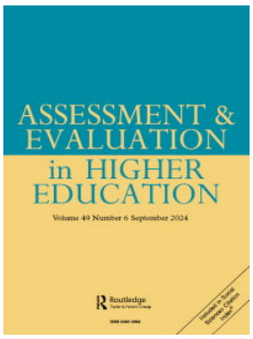

Assessment & Evaluation in Higher Education

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/caeh20

The language of assessment: identifying challenging terminology for students and exploring implications for practice

Harry West, Nicole Skidmore, Laura Bennett, Juliet Eve, Cathy Minett-Smith & Luke Rudge

To cite this article: Harry West, Nicole Skidmore, Laura Bennett, Juliet Eve, Cathy Minett-Smith & Luke Rudge (16 Sep 2024): The language of assessment: identifying challenging terminology for students and exploring implications for practice, Assessment & Evaluation in Higher Education, DOI: 10.1080/02602938.2024.2402960

To link to this article: https://doi.org/10.1080/02602938.2024.2402960

ASSESSMENT & EVALUATION IN HIGHER EDUCATION
https://doi.org/10.1080/02602938.2024.2402960

The language of assessment: identifying challenging terminology for students and exploring implications for practice

Harry West[a], Nicole Skidmore[b], Laura Bennett[c], Juliet Eve[d], Cathy Minett-Smith[e] and Luke Rudge[f]

[a]School of Architecture & Environment, University of the West of England, Bristol, UK; [b]School of Social Sciences, University of the West of England, Bristol, UK; [c]HE Quality, University College Birmingham, Birmingham, UK; [d]Learning & Teaching Hub, University of Brighton, Brighton, UK; [e]College of Business & Law, University of the West of England, Bristol, UK; [f]Bristol Centre for Linguistics, University of the West of England, Bristol, UK

ABSTRACT

Many interventions have been developed by staff to support students in understanding assessment task requirements and summative and formative feedback. However, to various extents, these approaches assume a shared understanding of the, often indeterminate, language used in assessment. This article reports on a project, which aimed to explore students' confidence in interpreting frequently used language in assessment and feedback documentation and explore ways in which staff can work with students to address misunderstanding and misinterpretation. To do this, we undertook a corpus-based linguistic analysis of assessment documentation to identify frequently used and potentially problematic language. We then used a survey and focus groups to explore with students their confidence in interpreting these terms. In doing so, we identify commonly used assessment and feedback terminology, which students find challenging to interpret, potentially resulting in frustration and demotivation. We conclude by exploring different approaches reported by students in the focus groups which help them interpret challenging terminology, notably different forms of dialogic-based intervention, but note, however, that time and space are needed within the curriculum for a shared (or ideally co-created) understanding of language to be developed between students and staff.

KEYWORDS

Assessment terminology; language; feedback; dialogue

Introduction

Universities frequently receive relatively low student satisfaction scores for assessment (Yang and Carless 2013; MacKay et al. 2019) and perceptions of inconsistency are a cause of concern for students (Bloxham et al. 2016). However, a truly standardised assessment system is unlikely, not least due to the often indeterminate nature of many terms used in Higher Education contexts (Morrish and Sauntson 2020). There have been efforts to reduce inconsistency, such as the development of standards-based assessment through publication of marking criteria which are generic

CONTACT Harry West harry.west@uwe.ac.uk
Supplemental data for this article can be accessed online at https://doi.org/10.1080/02602938.2024.2402960.

and student focused (Alonzo et al. 2019) and social calibration amongst staff to highlight differences and form a common consensus of standards (Wyse et al. 2020).

Many practices have been proposed to help students interpret assessment documentation, including supporting students to self-evaluate (Nicol and Kushwah 2023; Gladovic et al. 2024) and generate internal feedback (Nicol 2021), developing glossaries of assessment terminology (Richards and Pilcher 2014), using exemplars (Bell et al. 2013; Worth 2014), adopting dialogic feedback (Hill and West 2020, 2022), or through co-creation of marking criteria (Orsmond et al. 2000; Meer and Chapman 2015) and other forms of student-staff partnership (Matthews et al. 2023a). When students meaningfully engage with and reflect upon assessment documentation, performance is improved, and the benefits of assessment/feedback literacy are realised. Carless (2015) proposed this as new paradigm of feedback that prioritised active student roles in feedback as a dialogic process. Winstone and Carless (2019) expand on this and highlight three student-focussed elements: sense-making, learner generation of inputs and emphasis on student action. However, many of these practices assume a common understanding of the language used in assessment and feedback. This is critical to the success of many of these interventions, especially those involving dialogue (Matthews et al. 2023b).

Research has identified that students' confidence in interpreting assessment and feedback language can vary. There are elements of criteria and standards which are socially constructed, and interpretation will differ between individuals and groups (Ajjawi and Bearman 2018). There are terms and phrases used which students struggle to interpret, such as 'synthesise, 'analyse', 'evaluate' and 'discuss' (O'Donovan et al. 2004; Richards and Pilcher 2014). Boyle et al. (2020) also identify that language inferring the quality of a piece of work, such as 'good' or 'better', can be difficult for students to distinguish between, given inherently subjective or context-dependent interpretations. Challenges also arise where familiar terminology is used in unfamiliar assessment-related contexts, for example 'meeting' learning outcomes (Butcher et al. 2017).

O'Donovan et al. (2004) distinguish between explicit and tacit knowledge in interpreting assessment and feedback terminology. Students find it challenging to decode tacit assessment requirements (Boyle et al. 2020) such as 'developing an argument' (Butcher et al. 2017), using 'the literature' (Matshedisho 2020), or being 'critical' (Richards and Pilcher 2014). Similar statements which often require such tacit understanding are used in feedback such as 'too descriptive', 'superficial analysis' or 'fails to answer the question' (Adcroft 2010), contributing to the common misalignment between the message from the feedback provider and receiver (Jones and Ellison 2021).

Different demographic groups of students are more likely to be experience challenges with the language of assessment as interpretation takes place through cultural and linguistic lenses (Rossiter 2023). A lack of clarity in assessment criteria has a disproportionate effect on already disadvantaged groups of students (Balloo et al. 2018), such as widening participation students (Butcher et al. 2017). Gonsalves (2023) found that international students (for whom English may be not their first language) can struggle with interpretation of assessment rubrics and that consideration is needed in how criteria are presented and used within task to enhance accessibility.

Whilst there are practices that help to minimise inconsistencies, intuitive and tacit variability is unavoidable (Yorke 2011; Bloxham et al. 2016). It is, therefore, important that staff develop an understanding of which frequently used language students struggle to understand, potentially limiting their engagement with assessment tasks, inhibiting their development of assessment literacy, and undermining their agency as learners (Lee et al. 2024). Understanding which student groups are likely to experience challenges in interpreting assessment terminology due to social and cultural differences (Rossiter 2023) is also timely to address awarding gaps. Being aware of the potentially ambiguous language of assessment and developing strategies within our teaching to acknowledge and address it underpin the development of inclusive learning environments, which enable students to succeed.

