

**The Collective Work of Teachers for the Assessment
of Learning as a Professional Norm? A Literature
Review to Examine this Emerging Trend¹**

**Le travail collectif des enseignants pour l'évaluation
des apprentissages comme norme professionnelle?
Une revue de la littérature pour interroger
cette tendance émergente**

**O trabalho coletivo dos professores para a avaliação
das aprendizagens como norma profissional?
Uma revisão da literatura para questionar essa
tendência emergente**

Jean-Pascal Ochelen

ID ORCID: 0009-0004-9860-1142

University of Fribourg & University of Geneva (Switzerland)

Gonzague Yerly

ID ORCID: 0000-0002-5264-9720

University of Fribourg

Lucie Mottier Lopez

ID ORCID: 0000-0002-9374-7300

University of Geneva

1 The French version was published in issue 47(1) 2024: <https://doi.org/10.7202/1114565ar>



KEY WORDS: assessment, collaboration, teachers, collaborative assessment work, review of literature

This paper presents a rapid review of the Anglophone and Francophone literature on collaboration among teachers to assess student learning. This kind of collective work tends to become a norm in education systems, raising major issues in the field of educational assessment. The literature already contains summaries of empirical research dealing with teacher collaboration in general, but none dealing specifically with collective assessment practices. A systematic literature search led to the selection of 30 articles, in primary and secondary education. Analysis of these articles allows us to 1) identify the dynamics at work in the schools and 2) observe the consequences for all those involved in the school. The results allow us to discuss certain tensions linked to this trend, from the point of view of steering policies, but also about individual and collective evaluation practices.

MOTS CLÉS : enseignants, évaluation des apprentissages, revue de littérature, travail

L'article présente une rapid review de la littérature anglophone et francophone sur le travail collectif des enseignants pour l'évaluation des apprentissages. Ce genre de travail collectif tend à devenir une norme dans les systèmes éducatifs, soulevant des enjeux majeurs dans le champ de l'évaluation en éducation. La littérature compte déjà des synthèses et des recherches empiriques qui traitent de la collaboration des enseignants de manière générale, mais aucune ne traite spécifiquement des pratiques évaluatives collectives. Cette recherche documentaire dans la littérature conduit à la sélection de 30 articles, aux niveaux primaire et secondaire. L'analyse de ces articles permet 1) d'identifier les dynamiques mises en place dans les établissements et 2) d'observer les conséquences pour tous les acteurs de l'école. Les résultats permettent de discuter de certaines tensions liées à cette tendance, au point de vue des politiques de pilotage mais également sur le plan des pratiques évaluatives individuelles et collectives.

Author's note: All correspondence concerning this article should be addressed to: jean-pascal.ochelen@unifr.ch

This article contributes to a doctoral thesis conducted as part of a broader research project funded by the Swiss National Science Foundation (SNSF project 100019_207804, Yerly, 2022-2026)

PALAVRAS CHAVE: professores, avaliação das aprendizagens, revisão de literatura, trabalho coletivo

O artigo apresenta uma rapid review da literatura anglófona e francófona sobre o trabalho coletivo dos professores para a avaliação das aprendizagens. Este tipo de trabalho coletivo tende a tornar-se uma norma nos sistemas educativos, levantando questões importantes no campo da avaliação em educação. A literatura já inclui sínteses e investigações empíricas que abordam a colaboração entre professores de forma geral, mas nenhuma trata especificamente das práticas avaliativas coletivas. Esta investigação documental na literatura resulta na seleção de 30 artigos, nos níveis primário e secundário. A análise destes artigos permite 1) identificar as dinâmicas implementadas nas instituições e 2) observar as consequências para todos os atores da escola. Os resultados possibilitam discutir algumas tensões associadas a esta tendência, tanto do ponto de vista das políticas de gestão como no que respeita às práticas avaliativas individuais e coletivas.

Introduction

Over the past three decades, collaborative work has become a normative expectation in education (Borges & Lessard, 2007; Vangrieken et al., 2015). The concept of collaborative assessment was first introduced in anglophone literature by Hargreaves (2007), who examined collaboration between teachers and students in England, then Australia, addressing collaborative processes of social moderation among teachers (Adie et al., 2012; Wyatt-Smith et al., 2010). Collaborative work for assessment is not new (Adie et al., 2012; Black et al., 2011; Laveault & Yerly, 2017; Mottier Lopez et al., 2012), formal institutionalization of the process is a more recent development. These policy initiatives aim to enhance the quality of assessment outcomes by harmonizing teacher assessment practices in schools (Laveault & Yerly, 2017).

A substantial body of research, conducted over more than three decades, has examined collaborative work among teachers (Hargreaves, 1994, 2021; Vangrieken et al., 2015). The literature suggests that we now have a solid understanding of the effects of collaborative work in general. However, Yerly (2021) notes there is less work focusing on collaborative work for assessment practices. This article therefore addresses the following questions: What is the current state of research knowledge on teachers'

collaborative work on assessing student learning? What conclusions can be drawn from empirical studies conducted over the past two decades on this topic? To address these questions, we conducted a rapid review, a type of literature review that combines certain advantages of systematic reviews with a more expedited article selection process (Grant & Booth, 2009, cited in Dachet, 2024). The review draws on both anglophone and francophone educational literature. This novel study aims to shed light on a blind spot in both research literature and the field of educational assessment.

Context and Conceptual Framework

Teachers' collaborative work: a contemporary professional norm

Since the 1980s and policy efforts to decentralize schooling, schools have increasingly emerged as distinct entities and as key components of the education system (Maroy, 2005). According to Meirieu (2023), this renewed focus on schools has been combined with increased institutional autonomy. Schools have greater latitude for defining their institutional projects and objectives. Within this configuration, outputs, particularly students' results on standardized assessments, are the main tools governing the system.

Whether at the international level (OECD, European Union, UNESCO) or at the national level (policy circulars and decrees), calls for collective work among teachers appear largely consensual (Gilbert, 2018). Teachers are now expected to be effective collaborators. While teamwork is becoming increasingly prominent across society generally, teachers' collective work is also emerging as a professional norm in education (Decuyper et al., 2010). However, research examining how teachers translate directives to work collectively reveals considerable variation in how such expectations are applied across schools. Spillane et al. (2016) argue that the degree of integration between policy directives, collective practices, and teaching practices tends to be stronger in countries that have adopted robust accountability policies, such as the United States or England (Maroy & Voisin, 2014). Nevertheless, the norm surrounding teamwork remains relatively ambiguous, and multiple models coexist (Gilbert, 2018). Hence, collective practices may take different forms.

Modalities of collective work: from informal exchanges to learning communities

In their work, Marcel et al. (2007) distinguished three modes of collective work according to the intensity of task sharing: coordination, collaboration, and cooperation. According to Grangeat (2011), collective work modes (collaboration, distributed cooperation, co-action) mainly depend on the objectives pursued. Hargreaves (2019) distinguished between flexible, democratic arrangements (collaborative cultures) and formal, highly institutionalized arrangements (contrived collegiality). Vangrieken et al.'s (2015) literature review synthesized the modalities documented in the anglophone literature and clarified the associated terminology.

According to Barrère et al. (2008), informal interaction, the minimal level of collaboration, is mobilized constantly, often unconsciously. For example, exchanges around the coffee machine or during meals are not trivial. Through such interactions, teachers build ties, even when they do not directly serve their work (Letor, 2010). These ties foster a sense of collective belonging and constitute a basis for professional cooperation.

Engaging in team-based work is more formal. According to Cohen and Bailey (1997), this type of work is characterized by task interdependence, with teachers sharing responsibility for outcomes. Salas et al. (2000) add that teams are defined by their own structure and by a specified duration. In some cases, teacher teams reinforce formalized collective work. Literature has therefore developed different models, both dynamic and structured, to conceptualize teacher collectives. We present three briefly: the professional learning community (PLC), the teacher learning community (TLC), and the community of practice (CoP). According to Gilbert (2018), PLC refers to “a group of people who share a collaborative learning process within a reflective practice for a common, time-bound goal” (p. 5). Exchanges and learning are horizontal, with members sharing a set of norms, values, and orientations regarding teaching practices (Vangrieken et al., 2015). For Wenger (2005), a CoP develops within groups of practitioners working together to address problems encountered in their professional practice who develop a shared repertoire of resources for shared practice. When teachers form a community of practice, they exchange ideas and adopt shared terms for phenomena they previously interpreted or identified differently. They mutually engage in the joint enterprise that brings them together. In a TLC, teachers work collectively in a single content area to reflect on their practices, examine student outcomes, and implement

changes aimed at improving the quality of teaching and learning at their school (McLaughlin & Talbert, 2006). Across these different communities, learning is not limited to the individual. Influenced by the environment, learning is understood as situated, rooted in social and cultural practice (Mottier Lopez, 2008, 2021). Interaction and collaborative relationships are considered crucial for developing individual and collective knowledge.

The role of negotiation

According to Wenger (2005), practice within communities of practice is continuously negotiated, particularly for the construction of a shared repertoire. Many authors have examined the role of negotiation in collaborative work arrangements. Strauss (1992) distinguished between two types of negotiation: conflictual negotiation and integrative negotiation. The former involves a confrontation of viewpoints and aggressive tactics, whereas the latter is characterized by cooperation and an intention to reach a mutually beneficial agreement. Beuret (2010) explained that parties in conflict may move from one type of negotiation to another by identifying a “transactional passage point” (p. 45), such as mediation. Allenbach (2016) specified that processes of agreement-building through the negotiation of meaning and alliances are crucial factors for the success of collective work. In their empirical study, Germier and Marcel (2016) also showed that negotiation is important for building agreements within teacher collectives. In summary, negotiation plays an important role in structuring collective work and in shaping how it operates.

