Evaluating feedback alignment in language assignments: tutor and student perspectives

María Fernández-Toro and Concha Furnborough

School of Languages and Applied Linguistics, The Open University, Milton Keynes, United Kingdom

Feedback alignment is key to the effectiveness of formative feedback, but often tutors can only guess whether their feedback is consistent with students’ needs and expectations. This study aims to identify areas of potential misalignment by bringing together self-reported data and feedback analysis.

Two parallel surveys were conducted with 736 students and 96 tutors from a distance learning undergraduate programme in Language Studies, to compare their self-reported attitudes and behaviours. Survey data were then triangulated through the analysis of 216 marked assignments using the Feedback Analysis Categorisation Tool (FACT) according to the orientation (focus on a weakness or a strength) and depth (layers of scaffolding) of feedback comments. Misalignment was most apparent in relation to feedback on strengths and mutual assumptions regarding feedback ownership. The findings support the need for a more dialogic approach to feedback and confirm the value of feedback analysis as a means of evaluating feedback alignment for sustainable assessment.

Keywords: distance learning, language learning, assignment feedback, feedback alignment, scaffolding, sustainable feedback

Introduction

In recent years feedback analysis has often been used in an attempt to identify the features that make up “effective” feedback (Chase & Houmanfar, 2009; Chetwynd & Dobbyn, 2011; Donovan, 2014; Hamer, 2015; Hughes, 2011; Hyatt, 2005). However, feedback itself is only as effective as the students’ engagement with it (Handley, Price, & Millar, 2011), and such engagement is only possible if students and tutors share a common understanding of the role of feedback in the learning process. Rather than the
intrinsic quality of a tutor’s feedback, it is the extent to which such feedback is aligned with the students’ needs and expectations that will determine its effectiveness (Orsmond & Merry, 2011).

Furthermore, the last decade has seen a shift towards sustainable models of assessment where the nature and role of feedback is increasingly driven by learners themselves (Beck, Skinner, & Schwabrow, 2013; Boud, 2000; Boud & Soler, 2016; Carless, Salter, Yang, & Lam, 2011; Everhard, 2015b; Fastré, Klink, Sluijsmans, & Merriënboer, 2013; Nguyen & Walker, 2014). It is therefore essential to gain a better understanding of the beliefs and perceptions of students and their tutors regarding feedback, and to identify those areas where better alignment would be needed.

1 Literature review

1.1 The role of feedback

Assessment feedback has been defined as advising students how to close “the gap between [their] actual and desired levels of performance” (Black & William, 1998; Hunt, 2001; Ramaprasad, 1983). This requires learners to both perceive such a gap and appreciate that action on their part as well as by their tutors is needed to close it (Black & William, 1998; Ramaprasad, 1983).

Action to close this gap largely depends on alignment of perspectives between students and tutors. In their study of the relationship between self-regulation and feedback Nicol and Macfarlane-Dick (2006) make clear the responsibilities of both parties for achieving this, drawing a distinction between the feedback offered by tutors on their students’ assignments, characterised as external feedback, and the internal feedback that students produce as they interpret and build on this information at different levels, e.g. cognitive, metacognitive and socio-affective or motivational. Nicol
and Macfarlane-Dick argue that feedback will only be effective if students are able to incorporate it into their future learning processes. This requires appropriate feedback from tutors, which relies on shared perspectives between tutors and students.

Such alignment of perspectives is crucial for distance learners (Hurd, 2000, 2006; White, 2003) as in many cases assignment feedback constitutes their primary or sole contact with their tutor, and a unique opportunity to obtain informed and supportive information about their performance (Hyland, 2001; Ros i Solé & Truman, 2005; Truman, 2008). For these students the stakes are therefore particularly high.

The role of feedback on motivation is also significant (Walker & Symons, 1997, pp. 16–17). Nicol and Macfarlane-Dick identify “encouraging positive motivational beliefs and self-esteem” (2006, pp. 211–212) as one of their principles of good feedback practice for facilitating self-regulation. Other studies (Nicol & Milligan, 2006) have shown that the combination of external and internal feedback can assist students to take more responsibility for managing their own learning. Furthermore, the ability to use a combination of external and internal resources has been identified as a defining element of learner autonomy (Everhard, 2015b).

Everhard positions the relationship between assessment and learner autonomy as a continuum between heteronomy and autonomy. Summative assessment (assessment of learning) represents the highest degree of heteronomy, as the regulation of learning is the remit of “more knowledgeable others”. Between the two extremes, formative assessment (assessment for learning) represents varying degrees of interdependence where control is shared. The highest degree of learner autonomy occurs in sustainable assessment (assessment as learning), where students regulate their own learning through the use of internal and external resources. (Everhard, 2015b, p. 19)
Learning a language presents additional challenges. Unlike other subjects, language learning demands a great deal of self-expression and involves the learner’s self-concept (Horwitz, Horwitz, & Cope, 1986), which can result in what is known as ‘language anxiety’ (cf. Ehrman, 1996; Gardner & MacIntyre, 1993). The emotional dimension has been found to be key to success or failure in learning a language (Arnold, 1999; Hurd, 2008; Oxford, 1990). Such effects are potentially heightened in the case of distance language learners (Hurd, 2005), and must therefore be taken into account in assignment feedback (Truman, 2008).

