Evaluation of a New Assessment Scheme for a Third-year Concurrency Course

Paul Strooper  Larissa Meinicke
School of ITEE, The University of Queensland, Brisbane, Qld 4072, Australia, pstroop@itee.uq.edu.au larissa@itee.uq.edu.au

Abstract

In this paper, we describe and evaluate a change that was made to the assessment scheme for a third-year course on concurrent systems. The original assessment scheme consisted of tutorials, three assignments, and a final examination. The change was motivated by the fact that students often performed poorly, especially on the third assignment, and yet the average mark that they received for this and the other assignments contributed to what we felt were inflated course grades. In the changed scheme, the third assignment was made more complex and open-ended, but it was also made optional in that students can pass the course without having to do this third assignment. We evaluated the new assessment scheme by comparing the performance of students on Assignment 3 and the course for the last two years and by analysing the results of a student questionnaire.

Keywords: student assessment, concurrency.

1 Introduction

Lister and Leaney (2003) present and discuss an assessment scheme based on Bloom’s taxonomy of educational objectives (Bloom et al. 1956) for a first-year programming course. The assessment scheme is motivated by the fact that in many first-year programming courses, the weaker students flounder and the stronger students are not challenged. To maximise the potential of students in such a course, Lister and Leaney argue that different assessment tasks must be used for students of different abilities and they use Bloom’s taxonomy to motivate their choices. To pass the course, the students had to display knowledge and comprehension through lab exercises, a lab exam and a multiple choice exam. For a credit or distinction, students had to display application ability as well, through the completion of traditional, fairly well-defined programming assignments. Finally, for a high distinction, the students had to display the ability to synthesise and evaluate through an open-ended individual project and peer review of other students’ code. Students were required to complete only the assessment tasks for the level of achievement they desired or were capable of, thus reducing the assessment load and leaving only a relatively small number of motivated students to attempt the individual project.

Box (2004) documents her experience in applying the assessment scheme used by Lister and Leaney to a second-year object-oriented programming course. Her motivation was that her class consisted of students at different skill and motivational levels, that some students felt they were not being challenged by the course material, and that many students were achieving inflated grades because of grace marks that they earned on assignments. Box does not provide a detailed evaluation of the course outcomes.

The first author has been involved in teaching a third-year course on concurrent systems since 1993. Up until 2003, this course was assessed in a fairly traditional way, incorporating an optional tutorial participation mark, three fairly well-defined assignments, and a final exam. However, this assessment scheme was unsatisfactory for a number of reasons:

- The students performed poorly on some of the assignments. This was especially true for the third assignment, which tended to be slightly more complex than the first two assignments and which was due at a busy and stressful time for students. The poor student performance on the third assignment also made marking this assignment a frustrating task for the teaching staff.
- Despite the poor performance on assignments, we felt that the marks for the course were inflated due to an overemphasis on relatively easy marks that could be collected by students on tutorial participation and assignments (even if students failed the third assignment, it was relatively straightforward to pick up some marks that would help them with their final grade).
- On the course and teaching evaluations, a number of (presumably stronger) students complained that the course was not challenging enough.

Motivated by the work of Lister and Leaney, a number of changes were made to the assessment scheme in 2004. The third assignment was replaced by a much more open-ended assignment, which was made optional. To do well in the course, the students had to complete all three assignments and perform well on the exam, and the tutorial participation mark was no longer used in the calculation of the final grade. For the lower grades, tutorial participation, the first two assignments, and the exam were taken into consideration, but the third assignment was not.

Note that students had to decide whether to work on the third assignment during the semester, without knowing exactly how well they would end up doing in the course. However, we felt that it would not be a disadvantage to do the third assignment in terms of gaining a better understanding of the course material, even if the Assignment 3 mark did not directly count in the calculation of the final grade.

The anticipated advantages of the changes were:

- A more balanced and representative distribution of grades.
• A less stressful (but just as relevant) assessment scheme for weaker students.

• More challenging and interesting assessment tasks for stronger students.

• For teaching staff, a reduced marking load and more enjoyable marking experience for the third assignment. In addition, the more open-ended nature of the third assignment would make it much more difficult for students to plagiarise on this assignment.

In Section 2, we present a brief overview of the course and the assessment scheme that was used in the past. In Section 3, the new assessment scheme is presented, including a description of the third assignment. The evaluation of the new assessment scheme is discussed in Section 4.

