A Quest for Helpful Feedback to Programming Coursework

Shun Ha Sylvia Wong (s.h.s.wong@aston.ac.uk), Anthony J. Beaumont (a.j.beaumont@aston.ac.uk)

Aston University, UK

Abstract: The 2011 National Student Survey (NSS) revealed that 40% of full-time students in England did not think that the feedback on their work has been helpful, even though 66% of these students agreed that the feedback was detailed and 62% of them agreed that the feedback has been prompt. Detailed feedback that is not considered helpful by students means a waste of tutors’ time while students continue to struggle with their learning.

What do students consider as helpful feedback? What are the qualities of helpful feedback? What are the preferred forms of feedback? How should tutors write feedback so that students will find it helpful? Can ICT help to improve the quality of feedback?

In our ongoing search for answers to the above questions, we have trialled the use of a novel Internet application, called eCAF, to assess programming coursework from Engineering, Mathematics and Computing students and have collected their views on the feedback received through a survey. The survey reveals that most students prefer electronic feedback as given through eCAF, with verbal feedback ranked second and hand-written feedback ranked even lower. The survey also indicates that the feedback from some tutors is considered more helpful than others. We report on the detailed findings of the survey. By comparing the kinds of feedback given by each tutor who took part in the trial, we explore ways to improve the helpfulness of feedback on programming coursework in a bid to promote learning amongst engineering students.

Introduction

The 2011 National Student Survey (NSS) revealed that 40% of full-time students from Higher Education Institutes (HEIs) in England did not think that the feedback on their work has been helpful, even though 66% of these students agreed that the feedback was detailed and 62% of them agreed that the feedback has been prompt (Times Higher Education, 2011). While full-time students from HEIs in Wales expressed a very similar view, those in Scotland were less satisfied. Nevertheless, the 2011 NSS data also revealed that full-time students from Further Education Colleges (FECs) in England were more satisfied with the feedback they received: only 30% did not find the feedback helpful and only 26% did not agree that the feedback on their was sufficiently detailed. This shows that assessment and feedback remains a great challenge faced by HEIs throughout the country.

The national issue of relatively low student satisfaction in assessment and feedback has led to various attempts to alleviate the problem. For example, the Department of Aeronautical and Automotive Engineering at Loughborough University has put in place a departmental strategy to ensure that the timeliness and the quality of feedback in all modules meet the minimal required standard (Horner, 2010). However, without addressing the issue of how to prepare personalised, individual feedback efficiently, such an approach can only bring about a limited level of improvement and the helpfulness of feedback is not guaranteed. To help tutors identify the kinds of feedback required by their students, Tudor & Perera (2010) have trialled an approach where Mechanical Engineering students at Northumbria University indicate on a checklist what form of feedback they required when taking a class test. This approach has led to a good level of student engagement in the feedback process, with a relatively high proportion of students reported to have used the feedback to improve their learning. Tudor & Perera's (2010) method, however, did not seem to unpack the meaning of detailed feedback
sufficiently. Many students reported to have chosen the item "detailed feedback on question number..." on the feedback request labels. What did they mean by "detailed feedback"? How to prepare the requested feedback in a timely fashion? How to write feedback so that students will find it helpful to their learning? Such issues cannot be addressed by a small checklist. Without answering such questions, the feedback provided by tutors may still not be helpful to students.

Detailed feedback that is not considered helpful by students means a waste of tutors' time while students continue to struggle with their learning. Some literature on how to prepare effective feedback exists. For example, Bloxham & Boyd (2007, chapter 7) suggested that tutors should "replace apparently confident statements with more questioning comments", should not "over-correct written work" and should "consider writing just three or four comments" only. Wong (2010), however, reported that students doing Computer Science modules prefer feedback which highlights all mistakes made and provides explanation on how to rectify each mistake. Burke & Pieterick (2010) summarised a wealth of studies on how to prepare effective formative feedback. Those studies focussed on writing feedback for essay-type work. Some of the ideas are not applicable for practical coursework such as programming.