1.2 Alignment between tutor and student perceptions

The discussion so far shows that tutors’ and students’ perceptions regarding the purposes and practice of assignment feedback need to be aligned with each other. However, research into students’ perceptions of and reactions to feedback has variously indicated a limited concept of the purposes of assessment (Maclellan, 2001). Weaver (2006) related the students’ uncertainty about what was required of them to a lack of guidance on interpreting and using feedback. The findings of McDowell (2007) and Burke (2009) also confirmed areas of misalignment between staff expectations and student awareness, and the danger of tutors assuming unrealistic levels of metacognitive awareness in their students. Other studies have confirmed a mismatch between students’ own needs and expectations regarding feedback on the one hand, and tutors’ assumptions and practices on the other (Cohen & Cavalcanti, 1990; Orsmond & Merry, 2011).

Crucially, the students’ actual responses to the feedback that they receive remain largely invisible (Price, Handley, & Millar, 2011) and therefore unknown to their tutors. This is especially likely to occur in distance learning because opportunities for direct contact between students and tutors are more limited than in classroom learning, making
their respective ideas about feedback even more prone to differ. In his Theory of Transactional Distance, Moore (1997; 1980, 2012) posited that in distance education, ‘distance’ not only refers to space and time, but to “a psychological and communications space” between learners and their instructors, which he termed “transactional distance”, that could potentially result in misunderstandings (Moore, 2012, p. 22).

The first aim of this paper is therefore to identify specific areas of misalignment between the students’ and tutors’ ideas about feedback that may hinder its effectiveness.

1.3 Analysing tutor feedback

Even in those cases where the beliefs and perceptions of students and their tutors appear to be aligned, their respective behaviours when giving and receiving actual feedback may tell a different story. It is therefore important to understand how real-life feedback is used, and to what extent actual feedback behaviour matches the assumptions and expectations reported by students and tutors. This is the other purpose of the present study.

Glover and Brown (2006, pp. 83–85) devised a systematic approach for analysing tutors’ comments and annotations on written assignments in order to enhance the quality of formative feedback on Science assignments. Two sets of categories were proposed: Type and depth. Type referred to comments focusing on content, on skills, on strategies for future learning and on comments of a motivational nature. Comments focusing on content or skills were further categorised in terms of their depth. For comments about weaknesses this included three possible levels: (1) “An issue acknowledged”, (2) “A correct response provided”, (3) “The reason why a student's answer was inappropriate or why the preferred answer was appropriate”. A similar classification was applied to feedback on strengths, although categories 2 and 3 were
not as clearly differentiated: Praise alone (“good”, “well done”) was interpreted as category 1; and “the basis for pride and encouragement” as category 2 or 3, depending on the extent of the explanation provided.

Brown and Glover’s 2006 study, which focused on science assignments, was replicated for technology assignments with broadly comparable results (M. Walker, 2009). The method was later extended beyond the STEM (Science, Technology, Engineering and Mathematics) subject area, through a comparative study of assignment feedback in technology and languages (Fernández-Toro, Truman, & Walker, 2013). This revealed significant differences when providing feedback on language assignments, notably that comments on skills far outweighed comments on content. The depths at which feedback was offered were also different in Languages, with a greater emphasis on correction (category 2), and proportionally less emphasis on indicating (category 1) and explaining (category 3). Correction still constituted the most common type of comment in both subject areas. It was also apparent that in language assignments certain errors were not only indicated, but also categorised without however being corrected. This is common practice for frequently made language mistakes, where an indication of the type of error made (e.g. ‘spelling’) is often all that students need in order to supply the correction themselves (Ros i Solé & Truman, 2005). This practice is acknowledged in the modified version of Glover and Brown’s original classification used in the present study, where an additional level entitled “categorising” has been added.

2 Feedback Analysis Categorisation Tool (FACT)

The “Feedback Analysis Chart for Tutors”, later renamed Feedback Analysis Categorisation Tool (FACT) was developed to provide an observational method for evaluating tutor feedback on language assignments as an alternative to commonly used
evaluation methods involving self-reported data. While it draws on Glover and Brown’s (2006) notion of feedback depth, the FACT approach does not consider the different categories as mutually exclusive levels of depth, but as available layers of scaffolding that may or may not occur together within a given feedback comment. Scaffolding consists of “‘controlling’ those elements of the task that are initially beyond the learner's capacity, thus permitting him [or her] to concentrate upon and complete only those elements that are within his [or her] range of competence.” (Wood, Bruner, & Ross, 1976, p. 90). This enables the learner to operate within what Vygotsky defined as the zone of proximal development (Vygotsky, 1978, p. 86), which is the gap between a learner’s current ability and what she/he is able to accomplish with support from others. Scaffolding is then gradually removed until the learner is able to perform the task independently.

The revised instrument used in this study adds two layers of scaffolding to the ‘levels of depth’ identified by Glover and Brown’s (2006). These are categorisation of errors or strengths as layer 2; and feed forward (Walker, 2009) as layer 5. The five layers of scaffolding are exemplified in Table 1.

