

Author's accepted manuscript

Adalberon, E. (2020). Providing assessment feedback to pre-service teachers : a study of examiners' comments. *Assessment & Evaluation in Higher Education*, 46(4), 601-614. <https://doi.org/10.1080/02602938.2020.1795081>.

Published in: *Assessment & Evaluation in Higher Education*

DOI: <https://doi.org/10.1080/02602938.2020.1795081>

AURA: <https://hdl.handle.net/11250/2771738>

This is an Accepted Manuscript version of the following article, accepted for publication in *Assessment & Evaluation in Higher Education*. Adalberon, E. (2020). Providing assessment feedback to pre-service teachers : a study of examiners' comments. *Assessment & Evaluation in Higher Education*, 46(4), 601-614. <https://doi.org/10.1080/02602938.2020.1795081>. It is deposited under the terms of the Creative Commons Attribution-NonCommercial License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited.

Providing assessment feedback to pre-service teachers: A study of examiners' comments

Written by Erik Adalberon (ORCID: <https://orcid.org/0000-0002-0281-2307>)

Abstract

This article reports a study of written feedback comments in the context of teacher education. While feedback is believed to have the potential to improve students' learning, the concrete provision will rest upon educators' and examiners' ability and means to convey details about their assessment. In the context reported here, compulsory feedback beyond the regular grade was introduced to strengthen a teacher education programme at a Norwegian university. The interest of this study is thus to investigate how the examiners managed this task and will focus on their written comments during three consecutive years. A content analysis of 411 individual feedback comments reveals that most of them are written in a formulaic fashion closely related to the grade descriptors for the course. The discussion raises questions about inherent problems with written assessment feedback and how the standardisation of it may limit the informational value for students' learning

Keywords:

Feedback, Higher education, Teacher education

Assessment feedback in teacher education

Providing feedback to students has been described as a vital part of a formative assessment cycle (Hattie & Timperley, 2007; Villaroel, Boud, Bloxham, Bruna, & Bruna, 2019), and this type of response from educators can be conveyed in several ways. For instance, feedback may be presented through dialogue, written comments, or various types of digital media. The purpose behind the provision of assessment feedback will differ according to the context. When it comes to examinations, the aim is to give the students some insight into the assessment process and how to improve their work. While the students' uptake of this information may be challenging (Sutton, 2012; Carless & Boud, 2018), the point is to give directions for how the learners should structure their future learning processes.

The practice of providing feedback varies. In the context investigated here, namely Norwegian higher education, no feedback beyond a grade is normally provided unless the individual student requests it after the examinations (Universitets- og høyskoleloven, 2005, §5-3). The common procedure for examiners is to assess the students' performances on the examination according to a set of pre-defined criteria and award a grade as the only form of feedback. The legal right to feedback from the examiners is not specified as a formative comment the students can learn from, but is rather a justification of how the grade was determined. In other words, in this context there is no system for giving the students deeper insight into how they could improve themselves and their learning strategies.

A few remarks about the Norwegian higher education system are necessary to clarify the context. Most of the education programmes, and teacher education in particular, are influenced by political processes, such as reforms, policies and white papers (Afdal, 2013). As in many other Western countries, recently there has been a more discernible focus on increased quality in education, and this has paved the way for various small yet significant changes in the Norwegian higher education system (Elken & Frølich, 2017).

The aim of this article is thus to investigate a local initiative where the provision of written feedback was made obligatory. At the University of Oslo a recent reform in the teacher education programme entailed several changes. For example, it was decided that each pre-service teacher should receive written feedback in a particular course. After the completing a novel video examination design, examiners would provide elaborate comments on their assessments (Lund & Engelién, 2015). This practice was meant to bolster the overall strategy of constructive alignment, and harmonise the various elements in the course. In other words, they wanted to renew the assessment cycle and improve the alignment between the examination, feedback and other components in the course, such as teaching and practice. Providing feedback would potentially give the students insight into the assessment process, and there would be learning benefits as suggested in research literature.

Considering the local initiative as a rare and novel case of providing mandatory feedback, this provides a rich opportunity to broadly investigate feedback practice and its role in the learning process.

Research interest

The objective of this study is to shed light on a case where feedback was given to student teachers after completing a recently developed video-examination design. The investigation will cover three years of practice, and the overriding research question is: How did the examiners approach the task of providing feedback to the pre-service teachers?

The following sub-questions guide the analysis:

- How was the feedback organised?
- What characterises the written content?
- What kind of information does the feedback provide?

Categorisation of feedback is here inspired both in the findings from studies conducted by Ivanîc, Clark and Rimmershaw (2000) and Hughes, Smith and Creese (2015).

Theoretical framework

The socio-cultural perspective has been chosen for the analysis. This perspective will direct attention on the various tools used in this context. Higher education is known to be characterised by its various literacies (Ivanîc et al., 2009) and the use of specialised concepts, often described as ‘academic’ or ‘scientific’ (Vygotsky, 1978, p. 103). Writing to an unknown reader, such as when conveying feedback, involves a high level of abstraction (John-Steiner, 2007), and writing under these circumstances of full anonymity implies that the examiners have to be precise and clear in their formulation. The feedback will also rest upon a shared understanding of what the involved concepts mean.

