_____

Exploring content vs writer focus in written student feedback: Perceived impact on student emotional response, attention, and usefulness

Judith L. Stevenson, Kalliopi Mavromati, James E. Bartlett, Lorna I. Morrow, Linda M. Moxey and Eugene J. Dawydiak

Abstract

Emotional reaction to negative academic feedback can be a barrier to students' engagement with feedback. One factor that may impact students' emotional response is whether feedback references the student (e.g., "Your writing") or the content (e.g., "The writing"). This exploratory study investigated how form of reference in feedback impacts emotional reaction. Student participants (N=106) read simulated feedback statements that varied according to reference type (pronominal "your", neutral "the") and feedback polarity (positive, negative) and provided ratings for emotional response, attention paid to, and usefulness of the feedback. An open response question queried participants' perceptions regarding personal or neutral reference in feedback. In three separate 2x2 within-subject ANOVA, there was no significant effect of reference type on quantitative feedback ratings for emotion, attention, or usefulness, but content analysis on qualitative responses revealed that half the sample preferred neutral reference, largely to mitigate the emotional impact of negative feedback. Feedback polarity had a consistent significant effect where students perceived positive feedback as higher in happy emotion and usefulness compared to negative feedback. The results for attention were less consistent with a small decrease in attention for positive compared to negative feedback for neutral references but a smaller non-significant difference for positive references. Our results inform practice in terms of how written feedback can be framed to facilitate engagement.

Keywords

_____

[1] University of Glasgow, Glasgow, UK

Corresponding Authors:

Judith Stevenson and Eugene Dawydiak, School of Psychology and Neuroscience, University of Glasgow, Glasow, G12 8QB, UK

Email: Judith.Stevenson@glasgow.ac.uk, Eugene.Dawydiak@glasgow.ac.uk

Introduction

Educators provide feedback to students as a means of assessment and to direct further academic development. Therefore, students should find feedback clear and understandable to feed forward into future work. Significant staff time goes into feedback provision, but staff are often left wondering just how much engagement is taking place with the feedback (Jonsson & Panadero, 2018; Rowe, 2017). Both staff and students seem discouraged about how, and if, feedback is really working (Price et al., 2010).

Emotional barriers to feedback

One component that may be a barrier in students' engagement with feedback is their emotional response to it (Handley et al., 2011; Harris et al., 2014; Molloy et al., 2012; Ryan & Henderson, 2018). Students tend to *immediately* interpret feedback as either *positive* or *negative* depending on whether the feedback is noncritical (when the marker comments on what has been done well) or critical (when the marker comments on what has not been done well) (Holmes, 2023; Price et al., 2010). For example, students can feel pride, enjoyment, contentment, and relief in response to feelings of achievement rewarded by positive feedback (Pekrun, 2006). They also engage more with positive feedback and interpret it as more useful (Winstone et al., 2017). However, when feedback is negative, feelings of shame, anxiety, frustration and disappointment are common (Pekrun, 2006). Negative emotional responses do not just impact the *initial* processing of the feedback, but they can also detrimentally impact on *subsequent* engagement with the feedback (Hill et al., 2021) with engagement being delayed or prevented altogether (Holmes, 2023). Laudel and Narciss (2023) suggest emotional negative responses to negative feedback are an example of threat detection, and these emotional responses can impact on students' cognitive processing (Hill et al., 2021; Värlander, 2008). In other words, unhappy, emotional responses towards negative feedback may block subsequent engagement with feedback, while happy, emotional responses may facilitate students' engagement with their feedback. Therefore, it is important to investigate what factors have an impact on emotional reaction to feedback.

Reference type: pronominal vs neutral

Careful communication of feedback, particularly when it is negative, is vital, so that negative emotional reactions like the ones discussed above are minimised (Hill et al., 2021). Arguably, one of the most important aspects to consider in terms of how students experience written feedback is the tone in which the feedback is written (Lipnevich et al., 2016). However, markers are still in pursuit of how best to deliver this. Anecdotally, when marking student work, some markers report a tendency to write feedback comments either with pronominal reference to the reader or more neutrally. For example: "your evidence is very strong/weak here"; vs "the evidence is very strong/weak here". In natural language, "you" is the most frequently used personal pronoun (followed by "I" and "we") (Yeo & Ting, 2014). However, we are unaware of any literature investigating how reference type (pronominal vs neutral) does or does not affect student perceptions of feedback, leaving the potential impact of reference type on students' responses to feedback currently unknown.

The current study

In our study, when manipulating for the independent variable (IV) of feedback polarity, we use the term *positive feedback* to refer to the identification of something that has been done well (e.g., "The writing here is very clear"); and the term *negative feedback* to refer to feedback that identifies something that has not been done well (e.g., "The writing here is very unclear").

A second IV that may impact students' responses to feedback is whether the student is addressed using pronominal (e.g., "Your writing here is very clear/unclear") or neutral (e.g., "The writing here is very clear/unclear") reference. When considering how reference type may interact with feedback type, we suggest that it may play a role in students' experience of written feedback, as illustrated in see Table 1.

Table 1. Example positive and negative feedback using pronominal and neutral reference

Feedback examples	Reference type	Feedback polarity	Interpretation
Your writing here is very clear"	Pronominal	Positive	Positive comments interpreted as personal
"Your writing here is very unclear"	Pronominal	Negative	Negative comments interpreted as personal
"The writing here is very clear"	Neutral	Positive	Positive comments interpreted less personally
"The writing here is very unclear"	Neutral	Negative	Negative comments interpreted less personally

Further measures of engagement: attention and perceived usefulness

Although the main focus of our study relates to investigating how reference type may impact on how students respond emotionally to feedback, we feel it is also important to consider two other measures of engagement: attention paid to feedback and perceived usefulness of feedback. One of the main objectives of feedback is so that students can understand how they can bridge achievement and the desired outcome (Brown, 2007). In other words, feedback should be used to support the maintenance of what has been done well, and/or as a means of developing aspects that require improvement. For students to use feedback in this way, they need to: 1) *pay attention to it*; and 2) interpret the feedback as *useful* if they are going to apply it to future assessments. When we process information, interpretation and attention are mutually inclusive, in other words, they work together (Blanchette & Richards, 2010). This aligns with Winstone et al.'s (2017) argument that feedback polarity (positive, negative) and usefulness are experienced in tandem. We suggest that if feedback is processed at a superficial level, subsequent engagement with it and interpretation of how useful it is may be lower. If feedback can be processed on a deeper level, for example, being interpreted as worthy of attention and useful, students can take it forward and apply it to future work.

