Cambridge Assessment

Learning to think alike: A study of professional examiners' feedback interactions in a UK Qualification Awarding Organisation

Conference Paper

Dr Martin Johnson

Presented at the 9th international conference of the EARLI SIG 14 - Learning and Professional Development

University of Geneva, Switzerland

September 2018

Author contact details:

Dr Martin Johnson
Assessment Research and Development,
Research Division
Cambridge Assessment
The Triangle Building
Shaftesbury Road
Cambridge
CB2 8EA
UK

Johnson.M2@cambridgeassessment.org.uk

http://www.cambridgeassessment.org.uk

As a department of Cambridge University, Cambridge Assessment is respected and trusted worldwide, managing three world-class examination boards, and maintaining the highest standards in educational assessment and learning. We are a not-for-profit organisation.

How to cite this publication:

Johnson, M. (2018, September). *Learning to think alike: A study of professional examiners' feedback interactions in a UK Qualification Awarding Organisation.* Paper presented at the 9th international conference of the EARLI SIG 14 - Learning and Professional Development, University of Geneva, Switzerland.

Abstract

In the UK, many school and professional qualifications are delivered and administered by awarding organisations. The administration of qualifications involves a great number of examiners who are responsible for assessing individual performances according to pre-specified mark schemes.

In the case of the awarding organisation in this study, groups of examiners work under the remote supervision of a senior examiner (team leader). A team leader is responsible for overseeing the training of a group of examiners, and for ensuring that high quality marking performance is maintained throughout the marking period for that group of examiners. Feedback communication has an important role in this training and monitoring process. During the training and supervision period team leaders give feedback to each examiner on their work quality via electronic mail or telephone messages.

My study data comprises of 991 professional feedback interactions that take place between three TLs and 27 examiners working for a UK-based awarding organisation, and my research sets out to explore which features of professional feedback make it effective for professional learning.

For my study I adopt a sociocultural perspective on learning that suggests that interaction supports the development of reasoning through the alignment of culturally appropriate collective thinking (Littleton & Mercer, 2013). In other words feedback affords professional learning by allowing examiners to gain insights into the marking procedures and linguistic interpretations that characterise the professional examiner community of practice.

As a consequence, this perspective suggests that it is crucial to understand the nature of interaction as intrapersonal development is a function of the quality of interpersonal communication. In addition, I am also interested in the way that the team leaders use feedback discourse to coordinate divided labour across their marking teams. Drawing on the notion of *Articulation Work* (Strauss, 1985), I analyse how feedback involves the often unnoticed and taken for granted work that is carried out during interactions to ensure that coordinated task completion is managed.

I analyse the feedback messages using a *Sociocultural Discourse Analysis* approach (Mercer 2004) which employs methods informed by the traditions of Discourse Analysis, Ethnomethodology, and Corpus Linguistics. This mixed methodology enables me to develop a picture of feedback practice that is both contextually and theoretically grounded.

My study provides a methodology for analyzing devolved, remote communication and my outcomes show that effective feedback has a number of both transactional and interactional qualities (e.g. Brown & Yule, 1983). These outcomes are useful for supporting feedback giving practices in other professional environments where remote interactions are used as a mechanism for inducting mew members into a community of practice.

Learning to think alike: A study of professional examiners' feedback interactions in a UK Qualification Awarding Organisation

Introduction

The large scale delivery of school examinations in the UK is a highly regulated practice. In England, Wales and Northern Ireland, educational qualifications are offered by designated awarding organisations. These organisations are recognised as being eligible to award qualifications by the national body that regulates qualifications and examinations, the Office of Qualifications and Examinations Regulation (Ofqual). The Oxford, Cambridge & RSA (OCR) awarding organisation that is the focus of this study is one of the main three awarding organisations in England, Wales and Northern Ireland.