What do students learning programming consider as helpful feedback? What are the qualities of helpful feedback? What are the preferred forms of feedback? How should tutors write feedback so that students will find it helpful? Can ICT help to improve the quality of feedback? In our ongoing search for answers to the above questions, we have trialled the use of a novel Internet application, called eCAF (Wong et al., 2008), to assess programming coursework from Engineering, Mathematics and Computing students and have collected their views on the feedback received through a survey. The survey revealed that the feedback from some tutors was considered more helpful by students than others. We have studied the characteristics of the feedback written by these tutors. This paper reports the results and findings of this study. In a bid to promote learning amongst engineering students, this paper explores ways to improve the helpfulness of feedback on programming coursework.

Methodology

A study to collect views on assessment feedback was conducted in the academic year 2010/1. The study was based on the assessment and feedback data collected from four computing science modules. Tutors of these modules set up their coursework as required, but they were asked to collect student submissions electronically via the institution's Virtual Learning Environment (VLE). Tutors were then asked to use the in-house electronic coursework assessment and feedback system eCAF to assess the submissions and to provide each submission with feedback. No further instructions were given to the tutors and they were expected to assess each submission in the usual way.

In summer 2011, students who had received feedback through eCAF were surveyed in order to gauge how satisfied they were with the feedback received from their tutors. The tool of the survey was a questionnaire. It was primarily designed to obtain opinions from students regarding the quality of coursework feedback received via eCAF. It also surveyed the students' general view on assessment and feedback. The questionnaire is divided into two sections:

• Section A investigates the students' satisfaction level with the feedback received and it includes questions on assessment and feedback from the NSS.
• Section B investigates the students' general view on what was considered as timely and helpful feedback. It was designed to discover the preferred way to receiving feedback on coursework based on the findings from a related study conducted in 2009 (Wong, 2010).

To help the data analysis, the questionnaire also includes questions about the respondents' profile, including information about the module taken, the tutor who assessed the coursework, their degree programme, etc.

The questionnaire contains 30 Likert-type, multiple-choice, yes/no and open-ended questions. The participants were asked to state who assessed their coursework in a multiple-choice question. Based on the feedback received from the stated tutor, the participants were asked to express their opinion on 12 five-point Likert-type statements regarding the timeliness, quality and helpfulness of the feedback received. The score ranges from 1 to 5, with 5 being the highest score.

Participants

The study was carried out in the School of Engineering and Applied Science at Aston University in 2010/1. The four modules were taken by different cohorts of undergraduate students and their degree
programmes were: Computing Science (CS), Computing for Business (CB), Communications Engineering (CE), Multimedia Computing (MC), Mathematics (Maths), Mathematics with Computing (MathC), Combined Honours (CH), Electronic Engineering and Computing Science (EECS), Internet Systems (IS). The assessment of coursework in two first year modules (ie CS1020, CS1310) was divided amongst various tutorial group tutors, with each tutorial group having a class size of 22-34 students. Details about all four modules are shown in Table 1.

Table 1: Modules involved in this study

<table>
<thead>
<tr>
<th>Module</th>
<th>Number of Credits</th>
<th>Teaching Period</th>
<th>Number of Programming Assignments</th>
<th>Degree Programmes</th>
<th>Year of Study</th>
<th>Class Size</th>
<th>Tutor ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>CS1020 Java Programming Foundations</td>
<td>20</td>
<td>1 &amp; 2</td>
<td>3</td>
<td>CB, CH, IS, MathC, Maths &amp; MC</td>
<td>1</td>
<td>171</td>
<td>D - I</td>
</tr>
<tr>
<td>CS1310 Java Programming Foundations</td>
<td>20</td>
<td>1</td>
<td>2</td>
<td>CE, CS &amp; EECS</td>
<td>1</td>
<td>90</td>
<td>A - C</td>
</tr>
<tr>
<td>CS2310 Data Structures and Algorithms with Java</td>
<td>10</td>
<td>1</td>
<td>1</td>
<td>CS &amp; EECS</td>
<td>2</td>
<td>54</td>
<td>J</td>
</tr>
<tr>
<td>CS3250 Distributed Systems</td>
<td>10</td>
<td>1</td>
<td>1</td>
<td>CS, EECS &amp; IS</td>
<td>Final</td>
<td>19</td>
<td>K</td>
</tr>
</tbody>
</table>