The state of knowledge on teachers’ collective work

The most recent review of the literature we identified is Vangrieken et al.’s (2015) systematic review of anglophone literature featuring 82 empirical articles analyzed from 2000 onwards. This review showed that collective work is organized in different ways and that the terminology used can be confusing. Terms such as collaboration, cooperation, coordination, shared work, group work, and teamwork are often used interchangeably. The intensity of collaboration varies considerably, including within the same modality of work. The review highlighted the benefits of collective work for students, schools, but especially teachers (e.g., time saving, professional development, efficiency, and engagement). Certain limitations and disadvantages were also identified (e.g., tensions, deterioration of interpersonal relationships, social loafing, and lost autonomy). Hattie’s (2017) work on teachers’ collective effectiveness showed that the effect

of teachers' collective work on student learning ($d = 1.57$) is double the effect of feedback ($d = 0.70$) and nearly triple the effect of classroom management ($d = 0.62$). Teachers' collective work therefore appears to have a strong impact on student learning.

Collective assessment: a situated and social practice

Considering emerging trends in the research literature, for this paper, we approached assessment as a collective practice from a situated perspective (Lave, 1988; Wenger, 2005). We focus on teachers' professional judgement and related issues concerning actions and operations constitutive of assessment, regardless of its function (Allal & Mottier Lopez, 2009; Wyatt-Smith et al., 2010). For Mottier Lopez and Allal (2008), professional judgement from a situated perspective is both a cognitive act and a social practice grounded in norms, tools, languages, and cultures, with a specific social context. This definition reflects the complex nature of student assessment. We conceptualized professional judgement in assessment in the collective context of teachers' work by drawing on the three main conditions developed by Wenger to define a community of practice: joint work that creates the conditions for all actors involved to engage mutually and actively (mutual engagement); the pursuit and achievement of shared goals in relation to a common project (joint enterprise); and the construction and negotiation of meaning, tools, actions, and effective and responsible ways of working to achieve fair and accurate assessment (shared repertoire).

Research questions and specific objectives

Given the agenda of international education policies and pedagogical orientations increasingly encouraging collective assessment practices, our literature review aimed to provide an overview of the state of scientific knowledge on teachers' collective assessment practices. Considering the potential issues at the micro-, meso-, and macro-levels, their implementation and effects may have significant consequences for all stakeholders¹.

1. The micro, meso, and macro levels refer to the categorization proposed by De Ketele (2006). The author situates the micro-level within teaching and learning processes in a classroom or training group (in this case, students and teachers). The meso-level is located at the level of a school, educational institution, or training institute (in this case, the school). The macro-level corresponds to the level of an education or training system (in this case, the education system).

Given that assessing student learning is becoming an increasingly collective rather than individual practice (Yerly, 2021), our study aimed to address the following four general research questions:

1. Why do teachers collaborate to assess student learning?
2. What tools and methods do teachers use when assessing student learning together?
3. Which dimensions of assessment do teachers work on collectively?
4. What consequences does this collaborative work have for students and teachers, for schools, and for the education system?

To address these questions, we used the four following categories for analyzing the content of the articles:

1. Identify the objectives of collective work arrangements employed by teachers to assess student learning.
2. List the tools and methods used by teachers when they work together to assess student learning.
3. Identify objects of joint attention when teachers assess student learning collectively.
4. Highlight the effects of collective work on assessing student learning for students and teachers, for schools, and for the education system.

Methodology

Research design

We opted for a rapid review of the literature (Grant & Booth, 2009, cited in Dachet, 2024) to ensure a rigorous study within our time constraints. According to Ganann et al. (2010, cited in Dachet, 2024), this method combines a range of approaches used to accelerate and optimize the systematic literature review process, with the aim to save both time and human resources. Our methodology incorporated certain procedures typical of systematic reviews, including clearly defined research questions, the use of a replicable search protocol, explicit objectives, and well-defined criteria. However, we made certain methodological concessions compared with a full systematic review: the absence of an independent dual-reviewer decision process for article inclusion or exclusion; a reduced number of

databases; and no statistical analysis. Consequently, we conducted a narrative analysis and presentation of the results rather than synthesizing findings using statistical or quantitative methods.

Literature search procedure

We collected scientific articles from both anglophone and francophone literature in May and June 2023. For the anglophone literature, we first searched the ERIC database, which is specific to education research. To optimize the comprehensiveness of the study, we also used APA PsycInfo to cover related fields: psychology, social sciences, and humanities. For the francophone literature, we used the search platforms HAL Open Science, Erudit, and CAIRN, which are known for their rigorous and selective curation of academic publications. To initiate database searches, we drew on several foundational texts aligned with our topic and research criteria (Adie, 2014; Allal & Mottier Lopez, 2014; Heredia et al., 2016; Mottier Lopez & Pasquini, 2017) which helped identify the keywords for thesaurus searches.

For ERIC and PsycInfo, search streams were constructed using the index terms drawn from these thesauri. We also included additional keywords from free-text vocabulary (non-standardized terms used to describe article content). Table 1 presents the index terms and free-text keywords entered into the thesauri. Depending on the language, coding systems, and conventions specific to each database, all terms were then used to conduct searches across titles, abstracts, and full texts. Given the absence of thesauri for the French-language search platforms, we used translated keywords drawn from the free-text vocabulary.

To ensure the quality of our review, we applied the following inclusion criteria to our article selection:

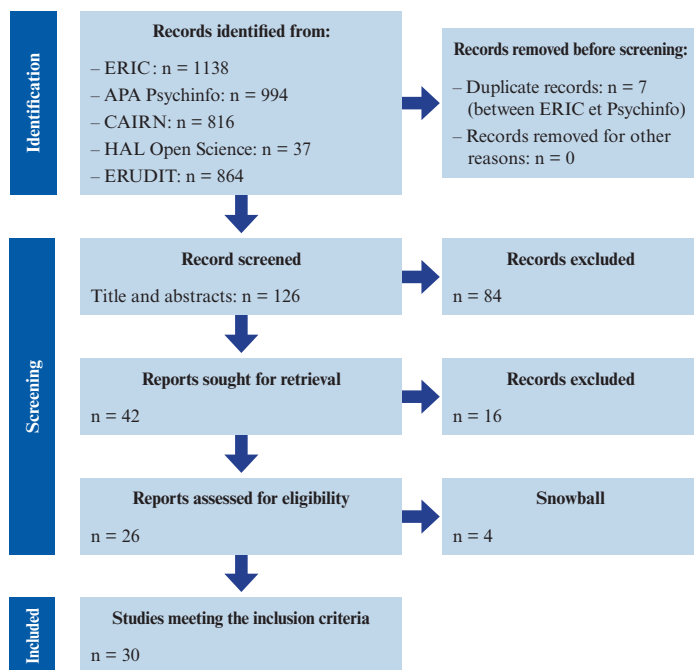
- written in English or French;
- published in peer-reviewed scientific journals;
- published from 2003 onwards;
- empirical studies;
- address a form of collective work among teachers specifically related to assessing student learning;
- concern school levels between Grades 1 and 12.

Table 1
Free-text vocabulary and index terms from the PsycInfo and ERIC thesaurus

	Concept 1: Assessing student learning	Concept 2: Teachers' collective work
Free-text vocabulary	assessment, evaluation, grading, social moderation, teachers' professional development	community of practice, cooperation, teacher collaboration, shared work
Index terms from the PsycInfo thesaurus	educational measurement, evaluation criteria, formative assessment, grading (educational)	collaboration, cooperation, teamwork
Index terms from the ERIC thesaurus	accountability, educational assessment, educational policy, educational practices, formative assessment, grading, interrater reliability, student evaluation, summative assessment, professional development	collaboration, collegiality, teacher collaboration, cooperation, cooperative plannings, peer relationship, teamwork, teacher attitudes

Although relevant, literature on special education was excluded, as it involves different dynamics and external paraprofessional actors that fall outside the scope of our study. We uniformly applied the various filters corresponding to the criteria across all search platforms. Among the articles retrieved, we conducted an initial screening based on article titles, excluding articles irrelevant to our study or not meeting our selection criteria. The second screening, based on abstracts, was conducted in the same way. For the final step, we read the entire texts. Finally, using the snowballing technique (Vangrieken et al., 2015), we analyzed the reference lists of the selected articles to identify additional sources not captured in the consulted databases. The same criteria were applied to the snowballing process. This manual search was conducted using Google Scholar and Taylor & Francis. Figure 1 below presents the flowchart, inspired by the PRISMA model (Page et al., 2021), illustrating the selection process and the number of articles included.

Figure 1
Flowchart based on the PRISMA model, illustrating the studies selection process



Data analysis

To analyze their content, we read, categorized, and organized the selected articles using a review matrix as proposed by Dumez (2011, adapted from Hart, 2009). The matrix included publication dates, authors, titles, abstracts, countries of origin, types of research, methodologies, and school levels. It was subsequently expanded to include four additional categories linked to our research questions and analytical framework: the objectives of collective work; the assessment-related objects addressed; the tools and methods used; and, finally, the effects. As the reading progressed, we developed a synthesis table to inventory the general characteristics of the selected studies. This approach facilitated sorting, grouping, and an “open, exploratory rereading of the assembled theoretical material” (Dumez, 2011, p. 24).

Initial general findings of our literature review

Our literature review process resulted in the selection of 30 empirical articles published between 2007 and 2022. Table 2 presents the corpus of selected articles². The breakdown is as follows: 27 qualitative studies, two quantitative studies, and one mixed-methods study. For the qualitative studies, all research designs focus on specific cases and use the following methodological tools: semi-structured individual interviews, focus groups, team observation, and document analysis. Limited information is available on qualitative data analysis procedures. Only two studies specify the use of NVivo software³. For the quantitative studies, no analytical software is reported. The most commonly used statistical tools include frequency counts, correlations, means, standard deviations, and linear regressions.

