

Developing an Assessment Checklist

Maria Maistro

Department of Computer Science
University of Copenhagen

Summary. The goal of this project was to prepare an assessment checklist for the Information Retrieval (IR) course at the department of Computer Science in the academic year 2019/2020. This project was motivated by some observations regarding the previous edition of the IR course: in the 2018/2019 edition there was a clear mismatch between students and teachers expectations regarding the assignment. The students were struggling in understanding how to structure and write a good quality assignment. Furthermore, even if the students were instructed with guidelines on how to give feedback, they were struggling also in providing useful feedback to their peers. With the proposed assessment checklist, we aimed at guiding and helping students in structuring their assignment and peer reviews.

This paper is organised as follows: Section 1 describes the course during which the project was carried out; Section 2 presents the project goals and motivations; Section 3 carefully describes how the project was conducted; Section 4 reports some analysis about the project results; and Section 5 presents conclusions and future challenges.

Context of the Project

The project was carried out during the Information Retrieval (IR) course¹ at the Computer Science Department in the academic year 2019/2020. The

¹ The official course description is available at: <https://kurser.ku.dk/course/ndak15005u/2019-2020>

course is part of the Master Programme in Computer Science, corresponds to 7:5 ECTS, and takes place in block 4. There were 33 students enrolled in the course, around 16–18 students were actively participating in lectures and labs, and 22 students submitted the complete portfolio at the end of the course.

Due to COVID-19 lockdown, the course was held completely online with live lectures and labs, that were recorded and made available for students, who could rewatch them if needed. This did not significantly affect Intended Learning Outcomes (ILO)s, but it affected the learning process since socialisation among students was missing: it was impossible for students to gather together and discuss their projects or the course content.

The course assessment was performed through a portfolio exam. The portfolio included: an individual assignment, which requires practical work, i.e. to implement and test several IR algorithms; the participation in an online student competition hosted on CodaLab²; and a written report, where the students have to reflect on the task and to analyse the results in relation to what they learnt during the theoretical lectures. The students had to include both the written report and the code in the submitted portfolio, as well as to describe their results in the online competition. This part contributed roughly 90% of the final grade.

In addition to the source code and report, the portfolio included also two peer reviews and a response letter. Around the middle of the block, a preliminary version of the report was submitted for peer review. Each student received two reports from their colleagues and had to provide feedback. The whole review process was double-blind, to allow the students to freely express their opinion. One lab lecture was devoted to explain how to structure peer reviews to provide useful insights and feedback. After receiving the peer reviews, the students had to prepare a response letter: they had to explain how they addressed each comment or motivate why they decided to ignore some comments. Peer review contributed 10% of the final grade.

A total of 3 examiners (2 teachers and 1 teaching assistant) graded the final portfolios. Students portfolios were graded with 2 phases: first each portfolio was independently graded by two examiners, then the examiners compared their grades and decided the final grade.

² https://competitions.codalab.org/competitions/24407?secret_key=60713ae9-26ac-4dc2-88cf-7e0e482bc657

Motivations and Goal of the Activity

The portfolio exam allows to assess a diverse set of learning outcomes in a more integrated way (Klenowski et al., 2006; O'Sullivan et al., 2012). In the case of the IR course, through the portfolio teachers can assess whether the students achieve skills such as the ability to transfer theoretical knowledge to real-world tasks, to identify problems, and to design and calibrate appropriate solutions. Furthermore, some of the required competencies can be tested as well: the usage of standard procedures and practices when designing or implementing specific solutions; and the ability to present evaluation analyses and results in a written report, such that a technically qualified person can follow and obtain similar findings.

One of the main drawbacks of the portfolio exam, is that it requires a considerable amount of work and time during the marking phase. Indeed, even if the scope of the portfolio is limited to specific tasks, each portfolio exam is the individual expression of the student learning process, and requires subjective judgements to be graded. Therefore, it can be challenging for teachers to assess them in a fair and reliable way.

One solution for this problem is to develop an assessment rubric (Andrade, 2000; Reddy & Andrade, 2010), this will guide teachers when assigning the final grade. Furthermore, if multiple teachers are involved in grading, as in the case of the IR course, this will help them to be fairer, better align their grades and check for high inter-rater reliability issues.

