

What to assess when assessing collaborative assignments: the practitioner perspective

Peter Levrai

To cite this article: Peter Levrai (13 Oct 2025): What to assess when assessing collaborative assignments: the practitioner perspective, Assessment & Evaluation in Higher Education, DOI: [10.1080/02602938.2025.2571219](https://doi.org/10.1080/02602938.2025.2571219)

To link to this article: <https://doi.org/10.1080/02602938.2025.2571219>



© 2025 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group



Published online: 13 Oct 2025.



Submit your article to this journal [↗](#)



Article views: 126



View related articles [↗](#)



View Crossmark data [↗](#)

What to assess when assessing collaborative assignments: the practitioner perspective

Peter Levrai 

Centre for Language and Communication Studies, University of Turku, Turku, Finland

ABSTRACT



Collaboration is a complex construct and collaborative assignments at university are no easy thing to assess. Various approaches have been proposed, mainly addressing formative assessment of collaborative assignments. When it comes to summative assessment, less has been said. This paper presents the findings of a constructivist Grounded Theory (GT) study into English for Academic Purposes (EAP) practitioners' evaluations of four different models of assessment for a group essay assignment. These evaluations consider the use of various assessment lenses to assess student collaborative work: assessing the product, assessing individual complementary assignments, and assessing the collaboration itself through measures like observation, teacher/group meetings, self- and peer-assessment, and learner reflection. Participants provided sound rationales for and against each assessment approach, further illustrating the complexity of assessing student collaboration. While assessing the product and individual complementary assignments found most favour, there was a clear appetite for assessing collaboration itself, even though there was no clear means of doing so. Although there is not a single assessment approach to collaboration which will fit all circumstances, a multi-lens approach, where assessment lenses are selected on the basis of course learning outcomes and potential learning gains, would allow for fully rationalised assessment schemes.

KEYWORDS

Collaborative assessment;
grounded theory;
assessment schemes;
English for Academic
Purposes

Introduction

Teamwork. Group work. Collaboration. We ask university students in higher education to work together on projects and assignments to develop vital skills valued by the world inside and outside of the academy (World Economic Forum 2016). Students engage in collaborative assignments (e.g. group essays, group presentations, or group projects) in faculties across an institution, sometimes working in a team within their discipline and sometimes operating in multidisciplinary teams. Group assignments have many learning affordances, but assessment remains a major concern

CONTACT Peter Levrai  peter.levrai@utu.fi  Centre of Language and Communication Studies, University of Turku, Turku, Finland

© 2025 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

(Strijbos 2016). Collaborative assessment is not as widely discussed in the literature as other forms of assessment and, when it is discussed, it is generally considered from a formative rather than summative perspective. Nonetheless, collaborative assignments are used for summative purposes in higher education and, consequently, how to best summatively assess them becomes an important issue.

The practitioner greatly impacts the experience of students during a collaboration (Ferreira, Zabolotna, and Lee 2024) and, consequently, we need a better understanding of the attitudes and beliefs of practitioners when it comes to student collaboration (Bolster 2025). This paper discusses a qualitative constructivist Grounded Theory (GT) study into English for Academic Purposes (EAP) practitioner approaches and attitudes towards assessing a collaborative assignment (Levrai 2025). This involved EAP practitioners working in different contexts evaluating and discussing four different assessment schemes for a group essay assignment, bringing to light prevailing concerns regarding the assessment of collaborative assignments.

Choosing what to assess is not the only concern when designing a collaborative assignment assessment scheme, but it is vital (Levrai 2025). The general finding of the research is that there is no standard collaborative assessment scheme which will be applicable in different contexts, meaning the selection of assessment lenses must be rationalised *in situ*. However, EAP practitioners favour a multi-lens approach to assessing collaborative assignments, which brings the collaboration into sharper focus and allows for the assessment of various learning outcomes.

Literature review

Assessment and learning outcomes

Assessment does not exist in a vacuum. We assess for a purpose and that purpose is typically to gather evidence to demonstrate students' level of knowledge, if learning has taken place, and/or if learning outcomes have been met (Boud and Bearman 2024). The assessment of learning outcomes is important to a range of stakeholders with an outcomes-based approach to learning common in most higher education institutions (Queen Mary Academy, n.d.). Graves (2016) highlights the centrality of learning goals and objectives in the planning of instruction and assessment, which is in line with Richards (2013) conception of Backward Design, with the means of assessment designed to demonstrate that learning outcomes have been met.

The importance of assessment to students and how it focuses their attention is well recognised since it 'shapes the experience of students and influences their behaviour more than the teaching they receive' (Bloxham and Boyd 2007, p. 3) and influences where they direct their effort (Gibbs 1999, p. 42). What is assessed gains most student attention and provides 'important motivating pressure' (Brown 2011, p. 46), which is relevant when considering whether to assess collaboration itself.

Assessing collaboration

The focus of this paper is the assessment of collaborative writing assignments in EAP. The role of EAP is to prepare students for their studies and future work,

expanding the scope of EAP beyond language to academic literacies (Bell 2024). Considerable work has been done to define ‘collaborative writing’. A collaborative writing task means all the students in the group are engaged in each aspect of the development of the written product, resulting in an indivisible final text for which they are all responsible and all have ownership of (Bolster and Levrai 2022; Storch and Knoch 2023). This is a closer means of working than cooperation, in which students can take a ‘divide and conquer approach’

One of the tensions around collaborative assignments is the question of assessment, which is recognised as a complex issue (Strijbos 2016). It is an area where we are still lacking solutions and which remains under-researched (Meijer et al. 2020), with assessment identified as one of the key research agenda areas identified by Li and Zhang (2023) in their exploration of current and future second language (L2) collaborative writing research.

Where assessing collaboration is discussed in the literature, it is commonly located as formative rather than summative assessment (Meijer et al. 2020; Storch and Knoch 2023), not least because all the challenges of assessing collaboration are compounded in the case of summative assessment. Nonetheless, the assessment of collaborative assignments matters because the assessment system impacts how fully university students engage with, participate in and behave in group assignments (Beard, Rymer, and Williams 1989).