Notably, the countries represented in the 30 articles are federal states where authority over education lies with local jurisdictions (Australia, New Zealand, the United States, the United Kingdom, and Switzerland) (n = 26), or unitary states that have implemented reforms to increase decentralization of their education systems (Norway, Sweden, and the Netherlands) (n = 4). Most of the articles are written in English (n = 27), with the remainder in French (n = 3).

In our sample of articles, collective work is most frequently organized horizontally i.e., among teachers working at the same grade level: in secondary education (n = 19) and primary education (n = 5) than vertically (between primary and secondary teachers) (n = 5). Negotiation is simpler in a horizontal configuration which is more homogeneous. Vertical arrangements involve the meeting of two worlds and therefore require a more extended and complex process. These findings align with those of Doppenberg et al. (2012), who reported that teachers prefer working with colleagues at the same level.

In the 19 studies conducted at the secondary level, collective work is always organized among teachers from the same subject (or subject area). These findings diverge from the results of Vangrieken et al.'s (2015) literature review, which reported that collaboration in secondary education mainly occurs among teachers from different subject areas. In most of the selected studies, teacher profiles were heterogeneous (gender,

2. A full version of this table is presented in Appendix A.

3. Smail (2020) and McFadden et al. (2022).

experience, and discipline). However, in terms of seniority, two studies (Allal & Mottier Lopez, 2014; Matre & Solheim, 2016) focused exclusively on experienced teachers.

Results

In the remainder of this article, we present the results of our content analysis of the 30 selected studies. The findings are organized according to the structure of our analytical categories.

The purposes of teacher collaboration for assessing student learning

This section addresses the following question: Why do teachers collaborate to assess student learning? Table 2 presents the purposes identified in the 30 studies. We also indicate the authors, publication dates, countries, and, where applicable, the local authorities responsible for education.

The most frequently reported purpose is social moderation of assessment results ($n = 20$). According to Gipps (1994), social moderation refers to teachers working together to discuss and compare the grades assigned to students to reach consensus and develop a shared understanding of the quality of student work. It involves confronting viewpoints, interpretations, and professional judgements regarding students' work (Linn, 1993). In these 20 studies, social moderation is sometimes mobilized for accountability ($n = 6$). In such cases, teacher collaboration is used to improve the quality of assessments ($n = 3$) or to monitor the practical implementation of reforms ($n = 3$). Social moderation is also mobilized for improvement purposes ($n = 14$), to enhance professional judgement ($n = 6$); develop competence in summative ($n = 1$) or formative assessment ($n = 3$); examine negotiation processes in grading ($n = 1$); or improve student learning ($n = 1$). It is also used to establish communities of practice and foster vertical collaboration ($n = 2$) to facilitate exchanges between primary and secondary education actors.

In some studies, teachers are also encouraged to collaborate when conducting assessments, to collectively build knowledge related to assessment for learning (AfL) ($n = 4$). Collaborative work also promotes exchanges among teachers for collectively designing assessments for formative ($n = 3$) or summative purposes ($n = 2$). Finally, teachers also collaborate when assessing to implement pertinent recommendations issued by local ($n = 2$) or federal authorities ($n = 1$).

Table 2
Full corpus of articles: purposes of collaborative work, authors, year of publication, and counties of origin

Purposes of collaborative work	Authors (year)	Countries (local jurisdictions)
Social moderation (n = 20)	Adie (2013)	AUS (Queensland)
	Adie (2014)	AUS (Queensland)
	Adie et al. (2012)	AUS (Queensland)
	Allal and Mottier Lopez (2014)	CH (Genève)
	Black et al. (2011)	GB
	Colbert et al. (2012)	AUS (Queensland)
	Connolly et al. (2012)	AUS (Queensland)
	Crisp (2017)	GB
	Grant (2012)	GB (SCT)
	Hipkins and Robertson (2012)	NZ
	Klenowski et al (2007)	AUS (Queensland)
	Matre and Solheim (2016)	NO
	Morales Villabona and Mottier Lopez (2016)	CH (Genève)
	Mottier Lopez and Pasquini (2017)	CH (Geneva and Vaud)
	Mottier Lopez et al. (2012)	CH (Geneva)
	Nilsberth and Sandlund (2020)	SE
Reid (2007)	GB (SCT)	
Smail (2020)	NZ	
Van Der Schaaf et al. (2012)	NL	
Wyatt Smith et al. (2010)	AUS (Queensland)	
Development of Afl practices (n=4)	Furtak et al. (2016)	USA
	Hargreaves (2013)	GB
	Hargreaves et al. (2012)	GB et HKG
	Stewart and Houchens (2014)	USA (Kentucky)
Development of tools (n = 5)	Heredia et al. (2016)	USA
	Mc Fadden et al. (2022)	USA (Kentucky)
	Buckley and Liscompe (2022)	AUS (New South Wales)
	Mottier Lopez and Pasquini (2017)	CH (Vaud)
	Morales Villabona and Mottier Lopez (2016)	CH (Geneva)
Implementation of recommendations (n = 3)	Hemansen (2014)	NO
	Hermansen & Nerland (2014)	NO
	Yerly (2021)	CH (Fribourg)

Tools and methods used in teacher collaboration for assessing student learning

The literature documents a wide range of tools and methods used to achieve the objectives outlined above. This section addresses the following question: What tools and methods do teachers use when assessing student learning together?

When collaboratively designing assessments, teachers used shared design guidelines for the procedures to be implemented ($n = 10$). These guidelines were either developed by authorities ($n = 7$) or already produced collectively by teachers ($n = 3$). Teachers also drew on personal resources and used their own assessments as models ($n = 3$). Curricula was likewise mobilized in the joint design of assessments ($n = 4$). Mottier Lopez et al. (2012) further observed that the primary concern of the teams in their study was to ensure curricular alignment between assessment, teaching practices, and learning objectives. To construct shared assessment tools, teachers also referred to examples of previously administered external assessments ($n = 3$).

When collaboration was initiated by researchers ($n = 3$), the development of assessment tools was also guided by a structured scenario ($n = 2$) or informed by a theoretical model ($n = 1$). In some cases, technological tools were used, including communication platforms with audio functions that allow geographically distant participants to interact ($n = 3$). In Australia, online moderation sessions were organized. Technological tools were also used in face-to-face self-observation sessions ($n = 1$), when members of a teacher community jointly reviewed how they implemented AfL practices previously developed collectively in their classrooms.

To compare the quality of interpretations of student work, the most frequent tools in our corpus were samples of previously graded assessments ($n = 22$). These assessments included both formative ($n = 7$) and summative tests, with or without certification purposes ($n = 15$). Assessment frameworks were another commonly used tool in collective sessions ($n = 17$). These documents contained the standards, criteria, and marking guides to be applied and served as reference points for building consensus around student work.

The methods teachers used varied across the 30 studies. In most studies, regular meetings were scheduled to enable collaborative work ($n = 19$), and were generally face-to-face ($n = 16$), and only sometimes

remote ($n = 3$). They were built into teachers' schedules and formed part of formal assessment procedures, often mandated by school leadership. The number and frequency of meetings depended on the purposes of collaboration. For instance, the Australian studies ($n = 8$) were all in the context of major curriculum and assessment reforms, particularly through the establishment and implementation of standards (Klenowski et al., 2007). Social moderation procedures were standardized and followed specifications set by national authorities, with meeting number and frequency aligned with assessment cycles. When procedures were not standardized, the regularity of collective sessions depended on team-specific organization ($n = 2$).

In some cases, teachers were encouraged to collaborate through government-funded projects ($n = 4$) or privately funded initiatives ($n = 2$). The grants and investments provided generated more or less explicit forms of policy pressure, with clearly stated objectives.

Several studies also reported that teachers are brought together by participation in professional networks ($n = 12$). Drawing on sociocultural learning theories, several studies ($n = 6$) mobilized Wenger's (1998, 2005) concept of communities of practice. Professional learning communities ($n = 1$) and teacher learning communities ($n = 5$) were also used, where collaborative work was sometimes organized through workshops ($n = 2$) or followed a predefined operational sequence ($n = 3$).

Finally, collaborative assessment work was also organized for research purposes. In such cases, sessions were held on a more occasional basis ($n = 6$), typically over one or two days as part of research seminars ($n = 3$), continuing professional development activities ($n = 2$), or exploratory research projects ($n = 1$).

Shared objects of attention in collaborative assessment

This section addresses the following question: Which dimensions of assessment do teachers work on collectively? We identify the objects of assessment – in other words, the dimensions - shaped by teachers' collaborative work. This section is organized into two parts: first, assessment, then, marking.

Collaborative work in assessment design

In 13 studies, collaborative work focused on the assessment design phase. This mainly concerns the referential framework of assessment, i.e., the set of guidelines defining how assessment should be designed, implemented, and interpreted. For example, teachers collectively discussed the learning objectives to be assessed, criteria, achievement thresholds, weightings, marking schemes, standards, and assessment tasks.

In these discussions, leadership was generally shared among teachers ($n = 9$), and in rarer cases, by members of school leadership (principals, deputy principals, heads of department) ($n = 2$), or external actors such as researchers ($n = 2$). Several authors reported that teachers experience difficulties in developing a shared understanding of assessment referents (e.g., objectives and criteria) ($n = 3$). In one study, consensus and shared expectations emerged through deliberate efforts to make professional discourse more transparent. Several authors also highlighted the constructive role of professional controversy in the collective design of assessments ($n = 2$). In some studies, discussions focused exclusively on developing a shared understanding and consistent application of criteria ($n = 3$). In certain cases, frameworks were treated as prescriptions to be applied in assessment, and collaboration leads to adjustments within these frameworks. In Colbert et al. (2012), for instance, such procedures consisted of four checkpoints to ensure assessment quality. These checkpoints were collectively reviewed by teachers and revised in light of the results. In the Norwegian study of Hermansen and Nerland (2014), the procedures concerned the implementation of assessment for learning practices. Through collaborative work, teachers revisited historically established principles, negotiating and recontextualizing them to develop new shared AfL practices.