While much of the literature on assessment and feedback focuses on theories and models of feedback (for a review, see Lipnevich & Panadero, 2021), the purpose of our study is not to extend on this but to investigate how reference type may affect student perceptions of feedback. Our present exploratory study aims to investigate reference type in student feedback, contribute to knowledge on how students perceive feedback, and further understand how we can facilitate engagement with feedback, using a mixed methods approach. Given the exploratory nature of the study, we had no confirmatory hypotheses to test, but we had three research questions:

1. How do pronominal reference and feedback polarity affect students' emotional response towards feedback?
2. How do pronominal reference and feedback polarity affect students' perceived attention towards feedback?
3. How do pronominal reference and feedback polarity affect students' perceived usefulness toward feedback?

Method

Participants

We recruited an opportunity sample of 125 students currently studying at UK Higher Education institutions. This was done by sharing an advert on University of Glasgow Microsoft teams channels, social media (Facebook, X), and by asking contacts at other institutions to share the advert. We excluded data from five participants who asked for their data to be removed and 14 who had missing data in our key variables, leaving 106 as our final sample size. Table 2 shows full demographic data for our sample. Ethical approval was granted by the University of Glasgow College of Science and Engineering ethics committee (application 300210153).

Table 2. Demographic characteristics

Demographic characteristic	Frequency	Percentage (%)
Gender		
Woman	90	84.91
Man	13	12.26
Non-binary	3	2.83
University		
U of Glasgow	103	97.17
Other UK	2	1.89
Other Scottish	1	0.94
UoG Psychology		
Yes	94	88.68

Demographic characteristic	Frequency	Percentage (%)
No	12	11.32
Year of Study		
1st Year	16	15.09
2nd Year	32	30.19
3rd Year	31	29.25
4th Year	17	16.04
5th Year	10	9.43

Design

We used a 2 x 2 within-subjects design (Bobbitt, 2021). Participants completed each combination of the IVs reference type (pronominal or neutral) and feedback polarity (positive or negative), producing four conditions: (1) pronoun-positive; (2) pronoun-negative; (3) neutral-positive; (4) neutral-negative. We presented participants with two items from each condition meaning they each completed 8 trials: 2 pronoun-positive; 2 pronoun-negative; 2 neutral-positive; 2 neutral-negative. The three dependent variables were the mean student ratings across two versions of emotional response, attention, and usefulness.

Materials

Stimuli

We generated an initial set of items and counterbalanced them so that the items were worded identically, except for reference type and feedback polarity. Counterbalanced items are presented in Table 3. We selected eight items that best represented written feedback statements that students could understand studying across multiple academic disciplines. Since each individual item could appear in one of four conditions, we divided the set of 8 generated feedback items into 4 four blocks (2 items per condition) and rotated, following a Latin Square design, to form 4 files. For example, items X and Y appeared in condition A in file 1, condition B in file 2 etc. This meant that each participant saw all 8 items (2 in each condition), but no item appeared more than once for them. The full stimulus list is available online: https://osf.io/v8gdx/.

In each trial, we presented participants with the prompt "Imagine you received the following feedback comment on a highlighted piece of coursework text". Participants saw one of four examples of coursework feedback relating to a condition in the design. For example, in the pronoun-positive condition, participants saw the feedback "Your writing here is very clear" as demonstrated in Table 3. After reading the feedback, participants completed three measures to indicate their opinion (see Figure 1).

Table 3. Item files counterbalanced for reference type and feedback polarity.

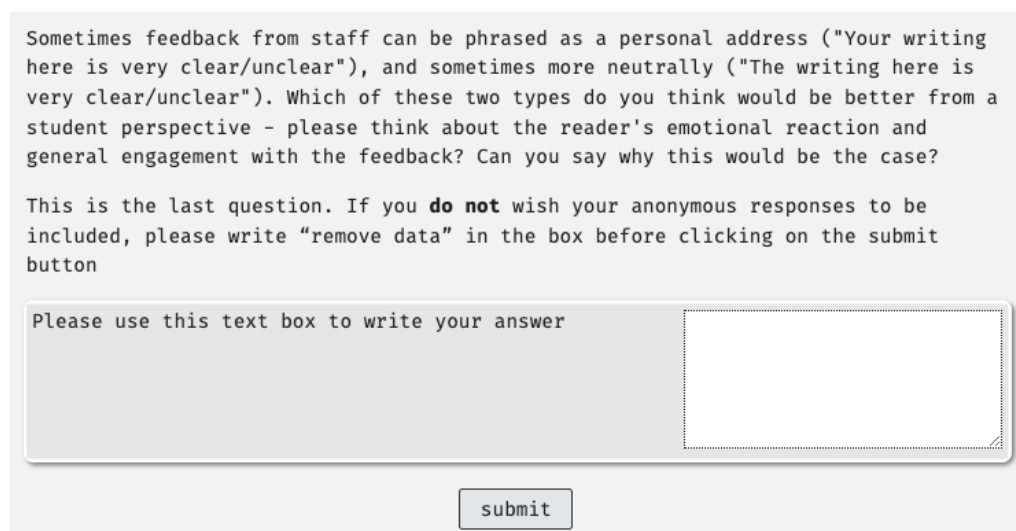| File 1	File 2	File 3	File 4
Your writing here is very clear	Your writing here is very unclear	The writing here is very clear	The writing here is very unclear
Evidence cited for your claim here is weak	Evidence cited for the claim here is strong	Evidence cited for the claim here is weak	Evidence cited for your claim here is strong
The argument here is very strong	The argument here is very weak	Your argument here is very strong	Your argument here is very weak
The interpretation of the literature here is very inaccurate	Your interpretation of the literature here is very accurate	Your interpretation of the literature here is very inaccurate	The interpretation of the literature here is very accurate
Your citation here is correctly formatted	Your citation here is incorrectly formatted	The citation here is correctly formatted	The citation here is incorrectly formatted
Your explanation here is very uninformative	The explanation here is very informative	The explanation here is very uninformative	Your explanation here is very informative
The evaluation here is very strong	The evaluation here is very weak	Your evaluation here is very strong	Your evaluation here is very weak
The writing here is very unprofessionally articulated	Your writing here is very professionally articulated	Your writing here is very unprofessionally articulated	The writing here is very professionally articulated

Figure 1. Example experimental trial.

Measures

For each dependent variable, we used a visual analogue scale (Crichton, 2001) with a negative response at the start (representing 0) and a positive response at the end of the scale (representing 100). The start and end points of the scale were "Very unhappy" and "Very happy" to measure students' emotional response to each feedback item, "I would definitely not pay attention to it" and "I would definitely pay attention to it" to measure how much attention students would pay to each written feedback item, and "Not at all useful" to" Very useful" to measure students' perceptions of how useful they rated each written feedback comment to be in guiding their future work.