On the student side, a portfolio exam can be very demanding. In particular, for the IR course, students are required to complete many tasks and to meet different deadlines. Indeed a preliminary version of each assignment needs to be submitted before the final deadline to allow the peer review process to start.

Furthermore, a well designed portfolio exam should give some freedom to students, so that they can express their creativity. On one hand, this can help to motivate and engage students with the content, but on the other hand students may struggle in understanding what are the actual requirements of the portfolio. For example, in the 2018/2019 IR course, the majority of the submitted assignments lacked of critical reflections and insights, showing that the students missed one of the portfolio goals, that is to interiorise the acquired knowledge and the sense of the assignment (Ross, 2009). Thus, there was an evident mismatch between teachers and students expectations, and this resulted in many mediocre grades, even though students put a great effort in completing the portfolio.

As suggested by Baume, 2003, an assessment rubric can be useful to address this problem, since students will know in advance how their work is going to be evaluated and they can design their portfolio consequently. Furthermore, students' ideas can be taken into consideration to define the rubric, for example by organising a group activity where students can propose their own assessment criteria. This will make students partially responsible for the grading phase, and will help them to better understand assessment requirements.

When it comes to peer feedback, it can be a positive experience for students. Indeed, Cole, 1991 claims that students behave differently when they receive a comment from a teacher and a peer. Teacher comments are perceived as authoritative and reliable, but also complex and hard to understand. On the other side, peer comments are seen as an opportunity to discuss and negotiate, especially when provided orally, and can be more detailed and easier to understand. However, as reported in Weaver, 1995, some students may not understand the purpose and benefits of peer feedback. This also happened during the IR course in 2018/2019: peer feedback was not perceived as a useful activity and the final quality of the submitted assignments was generally poor, showing that peer feedback did not help the students in identifying potential flaws in their assignment.

If an assessment rubric is developed in collaboration with the students, peer feedback allows to involve students in the marking process. Furthermore, the rubric helps the students in better understanding the assessment process performed by the teacher, and to gain insights related to the complexities and ambiguities when evaluating students portfolios. As stated by (Nulty, 2008), students will "feel ownership of the assessment (and learning) process rather than being alienated or victimised by it".

With these observations in mind, the main goal of this pedagogical project was to prepare an assessment rubric for the IR course. The rubric aimed at guiding the students to meet the expectations of their teachers both for the peer-feedback process and the written assignment.

Before its finalisation, the project was discussed with the course responsible in 2018/2019 and one of the co-teachers of the course, who was teaching the IR course both in 2018/2019 and 2019/2020. Both of them were enthusiastic about the project and thought that an assessment rubric could be very useful for the students. Furthermore, as source of inspiration and starting point, the co-teacher shared an assessment rubric that he developed for another course at the Computer Science Department.

Description of the Activity

To organise the peer feedback process and develop the assessment rubric, we followed some of the suggestions presented in Topping, 2009:

1. Clarify with the students the purpose and the learning outcomes;
2. Involve the students in the decision of assessment criteria;
3. Provide training, examples and practice;
4. Provide guidelines and checklists;
5. Examine the quality of peer feedback.

The first lab was designed to put into practice some of the aforementioned suggestions from Topping, 2009. Specifically, the lab was divided in 3 parts: firstly we presented the portfolio; secondly we instructed students how to conduct peer feedback; finally we organised a group activity to involve students in the development of the assessment rubric.

The first part of the lab was devoted to carefully explain the requirements and tasks of the portfolio exam. In particular, the assignment description is very long and detailed (7 pages excluding references), thus it is fundamental to give the students an overview of the most important requirements. The learning outcomes of the student assignment were clearly stated at the beginning of the assignment description and they were also orally presented during the lab.

After explaining the practicalities related to the assignment, we presented the peer feedback process. From last year, we noticed that the majority of the students were not familiar with peer feedback and they were struggling in providing useful feedback to their peers. Therefore, we prepared a set of slides that the students could follow to structure their feedback.

Finally, the first lab lecture ended with a group activity. The purpose of the group activity was to involve the students in the creation of the assessment rubric. We introduced the group activity with a couple of slides to explain the students what an assessment rubric is and what they needed to do during the group activity.

Around 10 students participated in the group activity, they were divided in 2 groups with 5 students each. The group activity was carried out online by means of Zoom breakout rooms. The students were provided with a link to a shared Google spreadsheet, where they could jointly work on the rubric.