Collaborative work in marking

The analysis of the article corpus showed that in most studies ($n = 17$), collaboration occurred during the marking of student work or immediately afterwards. Grading was the primary object addressed. In several articles, the final grades assigned to internal assessments were discussed and, where necessary, revised ($n = 9$). In other studies, collective discussions focused on grades awarded in external, standardized, high-stake assessments for certification ($n = 3$). In some cases, collaboration focused on the content of students' completed tasks ($n = 5$). Interpretations of evidence of learning were discussed and compared using portfolios ($n = 3$) or directly through

tasks in internal assessments ($n = 2$). In several studies, during collective sessions, teachers reported adopting a comprehensive (flexible) approach to criteria ($n = 5$). They justified this choice by arguing that a strictly instrumental application would be unfair in certain cases. They described the circumstances under which they adjust their assessment decisions, particularly when a student is close to the pass threshold, when working with students with special educational needs, or in response to specific classroom contexts (e.g., following teacher absence). Discussions about the use of criteria sometimes gave rise to intense debates, with some teachers - often the more experienced ones - adopting more rigid positions ($n = 2$).

Grades recorded in reports were also discussed collectively ($n = 1$). Teachers participating in this study reported that discussions with colleagues immediately before entering grades were particularly useful. They explained that the information exchanged enabled them to view students differently and to interact more closely with them. This also contributed to improving consistency in the qualitative sections of reports.

Effects of collaborative work on the micro, meso, and macro levels

This section addresses the following question: What consequences does this collaborative work have for students and teachers, schools, and the education system? To provide a clear overview, we present the results of the analysis of the 30 studies in two tables organized in three levels: micro with two sublevels for students and teachers; meso for schools; and macro for the education system. Following Vangrieken et al. (2015) in their systematic review, we distinguished between positive and negative effects, while acknowledging that this dichotomy is not always objective. Indeed, such a classification does not fully capture the complexity of the effects. Nevertheless, with this choice, we aimed to enhance clarity and readability by grouping the different effects in a clear and intuitive way for the reader.

Positive effects

As noted above, the literature highlights numerous positive effects of teacher collaboration overall. Drawing on the 30 studies, Table 3 presents these effects within the specific context of assessing student learning.

Table 3
Positive effects of collaborative assessment work for students-teachers, schools, and education systems

	Micro level		Meso level		Macro level	
	Students	Teachers	Schools	Education systems	Education systems	Education systems
– Student involvement in the assessment process (Buckley & Liscombe, 2022; Hargreaves, 2012; Hermansen, 2014; Mottier Lopez & Pasquini, 2017; Smail, 2020)	– Development of professional judgement (Adie et al., 2012; Allal & Mottier Lopez, 2014; Connolly et al., 2012; Furtak et al., 2016; Klenowski et al., 2007; Mottier Lopez et al., 2012; Nilsberth & Sandlund, 2020; Van der Schaaf et al., 2012; Wyatt-Smith et al., 2010)	– Mutual learning (Reid, 2007; Stewart & Houchens, 2014)	– Shared assessment culture (Hargreaves, 2013; Mottier Lopez & Pasquini, 2017; Smail, 2020)	– Improvement of assessment quality (validity, consistency, reliability) (Allal & Mottier Lopez, 2014; Black et al., 2011; Buckley & Liscombe, 2020; Connolly et al., 2012; Furtak et al., 2016; Hermansen, 2014; Hipkins & Robertson, 2012; Mottier Lopez & Pasquini, 2017; Smail, 2020; Van der Schaaf et al., 2012; Wyatt-Smith et al., 2010; Yerly, 2021)	– Improvement of assessment quality (validity, consistency, reliability) (Allal & Mottier Lopez, 2014; Black et al., 2011; Buckley & Liscombe, 2020; Connolly et al., 2012; Furtak et al., 2016; Hermansen, 2014; Hipkins & Robertson, 2012; Mottier Lopez & Pasquini, 2017; Smail, 2020; Van der Schaaf et al., 2012; Wyatt-Smith et al., 2010; Yerly, 2021)	– Improvement of assessment quality (validity, consistency, reliability) (Allal & Mottier Lopez, 2014; Black et al., 2011; Buckley & Liscombe, 2020; Connolly et al., 2012; Furtak et al., 2016; Hermansen, 2014; Hipkins & Robertson, 2012; Mottier Lopez & Pasquini, 2017; Smail, 2020; Van der Schaaf et al., 2012; Wyatt-Smith et al., 2010; Yerly, 2021)
– Support for students' progress (Hargreaves, 2012)	– Harmonization of practices (Allal & Mottier Lopez, 2014; Yerly, 2021)	– Improved pedagogical use of assessment (Buckley & Liscombe, 2022; Hermansen, 2014; Hermansen & Nerland, 2014; McFadden et al., 2022; Mottier Lopez & Pasquini, 2017)	– Improved interinstitutional relationships (Adie, 2014; Crisp, 2017; Hargreaves, 2012; Morales Villabona & Mottier Lopez, 2016; Reid, 2007)	– Improved interinstitutional relationships (Adie, 2014; Crisp, 2017; Hargreaves, 2012; Morales Villabona & Mottier Lopez, 2016; Reid, 2007)	– Improved interinstitutional relationships (Adie, 2014; Crisp, 2017; Hargreaves, 2012; Morales Villabona & Mottier Lopez, 2016; Reid, 2007)	– Improved interinstitutional relationships (Adie, 2014; Crisp, 2017; Hargreaves, 2012; Morales Villabona & Mottier Lopez, 2016; Reid, 2007)
– Communication of results (Mottier Lopez & Pasquini, 2017)	– Change/transformation of assessment practices (Adie, 2014; Allal & Mottier Lopez, 2014; Black et al., 2011; Buckley & Liscombe, 2022; Grant, 2012; Hermansen, 2014; McFadden et al., 2022; Morales Villabona & Mottier Lopez, 2016; Mottier Lopez & Pasquini, 2017; Smail, 2020; Stewart & Houchens, 2014; Yerly, 2021)	– Change/transformation of assessment practices (Adie, 2014; Allal & Mottier Lopez, 2014; Black et al., 2011; Buckley & Liscombe, 2022; Grant, 2012; Hermansen, 2014; McFadden et al., 2022; Morales Villabona & Mottier Lopez, 2016; Mottier Lopez & Pasquini, 2017; Smail, 2020; Stewart & Houchens, 2014; Yerly, 2021)	– Production of pedagogical resources (Black et al., 2011; Hipkins & Robertson, 2012; Matre & Solheim, 2016; Smail, 2020; Stewart & Houchens, 2014)	– Production of pedagogical resources (Black et al., 2011; Hipkins & Robertson, 2012; Matre & Solheim, 2016; Smail, 2020; Stewart & Houchens, 2014)	– Production of pedagogical resources (Black et al., 2011; Hipkins & Robertson, 2012; Matre & Solheim, 2016; Smail, 2020; Stewart & Houchens, 2014)	– Production of pedagogical resources (Black et al., 2011; Hipkins & Robertson, 2012; Matre & Solheim, 2016; Smail, 2020; Stewart & Houchens, 2014)
– Improved academic achievement (Furtak et al., 2016; Stewart & Houchens, 2014)	– Emergence of professional debate (Mottier Lopez et al., 2012; Mottier Lopez & Pasquini, 2017; Reid, 2007; Smail, 2020)	– Emergence of professional debate (Mottier Lopez et al., 2012; Mottier Lopez & Pasquini, 2017; Reid, 2007; Smail, 2020)	– Links with extracurricular providers (Allal & Mottier Lopez, 2014)	– Links with extracurricular providers (Allal & Mottier Lopez, 2014)	– Links with extracurricular providers (Allal & Mottier Lopez, 2014)	– Links with extracurricular providers (Allal & Mottier Lopez, 2014)

Table 3 (continued)
Positive effects of collaborative assessment work for students-teachers, schools, and education systems

	Micro level	Teachers	Meso level	Macro level
Students			Schools	Education systems
	<ul style="list-style-type: none"> - Development and support of new professional knowledge (Crisp, 2017; Hermansen & Nerland, 2014; Mottier Lopez et al., 2012; Mottier Lopez & Pasquini, 2017; Reid, 2007; Stewart & Houchens, 2014) - Reflection on assessment practices, metacognition (Hargreaves, 2013; Heredia et al., 2016; McFadden et al., 2022) - Stronger relationships (Adie, 2013; Hargreaves, 2013) - Development of shared understanding of standards (Adie, 2013, 2014; Buckley & Liscombe, 2022; Connolly et al., 2012; Crisp, 2017; Hermansen, 2014; Hipkins & Robertson, 2012; Klenowski et al., 2007; Matre & Solheim, 2016; Smail, 2020; Wyatt-Smith et al., 2010) - More flexible use of standards (Connolly et al., 2012; Klenowski et al., 2007; Matre & Solheim, 2016; Mottier Lopez et al., 2012; Wyatt-Smith et al., 2010) - Increased accountability (Adie, 2014; Grant, 2012; Smail, 2020) - Development of a shared repertoire (resources, strategies, knowledge) (Adie, 2014; Allal & Mottier Lopez, 2014; Hermansen & Nerland, 2014; Klenowski et al., 2007; Matre & Solheim, 2016; Mottier Lopez et al., 2012; Smail, 2020; Stewart & Houchens, 2014) 		<ul style="list-style-type: none"> - Shared leadership between teachers and school leadership (Hipkins & Robertson, 2012) - Process harmonization (Allal & Mottier Lopez, 2014; Buckley & Liscombe, 2022; Matre & Solheim, 2016; Stewart & Houchens, 2014; Yerly, 2021) - Facilitation of continued collaborative work (Grant, 2012) 	