All students who submitted a coursework to a module listed in Table 1 were invited to fill in the questionnaire, which amounted to 284 students. The majority were undergraduates doing a degree programme major in Computing, Mathematics or Electronic Engineering, with about 6.7% of students doing Combined Honours with Computing, eg Business & Administration, Psychology and Mathematics. 75 students responded which amounted to a response rate of 26.4%. Table 2 shows the profile of the respondents.

Table 2: Profile of the respondents

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Distribution of the Respondents</th>
</tr>
</thead>
<tbody>
<tr>
<td>Year of study</td>
<td>Year 1: 70.8%; Year 2: 22.6%; Final Year: 6.6%</td>
</tr>
<tr>
<td>Degree programmes</td>
<td>Computing for Business: 16.0%; Computing Science: 42.6%; Combined Honours: 6.7%; Electronic Engineering major: 7.9%; Maths major: 21.3%; Multimedia Computing: 5.3%</td>
</tr>
<tr>
<td>Source of funding</td>
<td>Local Authority: 51.4%; Family Member: 17.6%; Self: 16.2%; Other: 14.9%</td>
</tr>
<tr>
<td>Ethnic origins</td>
<td>Asian: 46.7%; White: 40.0%; Black: 4.0%; Chinese: 4.0%; Mixed: 2.6%; Others: 2.7%</td>
</tr>
</tbody>
</table>

Data Analysis

The questionnaire was designed and administered using the Bristol Online Surveys (BOS) system (ILRT, 2011). The data was analysed using LibreOffice version 3.3.4 (LibreOffice, 2012) and the BOS system, including correlations and ranking statistics (ie mean and standard deviation). The answers to the open-ended questions were collated and categorised manually.
In the data analysis, the data relating to consistency in marking was grouped by module, whereas the remaining data was specific to individual tutors. Table 3 shows the profile of each tutor group.

<table>
<thead>
<tr>
<th>Module</th>
<th>Tutor ID</th>
<th>Group ID</th>
<th>Submitted a Coursework</th>
<th>Participated in Survey</th>
<th>Response Rate (%)</th>
<th>Degree Programmes</th>
</tr>
</thead>
<tbody>
<tr>
<td>CS1310</td>
<td>A</td>
<td>28</td>
<td>24</td>
<td>9</td>
<td>37.50</td>
<td>CS</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>34</td>
<td>29</td>
<td>2</td>
<td>6.90</td>
<td>CS</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>29</td>
<td>21</td>
<td>1</td>
<td>4.76</td>
<td>CE, CS, EECS</td>
</tr>
<tr>
<td>CS1020</td>
<td>D</td>
<td>31</td>
<td>20</td>
<td>6</td>
<td>30.00</td>
<td>CB</td>
</tr>
<tr>
<td></td>
<td>E</td>
<td>31</td>
<td>26</td>
<td>7</td>
<td>26.92</td>
<td>CB, Maths</td>
</tr>
<tr>
<td></td>
<td>F</td>
<td>29</td>
<td>25</td>
<td>7</td>
<td>28.00</td>
<td>IS, Maths</td>
</tr>
<tr>
<td></td>
<td>G</td>
<td>26</td>
<td>23</td>
<td>4</td>
<td>17.39</td>
<td>IS, MathC, Maths, MC</td>
</tr>
<tr>
<td></td>
<td>H</td>
<td>22</td>
<td>19</td>
<td>7</td>
<td>36.84</td>
<td>CH, Maths</td>
</tr>
<tr>
<td></td>
<td>I</td>
<td>32</td>
<td>29</td>
<td>10</td>
<td>34.48</td>
<td>CB, MathC, Maths</td>
</tr>
<tr>
<td>CS2310</td>
<td>J</td>
<td>54</td>
<td>51</td>
<td>17</td>
<td>33.33</td>
<td>CS &amp; EECS</td>
</tr>
<tr>
<td>CS3250</td>
<td>K</td>
<td>19</td>
<td>17</td>
<td>5</td>
<td>29.41</td>
<td>CS, EECS &amp; IS</td>
</tr>
</tbody>
</table>

