Facilitating Timely Feedback in the Biomedical Sciences

Kay Colthorpe[a], Shaohong Liang[b] and Kirsten Zimbardi[a]

Corresponding author: Kay Colthorpe (k.colthorpe@uq.edu.au)
[a]School of Biomedical Sciences, The University of Queensland, Brisbane QLD 4072, Australia
[b]School of Population Health, The University of Queensland, Brisbane QLD 4006, Australia

Keywords: feedback; allied health science; biomedical science

International Journal of Innovation in Science and Mathematics Education, *21*(3), 60-74, 2013

Abstract

Feedback is one of the most influential factors on student learning gains (Hattie & Timperley 2007). However, studies also show that when students do receive feedback it is often too brief, too broadly stated, and is often misinterpreted by students (Nicol & Macfarlane-Dick, 2006; Stern & Solomon, 2006). Furthermore, evaluating the actual extent to which students engage with and utilise feedback is difficult.

This study evaluated a method of providing detailed, specific and timely feedback to allied health science students studying biomedical sciences in large class settings at a higher education institution in Australia. We investigated the extent and quality of feedback provided through analysis of annotated drafts, and examined how students interpreted and used the feedback received, by identifying how student work was modified in response to each item of feedback. This study has demonstrated that for feedback to elicit positive changes in student writing it must be specific, detailed and directed. The results indicate that the majority of the feedback given in the assignments analysed had a positive effect on subsequent student work, but also highlights that student responses to feedback can differ based on the type of feedback that is given.

Introduction

Assessment reforms in higher education recognise the need for effective feedback that is timely, informative, and encourages positive attitudes towards future learning amongst students (Boud et al, 2010). It is well established that feedback to students is one of the educational practices with some of the largest positive impacts on student learning (Hattie & Timperley, 2007), However, studies also show that when students do receive feedback it is often too brief, and too broadly stated, and is often misinterpreted by students (Nicol & Macfarlane-Dick, 2006; Stern & Solomon, 2006). Furthermore, evaluating the actual extent to which students engage with and utilise feedback is difficult.

Generally, academics recognise the value of formative feedback but often lack information on what are the most effective feedback practices, resulting in the provision of feedback which is inefficient, creates confusion and commonly communicates incorrect or unrealistic expectations to students (Stern & Solomon, 2006; Underwood, 2008). Overall, academics typically do not give positive feedback, address students' weaknesses and strengths, or provide comments which encourage critical thinking in students, but rather focus on ‘surface-level’ feedback such as correcting simple technical writing errors (Stern & Solomon, 2006; Underwood, 2008; Turnitin, 2012).

Although research has extensively characterised the types of feedback academics provide on student papers (Connors & Lunsford, 1993, Stern & Solomon, 2006, Turnitin, 2012), much of this work has been limited to categorising and analysing feedback provided on summative assessment items, and does not determine the impact of formative feedback on subsequent student work. The study reported here investigated the relationship between the types of individual items of feedback provided by academics on formative submissions and the type and quality of changes students made to the subsequent summative submission in response to that feedback. This analysis involved modifying the original simple category system of Stern and Solomon (2006) into a more complex multi-level system which enabled a more detailed analysis of the impact of feedback, both in terms of the extent and the effect of change elicited by the feedback. The implications of these findings for the development of key criteria for effective feedback practice are then discussed.

Methods

Undergraduate students (n=220) in the Bachelor of Physiotherapy (n= 112), Bachelor of Speech Pathology (n=95) or Masters of Speech Pathology (n=13) studying a second level Physiology course at the University of Queensland in Australia, were given the option to submit a draft of their major written assignment of semester 1, 2012. Assignments consisted of a 1500 word scientific literature review covering one of four available topics drawn from the lecture modules. Topics included rehydration strategy; the effect of *Botulinum* toxin; the impact of mould on respiratory function; and mechanisms, treatment and prevention of deep vein thrombosis. Students had access to both a guideline for the assessment task and a marking criteria sheet prior to submission of the draft assignment. Of these students, approximately 69% chose to submit a draft. Each draft was provided with individualised handwritten feedback, placed *in situ* on a hard copy of each student's work by the four academics who lectured within the course, as part of their normal contribution to this course. No marks were provided for the drafts. Two weeks after draft submission, the drafts were collected by the students, who were then required to submit their final assignment a further two weeks later. Marking of the final assignments was criteria-based, and all students received a marked criteria sheet, with a final grade. No feedback was given on any of the final assignments, regardless of whether or not that student had submitted a draft.