2 Old assessment scheme

2.1 Course overview

COMP3402 Concurrent and Real-Time Systems is an introductory course on concurrent systems. Despite its title, only a small portion of the course is devoted to real-time systems. The course is typically taken by third- or fourth-year Information Technology or Software Engineering students, and can also be taken by Masters students (but typically the number of Masters students enrolled is small).

Since 2001, we have used the textbook Concurrency: State Models and Java Programs by Magee and Kramer (1999) in the course. The book takes a model-based approach to the design and implementation of concurrent programs, with models specified in FSP (a variant of CSP) that can be analysed using the LTSA tool and then implemented in Java.

2.2 Assessment

The contact hours consist of three lectures per week, one group tutorial, and one open tutorial. In addition, there are three assignments and a two-hour final exam. In the group tutorials, which form part of the assessment, students have to hand in an individual solution to a “warm-up” exercise (typically these are very simple problems, which the students should be able to solve in 5–10 minutes) at the start of each tutorial, and then a group solution to one of the problems on the tutorial handout. There is no tutorial handout for the open tutorials, which can be used to get assistance with any issues related to the course.

Assignment 1 involves modelling a fairly well-defined problem (in 2004, a taxi dispatcher and a number of taxis and customers) in FSP and then analysing it using LTSA. Assignment 2 is a programming assignment, in which the students have to implement a bounded buffer using semaphores (the semaphore implementation is given to them) and then to use these bounded buffers to implement a datalow network. Assignment 3 is also a programming assignment, in which the students typically have to implement the system that was modelled in Assignment 1. Up until 2003, the design of the implementation and the detailed input and output was specified in the assignment handout, and parts of the solution were provided to the students so that they could focus on the “interesting” aspects.

In 2003, we presented a new assessment scheme. The main change to the assessment tasks in 2004 was that Assignment 3 was made more open-ended and a small component of its assessment was associated with the students evaluating their own design and implementation.

2.3 Determining grades

To calculate the student grades up until 2003, the three assignments contributed 10% each, the tutorial participation mark contributed 10%, and the final exam mark contributed 60%. If students performed better on the final exam than on tutorial participation (e.g., if they did not attend any of the group tutorials), then the final exam contributed 70% and the tutorial participation mark was not counted. The resulting combined mark was then converted to a grade as shown in Table 1.

<table>
<thead>
<tr>
<th>Combined Mark</th>
<th>Grade</th>
</tr>
</thead>
<tbody>
<tr>
<td>85–100</td>
<td>7 (high distinction)</td>
</tr>
<tr>
<td>75–84</td>
<td>6 (distinction)</td>
</tr>
<tr>
<td>65–74</td>
<td>5 (credit)</td>
</tr>
<tr>
<td>50–64</td>
<td>4 (pass)</td>
</tr>
<tr>
<td>45–49</td>
<td>3 (conceded pass)</td>
</tr>
<tr>
<td>20–44</td>
<td>2 (fail)</td>
</tr>
<tr>
<td>0–19</td>
<td>1 (fail)</td>
</tr>
</tbody>
</table>

Table 1: Conversion from marks to grades

2.4 Bloom’s taxonomy

In terms of Bloom’s taxonomy, the tutorial and the final (open-book) exam questions emphasise comprehension and application, with many questions asking the students about properties of small concurrent models or programs, and more complicated questions asking them to modify existing or write new (small) concurrent models and programs.

The assignments are relatively well-defined and address the students’ application and analysis abilities on medium-sized problems. Schneide and Gladkich (2006) have indicated that it can be difficult to consistently use Bloom’s taxonomy to determine levels in programming assignments. While it could be argued that the assignments contain some tasks at the evaluation and synthesis levels, it was felt that the tasks on the assignments are not challenging or creative enough for third-year students to be classified at higher levels.

As indicated above, there were a number of perceived problems with this assessment scheme, which motivated the changes introduced in 2004 and discussed below.

3 New assessment scheme

The main change to the assessment tasks in 2004 was that Assignment 3 was made more open-ended and a small component of its assessment was associated with the students evaluating their own design and implementation.

3.1 New Assignment 3

The following description is taken from the assignment handout (recall that Assignment 1 in 2004 involved the modelling of a taxi dispatcher and a number of taxi and customers): For this assignment, you must design and implement a system that meets at least the requirements from Part 3 of Assignment 1, but extend them where appropriate (for example, your dispatcher should be able to queue many calls from customers when no taxis are available). One extension that you must make is to add time to the simulation so that the output shows at what point in time events (such as taxis going on/off duty, calls from
customers to the dispatcher, etc.) occur in the simulation.