We looked at the scores for each tutor on each of the questions in order to select the highest and lowest scoring tutors for further analysis. We selected those tutors who scored at least 4.0 on a particular question or, if no tutors have scored at least 4.0, we selected the highest scoring tutor. We also selected tutors who scored less than 3.0 on any of the questions. We then analysed the style of feedback provided by the selected tutors to try and identify any characteristics common to higher or lower scoring tutors. Note that the response rate for Tutors B & C was very low (less than 7%) and so we have excluded those tutors from further analysis because any conclusions relating to that data may be unreliable.

**Results and Findings**

**General Views on Assessment and Feedback**

To help establish which qualities of assessment and feedback were considered important and helpful, the respondents were asked to express their views on what was considered to be:

- the most important quality of assessment and feedback,
- the most helpful form of feedback,
- the most important form of feedback, and
- prompt feedback for coursework/assignments.

The respondents were also asked to indicate their preferred way of receiving feedback on their coursework.

Table 4 shows the qualities of assessment and feedback that were most frequently identified as being the most important. Figure 1 shows the forms of feedback considered as most helpful and important by the majority of the respondents. Table 5 shows the preferred ways of receiving feedback.
### Table 4: The most important quality of assessment and feedback

<table>
<thead>
<tr>
<th>The Most Important Perceived Quality</th>
<th>Responses (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>The awarding of marks to each submission is fair.</td>
<td>18.7</td>
</tr>
<tr>
<td>Feedback on my coursework is sufficiently detailed.</td>
<td>17.3</td>
</tr>
<tr>
<td>The marking is consistent between different students.</td>
<td>16.0</td>
</tr>
<tr>
<td>Feedback on my coursework helps me learn from my mistakes.</td>
<td>14.7</td>
</tr>
</tbody>
</table>

### Table 5: Preferred way of receiving coursework feedback

<table>
<thead>
<tr>
<th>Preferred Way to Receiving Feedback</th>
<th>Responses (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Online (ie via a coursework feedback system such as eCAF)</td>
<td>56.0</td>
</tr>
<tr>
<td>Verbally (eg via a face-to-face meeting or a tutorial)</td>
<td>16.0</td>
</tr>
<tr>
<td>Electronically (eg as a Word/PDF document or in plain text)</td>
<td>14.7</td>
</tr>
<tr>
<td>Paper-based</td>
<td>9.3</td>
</tr>
<tr>
<td>As a recording on a sound file</td>
<td>2.7</td>
</tr>
<tr>
<td>As a recording of screen capture with voice-over commentary</td>
<td>0.0</td>
</tr>
<tr>
<td>Other (eg a combination of the above)</td>
<td>1.3</td>
</tr>
</tbody>
</table>

### Most Helpful to Learning vs Most Important to Student Perception of various forms of feedback

#### Figure 1: Perceptions on the helpfulness and importance of various forms of feedback

- **The feedback highlights the exact location in my submitted work where I have made mistake(s).**
- **The feedback shows a summary of the tutor's comment to my work.**
- **The feedback shows my overall score for the assessment.**
- **The feedback highlights the area(s) in which I have made mistake(s).**
- **The feedback shows a detailed breakdown of my overall score for the assessment.**
- **The feedback shows the marking scheme and my score against each marking criterion.**
- **The feedback highlights the exact location in my submitted work where I have made mistake(s) AND gives advice on how to rectify the mistake(s).**
Most respondents (37.3%) expressed that feedback to coursework received within 2 working weeks was considered to be prompt, while 29.3% indicated that prompt feedback should be returned within 1 working week.

### Satisfaction Levels on Feedback Received

Twelve Likert-type statements were used to gauge how satisfied our respondents were regarding the timeliness, quality and helpfulness of the feedback received. The score ranges from 1 to 5, with 5 being the highest score. Figure 2 shows the results by the tutor groups.