Study context

This article reflects on work undertaken over the 2020–2021 and 2021–2022 academic years to investigate students' confidence in interpreting assessment and feedback language and to identify ways in which staff can work with students to reduce feelings of uncertainty and clarify meaning. This work forms part of a wider QAA (Quality Assurance Agency) Collaborative Enhancement Project. The project was undertaken across four teaching-oriented universities based in central and southern England. This research was approved by all four institutional research ethics committees. The project team was comprised of staff representatives (including academic staff developers, professional services, and academics) and student partners.

The present study is split into three stages, which combined represent a mixed-methods analysis of assessment terminology and associated student understanding. Stage 1 utilised corpus linguistic analysis to identify common assessment language from a range of documentation. This informed Stages 2 and 3 which explored students' confidence in interpreting the identified terminology, and ways staff can minimise confusion and misinterpretation through a survey and focus groups. The overall aim was to identify problematic terminology used in assessment so that staff can reflect on when and where certain terminology is used in the assessment process and support students to clarify meaning and enhance assessment/feedback literacy.

Methods

Stage 1: linguistic analysis of assessment documentation

The project aim was to determine the extent to which there is a variation in students' confidence in interpretation of language used in assessment documentation. However, this could only be completed in a robust manner by initially following sound linguistic methodology to identify common instances of language use. Each institution performed equivalent corpus analyses of their analogous assessment documentation, with the goal being to apply corpus-based linguistic methodologies to query the language used (McEnery and Hardie 2012). This provided insight into identifying common terms and phrases relating to assessment, marking and regulation, and the level of performance. Each institution complied documentation from across seventeen subject areas, mapped to the UK HECoS CAH1 codes (the Common Aggregation Hierarchy codes provide a standardised approach to grouping subjects/disciplines across the higher education sector in the UK).

The documentation was compiled and analysed using Sketch Engine (Kilgarriff et al. 2014), made available via the EU Horizon 2020-funded ELEXIS project. Each institution complied one main corpus (a total of four across the project), and each also consisted of four subcorpora: 1) assessment briefs, 2) learning outcomes, 3) marking criteria and grade descriptors, and 4) assessment policy documentation. The smallest of the four institutional corpora was 84,224 words and the largest 646,776 words. The combined corpora totalled 1,390,676 words.

First, a word frequency analysis of common parts of speech (i.e. nouns, verbs, adjectives, and adverbs) was undertaken to identify recurring vocabulary. Word frequencies were normalised by corpus length to ensure comparability across corpora. Second, keyword analysis identified statistically salient single and multi-word terms relative to a corpus of everyday British English usage: the English Web 2020 corpus. Manual confirmation was employed to remove false positives and to investigate how the identified vocabulary is used in context via collocational analysis.

The results were then compared to identify lexical patterns, discrepancies and the terminology that would feed into Stage 2. The top 100 words/phrases from each institution were compiled and duplicates were removed. The project team systematically went through this list, first individually and then collaboratively to define a list which was categorised into 'assessment language' that provides instruction to students (e.g. define, analyse, explain), 'feedback language'

which would be used in formative or summative feedback to students (e.g. superficial analysis, descriptive writing, relevance), and 'quality descriptors' (e.g. good, excellent, poor). Examples of concordance (Gablasova et al. 2017) were also extracted and can be found in Supplementary Materials for this article.

Stage 2: student survey and focus groups

Stage 2 of this research sought to explore students' confidence in interpreting the assessment language identified in Stage 1. First, an online survey was launched across the four institutions; distributed via institutional academic development units and professional services departments to undergraduate module/unit leaders who were asked to promote the survey to students. Prior to participating, students were given an overview of the project and asked to provide informed consent.

Student respondents were presented with Likert scale questions asking them to indicate on a scale of 1–6 their confidence in interpreting the word or phrase (where 1 was not at all confident in interpretation, and 6 was certain in understanding). Demographic information was also recorded (Table S1).

In total, there were 318 survey responses across the four institutions. Most respondents were below the age of 24, did not report a disability, stated English as their first language and reported their ethnicity as white. However, there was a greater range of respondents by year of study, with a relatively even split across FHEQ Levels 4–6 (Years 1–3 of full-time undergraduate study in England).

The Likert-scale responses were analysed, with the words and phrases ranked according to the average confidence score across the four-partner institutions. Additionally, we compared the responses between different demographic groups using Kruskal–Wallis statistical tests for the word groups individually, and for a combined set of responses.

Stage 3: Focus groups and thematic analysis

Focus groups and one-to-one interviews with 31 students were conducted using Microsoft Teams and were recorded and transcribed using automatic transcription software. Reflexive thematic analysis (Braun and Clarke 2019, 2021) was used to analyse the transcripts and identify themes, with the aim of capturing patterns in participants' accounts and identifying areas in which language used in assessment was confusing, unclear, or unhelpful, or where the opposite was true. Codes were clustered together at each institution, and these were later shared with an additional team member so that they could determine whether patterns were evident across the dataset. Finally, representatives from each institution discussed the final themes in detail and ensured that each theme cohered around a central organizing concept, which is the key idea that underpins thematic explanation of data (Terry et al. 2017), and that theme names captured these concepts. To keep participants in the focus groups anonymous, quotations in this paper will not report any identifiable information. To aid readability and comprehension, verbal nods (e.g. pauses and repeated words) have been removed from the data: '[…]' indicates omitted data.

Results

Stage 1: linguistic analysis of assessment documentation

Initially, we identified the level of overlap in the terms used in the assessment documentation. The top 100 nouns, verbs, adjectives, and adverbs were extracted from each institutions corpus and duplicates removed. On average there were 173 individual terms and phrases from a

potential total of 400 (100 per institution), representing a 43.4% overlap in assessment terminology across the four institutions. A total of 54 individual terms/phrases were selected to take forward. These terms had a high frequency of use across the institutional corpora and were categorised as 'assessment language', 'feedback language' or 'quality descriptors'. The results from this analysis and examples of concordance can be found in the Supplemental Material.

Stage 2: student survey results

This section presents survey results where respondents were asked to rate their confidence in interpreting the terminology identified in Stage 1. The average interpretation score and variance are presented for each institution and from the combined set of results.

Figure 1 shows the average interpretation score and variance for the 'assessment' focused language. Several terms received low interpretation scores such as 'critical' and 'examine'. In particular, the term 'synthesise' received very low scores. These terms also had high variance indicating there is greater variability in students' ability to interpret these terms when completing an assessment task. Conversely language such as 'explain' and 'summarise' had higher average scores and lower variance indicating that these are terms students interpret with greater confidence.

Figure 2 presents the interpretations scores and variance for the feedback terms. The terminology which received lower scores and high variance include 'superficial analysis' and 'abstract'. Other low-scoring terms include 'reflective' and 'application of theory'. These also had high variance suggesting greater variability in students' confidence in understanding this language when used in feedback. Conversely, terms/phrases with higher confidence and lower variance (i.e. where students consistently feel confident in their understanding) include 'successfully', 'fails to answer the question' and 'relevance'.