Ofqual sets out its expectations for the quality assurance (QA) processes that awarding organisations need to adhere to for qualification accreditation in its General Conditions of Recognition documents (Ofqual, 2016, 2017). These QA expectations are an important influence on the ways that examiners interact. In practice, awarding organisations tend to use hierarchic marking arrangements, where the mark scheme interpretations of the most senior examiners (sometimes known as 'principal examiners') are cascaded down the examiner hierarchy, via other senior examiners (such as 'team leaders'), through to the broader examiner body. This process seeks to ensure that marking decisions are consistent across examiners. Figure 1 is a graphical representation of the way that large scale marking is structured.

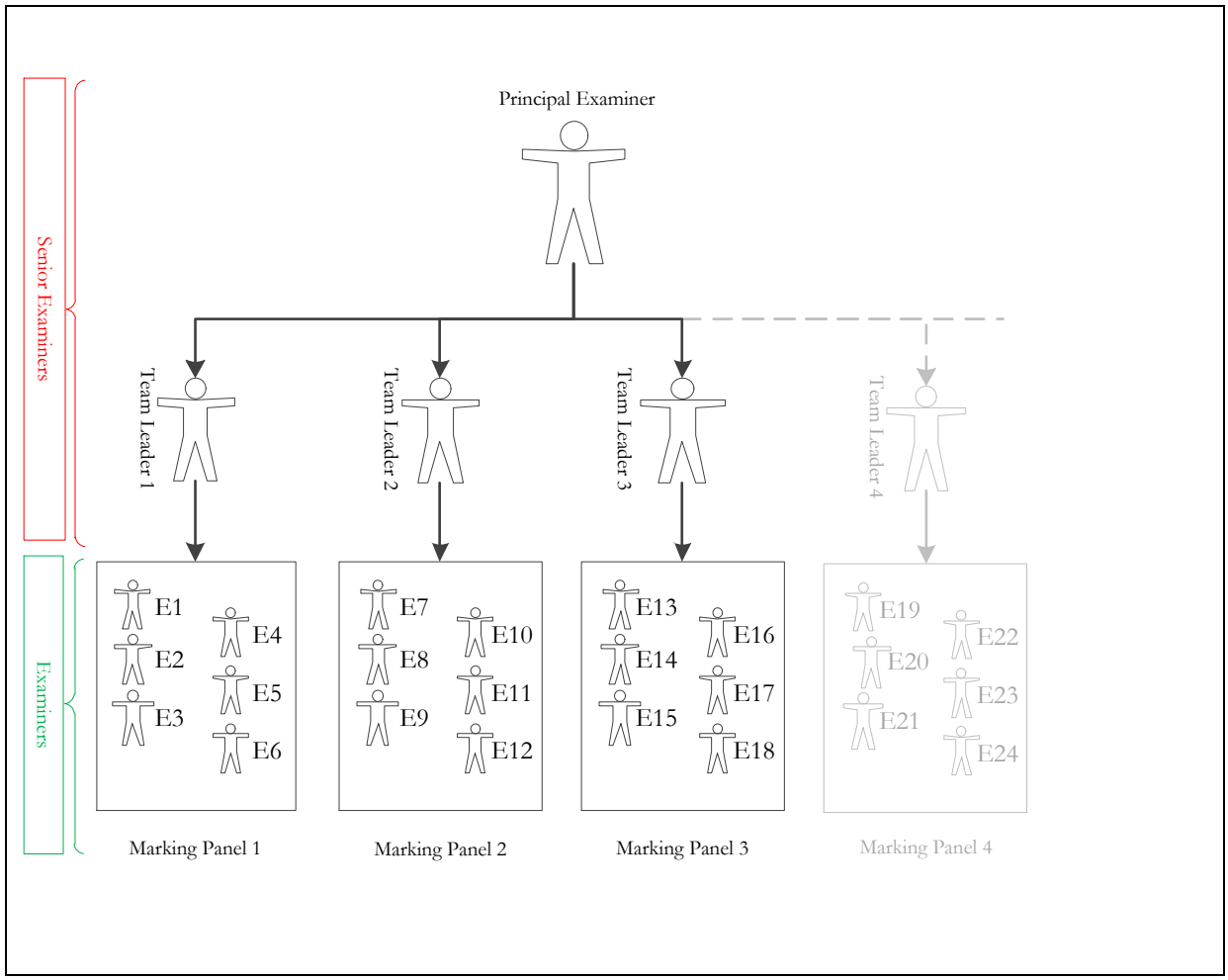

Figure 1: Senior examiners and examiners

UK awarding organisations commonly use technology to distribute digital copies of students' examination scripts to examiners for marking. This process allows senior examiners (team leaders) to monitor the marking quality of the examiners under their supervision throughout the marking period.

As part of this monitoring activity, team leaders are also required to give examiners feedback on their marking. This monitoring and remediation function is an important component of the awarding organisation's QA arrangements that are in place to ensure that the marking process results in fair and equitable assessment outcomes.

Figure 2 is a screen shot of the view of part of the monitoring and feedback process from a team leader's perspective. In common with examiners working for other large awarding organisations in the UK, examiners at OCR communicate feedback to the examiners in their team via the electronic messaging systems that are built into their digital marking environment software or via telephone conversations (AlphaPlus, 2014; Johnson, 2016a;

Ofqual, 2013). At the top of the figure there is a view of an examiner's marked script, with the team leader then shown responding to this marking through a feedback message to the examiner (at the bottom of the figure).