- 2a. Feedback on my coursework has been prompt.
- 2b. The marking is consistent between different students (ie as far as I know, different student’s work with similar mistakes are given similar level of feedback and is awarded with similar marks).
- 2c. The awarding of marks to each submission is fair.
- 2d. The feedback (ie tutor's comments on my work) is easy to read.
- 2e. The feedback (ie tutor's comments on my work) is easy to understand.
- 2f. I have received sufficiently detailed feedback on my coursework.
- 2g. Feedback on my coursework has highlighted areas which I have made mistakes.
- 2h. Feedback on my coursework has helped me clarify things I did not understand.
- 2i. Feedback on my coursework has helped me learn from my mistakes.
- 2j. Feedback on my coursework has helped me learn my module.
- 2k. Feedback on my work has helped me improve my grade.
- 2l. On the whole, I am satisfied with the feedback received through e-CAF.

![Figure 2: The respondents' satisfaction level on the feedback received](image)

### Promptness

There were three tutors, F, G and K, who scored over 4.0 on promptness of their feedback. These tutors returned their feedback within 3 weeks of the submission date, achieving an average turnaround time of 12, 12.7 and 9 working days respectively. These results indicate that, for programming coursework, a turnaround time of 3 weeks is acceptable to students. Respondents in the survey have said that prompt feedback should be within 1 to 2 weeks and an acceptable waiting time is 2 to 3 weeks, which is consistent with the satisfaction rates for these three tutors.
Furthermore, Tutor J scored 4.0 on promptness of their feedback. The result from this tutor is interesting because in terms of the actual turnaround time, Tutor J's performance was similar to the low scoring tutors such as Tutors E & I. However, the respondents for Tutor J perceived their feedback to be prompt. The entire CS2310 cohort of 51 submissions were assessed by Tutor J and they received their feedback at the same time; whereas the respondents for Tutors E & I belonged to a larger cohort and other groups from the same cohorts had received their feedback much earlier. Hence, in comparison, feedback from Tutors E & I was perceived to be less timely. Moreover, the cohorts for Tutors E & I were significantly smaller than Tutor J's cohort. This suggests that when students recognise the assessment load is large, they might be more willing to wait.

Consistency and Fairness

In terms of consistency, no tutor scored a mean value of 4.0 or above. Tutors K & I achieved the best scores, ie 3.8 and 3.7 respectively. In terms of fairness, Tutors J & K scored at least 4.0, ie 4.00 and 4.20 respectively. Tutor K had the smallest class size (19 students); whereas Tutor I was one of the 6 tutors involved in assessing CS1020, which was the biggest module in this study (cf Tables 1 & 3). Despite the fact that Tutor I’s group comprised students from various degree programmes and some of their peers were assessed by different tutors in the module, the score for consistency in marking was fairly close to 4.0. This result indicated that consistency in marking can be achieved when suitable marking schemes and ICT tools are used.

All submissions for the coursework on CS2310 and CS3250 were marked by a single tutor (ie Tutors J and K respectively), hence it is not surprising that these tutors were considered fair in their marking. Our results showed that, with an overall satisfaction rate of 3.75 for fairness, none of the tutors were considered unfair in their marking. This suggests that, with an unambiguous marking scheme and a suitable ICT tool to facilitate the marking process, fairness in marking can be achieved even when multiple tutors are involved.

The general perception about manual assessments is that the larger the class size, the more difficult it is to achieve consistency, especially when multiple tutors are involved in the marking. To verify this perception, we have also computed the mean score on consistency in marking by modules. This is done by averaging the perception of consistency in marking (cf Question 2b) reported for each module. The results are shown in Table 6.