The students had 30 minutes to work on the rubric and then there was a 10 minutes classroom discussion to analyse differences and similarities between the two rubrics.

After this lab lecture, we analysed the evaluation criteria proposed by students. Due to the short time assigned to the group activity, the assessment rubrics developed by the students were not complete. The students did not have enough time to describe each evaluation criteria and their rubric were mainly checklists. As mentioned before, the assignment description is very detailed and includes different sections with different tasks and goals, thus identifying overall evaluation criteria and describing them is a complex task. Therefore, instead of developing an assessment rubric, we decided to create a checklist. The checklist has 2 main advantages: first it guides the students to structure their report, indeed each section of the checklist corresponds roughly to a section in the report; second it is easy to use by students, who can simply check if they fulfilled each bullet point.

At the end of the course, when the students submitted their portfolio, we examined both the report and the peer reviews, since both of them were compulsory parts of the portfolio and contributed to the final grade. The outcomes are summarised in the following section.

Outcome of the Activity

From the examination of the portfolios submitted by students, it is clear that many students actually exploited the assessment checklist that we provided. Indeed, the majority of the students followed the structure of the checklist in their report. Furthermore, 5 out of 22 students followed exactly the checklist when preparing their reviews. By examining the reviews given by students, the majority of them were clearly written, well structured and included many useful suggestions to help peer students in improving their report.

In Figure 18.1 we compare the distributions of grades from the IR course in 2018/2019 and this year (2019/2020). When comparing the two distributions of grades we need to keep in mind that: (1) those distributions refer to different students and different assignments (even if the assessment of the IR course in 2018/2019 was also through a portfolio, the actual assignment was different); (2) different examiners graded the portfolios this year and last year (just 2 examiners graded the portfolios both in 2019 and 2020); (3) due to the switch to full online teaching, the students this year

were more challenged than last year. Nevertheless, from the distribution of grades in Figure 18.1, we can see that the grades in 2019/2020 tend to be overall higher than the previous year, and the distribution is shifted towards the top of the scale (7, 10 and 12 in the danish grading scale). Even if we can not conclude that the cause of this improvement in students grades is completely due to the assessment checklist that we provided, the checklist might have positively contributed in improving the quality of the submitted portfolios.

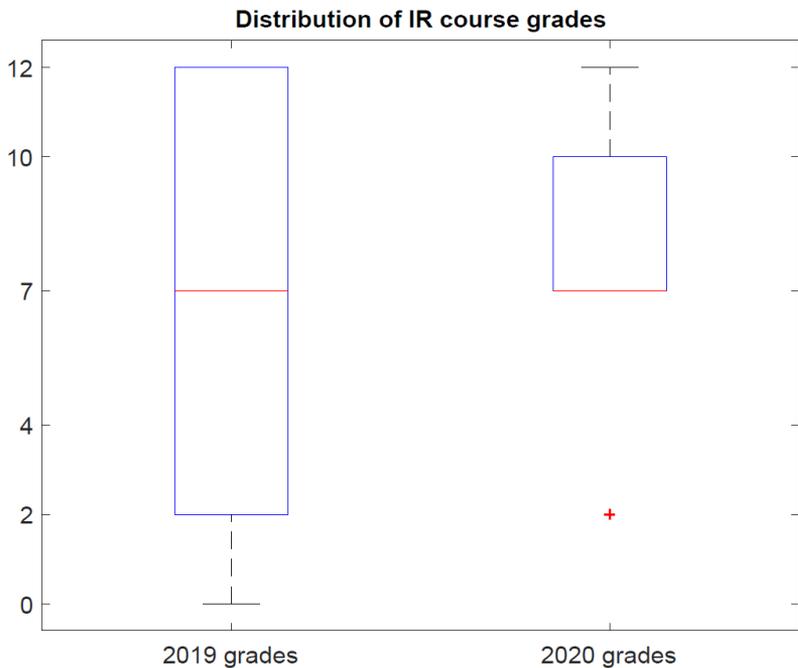


Fig. 18.1. Boxplots with the distribution of grades for the IR course in 2018/2019 and 2019/2020.

Moreover, the assessment checklist was given to the 2 co-examiners of the IR course, who could use it as a guide to grade the portfolios. They both said that they used the checklist for marking and they found it helpful.