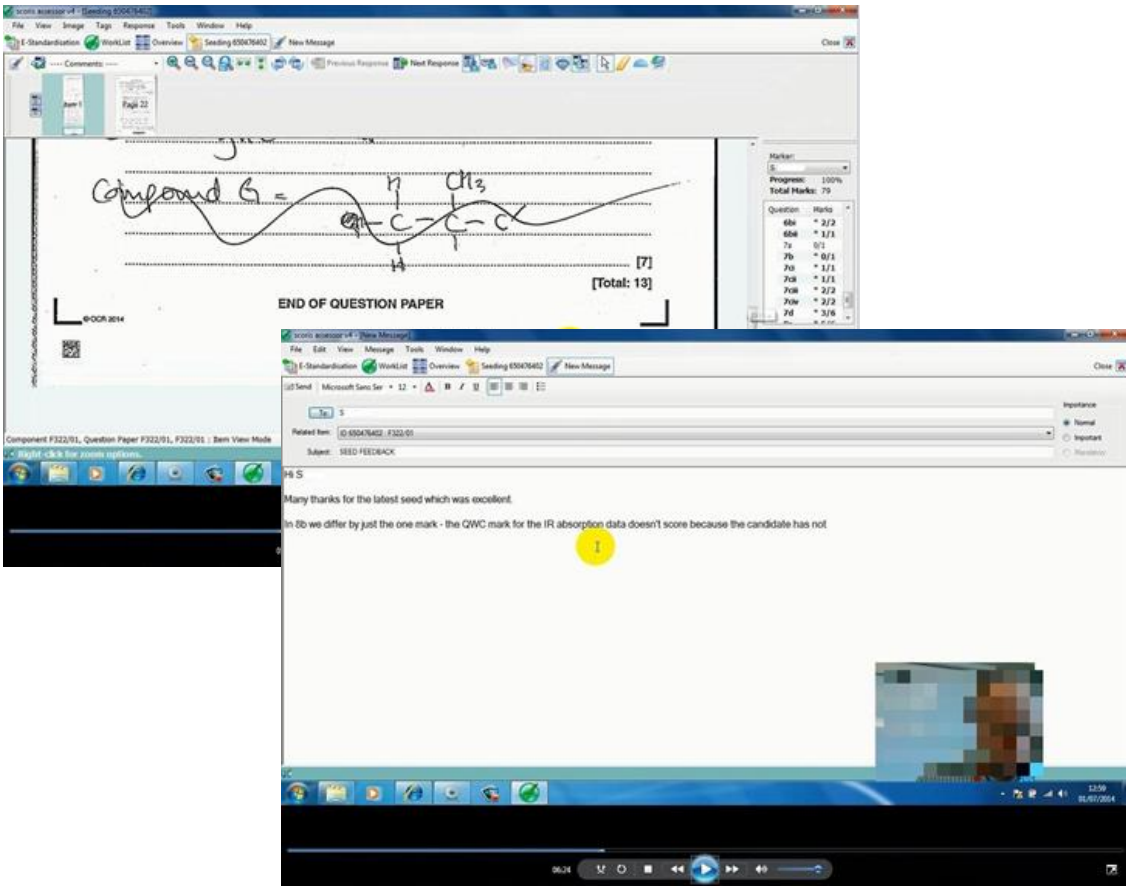

Figure 2: Team leader feedback on an examiner's marking

As well as having a crucial QA function, previous work has suggested that feedback that seeks to support the alignment of less senior examiners' marking decisions with those of more senior examiners can also be conceptualised as having an expansive developmental potential for the less senior examiners (Johnson & Black, 2012). Expansiveness is a concept that describes how some contexts help new participants in a professional community to gain access to the important knowledge and values that then allow them to go on to become more independent participants in an activity (Fuller & Unwin, 2003). I argue, in line with Beighton, Poma, & Leonard (2015), Dennen (2004), and some situated learning theorists, that this concept of expansion has important links to learning, since a development in the

understanding of professional practice in an area is synonymous with *learning to be a professional*. This expansiveness includes the type and extent of knowledge transfer, the quality of emotional and practical support for participants, and the appropriate alignment of individual objectives.

Literature Review

Previous literature suggests that communication context has implications for communication quality, and that this in turn has an influence on the potential quality of professional learning and development. For example, some research that looks at the connections between communication mode and communication quality has shown that the ability to provide information in various formats (such as simultaneous verbal and non-verbal forms) helps to convey content that is more nuanced, and that can enhance participants' development of a shared view (Bietz, 2008; Münzer & Holmer, 2009; Stannard, 2008).

The notion of shared view is important as it is a component of the expansiveness (mentioned earlier) that supports both the alignment of perspectives and the learning that comprises professional development. It is possible that a lack of alignment is related to the way that the affordances of a communication mode influences the transactional distance between examiners. Transactional distance describes the psychological and communicative space of potential misunderstandings that can exist between participants (Ackerman & Gross, 2010; Moore, 1993; Murphy & Rodriguez, 2008; Rovai, 2000; Vonderwell, 2003).