Four complete drafts from each of the four topics (16 in total) were randomly selected for extensive in-depth analysis of the feedback given, and the corresponding final assignment was examined to determine the impact of that feedback using the method described below. As each topic was the responsibility of a different contributing academic, there was variation in the type and extent feedback represented within this selection. In addition, the final assignment grade achieved by these students ranged from 57% to 97%, so they were representative of a broad academic standard.

The feedback annotations on each assignment were coded within *NVivo* 10^{TM} software (QSR International, MA, USA) based on the 23 categories developed by Stern and Solomon (2006; Table 1), who had extended the earlier, simpler category system of Connors and Lumsford (1993). Coding was performed by one of the authors (SL) who was not involved in the teaching or marking of the assessment for this course. In the context of this study, one category 'Rubric or grading sheet' was considered not applicable as no grading took place on the draft assignments, and the category described as 'Technical writing style' was interpreted as 'Scientific writing style'.

Table 1: Categories of feedback: Details of the full list of possible categories against which feedback annotations were coded, with descriptors and examples of each category (Adapted from Stern & Solomon, 2006), and designated numbers. Category 21 was considered not applicable in this study, and category 14 was modified from 'technical' to 'scientific' writing style.

No.	Category	Examples from Stern and Solomon (2006)
1	Overall quality	need work, good writing style, great paper
2	Paper structure and organisation	poorly organized, hard to follow, good flow, well integrated lit review
3	Creativity	creative!
4	Voice	eliminate passive voice, write in action, 1st or 2rd person
5	Quality of specific thoughts and claims	No!, this is an extreme claim, good reasoning, interesting idea
6	Procedure and technique	incorrect measurement, wrong tool, good technique, nice choice of method
7	Support or evidence for claims	insufficient data or proof, give an example, good data, great support
8	Request for content clarification	what does this mean?, what is the point?, define, why?
9	Paragraph and sentence structure or style	paraphrase rather than quote, repetitive, effective summary, good paragraph
10	Word choice or phrasing	awk, wordy, wrote in new word and cross out one of students, reword
11	Missing words and pieces	add sentence, word needed, wrote in word or phase
12	Grammar or punctuation	noun or verb agreement, deleted commas, fixed grammar, frag, run-on, verb tense
13	Spelling or typo	Spelling mistakes and typo
14	Scientific writing style	corrected format, corrected citation style, location of page break
15	References or citations	need citation, source?, cite your source, good cite, good sources
16	Invitations to discuss paper	if you need further clarification come see me, see me
17	Personal expressions and advice	Wow!, unbelievable, i found the same things
18	Scholarly advice	refer to chapter 2 in book, for further study see
19	Road maps	see above notes, ditto, same as above
20	Tracking marks	underline with no comments, check marks, "Late"
21	Rubric or grading sheet	"Grades and Criteria Scores"
22	Unidentifiable	anything illegible
23	Others	(anything that does not fit within any of the categories)

The change(s) made in the final assignment in response to each annotation were scored by determining the extent to which the student changed their work on a scale ranging from none to major (Table 2). Change(s) were then further qualified by identifying the effect of change, which ranged from positive to negative, with positive representing a substantive improvement in the academic standard, negative representing a substantive reduction in academic standard, and neutral representing either no change or a change which neither improved nor worsened the quality of that section of the student's assignment (Table 2). In addition, the final assignments were examined for correlations between the final grades awarded and both the amount of feedback provided, and the extent and effect of changes in response to that feedback, using the statistical analysis software Prism™ (GraphPad Software, Inc, CA, USA). This detailed analysis, based on comparison between subsequent submissions from individual students over time, allowed the elucidation of the relationships between individual items of feedback and student response.

Table 2: Multi-level analysis: Feedback annotations from the draft submissions were individually categorised using the 23 categories of Stern and Solomon (2006; Level A), then the response to each annotation in the corresponding final assignment was quantified (Level B) and qualified (Level C).

Level A	**23 Categories of Feedback (Table 1)** Developed by Stern and Solomon (2006), items of feedback on draft assignments are classified into 23 categories.
Level B	**Scale of change** Extent of change made in the final assignment in response to each feedback annotation, classified into Major, Moderate, Minor and None.
Level C	**Effect of change** Classification of the quality of change made in the final assignment in response to each feedback annotation, ranging from Positive, Slightly positive, Neutral, Slightly Negative or Negative.

