

Assignment 3B

Jeroen Rouwkema

From: Examination Board BMT/BME
To: Program Director BMT/BME
Date: 16-11-2023
RE: Use of summative peer assessment in the BMT/BME program

Dear Program Director of the program BMT/BME,

The Examination Board is of the opinion that the use of summative peer assessment can be a potential threat to the quality of assessment within a course. Even though it is possible to use summative peer assessment in a way that safeguards the quality of assessment, doing so is complex and will require additional training of the examiner. If this current situation is not addressed, there is a risk that the Examination Board will not be able to validate the quality of examination within courses that use this assessment method. As there is no clear overview on the use of summative peer assessment in the BMT/BME program, it is unclear whether the risk that the use of summative peer assessment negatively influences assessment quality in courses within the BME program is hypothetical or actually present. To ensure that the risks regarding the use of summative peer assessment are minimized, program management should obtain a good overview of courses that use summative peer assessment. Additionally, examiners that do use summative peer assessment should be supplied with an instruction leaflet on the proper use of assessment method. Combined, this will result in summative peer assessment that is valid and reliable, which will allow the Examination Board to validate the quality of examination.

Assessment competence is one of the key elements governing quality of assessment within a program. Sufficient expertise must be present at the organization level and amongst examiners to ensure quality of assessment at all levels of the quality of assessment pyramid. Within the BMT/BME program, the Rules and Guidelines of the Examination Board stipulate that the examination board assigns examiners, who are responsible for the grading of exams and tests. Further, the Examination Board has defined that a) the educational level of the examiner is one level higher than the level of the course for which this person is appointed as an examiner, and b) the examiner must possess sufficient teaching qualifications, in particular in the area of examination.¹ However, even though it is clear that a certified examiner is responsible for assessment within a course, the Examination Board has noted that for several courses summative peer assessment (here defined as assessment by other student participants within a course, so excluding assessment by for instance student assistants) is included as an assessment method. The Examination Board is of the opinion that the use of summative peer assessment can be a threat to the quality of assessment within a course, if the formal examiner is not sufficiently aware of the risks and limitations connected to this form of assessment. If this current situation is not addressed, there is a risk that the Examination Board will not be able to validate the quality of examination within these courses.

Peer assessment is often used as an aid in learning, by providing insight in how student peers approach a problem or an assignment. Additionally, peer assessment can be used as a tool to alleviate the workload of a teacher, especially when large groups of students need to be evaluated. For both goals, peer assessment can be used as a formative or summative assessment tool. Peer assessment can offer clear advantages and is regularly used in higher education, but generally little attention is given to the potential risk that the quality of assessment will be more difficult to ensure and safeguard. By delegating part of the assessment tasks to students who are not trained or skilled in assessment, proper implementation of peer assessment is needed to ensure assessment validity and reliability. Even though it is possible to use summative peer assessment in a way that safeguards the quality of assessment, doing so is complex and will require training of the examiner. As peer assessment is not a topic regularly addressed in the standard training of teachers, it cannot be expected that all examiners are aware of the boundary conditions that need to be taken into account to properly implement summative peer assessment.

¹ Rules and Guidelines of the Examination Board for the Master's Programme BME, July 14th 2022

There is no clear overview on the use of summative peer assessment in the BMT/BME program, especially for the MSc curriculum. Due to this, it is unclear whether the risk that the use of summative peer assessment negatively influences assessment quality in courses within the BMT/BME program is hypothetical or actually present. Even when the risk is hypothetical, proactive action should still be taken, especially since new courses can be developed or existing courses can be adapted to contain summative peer assessment. Educational developments such as an increasing focus on Challenge Based Learning can also increase the use of peer assessment or other forms of assessment by non-qualified individuals such as commissioning entities.

In order to ensure that the risks regarding the use of summative peer assessment within the BMT/BME program are minimized, the following steps should be taken:

1. Develop an instruction leaflet for the proper use of summative peer assessment.

An instruction leaflet should be developed containing rules and best practices for the use of summative peer assessment. By complying to these rules and best practices, examiners can implement summative peer assessment in such a way that the validity and reliability of the assessment task can be safeguarded. As the instruction leaflet is intended to enhance the assessment competence of examiners, it is the primary responsibility of the program management that the leaflet is prepared. The expertise of the Centre of Expertise in Learning and Teaching (CELT) can be sought to aid in the drafting of the leaflet. Further, the following can be taken into account:

- For a good assessment validity, it is important that the peer-assessment focuses on the correct content, at the correct level. This is especially important if the peer-assessment is used as a summative evaluation. Even though it has been noted that the traditional criterium of validity is not always relevant when the assessment is used as a formative tool,² a proper alignment with the learning objectives and the intended level of mastery is still important for an optimal learning experience. A misalignment may result in an over- or underestimation of their own learning progress by the assessors. All in all, for good assessment validity, it is important that the assessment process and the assessment criteria are clearly defined by the examiner, in such a way that it is clear for the student assessors what should be assessed according to which criteria.
- Additional measures and boundary conditions may need to be defined in order to ensure that students are awarded their study credits based on a just and fair evaluation instead of unfair or biased peer assessment. It is also important that there is consistency in the appraisal and that the same criteria and conditions are used for all students who are evaluated. In order to ensure this, the following should be considered:
 - Literature showed that reliability of peer-assessment by students can approach assessment by experts when assessment criteria are clearly defined.³ As such, clear instructions and a clear rubric are crucial.
 - A blinded or double-blinded assessment approach can help to reduce biased scoring where students for instance score the work of friends higher than the work of non-friends, or where student assessors may not feel comfortable to provide critical feedback to specific individuals.
 - As multiple biases can arise when using summative peer assessment, every student should be assessed by multiple other students to ensure a high level of reliability. A study by Cho et al suggests that every student should be assessed by four to six peers to achieve high reliability.⁴ It has further been pointed out that making students aware of their biases decreases the bias and improves assessment reliability.⁵
 - Students peer assessment can result in different results, including rating (cardinal evaluation), ranking (ordinal evaluation), or the provision of feedback. Even though cardinal evaluation generally most resembles the evaluation situation when peer-assessment is not used, it may not be the optimal option for peer assessment. In a study by J.T. Venderink, students recurrently expressed

² The interplay of evidence and consequences in the validation of performance assessments, S Messick (1994), *Educational Researcher*, 23; p13-23

³ The importance of marking criteria in the use of peer assessment, P Osmond, S Merry, K Reiling (1996), *Assessment and Evaluation in Higher Education*, 21; p239-250

⁴ Validity and reliability of scaffolded peer assessment of writing from instructor and student perspectives, Cho et al, (2006), *Journal of Educational Psychology*, 98(4); p891-901

⁵ Peer and self assessment in massive online classes, C Kulkarni et al (2013), *ACM Transactions on Computer-Human Interaction*, 20(6); p1-31

negative perceptions on the use of cardinal evaluation in peer assessment.⁶ Even if there is still a teacher involved in establishing the final grade, this sentiment is not beneficial for the willingness of students to take part in peer-assessment, which can negatively influence assessment reliability. Other papers have also stipulated that ranking is better than rating, as this form can help to reduce noise and bias in the assessment.⁷

Based on this, it is our opinion that the leaflet should at least include the following requirements when summative peer assessment is used:

- The availability of clear instructions and a clear rubric for the student assessors.
- The drafting of a peer assessment plan providing a substantiated discussion of the following:
 - How students will be prepared for their assessment task
 - Whether a non-blinded, a single blinded, or a double blinded approach will be used
 - Whether rating (cardinal evaluation), ranking (ordinal evaluation), provision of feedback, or a combination of these will be used as the assessment outcome
 - How the examiner will go from the peer assessment results to the final assessment of the student
- A minimal number of peer assessors in the case of cardinal or ordinal evaluation

2. Identify courses in the BMT/BME curriculum that use summative peer assessment

Program management should obtain a good overview of the courses that use summative peer assessment in the BMT/BME program. This will not only be important for determining which examiners should be supplied with the instruction leaflet for the proper use of summative peer assessment, but will also provide data that can be included in the assessment program. For this, program management should actively gather the assessment plans of courses that have not provided such a plan recently. Additionally, the Program Committee (OLC) and Quality Assurance Committee (OKC) can be asked to carefully assess whether (new) courses use summative peer assessment. This can be part of the regular course assessment cycle.

3. Supply examiners that use summative peer assessment with the instruction leaflet

As discussed with you during a meeting on the 2nd of November 2023, the goal of the information leaflet should not be to increase the overall use of summative peer assessment by drawing attention to this assessment method, but to ensure that examiners who already use summative peer assessment do so in a proper way. For this reason, our advice is that the program management initially only supplies the examiners that are actively using summative peer assessment in their courses with the instruction leaflet. Ideally, an active follow-up system should also be set in place to ensure that the examiners are properly incorporating the requirements communicated in the leaflet. Additionally, it will be beneficial to include a view on the use of summative peer assessment in the assessment policy of the program, which should mention the availability of the instruction leaflet. This way, the position of the program management on the use of summative peer assessment will be clarified, and (new) teachers who are thinking about implementing summative peer assessment in their course can actively reach out to obtain the instruction leaflet.

Combined, these steps will result in summative peer assessment that is valid and reliable, which will allow the Examination Board to validate the quality of examination within courses that use this assessment method.

The Examination Board would appreciate your timely response.

With kind regards,

Jeroen Rouwkema
Member Examination Board BMT/BME

⁶ Using peer assessment in online courses: a qualitative approach to deciding which behavioural and structural variables to incorporate in MOOC design, J.T. Venderink (2022) MSc thesis Utrecht University

⁷ A Case for Ordinal Peer-evaluation in MOOCs, N.B. Shah *et al*, (2013), 1-10