In this section I introduce and bring together two sets of theory that are relevant for exploring how feedback supports alignment and examiner development; *Intermental Activity* (Vygotsky, 1978), and *Articulation Work* (Strauss, 1985). Although emerging from different disciplines, the former from education psychology and the latter from workplace sociology, both share a common focus on the way that social interaction influences individual thinking and action. In this way, these theories relate to models of learning which consider education to be both 'an interpersonal and intrapersonal process' (Mercer, Littleton, & Wegerif, 2004, p. 203).

Looking first at *Intermental Activity*, this notion is underpinned by a sociocultural perspective. This perspective suggests that communication and language use is a crucial area of study for understanding the development of individual thinking and learning that derives from social interaction. Although sociocultural research is still exploring how language influences the transformation of reasoning (Littleton & Mercer, 2013, p. 99), the perspective suggests that language supports the development of participants' reasoning through the alignment of culturally appropriate collective thinking.

Building on this sociocultural perspective there is a belief that the quality of an individual's learning and development is related to the quality of interpersonal communication during learning, e.g. the quality of teacher and learner or peer learner communication (Johnson, 2016b; Mercer, Wegerif, & Dawes, 1999). A growing body of research evidence now suggests that productive learning communication (i.e. communication that attains its learning purpose) relies on the participants developing and maintaining common ground through their discourse (Beers, Boshuizen, Kirschner, & Gijselaers, 2007; Clark, 1996; Clark & Brennan, 1991; Edwards & Mercer, 1987; Mercer, 2000; Whittaker, 2003).