To encourage the students to start thinking about the assignment and to provide feedback on their initial thoughts, the assignment was split in two parts. In Part 1, the students had to describe the proposed user interface and the design. For the user interface, they had to describe the exact format of the input to the program and the output produced. They were encouraged to specify a GUI, but were not forced to do so. The design had to be described in terms of the classes and methods they planned to implement, as well as the concurrent aspects of the implementation.

Part 1 of Assignment 3 was actually due before Assignment 2 was due, and was marked on a pass/fail basis. If the student failed Part 1, then they had one week to resubmit it based on the feedback we had given them (and a general message that we posted to indicate common errors and problems in Part 1).

Part 2 of Assignment 3 consisted of the implementation of the system, an updated description of the user interface and design, a list of changes, and an evaluation of their own solution (e.g. describing any known flaws or deficiencies, what worked well, what they would do differently, etc.). The final mark for Assignment 3 was determined by their mark on Part 2, which was assessed based on the quality of the presentation (10%), the user interface (10%), the design (10%), the evaluation of their own solution (15%), the quality of the code (10%), a demonstration of their program (30%), and our overall impression of their “product” (15%). More detailed descriptors for each of these criteria were handed out to the students.

Part 2 was due in the penultimate week of the semester and all assignments were marked in demo sessions during the last week of semester. In the demo sessions, which typically lasted about 30 minutes, students were given the opportunity to briefly demonstrate their programs and then the markers asked questions about various aspects of their system (e.g. running through certain scenarios, questions about aspects of the code, etc.). To ensure consistency, all demo sessions were assessed with two markers present. Although all the assignments were assessed by the end of the week, it did take a lot of effort (certainly more effort than marking Assignment 3 in previous years).

### 3.2 New grading scheme

The main changes to the grading scheme in 2004 were as follows:

- Assignment 3 was made optional, but students had to perform well on it to get a grade of 6 or 7 on the course. Tutorial participation was not taken into account to determine if students should get a grade of 6 or 7.

- For grades of 5 or lower, Assignment 3 was not used in the calculation of the combined mark for the course, but tutorial participation could be counted for up to 10%.

- Cut-offs were introduced for the final exam mark (for most grades) and for Assignment 3 (for grades of 6 and 7).

In particular, there were two ways to calculate a combined mark for the course:

\[
M_1 = 0.15A_1 + 0.15A_2 + 0.6Ex + 0.1max(Ex, T)
\]

\[
M_2 = 0.1A_1 + 0.1A_2 + 0.2A_3 + 0.6Ex
\]

where \(A_1\)–\(A_3\) represent the marks on Assignments 1–3, \(Ex\) represents the exam mark, and \(T\) represents the tutorial participation mark. Then the final grade was determined by the highest grade in Table 2 for which the component results satisfy all the criteria.

### 3.3 Bloom’s taxonomy

In terms of Bloom’s taxonomy, the change was intended to address the the synthesis and evaluation aspects of Bloom’s taxonomy through the modified Assignment 3, while maintaining the other assessment tasks (tutorials, first two assignments, and the final exam) to address the lower levels. To address the synthesis and evaluation aspects, Assignment 3 was left much more open-ended than the other assignments and included a small assessment component in which the students had to evaluate their own design and implementation.

Our new approach to assessment differs from the one adopted by Lister and Leaney (2003) and Box (2004), since we require all students, even those at a pass level, to be able to demonstrate application and analysis level skills. This difference is motivated by the fact that our assessment scheme, unlike that of Lister and Leaney (2003) and Box (2004), is for a third-year, not a first- or second-year, course. Students who are at a more advanced stage of their program should be capable of demonstrating these more sophisticated characteristics.

### 4 Evaluation

In this section, we present the results of comparing the students’ performance on assignment 3 and the course in 2003 and 2004. We also present the results of an in-class survey carried out in 2004. Finally, we present our own perceptions of the changes and the subsequent changes that we made in 2005.