FACT analysis comprises five possible layers of scaffolding. The same categories are used for feedback on weaknesses and strengths: (1) indicated; (2) categorised/described; (3) corrected/exemplified; (4) explained; (5) advice relating to future action to avoid an error or build on an existing strength.
Table 1. Layers of scaffolding in language assignment feedback

<table>
<thead>
<tr>
<th>Layer ↓</th>
<th>Feedback focusing on weaknesses</th>
<th>Feedback focusing on strengths</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Error identified</td>
<td>Strength identified</td>
</tr>
<tr>
<td></td>
<td><em>Examples: Error underlined; question mark next to a sentence, word or paragraph.</em></td>
<td><em>Examples: ‘Good’; tick written next to a sentence, word or paragraph.</em></td>
</tr>
<tr>
<td>2</td>
<td>Error categorised</td>
<td>Strength categorised or described (as per marking criteria)</td>
</tr>
<tr>
<td></td>
<td><em>Example: Gender agreement</em></td>
<td><em>Example: ‘You use a wide range of language structures, including subjunctive clauses.’</em></td>
</tr>
<tr>
<td>3</td>
<td>Error corrected</td>
<td>Strength illustrated with specific example from student’s work</td>
</tr>
<tr>
<td></td>
<td><em>Example: Student writes ‘Manos blancos’ and tutor corrects the ending to ‘blancas’</em></td>
<td><em>Example: ‘Good use of the subjunctive’ written next to the sentence where the student used it.</em></td>
</tr>
<tr>
<td>4</td>
<td>Explanation given</td>
<td>Explanation given</td>
</tr>
<tr>
<td></td>
<td><em>Example: ‘Although it ends in ‘o’ mano is a feminine noun.’</em></td>
<td><em>Example: ‘This connector makes it very clear to the reader that a new section is starting here.’</em></td>
</tr>
<tr>
<td>5</td>
<td>Advice given on how to prevent errors in future performance</td>
<td>Advice given on how to develop existing strengths in future</td>
</tr>
<tr>
<td></td>
<td><em>Example: ‘Revise correct verb endings in section 6.1 of your grammar book.’</em></td>
<td><em>Example: ‘Excellent example, see also […] for an interesting discussion of a similar case.’</em></td>
</tr>
</tbody>
</table>

It should be emphasised that although the layers of scaffolding are presented linearly in Table 1 and occasionally referred to as “levels”, they do not necessarily occur sequentially. For example, an error may well be identified (1) and corrected (3) without being categorised (2), or a strength may be described (2) and explained (4) without making reference to specific examples (3) from the student’s work.
Furthermore, a simple correction may well elicit deeper thought processes than a very general explanation. In other words, layer 4 is not necessarily more cognitively deeper than layer 3.

FACT analysis can be applied to spoken as well as written assignments, and the feedback itself may also be in spoken or written form, provided that oral feedback is transcribed for analysis purposes.

3 The study

This study was designed to evaluate student and tutor attitudes to and perceptions of language assignment feedback in order to identify to what extent these sets of attitudes and perceptions were aligned (1) with each other, and (2) with actual feedback-related behaviours:

(1) Comparative survey data from students and tutors were expected to cast light on potential misalignments between their respective perceptions of key aspects of the assignment feedback process.

(2) In addition to tutor and student surveys, FACT analysis of tutor feedback was used to verify through triangulation any issues identified in survey data. Special attention was given to possible misalignments between tutors’ perceptions of what they were doing, as revealed in their questionnaire responses, and what they were actually doing, as evidenced by the FACT analysis.

The study was conducted in a distance Higher Education institution (Open University) where the use of both audio-recorded and written e-feedback had been standard practice in the Department of Languages for a number of years.

The research focused on two broad questions:
What are the differences and similarities between students’ and tutors’
perceptions about assignment feedback?

To what extent are these perceptions reflected in the tutors’ actual feedback?

4 Method

This project comprised two steps. The first consisted of establishing students’ beliefs
about and perceptions of feedback, and comparing this with tutor perceptions of their
own practice in order to answer question 1. The second step was to use the FACT tool
to analyse quantitatively the orientation of the tutors’ feedback comments (focus on
strengths vs. focus on weaknesses) and the layers of scaffolding provided in each case
in order to answer question 2.

4.1 Data collection

Three instruments were used to collect the study data:

(1) Student survey

(2) Tutor survey

(3) FACT analysis of actual tutor feedback

Surveys

The student and tutor surveys were conducted online, at around the mid point of the
courses. They consisted of two parallel questionnaires comprising similar items for
students and tutors (see supplemental material: Student Survey and Tutor Survey). These
included a set of questions or statements predominantly using Lickert scales to elicit
perceptions of the feedback process, how students perceived and responded to the feedback they received, and how tutors reflected on their own feedback. The questions were generic and did not refer to a particular assignment. All survey responses were anonymous.

The student questionnaire was completed by 736 students of all languages at different levels (Table 2) and represented a good cross-section of the student cohort with an overall response rate of 8%.

Table 2. Summary of student respondents

<table>
<thead>
<tr>
<th>Language</th>
<th>Beginner</th>
<th>Lower inter.</th>
<th>Upper inter.</th>
<th>Advanced</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chinese</td>
<td>27</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>27</td>
</tr>
<tr>
<td>French</td>
<td>106</td>
<td>72</td>
<td>23</td>
<td>31</td>
<td>232</td>
</tr>
<tr>
<td>German</td>
<td>63</td>
<td>41</td>
<td>16</td>
<td>9</td>
<td>129</td>
</tr>
<tr>
<td>Italian</td>
<td>85</td>
<td>46</td>
<td>-</td>
<td>-</td>
<td>131</td>
</tr>
<tr>
<td>Spanish</td>
<td>90</td>
<td>80</td>
<td>18</td>
<td>22</td>
<td>210</td>
</tr>
<tr>
<td>Welsh</td>
<td>7</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>7</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>378</strong></td>
<td><strong>239</strong></td>
<td><strong>57</strong></td>
<td><strong>62</strong></td>
<td><strong>736</strong></td>
</tr>
</tbody>
</table>

The tutor questionnaire was sent to all tutors in the Language Studies degree programme. 105 tutors took the survey, providing 96 complete responses.

*Actual tutor feedback*

Tutor feedback was obtained from tutors in Spanish, who were invited to make their marked assignments, both oral and written, available for analysis using the FACT tool.