Research on feedback

Research on feedback practices has grown markedly (Carless, 2015, p. 190) in recent years, which has led to several reviews on the topic (Hattie & Timperley, 2007; Evans, 2013; May 2013; Li & De Luca, 2014). Issues that have been treated frequently recently are, for instance, feedback as part of formative assessment (Knight, 2002; Shute, 2008), digital systems for feedback (Henderson & Phillips, 2014) and the involvement of students as active participants in the feedback process (Boud & Molloy, 2013). These various analyses provide a rich background that outlines the many roles feedback may have in the assessment process.

Feedback as a term is used to describe various practices, and it appears to be difficult to find a definition that completely covers all the contexts where feedback occurs (Price, Handley, Millar, & O'Donovan, 2010). Currently, research seems to be oriented towards how feedback

functions as a practice that pertains to assessment, which means that students are supposed to learn from the information they are given (Evans, 2013; Carless & Boud, 2018). But improving formative aspects of assessment through feedback has proven to be a challenging task. According to a number of researchers, the learning potential of feedback depends on the presence of certain factors, where some potential challenges have also been mentioned (Lizzio & Wilson, 2008; Shute, 2008). For instance, Higgins, Hartley and Skelton (2001) found that students rarely use the feedback they receive and point out that the communication between examiners and students is ‘inherently problematic’ (Higgins et al., 2001, p. 272).

There are many varieties when it comes to the format of the feedback, and each one has its possibilities and limitations. Feedback in the form of a personal dialogue seems to be preferred by both students and lecturers (Handley et al., 2007; Mulliner & Tucker, 2017). Although this format can be time-consuming, it facilitates communication between students and educators (Blair, Wyburn-Powell, Goodwin, & Shields, 2014; Carless, 2016). Rather than giving unidirectional and closed comments, the dialogue invites one into an in-depth conversation about the performance.

Feedback is often presented in the form of written comments (Jolly & Boud, 2013), as has been done in the case reported here. One related example is reported in the study of Bailey and Garner (2010), where educators were given the opportunity to reflect on their feedback practices in qualitative interviews. This study finds a tension between the idealised role of feedback as an element of formative learning and the practical aspect of being accountable for the grading (Bailey & Garner, 2010). In short, the educators write their feedback with more than one reader in mind. Of course, they address the candidate, but also any other potential authority that may hold them accountable for the content. As a result, some examiners reduce their comments to a minimum.

Written feedback commentaries may contain various types of information. In research on this topic it is not uncommon to identify different categories that reflect the intention behind the text. In a study by Hughes, Smith and Creese (2015), written comments on both drafts and final work in various study programmes were investigated through a profiling tool consisting of feedback categories. These categories were a) praise, b) recognising progress, c) critique, d) advice and e) clarification (Hughes et al., 2015, pp. 1083-1084). They found several possible patterns in the material, for instance, that the comments on the final work were dominated by what could be characterised as ‘closed’ remarks. This finding implies that

examiners tend to write their feedback as a concluding comment, rather than suggestive notes for further negotiation. In a similar study on feedback, Ivančić, Clark and Rimmershaw (2000) investigated comments on finalised texts and found that different tutors tended to have different styles of feedback. Finding tendencies across the various examiners, they identified some broad categories of feedback defined as: a) explaining the grade in terms of strengths and weaknesses, b) correcting or editing the student's work, c) evaluating the match between the student's essay and an ideal answer, d) engaging in a dialogue with the student, e) giving advice which will be useful in writing the next essay, and f) giving advice on rewriting the essay (Ivančić, Clark & Rimmershaw, 2000, p. 55).

The practice of providing feedback as written comments has certain limitations. Considering the effects of written comments, Orsmond and Merry (2011) and Glover and Brown (2006) investigated the depth and characteristics of such comments and found that the content focused on details about the performance. In their view, there are several shortcomings to such feedback, and their overall experiences can be summarised under three points that are also rooted in other relevant literature:

- Students do not understand the academic discourse used to underpin assessment criteria (Glover & Brown, 2006, p. 12)
- Feedback was 'focused on assignment content rather than feeding forward' (Orsmond & Merry, 2011, p. 133)
- Tutors do not expect the students 'to build on the piece of work in future models' (Hughes, Smith, & Creese., 2015, p. 1090)

These results stand in contrast to the findings in Ferguson (2011), where students report that they prefer written feedback when it is constructive. This response has been supported in a later study by Dowden, Pittaway, Yost & McCarthy (2013).

To summarise, research shows that feedback is commonly provided as written comments, even though there are potential pitfalls with this format. Examiners should pay attention to how such written comments are formulated so they can serve as an effective and valuable part of the students' learning process. In the context of professional education, where students are supposed to learn 'about,' 'for' and 'through' practice, special attention should be paid to the pedagogical aspects of feedback (Warhuus et al., 2018).

Background: The assessment process

A video-examination design was implemented in one of the courses to create a bond between the various knowledge domains taught in the course and what can be displayed in the assessment format (Lund & Engelién, 2015).