Approaches to assessing collaborative assignments

There are multiple approaches to grading collaborative assignments, as outlined in Race (2001), including the use of self- and peer-assessment, the application of an individual weighting factor based on individual contribution (see Nepal 2012) and the use of additional individual tasks or assessments. As yet there is no agreed principle in the literature as to who should be grading collaboration (e.g. teacher led-, self- or peer-assessment), how it should be done, and little discussion of how collaborative assignments are assessed in EAP specifically. The perspective and experience of practitioners regarding collaborative assessment is also missing. A recent proposal from Storch and Knoch (2023) identifies three elements that could be assessed in an L2 collaborative writing activity: the written product, contributions to the written product, and wider collaboration, e.g. joint responsibility for and engagement with the work.

Assessing the product is potentially the least problematic of these as it is a single piece of written work and there are similarities to assessing a piece of writing by an individual, given collaborative writing should be indivisible. Assessing the process of writing poses more challenges, not least in terms of the visibility of student contribution to the teacher, particularly in an assignment which may be written over an extended period of time, both synchronously and asynchronously. While technology gives us some insight into the process of online writing, attempting to review version history of online documents to see who did what would be time-consuming (Zhang and Chen 2022). While a computer-based assessment tool could potentially process a text to help determine who typed what, it would not be able to capture the quality of each students’ contributions,

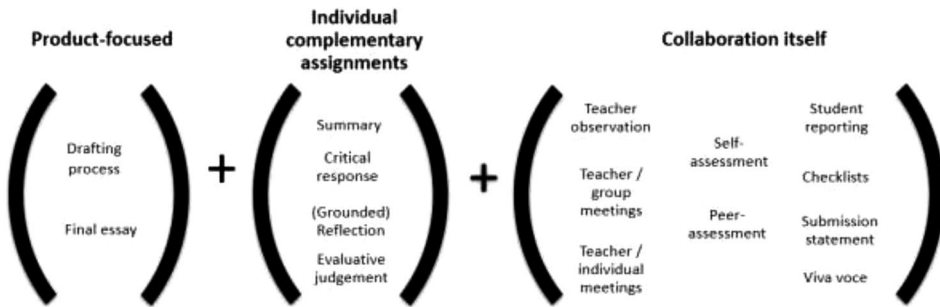


Figure 1. Potential assessment lenses for a collaborative assignment adapted from Levrai (2025).

particularly those not captured in keystroke (Van Aalst 2013). A challenge for the teacher when assessing collaboration is that collaboration is an opaque process, as recognised by Gammie and Matson (2007) and Johnston and Miles (2004).

Collaborative assignments, however, provides flexibility in ‘assessment lenses’, meaning assessment tools that provide opportunities to assess students. In the context of a written collaborative assignment, these assessment lenses could include ‘peer and individual reflective diaries and teacher observations’ (Storch 2017, p.139) and, in conjunction, they bring the collaboration into sharper focus. Figure 1 shows a range of potential assessment lenses that could be applied to evaluate a collaborative assignment: product-focussed, individual complementary assignments, and assessing collaboration itself.

Meijer et al. (2020) called for research into ‘teachers’ actual rationale for and practices of the assessment of collaborative learning in higher education’ (p. 1235), which is an area this paper hopes to address. Given the range of assessment lenses and approaches available, there needs to be better understanding of the approach(es) EAP tutors take when assessing collaborative assignments, which assessment lenses are employed, and why. Since collaborative assignments could address varied learning outcomes, the selection of assessment lenses becomes a key concern in shaping a course to meet those outcomes. The over-arching research question for the study is ‘How can we, if we can, best assess collaborative assignments in EAP?’ with this paper focused on practitioner evaluation of different assessment schemes. This allows for the development of an approach to developing an assessment scheme for collaborative assignments informed by both theory and practice.

Methodology

Employing grounded theory (GT)

To explore EAP practitioner approaches and attitudes to collaborative assessment, a qualitative constructivist GT approach (Charmaz 2014) was followed. GT was selected due to its appropriacy for examining under-researched phenomena and because it can also lead to a practical outcome, in this case a rationalised approach for assessing collaborative assignments. GT is an iterative research method, where research tool development leads to data collection, which in turn leads to data analysis and

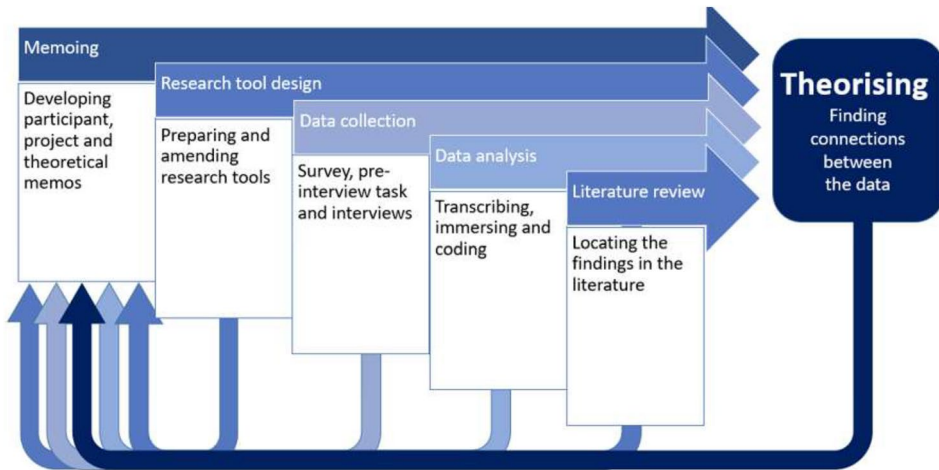


Figure 2. Concurrent processes in GT (Levrai 2025, 90).

Table 1. Participant HE institution locations.