Open-ended question

We also asked participants an open-ended question (Figure 2) that asked them to comment on the use of personal and neutral reference in written feedback, and the potential impact on engagement with feedback: "Sometimes feedback from staff can be phrased as a personal address ("Your writing here is very clear/unclear"), and sometimes more neutrally ("The writing here is very clear/unclear"). Which of these two types do you think would be better from a student perspective - please think about the reader's emotional reaction and general engagement with the feedback? Can you say why this would be the case?" Open-ended questions allow participants the opportunity to reflect, provide their own perspective, and use their own language (Harland & Holey, 2011).

```
Sometimes feedback from staff can be phrased as a personal address ("Your writing
here is very clear/unclear"), and sometimes more neutrally ("The writing here is
very clear/unclear"). Which of these two types do you think would be better from a
student perspective – please think about the reader's emotional reaction and
general engagement with the feedback? Can you say why this would be the case?

This is the last question. If you **do not** wish your anonymous responses to be
included, please write "remove data" in the box before clicking on the submit
button

 Please use this text box to write your answer

                                                submit
```

Figure 2. Open-ended question

Procedure

Participants completed the study individually using the online experiment platform Experimentum (DeBruine et al., 2020). After providing informed consent, each participant provided demographic information for age, gender identity, institution of study, if they study psychology at the University of Glasgow, and their year of study.

After reading task instructions, we presented one item per page and participants completed the three visual analogue scales as shown in Figure 1. After completing all eight pages, participants could provide a free-text response to the open-ended question (Figure 2). Finally, we presented participants with a debrief sheet explaining the rationale behind the study and contact details of the lead researchers.

Data Analysis

Quantitative Data

Per standard practice, alpha was set at $\alpha = .05$. Due to the lack of previous studies, we could not perform an a priori power analysis. Instead, we conducted a sensitivity power analysis (Bartlett & Charles, 2022), suggesting our final sample size of 106 would be sensitive to detect effects of Cohen's $d = 0.32$ for any pairwise comparison comparing two conditions and $r = .31$ for any correlation (two-tailed, 90% power).

We used R (Version 4.1.3; R Core Team 2022) and R Studio for the data analysis. The data and scripts are openly available on our OSF project (https://osf.io/v8gdx/), including a reproducible R Markdown version of the method and results. After checking parametric assumptions, we used three 2x2 within-subject ANOVA - one for each dependent variable - and Pearson's correlations for the relationship between conditions. For each ANOVA, we report partial eta squared ($\eta2p$), and for each pairwise comparison, we report the mean difference (Mdiff) and Cohen's d alongside their 95% confidence interval.

Content Analysis

One hundred participants answered the open-ended question that allowed them to express in their own words how they thought personal or neutral reference can affect emotional response to, and engagement with, written feedback. Two independent members of the team (ED and JS) coded the responses to the open-ended question using inductive content analysis. This technique is used to code patterns of words in text data and assign them to distinct categories (Elo & Kyngäs, 2008). This then allowed us to observe the frequency of each category. The unit of analysis was each entire individual text response (as opposed to the frequency of independent words) and we allocated preferences for reference type in written feedback to one of six categories which are presented in Table 4. Authors ED and JS then recoded the textual responses together to determine intercoder agreement. The textual data is openly available on our OSF project.

Table 4. Content analysis coding categories

Category	Coding
Pronominal	Preference expressed for personal reference in feedback
Neutral	Preference expressed for neutral reference in feedback
Polarity dependent	Preference expressed for personal reference when feedback is positive, and neutral reference when feedback is negative
Mixed	The benefit of both forms of reference was acknowledged
No preference	No preference expressed for personal or neutral reference in feedback
Unclear	Response unclear

Results

Quantitative results

Table 5 contains summary statistics for the combination of polarity and condition for each dependent variable. We applied a 2x2 within-subjects ANOVA to each dependent variable to investigate how students' emotional response, attention, and perception of usefulness differed across conditions. In our online supplementary reproducible analyses, we also checked the robustness of our findings if we apply linear mixed effects models instead of averaging over the different stimuli. We focus on reporting the results of our 2x2 ANOVA here as our conclusions are robust across analyses, apart from attention which we note below.

Table 5. Descriptive statistics

Reference	Polarity	Condition	Mean	SD	Range
Pronoun	Positive	Emotion	81.30	15.39	27 - 100
		Attention	68.17	18.21	14.5 - 99.5
		Usefulness	66.17	19.31	0 - 100
	Negative	Emotion	23.52	13.97	1 - 56
		Attention	69.73	19.24	14 - 100
		Usefulness	50.57	22.23	0 - 97
Neutral	Positive	Emotion	80.72	14.92	32 - 100
		Attention	65.05	20.46	10.5 - 100
		Usefulness	65.52	19.45	17.5 - 100
	Negative	Emotion	27.59	15.56	1 - 71
		Attention	71.96	18.22	7.5 - 100
		Usefulness	53.76	21.99	4 - 100

Emotion

There was no significant main effect of reference type (personal vs neutral; F (1, 105) = 3.18, p = .077, η2p = .029) but there was a significant main effect of feedback polarity (positive vs negative) on emotion ratings where mean happiness ratings were higher for positive polarity feedback (F (1, 105) = 811.51, p < .001, η2p = .885). There was also a significant but weak interaction between reference type and feedback polarity on emotion ratings (F (1, 105) = 4.89, p = .029, η2p = .044). Pairwise comparisons show that- reassuringly - there were very large increases in emotion ratings towards happier responses for positive feedback compared to negative feedback for both pronoun (Mdiff = 57.78 [53.39, 62.17], d = 2.54 [2.15, 2.94]) and neutral references (Mdiff = 53.13 [48.74, 57.51], d = 2.33 [1.97, 2.71]). In comparison, there were very small to small differences between pronoun and neutral references for both positive (Mdiff = 0.58 [-2.35, 3.52], d = 0.04 [-0.15, 0.23]) and negative feedback (Mdiff = -4.07 [-6.82, -1.31], d = -0.28 [-0.48, -0.09]), as shown in Figure 3.
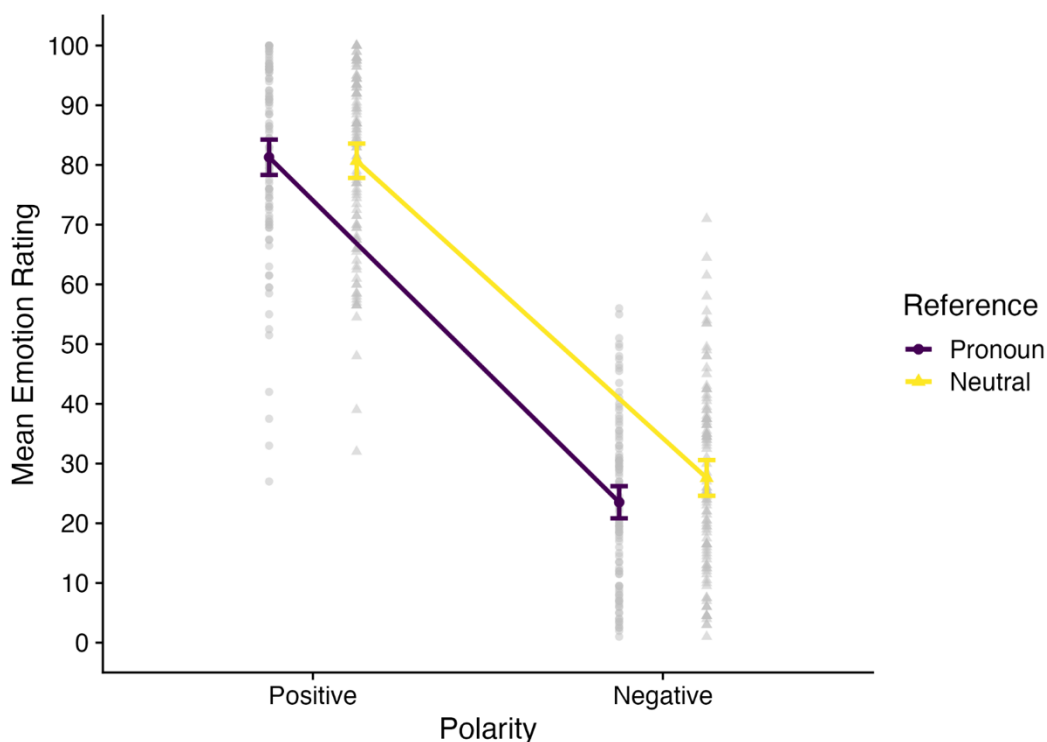