The video-examination design, introduced in January 2103, involved a video case as the focus for the tasks summed up below:

- Make observations and analyse them by using relevant theory
- Formulate a thesis statement based on the observations
- Discuss the thesis statement by using relevant theory

A video case will be +/- 10 minutes of video-recorded classroom activities from an authentic setting. The pre-service teachers are invited to a video analysis seminar prior to the examination and may test themselves on example cases that are available online. Although the examination will be submitted individually, the pre-service teachers may sit anywhere they like, collaborate with peers and use such tools like notes and course literature to compose their answers (for elaborate descriptions, see Adalberon, 2020). The point is to make a qualified observations and analysis of the video where they display their ability to notice and reflect over professionally relevant situations rather than just displaying knowledge of theories alone (Lund & Engelién, 2015; Adalberon, Hauge, & Säljö, 2019). The assessment process is concluded when the pre-service teachers receive their grade and feedback comments. This video examination format is not utilised in other courses.

The grade descriptors for the video examination are identical during the three years investigated here and were made available on the course web page from the beginning of each year the examination was scheduled. The content is presented in a table where general characteristics of each grade are defined, and where three categories with specific descriptors are given to indicate what is vital for a very good performance. These categories are:

1. Observation, interpretation and formulation of the thesis statement
2. Ability to assess and independence
3. Language, structure and formalities

Grade A is described as:

Outstanding

An excellent performance that clearly stands out. The candidate shows very good assessment skills and a high degree of independent thinking.

The other categories include more details about the criteria for an A:

1. The candidate gives a very precise description of key observations in the case and formulates a relevant thesis statement that is focused on Subject Didactics II (name of the course). The grasp of pedagogical and didactic theory is very relevant to the problem. The candidate shows excellent knowledge and understanding of the fields of study. Experiences drawn from the teacher training are very relevant to the discussions.
2. The candidate argues very clearly and assertively and substantiates arguments with reference to the course literature. The candidate is clearly able to see the connection between pedagogy and subject didactics. Examples from theory and practice are used with clear critical objectivity.
3. The language is very accurate and assertive with a minimum of paraphrasing. The text is well structured. Formalities are satisfied.

(Universitetet i Oslo, 2016) [my translation]

Many of these phrases and the wording in general are similar, if not identical, to the national grade description system (Universitets- og høyskolerådet, 2011).

The descriptors are presented in different ‘boxes’ where each category for the various grades is described, A through F. In comparison to the example above, grade B indicates *very good performance*, and *the candidate shows very good assessment skills and independence*. The difference between an A and a B is thus the variation between an *outstanding* and *very good* performance, respectively. Following the same pattern, grade C is *satisfactory* with respect to the mentioned key abilities, grade D is *acceptable*, and so on.

The examiners who have assessed the examination and provided feedback comprise a team of educators with relevant backgrounds. Some are internal examiners and about half are external to the institution where the examination is to be given. They have not met beforehand, and all have received an e-mail with the necessary information about the examination. After grading the exams individually, they discuss any differences between the awarded grades and calibrate their assessments. The feedback, the result of a similar discussion, is written as a separate text, anonymously, and then mailed to each student.

Methodology

This investigation is based on the written feedback given to each pre-service teacher after completing the video examination over a period of three consecutive years from 2014-2016. The material is summarised in Table 1:

(Table 1: Overview of the data, year of examination, number of students, number of words, and average number of words per student.

Table 1: General information about feedback data			
Year when feedback is given	# of students	Total # of words in the feedback document	# average word per feedback
2014	136	11247	~83
2015	142	15457	~109
2016	143	16693	~117

The analysis has been performed in three stages: a) an initial exploratory read-through of the material, b) a software-driven analysis of the material and c) a qualitative reading of idiosyncratic responses.

Initial reading

This exploratory stage was meant to reveal some general features of the material, such as its format, structure and style. Without paying much attention to details, it provided a necessary overview for developing the rest of the analysis.

Software-driven analysis

After the first stage, an appropriate analytical model was chosen to categorise the material. It was evident that the examiners' comments contained many of the same phrases that were iterated and rewritten in various forms, even though the material was made over three years with different examiners. An approach similar to corpus-based analysis (Stubbs, 1996; Randall & Mirador, 2003) was deemed viable. Although this approach is usually applied to investigate large amounts of text for linguistic purposes (Stubbs, 1996; Nesi & Gardner, 2012), the general principle of finding keywords and clusters of words was applied to this material to find patterns that characterise the content.

The software used to aid the process was QSR's N-Vivo 11. The functionalities in this programme proved helpful in organising and coding the material. Such functions as 'word frequency,' 'text search' and 'word tree' were also used to obtain an overview of the data.

Qualitative analysis

The final stage of the analysis was a qualitative investigation of the written feedback, where the aim was to find idiosyncratic responses from the examiners. Since the majority of comments followed a formulaic pattern, this stage of the analysis focused on the content that deviated from this line. The findings were categorised and summarised afterwards in what might be called an abductive process. The feedback comments were first organised exploratively with the node-functionality in N-vivo, and given a temporary descriptive label. Their content was to be considered against previous research. Suggested categories by Hughes, Smith and Creese (2015) and Ivančić, Clark and Rimmershaw (2000) were deemed sufficient to cover the limited number of elaborate comments found in the material. It was somewhat difficult to apply the categories as they were originally proposed. As the examination papers were finalised texts, some categories, such as ‘advice on rewriting’, were not deemed relevant in this context. Moreover, some categories, such as ‘praise’/‘criticism’ and ‘explain the grade in terms of strengths/weaknesses’ tend to overlap, so a modified set was established to cover the essence of the idiosyncratic feedback.