A sample of 36 tutors was selected, spread evenly across the four course levels (9 tutors
A total of 108 written and 108 spoken assignments (i.e. 3 samples each per tutor, reflecting a broad range of marks awarded) were used in the study. The project complied with the Open University’s ethical standards and was scrutinised by the institution’s Human Research Ethics Committee. Informed consent was obtained from the tutors to use their feedback as data. Student consent was not required as it was the tutor’s feedback, not the student’s written output that was to be analysed. All tutors’ and students’ names were removed prior to analysis and audio files were transcribed.

4.2 *Data analysis*

4.2.1 *Survey data*

The two questionnaires were analysed for areas of alignment or misalignment between student and tutor perceptions, focusing on four broad areas:

- The importance that students attach to feedback (as self-reported by students/as perceived by tutors)
- Perceptions of the purpose and nature of feedback (among students/among tutors)
- The students’ understanding of the feedback received (as self-reported by students/as perceived by tutors)
- The tutors’ approaches to giving feedback (as self-reported by tutors/as perceived by students)

The final area focused more specifically on the balance between comments on weaknesses and comments on strengths, as well as the level of scaffolding provided, to reflect the FACT categories subsequently used in the analysis of tutor feedback.
4.2.2 Tutor feedback data

The FACT method was then applied to code the comments given by tutors on a total of 108 written and 108 spoken assignments together with their accompanying e-feedback summary forms. Each comment or annotation was coded according to its orientation (focus on a weakness or a strength) and layers of scaffolding (1: Indicated; 2: Categorised/Described; 3: Corrected/Exemplified; 4: Explained; 5: Feeding-forward).

A sample of 6 sets (3 speaking assignments and 3 written assignments) was analysed independently by two researchers and subsequently tested for inter-rater reliability. As it was the first time the FACT tool was tested, several adjustments were required in order to clarify the following issues:

- Defining what constitutes a feedback comment or annotation (e.g. a single sentence may refer to more than one issue)
- Deciding whether comments including more than one layer of feedback should be coded twice
- Deciding whether comments referring to both a strength and a weakness should be coded twice

The general approach was to code items under more than one category where appropriate, to provide a more accurate account of the strategies used. This approach was adopted in order to avoid loss of information and reliability resulting from oversimplification. After three rounds of tests a highly reliable version of the coding conventions was produced, with an inter-rater reliability index of 0.97. Hereafter all feedback items will be referred to as 'comments', whether they consist of actual comments or ad-hoc annotations (such as ticks, words underlined, letters crossed out, etc.) on a written script.
Each tutor comment was then coded by an experienced educational researcher who was a native speaker of Spanish, and transferred to a spreadsheet. Statistical analysis of the coded data was then carried out by an external consultant in order to establish the frequencies of different types of feedback in the three media used in the Language Studies programme: tutors’ annotations on written scripts, audio-recorded feedback on speaking assignments, and electronic feedback summary forms (used for both speaking and writing assignments).

5 Results

5.1 Student and tutor surveys

5.1.1 Importance attached to feedback

The survey confirmed informal tutor comments which assumed that students were primarily concerned with the mark awarded (Table 3). Students themselves on the other hand placed greater emphasis on ‘seeing what they got wrong’.

Table 3. Most important aspect of feedback for students

<table>
<thead>
<tr>
<th>According to the tutors</th>
<th>(%)</th>
<th>According to students themselves</th>
<th>(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>good mark</td>
<td>33</td>
<td>seeing what they got wrong</td>
<td>35</td>
</tr>
<tr>
<td>feeling supported</td>
<td>27</td>
<td>good mark</td>
<td>21</td>
</tr>
<tr>
<td>advice for future</td>
<td>14</td>
<td>gauging own progress</td>
<td>19</td>
</tr>
<tr>
<td>gauging own progress</td>
<td>13</td>
<td>advice for future</td>
<td>17</td>
</tr>
<tr>
<td>seeing what they got wrong</td>
<td>13</td>
<td>feeling supported</td>
<td>8</td>
</tr>
</tbody>
</table>
The vast majority of students were keen to see their feedback (84% said they looked at it on the same day it was returned), and 95% responded that they always looked at all feedback they received. In contrast tutors’ believed that only 73% of students looked at all the feedback. Between 30% and 40% of tutors also considered that there were some parts of the feedback that a ‘significant minority’ of students did not look at.

5.1.2 Perceptions of the purpose and nature of feedback

Student and tutor respondents were asked to comment on 14 statements regarding feedback.

Areas of alignment between student and tutor perceptions. As might be expected, there were certain fundamentals where there was a broad degree of consensus (Table 4).

Table 4. Areas of consensus/alignment between students and tutors

<table>
<thead>
<tr>
<th>Statement</th>
<th>Tutors (%)</th>
<th>Students (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Feedback on students’ use of language was useful</td>
<td>97</td>
<td>93</td>
</tr>
<tr>
<td>3. Feedback helps students see what they got wrong</td>
<td>99</td>
<td>94</td>
</tr>
<tr>
<td>2. Feedback on content and structure was useful</td>
<td>93</td>
<td>87</td>
</tr>
<tr>
<td>10. Feedback does help students to learn</td>
<td>84</td>
<td>87</td>
</tr>
</tbody>
</table>

Areas of misalignment between student and tutor perceptions. In other areas tutors were very confident that feedback was achieving its goals, whereas students were markedly less certain (although percentages were still high). Note that although there was almost universal agreement that feedback helped students see what they got wrong...
(Table 4), students were rather less sure than their tutors that it helped them identify what they had got right or how they could improve their performance (Table 5).