Table 3 (continued)
Positive effects of collaborative assessment work for students-teachers, schools, and education systems

Students	Micro level	Teachers	Meso level	Macro level
			Schools	Education systems
		<ul style="list-style-type: none"> - Planning next steps in the teaching process (Smail, 2020) - Development of shared professional identity (Adie, 2013; 2014; Klenowski et al., 2007) - Less disparities (Allal & Mottier Lopez, 2014) - Assessment adjustments (Mottier Lopez et al., 2012; Nilsberth & Sandlund, 2020) - Improved assessment tasks (Black et al., 2011; Mottier Lopez & Pasquini, 2017) - Construction of consensus (Buckley & Liscompe, 2022; McFadden et al., 2022; Morales Villabona & Mottier Lopez, 2016; Mottier Lopez et al., 2012; Wyatt-Smith et al., 2010; Yerly, 2021) - Increased engagement (Connolly et al., 2012; Hargreaves et al., 2013) - Stronger leadership (Stewart & Houchens, 2014) - Sense-making and links between theory and practice (Hargreaves, 2013; Mc Fadden et al., 2022) - Easing tensions (Morales Villabona et Mottier Lopez, 2016) 		

Although collaboration appeared beneficial at all levels, most of its positive effects were observed at the micro level, particularly for teachers. The most frequently reported benefits were developments in assessment practices ($n = 12$); enhanced professional judgement ($n = 9$); new professional knowledge ($n = 6$); shared understanding of criteria and standards ($n = 11$); a common repertoire ($n = 8$); and reaching consensus ($n = 6$). Like the other effects reported in this section of the table, they all contributed to improved professional development.

At the school level, collaboration across educational stages (primary and secondary) enabled teachers to better understand their colleagues' working realities and the assessment policies implemented in their respective institutions ($n = 5$). Collaboration also contributed to the harmonizing processes, for assessment directly ($n = 2$) but also for related aspects such as organizing parent meetings or formulating direct feedback in assessments ($n = 2$) or in reports ($n = 1$). Several studies reported the development of shared assessment culture ($n = 3$). According to Mottier Lopez (2013), sharing assessment culture is important because it provides a foundation for more equitable assessing student learning in a school.

Another finding emerging from the results concerns the macro level, with 12 studies highlighting an accountability function. Across these studies, authors reported improvements in assessment relevance, validity, and/or reliability⁴. With the development of assessment practices ($n = 12$), this is the most frequently reported effect across all the articles.

Negative effects

As in the previous section, we present a table of the disadvantages of collaborative work. The findings in Table 4 showed there is no negative effect on the macro level at all. Once again, most effects were at the micro level, particularly for teachers. Unlike the previous table, we observed a greater prevalence of socio-emotional effects. Collaborative work focused on assessing student learning triggers interpersonal tensions ($n = 2$), shyness ($n = 3$), strong emotional reactions (unspecified) ($n = 1$), and fear

4. According to De Ketele and Gérard (2005), "relevance refers to the extent to which an assessment is appropriate, insofar as it aligns with the intended objectives [...]; validity refers to the degree of correspondence between what is claimed to be assessed (such as a given dimension) and what is actually assessed, that is, between what the instrument measures and what it is intended to measure [...]; and reliability refers to the degree of confidence that can be placed in the observed results [...]" (p. 3).

of exposing personal assessment practices ($n = 2$). Discussions around assessment also provoked varying degrees of reluctance. For example, some teachers refused to adopt new assessment practices ($n = 1$) or avoided actively engaging in exchanges ($n = 3$). Several studies also indicated that collaboration reinforces some teachers' tendency to adopt an instrumental approach to criteria ($n = 2$), referring to them strictly and justifying their choices by emphasizing equality of treatment.

At the school level, few negative consequences were reported across the 30 articles. In their study, Black et al. (2011) argued that collective assessment practices were overly influenced by external assessments during the joint design of tests. They observed that teachers tended to reproduce their format without fully understanding the underlying rationale, which did not meaningfully support their professional development. They also found that disagreements and controversy led to misalignment between assessment and instruction, resulting in some assessments suffering from significant bias. In another study, where collaboration was initiated between primary and secondary teachers, exchanges contributed to resurgence of pre-existing tensions between schools.

Discussion

The issues raised throughout this review reflect contemporary challenges intrinsically linked to an emerging trend in education: teachers working collaboratively for assessing student learning. We discuss the findings by examining several tensions that emerged from our documentary analysis.

First, we present how the distinction between control and evaluation helps clarify the objectives and implications of social moderation. This distinction highlights the different dimensions of the process and encourages reflection on practices and interpretations. We then discuss the challenges associated with formal and informal work in assessment, observing how the transition from informal to formal practices can promote situated professional judgement. The discussion also addresses the challenge of balancing assessment training needs for teachers with professional development benefits generated by collaborative work. Finally, the concept of situated professional judgement is examined in light of the results of the documentary analysis.

Table 4
Negative effects of collaborative assessment work for students-teachers, schools, and education systems

	Micro level	Teachers	Meso level	Macro level
Students			Schools	Education systems
– Unfavorable criteria (Mc Fadden et al., 2020)	– Professional judgement more strongly influenced by personal opinions (Van Der Schaaf et al., 2012)	– Professional judgement more strongly influenced by personal opinions (Van Der Schaaf et al., 2012)	– Misuse of external tests (Black et al., 2011)	
	– Mechanical threshold setting (Mottier Lopez & Pasquini, 2017)	– Mechanical threshold setting (Mottier Lopez & Pasquini, 2017)	– Misalignment between assessment and instruction (Black et al., 2011)	
	– Emotional reactions (Crisp, 2017)	– Emotional reactions (Crisp, 2017)	– Strained interinstitutional relationships (Morales Villabona & Mottier Lopez, 2016)	
	– Misunderstood criteria (Reid, 2007)	– Misunderstood criteria (Reid, 2007)		
	– Fear of exposing personal practices (Adie, 2013; Nilsberth & Sandlund, 2020)	– Fear of exposing personal practices (Adie, 2013; Nilsberth & Sandlund, 2020)		
	– Less involvement, reluctance (Adie, 2013; Connolly et al., 2012; Wyatt Smith et al., 2010)	– Less involvement, reluctance (Adie, 2013; Connolly et al., 2012; Wyatt Smith et al., 2010)		
	– Shyness and discomfort (Adie, 2013; Nilsberth & Sandlund, 2020; Wyatt Smith et al., 2010)	– Shyness and discomfort (Adie, 2013; Nilsberth & Sandlund, 2020; Wyatt Smith et al., 2010)		
	– Wasted time and increased workload (Klenowski et al., 2007)	– Wasted time and increased workload (Klenowski et al., 2007)		
	– Power struggles (Morales Villabona & Mottier Lopez, 2016; Nilsberth & Sandlund, 2020)	– Power struggles (Morales Villabona & Mottier Lopez, 2016; Nilsberth & Sandlund, 2020)		
	– Instrumental (inflexible) approach to standards (Hipkins & Robertson, 2012; Nilsberth & Sandlund, 2020)	– Instrumental (inflexible) approach to standards (Hipkins & Robertson, 2012; Nilsberth & Sandlund, 2020)		
	– Interpersonal tensions (Black et al., 2011; Yerly, 2021)	– Interpersonal tensions (Black et al., 2011; Yerly, 2021)		
	– Assessment arrangements (Allal & Mottier Lopez, 2014; Colbert et al., 2012; Hipkins & Robertson, 2012; Wyatt-Smith et al., 2010)	– Assessment arrangements (Allal & Mottier Lopez, 2014; Colbert et al., 2012; Hipkins & Robertson, 2012; Wyatt-Smith et al., 2010)		
	– Analytic grading (Klenowski et al., 2007)	– Analytic grading (Klenowski et al., 2007)		
	– Difficulty reaching consensus (Mottier Lopez et al., 2012)	– Difficulty reaching consensus (Mottier Lopez et al., 2012)		
	– Persisting traditional practices (Heredia et al., 2016)	– Persisting traditional practices (Heredia et al., 2016)		

Our findings show that social moderation is the most widespread aim of collaborative work for assessment. In the early 2000s, Maxwell (2001) considered the professional development function of social moderation to be secondary to its accountability function. This study, 20 years later, suggests the opposite trend: the professional development purpose ($n = 13$) is more frequent than accountability ($n = 3$). In this context, following Ardoino and Berger (1990), who approached educational evaluation through foundational perspectives, for the purposes of this study, we distinguish between control and evaluation. Through an etymological approach based on a French wordplay between *plus-valeur* and *plus-value* - referring respectively to value-based judgment and added value - Ardoino and Berger define control as a set of procedures aimed at verifying the implementation of plans or programs, closely linked to accountability and coordination. By contrast, they conceptualize evaluation as a process aimed at extracting the value of a task, with a focus on improvement and development. This distinction sheds light on the different purposes of collaborative work in assessment, showing that the emphasis is increasingly placed on improving practices rather than merely controlling outcomes.