Figure 3. Polarity by reference for emotion. The interaction plot shows the mean and within-subject 95% confidence interval per condition.

Attention

There was no significant main effect of reference type (F (1, 105) = 0.13, p = .719, $\eta2p$ < .001) but there was a weak significant main effect of feedback polarity on attention ratings (F (1, 105) = 4.23, p = .042, $\eta2p$ = .039) such that mean attention ratings were higher for negative polarity feedback. There was also a significant but weak interaction between reference type and feedback polarity (F (1, 105) = 5.39, p = .022, $\eta2p$ = .049). In our supplementary linear mixed effects model, both weak effects are non-significant, so we treat these differences cautiously. Pairwise comparisons showed little difference in attention between negative and positive feedback for pronoun references (Mdiff = -1.56 [-6.13, 3.01], d = -0.07 [-0.26, 0.13]), but there was a small decrease in attention for positive compared to negative feedback for neutral references (Mdiff = -6.92 [-11.71, -2.12], d = -0.28 [-0.47, -0.08]). For the comparison between pronoun and neutral references, the effect sizes were very small without a clear direction for both positive (Mdiff = 3.12 [-0.34, 6.58], d = 0.17 [-0.02, 0.37]) and negative feedback (Mdiff = -2.24 [-5.46, 0.99], d = -0.13 [-0.33, 0.06]), as shown in Figure 4.
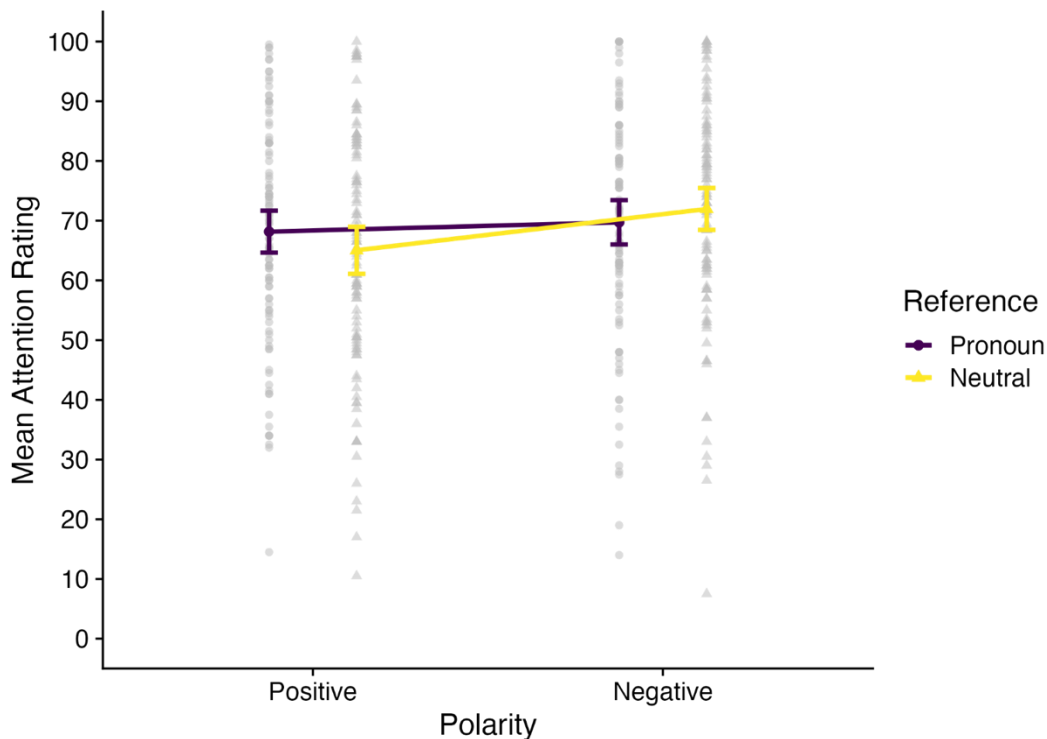

Figure 4. Polarity by reference for attention. The interaction plot shows the mean and within-subject 95% confidence interval per condition.