Institution location	Participant pseudonym
Great Britain	Andrea, Barry, Bertha, Kyle, Van, Wendy
Asia	Ava, Dolly, Fernand
Middle East	Cristobal, Fay, Jerry
Europe	Dorian, Humberto, Karen, Laura, Lorenzo, Olga, Tanya
USA	Gonzalo

theorising, which inform further research tool development and the next cycle of data collection and analysis (see Figure 2) and there were four rounds of data collection and analysis in this study.

The participants

The 20 participants in the study, who will be referred to by pseudonym throughout, were selected through purposeful then theoretical sampling (Birks and Mills 2023) as active practitioners who would be able speak to the assessment of collaborative assignments in EAP. They were all experienced EAP professionals with an average 10–15 years' experience teaching EAP and the majority held MAs with some holding PhDs. Rather than coming from a single institution, participants were selected from various teaching contexts, drawing on 15 different higher education institutions in various geographical locations (see Table 1).

Data collection tools: pre-interview task & intensive interviews

In a pre-interview task, participants were asked to evaluate four different models of assessment for a group essay assignment (3 students per group following a process writing approach), which were drawn from Race (2001) and are outlined in Table 2.

For the purpose of the models, the group product of the collaborative assignment (the essay and drafting process) received a grade of 70% and each of the students

Table 2. Models of collaborative assessment used in the pre-interview task.

Model A	Group Grade	All students in the group receive the same grade, based on the quality of the group product
Model B	Group grade + individual complementary assignments	All students in the group receive the same grade, based on the quality of the group product + an individual grade for complementary assignments
Model C	Group grade + individual collaboration grade	All students in the group receive the same grade, based on the quality of the group product + an individual collaboration grade, based on criteria
Model D	Group grade x weighted collaboration grade	Students receive the grade for the group project, adjusted by their relative contribution to the other group members

received a high, mid or low grade for the other component of the assessment scheme. The results of each student were then illustrated, showing the potential for differentiation within each assessment scheme. Intensive interviews (averaging just over an hour long each) followed the pre-interview task and were the principal source of data collection as a means to examine the topic by engaging with participants with relevant experience (Charmaz 2014).

Data analysis

Data analysis was driven by coding, memoing and diagramming, in line with GT methodology. Four pilot interviews provided the data set used for generating the initial codebook. A line-by-line coding approach was adopted, which would lay the foundation for future coding cycles. The aim of using line-by-line coding was to force close reading of the interview data and to avoid attention being directed by researcher bias. Simultaneous coding was allowed, with a single line of interview transcript potentially carrying multiple codes. Coding generally, and certainly within a GT approach, is cyclical, encouraging re-engagement with the data. The line-by-line coding led to open coding, during which patterns and nascent concepts emerged. These were followed up by memoing, diagramming and focused coding and allowed for theoretical sampling (Birks and Mills 2023; Charmaz 2014) which, in practical terms, means choosing where to get data (e.g. who to interview) and what data to collect (e.g. what to ask in the interview). This led to the final round of coding: theoretical coding. Theoretical coding in this context follows the Saldaña (2021) conception of bringing concepts together in a way that ‘suggests a theoretical explanation for the phenomenon’ (314).

Results and discussion

The following section provides an overview of the participants’ evaluation of the different models of assessment before discussing the issues around product-focused assessment, the use of individual complementary assignments and assessing the collaboration itself. It closes with discussion of the participants’ proposals for models of collaborative assessment, indicating the need for a multi-lens approach, directed by the learning outcomes of the course.

Evaluating models of assessment

One component of the pre-interview task was participants assigning a star rating out of 5 to each model. This was not intended as a robust statistical measure but more as a general overall impression rating. As seen in Figure 3, Model B, incorporating individual complementary assignments, was the most favoured approach but there was also considerable appetite for assessing the collaboration itself, albeit recognition that there was no clear means to do so. Model D was the least well received model and was seen as a more complex version of Model C so it will not be discussed in depth in this paper.

The models of assessment led to discussion of the final product, individual complementary assignments and the collaboration itself as assessment lenses, considering the affordances and reservations of each.

The final product as assessment lens

The first model of assessment discussed was product-focused assessment, with Figure 4 below illustrating the main arguments around the approach.

The model was most highly rated by Fay and Lorenzo. Fay viewed it positively, with the proviso the teacher is 'present' during the process of essay development in the students' shared document. One of the features of Model A that appealed to Lorenzo was that it was output driven, meaning the focus was on the work the students produced. As Cristobal suggested,

Essentially, if they do the research, and they work together, and they write it, it will come out as a good product. So, you know, ideally, the product should be the easiest way of and most effective and most accurate way of assessing their work.

As well as simplicity, there was also a realworldness to a product-focused approach, as that is how collaboration would likely be measured outside the academy. Fernand recognised it could be an acceptable model if the aim of the assignment was collaboration, with students aware they had to work together to get a good result. However, he still thought there had to be scope for some differentiation between students and an opportunity to 'reward them for effort', which was a theme that came up throughout his interview.

While acknowledging the efficiency of Model A, Bertha felt it could be the most problematic of the models due to potential unequal contributions by group members

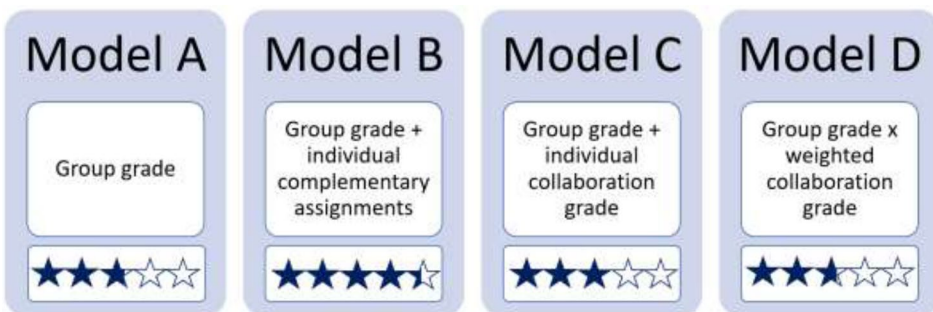


Figure 3. Participants' star rating for the four models of collaborative assessment.