Across the three terminology groups, the quality descriptors had higher levels of student confidence in interpretation (Figure 3). The quality descriptors with higher levels of confidence and lower variance include 'excellent' and the adverb 'extremely'. The descriptor 'competent' had the lowest average score and highest average variance.

Student Average Score (Assessment Words)

		Institution			
Word	Average	U1	U2	U3	U4
Synthesise	3.175	2.926	3.023	3.636	3.114
Critical	4.250	4.322	3.942	4.485	4.250
Examine	4.273	4.281	4.105	4.478	4.227
Inform	4.301	4.388	3.930	4.319	4.568
Assess	4.380	4.397	4.267	4.333	4.523
Explore	4.411	4.463	4.212	4.551	4.419
Interpret	4.429	4.322	4.163	4.696	4.535
Review	4.434	4.355	4.372	4.464	4.545
Carefully	4.445	4.331	4.337	4.565	4.545
Reflect	4.476	4.554	4.372	4.275	4.705
Develop	4.508	4.579	4.282	4.397	4.773
Meet	4.650	4.636	4.407	4.721	4.837
Evaluate	4.670	4.653	4.424	4.899	4.705
Analyse	4.724	4.702	4.605	4.884	4.705
Demonstrate	4.738	4.694	4.616	4.913	4.727
Discuss	4.843	4.752	4.733	5.000	4.886
Consider	4.869	4.702	4.791	4.868	5.114
Reference	5.099	5.157	4.860	5.014	5.364
Summarise	5.036	5.000	4.837	5.101	5.205
Define	5.067	5.017	5.058	5.217	4.977
Explain	5.119	5.008	5.023	5.217	5.227
Average	4.566	4.535	4.398	4.668	4.664

Variance of Student Score (Assessment Words)

		Institution			
Word	Average	U1	U2	U3	U4
Explain	1.076	1.042	0.988	1.026	1.249
Summarise	1.213	1.017	1.079	1.240	1.515
Reference	1.246	1.167	1.392	1.397	1.027
Consider	1.260	1.361	1.297	1.161	1.219
Discuss	1.305	1.188	1.469	0.971	1.591
Analyse	1.323	1.127	1.160	1.163	1.841
Demonstrate	1.402	1.397	1.133	1.198	1.877
Define	1.475	1.733	1.232	0.820	2.116
Develop	1.492	1.379	1.705	1.497	1.389
Assess	1.512	1.325	1.445	1.255	2.023
Evaluate	1.547	1.612	1.628	1.387	1.562
Explore	1.634	1.417	1.502	1.369	2.249
Review	1.644	1.564	1.578	1.458	1.975
Reflect	1.659	1.666	1.789	1.526	1.655
Examine	1.785	1.587	1.295	1.753	2.505
Interpret	1.824	1.487	1.456	1.862	2.493
Inform	1.879	1.723	1.524	1.926	2.344
Carefully	1.952	1.690	1.920	1.573	2.626
Meet	1.991	1.783	2.244	1.846	2.092
Critical	1.994	1.854	1.938	1.806	2.378
Synthesise	2.930	2.369	2.799	2.912	3.638
Average	1.626	1.499	1.551	1.483	1.970

Figure 1. Average student interpretation score and variance for assessment-related terminology. Terms are listed in ascending order based on the overall average value across the four institutions.

Student Average Score (Feedback Words)

Word	Average	Institution U1	U2	U3	U4
Superficial analysis	3.538	3.438	3.349	3.706	3.659
Abstract	3.906	3.967	3.547	4.090	4.023
Reflective	4.302	4.364	4.163	4.206	4.477
Application of theory	4.391	4.322	4.116	4.647	4.477
Concrete	4.526	4.661	4.093	4.735	4.614
Professional	4.628	4.587	4.360	4.676	4.886
Academic	4.647	4.711	4.256	4.826	4.795
Logical	4.675	4.760	4.176	4.809	4.955
The literature	4.682	4.934	4.337	4.618	4.841
Suitably	4.717	4.711	4.407	4.681	5.068
Properly	4.730	4.669	4.407	4.797	5.045
Indicate	4.735	4.727	4.384	4.899	4.930
Effective	4.742	4.744	4.453	4.794	4.977
Identification of key concepts	4.823	4.860	4.535	4.899	5.000
Descriptive	4.840	4.876	4.628	4.881	4.977
Vague	4.889	4.926	4.812	5.000	4.818
Wide	4.908	5.033	4.706	4.986	4.909
Appropriate	4.923	4.901	4.547	5.130	5.114
Relevance	4.942	5.017	4.616	5.043	5.091
Fails to answer the question	5.160	5.281	4.882	5.319	5.159
Successfully	5.208	5.355	4.965	5.217	5.295
Average	4.663	4.707	4.369	4.760	4.815

Variance of Student Score (Feedback Words)

Word	Average	Institution U1	U2	U3	U4
Successfully	1.167	0.898	1.375	1.114	1.283
Fails to answer the question	1.512	1.304	1.748	1.044	1.951
Identification of key concepts	1.515	1.222	1.593	1.387	1.860
Relevance	1.518	1.316	1.627	1.277	1.852
Indicate	1.526	1.433	1.416	1.475	1.781
Appropriate	1.555	1.373	1.686	1.056	2.103
Effective	1.615	1.359	1.733	1.300	2.069
Suitably	1.623	1.524	1.727	1.779	1.460
Logical	1.658	1.650	1.623	1.500	1.858
Wide	1.695	1.449	1.639	1.514	2.178
Descriptive	1.751	1.960	1.601	1.652	1.790
Academic	1.778	1.591	1.910	1.352	2.260
Application of theory	1.780	1.670	1.727	1.515	2.209
Vague	1.823	1.986	1.583	1.618	2.106
Properly	1.825	1.606	1.844	1.664	2.184
The literature	1.845	1.662	1.897	1.822	1.997
Professional	1.924	1.611	1.904	1.983	2.196
Concrete	1.929	1.426	2.085	1.869	2.336
Reflective	2.019	1.717	2.020	1.897	2.441
Abstract	2.186	1.932	1.898	2.052	2.860
Superficial analysis	2.776	2.548	2.583	2.629	3.346
Average	1.763	1.583	1.772	1.595	2.101

Figure 2. As per Figure 1, however, for feedback-related terms.