Our findings also show that when teachers use common criteria, informal (individual) practices cross more formal ones. This raises a key question: How do these two dimensions interact to support situated professional judgement? The findings indicate that formal work can reveal sometimes hidden informal practices such as adjusting assessments. Such practices may become a source of tension regarding the application of procedures, recommendations, or standards in assessment. Informal assessment, such as classroom observation, are a source of information for teachers, but are not necessarily planned, and such informal practices are influenced shaped by contextual considerations. In this sense, informal work around assessment is inherently situated. Informal practices can therefore serve as a starting point for formal discussions, fostering professional dialogue - for example, about assessment approaches and professional judgement. Moving from the informal to the formal may encourage teachers to reflect on their assessment practices. Formal work can thus provide opportunities to clarify and formalize professional norms in assessment, as well as to build a shared repertoire of tools, practices, and methods in line with institutional expectations. Formal discussions also deepen understanding of contextual considerations and create opportunities to integrate them into the assessment process, promoting situated professional judgement.

Regarding the consequences of collaborative work, our findings are consistent with those of Vangrieken et al. (2015). Although all levels are affected, the effects are concentrated on teachers, and more positively than negatively. Whereas positive effects mainly contribute to professional development, negative effects are largely socio-emotional in nature (e.g., shyness, discomfort). Teachers seem to find it difficult to publicly expose their assessment practices. According to Mottier Lopez and Allal (2014), such practices often remain a private matter. These findings resonate with research suggesting both that teachers feel insufficiently trained (Mertler, 2004) and that shortcomings persist in their assessment practices (DeLuca et al., 2016). A tension therefore emerges between the need to strengthen teachers' training in assessment and the professional development benefits associated with collaborative work. This tension highlights the challenge of balancing these two dimensions: Should greater emphasis be placed on teacher education (initial and continuing), or on more systematic forms of collective assessment?

One of the most frequently reported effects for teachers was the development of professional judgement ($n = 9$). In several studies ($n = 4$), participants adjust their professional judgement in response to the sociocultural context. From a situated perspective on assessment (Mottier Lopez, 2013; 2021), the classroom and school context, as well as the role of language, discourse, and cultural tools, are central. A closer analysis of the findings shows that, among the nine articles addressing the development of professional judgement, seven concern arrangements designed to moderate results. In such contexts, professional judgement tends to be conceptualized at the end of the process, mainly during grading. However, Allal and Mottier Lopez (2008) suggested that professional judgement operates throughout the entire assessment process. Their work highlighted that professional judgement is both a cognitive act and a situated social practice throughout all phases of assessment. This continuous conception of professional judgement contrasts with contexts where it is only institutionalized during marking and grading stages (e.g., moderation sessions in Australia) but aligns with studies conducted in natural settings by Yerly (2021) in a secondary school in French-speaking Switzerland and by Holmeier et al. (2017)⁵ in several secondary schools in German-speaking

5. This article was not included in our selection because it is written in German.

Switzerland. This raises a legitimate question: When should collaboration be introduced into the assessment process and are some stages more critical than others for achieving the intended objectives?

Conclusion

The main objective of this study is to provide an overview of the English- and French-language scientific literature on teachers' collaborative work in assessing student learning, focusing on both organization and the consequences. The review examined four main areas: (a) purposes, (b) tools and methods, (c) dimensions of assessment, and (d) effects for all participants involved in education.

The concept of collective assessment appears to be emerging and more prominent in the English-language literature than in the French-language literature. Education systems are less homogeneous in the French-speaking world. For example, the nationally centralized French system is very different to the more federal structures found in the Anglo-Saxon context. This leads to cultural differences, particularly regarding assessment policies. Collaborative assessment work serves several purposes: moderating results, designing tools, developing new practices, and implementing guidelines. When teachers work collaboratively, they use personal tools, collectively developed tools, or tools made available to them. When they come together, they work in teams or draw on communities. Grading is the dimension of assessment most affected by collaborative work. Assessment referents (objectives, tasks, criteria, thresholds, and weightings) are also strongly debated during collaborative assessment design. All levels benefit from collaborative work, with most positive effects on teachers (e.g., development of practices and knowledge, enhancement of professional judgement). Our overview also suggests that collaborative work enables education systems to improve the quality of their assessments. However, some negative effects are also observed, particularly for teachers, which manifest in shyness, concerns, and interpersonal tensions, for example.

Although conducted with the greatest possible rigor, this study has several limitations. First, we identify a methodological bias related to the research design. While a rapid review streamlines and accelerates the systematic review process, it also increases the risk of bias—for example, in the selection of studies or the inclusion of weaker studies. The use

of additional databases could have resulted in a richer sample, and it is therefore possible that relevant studies were overlooked. A strictly systematic approach would be likely to produce a more exhaustive and robust selection of sources with more structured decision-making. Our study also suffers from a clear linguistic bias. Our inclusion criteria limited the review to articles written in English and French, thereby excluding studies published in other languages. Finally, potential bias in the interpretation of the articles must also be acknowledged.

Our documentary analysis points to contemporary and emerging issues in assessing student learning. However, it also raises additional questions that we were unable to address. For instance, our results examine collaborative work in both low- and high-stake assessment without distinction. Future research exploring whether collective dynamics—such as attitudes, tools, methods, and effects—vary depending on context would be particularly valuable. The literature suggests that collaborative work is more constructive when teachers are not (overly) constrained (Hargreaves, 2012, 2019; Vangrieken et al., 2015). Examining this specific issue in the context of student assessment would therefore be another promising avenue for research.

Proofreading: Caroline Lefour

Formatting: Emmanuel Gagnon

Portugue abstract: Eusebio Andre Machado

LISTE OF REFERENCES

- Adie, L. (2013). The development of teacher assessment identity through participation in online moderation. In V. Klenowski (Ed.), *International Teacher Judgement Practices* (pp. 101-116). Routledge. <https://doi.org/10.1080/0969594X.2011.650150>
- Adie, L. (2014). The development of shared understandings of assessment policy: Travelling between global and local contexts. *Journal of Education Policy*, 24(4), 532-545. <https://doi.org/10.1080/02680939.2013.853101>
- Adie, L. E., Klenowski, V., & Wyatt-Smith, C. (2012). Towards an understanding of teacher judgement in the context of social moderation. *Educational Review*, 64(2), 223-240. <https://doi.org/10.1080/00131911.2011.598919>
- Allal, L. & Mottier Lopez, L. (2008). Mieux comprendre le jugement professionnel en évaluation: apports et implications de l'étude genevoise. In L. Lafortune & L. Allal (Eds.), *Jugement professionnel en évaluation: pratiques enseignantes au Québec et à Genève* (pp. 223- 239). Presses de l'Université du Québec.
- Allal, L., & Mottier Lopez, L. (2009). Au cœur du jugement professionnel en évaluation: des démarches de triangulation. *Les dossiers des sciences de l'éducation*, 22(1), 25-40. <https://doi.org/10.3406/dsedu.2009.1164>
- Allal, L., & Mottier Lopez, L. (2014). Teachers' professional judgment in the context of collaborative assessment practice. In C. Wyatt-Smith, V. Klenowski & P. Colbert (Eds.), *Designing assessment for quality learning* (pp. 151-165). https://doi.org/10.1007/978-94-007-5902-2_10
- Allenbach, M. (2016). *Négociation de pratiques collaboratives et construction du rôle des intervenant-es à l'école* [Thèse de doctorat, Université de Fribourg].
- Ardoino, J., & Berger, G. (1990). Etude critique – Vers une nouvelle conception de l'évaluation? *Revue française de pédagogie*, (90), 105-112. https://www.persee.fr/doc/rfp_0556-7807_1990_num_90_1_2466_t1_0105_0000_1
- Barrère, A., Saujat, F., & Lantheaume, F. (2008). Rendre visible le travail enseignant. Questions de méthodes: Entretien de Anne Barrère et Frédéric Saujat par Françoise Lantheaume. *Recherche et formation*, 57(1). <https://doi.org/10.4000/rechercheformation.861>
- Beuret, J. E. (2010). De la négociation conflictuelle à la négociation concertative: un «Point de Passage Transactionnel». *Négociations*, (1), 43-60. <https://www.cairn.info/revue-negociations-2010-1-page-43.htm>
- Black, P., Harrison, C., Hodgen, J., Marshall, B., & Serret, N. (2011). Can teachers' summative assessments produce dependable results and also enhance classroom learning? *Assessment in Education: Principles, Policy & Practice*, 18(4), 451-469. <https://doi.org/10.1080/0969594X.2011.557020>
- Borgès, C. & Lessard, C. (2007). Qu'arrive-t-il quand la collaboration enseignante devient une norme? In J-F. Marcel (Ed), *Coordonner, collaborer, coopérer: de nouvelles pratiques enseignantes* (pp. 61-74). De Boeck Supérieur. <https://doi.org/10.3917/dbu.marce.2007.01.0061>

