

## **"What's the point: Do annotations on students' work promote self-regulation?"**

**Cathrine Derham, Kieran Balloo, Maria Norman, and Naomi E. Winstone**

**University of Surrey**

### **Abstract**

As one component of the feedback process, annotations on student work should focus upon providing explanations and guidance, which encourage students to use the comments to develop their abilities to act as self-regulated learners; thus promoting what Carless (2015) refers to as the new paradigm of feedback practices. This is contrary to the old paradigm in which annotations merely serve to transfer information, characterised by evaluative statements and corrective advice. It is argued that it is not only the content of the message, but also the language used, which has an impact upon the sustainability of this form of feedback practice. The current study reports on an analysis of annotations in the form of 1760 in-text comments added by markers to 52 summative essays. Findings indicated that the majority of comments were directed at the level of task performance rather than relating to the process (i.e. giving students advice about their future work and regulation of their actions). Additionally, there were positive correlations found between grades and words expressing a positive emotional tone, as well as negative correlations between grades and words which had connotations of sadness, risk and were phrased as questions. It appears that all annotations encourage the old paradigm as they focus upon the delivery of information, which minimises the potential upon student learning. It is argued that markers' practices could be modified to incorporate appropriate language and direction which could have a more positive impact upon students learning, maximising the benefit of in-text comments.

### **Extended Summary**

Although the use of annotations on students' work in the form of in-text comments is frequently utilised, it is often used with little specific training and is not supported by an evidence-base. It has been argued that this practice can provide bite-sized pieces of information in the place to which it refers (Rand, 2017), encouraging active participation by students (Liu, 2006, cited in Ball, 2010). On the other hand, Arts et al. (2016) found that most comments on formative work related to that specific task (48%), with an overuse of corrections and little information about how to improve future work. Annotations which are characterised by evaluative statements and corrective advice fail to support the notion of sustainable feedback practices (Carless et al., 2011). It is argued that the use of annotations that promote student action and self-regulation, may facilitate a new paradigm approach where students use comments to support their learning (Carless, 2015). This research focuses upon the language and content of in-text comments, and how this differs based on the grade given to the work, to determine whether annotations are used to promote student action.

### **Method**

Annotations in the form of 1760 in-text comments by multiple markers on 52 undergraduate Psychology and Health Sciences summative essays were sampled. Essays covered each year of study and all grade ranges.

To categorise the content of the in-text comments, a coding scheme was developed building upon the work by Hattie and Timperley (2007) and Arts et al. (2016) (see Table 1).

Table 1. Model of Feedback to Enhance Learning Coding Scheme

<b>Model of Feedback to Enhance Learning</b>	<b>Category</b>	<b>Descriptor</b>
Feedback on self-regulation level	6	Informative developmental guidance
Feedback related to the process	5	Directive developmental guidance in relation to future work
Feedback related to task performance	4	Corrective or directive advice related to task
	3	Evaluative statement or comment related to task
	2	Edit
Feedback on the self or personal level	1	Personal Opinion

After training to reach a consistent level of agreement (Krippendorff's  $\alpha = .72$ ), one researcher coded the comments independently (10% reliability check: Krippendorff's  $\alpha = .76$ ).

Text analysis software, Linguistic Inquiry and Word Count (LIWC) (Pennebaker et al., 2015), was used to identify the percentage of words in each comment from particular linguistic domains.

## Results

Table 2 shows the average number of comments and average percentage of comments coded within each feedback category.

Table 2. Summary statistics

	<b>Min</b>	<b>Max</b>	<b><i>M</i></b>	<b><i>SD</i></b>
<b>Total number comments</b>	8	119	34.71	27.63
<b>Comments coded category 1</b>	0	0	0	0
<b>Comments coded category 2</b>	0	22	5	6
<b>Comments coded category 3</b>	0	93	40	21
<b>Comments coded category 4</b>	7	73	43	17
<b>Comments coded category 5</b>	0	42	11	9
<b>Comments coded category 6</b>	0	9	1	2

*Note.*  $N = 52$  essays.

On average, few comments on each essay were coded as category 2 (5%) or 6 (1%). No comments were coded as category 1. The number of comments significantly reduced as grades increased,  $r = -.29$ ,  $p = .041$ . As grades increased, the use of category 3 comments significantly increased,  $r = .63$ ,  $p < .001$ . The use of category 4 and category 5 comments significantly reduced as grades increased ( $r = -.50$ ,  $p < .001$  and  $r = -.43$ ,  $p = .001$ , respectively). There were no relationships between grades and the use of category 2 or category 6 comments.

Positive correlations were found between grades and words expressing: emotional tone,  $r = .46$ ,  $p = .001$ , positive emotions,  $r = .59$ ,  $p < .001$ , and reward,  $r = .61$ ,  $p < .001$ . There were negative correlations between grades and words expressing: interrogatory words,  $r = -.31$ ,  $p = .03$ , sadness,  $r = -.34$ ,  $p = .018$ , risk,  $r = -.63$ ,  $p < .001$ , words focusing on the present,  $r = -.39$ ,  $p = .007$ , and use of question marks,  $r = -.45$ ,  $p = .001$ .

### **Conclusions and implications**

In-text comments were mostly related to task performance, rather than process, and comments at the level of self-regulation were rarely identified. The language used in comments differed depending on the grade assigned to the work. More positive emotional language was used for better graded work. With poorer work, more questions appeared to be directed to the students in the comments. Since directive developmental guidance comments decreased for work receiving higher marks, it appears that comments on higher graded essays did not tend to encourage sustainable feedback practices that enable students to enhance future work (Carless et al., 2011).

The findings suggest that there is a need for teachers to reconsider the use of annotations, in terms of the message and the language used in order to ensure written comments can enable the student action, which is the feature of the new paradigm. There is a need to develop markers who are deep annotators, who write more critical and analytical comments, as opposed to surface annotators who recycle information and lack analytical engagement with the work and feedback process (Liu, 2006, cited in Ball, 2010) in order to encourage the same practice by students.

### **References**

- Arts, J. G., Jaspers, M., & Joosten-ten Brinke, D. (2016). A case study on written comments as a form of feedback in teacher education: so much to gain. *European Journal of Teacher Education*, 39(2), 159–173.
- Ball, E. C. (2010). Annotation an effective device for student feedback: A critical review of the literature. *Nurse Education in Practice*, 10(3), 138–143.
- Carless, D. (2015). *Excellence in University Assessment: Learning from Award-Winning Practice*. Abingdon, Oxford: Routledge.
- Carless, D., Salter, D., Yang, M., & Lam, J. (2011). Developing sustainable feedback practices. *Studies in Higher Education*, 36(4), 395–407.
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81–112.
- Pennebaker, J. W., Boyd, R. L., Jordan, K., & Blackburn, K. (2015). *The development and psychometric properties of LIWC2015*. Austin: University of Texas.
- Rand, J. (2017). Misunderstandings and mismatches: The collective disillusionment of written summative assessment feedback. *Research in Education*, 97(1), 33–48.