According to Mercer (2000) and Littleton & Mercer (2013), who have looked at productive discourse in terms of talk that is effective for learning, there are some forms of talk that are better suited for learning compared with others. *Disputational Talk* is the least productive form of discourse as it is dominated by assertions and counter-assertions, disagreement and individualised decision-making. On the other hand, *Cumulative Talk* and *Exploratory Talk* involves the participants making judicious linguistic choices that reference features of the shared context (Clark, 1992). These forms of discourse encourage collective thinking in the way that they allow 'ideas [to be] shared openly and possible explanations [to be] considered critically but in an atmosphere of trust' (Littleton & Mercer, 2013, p. 93).

In the specific context of examiner feedback, the concept of intermental activity anticipates that examiners develop their understanding of a mark scheme through sharing their (sometimes contrasting) perspectives with their team leaders on a marking performance. In developing this 'shared view', examiners draw on communal resources (such as mark scheme documents or candidates' exam scripts) that invoke concepts that they believe to reside within the cognition of each other. A consequence of this theory is that where common ground in feedback communication is weak it is possible that communication will break down, that examiners will fail to establish shared understandings, and that less senior examiners will not become a full participant in the professional examiner community.

The second area of relevant theory that I draw on links to the notion of *Articulation Work* (Schmidt, 1994, 2011; Strauss, 1985). This concept describes how communication helps to coordinate individuals whose work is professionally interconnected. According to Schmidt and Strauss, articulation work involves the often unnoticed and taken for granted work that is carried out by managers to ensure that those around them complete their own tasks, and thus attain mutually important strategic goals.

I find this phrase to be particularly useful for my study context because it simultaneously references the notion of *expressing* and the act of *coordinating interconnected work across*

individuals. Both of these elements appear to coalesce around feedback giving activity (Johnson, 2015).

In the examination marking context it is likely that team leaders will be very conscious of the need to ensure that competent examiners (or those who the team leader believes will become so) are motivated to complete their marking tasks. The consequence of losing competent examiners from the workforce before their marking is completed represents additional workload for the team leader, and a concomitant strain on their relations with other examiners who then share the additional workload.

The combination of theories that I have included in this study has implications for research method. This is because the focus on professional interaction draws attention to the importance of evidencing the minutiae of the professional behaviours that participants carry out and which often go unnoticed as they are generally taken for granted. Bringing together the two areas of theory, my study considers the articulation work that team leaders carry out through their feedback communication as they build and maintain common ground with the examiners in their team.

Method

My study focused on three Advanced level General Certificate of Education (GCE)[1] subjects (Chemistry, Economics and Geography). These subjects were chosen because they included examination performances (scripts) that incorporated subjective items. These items tend to invite performances that require higher order skills, and inevitably involve intricate decision making on the part of an examiner when applying the mark scheme. Such items are considered to be the most complex item type, and they tend to result in lower levels of examiner agreement (Bramley, 2008; Massey & Raikes, 2006), so I anticipated that they would produce rich between-examiner feedback interactions.

In my study I captured all of the feedback messages that were given by three team leaders to the 27 examiners that they worked with in their respective marking teams over two different examination sessions. The feedback data included all of the email messages that were conveyed through the examiners' digital marking system, as well as any messages that were communicated by telephone. These data comprised 991 messages. In addition to the raw

[1] Advanced level General Certificate of Education (GCE) courses are usually studied over a two-year period and are widely recognised in England, Wales and Northern Ireland as being the standard entry qualification for assessing the suitability of applicants for academic courses in UK Universities.

message data I was also able to observe and interview all of the team leaders during the feedback giving process, as well as interviewing a sample of 13 examiners. This allowed the participants to reflect on the rationales behind, and their reactions to, particular feedback practices. Table 1 outlines the characteristics of the examiners who took part in the study, along with how they were involved in data gathering.