- Buckley, K., & Lipscombe, K. (2022). Validity and the design of classroom assessment in teacher teams. *The Australian Educational Researcher*, 49(2), 425-444. <http://dx.doi.org/10.1007/s13384-021-00437-9>
- Cohen, S. G., & Bailey, D. E. (1997). What makes teams work: Group effectiveness research from the shop floor to the executive suite. *Journal of Management*, 23, 239–290. <https://doi.org/10.1177/01492063970230>
- Colbert, P., Wyatt-Smith, C., & Klenowski, V. (2012). A systems-level approach to building sustainable assessment cultures: Moderation, quality task design and dependability of judgement. *Policy Futures in Education*, 10(4), 386-401. <https://doi.org/10.2304/pfie.2012.10.4.386>
- Connolly, S., Klenowski, V., & Wyatt-Smith, C. M. (2012). Moderation and consistency of teacher judgement: teachers' views. *British Educational Research Journal*, 38(4), 593-614. <https://doi.org/10.1080/01411926.2011.569006>
- Crisp, V. (2017). The judgement processes involved in the moderation of teacher-assessed projects. *Oxford Review of Education*, 43(1), 19-37. <https://doi.org/10.1080/03054985.2016.1232245>
- Dachet, D. (2024). *Umbrella Review: de sa méthodologie aux considérations méthodologiques relatives à son application dans le champ des sciences de l'Éducation* [Thèse de doctorat, Université de Liège].
- Decuyper, S., Dochy, F., & Van den Bossche, P. (2010). Grasping the dynamic complexity of team learning: An integrative model for effective team learning in organizations. *Educational Research Review*, 5, 111–133. <https://doi.org/10.1016/j.edurev.2010.02.002>
- De Ketele, J.-M. (2006). La recherche en évaluation: propos synthétiques et prospectifs. *Mesure et évaluation en éducation*, 29(1), 99-118. <https://doi.org/10.7202/1086970ar>
- De Ketele, J.-M. & Gérard, F.-M. (2005). La validation des épreuves d'évaluation selon l'approche par les compétences. *Mesure et évaluation en éducation*, 28(3), 1-26. <https://doi.org/10.7202/1087028ar>
- DeLuca, C., Lapointe-Mcewan, D., & Luhanga, U. (2016). Teacher assessment literacy: A review of international standards and measures. *Educational Assessment, Evaluation and Accountability*, 28(3), 251-272. <https://doi.org/10.1007/s11092-015-9233-6>
- Doppenberg, J. J., Den Brok, P. J., & Bakx, A. W. E. A. (2012). Collaborative teacher learning across foci of collaboration: Perceived activities and outcomes. *Teachers and Teacher Education*, 28, 899–910. <https://doi.org/10.1016/j.tate.2012.04.007>
- Dumez, H. (2011). Faire une revue de littérature : pourquoi et comment ? *Le libellio d'aegis*, 7(2), 15-27. <https://hal.archives-ouvertes.fr/hal-00657381>
- Furtak, E. M., Kiemer, K., Circi, R. K., Swanson, R., de León, V., Morrison, D., & Heredia, S. C. (2016). Teachers' formative assessment abilities and their relationship to student learning: Findings from a four-year intervention study. *Instructional Science*, 44, 267-291. <https://doi.org/10.1007/s11251-016-9371-3>
- Germier, C., & Marcel, J. F. (2016). La négociation des accords. Approche ethnographique des pratiques de deux collectifs d'enseignants. *Les dossiers des sciences de l'éducation*, (35), 73-96. <https://doi.org/10.4000/dse.1271>
- Gilbert, A-F. (2018). *Le travail collectif enseignant, entre informel et institué*. Dossier de veille de l'IFÉ, n° 124.

- Gipps, C. (1994, April, 4-8). *Quality assurance in teachers' assessment* [Présentation]. Annual meeting of the American Educational Research Association, New-Orleans, USA.
- Grangeat, M. (2011). Le travail collectif enseignant : éléments de modélisation du développement professionnel. In M. Grangeat (Éd.), *Les démarches d'investigation dans l'enseignement scientifique Pratiques de classe, travail collectif enseignant, acquisitions des élèves* (pp. 79-106). École Normale Supérieure. <https://shs.hal.science/halshs-01205850>
- Grant, L. (2012). Cross-sector moderation as a means of engaging staff with assessment and teaching pedagogy. *Policy Futures in Education*, 10(4), 434-446. <https://doi.org/10.2304/pfie.2012.10.4.434>
- Hargreaves, A. (1994). *Changing teachers, changing times: Teachers' work and culture in the postmodern age*. OISE Press.
- Hargreaves, E. (2007). The validity of collaborative assessment for learning. *Assessment in Education*, 14(2), 185-199. <https://doi.org/10.1080/09695940701478594>
- Hargreaves, E. (2012). Assessment for learning and teacher learning communities: UK teachers' experiences. *Teaching Education*, 24(3), 327-344. <https://doi.org/10.1080/10476210.2012.713931>
- Hargreaves, E., Berry, R., Lai, Y. C., Leung, P., Scott, D., & Stobart, G. (2013). Teachers' experiences of autonomy in continuing professional development: Teacher learning communities in London and Hong Kong. *Teacher development*, 17(1), 19-34. <https://doi.org/10.1080/13664530.2012.748686>
- Hargreaves, A. (2019). Teacher collaboration: 30 years of research on its nature, forms, limitations and effects. *Teachers and Teaching*, 25(5), 603-621. <https://doi.org/10.1080/13540602.2019.1639499>
- Hattie, J. (2017). *L'apprentissage visible pour les enseignants. Connaître son impact pour maximiser le rendement des élèves*. PUQ.
- Heredia, S. C., Furtak, E. M., Morrison, D., & Renga, I. P. (2016). Science teachers' representations of classroom practice in the process of formative assessment design. *Journal of Science Teacher Education*, 27(7), 697-716. <https://doi.org/10.1007/s10972-016-9482-3>
- Hermansen, H. (2014). Recontextualising assessment resources for use in local settings: Opening up the black box of teachers' knowledge work. *Curriculum Journal*, 25(4), 470-494. <https://doi.org/10.1080/09585176.2014.956771>
- Hermansen, H., & Nerland, M. (2014). Reworking practice through an AfL project: An analysis of teachers' collaborative engagement with new assessment guidelines. *British educational research journal*, 40(1), 187-206. <https://doi.org/10.1002/berj.3037>
- Hipkins, R., & Robertson, S. (2012). The complexities of moderating student writing in a community of practice. *Assessment Matters*, 4, 30-52. <https://doi.org/10.18296/am.0099>
- Holmeier, M., Maag Merki, K., & Hirt, C. (2017). Holmeier, M., Maag Merki, K., & Hirt, C. (2017). *Gemeinsames Prüfen. Eine Fallanalyse in Gymnasien in der Schweiz*. Wiesbaden: Springer. <https://doi.org/10.1007/978-3-658-17060-8>

- Klenowski, V., Adie, L., Gunn, S., Looney, A., Elwood, J., Wyatt-Smith, C., & Colbert, P. (2007, November 25-29). *Moderation as judgement practice: Reconciling system level accountability and local level practice*. In P.-L. Jeffery (chair), *Conference: Research impacts: Proving or improving?* [Conference]. Australian Association for Research in Education, Fremantle, Australia.
- Lave, J. (1988). *Cognition in practice: Mind, mathematics and culture in everyday life*. Cambridge University Press.
- Laveault, D. & Yerly, G. (2017). Modération statistique et modération sociale des résultats scolaires : approches opposées ou complémentaires? *Mesure et évaluation en éducation*, 40(2), 91–123. <https://doi.org/10.7202/1043569ar>
- Letor, C. (2010). Moments de coopération entre enseignants: entre logiques institutionnelle, organisationnelles et professionnelles. *Travail et formation en éducation*, 7, 107-120. <http://journals.openedition.org/tfe/1458>
- Linn, R. L. (1993). Linking results of distinct assessments. *Applied Measurement in Education*, 6(1), 83–102. https://doi.org/10.1207/s15324818ame0601_5
- Marcel, J-F, Dupriez, V., Perisset Bagnoud, D., & Tardif, M. (2007). *Coordonner, collaborer, coopérer : de nouvelles pratiques enseignantes*. De Boeck.
- Maroy, C. (2005). Vers une régulation post-bureaucratique des systèmes d'enseignement en Europe? *Cahiers de la recherche en Education et Formation*, (49), 30p. <http://hdl.handle.net/2078.1/81045>
- Maroy, C., & Voisin, A. (2014). Une typologie des politiques d'accountability en éducation: l'incidence de l'instrumentation et des théories de la régulation. *Éducation comparée. Revue de recherche internationale et comparative en éducation*, 11, 31-58. <https://doi.org/1866/14074>
- Matre, S., & Solheim, R. (2016). Opening dialogic spaces: Teachers' metatalk on writing assessment. *International Journal of Educational Research*, 80, 188-203. <https://doi.org/10.1016/j.ijer.2016.07.001>
- Maxwell, G. S. (2001). *Moderation of assessments in vocational education and training*. Queensland Government Departement of Employment and Training.
- McFadden, J., Jung, K., Robinson, B., & Tretter, T. R. (2022). Teacher-developed Multi-Dimensional Science Assessments Supporting Elementary Teacher Learning about the Next Generation Science Standards. *Journal of Science Teacher Education*, 33(1), 55-82. <https://doi.org/10.1080/1046560X.2021.1905331>
- McLaughlin, M. W., & Talbert, J. E. (2006). *Professional communities and the work of high school teaching*. University of Chicago Press.
- Meirieu, P. (2023). Du bon usage de l'autonomie des établissements scolaires. *Administration et Éducation*, 1, 127-141. <https://doi.org/10.3917/admed.177.0127>
- Mertler, C. A. (2004). Secondary teachers' assessment literacy: Does classroom experience make a difference? *American secondary education*, 33(1) 49-64. <https://www.jstor.org/stable/41064623>
- Morales Villabona, F. & Mottier Lopez, L. (2016). Quelle évaluation collaborative dans la modération sociale entre enseignants? In L. Mottier Lopez & W. Tessaro (Eds.), *Le jugement professionnel, au cœur de l'évaluation et de la régulation des apprentissages* (pp. 289-312). Peter Lang.
- Mottier Lopez, L. (2008). *Apprentissage situé: la microculture de classe en mathématiques*. Peter Lang.