<table>
<thead>
<tr>
<th>Module</th>
<th>Number of Tutors involved in Marking</th>
<th>Population</th>
<th>Sample Size</th>
<th>Response Rate (%)</th>
<th>Mean Satisfaction</th>
<th>Standard Derivation</th>
</tr>
</thead>
<tbody>
<tr>
<td>CS1020</td>
<td>6</td>
<td>142</td>
<td>41</td>
<td>28.87</td>
<td>3.46</td>
<td>0.78</td>
</tr>
<tr>
<td>CS1310</td>
<td>3</td>
<td>74</td>
<td>12</td>
<td>16.22</td>
<td>3.50</td>
<td>0.67</td>
</tr>
<tr>
<td>CS2310</td>
<td>1</td>
<td>50</td>
<td>17</td>
<td>34.00</td>
<td>3.47</td>
<td>0.71</td>
</tr>
<tr>
<td>CS3250</td>
<td>1</td>
<td>16</td>
<td>5</td>
<td>31.25</td>
<td>3.80</td>
<td>0.84</td>
</tr>
</tbody>
</table>

The overall trend of our data supported the perception that marking done by a single tutor is more consistent than marking shared amongst multiple tutors. However, our data also supports the proposition that consistency in marking can still be achieved even with a relative large class involving multiple assessors (cf the results for CS1310 and CS2310).

Helpfulness

To help identifying the tutors whose feedback was considered helpful by our respondents, we focussed on the results from Questions 2d - 2k shown in Figure 2. We computed the mean score for these questions for each tutor. Overall, the feedback given by Tutors I & K was considered most helpful by our respondents with a mean score of 3.875 and 3.8, respectively. Feedback from Tutors D, E & G was considered least helpful, each with a mean score below 3.26.

To discover why feedback from some tutors is considered more helpful than others, we have analysed all feedback items written by each tutor and compared the results. The analysis comprises:
- categorising each feedback item manually based on its contents (eg whether it is a praise or it states the error, the consequence of the error and/or how to fix the error),

Innovation, Practice and Research in Engineering Education  EE2012

• counting the number of feedback items given to each marking criterion,
• tallying the total number of feedback items per submission,
• calculating the average number of words in a feedback item, and
• computing the percentage of feedback items linked to the exact location of the error.

Figure 3 shows the results of the categorisation. Table 7 shows the quantitative results.

Note that no reference to the actual student submissions was made during the feedback item analysis. The analysis only took account of the feedback items in isolation. If the content of a feedback item was unclear, e.g., containing a question mark only or some cryptic text, it was classified as Others.

Over 94% of the feedback items from Tutors K, I, E & D described the errors, consequences and/or fixes. 54.74% of the feedback items from Tutor G, however, was categorised as Others.

Table 7: Feedback Item Analysis: Quantitative results

<table>
<thead>
<tr>
<th>Tutor</th>
<th>Average comments per criterion</th>
<th>Average number of comments per submission</th>
<th>Average word count per comment</th>
<th>Total number of comments</th>
<th>Percentage of comments linked to exact location of the error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tutor K</td>
<td>1.58</td>
<td>7.88</td>
<td>22.90</td>
<td>126</td>
<td>67.46%</td>
</tr>
<tr>
<td>Tutor I</td>
<td>2.92</td>
<td>17.77</td>
<td>18.12</td>
<td>231</td>
<td>94.81%</td>
</tr>
<tr>
<td>Tutor E</td>
<td>2.12</td>
<td>9.17</td>
<td>16.71</td>
<td>110</td>
<td>93.64%</td>
</tr>
<tr>
<td>Tutor G</td>
<td>1.98</td>
<td>6.79</td>
<td>2.69</td>
<td>95</td>
<td>86.32%</td>
</tr>
<tr>
<td>Tutor D</td>
<td>3.72</td>
<td>23.25</td>
<td>11.83</td>
<td>279</td>
<td>80.65%</td>
</tr>
</tbody>
</table>

Discussion

Is satisfaction linked to the marks received?

We have compared the overall satisfaction of the feedback received with the actual coursework mark awarded for Tutors I & K. There appears to be no correlation between satisfaction and the marks received, i.e., a low coursework mark did not result in dissatisfaction. On the contrary, the submissions with relatively low marks indicated that the feedback given to their coursework helped them learn the module and improve their grade.