Table 5. Areas of discrepancy between students and tutors

<table>
<thead>
<tr>
<th>Statement</th>
<th>Tutors (%)</th>
<th>Students (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>12. Feedback helps students understand mark</td>
<td>97</td>
<td>87</td>
</tr>
<tr>
<td>11. Feedback makes students feel more supported</td>
<td>95</td>
<td>82</td>
</tr>
<tr>
<td>4. Feedback helps students see what they got right</td>
<td>97</td>
<td>85</td>
</tr>
<tr>
<td>5. Feedback shows students how they can improve</td>
<td>99</td>
<td>83</td>
</tr>
</tbody>
</table>

By contrast there were other areas (Table 6) where students felt more confident or certain than their tutors:

Table 6. Areas where students’ perceptions are more positive than their tutors’

<table>
<thead>
<tr>
<th>Statement</th>
<th>Tutors (%)</th>
<th>Students (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>7. Students make efforts to act on feedback</td>
<td>69</td>
<td>87</td>
</tr>
<tr>
<td>14. Students look forward to seeing feedback</td>
<td>60</td>
<td>87</td>
</tr>
<tr>
<td>6. Students are able to remember most of feedback</td>
<td>41</td>
<td>76</td>
</tr>
</tbody>
</table>

A substantial minority of both students (30%) and tutors (27%) appeared uncertain whether feedback actually raised student confidence in their learning.
There were also other major discrepancies that lay at the heart of feedback practice. The tutor’s role as provider of corrections and explanations of errors made (Table 7) was regarded as more important by students than it was by tutors themselves.

Table 7. Role of tutors in assignment feedback

<table>
<thead>
<tr>
<th>Statement</th>
<th>Tutors (%)</th>
<th>Students (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>8. Tutors should correct every single error</td>
<td>38</td>
<td>58</td>
</tr>
<tr>
<td>9. Tutors should explain every single correction</td>
<td>44</td>
<td>72</td>
</tr>
</tbody>
</table>

5.1.3 Understanding feedback received

In general students felt that they understood the feedback they had received; only 20% of students found feedback (‘occasionally’ / ‘often’ / ‘always’) unclear; if this were the case 62% said they would ask their tutor, although 56% of the tutors stated that they were never or rarely asked for clarification, and 38% only by a ‘significant minority’.

5.1.4 Approaches to giving feedback

Feedback on weaknesses. Students showed high satisfaction levels on all counts (Table 8); they felt satisfied that tutors indicated and corrected errors ‘just often enough’, although the majority of dissatisfied students (12% of respondents) would have liked more explanations.
### Feedback on weaknesses

<table>
<thead>
<tr>
<th>Statement</th>
<th>% Tutors</th>
<th>% Students satisfied</th>
<th>% Students not satisfied</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(In brackets: *See note below)</td>
<td>(In brackets: ‘Not often enough’)</td>
<td></td>
</tr>
<tr>
<td>1 Indicate errors made</td>
<td>97 (67; 30)</td>
<td>95</td>
<td>5 (2)</td>
</tr>
<tr>
<td>2 Correct errors made</td>
<td>94 (61; 33)</td>
<td>94</td>
<td>6 (3)</td>
</tr>
<tr>
<td>3 Give explanations</td>
<td>90 (50; 40)</td>
<td>85</td>
<td>15 (12)</td>
</tr>
</tbody>
</table>

*Note: Tutor figures in brackets are for ‘at every opportunity’; ‘generally (more than half of the time)’

### Feedback on strengths

The highest proportion of tutors (72%) tried to indicate ‘at every opportunity’ when students did well (Table 9). However, the figures for actually describing student strengths, and to a lesser extent telling students how to improve in future, were substantially lower, and this is also reflected in the higher percentage of students expressing dissatisfaction in these respects.

<table>
<thead>
<tr>
<th>Statement</th>
<th>% Tutors</th>
<th>% Students satisfied</th>
<th>% Students not satisfied</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(In brackets: *See note below)</td>
<td>(In brackets: ‘Not often enough’)</td>
<td></td>
</tr>
<tr>
<td>4 Indicate that did well</td>
<td>95 (72; 23)</td>
<td>90</td>
<td>10 (6)</td>
</tr>
<tr>
<td>5 Describe strengths specifically</td>
<td>81 (46; 35)</td>
<td>76</td>
<td>24 (22)</td>
</tr>
<tr>
<td>6 Explain how to improve in the future</td>
<td>88 (61; 27)</td>
<td>80</td>
<td>20 (17)</td>
</tr>
</tbody>
</table>

*Note: Tutor figures in brackets are for ‘at every opportunity’; ‘generally (more than half of the time)’
5.2 FACT analysis

108 spoken and 108 written assignments in Spanish were analysed within the study, with every item of feedback coded and recorded. A total of 12509 comments were tagged (Table 10).

Table 10. Summary of overall findings

<table>
<thead>
<tr>
<th>Orientation of Comment</th>
<th>Number of Comments</th>
<th>% of total</th>
<th>Avge. number per student per piece of submitted work</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strength</td>
<td>3635</td>
<td>29%</td>
<td>16.8</td>
</tr>
<tr>
<td>Weakness</td>
<td>8874</td>
<td>71%</td>
<td>41.1</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>12509</strong></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Feedback on weaknesses accounted for more than two thirds of all feedback (71%), whereas less than a third (29%) related to strengths in the students’ work.