Usefulness

There was no significant main effect of reference type (F (1, 105) = 0.76, p = .384, η2p = .007) but there was a significant main effect of feedback polarity on usefulness ratings (F (1, 105) = 49.2, p < .001, η2p = .319) where mean usefulness ratings were higher for positive polarity feedback. The interaction between reference type and feedback polarity was not statistically significant (F (1, 105) = 1.93, p = .168, η2p = .018). Pairwise comparisons focusing on differences in polarity showed students perceived feedback to be more useful for positive feedback compared to negative feedback for both pronoun (Mdiff = 15.6 [10.76, 20.44], d = 0.62 [0.41, 0.83]) and neutral references (Mdiff = 11.75 [7.11, 16.4], d = 0.49 [0.29, 0.69]), as shown in Figure 5.
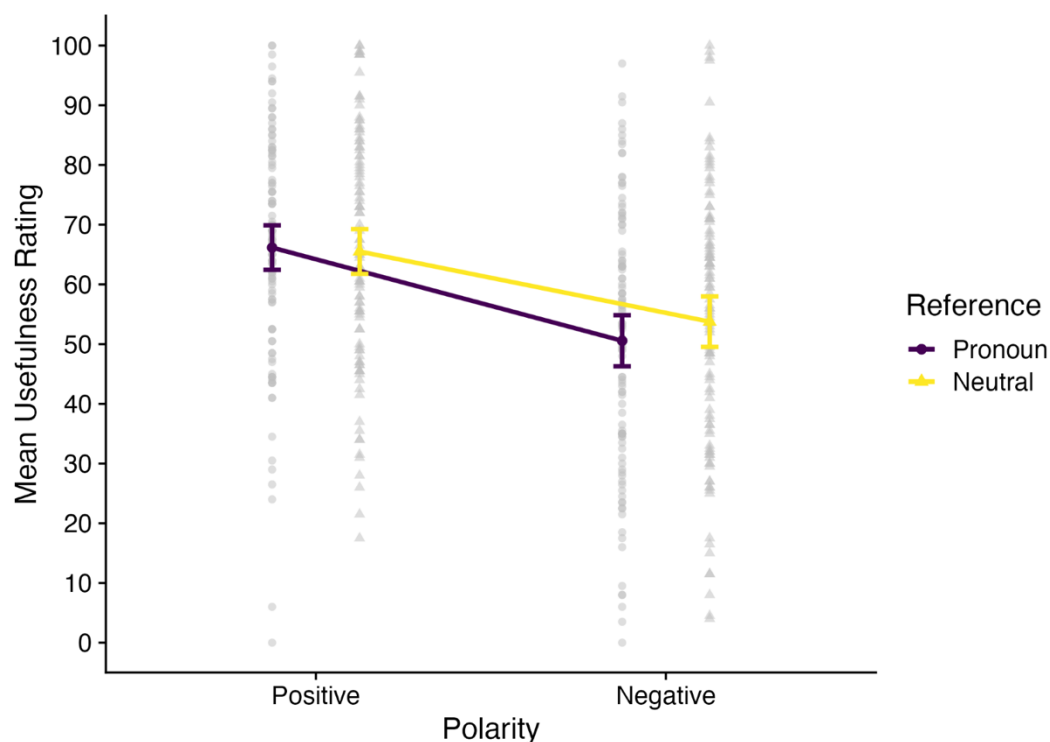

Figure 5. Polarity by reference for usefulness. The interaction plot shows the mean and within-subject 95% confidence interval per condition.

Content analysis

One hundred participants answered the open-ended question. After coding responses, we allocated preferences for reference type in written feedback to one of six categories (see Table 6 for frequency data). Fifty respondents indicated a preference for the use of neutral reference. Of these 50 responses, 44 explicitly indicated that this preference was driven by a desire to soften the blow of negative feedback. Twenty-four indicated that they would prefer personal reference. Eleven respondents preferred polarity dependent feedback, that is, personal reference if the feedback provided was positive, and neutral reference if it was negative. The benefits of both forms of reference were acknowledged by 5 respondents. No preference for either personal or neutral was expressed by 8 respondents and 2 provided unclear answers.

Table 6. Frequency statistics for preference for reference type in written feedback

Reference type	*N*=100
Neutral	50
Pronominal	24
Polarity dependent	11
No preference	8
Mixed	5
Unclear	2

Further exploratory analysis: correlations

After completing our planned analyses (ANOVA and content analysis), we decided to carry out a further exploratory analysis of our data. We explored associations between emotional reaction, perceived usefulness, and perceived attention in the contexts of positive and negative feedback polarity. Given the lack of any statistically significant effect of reference type (pronoun vs neutral), we calculated the mean value across feedback polarity and question for each participant. Table 7 shows the Pearson correlations and their 95% CI.

Table 7. Means, standard deviations, and correlations with confidence intervals

Variable	*M*	*SD*	1.	2.	3.	4.	5.
1. Positive feedback - Emotion	81.01	13.10					
2. Negative feedback – Emotion	25.56	12.94	-.18				
			[-.36, .01]				
3. Positive feedback - Attention	66.61	17.16	.43**	-.05			
			[.26, .58]	[-.24, .14]			
4. Negative feedback - Attention	70.84	16.76	.47**	-.05	.22*		
			[.30, .60]	[-.24, .14]	[.03, .39]		

Variable	*M*	*SD*	1.	2.	3.	4.	5.
5. Positive feedback - Usefulness	65.84	17.62	.61**	.01	.65**	.34**	
			[.47, .72]	[-.18, .20]	[.53, .75]	[.16, .50]	
6. Negative feedback - Usefulness	52.17	18.42	.32**	.25*	.10	.54**	.38**
			[.13, .48]	[.06, .42]	[-.09, .29]	[.38, .66]	[.20, .53]

Note. M and SD are used to represent mean and standard deviation, respectively. Values in square brackets indicate the 95% confidence interval for each correlation. * indicates p < .05. ** indicates p < .01.

Emotional reaction to positive feedback was significantly positively correlated with attention paid to both positive feedback (r (104) = 0.43, p < .001, 95% CI = [0.26, 0.58]) and negative feedback (r (104) = 0.47, p < .001, 95% CI = [0.3, 0.6]), and perceived usefulness of both positive feedback (r (104) = 0.61, p < .001, 95% CI = [0.47, 0.72]) and negative feedback (r (104) = 0.32, p = .001, 95% CI = [0.13, 0.48]). However, emotional reaction to positive feedback was not significantly correlated with emotional reaction to negative feedback (r (104) = -0.18, p = .058, 95% CI = [-0.36, 0.01]).

Emotional reaction to negative feedback was only significantly positively correlated with perceived usefulness of negative feedback (r (104) = 0.25, p = .01, 95% CI = [0.06, 0.42]). For the remaining variables, the correlations with emotional reaction to negative feedback were very small and not statistically significant.

Discussion

We investigated if reference type in written student feedback impacts student ratings of the feedback on three engagement-related dependent variables: (1) emotional response; (2) attention; and (3) usefulness. Our quantitative analyses revealed no significant effect of reference type (whether the feedback addressed the student ("you") or the content ("the") on ratings of emotion, attention, and usefulness. Feedback polarity (whether the feedback was positive or negative) had a consistent significant effect where happiness and usefulness ratings were higher in response to positive feedback in comparison to negative feedback. There was a more nuanced pattern for attention, where there was only a small decrease in attention for positive compared to negative feedback for neutral references.

Our open-ended question queried opinions about preferences for pronominal or neutral reference in written feedback on emotional responses. Content analysis revealed that 50 out of 100 respondents expressed a clear preference for neutral reference with 44 indicating that it "cushions the blow" of negative feedback. Negative emotional responses towards negative feedback can produce a "knee jerk" reaction where students see the feedback as a reflection of themselves and take it personally impairing or in some cases, preventing subsequent feedback engagement (Holmes, 2023). This demonstrates the

powerful impact of emotion (in this case, negative) on cognitive processing (Värlander, 2008). As stated in our results, variations of the word "attack" featured in 15 responses, for example, "I think having the neutral stance makes it less of a personal attack. I think people would be more likely to take the neutral feedback on board than personal." This is in line with Hill et al. (2021) who found that students expressed the desire for careful communication of negative feedback comments, due to the impact of such feedback on their self-esteem and self-confidence. In the context of our findings, careful communication could be a neutral reference (preferred by half of the sample) to mitigate the emotional effects of negative feedback.