Presenting a case

Among the various written examples of feedback one particular case from the data was selected to provide some insight into how the message in a typical feedback comment was formulated. During the qualitative analysis stage, five texts were found as representative, considering their a) length, b) content and c) organisation, of the overall feedback. One of these has then been chosen as a case that is representative of the rest and, at the same time, provides interesting details.

Ethics

The material used in this research project was anonymised, and thus no names have been mentioned, neither the students’ nor the examiners’. The required permissions has been procured from the involved parties, and the Guidelines from the Norwegian Centre for Research Data (NSD) have been complied with when gathering and analysing the data. Considering the original feedback was written in Norwegian, care (i.e. using three translators) has been taken to ensure accurate translation of the material to preserve the meaning.

Findings

This section will be organised into three subsections; a brief overview of the organisation and composition, a look at which phrases and expressions are commonly used by the examiners, and idiosyncratic comments where the examiners have given specific advice.

Organisation of the feedback

The analysed documents cover three consecutive years of feedback, where several examiners have been involved in the process of writing it. However, there is a striking uniformity to the material and its structure. Each document has been organised with each candidate's number in a row, where the feedback is given in five columns pertaining to these categories:

1. Formulation of the thesis statement
2. Use and integration of pedagogical and subject-didactics literature
3. Ability to assess, and independence
4. Language, structure and formalities
5. Other comments

Since categories 1, 3 and 4 are almost identical to the categories used in the grade descriptors presented above, it is reasonable to assume that the categories have been based on these descriptors.

Standard phrases in the feedback

The feedback contains a great number of expressions and phrases that are used repeatedly across all three years investigated here. Some of the recurring expressions, found in the material by using a search function, are presented in Table 2:

(Table 2: Expressions used in the feedback (first column), expressions found in the first, second and third year.)

Table 2: Frequently occurring expressions			
Expression (Translated from Norwegian)	1st year Per cent	2nd year Per cent	3rd year Per cent
The candidate describes (important) observations	4.4	24.6	25.9
The language is precise	5.1	15.5	10.5
Formulates a relevant issue	4.4	14.1	6.3
The grasp of pedagogical and didactic theory is relevant	11.0	16.2	15.4
The candidate argues well	0.7	17.6	21.0
Good language	11.0	18.3	11.9

Formalities are fulfilled	11.0	44.4	34.3
Good structure	13.2	11.3	6.3
Critical objectivity	9.6	62.7	11.9
Precise and assertive	5.1	29.6	9.1

The content reveals a tendency where some expressions, and variations of these, are repeated quite frequently over the three years. Considering that not all the examiners were the same over the three-year period, the choice of wording is strikingly similar. The provision of feedback seems to be following a pattern where certain expressions are utilised in the feedback.

The connection to the grade descriptors is evident. Keywords and expressions from Table 2, such as *critical objectivity*, *structure*, *assertive*, are also part of the earlier presented criteria set for this video-case examination. The expressions are almost identical and seem to be based on much of the same logic.

Furthermore, the many expressions used in the feedback tend to be concentrated on certain aspects of the students' texts. Some examples of this:

- The observations must be important ones
- The language must be precise
- The thesis statement must be relevant
- The grasp of theory must be relevant

These specific expressions point out various issues the examiners find vital for the students' performance, and the similarities between the feedback and grade descriptors are clear. What the examiners consider to be *important*, *precise* and *relevant* is not made explicit in these standardised expressions and has not been elaborated on further. Most of the data suggest that the bulk of the feedback is organised in a formulaic fashion, and strongly tied to the pre-defined criteria.

Case: Feedback on the individual level

The written feedback to one of the students commenced by commenting on her performance by pointing to issues that belong to the first category, namely observations and the formulation of a chosen thesis statement. Here the examiners write:

Relevant observations and assessments. The candidate describes observations in the case and formulates an thesis statement for subject-didactics [cites the thesis statement].¹

The first brief sentence serves to inform the student that her observations and assessments were ‘relevant’. This term is used to describe the quality of the candidate’s observations and assessments, but is not elaborated on or explained anywhere. Thus, the meaning behind the comment ~~thus~~ rests upon the premise that the addressee knows what it implies. The second sentence, which is mainly descriptive, confirms that the candidate has fulfilled a part of the task. No additional comments of an evaluative nature are included.

In the next section of the feedback, which is called ‘use and integration of pedagogical and subject-didactics literature’, the examiners point out that:

The amount of course literature and subject-didactics theory are relevant to the thesis statement. The candidate displays knowledge and understanding of the fields of study without obvious deficiencies.

Once again, the examiners state that the student has used ‘relevant’ literature in the examination and comment on the knowledge displayed.