Student Average Score (Quality Descriptors)

Word	Average	Institution U1	U2	U3	U4
Competent	4.431	4.537	4.035	4.493	4.659
Adequate	4.776	4.686		4.870	4.773
Original	4.915	5.066	4.581	4.942	5.070
Satisfactory	5.001	5.074	4.721	5.116	5.091
Inadequate	5.015	5.149	4.744	5.145	5.023
Weak	5.163	5.355	4.930	5.116	5.250
Good	5.244	5.298	5.116	5.290	5.273
Poor	5.290	5.380	4.965	5.348	5.465
Exceptional	5.325	5.512	5.047	5.333	5.409
Outstanding	5.345	5.421	5.163	5.362	5.432
Extremely	5.358	5.479	5.012	5.420	5.523
Excellent	5.418	5.603	5.163	5.406	5.500
Average	5.114	5.213	4.862	5.153	5.206

Variance of Student Score (Quality Descriptors)

Word	Average	Institution U1	U2	U3	U4
Extremely	1.056	0.885	1.470	0.777	1.092
Excellent	1.108	0.725	1.385	0.951	1.372
Outstanding	1.179	1.013	1.197	0.999	1.507
Poor	1.244	1.154	1.658	1.054	1.112
Good	1.349	1.311	1.233	1.209	1.645
Exceptional	1.352	0.919	1.715	0.990	1.782
Weak	1.581	1.114	1.666	1.398	2.145
Adequate	1.996	2.001		1.762	2.226
Satisfactory	1.711	1.786	1.756	1.310	1.992
Original	1.729	1.546	2.011	1.438	1.924
Inadequate	1.783	1.628	1.722	1.479	2.302
Competent	2.088	1.784	1.963	2.048	2.555
Average	1.504	1.322	1.616	1.284	1.805

Figure 3. As per Figure 1, however, for quality descriptors. NB no data were recorded for the term adequate for university 2 (U2).

Table 1. Kruskal–Wallis test results for all words/phases in the survey. PNS abbreviates 'prefer not to say' responses.

	Age		Disability		First Language		Ethnicity		Level	
Significance (*p* value)	0		0		0.513		0		0	
Groups (*z* scores)	<20	−8.67	Yes	−8.68	Yes	−0.78	Asian	−7.96	L4	−7.18
	21–24	4.74	No	7.42	No	0.93	Black	5.02	L5	−10.09
	25–29	5.54	PNS	0.14	PNS	−0.73	Mixed	0.38	L6	17.94
	30+	3.81					Other	4.16		
	PNS	−4.82					PNS	−4.42		
							White	2.23		

Table 1 presents the statistical comparison between demographic groups based on all terms in the survey. Significant differences were found between demographic categories for age, disability, ethnicity, and level of study. Students below the age of 20, reported a learning disability, reported their ethnicity as Asian or were in Level 4 or 5, statistically gave lower confidence scores in their ability to interpret the terminology. Students who were older, did not have a disability, were white/black or in their final (Level 6) year of study, reported higher confidence.

Table 2 presents the comparative analysis between demographic groups for the assessment-related terms in the survey. Significant differences were found for all demographic

Table 2. Kruskal–Wallis Test results for assessment-related terms. PNS abbreviates 'prefer not to say' responses.

	Age		Disability		First language		Ethnicity		Level	
Significance (*p* value)	0		0		0.024		0		0	
Groups (*z* scores)	<20	−5.82	Yes	−10.35	Yes	−2.72	Asian	−1.88	L4	−2.49
	21–24	1.68	No	8.33	No	2.7	Black	5.49	L5	−8.72
	25–29	3.9	PNS	1.05	PNS	0.22	Mixed	−0.13	L6	11.59
	30+	3.94					Other	1.91		
	PNS	−1.99					PNS	−0.97		
							White	−2.17		

Table 3. Kruskal–Wallis Test results for feedback-related terms. PNS abbreviates 'prefer not to say' responses.

	Age		Disability		First language		Ethnicity		Level	
Significance (*p* value)	0		0.004		0.781		0		0	
Groups (*z* scores)	<20	−7.01	Yes	−2.9	Yes	0.4	Asian	−5.03	L4	−7.05
	21–24	4.81	No	3.31	No	−0.28	Black	2.32	L5	−4.67
	25–29	3.58	PNS	−1.42	PNS	−0.63	Mixed	−0.33	L6	12.23
	30+	2.52					Other	2.42		
	PNS	−4.18					PNS	−3.96		
							White	2.51		

Table 4. Kruskal–Wallis Test results for quality descriptors. PNS abbreviates 'prefer not to say' respondents.

	Age		Disability		First language		Ethnicity		Level	
Significance (*p* value)	0.015		0.301		0.171		0		0	
Groups (*z* scores)	<20	−1.62	Yes	−1.53	Yes	1.31	Asian	−7.51	L4	−2.43
	21–24	1.95	No	1.12	No	−1.04	Black	0.64	L5	−4.15
	25–29	1.61	PNS	0.36	PNS	−1.52	Mixed	1.11	L6	6.83
	30+	−0.52					Other	3.27		
	PNS	−2.44					PNS	−3.4		
							White	3.94		

groups in line with the above description for Table 1, with the addition of those who report English as being their first language as having lower confidence in interpreting the assessment terminology.

Tables 3 and 4 show the statistical test results for the feedback and quality terminology. Similar demographic patterns as previously described are present; however, first language groups show no significant differences, and learning disability groups also show no significant difference for the quality descriptor group.

Stage 3: focus groups and thematic analysis

The thematic analysis led to the construction of three themes: challenging language; assessment criteria/feedback; and helpful staff interventions. Each of these themes will be discussed below, drawing on illustrative examples from the data.

Theme one: challenging language

Students expressed frustration with challenging and unclear language, with the words 'critical(ly)', 'assess', 'synthesise', and 'superficial' frequently noted as points of contention. This language arose in marking criteria, assessment questions, assessment guidance, and in feedback, meaning that for students it felt like there was no escaping these confusions. Students specifically discussed how language in assessments can be vague or ambiguous, which make expectations unclear. Students also referred to the relatability of the language used in

assessment and feedback and how it can be challenging when they cannot easily relate to certain terms:

> *So it's like "critical thinking". It's like throughout all of the assignments […]. But I think it's a bit hard to really use it in the right way. So even though I have now finished the two semesters […] if you ask me what is "critical thinking"? I mean, it's stupid hard to answer what it really is.*

> *For me, we've done a lot on critical writing, so we've had it explained […] I guess it would be having those words explained in a way that you can sort of relate to.*

Students also explained there are times when individual terms themselves are understandable; however, the unfamiliar or challenging context in which the term is used (sometimes alongside multiple other challenging terms) can make it hard for them to understand expectations. This can lead to students feeling overwhelmed and frustrated by the task and needing to re-read and reflect on the instruction:

> *It's not just […] one word. It's the sentence in which it's written. Often you will read it and you think, I don't know what that means straight away, you sort of have to read it a few times […]. So, for example, one of ours is "evidence of a highly structured approach with key information effectively identified, communicated and comprehensively analysed". So that sentence, you read that and don't think […] I know what that means.*

> *I don't think you need to make the question as complicated and wordy and long as possible. I don't think that's part of the test […], understanding the question shouldn't be part of your assessment.*

> *I didn't know how to assemble it, and I didn't know what was relevant […] so that was extremely frustrating for me.*

Students also highlighted how diverse groups of students, notably new (Year 1) students, or students for whom English is not their first language, might be more likely to encounter challenges in interpreting assessment terminology:

> *Within [assessment] tools, a lot of the language can be ambiguous, I think, particularly to a newcomer.*

> *I know there have been other students on the course where English isn't their first language, and they've particularly found sentences like that [a marking criteria example] difficult to understand.*

Theme two: assessment criteria and feedback

In our survey, the quality descriptor terms such as good or excellent, had high confidence scores. However, in the focus groups the students explained that whilst the individual descriptors themselves are not necessarily problematic to understand, there is a challenge for them in distinguishing between marking levels and how to move between them, especially when the explanatory criteria to achieve a certain level contains challenging terminology:

> *A 70 and 80, is for example, "good research", "excellent research", that's the difference. And I have no idea what that means […] how can I go from good to excellent then?*

> *I think the language itself […] is easy to understand. They don't use over complicated language […] but if that tool is to help me […] know what I need to do to go from 70 to 80, I don't think it does that.*

In common with previous studies our focus group data suggested that emotions can play a negative role and deter students from engaging with, conventionally written, feedback (Shields 2015, Hill et al. 2021a). This suggests that not only do staff need to pay attention to the language used in feedback, but also that there is a role to play in supporting student resilience to deal with the emotional aspects of assessment and feedback, i.e. to manage affect (Carless and Boud 2018):

> *The feedback I received on some work I didn't do as well at, was just praising how bad it was. That's just like, I'm not gonna want to do any more otherwise, that's the reason I've enjoyed my [other] module a lot more, because they were nice about it.*

Finally, students also explained how the use of challenging terminology as part of their feedback limited their ability to apply summative feedback formatively for future assessment tasks, limiting their ability to learn from mistakes and improve in the future:

> *It's not clear to understand what it is you can do with that feedback for the next steps.*

Theme three: helpful interventions

This final theme explores how students have overcome challenges in interpreting the language used in assessment and feedback. After receiving written summative feedback students explained that they frequently would need to meet with their tutor to understand the comments which have been provided and how to take feedback forward to the next assessment:

> *I know a lot of people on the course […] have had to speak with whoever's marked it to go over the feedback because it doesn't often make sense.*

> *That's why it's important to get that time with your personal tutor, because if you don't understand the feedback from the comments […], if you don't take the time to fully understand what they are saying, you will keep making mistakes.*

Clearly being able to engage with the marker and hear their rationale and explanation of terms helped the students engage with their feedback. In the focus groups one student also noted an occasion where audio feedback was provided which also helped meet this need:

> *There had been one occasion that I have been able to hear the person that was grading my paper, they gave me like an audio for the whole thing. I still think that she graded me a bit harsh but like, at least it's like peace of mind, you know. So at least it gives you more to work on.*

In a formative context students noted occasions where they have been able to submit drafts of their work to tutors and receive formative feedback as being helpful. These formative opportunities allowed students to check understanding of task requirements and the language used in the assessment documentation. Alongside submission of draft work, students spoke of how specific sessions and activities helped them to understand assessment terminology and expectations prior to submission. Students explained how they feel it is often assumed that they would understand terminology and assessment requirements, and when they did not, they may feel embarrassed to voice their confusion. However, having support built into modules whilst completing a task, such as lectures, recorded videos, or drop-in sessions, helped students to understand terminology and assessment requirements and feel less awkward about asking questions:

> *I like the approach of one of my module leaders, because she is recording this short video of going through the assignment brief, and then she is adding these little annotations […] And I think it's great, because it's giving us some ideas of how to approach the assignment. And it may help us to notice such things in the future when approaching other ones.*

> *I think for me it would be helpful if [staff] went over the marking criteria and those words were explained, and then it gives […] students the chance to question that if they're still not clear on certain words that come up.*

Conclusions and implications for practice

This project sought to identify commonly used terminology in assessment and feedback documentation and how students' confidence in interpretation might vary. Previous research has emphasized there are many assessment-related terms which students find challenging (O'Donovan et al. 2004; Richards and Pilcher 2014; Butcher et al. 2017; Boyle et al. 2020). Building on this, Table 5 highlights commonly used assessment and feedback terms where students reported low confidence in interpretation. Conversely, Table 6 notes words where students have higher

Table 5. Terms where more than 25% of respondents scored a term <3 (low levels of confidence).

Synthesise	Carefully
Superficial analysis	Examine
Abstract	Inform
Critical	Interpret
Reflective	Competent
Application of theory	Assess

Table 6. Terms where more than 70% of respondents scored a term >5 (high levels of confidence).

Excellent	Weak
Exceptional	Define
Extremely	Inadequate
Outstanding	Explain
Good	Reference
Poor	Summarise
Fails to answer the question	Satisfactory
Successfully	

confidence in interpretation. These tables reveal a distinction in instructional terms with parallels to Bloom's Taxonomy (Bloom et al. 1956), where students are more familiar with instructions where they need to fundamentally demonstrate knowledge, compared to terms such as critical, interpret, and assess where students need to apply knowledge and critically evaluate and reflect.

As highlighted in the focus groups, the more challenging terms (Table 5) might be used in a range of assessment and feedback contexts, such as assessment briefs, written formative or summative feedback, or marking criteria. As these are multi-faceted and common terms, staff should take steps to engage students in dialogue about their meaning and encourage exploration of the terms to promote understanding through, for example, use of exemplars or an individual or class discussion.

Our statistical analysis shows that different demographic groups might particularly find interpreting language difficult. We found that age and level of study (which in many cases might be related, however not exclusively) show notable differences; with younger students and those in lower years having lower levels of confidence in assessment language interpretation. To some extent, this is to be expected. As students' progress through their undergraduate studies, they will experience the assessment cycle frequently and therefore, where feedback is clear, understandable, and actionable for future tasks (the importance of which was noted in the focus groups), performance and assessment and feedback literacy should improve over time.

We also found other significant demographic differences. For example, across all three terminology groups those who identified as of Asian ethnicity reported lower confidence scores than other ethnic groups. Students who identified as of white or black ethnicity had more variance in their responses. For example, white students had high confidence in interpreting the feedback terminology, but lower confidence in interpreting assessment instructional terms. Black students however reported high confidence in the assessment and feedback groups, but relatively lower confidence in the quality descriptor group. Students who identified as disabled consistently reported lower confidence than students who did not. These differences highlight the complexity involved in understanding language, and how groups of students will variably experience this challenge which may contribute to awarding gaps.

In the focus groups, students reinforced that undertaking summative assessment is a high-stakes and emotional experience (Dowden et al. 2013), and receiving feedback can trigger both positive and negative emotions (Shields 2015; Winstone et al. 2017). Despite the perceived importance of feedback, there are situations where it might have negligible impact on student learning (Carless and Boud 2018). Our research demonstrates that where students struggle to interpret their feedback because of the terminology used, especially when in written form (Hill et al. 2021a), frustration and demotivation can occur.