Student performance on summative assessment was collated for both the final assignment submission and the end of semester examination, which contributed 25% and 60% toward the course grade respectively. Summative results for students who did, and those who did not, submit a draft of their assignment were compared using a t-test (Microsoft Excel, Microsoft Corporation, WA, USA). Results were considered significant in $p<0.05$. This study has been approved by the University of Queensland Human Experimentation Ethical Review Committee.

Results

Approximately 69% (n=156) of the student cohort chose to submit a draft of their assignment for feedback. The vast majority of the drafts were close to full length, with a small number being partial drafts or outlines. After receiving feedback, 100% of the submitted drafts were collected by the students.

Impact of feedback

A total of 140 feedback annotations were provided on the 16 draft assignments that had been randomly selected for in-depth analysis. Examples from 16 of the 22 applicable categories described by Stern and Solomon (2006) were present, with the vast majority (90.8%) falling within 10 categories (Table 3). Feedback ranged from simple symbols or single words to specific, detailed paragraphs of advice and expectations (Table 3). Some feedback annotations were a combination of categories, the most common example of this was where parentheses were used to indicate the area to which a comment referred (such as the examples in Table 3, Categories 4, 5, and 9), in these cases the annotation was only categorised once, based on the associated comment. There were no significant correlations identified between the amount of feedback students received and their final assignment grade, nor were there any correlations found between the extent or effect of the changes the students made in response to feedback and their final grade.

Quantitative and qualitative analysis of the changes made by the students on their final submissions in response to annotations showed that there were clear differences in the responses students made to the different categories of feedback (Table 4). For example, annotations from the most common category, ‘Request for content clarification’, represented 25.7% of all annotations (Table 3) and elicited both the most major changes and had the most

positive effect on student work (Table 4). Annotations which were categorised as 'Unidentifiable' represented 12.1% of the feedback provided (Table 3), but had only moderate effect if any, and majority of student responses were of little or no academic benefit, or were detrimental (Table 4).

The extent of student responses to feedback varied from none to major, and the effect varied from positive to negative (Tables 5-7). Variations in extent and/or effect of responses were apparent within all individual categories of feedback present, for example in category 8 'Request for content clarification' (Table 5). Overall, 77.9% of all the feedback provided elicited positive or slightly positive effects in student work.

Summative performance

Student performance on the final submission of the assignment was significantly higher for students who submitted a draft assignment (79.9±0.82; mean±SEM) compared to those who had not (76.85±1.40; $p<0.05$). However, students who had submitted a draft assignment also had significantly higher scores on the end of semester examination (67.5±1.10) than those who had not (61.4±2.46; $p<0.05$; Figure 1). Within the subset of students whose work was subjected to detailed analysis, there were no significant correlations between the number or type of feedback items students received and their final assignment mark, nor were there any correlations found between either the extent or effect of the changes the students made in response to feedback and their final mark.