Figure 3: Types of feedback items used by each tutor
Helpful Styles of Feedback

In terms of helpfulness of feedback, the tutors who scored at least 4.0 in the survey tend to write feedback in a feed forward style, with each feedback item stating:

1. What has been done inappropriately.
2. What is the consequence.
3. How to improve it.

The feedback given by the tutors who were considered most helpful (Tutors I and K) contains relatively more comments identifying errors, has relatively fewer comments identifying a fix without also identifying the errors, and relatively more comments containing some combination of the error, the consequence of the error and a fix for the error. This latter statistic is consistent with the survey feedback that shows 33.3% of respondents considered comments that address both the mistake and a way of rectifying the mistake to be helpful whereas only 5.3% of respondents found comments that address the mistake alone to be helpful.

A detailed explanation on how to fix every mistake made will have a detrimental effect on the timeliness of feedback. Our results show that, so long as the feedback is accompanied by a brief statement on how to rectify the problem or improve the work, students were satisfied with a general feedback statement such as “See sample solution for detail.”

The results of our analysis show that our respondents appreciated feedback that clearly identified the mistakes made, even when no further information on how to rectify the issue was given in the feedback. As shown in Figure 3, Tutor E, whose feedback was rated as the least helpful, stated what should be done to improve the work in 49.09% of his feedback items without actually stating the errors. Tutors K & I, whose feedback was considered more helpful, clearly identified the mistakes made by students in 73.02% and 88.74% of their feedback items respectively. These results show that, when giving feedback to programming coursework, it is very important for tutors to state the mistakes made in each submission clearly. This finding also suggests that feed forward without feedback is not an effective way to facilitate learning. Race’s (2007) “ripple on a pond” model to learning suggests that learning is driven by “wanting/need” and feedback has the capability to increase the wants and needs. Helping students to identify mistakes therefore has an effect of motivating them to further engage in their learning.

On average, the best-scoring tutor (ie Tutor K) in our study gave 7.88 feedback items to each student submission. Tutor D, who scored less than 3.0 in Question 2h (ie the feedback helped to clarify things which the respondents did not understand), gave an average of 23.25 feedback items to each submission. Our analysis revealed that the number of formative feedback items given to each student did not affect students’ perception on the helpfulness of feedback. The style of feedback, however, appeared to be a dominant factor for student satisfaction. The dominant features of formative feedback written by our best-scoring tutors are as follows:

- **Clarity**: The feedback was not wordy, but clearly stated errors in the submitted work.
- **Ease of understanding**: The message delivered in each feedback item is understandable even to those who have not read the coursework specification nor the actual submitted work.
- **Simplicity of language**: The feedback items were written in simple English akin to everyday discourse as opposed to academic language, free from unnecessary use of jargon or terminology that is not covered in the subject, eg “A service proxy should be reusable for subsequent uses - yours is a use-once-only proxy, which is not ideal”.
- **Friendliness of tone**: The feedback items were written in a friendly, neutral and unthreatening tone, expressing an honest opinion, eg “I do not like this trick very much - I would prefer to propagate the exception or throw another kind of exception”.
- **Genuine praise**: Suitable level of praise was given to work which showed signs of non-routine execution of knowledge or exceptional understanding, eg, “I like your method - very clean interface, also well documented”.

In writing feedback, our best-scoring tutor for Question 2h also used a higher proportion of qualified evaluations (Burke & Pieterick 2010, p. 48), which used “qualifiers to temper the tutor’s authority and imply less control”, as opposed to commands (Burke & Pieterick 2010, p. 48), which told students exactly what to do or should have been done. Our worst-scoring tutors gave feedback which was primarily composed of commands.

Bloxham & Boyd (2007, p. 109) suggested that questioning comments, which facilitate a dialogue between the student and the assessor, are considered to be more effective than confident statements.
Our data showed that, with programming coursework, the absence of questioning comments did not have a significant influence on the perception of the helpfulness of the feedback. This is perhaps due to the practical and objective nature of the assignments and the students’ eagerness to learn from their mistakes. The 2011 NSS results by subject areas showed that the mean satisfaction score for helpfulness of feedback in Computer Science discipline is 3.6, which is amongst the second quartile. While confident statements are considered to be the typical feedback style in Computer Science, students did not seem to show dissatisfaction over such style of feedback. However, further investigation is desirable to establish the helpfulness of questioning comments in practical programming coursework.