Challenges clearly exist related to students' confidence in the interpretation of assessment and feedback terminology, which if not addressed will inhibit their ability to improve within and between assessment tasks and develop feedback literacy. Our students described interventions which have been adopted by staff which have helped them understand the terminology used in briefs and marking criteria. Importantly, students explained how these approaches enabled their understanding of feedback, providing a sense of closure on a task with actions for future improvement.

Students appreciated occasions where staff created space within teaching sessions to work through and explain terminology and marking criteria. They also referred to where audio recordings of their tutor have been used to introduce a task and define expectations, but also for summative feedback. Being able to hear their tutor helped the students understand their task/feedback, allowing them to develop as learners. This aligns with previous research which has explored the benefits of audio recorded feedback (Lunt and Curran 2010; Voelkel and Mello 2014); however, it is important to acknowledge that audio-based feedback might not suit all students with differing learning styles (Sarcona et al. 2020) and so offering students flexibility or a combination of approaches might be needed.

Students in the focus groups explained the need to often speak with tutors after work has been marked to decode their written feedback comments and the terminology used within. Dialogue has proven to be a successful means of ensuring there is a shared understanding, as meaning is socially constructed and co-created between student and tutor (Nicol 2010; Boud and Molloy 2013), building trust and mutual understanding (Matthews et al. 2023b). It also allows for the foregrounding of emotions and through discussion a student-staff partnership is formed to positively work with the range of emotions assessment and feedback generates (Hill et al. 2021b), such as frustration or demotivation due to not understanding terminology. Hill and West (2020) demonstrate how dialogue can be implemented in a formative context, reducing the high-stakes nature associated with summative assessment. Such an approach might be effective to help students clarify the meaning of challenging terminology. However, this needs to be effectively designed into the student learning journey, and this approach is not without challenges related to engagement and resourcing (Hill and West 2022).

It is important for staff to recognise the barriers students might face in decoding language and consider how different pedagogic approaches might be used to support student learning. The above pedagogic interventions have a common theme around staff creating the time and space within task and/or post-task to ensure that terminology and meaning is clarified, and a shared (or co-created) understanding exists. We suggest that creating the time and space for these opportunities is a form of compassionate pedagogy where staff and students share uncertainty and vulnerability around language (Arai and Tepylo 2016), creating an affective and motivational learning experience (Jazaieri 2018). In other words, the opposite of a situation where a student independently cannot interpret assessment terminology, resulting in demotivation and an isolated learning experience.

It might be possible to implement a scaffolded approach to assessment and feedback interventions. For example, in the early years of study adopting a tutor-led dialogic approach, where instructional terms requiring higher level academic skills (such as criticality or reflectivity) are discussed and clarified, hopefully helping with the significantly lower confidence in understanding reported in this research amongst Level 4 students. This can then progress to more autonomous student-led approaches in the higher years as a student's learning and assessment/feedback literacy improves over time. Future research might explore how different support might be used in combination and implemented in a scaffolded, longitudinal model.

However, it is important to note that, as implied above, additional space and time is needed to support students in decoding assessment terminology. This ultimately has staff resourcing implications. Creating opportunities for socially constructive and meaningful learning can be challenging under the cumulative pressures of the contemporary neo-liberal higher education

academy, where increasingly performance-based metrics and competing demands can result in 'pedagogic frailty' where traditional assessment practices are maintained which do not fully support student learning and assessment/feedback literacy (Kinchin et al. 2016).

One area which unfortunately, it was not possible to explore in this research was to what extent there might be disciplinary differences in students' understanding of assessment terminology. Disciplines define ways of thinking and practicing (Hounsell and Anderson 2009). There are discipline-specific concepts (Meyer and Land 2003) and signature pedagogies (Shulman 2005), resulting in members of a discipline sharing a common language (Healey et al. 2023). Such disciplinary differences may result in students encountering specialist terminology, or interpreting the language identified in this research through their disciplinary lens resulting in diverse meaning across subjects.

Future studies might also continue to explore the differences in understanding of assessment and feedback language between students and staff, and between faculty. Understanding these differences is needed if challenges around assessment and feedback terminology (and the further issues they contribute to, such as awarding gaps) are to be addressed.

To conclude, this research has undertaken a broad exploration of how students' confidence in interpreting assessment and feedback terminology might differ for individual terms and phases and between different student groups. Our results show that there are several commonly used terms which might prove problematic for students such as 'synthesise', 'assess', or being 'critical' or 'reflective' in their writing. We also explore different pedagogic interventions that allow tutors to work with their students to develop a shared understanding within and post-task, which used effectively can act as pivotal moments for students in developing assessment/feedback literacy, but we note such practices require the time and space within the curriculum and staff workloads for delivery.

Acknowledgements

The authors would like to acknowledge the team who have contributed to the success of the language of assessment project:

- University of the West of England: Laura Bennett, Cathy Minett-Smith, Jackie Chelin, Anne Harbin, Sabine Hassler, Nhi Ho, Judith Ritchie, Luke Rudge, Nicole Skidmore, Harry West, Mihaela Bachvarova.
- University of Brighton: Lucy Chilvers, Juliet Eve, Jason Preston.
- University of Greenwich: Catherine Emeordi-Okeke, Rachel George, Nikki Makinwa, Gemma Mansi, Jennifer Marie, Robert Wilson.
- University of Hertfordshire: Marvelle Brown, Dorothe Diensthuber, Marjolein Groefsema.

The project team would also like to thank the QAA for their funding of the Collaborate Enhancement Project and their support with the dissemination of the work and development of the online resources linked above.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Notes on contributors

Harry West is a Senior Lecturer in Geography at the University of the West of England. He actively researches higher education pedagogies and is on the Editorial Boards of the Journal of Geography in HE and the International Journal for Students as Partners. Harry worked as part of the linguistics analysis and survey teams on this project.

Nicole Skidmore is a PhD student at the University of the West of England. She was one of the student partners working on this collaborative project. Nicole played a key role in the focus group aspect of the project.

Laura Bennett is based at University College Birmingham (and previously was Associate Director of Academic Practice the University of the West of England). Laura co-led this project.

Juliet Eve is the Head of the Learning & Teaching Hub at the University of Brighton. Juliet led Brighton's involvement in this collaborative project.

Cathy Minett-Smith is a Dean of Learning and Teaching at the University of the West of England. Cathy co-led this project.

Luke Rudge is a Visiting Fellow with the Bristol Centre for Linguistics. Luke was actively involved in the initial stages of this project helping develop the methodological approach.

ORCID

Harry West http://orcid.org/0000-0002-2704-5474

References

Adcroft, A. 2010. "Speaking the Same Language? Perceptions of Feedback Amongst Academic Staff and Students in a School of Law." *The Law Teacher* 44 (3): 250–266. doi:10.1080/03069400.2010.524029.

Ajjawi, R., and M. Bearman. 2018. "Problematising Standards: Representation or Performance?" In *Developing Evaluative Judgement in Higher Education: Assessment for Knowing and Producing Quality Work*, edited by Boud, D., Ajjawi, R., Dawson, P. and Tai, J. Routledge: Oxon.