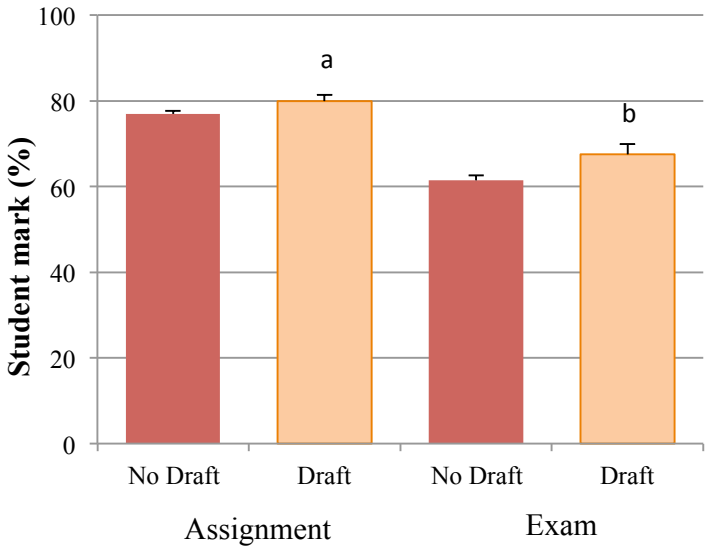
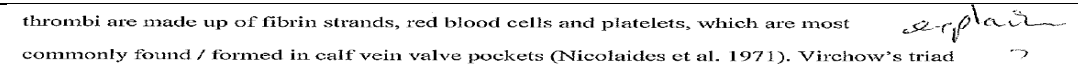
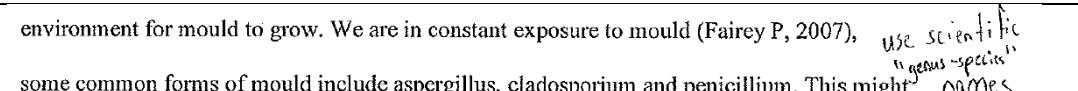
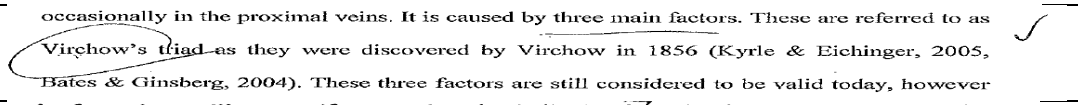
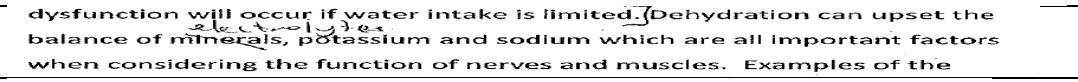
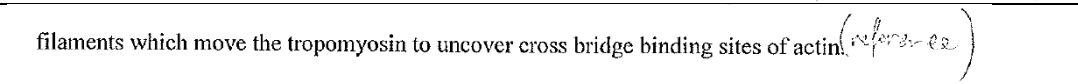
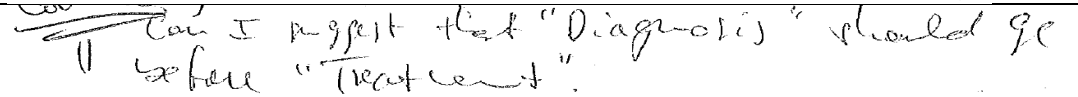
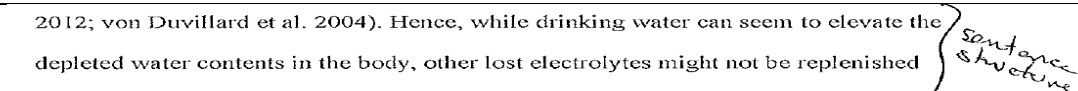
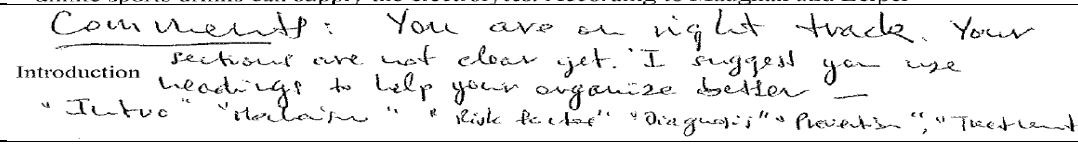


Figure 1: Student summative performance on final assignment submission and end of semester examination for those who did not submit a draft (n=64) or those who did (n=156) as percentage of available marks (Mean +/- SEM). A, significantly higher assignment mark than students who did not submit a draft; b, significantly higher examination mark than students who did not submit a draft, $p<0.05$

Table 3: Categories of feedback: The frequency of appearance of feedback annotations in each category, in descending order of usage, based on the 140 annotations identified from the 16 draft assignments. A representative example of a handwritten annotation is shown for each category. Category numbers correspond to those assigned to the categories of Stern and Solomon (2006; Table 1). Categories 3, 6, 16, 18, 20, 21 and 23 with 0% frequency have been omitted.

No.	Category	Overall frequency of feedback (%)	Examples retrieved from drafts
8	Request for content clarification	25.7	
14	Scientific writing style	12.9	
22	Unidentifiable	12.1	
10	Word choice or phrasing	8.6	
15	References or citations	7.9	
17	Personal expressions and advice	7.1	
9	Paragraph and sentence structure or style	5.0	
2	Paper structure and organization	4.3	

12	Grammar or punctuation	3.6	Injection of BoNT-A is relatively safe and comes with few serious side effects including swelling, bruising and flu-like symptoms (Klein 2004). The paralysis effects of
1	Overall quality	3.6	– A very good essay with lots of good information.
19	Road maps	2.9	Start with introductory sentence Stress Exercise → Heat → sweat to evaporate & cool. production Introduction Strenuous exercise usually results in large amounts of wa
13	Spelling or typo	2.1	and mould, there was no significant difference between between the exposure groups on the basis of sick leave or doctor visits.
7	Support or evidence for claims	1.4	– use facts + figures to back your statement eg. what is the evidence of DUT ?
11	Missing words and pieces	1.4	& electolyte amounts of water loss especially via
5	Quality of specific thoughts and claims	0.7	lower sodium concentration in their water intake may suggest that there were greater water movements into the intracellular fluid (ICF). Such water movement is called osmosis, where the water diffuses from an area of lower osmolarity to an area of higher osmolarity through a semi permeable membrane (Sherwood 2010). A higher rate of osmosis might have taken place because of the greater osmotic gradient between the ECF and ICF (Sherwood 2010). However, the amount of water that has entered the cells for rehydration might not be significant to offset the water loss from the ECF as urine really a repeat of what you say in previous paragraph
4	Voice	0.7	excretion through urine and feces, respiration and evaporation from skin (Hoffman, 2002). However, with exercise, water loss is accelerated by the critical thermoregulatory mechanism of sweat production (Dubnov-Raz et al., 2011). The Great opening paragraph