The number of students enrolled in 2004 (52) was significantly lower than the number of students enrolled in 2003 (77). We believe that this is attributed to an overall drop in the number of IT and SE students and the fact that there was a time clash in 2004 with the lectures from a big third-year IT course that was discovered too late. In other words, we do not think that the new assessment scheme had a detrimental affect on student numbers and we did not see a higher than normal attrition rate in the course either. Unless otherwise noted, the frequencies shown in the bar graphs in this section are expressed as percentages, to facilitate comparison between 2003 and 2004.

#### 4.1 Student performance on Assignment 3

Figure 1 shows the distribution of Assignment 3 marks in 2003 and 2004.

In 2004, 41 out of 52 (79%) students enrolled submitted Part 1 of Assignment 3. Only 16 of these passed and 25 failed. We felt that the two main reasons for this was that we had not given clear enough

<table>
<thead>
<tr>
<th>(M_1)</th>
<th>(M_2)</th>
<th>Ex</th>
<th>(A_3)</th>
<th>grade</th>
</tr>
</thead>
<tbody>
<tr>
<td>≥ 5</td>
<td>≥ 5</td>
<td>≥ 75</td>
<td>≥ 75</td>
<td>7 (high distinction)</td>
</tr>
<tr>
<td>≥ 5</td>
<td>≥ 50</td>
<td>≥ 50</td>
<td>-</td>
<td>6 (distinction)</td>
</tr>
<tr>
<td>≥ 60</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>5 (credit)</td>
</tr>
<tr>
<td>≥ 50</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>4 (pass)</td>
</tr>
<tr>
<td>≥ 40</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>3 (conceded pass)</td>
</tr>
<tr>
<td>≥ 20</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>2 (fail)</td>
</tr>
<tr>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1 (fail)</td>
</tr>
</tbody>
</table>

Table 2: Grading scheme for 2004
instructions on what we wanted and a number of students had not spent enough time thinking about their user interface and design. In addition to individual feedback to the students, we posted a general message about the common errors and what we expected in the resubmissions. The students were then given a week to resubmit, and 22 out of 25 students resubmitted. Of these, 15 passed the second time and 7 still failed (2 of these resubmitted a second time and finally passed). This meant that in total 33 students passed Part 1 of the assignment. Of those, 30 students submitted Part 2 (this represents 58% of the students enrolled in the course). Note that these are the only students who received a mark for Assignment 3 in 2004.

In 2003, 69 out of 77 (90%) students enrolled submitted Assignment 3. Out of the 8 who did not submit the assignment, 3 submitted no assignments, 4 submitted only 1 assignment, and 1 submitted the other 2 assignments, so it is fair to say that most students at least attempted Assignment 3.

As can be seen from this data, students performed fairly poorly on Assignment 3 in 2003, which was one of the motivations for changing the assessment scheme in the course. Of the 69 students who submitted, 18 failed the assignment and many others performed poorly. Despite the fact that the assessment scheme in 2004 was tougher, the students who submitted Part 2 of the assignment did better than in 2003. Only one student failed, three students received marks in the range of 20–21 out of 40, and the others received a mark of 25 out of 40 or better.

We also compared the performance of the students in both years for Assignments 1 and 2, and for the exam, to ensure that there were no significant differences between the two groups. With the exception of Assignment 1, on which the students performed slightly better in 2003 (which we believe to be due to the requirements for that assignment being more straightforward), the distributions for the marks were similar in both years.

4.2 Student performance on the course

We present and discuss the distribution of grades in 2003 and 2004. In order to understand the differences in performance we investigate possible differences in groups by comparing their results on assignment 1, assignment 2 and the exam, and we examine how changing the marking scheme in various ways may have affected the resulting grades.

4.2.1 Distribution of grades

The distribution of student grades for the course in 2003 and 2004 are shown in Figure 2.

The 2003 data shows a relatively large number of students with a grade of 6 (and also with a 7), compared to the students with a grade of 4 or 5. As indicated earlier, we felt that these grades were inflated and at least partially caused by the students being able to collect relatively “easy” marks during the semester. The grade distribution for 2004 is more what we would expect to see from an average class.

As can be seen from this data, students performed fairly poorly on Assignment 3 in 2003, which was one of the motivations for changing the assessment scheme in the course. Of the 69 students who submitted, 18 failed the assignment and many others performed poorly. Despite the fact that the assessment scheme in 2004 was tougher, the students who submitted Part 2 of the assignment did better than in 2003. Only one student failed, three students received marks in the range of 20–21 out of 40, and the others received a mark of 25 out of 40 or better.