In contrast, an explicit desire for personally referenced feedback was only expressed by 24 out of 100 participants. The role of emotion emerged less in these responses, with personal reference simply being seen as more customised to the individual and of a less general nature, for example, "it feels more tailored to my own work", and "it gives credit to the student". Personalised feedback tends to be perceived as less generic (Pitt & Winstone, 2018) and of higher value to students (Birch et al., 2016), but only a quarter of our sample expressed this preference. An explanation for this could be that a third category of responses expressed by 11 participants who indicated polarity dependence, that is, preference for personal reference when feedback is positive and neutral reference when feedback is negative (as illustrated in Table 1). This was followed by five respondents who acknowledged the potential role of both forms reference type with a mixed preference, and eight who had no preference and said that reference type did not really matter to them.

One potentially important finding of our qualitative content analysis is that students clearly had different perspectives on the use of personal vs neutral in written feedback, while this was not apparent in our quantitative results. Consistent with our initial reasoning, a small proportion of students felt that it was polarity dependent, indicating that they thought personal is better for positive feedback, while a larger proportion indicated that neutral can be a buffer against the impact of negative feedback. It is important to consider that our quantitative results for reference type could potentially be due to the different perspectives (expressed in response to the qualitative open-ended question) adding 'noise' to the quantitative data and preventing the emergence of a clear pattern, reflecting a lack of consensus. Future research could investigate this possibility by analysing groups based on these differing perspectives (e.g., personal, neutral, polarity dependent) individually. For now, however, it is helpful for feedback providers to be aware that there exists a substantial proportion of students who feel that negative feedback is better expressed neutrally.

Another plausible explanation for the divergent findings between our quantitative and qualitative results could stem from the rigorously controlled stimuli used in this study not being sensitive enough to capture the full spectrum of the participants' responses. The hypothetical nature of the stimulus question "imagine you received" is likely to have mitigated potential reactions in both the personal and neutral reference condition. Student participants were aware that they were taking part in a study and that this was clearly not a real assessment carrying real feedback and grade. This artificial context could have impacted any genuine potential effects of reference type on ratings. To keep the discourse (i.e., feedback items) consistent across conditions, and to limit the impact of extraneous variables such as variable text length and different lexical characteristics (e.g., word frequency (Wilson, 1988) and word valence (Bradley & Lang, 1999) across

conditions, our stimuli were carefully designed and controlled by applying methods from psycholinguistics (i.e., counterbalancing of matched lexical stimuli). This level of control would not be possible if we were using real example of real-world feedback. A second factor that we felt was important was to develop feedback items which were broadly relevant across different disciplines and coursework settings so that we could investigate the general impact of reference type in a way that was as little constrained by specific discipline and type of coursework as possible. A third consideration was that our feedback should reflect the practice of anonymous marking which is now standard in many institutions (O'Donovan et al., 2021). Consequently, our feedback items were relatively unrealistic and of low ecological validity in two important ways. First, it was not contextualised in terms of representing a response to a particular aspect of a student's specific coursework by a marker who had fully engaged with a piece of work and recognised the effort but, rather, was more "cut and paste" (O'Donovan et al., 2021). This will have limited participants' available information to base ratings on. Second, it did not include feedforward for improvement on how to use the feedback to enhance future work which is a factor which not only helps students improve (e.g., Fong et al., 2019; Lowe & Shaw, 2019) but can also mitigate the disengagement effect of negative feedback (e.g., Hill et al., 2021). Issues regarding the ecological validity of the simplified and uncontextualised feedback items may have constrained the participants' capacity to make real-life judgements of emotion, attention, and usefulness. This may in turn have had the effect of lowering positive appraisal of the feedback and potentially contributing to the lack of a clear pattern for reference type in the ANOVA. Although it was important for our exploratory study to start with rigorously controlled stimuli, future research could utilise more contextualised feedback and integrate elements of feedforward to improve the ecological validity of the stimulus materials.

Considering our findings as presented above, we propose some recommendations that can help markers inform their practice by considering how they word their feedback. 1) Focus on the submission: when providing feedback that is likely to be interpreted as negative, markers should direct comments towards the submission rather than towards the student. 2) We suggest that this can be achieved in one very simple and efficient way: by using neutral reference, particularly when providing negative feedback. This approach allows markers to deliver precise and necessary feedback while minimising the risk of students taking it personally. 3) By using neutral reference, markers can enhance engagement as students engage with the feedback more constructively.

Our unplanned, further exploratory correlations investigated potential associations between emotional reaction and the other engagement-related variables of attention and usefulness. As reference type ("your" vs "the") had small, non-significant effects on feedback ratings, we collapsed scores across reference type and examined correlations between our dependent variable ratings of emotional reaction, attention, and usefulness. Emotional reaction is an important factor for engagement with feedback where students appreciate positive feedback and experience increased self-confidence and feelings of self-worth when they received it (Hill et al., 2021). Based on this, we felt it was important to explore potential associations between emotional reaction and the other engagement-related variables of attention and usefulness. Any findings may be informative in suggesting the potential association between emotion and engagement and therefore identify areas where future research may be able to systematically investigate.

The primary finding of interest here in this exploratory analysis relates to the significant moderate to large positive correlations between emotional reaction to positive feedback, and perceptions of attention and usefulness for both positive and negative feedback. In other words, as happiness ratings increased, ratings of attention and ratings of usefulness also increased for positive feedback excerpts. This may indicate that the more a student enjoys (in this case, feels happy) affirmation of their work, it is associated with more positively generally engaging with the feedback (irrespective of the positive or negative polarity of the feedback) in terms of attending to it and seeing it as being useful. The importance of emotional reaction is also consistent with previous research. Emotional components of text are processed very early on (Herbert et al., 2008) and emotion has also been shown to influence many aspects of information processing, including attention and decision making (Blanchette & Richards, 2010; Tyng et al., 2017).