Table 4: Quantity and quality of student responses to feedback: 60% of the 140 feedback annotations provided by markers across the 16 student assignments fell within four categories. The scale and effect of the changes students made in response to those annotations differed between categories, with the most common category 'request for content clarification' eliciting the largest, positive responses.

(A) Category	(B) Scale of change	(C) Effect of change					
		No. of Items	Positive	Slightly positive	Neutral	Slightly negative	Negative
8. Request for content clarification (25.7%)	Major	14	14				
	Moderate	8	7	1			
	Minor	6	3	1		1	1
	None	8				4	4
14. Scientific writing style (12.9%)	Major	0					
	Moderate	2	2				
	Minor	16	16				
	None	0					
22. Unidentifiable or ambiguous (includes lines, symbols & question marks; 12.1%)	Major	0					
	Moderate	0					
	Minor	7	1	3	1	2	
	None	10			8		2
10. Word choice or phrasing (8.6%)	Major	0					
	Moderate	3	2	1			
	Minor	9	4	5			
	None	0					
Total		83	49	11	9	7	7

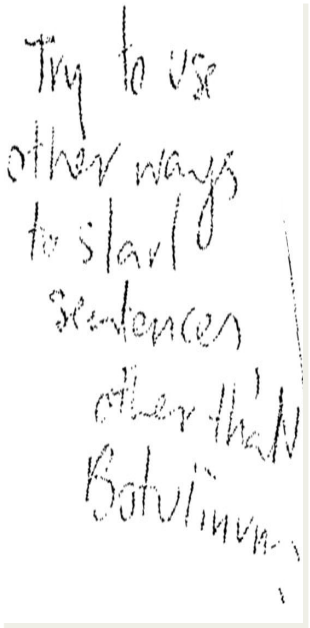
Table 5: Content clarification and the impact of descriptive feedback: Similar items of feedback on drafts assignments with varying levels of description (left column), and their corresponding response from students in final assignments (right column). Examples lacking description elicit (a) no response or (b) a minor, negative response. Only the most descriptive example (c) elicits a major, positive response.

Draft and feedback given	**Corresponding section in final report and changes made**
(a) thrombi are made up of fibrin strands, red blood cells and platelets, which are most commonly found / formed in calf vein valve pockets (Nicolaides et al. 1971). Virchow's triad is made up of three influences on the formation of a thrombus or clot. They are endothelial injury, venous stasis and hypercoagulabiltiy (Kumar et al. 2007). explain ?	**No change** strands, red blood cells and platelets, which are most commonly formed in calf vein valve pockets (Nicolaides et al. 1971). Virchow's triad is made up of three influences on the formation of a thrombus or clot. They are endothelial injury, venous stasis and hypercoagulabiltiy (Kumar et al. 2007).
(b) Prevention measures are well established in reducing the risk of DVT, a large variety of options are available to the general population, these include; flite tabs (pinokinase and pycnogenol), leg exercises and compression stockings. explain	**Minor change, negative (red)** Prevention measures are well established in reducing the risk of DVT. A large variety of options exist including: Flite Tabs (commercial medication), leg exercises and compression stockings.
(c) Botulinum, a protein consisting of a heavy and light polypeptide chain, causes flaccid muscle paralysis. It prevents muscle contractions by acting on autonomic or somatic motor neurons; inhibiting acetylcholine (Ach) release Explain the subtypes of botox A-G	**Major change, positive (red)** Botulinum, a protein consisting of a heavy and light polypeptide chain held by a disulphide bond, causes flaccid muscle paralysis. There are seven subtypes – A to G – that inhibit acetylcholine (Ach) release (Nigam & Nigam 2010) from cholinergic neurons with varying potencies by targeting different intracellular proteins (Eleopra et al. 2006). Such neurons include neuromuscular junctions (NMJ) of autonomic or somatic neurons, free nerve endings of autonomic ganglia, postganglionic parasympathetic and sympathetic neurons (Sellin 1985).