We also compared the performance of the students in both years for Assignments 1 and 2, and for the exam, to ensure that there were no significant differences between the two groups. With the exception of Assignment 1, on which the students performed slightly better in 2003 (which we believe to be due to the requirements for that assignment being more straightforward), the distributions for the marks were similar in both years.

4.2.2 Factors affecting grades

Since we made multiple changes to the assessment scheme simultaneously, we wanted to ensure that the change in grade distribution was not caused solely by the introduction of exam or assignment cut-offs or the fact that we no longer counted tutorial participation for a grade of 6 or 7.

We first checked what the grades would have been in 2004 if there had been no exam cut-offs. This would not have changed the grade distribution very much: there would be 1 less grade of 2, 2 less grades of 4, and 3 more grades of 5 (the same number of 1s, 3s, 6s and 7s). This is not surprising, as the exam cut-offs are really not that stringent. In fact, the exam cut-off of 75% for a grade of 7 is not really a cut-off at all, because it is impossible to get a combined mark of 85% or higher on the course if you do not get at least 75% on the final exam. An explicit cut-off was still included in the assessment scheme for simplicity and to indicate the importance that we attributed to the final exam.

We also checked the cut-offs for Assignment 3, but these did not make any difference to the grades for any of the students in the course.

Finally, we wanted to check what the significance of the removal of the tutorial participation marks for grades of 6 and 7 was. However, it was not clear how we could use the 2004 data for this, since it was not clear how the tutorial marks should be incorporated into the calculation for the final grade (drop the exam mark to 50%, reduce the weighting for one of the assignments, or some other way). Instead, we used the
data from 2003 and checked to see what the distribution of grades would have looked like if we would not have allowed tutorial participation to count for grades of 6 or 7 (thus to get a grade of 6 or 7, the exam was counted for 70% of the final mark). Although this change reduced the percentage of 6s and 7s from what it was previously, there still is a significant number of 6s in the course (close to 25%), so the change related to the tutorial participation mark does not seem to account for the difference in the grade distribution between 2003 and 2004.

4.2.3 Factors affecting Assignment 3 choice
We had a closer look at the students who did not submit Assignment 3. One of the concerns with the new assessment scheme is that some “good” students may simply be too lazy to submit Assignment 3 and underperform as a result. We checked the exam marks of all the students who did not submit Assignment 3, and there was only one who did very well on the exam with a mark of 54 out of 60. The next highest mark for this group of students was 36 out of 60. The student who got a mark of 54 out of 60 also did well on the first two assignments and clearly could have received a 6 or a 7 on the course if he had submitted the third assignment. However, his combined mark for the course including Assignment 3 (where we counted his mark for Assignment 3 as 0) was only 74%, which is only a grade of 5. In other words, the marking scheme did not penalize him, but clearly it may have discouraged him from attempting the third assignment.

We also checked that for those students who did Assignment 3, but still only received a grade of 4 or 5 on the course, if it would have made a difference if we had allowed them to use the second marking scheme ($M_2$ in Section 3) for grades lower than 6. There were two students whose combined mark $M_2$ was higher than $M_1$, but for both these students $M_2 = 61$ and $M_1 = 60$, so it did not make a difference in the grade that they would have received.

A number of students commented to us after their Assignment 3 demonstrations that they really enjoyed the experience of working on a more open-ended assignment. It allowed them to “get into trouble” and then they had to find their own way out. A number of students also commented that they felt that they had a much deeper understanding of the problems associated with concurrency. Most importantly, however, some of these students did not think they would get a 6 or 7 in the course, but still appreciated the experience of working on the third assignment.

4.3 Student survey
In the last week of semester, which was before most of the Assignment 3 demos, we carried out a survey to obtain feedback on the new assessment scheme. Out of the 52 students in the course, 28 completed the survey.

The first part of the questionnaire asked if the students completed Part 1 and/or Part 2 of Assignment 3. If not, they were asked why. Out of the 28 students who completed the survey, 24 had completed Part 1 and 21 had completed Part 2. The main reasons for not handing in the assignment were lack of time and failing Part 1 (or a combination of both).