In contrast to the pattern observed for emotional reaction to positive feedback in our exploratory analysis, we found no clear associations with negative feedback. Emotion ratings to negative feedback only significantly correlated with ratings of perceived usefulness of negative feedback. In other words, the more positive/happy the emotional reaction was in response to a negative feedback item, the more this feedback was perceived as useful. As emotional responses to negative feedback did not significantly correlate with any other dependent ratings with very small effect sizes, it is possible this could be due to negative feedback items generally being associated with a more negative emotional experience. Negative feedback items in our study elicited lower mean emotion ratings than positive feedback items, reflecting a general unhappy experience. This general negative emotional response is consistent with research indicating that negative feedback is perceived as a threat (Laudel & Narciss, 2023). In our study, negative feedback items perceived as conveying threat may lead to students disengaging from the feedback and "tuning out" if they see it as a threat to themself (Laudel & Narciss, 2023). This may explain why there is no clear pattern of results in relation to emotion ratings to negative feedback in this study.

Conclusion

In conclusion, we explored how reference type (personal, neutral) in student feedback affects student perceptions of feedback. We found no statistically significant effect of reference type quantitatively, but content analysis yielded an overall preference for neutral reference, particularly when feedback polarity is negative, to lessen the emotional impact on the self. We observed positive correlations between positive feedback and engagement with feedback, in that, if students enjoy their feedback, they tend to find it more useful and pay more attention to it. We encourage educators to explore how to provide feedback that students will engage with the most. The use of neutral reference can potentially mitigate the blow of negative feedback; however, our data is exploratory and requires replication and further investigation.

Disclosures

Funding

CRediT contributions

J.L.S: Conceptualisation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Supervision, Writing - original draft, and Writing - review & editing.

K.M: Data curation, Formal analysis, Methodology, Visualisation, and Writing - original draft.

J.E.B: Data curation, Formal analysis, Visualisation, and Writing - review & editing.

L.I.M: Investigation, Methodology, and Writing - review & editing.

L.M.M: Investigation and Methodology

E.J.D: Conceptualisation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Supervision, Writing - original draft, and Writing - review & editing.

R package citations

We created the results using R (Version 4.1.3; R Core Team, 2022) and the R-packages afex (Version 1.0.1; Singmann et al., 2021), apaTables (Version 2.0.8; Stanley, 2021), broom (Version 0.7.12; Robinson et al., 2022), car (Version 3.0.12; Fox & Weisberg 2019), cowplot (Version 1.1.1; Wilke, 2020), dplyr (Version 1.1.0; Wickham et al., 2023), effectsize (Version 0.5; Ben-Shachar et al., 2020), emmeans (Version 1.7.2; Lenth, 2022), ggplot2 (Version 3.4.1; Wickham, 2016), papaja (Version 0.1.1; Aust & Barth, 2022), pwr (Version 1.3.0; Champely, 2020), readr (Version 2.1.2; Wickham et al., 2022), stringr (Version 1.5.0; Wickham, 2022), tibble (Version 3.1.8; Müller & Wickham, 2022), tidyr (Version 1.3.0; Wickham et al., 2023), and viridis (Version 0.6.2; Garnier et al., 2021).

References

Aust, F., & Barth, M. (2022). *papaja: Prepare reproducible APA journal articles with R Markdown*. https://github.com/crsh/papaja

Bartlett, J., & Charles, S. (2022). Power to the people: A beginner's tutorial to power analysis using jamovi. *Meta-Psychology, 6*. https://doi.org/10.15626/MP.2021.3078

Ben-Shachar, M. S., Lüdecke, D., & Makowski, D. (2020). effectsize: Estimation of effect size indices and standardized parameters. *Journal of Open Source Software, 5*(56), 2815. https://doi.org/10.21105/joss.02815

Birch, P., Batten, J., & Batey, J. (2016). The influence of student gender on the assessment of undergraduate student work. *Assessment & Evaluation in Higher Education, 41*(7), 1065-1080. https://doi.org/10.1080/02602938.2015.1064857

Blanchette, I., & Richards, A. (2010). The influence of affect on higher level cognition: A review of research on interpretation, judgement, decision making and reasoning. *Cognition & Emotion, 24*(4), 561-595.

Bobbitt, Z. (2021, May 13). 2×2 factorial design. *Statology*. https://www.statology.org/2x2-factorial-design/

Bradley, M.M., Lang, P.J. (1999). *Affective Norms for English Words (ANEW)*. The NIMH Center for the Study of Emotion and Attention, University of Florida.

Brown, J. (2007) Feedback: The student perspective. *Research in Post-Compulsory Education, 12*(1), 33-51. https://doi.org/10.1080/13596740601155363

Champely, S. (2020). *pwr: Basic functions for power analysis*. https://CRAN.R-project.org/package=pwr

Crichton, N. (2001). Visual Analogue Scale (VAS). *Journal of Clinical Nursing, 10*(5), 706-6.

DeBruine, L., Lai, R., Jones, B., Mahrholz, G., Abdullah, A., Burns, N. (2020). Experimentum (Version v.0.2). *Zenodo*. https://doi.org/10.5281/zenodo.2634355

Elo, S., & Kyngäs, H. (2008). The qualitative content analysis process. *Journal of Advanced Nursing, 62*(1), 107-115. https://doi.org/10.1111/j.1365-2648.2007.04569.x

Fong, C. J., Patall, E. A., Vasquez, A. C., & Stautberg, S. (2019). A meta-analysis of negative feedback on intrinsic motivation. *Educational Psychology Review, 31*, 121-162. https://doi.org/10.1007/s10648-018-9446-6

Fox, J., & Weisberg, S. (2018). *carData: Companion to applied regression data sets*. Sage publications. https://doi.org/10.32614/CRAN.package.carData

Garnier, S., Ross, N, Rudis, R, Camargo, P.A, Sciaini, M. & Scherer, C. (2021). *viridis - colorblind-friendly color maps for R*. https://doi.org/10.5281/zenodo.4679424

Handley, K., Price, M., & Millar, J. (2011). Beyond 'doing time': Investigating the concept of student engagement with feedback. *Oxford Review of Education, 37*(4), 543-560. https://doi.org/10.1080/03054985.2011.604951

Harland, N., & Holey, E. (2011). Including open-ended questions in quantitative questionnaires—theory and practice. *International Journal of Therapy and Rehabilitation, 18*(9), 482-486. https://doi.org/10.12968/ijtr.2011.18.9.482

Harris, L. R., Brown, G. T., & Harnett, J. A. (2014). Understanding classroom feedback practices: A study of New Zealand student experiences, perceptions, and emotional responses. *Educational Assessment, Evaluation and Accountability, 26*(2), 107-133. https://doi.org/10.1007/s11092-013-9187-5

Herbert, C., Junghofer, M., & Kissler, J. (2008). Event related potentials to emotional adjectives during reading. *Psychophysiology, 45*(3), 487-498. https://doi.org/10.1111/j.1469-8986.2007.00638.x

Hill, J., Berlin, K., Choate, J., Cravens-Brown, L., McKendrick-Calder, L., & Smith, S. (2021). Exploring the emotional responses of undergraduate students to assessment feedback: Implications for instructors. *Teaching and Learning Inquiry, 9*(1), 294-316. https://doi.org/10.20343/teachlearninqu.9.1.20

Holmes, A. G. (2023). 'I was really upset and it put me off': The emotional impact of assessment feedback on first-year undergraduate students. *Journal of Perspectives in Applied Academic Practice, 11*(2), 62-74. https://doi.org/10.56433/jpaap.v11i2.529

Jonsson, A., & Panadero, E. (2018). Facilitating students' active engagement with feedback. In A. A. Lipnevich & J. K. Smith (Eds.), *The Cambridge handbook of instructional feedback*. Cambridge University Press.