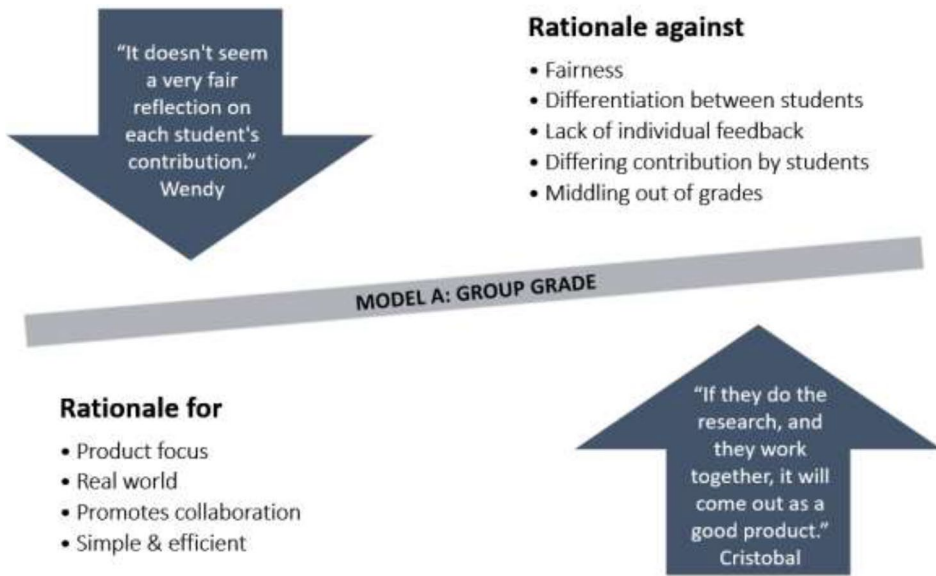


Figure 4. Summary of evaluation of model A: Group grade (Levrai 2025, 167).

and a lack of individual feedback. This highlights what was the most widely commented issue: the unfairness of an assessment scheme not allowing for differentiation. Indeed, differentiation was the most widely cited concern and, as Tanya put it, ‘it’s straightforward, it’s easy to understand, but I do think there’s some things lacking here as far as nuance between students’. Kyle considered the impact on motivation since it could encourage free-riders as all the students would get the same grade regardless of how much work they did. Connected to this, Laura also saw the potential for demotivation in laid back students in the face of a group grade since ‘she knows that she will receive the same grade with the high achievers’. These concerns illustrate the kind of behavioural misalignment Meijer et al. (2020) warned against if adopting this kind of assessment approach, which does not foster individual accountability.

Nonetheless, the final output of a collaboration has to be assessed, as the first assessment lens. If students are working together to produce a piece of work, the quality of that final work matters and if the work is not graded, students could have limited motivation to engage effectively with it. Even if the assignment’s prime concern is the collaboration itself, the product of that collaboration is an indicator of the work that was done, with Van arguing the coherence of a text is ‘a kind of inferred feature of the, how good or bad the collaboration was’. The final product is not, however, a perfect indicator. As Storch and Knoch (2023) and Van Aalst (2013) hold, an individual’s abilities are not easily assessed through a piece of collaborative writing. Nor is the final work a clear measure of collaboration, especially given good work could be produced through dysfunctional collaboration (see Figure 5)

A good quality final product could mask problematic collaboration (e.g. the students adopt a cooperative ‘divide and conquer’ approach, without really engaging with each other’s ideas or work). The quality of the final product-only evidences



Figure 5. How collaborations of differing quality can lead to a good product.

the quality of the final product, not how it was produced. As Barry explained, assessing the product is ‘limited in terms of actually producing any, if you like, outcome for recognising the value of collaboration’.

While it was not the preferred approach, the participants felt that adopting product-focused assessment could be appropriate if:

- it meets the course learning outcomes.
- there is focus on the process and ongoing teaching involvement and feedback, so students learn from the experience.
- it is a low stakes component of a wider assessment scheme which allows for assessment of the individual.

Individual complementary assignments as assessment lenses

The second set of assessment lenses encompass individual complementary assignments, which were generally favoured by the research participants. From the outset, it is important to note that the limitations inherent in solely addressing the final product were still recognised in an assessment scheme incorporating individual complementary assignments. As Wendy noted, ‘you’re not assessing their collaboration at all actually’. Although the individual components *may* be an indicator of engagement with the main collaborative task, they add little additional insight into the collaborative process unless specifically designed to do so. Nonetheless, utilising complementary assignments was the most favoured assessment scheme, with an overview of the pros and cons in [Figure 6](#).

The model was most highly rated by five participants. Olga appreciated the opportunity to see individual work and considered the assessment scheme was easy to justify to students. Fay felt there was a value to the individual element as it ‘can ensure that everyone is doing some work’. Ava noted the group grade aspect

