment stops with the assessment design. After that, when you assign a grade, when you choose the scale, it's no longer part of the curricular alignment" (Translator's translation). We hypothesize that this way of understanding the model, at the very end of the research, accounts for the trade-offs that teachers make all the time when they evaluate and that are by nature profoundly dynamic (Carless, 2011). This doesn't invalidates the relevance of the model, as we will argue in our conclusion.

A comparison with other results

The findings exemplified through Paul's situation form part of more transversal results that further enhance this new, expanded and systemic conceptualization of the model, while simultaneously highlighting areas of tension (Mottier Lopez & Pasquini, 2017).

For example, comparing the model with teachers' habitual practices has an impact on several dimensions of these practices. Thus, it facilitates (i) a deeper understanding of the curriculum, allowing for a better prioritization and expression of the learning objectives to be taught and assessed; (ii) greater homogeneity in the complexity of the assessment tasks (Guskey, 2006); (iii) scoring using points, criteria or mixed modality with reference to the objectives; and (iv) a reconsideration of certain teaching processes, something that would validate our new systemic understanding of the model.

However, it should be noted that Paul is not the only teacher to have difficulties in maintaining curricular coherence up to grading. The three other mathematics teachers also encountered obstacles when they had to construct a grade using criteria and, in the end, had to resort to scales that did not take into account the assessed learning. The same was true for two French teachers with regard to language structure and written comprehension tests.

A final reflection on the expanded model

We believe that these results strengthen Biggs' (2003) contribution regarding the high potential for misalignment of tools such as standardized scales that are external to actual learning or broad grading scales. We therefore hypothesize that this way of modelling curricular coherence in summative assessment provides a clearer understanding of the causes of these misalignments. Indeed, for a majority of the participants, the impossibility of constructing a learning-related grade - even if the coherence of the expanded curricular alignment is noticeable all the way to the scoring - underlines the impact of the grading tools on summative assessment practices.

It is also important to remember the role that context plays in assessment practices. The culture of the educational institution, its leadership, teachers' habits and teamwork methods, as well as its rules and regulations, are determining factors that greatly influence the implementation of different and even innovative assessment practices (Timperley & Parr, 2009). One of the limitations of our research is that we did not adequately thematize how the interactions between these environments and evolving practices might impact the model. We return to this point at the end of the article.

Conclusion

Overall, and based on the example provided, our results highlight the value of using the expanded model as a comprehensive tool for summative assessment practices, thereby helping us to better understand their complexity. As such, our research demonstrates that examining the coherence of summative assessment practices is relevant. Conceptualized in this light, the quest for coherence resembles the search for validity, a central concept in assessment. Bonner (2013) supports this view when she describes the characteristics of any valid assessment process: strong coherence in the curricular alignment between the teaching, learning and assessment processes; maximum avoidance of biases during the assessment process; judging the relevance of the results obtained; and perceptible process transparency for all actors involved.

However, it is also important to examine the relationships that could exist between effective practices and a theoretical model: under what conditions could the model remain comprehensive? To what extent does its use, notwithstanding - and over and above the intentions of the research - introduce a prescriptive dimension to the practice, as might be inferred from the compromises made by Paul and certain of his colleagues regarding their inability to grade on the basis of criteria? And what role does the research or training scheme play in this connection?

Furthermore, and as briefly mentioned, the systemic coherence that our model calls for is eminently contextual. Its conceptualization cannot therefore be generalized, since it depends to a large extent on the nature of the curriculum, the characteristics of the discipline and the grading used. In this regard, our results support the hypothesis that the problem is not so much the existence of a grade as the way it is constructed and used (Brookhart, 2017; Walvoord & Johnson Anderson, 2009) and that, under certain conditions, a grade can be constructive and highly informative (Hadji, 2016). But to what extent can this expanded theoretical model coexist with the assessment practices present in various school environments? Under what conditions can it interact with the specific features of various curricula, with the multiple frames of reference teachers use daily, and with the tools proposed for them - or even imposed on them – for grading?

It would be interesting to lead a similar study with teachers who are asked to use expanded ranges of grades (e.g., 1 through 20) or who are required to use lettered grading scales with specific ordinal logics.

Furthermore, how would this coherence be intelligible in a skill-based approach whose epistemology is very different from ours? And, based on the observation that disciplines and subjects play a leading role in summative assessment approaches, deeper analysis could explore the extent to which some contents allow for the construction of curricular coherence up to the point of grading (e.g., in written production), while others seem to pose significant obstacles to such practices, as in the case of Paul.

As Moss (2013) points out, future studies examining effective summative assessment practices in a variety of contexts is urgently needed. The challenge is twofold: to produce knowledge about the coherence of these practices up to the grading processes, and to improve teachers' skills and knowledge in the field of summative assessment which, to date, have been woefully inadequate.

Received: 11 December 2018

Final Draft: 5 September 2019

Acceptance: 6 September 2019

NOTES

1. This article is based on completed doctoral research.
2. Using this inductive-deductive approach, we created five categories: relationship to the curriculum, definition of the learning object, development of assessment tasks, scoring practices and grading practices.
3. Alias
4. www.plandetudes.ch/per

REFERENCES

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; French version available). Paris: OECD-CERI.

Anadon, M. & Guillemette, F. (2007). La recherche qualitative est-elle nécessairement inductive? *Recherches qualitatives*, 5, 26-37. Retrieved from 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: 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. & 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 ?* Brussels: De Boeck.

Denzin, N. K. (1978). *The research act: A theoretical introduction to sociological methods*.

New York, NY: McGraw 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. & 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. & Richard, M. (2005). *Écoles efficaces et réussite scolaire des élèves à risque : un état de la recherche* (Research report). 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 ? In B. Boquet (Ed.). *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. In A. Bentolila (Ed.). *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. & 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). *Evaluations formative et certificative des apprentissages*. Brussels: De Boeck.

Mottier Lopez, L. & 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 sommative 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. & 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. Retrieved from https://journals.openedition.org/questionsvives/1235

Paillé, P. & 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. Retrieved from https://orfee.hepl.ch/handle/20.500.12162/2423

Pasquini, R. (2017). Évaluation : première question vive de la pratique. *L'Éducateur*, *4*, 38-39. Retrieved from 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* (Unpuplished doctoral thesis). University of Geneva, Geneva. Retrieved from 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.

Sadler, R. 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. & 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), 103-121. Retrieved from http://usherbrooke.crifpe.ca/wp-content/uploads/2016/05/vnh-colo-jiref_2015.pdf

Vial, M. (2012). *Se repérer dans les modèles de l'évaluation*. Brussels: 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.

Winger, T. (2009). Grading what matters. *Educational Leadership*, *67*(3), 73.