Table 6. Examples of feedback that failed to elicit student responses: Feedback on draft assignments from category 22 'Unidentifiable' (a) & (b) and category 9 'Request for content clarification' (c) & (d) with differing levels of description (left column) and their corresponding sections of text from final assignments (right column). Each of these examples failed to elicit any response from the student in their final assignment.

Draft	Student response in final report
(a) mould-associated and what is damp-associated and what the combined effects are on the	**No change** factors. It becomes difficult to differentiate between what are mould associated symptoms and what are damp associated symptoms.
(b) location (Launius and Graham,1998). This disturbance is caused by the combination of three main factors identified by Virchow, which is known as the Virchow's triad that affects the individuals differently (Dipaola, 2008).	**No change** hemostasis in an abnormal location (Launius & Graham,1998). This disturbance is caused by the combination of three main factors identified by Virchow, which is known as the Virchow's triad that affects the individuals differently (Dipaola, 2008).
(c) It is possible for the botulinum toxin to spread to nearby muscles, causing unwanted weakness or paralysis (Nigam & Nigam, 2010). The botulinum toxin molecules can become displaced and travel, due to applied pressure at the site or increased blood flow (Nigam & Nigam, 2010). This can also occur by way of errors of injection placement made by the clinician (Levy & Lowenthal, 2010). Botulinum toxin treatments injected into the neck can cause temporary dysphagia (difficulty swallowing) (Anton, 2011), which comes with the risk	**No change** It is possible for the botulinum toxin to spread to nearby muscles, causing unwanted weakness or paralysis (Nigam & Nigam, 2010). The botulinum toxin molecules can become dislodged and travel, due to applied pressure at the site or increased blood flow (Nigam & Nigam, 2010). This can also occur by way of errors of injection placement made by the clinician (Levy & Lowenthal, 2010). Botulinum toxin treatments injected into the neck can cause temporary dysphagia (difficulty swallowing) (Anton, 2011), which comes with the risk
(d) 2010); and blood-thinning medications (Nigam & Nigam, 2010). Despite the dangers and side effects of botulinum toxin, it is a very important advance in medicine. Medical professionals need to be fully aware of the seriousness of the injection and ensure full precautions are taken in order to avoid severe reactions. Future research should target the question of whether botulinum toxin injections can be causal in the serious disease botulism. It would also be of use to develop a strategy to identify patients, pre-	**No change** hydroxychloroquine (Nigam & Nigam, 2010); calcium channel blockers (Nigam & Nigam, 2010); and blood-thinning medications (Nigam & Nigam, 2010). Despite the dangers and side effects of botulinum toxin, it is a very important advance in medical technology. Medical professionals need to be fully aware of the seriousness of the injection and ensure full precautions are taken in order to avoid severe reactions. Future research should target the question of whether botulinum toxin injections can be causal in the serious disease botulism. It would also be of use to develop a strategy to identify patients,

Table 7. Examples of feedback that address global issues: Feedback on draft assignments that addressed global issues were most often placed at either the end or beginning of the whole document or adjacent to the major section to which they referred. An example annotation (left column) commenting on the student's repetition of the same opening word for multiple paragraphs, with three of ten examples from draft assignment shown (centre column), elicited a global response (right column) in the final submission.

Feedback given	Draft	Final: global changes (red)
	**(a)** 3. **Skeletal muscle activity and movement** Botulinum can be used to treat movement disorders, for instance dystonias and spasticity by targeting two areas – regulating skeletal muscle	**Skeletal muscle activity and movement** The effect of botulinum on skeletal muscle and muscle spindle activity is efficacious in treating movement disorders.
	(b) Botulinum reduces activity in overactive muscles by preventing Ach transmission across the motor end plate (Rosales, Arimura, Takenaga & Osame 1996). Muscle overactivity, typical in dystonias, is caused by abnormal	Muscle overactivity – typical in dystonias – can be reduced by preventing Ach transmission across motor end plates (Rosales et al. 1996). Sustained, unwanted co-contractions are caused by abnormal sharing of pre-synaptic inputs to agonistic and
	(c) Botulinum can be used to promote delayed healing of ischaemic ulcers as a result of hypertonic muscle activity. For instance, hypertonic internal sphincters delay chronic anal fissure resolution because of reduced blood	By reducing hypertonic smooth muscle activity, botulinum can promote delayed ischaemic ulcer healing. For instance, hypertonic internal sphincters delay chronic anal fissure resolution because of reduced blood perfusion to the