With respect to the ability to assess and the independence of the candidate, it is mentioned that:

Experiences from practice that are drawn upon are fairly relevant for the discussion. The candidate argues with assertiveness, to a certain degree. The candidate can see connections between pedagogy and didactics. Theory and experiences from practice are used with critical objectivity. The summary and conclusion capture the thesis statement and the discussion. [my translation]

Again, these comments pertain to the grading criteria, and it is worth noting the expression ‘fairly relevant’. The examiners appear to distinguish between ‘relevant’ and ‘fairly relevant’ by using grading adverbs. As in the grade descriptors, they differentiate between the levels of performance by using graded descriptions.

In the final section, the examiners continue to use expressions that are similar to the previous ones:

¹ My translation from Norwegian in italics.

The language is fairly good, but from time to time it is characterised by paraphrasing.

The text has a clear structure. Formalities are, to a reasonable degree, fulfilled.

The adverb 'fairly' is used to moderate the description of the performance, and quantitative comments such as 'to a reasonable degree' have been added. This choice of words may have been used to indicate that something about the performance is not entirely satisfactory from the examiners' point of view. Here again, there is no elaboration on this point in the feedback.

To summarise, the case described here represents an average example of how feedback has been given to the students over the three years in question. As demonstrated, the feedback is connected to the grading criteria and tends to contain short sentences referring to the examiners' assessments. As a consequence, the feedback is not particularly rich in detail, and the assessment rests on keywords that are used to describe the quality of the students' performance, such as 'relevant' or 'important'. The message the comments convey would thus be dependent on a mutual understanding of such terms.

Idiosyncratic comments

Most of the above-mentioned comments are related to the grade descriptors and justification of the grade. However, the examiners would sometimes provide more elaborate comments about the candidates' performances beyond the formulaic examples. In the following, some of the instances of elaborate feedback will be investigated closer. The chosen examples have been categorised into two categories of feedback: praise/criticism and advice for writing.

Praise/criticism

Most of the feedback has been written in a neutral, descriptive format, but some examples show how the examiners also make value-laden assessments where both praise and criticism are included. For instance, one extract illustrates how the examiners start by praising the use of the case and at the same time offer some constructive criticism:

It is positive that the candidate chooses to use the case actively, but the discussions could have been elevated to a more theoretical level and been tied more closely to the subject-didactics themes that emerge from the thesis statement.

This evaluation by the examiners displays a different approach than seen in the previous section. They point to a specific instance in the examination report and analyse it in terms of strengths and weaknesses. In this instance, the examiners explain that using the case (actively) is a positive approach, but other elements could have been improved. These mentioned issues

are, however, not part of the formal grade descriptors. To some extent, the comment can be interpreted as a suggestion for improvement.

Another example of critique from examiners is seen in this extract:

Shows that he is not capable of seeing the large connections between pedagogy and subject didactics, which is clear to see through the lack of theoretical grounding.

Without referring back to the formal grade descriptors, the examiners believe that the student has not proven that he has any insight into the ‘large connections between pedagogy and subject didactics. Compared to the previous example, this statement is less constructive and represents an uncommon way of presenting criticism.

Advice for writing

In some of the feedback to the candidates, the examiners make specific suggestions as to how they could improve their examination reports. For instance, the examiner might give advice on key areas in the examination report that could be improved: ‘Has chosen a thesis statement comprising two questions; it would have been better with a simplification’. Although this is a brief suggestion, it clearly stands out when compared to the standardised way of providing feedback. This is a particular piece of advice that may be used for future reference.

In another similar example given to the same candidate the examiner says ‘The candidate could advantageously have utilised another reference in her discussion of start-ups (...).’ This is an even more specific comment, and direct advice to the student on how to improve her work. Without giving any examples of a better reference, the examiner hints that the candidate has based her discussion on a source that has weak relevance.

Other examples from the material show how the examiners could also be more specific in their advice:

For instance, Björkman (2010) is used as one of four texts from the curriculum covering the theme of oral communication in English. According to the thesis statement, it would have been desirable that three others had been drawn in (Rindal 2014; Swain 2001; Cogo, 2012).

This comment points to a perceived shortcoming in the examination report and gives examples of the other references that could have been used to gain a better result.

To summarise, examples from both categories illustrate that the examiners are sometimes inclined to comment on other aspects than the ones mentioned in the official grade descriptors and draw upon their own knowledge.

Discussion

This article has focused on a context where examiners are giving feedback to pre-service teachers after assessing their performance on a video-case examination in teacher education. Such provision of feedback may be perceived as a writing genre which requires comprehensible note-taking and involves being fair and concise. Hence, a balance between various considerations is required.

The first issue that should be pointed out here, and which is the most obvious aspect, is that the majority of the feedback in this study is characterised by the use of similar or even standardised phrases. This pattern would thus indicate an emerging standardised practice of providing feedback over the three years investigated. Various examiners were involved in this period, but apart from a slight increase in length, the feedback comments would remain quite uniform. Considering that educators commonly have their own style of providing feedback (Ivančić, Clark and Rimmershaw, 2000), such similarity across examiners strengthens the argument that a standardisation has taken place. The content of the repeated phrases and expressions can be traced directly to the wording found in the grade descriptors through a striking similarity in words and sentences. Most of the feedback seems to be information on whether the performance was in accordance with the grade descriptors. It is also possible to see a close connection with the national template for grade descriptors (Universitets- og høgskolerådet, 2011).