- Mottier Lopez, L. (2013). De la mesure à l'évaluation collaborative en éducation. *Revue française d'administration publique*, 148(4), 939-952. <https://doi.org/10.3917/rfap.148.0939>
- Mottier Lopez, L. (2021). Une évaluation continue pour apprendre durablement, une évaluation à visée inclusive. *Revue suisse de pédagogie spécialisée*, 11(4), 9-16. <https://ojs.szh.ch/revue/article/view/168>
- Mottier Lopez, L., & Allal, L. (2008). Le jugement professionnel en évaluation : un acte cognitif et une pratique sociale située. *Schweizerische Zeitschrift für Bildungswissenschaften*, 30(3), 465-482. <https://doi.org/10.25656/01.4229>
- Mottier Lopez, L., & Pasquini, R. (2017). Professional controversies between teachers about their summative assessment practices: A tool for building assessment capacity. *Assessment in Education: Principles, Policy & Practice*, 24(2), 228-249. <https://doi.org/10.1080/0969594X.2017.1293001>
- Mottier Lopez, L., Tessaro, W., Dechamboux, L. & Villabona, F.- M. (2012). La modération sociale : un dispositif soutenant l'émergence de savoirs négociés sur l'évaluation certificative des apprentissages des élèves. *Questions vives: Recherches en éducation*, 6(18), 159-175. <https://doi.org/10.4000/questionsvives.1235>
- Nilsberth, M., & Sandlund, E. (2021). On the interactional challenges of revealing summative assessments: Collaborative scoring talk among teachers and students in Swedish national tests. *Linguistics and education*, 61(2), 1-16. <https://doi.org/10.1016/j.linged.2020.100899>
- Page M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffman, T. C., Mulrow, C. D. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic review. *British Medical Journal*, 372(1). <https://doi.org/10.1136/bmj.n71>
- Reid, L. (2007). Teachers talking about writing assessment: valuable professional learning? *Improving Schools*, 10(2), 132-149. <https://doi.org/10.1177/1365480207077843>
- Salas, E., Burke, S. C., & Cannon-Bowers, J. A. (2000). Teamwork: Emerging principles. *International Journal of Management Reviews*, 2, 339-356. <https://doi.org/10.1111/1468-2370.00046>
- Smail, E. (2020). Using involvement in moderation to strengthen teachers' assessment for learning capability. *Assessment in Education: Principles, Policy & Practice*, 27(5), 522-543. <https://doi.org/10.1080/0969594X.2020.1777087>
- Spillane, J. P., Shirrell, M., & Hopkins, M. (2016). Designing and deploying a professional learning community (PLC) organizational routine: Bureaucratic and collegial arrangements in tandem. *Les dossiers des sciences de l'éducation*, 35, 97-122. <https://doi.org/10.4000/dse.1283>
- Stewart, T. A., & Houchens, G. W. (2014). Deep impact: How a job-embedded formative assessment professional development model affected teacher practice. *Qualitative Research in Education*, 3(1), 51-82. <https://doi.org/10.4471/qre.2014.36>
- Strauss, A. (1992). *La trame de la négociation*. L'Harmattan.
- Van der Schaaf, M., Baartman, L., & Prins, F. (2012). Exploring the role of assessment criteria during teachers' collaborative judgement processes of students' portfolios. *Assessment & Evaluation in Higher Education*, 37(7), 847-860. <https://doi.org/10.1080/02602938.2011.576312>
- Vangrieken, K., Dochy, F., Raes, E., & Kyndt, E. (2015). Teacher collaboration: A systematic review. *Educational research review*, 15, 17-40. <http://dx.doi.org/10.1016/j.edurev.2015.04.002>

- Wenger, E. (1998). *Communities of practice: Learning, meaning, and identity*. Cambridge University Press.
- Wenger, E. (2005). *La théorie des communautés de pratique*. Presses Université Laval.
- Wyatt-Smith, C., Klenowski, V., & Gunn, S. (2010). The centrality of teachers' judgement practice in assessment: A study of standards in moderation. *Assessment in Education: Principles, policy & practice*, 17(1), 59-75. <https://doi.org/10.1080/09695940903565610>
- Yerly, G. (2021). La collaboration des enseignants autour de l'évaluation des apprentissages comme moyen de piloter le système? Le cas de la politique «évaluation commune» dans le secondaire supérieur en Suisse. *Les dossiers des sciences de l'éducation*, 44, 77-97.

Appendix. Full corpus of articles: authors, titles, publication years, countries of origin, educational levels, and research methods

Authors	Titles	Years	Countries	Educational level	Research designs
Adie	<i>The development of teacher assessment identity through participation in online moderation</i>	2013	AUS (Queensland)	Primary and secondary	Qualitative
Adie	<i>The development of shared understandings of assessment policy: Travelling between global and local contexts</i>	2014	AUS (Queensland)	Secondary	Qualitative
Adie et al.	<i>Towards an understanding of teacher judgement in the context of social moderation</i>	2012	AUS (Queensland)	Secondary	Qualitative
Allal and Mottier Lopez	<i>Mieux comprendre le jugement professionnel en évaluation : apports et implications de l'étude genevoise</i>	2014	CH (Geneva)	Secondary	Qualitative
Black et al.	<i>Can teachers' summative assessments produce dependable results and also enhance classroom learning?</i>	2011	GB	Secondary	Qualitative
Buckley and Liscombe	<i>Validity and the design of classroom assessment in teacher teams</i>	2022	AUS (NSW)	Primary	Qualitative
Colbert et al.	<i>A systems-level approach to building sustainable assessment cultures: Moderation, quality task design and dependability of judgement</i>	2012	AUS (Queensland)	Secondary	Qualitative
Connolly et al.	<i>Moderation and consistency of teacher judgement: teachers' views</i>	2012	AUS (Queensland)	Secondary	Qualitative
Crisp	<i>The judgement processes involved in the moderation of teacher-assessed projects</i>	2017	GB	Secondary	Qualitative
Furtak et al.	<i>Teachers' formative assessment abilities and their relationship to student learning: Findings from a four-year intervention study</i>	2016	USA	Secondary	Quantitative

Appendix. Full corpus of articles: authors, titles, publication years, countries of origin, educational levels, and research methods

Authors	Titles	Years	Countries	Educational level	Research designs
Grant	<i>Cross-sector moderation as a means of engaging staff with assessment and teaching pedagogy</i>	2012	GB (SCT)	Primary and secondary	Qualitative
Hargreaves	<i>Assessment for learning and teacher learning communities: UK teachers' experiences</i>	2012	GB	Secondary	Qualitative
Hargreaves et al.	<i>Teachers' experiences of autonomy in continuing professional development</i>	2013	GB & HKG	Secondary	Qualitative
Heredia et al.	<i>Science teachers' representations of classroom practice in the process of formative assessment design</i>	2016	USA	Secondary	Qualitative
Hermansen	<i>Recontextualising assessment resources for use in local settings: Opening up the black box of teachers' knowledge work</i>	2014	NO	Secondary	Qualitative
Hermansen & Nerland	<i>Reworking practice through an AfL project: An analysis of teachers' collaborative engagement with new assessment guidelines</i>	2014	NO	Secondary	Qualitative
Hipkins & Robertson	<i>The complexities of moderating student writing in a community of practice</i>	2012	NZ	Secondary	Qualitative
Klenowski et al.	<i>Moderation as judgement practice: Reconciling system level accountability and local level practice</i>	2007	AUS (Queensland)	Secondary	Qualitative
Matre & Solheim	<i>Opening dialogic spaces: Teachers' metatalk on writing assessment</i>	2016	NO	Secondary	Qualitative

Appendix. Full corpus of articles: authors, titles, publication years, countries of origin, educational levels, and research methods

Authors	Titles	Years	Countries	Educational level	Research designs
Mc Fadden et al.	<i>Teacher-developed Multi-Dimensional Science Assessments Supporting Elementary Teacher Learning about the Next Generation Science Standards</i>	2022	USA (Kentucky)	Primary	Qualitative
Morales Villabona & Mottier Lopez	<i>Quelle évaluation collaborative dans la modération sociale entre enseignants?</i>	2016	CH (Geneva)	Primary and secondary	Qualitative
Mottier Lopez & Pasquini	<i>Professional controversies between teachers about their summative assessment practices: A tool for building assessment capacity</i>	2017	CH (Geneva and Vaud)	Primary and secondary	Qualitative
Mottier Lopez et al.	<i>La modération sociale : un dispositif soutenant l'émergence de savoirs négociés sur l'évaluation certificative des apprentissages des élèves</i>	2012	CH (Geneva)	Primary	Qualitative
Nilberth & Sandlund	<i>On the interactional challenges of revealing summative assessments: Collaborative scoring talk among teachers and students in Swedish national tests</i>	2021	SE	Secondary	Qualitative
Reid	<i>Teachers talking about writing assessment: valuable professional learning?</i>	2007	GB (SCT)	Primary and secondary	Qualitative
Smail	<i>Using involvement in moderation to strengthen teachers' assessment for learning capability</i>	2020	NZ	Primary	Qualitative
Stewart & Houchens	<i>Deep impact: How a job-embedded formative assessment professional development model affected teacher practice</i>	2014	USA (Kentucky)	Primary	Qualitative
Van Der Schaaf et al.	<i>Exploring the role of assessment criteria during teachers' collaborative judgement processes of students' portfolios</i>	2012	NL	Secondary	Quantitative

Appendix. Full corpus of articles: authors, titles, publication years, countries of origin, educational levels, and research methods

Authors	Titles	Years	Countries	Educational level	Research designs
Wyatt Smith et al.	<i>The centrality of teachers' judgement practice in assessment: A study of standards in moderation</i>	2010	AUS (Queensland)	Primary and secondary	Mixte
Yerly	<i>La collaboration des enseignants autour de l'évaluation des apprentissages comme moyen de piloter le système? Le cas de la politique « évaluation commune » dans le secondaire supérieur en Suisse.</i>	2021	CH (Fribourg)	Secondary	Qualitative