Table 11. Number of comments at each feedback level for strengths and weaknesses

<table>
<thead>
<tr>
<th>Level</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strength</td>
<td>1216</td>
<td>2070</td>
<td>280</td>
<td>45</td>
<td>24</td>
<td>3635</td>
</tr>
<tr>
<td>Weakness</td>
<td>84</td>
<td>1501</td>
<td>4284</td>
<td>2126</td>
<td>879</td>
<td>8874</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>1300</strong></td>
<td><strong>3571</strong></td>
<td><strong>4564</strong></td>
<td><strong>2171</strong></td>
<td><strong>903</strong></td>
<td><strong>12509</strong></td>
</tr>
</tbody>
</table>
There was also a considerable contrast between the levels of the feedback offered on strengths and on weaknesses respectively (Table 11).

This can be appreciated more clearly when the distributions are recalculated as percentages (Figure 1), since the adjustment controls for the fact that over twice as many weaknesses were tagged as strengths.

Figure 1. Distribution of feedback levels within strengths and weaknesses

<table>
<thead>
<tr>
<th></th>
<th>% of total strength comments</th>
<th>% of total weakness comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Indicated</td>
<td>33%</td>
<td>1%</td>
</tr>
<tr>
<td>Categorised /Described</td>
<td>57%</td>
<td>17%</td>
</tr>
<tr>
<td>Corrected /Exemplified</td>
<td>8%</td>
<td>48%</td>
</tr>
<tr>
<td>Explained</td>
<td>1%</td>
<td>24%</td>
</tr>
<tr>
<td>Feeding forward</td>
<td>1%</td>
<td>10%</td>
</tr>
</tbody>
</table>

Comments on strengths are restricted almost entirely to levels 1 (indicated) and 2 (categorised or described), with only 8% at level 3 (exemplified) and a negligible proportion (1%) at levels 4 (explained) and 5 (feeding forward). In contrast almost half of the feedback on weaknesses was tagged at level 3 (corrected), with most of the remainder spread in decreasing order between levels 4 (explained), 2 (categorised) and 5 (feeding forward).
6 Discussion

6.1 Importance attached to feedback

Whilst there was an apparent discrepancy between students’ initial responses and those of their tutors (Table 3), both parties regarded all the options listed as important to students, and nearly all the students reported looking at the feedback they received.

In general terms the 73% figure for tutors who thought that nearly all, or a significant majority, of their students “looked at” all the feedback closely matched the student data. However only half the students reported doing follow-up work, i.e. the kind of engagement that would indicate that feedback was being used as a learning tool (See Furnborough & Truman, 2009 for similar findings from responses to assessment feedback in beginner distance language learners).

6.2 Perceptions of the purpose and nature of feedback

There was broad agreement on certain fundamentals, suggesting that assignment feedback was indeed performing a valuable function (Table 4). However, although feedback was perceived as helping students to see what they got wrong, there was rather less certainty among both tutors and students, that the latter were necessarily learning from this process. Nor did students feel that it helped them identify what they had got right, or how they could improve (Table 5). These are crucial areas of misalignment, with tutors much more confident than students that feedback was working well. Tutors’ emphasis on the importance for their students of “feeling supported” through feedback (Table 3) also raises the question of whether tutors and students have the same understanding of what it means to feel “supported.
6.3 Understanding feedback received

Most students were confident that the feedback provided was generally clear to them, and reported that when this was not the case they would ask their tutor for clarification. However the tutors’ perception was that in most such cases their students did not approach them. The latter is consistent with the findings of an earlier study (Furnborough, 2012) where confident learners indicated a strong preference for sorting out problems on their own if at all possible, and only referring to the tutor as a last resort. In the present study 29% of student respondents indicated that they would “get on with the module and hope that it will become clear later”. Perhaps for this reason tutors also appeared to have a lower opinion of students’ efforts to engage with the feedback and of their ability to remember it than was warranted by the students’ own responses (Table 6).

However, students’ attempts to interpret tutor feedback independently are not necessarily successful. A study where 10 students were asked to talk through the feedback given by their tutors on a written assignment (Fernández-Toro & Furnborough, 2014) showed that even confident students have a tendency to misinterpret feedback. This typically occurred when the level of scaffolding provided did not match the student’s ability to infer the precise lessons to be drawn from it. Student’s questions are the most effective means for a tutor to gauge that ability and adjust the level of feedback accordingly. Bloxham and Campbell (2010) experimented with cover sheets where nine students were asked to identify specific aspects of their work on which they wanted feedback from their tutors. They found that “students’ limited understanding of staff expectations and standards can limit their ability to initiate meaningful dialogue with their tutors”, and that the conversation needs to be “seeded” through the tutors’ initial feedback (Bloxham & Campbell, 2010, p. 299).
In the absence of this kind of communication misalignment between tutors’ assumptions and a students’ ability is likely to be greater. Therefore there is a strong case for assessment designs that proactively create opportunities for students to respond explicitly to the feedback (e.g. Barker & Pinard, 2014; O’Siochru, 2011), particularly in a distance learning environment where there is little or no face-to-face contact.

6.4 Giving/receiving feedback on strengths and weaknesses

The analysis of feedback revealed that tutors commented on weaknesses 2.4 times more than they did on strengths. This is not entirely surprising in the context of a foreign language programme where a high occurrence of linguistic errors is to be expected; indeed ‘taking risks’ by using new words and language structures is recognised as a key strategy for language learning. Yang and Carless (2013, p. 285) point out that “the academic discipline profoundly influences the feedback process.” An earlier study comparing students’ attitudes to feedback in two subject areas (Fernández-Toro et al., 2013), showed that Language students find feedback comments primarily useful in ‘seeing where [they] went wrong’, whereas this is not the case with Technology students (e.g. 42% mentions in Languages vs. 16% in Technology). Nevertheless, our survey data suggest that tutors may not fully appreciate the nature, extent and implications of this emphasis on weaknesses.