Alonzo, D., N. Mirriahi, and C. Davison. 2019. "The Standards for Academics' Standards-Based Assessment Practices." *Assessment & Evaluation in Higher Education* 44 (4): 636–652. doi:10.1080/02602938.2018.1521373.

Arai, S. M., and H. Tepylo. 2016. "Compassionate Pedagogy for Reflexive Community Practices." In *Community Development: Applications for Leisure, Sport and Tourism*, edited by Shape, E., Mair, H. and Yuen, F. State College, PA: Venture Publishing.

Balloo, K., C. Evans, A. Hughes, X. Zhu, and N. Winstone. 2018. "Transparency Isn't Spoon-Feeding: How a Transformative Approach to the Use of Explicit Assessment Criteria Can Support Student Self-Regulation." *Frontiers in Education* 3: 69. doi:10.3389/feduc.2018.00069.

Bell, A., R. Mladenovic, and M. Price. 2013. "Students' Perceptions on the Usefulness of Marking Guides, Grade Descriptors and Annotated Exemplars." *Assessment & Evaluation in Higher Education* 38 (7): 769–788. doi:10.1080/02602938.2012.714738.

Bloom, B. S., M. D. Engelhart, E. J. Furst, W. H. Hill, and D. R. Krathwhol. 1956. *Taxonomy of Educational Objectives: The Classification of Educational Goals. Handbook I: Cognitive Domain*. New York: David McKay Company.

Bloxham, S., B. den-Outer, J. Hudson, and M. Price. 2016. "Let's Stop the Pretence of Consistent Marking: Exploring the Multiple Limitations of Assessment Criteria." *Assessment & Evaluation in Higher Education* 41 (3): 466–481. doi:10.1080/02602938.2015.1024607.

Boud, D., and E. Molloy. 2013. "Rethinking Models of Feedback for Learning: The Challenge of Design." *Assessment & Evaluation in Higher Education* 38 (6): 698–712. doi:10.1080/02602938.2012.691462.

Boyle, Brendan, Rebecca Mitchell, Anthony McDonnell, Narender Sharma, Kumar Biswas, and Stephen Nicholas. 2020. "Overcoming the Challenge of "Fuzzy" Assessment and Feedback." *Education + Training* 62 (5): 505–519. doi:10.1108/ET-08-2019-0183.

Braun, V., and V. Clarke. 2019. "Reflecting on Reflexive Thematic Analysis." *Qualitative Research in Sport* 11 (4): 589–597. doi:10.1080/2159676X.2019.1628806.

Braun, V., and V. Clarke. 2021. *Thematic Analysis: A Practical Guide*. London, UK: Sage Publications.

Butcher, J., E. McPherson, I. Shelton, A. Clarke, L. Hills, and J. Hughes. 2017. "Unfit for Purpose? Rethinking the Language of Assessment for Widening Participation Students." *Widening Participation and Lifelong Learning* 19 (2): 27–46. doi:10.5456/WPLL.19.2.27.

Carless, D. 2015. *Excellence in University Assessment*, London: Routledge.

Carless, D., and D. Boud. 2018. "The Development of Student Feedback Literacy: Enabling Uptake of Feedback." *Assessment & Evaluation in Higher Education* 43 (8): 1315–1325. doi:10.1080/02602938.2018.1463354.

Dowden, T., S. Pittaway, H. Yost, and R. McCarthy. 2013. "Students' Perceptions of Written Feedback in Teacher Education: Ideally Feedback is a Continuing Two-Way Communication That Encourages Progress." *Assessment & Evaluation in Higher Education* 38 (3): 349–362. doi:10.1080/02602938.2011.632676.

Gablasova, D., V. Brezina, and T. McEnery. 2017. "Collocations in Corpus-Based Language Learning Research: Identifying, Comparing and Interpreting the Evidence." *Language Learning* 67 (S1): 155–179. doi:10.1111/lang.12225.

Gladovic, C., J. H.-M. Tai, K. Nicola-Richmond, and P. Dawson. 2024. "How Can Learners Practice Evaluative Judgement Using Qualitative Self-Assessment?" *Assessment & Evaluation in Higher Education* 49 (6): 755–766. doi:10.1080/02602938.2023.2291341.

Gonsalves, C. 2023. "Democratising Assessment Rubrics for International Students." *Assessment & Evaluation in Higher Education* 49 (5): 587–600. doi:10.1080/02602938.2023.2281237.

Healey, M., R. Healey, and H. West. 2023. "Reflections on Disciplinary Perspectives on Students as Partners." *International Journal for Students as Partners* 7 (1): 1–13. doi:10.15173/ijsap.v7i1.5444.

Hill, J., K. Berlin, J. Choate, L. Cravens-Brown, L. McKendrick-Calder, and S. Smith. 2021a. "Exploring the Emotional Responses of Undergraduate Students to Assessment Feedback: Implications for Instructors." *Teaching & Learning Inquiry* 9 (1): 294–316. doi:10.20343/teachlearninqu.9.1.20.

Hill, J., R. L. Healey, H. West, and C. Déry. 2021b. "Pedagogic Partnership in Higher Education: Encountering Emotion in Learning and Enhancing Student Wellbeing." *Journal of Geography in Higher Education* 45 (2): 167–185. doi:10.1080/03098265.2019.1661366.

Hill, J., and H. West. 2020. "Improving the Student Learning Experience through Dialogic Feed-Forward Assessment." *Assessment & Evaluation in Higher Education* 45 (1): 82–97. doi:10.1080/02602938.2019.1608908.

Hill, J., and H. West. 2022. "Dialogic Feedforward in Assessment: Pivotal to Learning but Not Unproblematic." *Teaching and Learning Inquiry*. doi:10.20343/teachlearninqu.10.20.

Hounsell, D., and C. Anderson. 2009. "Ways of Thinking and Practicing in Biology and History." In *The University and Its Disciplines: Teaching and Learning within and beyond Disciplinary Boundaries*, edited by Kreber, C., 71–83, London: Routledge.

Jazaieri, H. 2018. "Compassionate Education from Preschool to Graduate School: Bringing a Culture of Compassion into the Classroom." *Journal of Research in Innovative Teaching & Learning* 11 (1): 22–66. doi:10.1108/JRIT-08-2017-0017.

Jones, D., and L. Ellison. 2021. "When is a Word Not Just a Word? An Investigation into the Dissonance and Synergy between Intention and Understanding of the Language of Feedback in Legal Education." *The Law Teacher* 55 (2): 155–168. doi:10.1080/03069400.2020.1729012.

Kilgarriff, A., V. Baisa, J. Bušta, M. Jakubíček, V. Kovář, J. Michelfeit, P. Rychlý, and V. Suchomel. 2014. "The Sketch Engine: Ten Years on." *Lexicography* 1 (1): 7–36. doi:10.1007/s40607-014-0009-9.

Kinchin, I. M., E. Alpay, K. Curtis, J. Franklin, C. Rivers, and N. E. Winstone. 2016. "Charting the Elements of Pedagogic Frailty." *Educational Research* 58 (1): 1–23. doi:10.1080/00131881.2015.1129115.