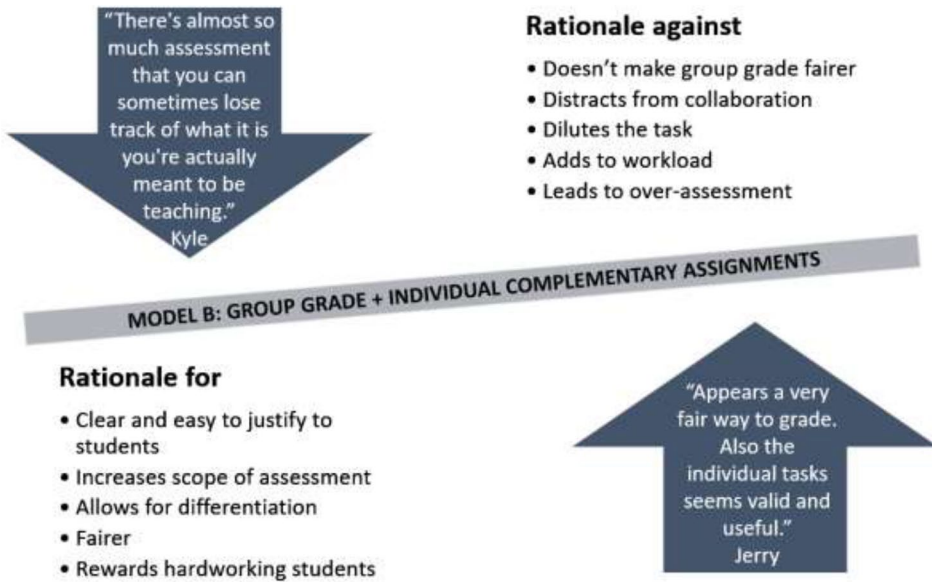


Figure 6. Summary of evaluation of model B: Group grade+ individual complementary assignments (Levrai 2025, 176).

would encourage collaboration since the student would work together to produce a shared piece of work but appreciated the individual component 'merits those who perhaps work a little bit harder'. Van valued the individual work for a slightly different reason, recognising it could help identify students who may need additional support. The lowest rating of three stars was given by Fernand. He still expressed concern around the group grade aspect and the potential of that to artificially inflate some students' results but appreciated the opportunity for giving individual grades and, in an echo of Ava, rewarding students for how hard they worked.

One of the benefits of Model B was that it was 'clear to follow' (Olga) and more comprehensive, thereby addressing the shortcomings of Model A. Cristobal noted, 'I like the way it brings in other skills ... it makes the course bigger, because it brings in other skills'. The addition of the individual element was seen as something that worked effectively with the group task as it could increase student motivation and incentivise students, although Bertha was concerned an individual component could send 'mixed messages'. The main appeal of Model B, however, lay in the individual component providing the opportunity for differentiation between students and the opportunity to see individual work that students produced. Humberto considered 'this has more credibility perhaps to the students, because they're seeing that their own efforts are getting a bit more emphasis'. This resonates with Karen's point that students 'want to be *noticed* for what they've contributed'. Wendy valued the assessment being based on something concrete, the individual work produced by students, rather than 'just a teacher's 'sense' of how much effort they've made'.

Although Model B was generally well received, it was not without criticism. While Kyle acknowledged there was still a simplicity to Model B, he felt that it was 'more difficult to manage from an operational point of view' in comparison to Model A.

He also had a concern about over-assessment, explaining ‘at a certain point, there’s almost so much assessment that you can sometimes lose track of what it is you’re actually meant to be teaching’. Lorenzo also questioned the number of assessments and Andrea disliked bringing in other tasks which assess ‘something different other than the assignment itself’. Related to this, Lorenzo also shared a concern expressed by Bertha and Kyle, considering excessive tutor workload with additional items to mark. In addition to the grading, Bertha voiced concern about having enough time to prepare students for the individual tasks.

Despite concerns, the use of individual complementary assignments was the preferred approach, on condition that:

- all assignment components must lead into learning aims
- individual tasks are supportive of the main collaborative assignment
- assessment is weighted towards the collaborative product
- there is time for relevant learner training and the assessment load

Individual complementary assignments should have a purpose that either a) helps the main collaborative task, b) meets a distinct learning aim or c) ideally, both of the former. The workload, for teachers and students, would need to be considered to ensure no overload and the relationship between assignments and the positive washback from one to the other would need to be clear to students, and teachers, from the outset.

Collaboration itself as an assessment lens

The third set of assessment lenses are those focused on the collaboration itself. The assessment of collaboration is particularly challenging, although, as Humberto recognised ‘If you *can* do it without screwing up the entire operation and making it too surgical, then yes, I think it could and should [be assessed]’. [Figure 7](#) illustrates the main arguments for and against assessing collaboration itself.

While no participants awarded Model C five stars, five of the participants gave it a four-star rating, albeit with considerable qualification. One of the positives about Model C was grading the collaboration, in the sense that grading collaboration ‘gives a message to students that, you know, your collaboration matters’ (Laura). Dolly appreciated the opportunity to ‘reward collaboration, because I’m a believer in collaboration’ and Lorenzo liked being able to reward an individual for the quality of their collaboration. Bertha advocated for the evaluation of the collaboration to be done through peer-assessment rather than teacher-assessment, questioning whether the teacher should ‘assess the collaboration if they weren’t actively involved in the collaboration’. Her previous experience with peer-assessment had shown that it could be effective and was a better measure of collaboration than teacher-assessment, which could be quite arbitrary.

The biggest issue with Model C was the question of how to assess collaboration, which was raised by all participants, and this was encapsulated in Wendy’s pre-interview task, where she wrote,

Assessing collaboration is quite subjective. How can the teacher know how much each person put into the groupwork? So much of it happens outside the classroom, and manifests itself in ways the teacher can’t access (i.e. replying to messages, being open to other’s ideas, easy to work with etc).