**Unhelpful Styles of Feedback**

Our analysis on the feedback items showed that formative feedback which simply stated how to fix an error in the submitted work was considered unhelpful to students. Tutors E & D scored less than 3.0 in Question 2h (ie the feedback helped to clarify things which the respondents did not understand). A high proportion of the feedback items from these tutors identified fixes only, without stating exactly what the error was and why the work was inappropriate. Without being informed of the error made, some students therefore struggled to see why they should employ a different approach in their work and hence did not find the feedback helpful. This style of feedback also puts the responsibility on the student to identify the errors made when they might not possess sufficient knowledge and understanding to do so.

Feedback items written as complete sentences, rather than phrases, were considered easy to understand by our respondents. This is supported by the results that both Tutors K & I scored at least 4.10 on ease of understanding (cf Question 2e) and both tutors mainly wrote feedback in complete sentences. However, our results showed that complete sentences did not guarantee helpfulness of feedback. For example, feedback from Tutor E was written in complete sentences and they were considered easy to understand (ie with a satisfaction level of 3.71 for Question 2e). However, our respondents were dissatisfied with such feedback as the feedback failed to clarify things that they did not understand (ie a satisfaction level of 2.71 for Question 2h). A close inspection of the feedback items showed that 49.09% of feedback items from Tutor E only identified how to correct errors without actually stating the error and its consequence. Students therefore could not appreciate the relevance of the feedback.

Our analysis of the feedback styles from the worst-scoring tutors (ie Tutors E, G & D) revealed the following unwelcome features in formative feedback:

- **Cryptic messages**: eg "remaining credits", "table header", "equal" or even "??". The main purpose of feedback is to help students identify any gaps in learning and understanding so that the gaps can be closed (Burke & Peiterick, 2010). If the feedback cannot be deciphered, it serves no purpose. Hence, cryptic messages must be avoided.

- **Rhetorical questions**: eg "What if it is negative?", "What happens if X?", "Why X, but not Y?". Such questions assumed a level of understanding which the students might not possess and therefore leaving them feeling confused or even mystified.

- Simply pointing out that something should not have been used or done, without explaining how else to accomplish the task, eg a brief statement such as "Should not use X". In some cases, students chose to use certain methods in their work or failed to address certain requirements of the coursework because they did not know how else to accomplish the task. Hence, feedback which simply states something should not be done is considered unhelpful.

- Grammatical and spelling mistakes, which hindered the readability of the feedback.

- Feedback written in short phrases rather than complete sentences, eg "No need", "Not efficient". Such short phrases are less likely to convey the precise issue identified in the student's work.

**Conclusion**

We have presented the results and findings of our first study into how to prepare feedback to programming coursework that would facilitate learning. Our results showed that suitable use of ICT help to improve the timing and the quality of feedback. Our students preferred receiving feedback electronically, in particular via a purpose-built, online coursework feedback system such as eCAF. They found feedback that highlighted the exact location in their submitted work where mistakes were made and gave advice on how to rectify the mistakes most helpful to their learning.
Based on the results of our analysis on all feedback items collected in this study, we have identified some qualities of programming coursework feedback that were regarded as helpful and unhelpful. Our results suggested that clarity, ease of understanding, simplicity of language used, friendly tone and genuine praises were the key features of feedback to programming coursework that helped our students to close the gap in their learning. In terms of the contents of each feedback item, students found it more helpful when a significant proportion of feedback identified the errors and provided a fix for each.

We recognise that this study is our first step in our quest to understand how to prepare feedback to programming coursework that would improve students' learning experience. Our next step is to extend this study to include students from other Science, Technology, Engineering and Mathematics (STEM) programmes and to perform a more detailed analysis on the feedback given to students.

References


Copyright statement

Copyright © September 2012, authors as listed at the start of this paper. This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivs 3.0 Unported License (CC BY-NC-ND 3.0).