This pattern leads to a second and related point, namely a potential tension between the standardisation of feedback, on the one hand, and the need to provide sufficient information about the assessment processes on the other. These two concerns are not necessarily mutually exclusive, but may be problematic for examiners in some cases. Following the logic of Bailey and Garner (2010), who inferred that feedback is written for more than one reader, it would seem that some of the same conflicting interests are present in this material. In the emerging standardisation seen in a great portion of the feedback, it is possible to recognise an element of accountability, where feedback is linked closely to grading guidelines and other official documents. The idiosyncratic comments demonstrate a concern for the student's learning outcome, and challenges standardisation.

This tension is visible in the instances where the examiners elaborate on their feedback and add other aspects than the ones covered by the grade descriptors. Breaking the pattern in this way indicates that the examiners have seen the need to comment on other things than the pre-defined aspects. This approach is closer to the notion of ‘double duty’ (Boud, 2000), which means that feedback should justify the grade *and* inspire future learning. The formalised and standardised format appears to cover the former, while the few instances of elaborate feedback cover the latter. The formulaic feedback will have the potential to convey some meaning, but will seldom include constructive comments requested by students (Dowden et al., 2013; Ferguson, 2011)

The third point in this discussion is that the investigated feedback relies on a shared understanding of certain frequently used concepts. Following the principle of deliberate semantics, the original meaning behind the examiners’ feedback to the candidate rests upon a shared understanding of the terms used in the comments, which is also at the core of general language appropriation (Vygotsky, 1978). For instance, terms such as ‘relevant’ and ‘good’ are used quite often, but no examples or explanations are provided to tell the student what these terms mean. In other words, the use of such terms is insufficiently explicit, and hence the feedback may become ambiguous and can be interpreted in various ways. Written feedback is thus prone to the well-described and inherent dilemma of writing, namely how the original meaning is conveyed and what the addressee ultimately interprets (Bal, 2009). Since the investigated feedback also follows the wording of official standards closely, the use of criteria or grade descriptors will represent yet another problem; how should generalised terms be understood across levels in the higher-education hierarchy?

Written feedback has some inherent limitations that have been described in previous research (Glover & Brown, 2006; Orsmond & Merry, 2011; Hughes, Smith, & Creese, 2015). Such issues as understanding the academic discourse, giving feed-forward messages and the educators’ attitudes are all concerns that need to be addressed.

In short, this study provides some insight into a process where a chosen feedback model was implemented in order to inform and guide the pre-service teachers after completing a new video-examination format. In light of such ideals as stimulating students’ assessment literacy (O’Donovan, Rust, & Price, 2016) and handling future learning processes (Hattie & Timperley, 2007), the feedback, as it is formulated in this study, has obvious limitations. Even though there are examples of giving individual hints and tips, most of the comments are written in close accordance with a set of grade descriptors. The result is feedback that will

lack the depth suggested in the literature, and does not then harmonise with the original intent of ‘constructive alignment’ in the course.

Conclusion

Before any conclusion can be drawn it is important to point out some limitations of this study. Since this research is based on written content, little information about the examiners and their working conditions is available. As a consequence, it is not possible to make any strong claims about the writing processes, such as the examiners’ collaboration, discussions and special considerations they may have made.

This article aimed to investigate how feedback was given in a recently reformed teacher education programme. As a potentially valuable source for student learning, the feedback was here intended to improve the coherence between the different elements in the course. The findings indicate that the examiners chose a formulaic approach to the task, with a close relationship between the comments and the national grade descriptors. In further research, it would be interesting to investigate the examiners’ decision-making process when providing feedback.

In Norwegian higher education, feedback is not given often, and thus most examiners have little experience in writing feedback comments. Being an examiner is not a certified position, and each educator has to find their way to assess student performances. Carless and Boud (2018) suggest more training for students in what they call ‘feedback literacy’, and it is tempting to propose similar training for examiners. Writing understandable and useable feedback is a challenging task that requires systematic practice and builds on experience over time. The education institutions are responsible for the practice of providing feedback, and should, therefore, also facilitate and encourage training and systematic improvement.

Last but not least, it is worth raising the question as to whether the current focus on standardisation reduces the examiners’ opportunities to give a proper feed-forward message to students, particularly in a written format. The notion of establishing standards or criteria independent of the examiners is troublesome. Standards set a frame that will define what dimensions should be included or excluded from the assessment provided. Moreover, standards will entail some degree of interpretation. For instance, a certain understanding of a

'good' or 'outstanding' performance will be the outset when writing standards. The examiners will later have to interpret and conceptualise the terms before they remediate their understanding to the students through writing. This issue relating to the establishment of standards has already been noted in other nation-wide projects (Hopfenbeck, Throndsen, Lie & Dale, 2012). At the final stage, the students will have to establish some understanding of the terms, and here everything could become a chain of possible misinterpretations. Reducing the complex process of quality recognition into a simple set of standards seems to result in a feedback-format with little depth, as seen in this study.