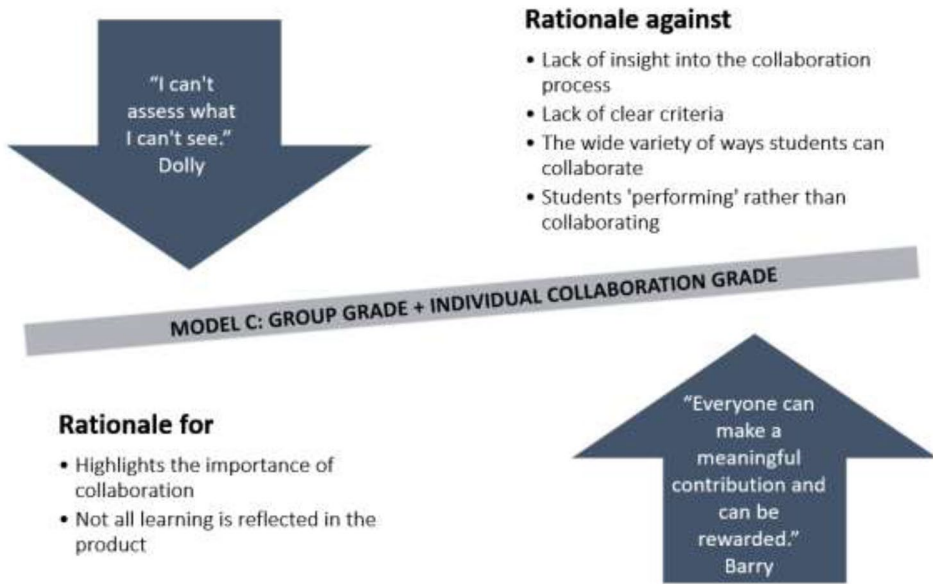


Figure 7. Summary of evaluation of model C: Group grade + individual collaboration grade (Levrai 2025, 181).

Both Cristobal and Olga gave Model C the lowest possible one-star rating. In his pre-interview task, Cristobal claimed, 'it would be too difficult to assess collaboration - the teacher could not do it and the students could be too biased!'. He expanded in his interview, discussing the difficulty of knowing what someone had contributed to a collaboration and how this was complicated by the various ways in which people could contribute (e.g. through research, ideation, logistics). Olga also drew attention to the varied ways students could collaborate, questioning if it should be defined 'in terms of how much percentage of the writing they do, would that count as a collaboration or the background research, should that count as collaboration?'. In terms of facilitation, Cristobal was probably the most deeply embedded of the participants in the student groups, taking an active role in their shared work documents, providing ongoing feedback during the drafting process and operating almost as another group member. However, he still felt there was too much he did not see to be able to reliably assess collaboration. He also expressed concern that any mechanisms to try to track contribution could just become an exercise in box ticking with 'the teachers doing all this admin having to collect all these different things' when the product was already there to be assessed.

Criteria driven assessment of collaboration. During the interviews, participants spoke of the need for clear assessment criteria which are shared with students so that they understand the expectations of the assignment. While it was felt that collaboration criteria would provide clarity for students about what was expected, the participant's shared Strijbos (2016) concern about performative collaboration and that criteria shape the student experience and could ultimately be reductive (Davidson 2004; Lantolf and Frawley 1985). When behaviours are codified in criteria they become

the target behaviour for both teachers and students potentially at the expense of other natural and beneficial behaviours. However, the process of teachers negotiating collaboration criteria will be highly beneficial in leading to careful thought of what ‘good’ collaboration looks like in their context. Olga’s suggestion that ‘perhaps, there must be more studies into [assessment criteria for collaboration]’ also points the way towards future research.

Utilising self- and peer-assessment to evaluate collaboration. Two other potential lenses for assessing collaboration are self- and peer-assessment. Within the high-stakes realm of EAP assessment, self- and peer-assessment need to be treated with caution. In a context where students have the assessment literacy to self- or peer-assess, such assessment lenses could have great value. In contexts where students may lack the assessment literacy to engage in such assessment a cost/benefit analysis of the learner training required would need to be made. In assessment literate cohorts, self- and peer-assessment would provide a straightforward assessment mechanism the students are already au fait with. However, the initial caveat of ‘In assessment literate cohorts’ cannot be understated and EAP practitioners may not find themselves with such assessment literate students. As Fernand explained when talking about his context, self- and peer- assessment would be ‘completely alien to the students’.

Despite the challenges, there was an appetite for assessing collaboration if:

- collaboration is a key learning outcome of the course
- there is a clear conception of how the students are expected to collaborate
- those expectations are clear, with explicit discussion of collaboration at the outset
- there are assessment criteria, which are clear to the students (or which are developed in discussion with the students)
- a minor proportion of the grade is awarded to collaboration.

Participants’ proposals of a model of assessment

In the early iterations of the interview, after discussion of the four models of assessment, participants were asked which assessment lens(es) they would use to assess a collaborative assignment. This was with a view to asking participants to give a weighting to each component to construct an assessment scheme for collaborative assignments. However, during the pilot interviews it became clear that expecting participants to spontaneously provide a well-considered assessment scheme with assessment lenses and weighting decisions was unrealistic. This is not least because it quickly became clear that for the participants any assessment scheme would be dependent on the learning aims of the course in question, meaning there is no single answer to the question.

When discussing how to assess collaborative assignments, only two participants did not favour assessing the collaboration. Tanya felt students in her context could be relied upon to work together so there was no need for direct assessment of collaboration (although other participants working in her context disagreed).

Van was the most resolutely against assessing collaboration as he believed that it is already done implicitly through assessing the work the group produces, meaning the product is the only assessment lens needed. For all the other participants, the feeling was collaboration should be assessed in collaborative assignments if it is a learning outcome. Further to this, 10 of the participants favoured an assessment scheme comprising the three components of the final work, individual complementary assignments, and the collaboration itself. There were caveats:

- the lenses selected and the weightings were outcome dependent.
- collaboration and individual complementary assignments were valuable but would necessarily be the minor part of the assessment.
- assessing collaboration is desirable but our scope for successfully assessing collaboration is limited.

Nonetheless, this does indicate both practitioners' acceptance of multi-lens assessment for collaborative assignments and a need for collaboration to be assessed.

It is important to note, however, that while there are an array of assessment lenses available to evaluate a collaborative assignment, which lenses to employ will depend on wider course outcomes, the profile and assessment literacy of the student cohort, and what is practically possible for students and teachers for that particular assignment as part of the fuller workload of the whole course. Each assessment lens will have a learner training opportunity cost, which will take time away from the learning training for the assignment itself and potentially lead to task dilution. Multiple lenses *could* be employed in service of assessing a collaborative assignment, all of which contribute to a single overall grade. To assuage Van's fears of duplication and unnecessary assessment, there must be value to an assessment lens in and of itself, with each lens aligned to the learning outcomes of the course and leveraged for positive washback and increased learning gains.