Kissler, J., & Herbert, C. (2013). Emotion, etmnooi, or emitoon? – Faster lexical access to emotional than to neutral words during reading. *Biological Psychology, 92*(3), 464-479. https://doi.org/10.1016/j.biopsycho.2012.09.004

Laudel, H., & Narciss, S. (2023). The effects of internal feedback and self-compassion on the perception of negative feedback and post-feedback learning behavior. *Studies in Educational Evaluation, 77*, 101237. https://doi.org/10.1016/j.stueduc.2023.101237

Lenth, R. V. (2022). *emmeans: Estimated marginal means, aka least-squares means*. https://CRAN.R-project.org/package=emmeans

Lipnevich, A. A., Berg, D. A. G., & Smith, J. K. (2016). Toward a model of student response to feedback. In G. T. L. Brown & L. R. Harris (Eds.), *Human factors and social conditions of assessment* (pp. 169–185). Routledge.

Lipnevich, A. A., & Panadero, E. (2021). A review of feedback models and theories: Descriptions, definitions, and conclusions. *Frontiers in Education, 6*, 720195. https://doi.org/10.3389/feduc.2021.720195

Lowe, T., & Shaw, C. (2019). Student perceptions of the" best" feedback practices: An evaluation of student-led teaching award nominations at a higher education institution. *Teaching & Learning Inquiry, 7*(2), 121-135. https://doi.org/10.20343/teachlearninqu.7.2.8

Molloy, E., Borrell-Carrio, F., & Epstein, R. (2012). The impact of emotions in feedback. In D. Boud & E. Molloy (Eds.), *Feedback in higher and professional education* (pp. 60-81). Routledge. https://doi.org/10.4324/9780203074336

Müller, K., & Wickham, H. (2022). *tibble: Simple data frames*. https://CRAN.R-project.org/package=tibble

Pekrun, R. (2006). The control-value theory of achievement emotions: Assumptions, corollaries, and implications for educational research and practice. *Educational Psychology Review, 18*, 315-341. https://doi.org/10.1007/s10648-006-9029-9

Pitt, E., & Winstone, N. (2018). The impact of anonymous marking on students' perceptions of fairness, feedback and relationships with lecturers. *Assessment & Evaluation in Higher Education, 43*(7), 1183-1193. https://doi.org/10.1080/02602938.2018.1437594

Ponz, A., Montant, M., Liegeois-Chauvel, C., Silva, C., Braun, M., Jacobs, A. M., & Ziegler, J. C. (2014). Emotion processing in words: a test of the neural re-use hypothesis using surface and intracranial EEG. *Social Cognitive and Affective Neuroscience, 9*(5), 619-627. https://doi.org/10.1093/scan/nst034

Price, M., Handley, K., Millar, J., & O'Donovan, B. (2010). Feedback: All that effort, but what is the effect? *Assessment & Evaluation in Higher Education, 35*(3), 277-289. https://doi.org/10.1080/02602930903541007

O'Donovan, B. M., Den Outer, B., Price, M., & Lloyd, A. (2021). What makes good feedback good?. *Studies in Higher Education, 46*(2), 318-329. https://doi.org/10.1080/03075079.2019.1630812

R Core Team. (2022). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Robinson, D., Hayes, A., & Couch, S. (2020). *broom: Convert statistical objects into tidy tibbles*. https://CRAN.R-project.org/package=broom

Rowe, A. D. (2017). Feelings about feedback: The role of emotions in assessment for learning. In D. Carless, S. M. Bridges, C. K. Y. Chan, and R. Glofcheski (Eds.), *Scaling up assessment for learning in higher education* (pp. 159-172). Springer. https://doi.org/10.1007/978-981-10-3045-1_11

Ryan, T., & Henderson, M. (2018). Feeling feedback: Students' emotional responses to educator feedback. *Assessment & Evaluation in Higher Education, 43*(6), 880-892. https://doi.org/10.1080/02602938.2017.1416456

Singmann, H., Bolker, B., Westfall, J., Aust, F., & Ben-Shachar, M. S. (2021). *afex: Analysis of factorial experiments*. https://CRAN.R-project.org/package=afex

Stanley, D. (2021). *apaTables: Create American Psychological Association (APA) style tables*. https://CRAN.R-project.org/package=apaTables

Tyng, C. M., Amin, H. U., Saad, M. N., & Malik, A. S. (2017). The influences of emotion on learning and memory. *Frontiers in Psychology, 8*, 235933. https://doi.org/10.3389/fpsyg.2017.01454

Värlander, S. (2008). The role of students' emotions in formal feedback situations. *Teaching in Higher Hducation, 13*(2), 145-156. https://doi.org/10.1080/13562510801923195

Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis*. Springer-Verlag. https://doi.org/10.1007/978-3-319-24277-4_9

Wickham, H. (2022). *stringr: Simple, consistent wrappers for common string operations*. https://CRAN.R-project.org/package=stringr

Wickham, H., François, R., Henry, L., Müller, K., & Vaughan, D. (2023). *dplyr: A grammar of data manipulation*. https://CRAN.R-project.org/package=dplyr

Wickham, H., Hester, J., & Bryan, J. (2022). *readr: Read rectangular text data*. https://CRAN.R-project.org/package=readr

Wickham, H., Vaughan, D., & Girlich, M. (2023). *tidyr: Tidy messy data*. https://CRAN.R-project.org/package=tidyr

Wilke, C. O. (2020). *cowplot: Streamlined plot theme and plot annotations for 'ggplot2'*. https://CRAN.R-project.org/package=cowplot

Wilson, M.D. (1988). The MRC psycholinguistic database: Machine readable dictionary, version 2. *Behavioral Research Methods, Instruments, & Computers, 20*, 6-10. https://doi.org/10.3758/BF03202594

Winstone, N. E., Nash, R. A., Parker, M., & Rowntree, J. (2017). Supporting learners' agentic engagement with feedback: A systematic review and a taxonomy of recipience processes. *Educational Psychologist, 52*(1), 17-37. https://doi.org/10.1080/00461520.2016.1207538

Yeo, J. Y., & Ting, S. H. (2014). Personal pronouns for student engagement in arts and science lecture introductions. *English for Specific Purposes, 34*, 26-37. https://doi.org/10.1016/j.esp.2013.11.001