Table 1: Team leader and examiner demographic data

Chemistry					Economics					Geography				
Team leader 1: *Male; Age: 50s; Examiner for 20+ years; Team Leader for 8 years; a retired teacher of Chemistry from a Comprehensive School; still continues to tutor students privately*					Team leader 2: *Male; Age: 20s; Examiner for 5 years; Team Leader for 2 years; currently a teacher of Economics in a Comprehensive School*					Team leader 3: *Female; Age 50s; Examiner for 20+ years; Team Leader for 5 years; a retired teacher of Geography from a Comprehensive Sixth Form College*				
Examiner	♀♂	Exp[A]	Fam[B]	Int[C]	Examiner	♀♂	Exp[A]	Fam[B]	Int[C]	Examiner	♀♂	Exp[A]	Fam[B]	Int[C]
1	F	*	*		8	M	*	*	†	14	M			
2	M	*	*	†	9	F				15	M	*	*	
3	F				10	F	*	*		16	F	*	*	
4	F	*	*		11	M			†	17	M			
5	M			†	12	M	*		†	18	M			*
6	M	*	*		13	F				28	M	*	*	
7	F				24	M			†	29	F			*
19	F			*	25	M	*	*	†	30	M			
20	M	*	*		26	M	*	*	†					
21	M			†	27	M	*	*						
22	M	*												
23	F	*	*	*										

[A] Experience (the examiner has examined in a previous marking session); [B] Familiar to the team leader (worked together previously); [C] Interviewed examiner (*Face-to-face/†Virtual)

As I noted earlier, my choice of analytical approach was influenced by my theoretical perspective, which culminated in my adoption of a method that foregrounds the fine grained analysis of professional communication and interaction behaviours. In brief, sociocultural approaches seek to gain insights into the social interactions that underpin learning and development. This entails evidencing the interplay between externalised concepts expressed through language, and the impact that this has on internalised thinking processes. In this way sociocultural approaches study the ways that meanings and understandings are constituted by participants in social action, and take into consideration both the material and the semiotic elements of language use (Gee & Green, 1998). Mercer (2004) and Littleton & Mercer (2013) have distilled these elements into a sociocultural discourse analysis approach (SCDA) that can be distinguished from other forms of discourse analysis through its focus, amongst other things, on how common ground is established through the dimensions of communication *content* and its development *over time* (Mercer, 2008). According to sociocultural theory, discourse change over time can be indicative of learning. The things that come into

discourse may suggest that they are addressing a need, whilst things that fall away from discourse may represent areas of common ground that have been established (and no longer need stating). This broad analytical focus invites a critical reflection on methods use, and encourages an eclectic approach to data gathering that is focused on the specific elements of theoretical interest.

To analyse the feedback data content I employed methods that were closely aligned with three different methodological approaches to discourse analysis; *Thematic Content Analysis*, *Conversation Analysis*, and *Corpus Linguistics*. Thematic content analysis is a 'respected…well-established' (Boje, 2001, p. 122) and 'widely used' (Braun & Clarke, 2006, p. 101) approach that is assumed to be a relatively unproblematic method for gathering and representing language units of conveyed meaning. The identification of themes rests on a coding process which involves the systematic application of codes to a text to capture the dominant ideas or constructs within it.

Conversation Analysis tends to focus on very specific interaction episodes that are relatively limited in scope, and with 'close and detailed analysis of small extracts or fragments [helping] to develop a disciplined understanding of these episodes' (Llewellyn & Hindmarsh, 2010, p. 26). In this way it is possible to consider all aspects of the interaction, since apparently irrelevant or mundane features potentially perform important functions.

Corpus Linguistics is an approach that uses computer software to identify text characteristics. Generally working with large data sets it compares text features and uses statistical measures to validate the robustness of its outcomes. A perceived advantage of the approach is that it uses software to search large corpora, and that this is more efficient than manual human search functions and minimizes the errors associated with manual human coding.