Discussion

This study has demonstrated that for feedback to elicit positive changes in student writing it must be specific, detailed and directed. The results indicate that the majority of the feedback given in the assignments analysed had a positive effect on subsequent student work, but also highlights that student responses to feedback can differ based on the type of feedback that is given (Table 4). Furthermore, these results provide empirical evidence for the types of feedback which elicit both positive and negative changes in student writing. The assessment design used in this study conformed to a number of the conditions which support student learning, as outlined by Gibbs and Simpson (2004). These conditions included that the feedback was timely, it was received by the students, as all students collected their drafts, and it was appropriate to the purpose of the task, particularly for these allied health science students who were writing in a genre that was relatively unfamiliar to them.

A clear example of the importance of description to elicit positive student responses can be seen where requests for content clarification occur. This category of feedback was the one most frequently found in this study, but examples of it ranged from detailed, lengthy annotations, to just single words. For example, in Table 5a and 5b, where "explain" or "explain?" were used (with associated underlining of relevant text), students responded with either no change or a small but negative change, whereas when more detail was included in "explain the subtypes of botox A-G", the student responded with a major, positive change (Table 5c). Despite these items of feedback being very similar, as all are requests for content clarification, only the more descriptive annotation elicited a desirable response. Dinneen (2010) suggests that descriptive feedback is necessary to help students gain better perceptions of a given task, the examples given here highlight the value of even a small addition to description for guiding the student to respond to the specific task in the appropriate way.

The third most common category of feedback found in this study was 'Unidentifiable' representing 12.1% of all feedback given (Table 3). These indicative marks, such as question marks or underlined or circled sections, where not associated with comments, did not elicit as strong positive responses. Indeed such feedback caused only minor or no responses from the students, the majority of which were of either neutral or negative effect (Table 4). The combination of the common usage of these types of indicative marks, and the finding that they are of little value as feedback, has important implications for the provision of effective feedback, and suggests that, while these types of annotations are quick and easy to add to student work, they are essentially worthless. The exception to this is where these types of marks are used in combination with comments, for example the use of parentheses or arrows, to direct the comment to a specific area of student work (Table 3, category 5; Table 3, category 9). In these cases, they add value by enabling the feedback providers to highlight the area of student work to which a comment specifically refers.

Indeed, another key criterion for effective feedback identified by this study is the positioning of the feedback. Specific guidance such as arrows and parentheses, or placing of feedback in a position that clearly links that feedback and the area of their work to which it refers, appear to aid the students understanding of the task given to them. Where this type of guidance was absent, and there was no clear indication of where the feedback is targeted, there appeared to be a reduced likelihood of the student responding in the desired way. An example of this can be seen in Table 6c, where a poorly positioned request for content clarification failed to elicit any changes in the final report. Another example of the value of positioning is in the addressing of global issues, such as overall writing style or quality. Ideally, global comments

should be positioned at the beginning or end of the page, section or report. The value of this is demonstrated by the example of a well-positioned feedback annotation (Table 7, left column), that led to global changes in the final report (Table 7, right column).

There were, however, examples within this study of feedback that, despite meeting the requirements of being specific, detailed and directed, failed to elicit responses from the students. For example, the feedback "Expand upon why some of these are contraindicated with reference to molecular mechanisms" (Table 6d) failed to elicit any response from the student. Clearly, in this case, an appropriate response to this feedback would have required considerable further reading and editing on behalf of the student. While there is insufficient evidence available from this study to fully elucidate the cause of this failure, it may reflect that the complexity of task required to respond to feedback plays a part in the extent and nature of the students' response. It is feasible that in such cases, the task may have been either beyond the capabilities of the student, or required more than the available time to complete, and consequently the student failed to respond. Perhaps in these cases, the onus is on the feedback giver to either have realistic expectations of student capabilities, or to provide further guidance to encourage students to tackle these more complex issues. Ideally, further information should be sought to elucidate why feedback failed to elicit student responses on these occasions, despite meeting the desirable characteristics, and this should be a focus of future research in this field.

This study was designed as an extension of the earlier work of Stern and Solomon (2006), who developed their category system from the much simpler version of Connors and Lunsford (1993). However, this study differs from those earlier studies in both sample size and selection strategy, and was a more detailed study, with the direct comparison of draft and final versions of the assignment allowing detailed analysis of the direct impact of feedback on subsequent student work. While direct comparison to results from Connors and Lunsford (1993) is difficult, as feedback was categorised somewhat differently, direct comparisons can be made with all categories from Stern and Solomon (2006). The findings from this study show similarities and some notable differences in feedback category distribution when compared to that study.