Lee, M., D.-T. Chen, R. J. Y. Tan, and W. L. D. Hung. 2024. "Re-Conceptualising Learner Feedback Agency: A Situational, Deliberative and Entangled Perspective." *Assessment & Evaluation in Higher Education* 49 (5): 680–692. doi:10.1080/02602938.2024.2318281.

Lunt, T., and J. Curran. 2010. "Are You Listening Please?' The Advantages of Electronic Audio Feedback Compared to Written Feedback." *Assessment & Evaluation in Higher Education* 35 (7): 759–769. doi:10.1080/02602930902977772.

MacKay, J. R. D., K. Hughes, H. Marzetti, N. Lent, and S. M. Rhind. 2019. "Using National Student Survey (NSS) Qualitative Data and Social Identify Theory to Explore Students Experiences of Assessment and Feedback." *Higher Education Pedagogies* 4 (1): 315–330. doi:10.1080/23752696.2019.1601500.

Matshedisho, K. R. 2020. "Straddling Rows and Columns: Students' (Mis)Conceptions of an Assessment Rubric." *Assessment & Evaluation in Higher Education* 45 (2): 169–179. doi:10.1080/02602938.2019.1616671.

Matthews, K. E., C. Sherwood, E. Enright, and A. Cook-Sather. 2023a. "What Do Students and Teachers Talk about When They Talk Together about Feedback and Assessment? Expanding Notions of Feedback Literacy through Pedagogical Partnership." *Assessment & Evaluation in Higher Education* 49 (1): 26–38. doi:10.1080/02602938.2023.2170977.

Matthews, K. E., J. Tai, E. Enright, D. Carless, C. Rafferty, and N. Winstone. 2023b. "Transgressing the Boundaries of 'Students as Partners' and 'Feedback' Discourse Communities to Advance Democratic Education." *Teaching in Higher Education* 28 (7): 1503–1517. doi:10.1080/13562517.2021.1903854.

McEnery, T., and A. Hardie. 2012. *Corpus Linguistics*, Cambridge: Cambridge University Press.

Meer, N., and A. Chapman. 2015. "Co-Creation of Marking Criteria: Students as Partners in the Assessment Process." *Business and Management Education in HE*. https://www.tandfonline.com/doi/abs/10.11120/bmhe.2014.00008?journalCode=rhep18. doi:10.11120/bmhe.2014.00008.

Meyer, J. H. F, and R. Land. 2003. "Threshold Concepts and Troublesome Knowledge: Linkages to Ways of Thinking and Practicing within the Disciplines." In *Improving Student Learning: Ten Years on*, edited Rust, C., 412–424, Oxford Centre for Staff and Learning Development, Oxford Brookes University.

Morrish, L., and H. Sauntson. 2020. *Academic Irregularities: Language and Neoliberalism in Higher Education*, London: Routledge.

Nicol, D. 2010. "From Monologue to Dialogue: Improving Written Feedback Processes in Mass Higher Education." *Assessment & Evaluation in Higher Education* 35 (5): 501–517. doi:10.1080/02602931003786559.

Nicol, D. 2021. "The Power of Internal Feedback: Exploiting Natural Comparison Processes." *Assessment & Evaluation in Higher Education* 46 (5): 756–778. doi:10.1080/02602938.2020.1823314.

Nicol, D., and L. Kushwah. 2023. "Shifting Feedback Agency to Students by Having Them Write Their Own Feedback Comments." *Assessment & Evaluation in Higher Education* 49 (3): 419–439. doi:10.1080/02602938.2023.2265080.

O'Donovan, B., M. Price, and C. Rust. 2004. "Know What I Mean? Enhancing Student Understanding of Assessment Standards and Criteria." *Teaching in Higher Education* 9 (3): 325–335. doi:10.1080/1356251042000216642.

Orsmond, P., S. Merry, and K. Reiling. 2000. "The Use of Student Derived Marking Criteria in Peer and Self-Assessment." *Assessment & Evaluation in Higher Education* 25 (1): 23–38. doi:10.1080/02602930050025006.

Richards, K., and N. Pilcher. 2014. "Contextualising Higher Education Assessment Task Words with an 'anti-Glossary' approach." *International Journal of Qualitative Studies in Education* 27 (5): 604–625. doi:10.1080/09518398.2013.805443.

Rossiter, M. P. 2023. "What You Mean versus What You Say' – Exploring the Role of Language and Culture in European Students' Interpretation of Feedback." *Assessment & Evaluation in Higher Education* 48 (4): 544–555. doi :10.1080/02602938.2022.2097197.

Sarcona, A., D. Dirhan, and P. Davidson. 2020. "An Overview of Audio and Written Feedback from Students' and Instructors' Perspective." *Educational Media International* 57 (1): 47–60. doi:10.1080/09523987.2020.1744853.

Shields, S. 2015. "My Work is Bleeding': Exploring Students' Emotional Responses to First Year Assignment Feedback." *Teaching in Higher Education* 20 (6): 614–624. doi:10.1080/13562517.2015.1052786.

Shulman, L. 2005. "Signature Pedagogies in the Professions." *Daedalus* 134 (3): 52–59. doi:10.1162/0011526054622015.

Terry, G., N. Hayfield, V. Clarke, and V. Braun. 2017. "Thematic Analysis." *In The SAGE Handbook of Qualitative Research in Psychology*, edited by Willig, C. and Stainton-Rogers, W., 17–37. London, England, County: SAGE.

Voelkel, S., and L. V. Mello. 2014. "Audio Feedback – Better Feedback?" *Bioscience Education* 22 (1): 16–30. doi:10.11120/beej.2014.00022.

Winstone, N., and D. Carless. 2019. *Designing Effective Feedback Processes in Higher Education: A Learning-Focused Approach*, London: Routledge.

Winstone, N., R. A. Nash, J. Rowntree, and M. Parker. 2017. "It'd Be Useful, But I Wouldn't Use It': Barriers to University Students' Feedback Seeking and Recipience." *Studies in Higher Education* 42 (11): 2026–2041. doi:10.1080/03075079.2015.1130032.

Worth, N. 2014. "Student-Focused Assessment Criteria: Thinking through Best Practice." *Journal of Geography in Higher Education* 38 (3): 361–372. doi:10.1080/03098265.2014.919441.

Wyse, S., B. Page, H. Walkington, and J. Hill. 2020. "Degree Outcomes and National Calibration: Debating Academic Standards in UK Geography." *Area* 52 (2): 376–385. doi:10.1111/area.12571.

Yang, M., and D. Carless. 2013. "The Feedback-Triangle and the Enhancement of Dialogic Feedback Processes." *Teaching in Higher Education* 18 (3): 285–297. doi:10.1080/13562517.2012.719154.

Yorke, M. 2011. "Summative Assessment: Dealing with the Measurement Fallacy." *Studies in Higher Education* 36 (3): 251–273. doi:10.1080/03075070903545082.