Efforts must also be made to maximise the learning affordances of each assessment lens. For an example of maximising an assessment lens, consider groupmates being required to give feedback to each other on the work they produce. There is potential for this to descend into busy work if the task is phrased as 'Provide feedback on the contributions of the other members of the group'. However, if it is phrased as 'Provide feedback which should have a positive impact on the quality of the final product' then students will necessarily be thinking about the consequence of their feedback and be orientated towards providing feedback to improve the essay.

Learning outcome directed lens selection

A variety of different assessment schemes can be justified when evaluating a collaborative assignment, with the learning outcomes of the course directing the assessment lenses selected in each case. Decision making for the design of a collaborative assignment marking scheme should be driven by the importance of collaboration as a learning outcome, as that is what determines how thoroughly collaboration needs to be assessed, with the need to assess collaboration increasing the more central collaboration is as a stated outcome.

If collaboration is not a learning outcome, the recommendation is that collaborative assignments are avoided. This is because collaboration takes time and resources and, in already busy EAP courses, the work that would go into facilitating collaboration could well be better spent elsewhere. In the case where collaboration is a secondary learning outcome and potential assessment lenses do not help meet other learning outcomes, the recommendation is to base assessment solely on the product. This is an unexpected recommendation but if there are no relevant learning gains that can come from individual complementary assignments or assessing the collaboration itself, then a product-only approach is justified. However, that does not mean collaboration can be neglected. It would still have to be facilitated and supported, with mechanisms in place to address group dysfunction. This reflects Fay's approach when talking about the need for teacher-group meetings, classroom observation and potentially group mediation in case of a problem during the facilitation of collaborative assignments, seeing them as vital elements despite their not being graded. There should likely also be scope for individual submission of final work in the case of full group breakdown.

Where collaboration is a secondary learning outcome and other assessment lenses (see [Figure 1](#)) would help meet other learning outcomes, then lenses which have learning benefit can be employed alongside the main group product. This would have to be clearly rationalised to students so that they understand the purpose of and need for each part of the assessment. When collaboration becomes a core course outcome, it necessarily carries more weight and will require more robust assessment. If other assessment lenses do not contribute to expected learning outcomes, assessment will be product-focused again. However, collaboration will also have to be assessed as a core learning outcome and there are two ways this could be done. The first is assessing the collaboration that happened, which can be done through observation and student reporting. Given that these will give incomplete insight, the collaboration grade would necessarily be the smaller part of the assignment. Alternatively, rather than assessing the collaboration that was, attention could turn to awareness of collaboration, incorporating an individual grounded reflection. In this type of reflection task students are given some learning theories, with the reflection geared towards how theory explains their experiences, or where their experiences differed from theory, and why. If a student is able to demonstrate understanding of their own roles, contributions and behaviours and how their team worked, this would indicate their awareness of collaboration. If collaboration is a core learning outcome and potential assessment lenses also have utility in addressing other learning outcomes, a multi-lens assessment approach can be used, with each lens rationalised to students so they know the purpose of each and how they all cumulatively fit together as a full scheme of assessment.

Conclusion

As a context-free study, this paper speaks to global concerns held by EAP practitioners around different assessment approaches for collaborative assignments. What it lacks is focus on a particular teaching context to examine the impact of institutional assessment culture on practitioner attitudes and beliefs and the evaluation

of assessment schemes in relation to an actual collaborative assignment the participants are working with. Nonetheless, it has brought the practitioner voice into the discussion of what to assess when assessing collaboration, a necessary step in bringing theory and practice closer.

We know that the assessment of collaboration is going to be flawed and necessitate compromises, no matter how many assessment lenses are employed or who conducts the assessment, be that the teacher-assessment, self-assessment, peer-assessment or teacher-mediated self- and/or peer-assessment. As is clear, no single assessment scheme will fit every context, course, or assignment. In each case the appropriate approach will have to be rationalised ‘in-house’. The benefit of this is that the assessment scheme will be fully understood by the teachers enacting it and the negotiation of it could prompt fuller engagement and investment from teachers. We also know the assessment of collaboration could include a lot of work and effort for very marginal gains. This necessarily means collaboration cannot be weighted too heavily in an assessment scheme, but it should be weighted at a significant enough level to matter. The evaluation of the collaboration itself is not going to be perfect, but, taken in concert with assessment of the final product and, potentially, individual complementary assignments, a collaborative assignment can productively sit amongst other course assessments.

Acknowledgments

Thanks to the research participants for sharing their time and expertise, to Dr Maria Del Pilar Garcia Mayo and Dr Averil Bolster for their valuable feedback.

Ethics

The plan adhered to the ethical principles of research with human participants and ethical review in the human sciences in Finland. Informed consent was obtained from all participants, who were informed of the purpose of the study at the outset. Participation was entirely voluntary, and participants were free to withdraw at any time. Their anonymity is guaranteed.

Disclosure statement

No funding was provided, there are no competing interests to declare and AI was not used in the development of this work.

ORCID

Peter Levrai  <http://orcid.org/0000-0002-1917-3695>

References

Beard, J. D., J. Rymer, and D. L. Williams. 1989. “An Assessment System for Collaborative-Writing Groups: Theory and Empirical Evaluation.” *Journal of Business and Technical Communication* 3 (2): 29–51. <https://doi.org/10.1177/105065198900300203>.