Furthermore, they were impressed by the high quality of the portfolios submitted by students. They commented that it was remarkable for students to prepare such good portfolios, considering all the challenges and difficulties of online teaching.

Finally, the original plan of this project included some interviews with students after the end of the course to ask their opinion about the assessment checklist. Unfortunately, due to online teaching and the closure of university buildings it was hard to arrange those interviews, thus no students interviews were included in the evaluation of this project.

Conclusions and Future Work

The current project was carried out during the IR course in the academic year 2019/2020. The aim of the project was to develop a checklist that could guide the students to organise their assignment and structure the peer feedback process. The checklist could also help students in self-assess their own assignment and it was provided to all the course examiners to grade students portfolios. Overall, the examiners found the checklist useful for grading and the quality of students portfolios improved with respect to the previous year. Therefore, the checklist might have helped students in better understanding teachers expectations.

There are still some challenges that can be addressed in relation to the IR course for the following years. From this year course evaluation, some comments were: “I felt like the peer review did not help much”, “No peer-review” and “I would maybe prefer some feedback from the TA’s besides the student reviews as I believe they may be more suited for guiding me in the right direction”. Therefore, the students are still struggling in understanding the value of the peer review process, even if they were instructed how to conduct peer reviews, they were provided with the assessment checklist and the quality of the submitted reviews was quite high.

This might be due to many concurrent reasons. First, it is hard to find an optimal time to set the peer review deadline. If the deadline is too early in the block, the students do not have enough time and material to fill the assignment. If the deadline is too late in the block, then the students do not have enough time to adjust their assignment in relation to the feedback they received. In this edition of the course, the peer review deadline was set around the middle of the course and the students were able to complete half of the assignment at that time.

Second, some students submitted a very poor assignment for the peer review process. Since the beginning of the course, we advised students to write as much as possible in the assignment for the peer review submission. We explained students that the more they could put in the preliminary submission of their assignment, the more they could get from the peer review process. We also specified that a blank submission was considered as a missing submission, thus a missing part in their final portfolio. Nevertheless, some students submitted just a bullet point list with a short description of what they planned to do for the assignment. This sort of submissions, are not useful neither for students who authored them, since they can not get any useful feedback, nor for students who received them, since they can not compare their own work with peers work.

Next year, we will stress even more the importance of peer feedback and the benefits that students can get from it. Furthermore, we will add extra questions on the course evaluation about peer feedback, to better understand the opinion of students in relation to peer feedback. We will also consider to set a mid-term deadline for the first half of the assignment, so students will be forced to work on the assignment before the preliminary submission and peer review will be done on a partial assignment. Finally, in addition to the checklist, we will provide an assessment rubric, developed upon the current checklist, that students can use for peer feedback and for self-assess their own assignment.

References

- Andrade, H. G. (2000). Using rubrics to promote thinking and learning. *Educational leadership*, 57(5), 13–19.
- Baume, D. (2003). Supporting portfolio development continuing professional development series. *York, Learning and Teaching Support Network Generic Center*.
- Cole, D. A. (1991). Change in self-perceived competence as a function of peer and teacher evaluation. *Developmental Psychology*, 27(4), 682.
- Klenowski, V., Askew, S., & Carnell, E. (2006). Portfolios for learning, assessment and professional development in higher education. *Assessment & Evaluation in Higher Education*, 31(3), 267–286.

- Nulty, D. (2008). A guide to peer and self assessment: Approaches and practice strategies for academics. *Griffith Institute for Higher Education. Griffith University.*
- O'Sullivan, A. J., Harris, P., Hughes, C. S., Toohey, S. M., Balasooriya, C., Velan, G., Kumar, R. K., & McNeil, H. P. (2012). Linking assessment to undergraduate student capabilities through portfolio examination. *Assessment & Evaluation in Higher Education, 37*(3), 379–391.
- Reddy, Y. M., & Andrade, H. (2010). A review of rubric use in higher education. *Assessment & evaluation in higher education, 35*(4), 435–448.
- Ross, P. (2009). Using a portfolio to assess the key learning outcomes of practical classes. In *Bio-assess collection: Enhancing assessment in the biological sciences: Ideas and resources for university educators.*
- Topping, K. J. (2009). Peer assessment. *Theory into practice, 48*(1), 20–27.
- Weaver, M. E. (1995). Using peer response in the classroom: Students' perspectives. *Research and Teaching in Developmental Education, 31*–37.