Since students have to decide whether or not to work on Assignment 3 before they knew how well they would do in the course, we were interested to find out how accurately students could predict their final grade (which could influence their decision whether or not to do the assignment). So we asked students about the grade that they expected to get in the course and the grade they were hoping to get. This data is shown in Figure 3 (note that in this case, the vertical axis represents the number of students, and not the percentage of students).
We also asked the students if they were willing to identify themselves on the questionnaire, so that we could compare their predicted with their actual grades (22 out of 28 students allowed us to do this). When we compare the expected with the actual grades, one student overestimated their final grade by 3 (expected a 5 and received a 2), another one by 2 (expected a 7 and received a 5), and all others were within 1 grade (8 overestimated by 1 grade, 10 accurately predicted their grade, and 2 underestimated by 1 grade). This data suggests that most students were able to predict fairly accurately how well they will do in this course before knowing how well they did on Assignment 3 or the exam.

The last two sections of the questionnaire asked the student to rate the assessment schemes in 2003 and 2004 (the 2003 assessment scheme was explained on the questionnaire). For both schemes, we asked them to rate their level of agreement for the following statements:

1. I think the assessment scheme is fair, and
2. I like the assessment scheme.

For both questions, we used a five-point scale: strongly agree (SA), agree (A), neutral (N), disagree (D), or strongly disagree (SD). In addition, they were asked to provide comments on either of the assessment schemes, or any other comments related to assessment.

Note that while it is normally not reasonable to ask students to evaluate an assessment scheme to which they have not been exposed, we felt that this was not an issue in this case because the old assessment scheme is a standard assessment scheme that is used in a number of our courses.

The data for the ratings of the above statements is shown in Table 3. The students rated the old scheme higher both in terms of thinking it is fair and liking it. The difference in the ratings is somewhat bigger for the issue of liking the assessment schemes than it is for rating the fairness. However, even for the new scheme, 18 out of 28 (64%) students agreed or strongly agreed with the statement that they thought the scheme was fair, and 15 out of 28 students (53%) agreed or strongly agreed that they liked it. A more detailed analysis of the data reveals that 7 students thought the new scheme was fairer than the old, 14 thought the old scheme was fairer, and 7 rated them the same. Similarly, 6 students liked the new scheme better, 15 liked the old scheme better, and 7 rated them the same.

The following is a summary of the most interesting general comments by the students. The students could write any comments they wanted, and several students included multiple comments. Below we paraphrase and group their comments, so that we can indicate the number of students who included similar comments in their feedback:

- 5 students liked the more open-ended and challenging nature of Assignment 3;
- 3 liked the flexibility offered by making Assignment 3 optional;
- 4 thought Assignment 3 should have been defined better;
- 3 thought Assignment 3 should also be counted for grades lower than 6;
- 3 did not like the exam cut-offs;
- 1 thought the exam was weighted too heavily;
- 3 thought it was (too) hard to achieve a 6 or a 7 with the new scheme;
- 4 thought tutorial participation should count for all grades; and
- 2 thought tutorial participation should never be counted.

4.4 Own evaluation

The main concerns we had with the changes were the large number of students who failed Part 1 of Assignment 3 and the fact that some of the user interfaces in the final submissions were hard to use and comprehend.

We were satisfied with most other aspects of the changes:

- many of the students who did the third assignment seemed to enjoy the experience,
- although it was even more time-consuming to mark the third assignment than in 2003, it was more enjoyable because of the interactive nature of the demo sessions and the fact that the standard of solutions was better, and
- we felt that the distribution of final grades better reflected the students’ ability than in previous years.

4.5 Changes for 2005

We continued with the new assessment scheme in 2005 with the following changes.

1. We dropped the cut-offs for Assignment 3, because they did not seem to make a big difference in the final grades.

2. We also allow students to count M2 (and thus Assignment 3) for any grade, not just for grades of 6 and 7. As indicated above, this would not have made a difference in 2004 because only 2 students had a slightly higher mark M2 than M1, but it could make a difference in theory. Moreover, this addressed a concern raised by some of the students and we also felt that by making this change, it was less likely that the assessment scheme would discourage students from attempting Assignment 3 because it would not count for them anyway.

3. In response to a requirement from our University to distinguish the assessment requirements for undergraduate and masters courses, we always used marking scheme M2 to calculate the final grade for masters students. We thus always counted Assignment 3 for masters students, and we did not consider tutorial participation in their final grade.

4. For Assignment 3 itself, we provided more detailed instructions for Part 1 the assignment, including a template document to clarify the type of information that we were expecting. We also increased the portion of the assessment attributed to the user interface from 10% to 15%.