The required understanding of quality is essential in any assessment activity and should be negotiated among educators who are responsible for the overriding learning process, and discussed with students to create a better understanding of the discourse. It is thus problematic to make central guidelines for this understanding. An approach like this, without the involvement of the implicated parties, will probably result in a ritualised method for providing feedback that has limited value.

References

- Adalberon, E. (2020). Affordances of a video examination: Opportunities for pre-service teachers to demonstrate professional knowledge of teaching and learning. *Educational theory and practice*, 42(1), 47-64.
- Adalberon, E., Hauge, T. E., & Säljö, R. (2019). Pre-service teachers' experiences with a digital examination design: The inter-relation between continuity and change in an institutional context. *ACTA Didactica*, 13(3), 1-19.
doi:<http://dx.doi.org/10.5617/adno.6864>
- Afdal, H. W. (2013). Policy making processes with respect to teacher education in Finland and Norway. *Higher Education*, 65(2), 167-180.
- Bailey, R., & Garner, M. (2010). Is the feedback in higher education assessment worth the paper it is written on? Teachers' reflections on their practices. *Teaching in Higher Education*, 15(2), 187-198.
- Bal, M. (2009). Working with concepts. *European Journal of English Studies*, 13(1), 13-23.
- Blair, A., Wyburn-Powell, A., Goodwin, M., & Shields, S. (2014). Can dialogue help to improve feedback on examinations? *Studies in Higher Education*, 39(6), 1039-1054.
doi:[10.1080/03075079.2013.777404](https://doi.org/10.1080/03075079.2013.777404)
- Boud, D. (2000). Sustainable Assessment: Rethinking assessment for the learning society. *Studies in Continuing Education*, 22(2), 151-167. doi:[10.1080/713695728](https://doi.org/10.1080/713695728)
- Boud, D., & Molloy, E. (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](https://doi.org/10.1080/02602938.2012.691462)
- Carless, D. (2015). *Excellence in university assessment: Learning from award-winning practice*. New York, NY: Routledge.

- Carless, D. (2016). Feedback as dialogue. In Peters, M. A. (Ed.), *Encyclopedia of Educational Philosophy and Theory* (pp. 1-6): Springer Singapore.
- Carless, D., & Boud, D. (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., Pittaway, S., Yost, H., & McCarthy, R. (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
- Elken, M., & Frølich, N. (2017). Pulling the plug in a bathtub: The big consequences of a small change in Norwegian higher education. In *Policy Analysis of Structural Reforms in Higher Education: Processes and Outcomes* (pp. 95-117). Cham: Springer International Publishing.
- Evans, C. (2013). Making sense of assessment feedback in higher education. *Review of Educational Research*, 83(1), 70-120. doi:10.3102/0034654312474350
- Ferguson, P. (2011). Student perceptions of quality feedback in teacher education. *Assessment & Evaluation in Higher Education*, 36(1), 51-62. doi:10.1080/02602930903197883
- Glover, C., & Brown, E. (2006). Written feedback for students: too much, too detailed or too incomprehensible to be effective? *Bioscience Education*, 7(1), 1-16. doi:10.3108/beej.2006.07000004
- Handley, K., Szwelnik, A., Ujma, D., Lawrence, L., Millar, J., & Price, M. (2007). When less is more: Students' experiences of assessment feedback. *Paper presented at the Higher Education Academy*.
- Hattie, J., & Timperley, H. (2007). The Power of Feedback. *Review of Educational Research*, 77(1), 81-112.
- Henderson, M., & Phillips, M. (2014). *Technology enhanced feedback on assessment*. Paper presented at the Australian Computers in Education Conference.
- Higgins, R., Hartley, P., & Skelton, A. (2001). Getting the Message Across: The problem of communicating assessment feedback. *Teaching in Higher Education*, 6(2), 269-274. doi:10.1080/13562510120045230
- Hopfenbeck, T. N., Throndsen, I., Lie, S., & Dale, E. L. (2012). Assessment with Distinctly Defined Criteria: A Research Study of a National Project. *Policy Futures in Education*, 10(4), 421-433. doi:10.2304/pfie.2012.10.4.421
- Hughes, G., Smith, H., & Creese, B. (2015). Not seeing the wood for the trees: developing a feedback analysis tool to explore feed forward in modularised programmes. *Assessment & Evaluation in Higher Education*, 40(8), 1079-1094. doi:10.1080/02602938.2014.969193
- Ivančić, R., Edwards, R., Barton, D., Martin-Jones, M., Fowler, Z., Hughes, B., . . . Smith, J. (2009). *Improving Learning in College: Rethinking Literacies Across the Curriculum*. New York NY: Routledge.
- Ivančić, R., Romy, C., & Rimmershaw, R. (2000). What am I supposed to make of this? The messages conveyed to students by tutors' written comments. In M. R. Lea & B. Stierer (Eds.), *Student writing in higher education* (pp. 47-65). Philadelphia, PA: The Society for Research into Higher Education & Open University Press.
- John-Steiner, V. (2007). Vygotsky on thinking and speaking. In H. Daniels, M. Cole, & J. V. Wertsch (Eds.), *The Cambridge companion to Vygotsky* (pp. 136-152). Cambridge: Cambridge University Press.
- Jolly, B., & Boud, D. (2013). Written feedback. In D. Boud & E. Molloy (Eds.), *Feedback in higher and professional education*. London: Routledge (pp. 104-124). New York, NY: Routledge.