Feedback on weaknesses. Given the key role played by errors in language development it is not surprising that almost half (48%) the feedback on weaknesses was found to be error corrections (Level 3). Survey data confirmed that tutors were generally aware of this (61% reported correcting errors ‘at every opportunity’, 33%‘more than half of the time’), and students were generally satisfied (95%) with the amount of corrections provided. However students seemed to attach more importance to the tutor’s
responsibility for correcting and explaining errors than did tutors themselves (Table 7). This would suggest that tutors expect their students to take on more responsibility than students themselves are prepared to assume.

The proportion of explanations provided was much lower (24%), half as many as error corrections. Tutors were generally aware that they gave explanations less often than corrections, but the proportion reported (50% ‘at every opportunity’; 40% ‘more than half of the time’) does not reflect the extent of the actual difference found. Furthermore, 12% students reported that their tutors did not explain errors often enough. Tutors may of course have very sound pedagogical reasons to be selective in the errors they choose to explain. For example, explanations are not necessary where students have made a careless error which they are able to self-correct (Ros i Solé & Truman, 2005). However, even those students who believe that they understand their tutors’ annotations may in fact misinterpret them. Results from the talk-aloud study by Fernandez-Toro and Furnborough (2014) mentioned above suggest that this could be happening more often than students and tutors both assume. In gauging their students’ ability to interpret and follow up feedback effectively, tutors inevitably make assumptions, and there is little opportunity to test those assumptions in order to adjust the level of subsequent feedback. Such opportunities need to be proactively created through activities that elicit communication between tutors and students, for example guided activities where students go through the feedback and tell others what further corrections and explanations they figured out, and what strategies they will follow to improve. Tutors who took part in a follow-up trial of the talk-aloud approach (Fernández-Toro & Furnborough, 2014) with their own students providing ‘feedback on feedback’ reported that the experience had an impact on their feedback practices.

*Feedback on strengths.* A high proportion of tutors reported using “every opportunity”
to indicate when students did well (Table 9); this is consistent with their belief that support and encouragement are particularly important to students (Table 1). The fact that they provide on average 17 comments on strengths per student submission (Table 10) confirms their desire to indicate good performance.

However, the level of feedback provided within such comments reveals a considerable misalignment between the tutors’ assumptions, the students’ expectations, and the level of feedback that tutors actually give on strengths. Firstly, it could be argued that, if tutors truly did seize “every opportunity” to acknowledge strengths (as 72% claim to do), the analysis of tutor feedback should result in a well-balanced proportion between strengths and weaknesses, which clearly is not the case. Secondly, FACT analysis reveals that the highest proportion of feedback on strengths is in the form of indications (coded as level 1 – e.g. ticks or general acknowledgements such as ‘good’, ‘well done’, etc.), or mere descriptions of the strength in question (coded as level 2 – e.g. ‘you use a good range of vocabulary’). Specific examples, explanations and advice on how to build on strengths are much rarer. Self-reported data from tutors appear to confirm this pattern even though it is not as pronounced.

As for the students’ preferences, survey results clearly indicate that 22% would like their tutors to provide more descriptions of the strengths indicated in the feedback, and 17% would like more advice on how they could build on those strengths. While a large majority of students did consider the amount of feedback provided within each of these categories to be “just enough”, the discrepancies observed should not be ignored. Failing to provide sufficiently precise information about strengths can impact the students’ learning in at least two ways: Firstly, praise without evidence may be dismissed as a mere pat on the back, or at worst perceived as patronising. An experiment involving 115 physical education students (Vallerand & Reid, 1984)
confirmed that, while positive feedback did contribute to intrinsic motivation, these effects were mediated by perceived competence (Vallerand & Reid, 1984, p. 99). The best way to make students aware of their ability to perform a task is to provide evidence of their competence. The use of specific examples to substantiate praise is therefore important. Secondly, by not providing specific examples, explanations and feed-forward advice, tutors are depriving competent students of the opportunity to build on their strengths. While less competent students receive plenty of feedback at level 3 and above (especially corrections), a student who would like to raise her grade from 90 to 100 may get very little guidance as to how this could be achieved. Adding the missing layers of scaffolding would make feedback on strengths actually useable when students do well.

6.5 The FACT analysis method

These findings demonstrate the potential of FACT analysis as a research instrument to triangulate self-reported data from tutors and students. With an inter-rater reliability index of 0.97, the final version of the tool proved to be a valuable means to operationalising assignment feedback on the basis of its observable features. However, the complexity of the criteria used and the amount of training required in order to achieve that level of reliability make it unsuitable for other uses such as teacher training. For that purpose a simplified version of the FACT analysis criteria, presented in the form of a checklist, was produced as a self-reflection tool for tutors (see supplemental material: Handout – Reflecting on your own practice). Equivalent versions could also be developed for peer review and self-assessment purposes.