<table>
<thead>
<tr>
<th>I think this year’s scheme is fair</th>
<th>SD</th>
<th>D</th>
<th>N</th>
<th>A</th>
<th>SA</th>
</tr>
</thead>
<tbody>
<tr>
<td>I like this year’s scheme</td>
<td>0</td>
<td>2</td>
<td>8</td>
<td>14</td>
<td>4</td>
</tr>
<tr>
<td>I think last year’s scheme is fair</td>
<td>0</td>
<td>7</td>
<td>6</td>
<td>11</td>
<td>4</td>
</tr>
<tr>
<td>I like last year’s scheme</td>
<td>0</td>
<td>0</td>
<td>6</td>
<td>16</td>
<td>6</td>
</tr>
</tbody>
</table>

Table 3: Student rating of old and new assessment schemes
and correspondingly decreased the proportion attributed to the students evaluation of their own solution from 15% to 10%.

Despite the comments by some of the students, we kept the exam cut-offs because we felt they allowed us to ensure that students achieve at least a certain standard in the course (e.g., for a grade of 4 or 5, up to 40% of a student's combined mark $M_1$ consists of their marks on the first two assignments, on which students can “collaborate” extensively, and the tutorial participation mark). Similarly, we continued with not counting tutorial participation in the $M_2$ mark, because we felt that it is reasonable for grades of 6 or 7 to not include any “easy” marks.

### 4.6 Results from 2005

We did not perform a student survey or a detailed analysis in 2005, but we briefly discuss the results from the 2005 version of the course below. The distribution of marks for Assignment 3 and the distribution of final grades for the 2005 version of the course are shown in Figure 4.

In 2005, there were 41 students enrolled in the course, 33 undergraduate students and 8 masters students (compared to 52 in 2004, of whom 6 were masters students). Of those, 29 submitted Part 1 of Assignment 3, 26 of whom passed the first time and 2 passed after a resubmission, for a total of 28 students who passed Part 1. This compares favourably to the 2004 data, when only 16 out of 41 students passed Part 1 the first time and 33 out of 41 students passed eventually. Of the 28 students who passed Part 1, 23 submitted Part 2 of the assignment (56% of the students enrolled, compared to 58% in 2005). The average assignment mark in 2005 was 68.5% compared to 76.2% in 2004. Part of the reason for this was that 5 out of 23 students failed the assignment, including 2 out of 8 masters students (one of whom received a mark of only 8 out of 40). Part of this may be attributable to the fact that masters students were required to submit Assignment 3 as part of their assessment.

The overall performance of the students in 2005 was similar to that in 2004. The average mark on the final exam was similar (35.3 out of 60 in 2005 compared to 34.8 out of 60 in 2004), and the grade distribution for the course in 2005 was also similar.

We again checked if there were students in 2005 who we thought could have done much better if they had chosen to do Assignment 3. There were 3 students who scored at least 40 out of 60 on the final exam and who had not attempted Assignment 3. Of those, one student would have received a grade of 6 if they had just passed Assignment 3, the second student would have only received a grade of 6 if they had received full marks (40 out of 40) for Assignment 3, and the third student would not have received a 6 even with full marks for Assignment 3.

Based on these results, we are currently going ahead with the same assessment and grading scheme in the 2006 version of the course and a similar scheme was used in a third-year Compilers course at our university.

### 5 Concluding remarks

We have motivated, presented, and discussed the change in the assessment scheme for a course on concurrent systems. The main changes were making the third assignment much more open-ended and optional, using different ways of combining marks for different grades, the introduction of cut-off scores for the exam and third assignment, and not counting tutorial participation for grades of 6 or 7. In terms of the anticipated advantages discussed in the introduction, we feel that the new assessment scheme has resulted in a more balanced and representative distribution of grades, the feedback of both weaker and stronger students on the third assignment has been positive, and although the scheme has not resulted in a reduced marking load, the marking has become more enjoyable (and we are fairly certain that there was no problem with plagiarism on the third assignment). The student feedback on the overall assessment scheme was not quite as positive, but we did not feel that the criticism was sufficient to warrant a substantial change. As a result, we continued with the new assessment scheme in 2005 and 2006 with only a few minor changes.

### Acknowledgments

We would like to thank Margaret Wojcicki and the anonymous referees for their valuable feedback on earlier versions of this paper. We would also like to thank the students enrolled in the course for their feedback on the assessment schemes.
References