There are similarities in the proportion of assignments in which feedback on paper structure, spelling and word choice appear, when compared to Stern and Solomon (2006), but feedback on grammar and sentence structure appears on far fewer papers in the current study. Most notably though, the frequency of feedback on 'writing style' and 'references and citations' in the current study differed markedly from the Stern and Solomon (2006) study. Their study had a similar frequency of feedback in both these categories of approximately 10%, whereas in the current study feedback from these categories appeared on nearly half of all assignments. In addition, in the current study 'requests for content clarification' appeared on twice as many assignments as in the Stern and Solomon (2006) study. Together, these three categories represent 46.5% of all feedback annotations in this study (Table 3). These clear differences may reflect the characteristics of the type of assignment analysed in this study, as a scientific literature review it was inherent that the students addressed the content of the topics in a detailed way; consequently there was a high representation of feedback requesting clarification of content. Further, feedback on writing was more specifically directed at scientific writing style, rather than writing in general, and on the use and citation of references. In the earlier study, a full range of written assignment types were deliberately chosen for inclusion in their analysis, potentially reducing the predominance of these categories.

While a large proportion of the feedback given was associated with positive improvements in the quality of student work, the higher marks achieved by students who submitted drafts and received feedback cannot be solely attributed to that feedback. Students who submitted drafts also performed significantly better on the end of semester exam (Figure 1). It is unlikely that performance on the assignment *per se* influenced examination performance, as the material covered within the assignment topics was of considerably greater depth and far less breadth than the examination material. Therefore, it is likely that students who took the option to submit a draft are those with the stronger academic focus and commitment and, consequently, would be expected to perform better. Given the optional nature of draft submission, and the lack of grading of the drafts, we cannot definitively conclude that the feedback improved overall student marks. However, the design of this study was such that we were able to identify the impact of individual items of feedback, in both the extent and quality of the student responses to each item, demonstrating that the majority of the feedback provided elicited positive responses, and allowing us to draw important conclusions on the nature and type of feedback that elicits the largest, most positive responses. Therefore, the important implications to take from this study are not that all instructors should provide feedback on drafts, but rather that in any instance of providing feedback, feedback which is specific, detailed and directed is more likely to produce improvements in student writing.

Conclusion

As shown in our examples, feedback with more description elicited stronger positive responses than those with no description, and indicative marks such as underlined or circled sections, or question marks, when not associated with comments, did not elicit strong positive responses. In addition, our examples show that positioning of feedback is important, both to link feedback to specific areas of work, or to create global responses. With these findings in mind, we have identified four key criteria to giving effective feedback:

- Students respond more positively to feedback with words than symbols alone.
- Annotations should be simple yet descriptive to achieve the desired responses.
- Feedback should be specific and directed at the region of work it refers to.
- In order to address global issues, feedback should be placed separately at the end or beginning of either the section or the document.

References

Boud, D. (2010) Assessment 2020: Seven propositions for assessment reform in higher education. Sydney: *Australian Learning and Teaching Council*. Viewed 20th February, 2013.

Connors, S. J., & Lunsford, A. A. (1993). Teachers' rhetorical comments on student papers. *College Composition and Communication, 44*, 200-223.

Dinneen, C. D. (2010). *Using Descriptive Feedback in an Assessment as Learning Context for Constructing the Way Forward*. (Doctoral dissertation, University of Melbourne, Department of Learning and Educational Development, Australia). Retrieved from http://repository.unimelb.edu.au/10187/9569

Gibbs, G., & Simpson, C. (2004). Conditions under which assessment supports students' learning. *Learning and teaching in higher education*, *1*(1), 3-31.

Hattie, J., & Timperley H. (2007). The Power of Feedback. *Review of Educational Research*, *77*(1), 81-112.

Nicol, D. J. & Macfarlane-Dick, D. (2006) Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. *Studies in Higher Education 31*(2): 199-218.

Stern, L. A., & Solomon, A. (2006). Effective faculty feedback: The road less travelled. *Assessing Writing, 11*, 22-41.

Turnitin. (2012) From the Margins: What Instructors Say on Student Papers - 30 Million Insights into Improving Instructor Feedback. *Turnitin White Paper*. Retrieved from http://www.turnitin.com/en_us/resources/white-papers

Underwood, J. S. (2008). Effective feedback: guidelines for improving performance. *Proceedings of the 8th International Conference for the Learning Sciences - Volume 2*. 415-422. International Society of the Learning Sciences, Utrecht, The Netherlands.