It must also be stressed that FACT analysis is merely a descriptive tool. Its unique strength is that it makes misalignments visible in terms of the nature and scaffolding provided in the feedback. However it does not in itself constitute a solution to misalignment. As a learning tool, feedback analysis is only useful as part of an array
of integrated strategies intended to create what Nicol (2010, p. 512) refers to as a “dialogic context”. Handley et al. (2011) stress the importance of integrating the feedback dialogue throughout the duration of the students’ learning journey, as the cumulative effect of previous experiences influences the student’s engagement with feedback on the current assignment, which will in turn influence engagement in future assessment. To implement this approach, Boud and Molloy (2013) shift the focus from assignments to the curriculum as a whole, and reposition feedback as “a key curriculum space for communicating, for knowing, for judging and for acting” (Boud & Molloy, 2013, p. 706).

Curriculum design plays a key role in providing a dialogic context for learning through feedback. However the structure set up for the feedback process must be flexible enough to allow meaningful dialogue to occur (Moore, 2012). It would therefore be ill-advised to use FACT analysis in a prescriptive manner, for example by instructing all tutors to correct every single error found in a particular assignment, given that certain students will be able to self-correct and others not.

As previously discussed, tutors are constantly making judgements as to a learner’s ability to contribute to the process of bridging the gap between actual and desired performance, and the fine adjustments they make on the basis of such judgements are key to the feedback’s effectiveness. However, judgements can be difficult to make, especially in a distance learning environment. In that respect, FACT analysis has proved to be particularly valuable as a means of gauging the level of feedback that individual students require. Fernández-Toro and Furnborough (2014) invited 10 students to talk through the written feedback received from their tutors, which had been FACT analysed. The depth of each feedback comment was then compared to the depth of the student’s comment about the feedback in question, thus
revealing any outstanding gaps. In some cases students were able to provide a deeper level of commentary (for example by explaining a correction that their tutor had made), but in other cases they were unable to do so (e.g. failing to notice a correction, or inferring an explanation that was incorrect). This approach, referred to as ‘feedback on feedback’ has since been developed as a learning activity aiming to give students an opportunity to drive the feedback process through a two-way dialogue with their tutors.

7 Conclusions

The main contribution of this project derives from the triangulation of self-reported data from tutor and student surveys against observational data from feedback analysis. This made it possible to evaluate the extent to which the nature and quantity of scaffolding provided in assignment feedback was in line with the students’ and tutor’s assumptions and expectations.

In some instances feedback analysis provided confirmation of the claims made by survey participants, while in others it revealed clear discrepancies between their beliefs and the actual feedback provided. The two most striking areas of misalignment related to feedback on strengths (helping students to appreciate what they got right), and feed-forward (giving practical suggestions for improvement in future). These areas are directly relevant to the seven principles of good feedback practice that Nicol and Macfarlane-Dick (2006) identify as necessary to develop self-regulation, especially to principle 1: “helps clarify what good performance is” and principle 6: “provides opportunities to close the gap between current and desired performance”.

As for feedback on weaknesses, FACT analysis revealed that tutors made extensive use of corrections while their use of other types of feedback, including explanation of errors (a function that students valued highly), was far less frequent.
Again, this approach may need to be reconsidered in line with principle 6 quoted above to help students “close the gap” through improved scaffolding.

However it would be over-simplistic to recommend feedback that includes all possible levels of scaffolding as the most effective. There is no ideal set of features of the feedback itself that would make it particularly effective. Early feedback research aiming to identify such ideal features through feedback analysis has since been criticised as ‘atomistic’ and lacking in validity by Handley et al. (2011, p. 546), who argue that the emphasis should be not on the feedback itself, but on “the relationship between feedback and student learning” (Handley et al., 2011, p. 546). In that sense, the present contribution is in line with Nicol and Macfarlane-Dick’s recommendations, as it “provides information to teachers that can be used to help shape teaching” (Nicol & Macfarlane-Dick, 2006, p. 205, Principle 7).

Provided that a suitably dialogic context has been created through curriculum design, FACT analysis (or some elements of it) can also be harnessed as a means of guiding students in constructing feedback for their peers. Our survey results revealed some discrepancies between the students’ and tutors’ respective beliefs about feedback ownership, with tutors expecting their students to engage more proactively than they were able or willing to do. Recent work on peer review and self-assessment has placed the emphasis on developing the students’ ability to make critical judgements about the quality of their own work and the work of their peers (Boud, Lawson, & Thompson, 2013; Everhard, 2015a; Nicol, 2011; Tai, Ajjawi, Boud, Dawson, & Panadero, 2017). Everard characterises the nature of formative assessment in terms of “critical judgement, using agreed transparent criteria; importance placed on feedback from multiple sources”, and conceives the relationship between learners and tutors as one of interdependence (Everhard, 2015b, p. 27). Feedback analysis is therefore an ideal tool
in formative assessment, because it can be used to fine-tune the feedback that tutors, learners themselves and their peers generate around a set of agreed criteria. If the learning dialogue is successful, learners gradually will move from interdependence to autonomy as they take ownership of the process and learn to utilise the tools at their disposal to meet their own goals. Assessment (and implicitly, feedback) becomes sustainable when it “meets the needs of the present without compromising the ability of students to meet their own future learning needs” (Boud, 2000, p. 151). Within a sustainable model of assessment, students are able to draw upon the lessons learnt through feedback analysis to monitor, plan and regulate their own learning.

The fundamental challenge is for educational institutions to develop and implement sustainable assessment models where feedback dialogue is allowed and encouraged to flourish.

Acknowledgments

This project was funded by the JISC (Joint Information Systems Committee), a UK public body aiming to provide leadership in the use of ICT in learning, teaching, research and administration. We also wish to thank Bea de los Arcos for her invaluable work in coding the FACT data, and Daniel Richards and Paul Grayson for their help with statistical analysis. Finally we would like to thank all the students and tutors who agreed to take part in the project.

References