- Bell, D. 2024. *English for Academic Purposes: Perspectives on the Past, Present and Future*. Vol. 122. Bristol, UK: Channel View Publications
- Birks, M., and J. Mills. 2023. *Grounded Theory: A Practical Guide*. 3rd ed. London, UK: Sage Publications.
- Bloxham, S., and P. Boyd. 2007. *Developing Assessment in Higher Education: A Practical Guide*. Maidenhead, England: McGraw-Hill Education.
- Bolster, A. (2025). "A Qualitative Grounded Theory Study of EAP Practitioners and Collaborative Learning: Identity and Beliefs." Unpublished doctoral diss., The University of the Basque Country.
- Bolster, A., and P. Levrai. 2022. "Are we Talking about the Same Thing? Researcher and Practitioner Perspectives of Student Collaboration." In *Innovation, Exploration and Transformation – Proceedings of the 2019 BALEAP Conference*, edited by M. Evans, B. Bond, and A. Ding, 96–104. Reading, UK: Garnet Education.
- Boud, D., and M. Bearman. 2024. "The Assessment Challenge of Social and Collaborative Learning in Higher Education." *Educational Philosophy and Theory* 56 (5): 459–468. <https://doi.org/10.1080/00131857.2022.2114346>.
- Brown, G. 2011. "Teachers' Conceptions of Assessment: Comparing Primary and Secondary Teachers in New Zealand." *Assessment Matters* 3: 45–70. <https://doi.org/10.18296/am.0097>.
- Charmaz, K. 2014. *Constructing Grounded Theory*. 2nd ed. London, UK: Sage Publications.
- Davidson, C. 2004. "The Contradictory Culture of Teacher-Based Assessment: ESL Teacher Assessment Practices in Australian and Hong Kong Secondary Schools." *Language Testing* 21 (3): 307–334.
- Ferreira, J. M., K. Zabolotna, and S. Lee. 2024. "Teaching Twenty-First-Century Skills: Examining Collaborative Learning in Initial Teacher Education in Finnish Universities." *Scandinavian Journal of Educational Research* 1–22. <https://doi.org/10.1080/00313831.2024.2419078>.
- Gammie, E., and M. Matson. 2007. "Group Assessment at Final Degree Level: An Evaluation." *Accounting Education* 16 (2): 185–206. <https://doi.org/10.1080/09639280701234609>.
- Gibbs, G. 1999. "Using Assessment Strategically to Change the Way Students Learn." In *Assessment Matters in Higher Education: Choosing and Using Diverse Approaches*, edited by S. Brown and A. Glasner, 41–53. Buckingham, UK: McGraw-Hill Education.
- Graves, K. 2016. "Language Curriculum Design: Possibilities and Realities." In *The Routledge Handbook of English Language Teaching*, 79–94. Abingdon, UK: Routledge.
- Johnston, L., and L. Miles. 2004. "Assessing Contributions to Group Assignments." *Assessment & Evaluation in Higher Education* 29 (6): 751–768. <https://doi.org/10.1080/0260293042000227272>.
- Lantolf, J. P., and W. Frawley. 1985. "Oral-Proficiency Testing." *The Modern Language Journal* 69 (4): 337–345. <https://doi.org/10.2307/328404>.
- Levrai, P. 2025. "Exploring Practitioner Collaborative Assessment Identity to Develop a Principled Multi-Lens Approach to Assessing Collaborative Assignments In English For Academic Purposes." Unpublished doctoral diss., University of the Basque Country.
- Li, M., and M. Zhang. 2023. "Collaborative Writing in L2 Classrooms: A Research Agenda." *Language Teaching* 56 (1): 94–112. <https://doi.org/10.1017/S0261444821000318>.
- Meijer, H., R. Hoekstra, J. Brouwer, and J. W. Strijbos. 2020. "Unfolding Collaborative Learning Assessment Literacy: A Reflection on Current Assessment Methods in Higher Education." *Assessment & Evaluation in Higher Education* 45 (8): 1222–1240. <https://doi.org/10.1080/02602938.2020.1729696>.
- Nepal, K. P. 2012. "An Approach to Assign Individual Marks from a Team Mark: The Case of Australian Grading System at Universities." *Assessment & Evaluation in Higher Education* 37 (5): 555–562. <https://doi.org/10.1080/02602938.2011.555815>.
- Queen Mary Academy. n.d. "Intended Learning Outcomes." Queen Mary Academy. <https://www.qmul.ac.uk/queenmaryacademy/educators/resources/curriculum-design/intended-learning-outcomes/>

- Race, P. 2001. *A Briefing on Self, Peer and Group Assessment*. York, UK: LTSN Generic Centre.
- Richards, J. C. 2013. "Curriculum Approaches in Language Teaching: Forward, Central, and Backward Design." *RELC Journal* 44 (1): 5–33. <https://doi.org/10.1177/0033688212473293>.
- Saldaña, J. 2021. *The Coding Manual for Qualitative Researchers*. 4th ed. London, UK: Sage Publications Ltd.
- Storch, N. 2017. "Implementing and Assessing Collaborative Writing in EAP Classes." In *Teaching Writing for Academic Purposes to Multilingual Students: Instructional Approaches*, edited by J. Bitchener, N. Storch, and R. Wette, 130–142. New York, USA: Routledge.
- Storch, N., and U. Knoch. 2023. "Assessing Collaborative Writing: Current Challenges and Future Possibilities." In *L2 Collaborative Writing in Diverse Learning Contexts*, edited by M. Li and M. Zhang, 229–247. Amsterdam, The Netherlands: John Benjamins.
- Strijbos, J. 2016. "Assessment of Collaborative Learning." In *Handbook of Social and Human Conditions in Assessment*, edited by G. T. L. Brown and L. R. Harris, 302–318. New York, USA: Routledge.
- Van Aalst, J. 2013. "Assessment in Collaborative Learning." In *The International Handbook of Collaborative Learning*, edited by C. Hmelo-Silver, C. Chinn, C. Chan & A O'Donnell, 280–296. New York, USA: Routledge.
- World Economic Forum. 2016. "New Vision for Education: Fostering Social and Emotional Learning through Technology." World Economic Forum. https://www3.weforum.org/docs/WEF_New_Vision_for_Education.pdf
- Zhang, M., and W. Chen. 2022. "Assessing Collaborative Writing in the Digital Age: An Exploratory Study." *Journal of Second Language Writing* 57: 100868. <https://doi.org/10.1016/j.jslw.2022.100868>.