- Knight, P. (2002). Summative assessment in higher education: practices in disarray. *Studies in Higher Education*, 27(3), 275-286.
- Li, J., & De Luca, R. (2014). Review of assessment feedback. *Studies in Higher Education*, 39(2), 378-393. doi:10.1080/03075079.2012.709494
- Lizzio, A., & Wilson, K. (2008). Feedback on Assessment: Students' Perceptions of Quality and Effectiveness. *Assessment & Evaluation in Higher Education*, 33(3), 263-275.
- Lund, A., & Engelién, K. (2015). Oppgaver og vurdering i digitale omgivelser [Tasks and assessment in digital environments]. In A. Lund, U. Rindal, & R. Jakhelln (Eds.), *Veier til fremragende lærerutdanning* [Paths to excellent teacher education] (pp. 136-147). Oslo: Universitetsforlaget.
- May, T. (2013). Identifying the characteristics of written formative feedback used by assessors in work-based qualifications. *Journal of Vocational Education & Training*, 65(1), 18-32. doi:10.1080/13636820.2012.727855
- Mulliner, E., & Tucker, M. (2017). Feedback on feedback practice: perceptions of students and academics. *Assessment & Evaluation in Higher Education*, 42(2), 266-288. doi:10.1080/02602938.2015.1103365
- Nesi, H., & Gardner, S. (2012). *Genres across the disciplines: student writing in higher education*. Cambridge: Cambridge University Press.
- O'Donovan, B., Rust, C., & Price, M. (2016). A scholarly approach to solving the feedback dilemma in practice. *Assessment & Evaluation in Higher Education*, 41(6), 938-949. doi:10.1080/02602938.2015.1052774
- Orsmond, P., & Merry, S. (2011). Feedback alignment: effective and ineffective links between tutors' and students' understanding of coursework feedback. *Assessment & Evaluation in Higher Education*, 36(2), 125-136.
- 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.
- Randall, M., & Mirador, J. (2003). How well am I doing? Using a corpus-based analysis to investigate tutor and institutional messages in comment sheets. *Assessment & Evaluation in Higher Education*, 28(5), 515-526.
- Shute, V. J. (2008). Focus on formative feedback. *Review of Educational Research*, 78(1), 153-189. doi:doi:10.3102/0034654307313795
- Stubbs, M. (1996). *Text and corpus analysis: Computer-assisted studies of language and culture*. Oxford: Blackwell Oxford.
- Sutton, P. (2012). Conceptualizing feedback literacy: knowing, being, and acting. *Innovations in Education and Teaching International*, 49(1), 31-40. doi:10.1080/14703297.2012.647781
- Universitetet i Oslo. (2016). 5-timers videoeksamen [5-hour video examination]. Retrieved from <https://www.uio.no/studier/emner/uv/ils/PPU3210/h16/5-timers-videocaseksamen.html>
- Universitets- og høyskoleloven (2005). Lov om universiteter og høyskoler [Act relating to universities and university colleges]. ((LOV-2005-04-01-15)). Retrieved from https://lovdata.no/dokument/NL/lov/2005-04-01-15/KAPITTEL_1-5#%C2%A75-3
- Universitets- og høyskolerådet. (2011). Karaktersystemet - generelle, kvalitative beskrivelser [The grading system - general, qualitative descriptions]. Retrieved from https://www.uhr.no/f/p1/i4bfb251a-5e7c-4e34-916b-85478c61a800/karaktersystemet_generelle_kvalitative_beskrivelser.pdf
- Villarroel, V., Boud, D., Bloxham, S., Bruna, D., & Bruna, C. (2019). Using principles of authentic assessment to redesign written examinations and tests. *Innovations in Education and Teaching International*, 1-12. doi:10.1080/14703297.2018.1564882

- Vygotsky, L. S. (1978). *Thought and language* (E. Hanfmann & G. Vakar, Trans. A. Kozulin Ed.). Cambridge, MA: MIT press.
- Warhuus, J. P., Blenker, P., & Elmholdt, S. T. (2018). Feedback and assessment in higher-education, practice-based entrepreneurship courses: How can we build legitimacy? *Industry and Higher Education*, 32(1), 23-32.
- Bailey, R., & Garner, M. (2010). Is the feedback in higher education assessment worth the paper it is written on? Teachers' reflections on their practices. *Teaching in Higher Education*, 15(2), 187-198.
- Warhuus, J. P., Blenker, P., & Elmholdt, S. T. (2018). Feedback and assessment in higher-education, practice-based entrepreneurship courses: How can we build legitimacy? *Industry and Higher Education*, 32(1), 23-32.
<http://search.ebscohost.com/login.aspx?direct=true&db=eric&AN=EJ1168035&site=ehost-live>
<http://dx.doi.org/10.1177/0950422217750795>