The rationale for this integrated approach was to enable my analysis to capture evidence of the global, generic elements of language use as well as the particular, contextualised aspects of language use. This approach also allowed me to integrate a qualitative dimension to my analysis, using a framework to consider why humans tend to interact in certain ways at specific times, as well as to employ specialist software to overcome some of the limitations that pertain to human analysts (e.g. the challenges of identifying patterns across a large dataset). Figure 3 outlines the areas of methodological overlap that I sought to exploit through this combined analysis (for more on the details of this analysis see Johnson, 2017).

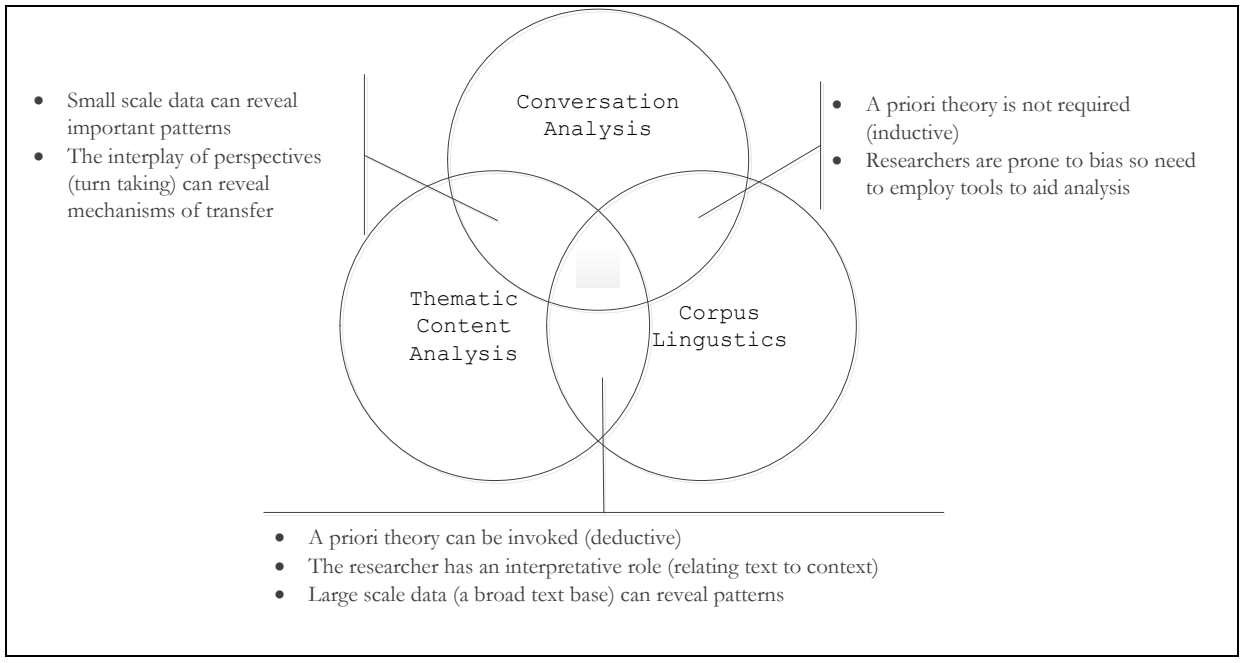

Figure 3: My integrated methodological approach for feedback analysis (from Johnson, 2017)

A final benefit of the adoption of this methodological approach was that it enabled me to capture both the transactional and the interactional dimensions of feedback discourse (Brown & Yule, 1983). Transactional dimensions include the actual content of feedback, whilst the interactional dimensions include stylistic choices around the ways that such content were presented and the intentions behind these choices.

In the next section I outline the outcomes of my analysis. These principally focus on the content features of the feedback communication as well as consideration of the dimension of content change over time.

Outcomes

Content

One of the principal outcomes of my analysis was to be able to identify the types of content that were included in feedback messages. Table 2 outlines the five types of information that were most commonly found (in descending order of prevalence).

Table 2: Common information types found in feedback messages

Content	*Definition*	Example
Locating Credit	*Pointing out that there is a difference without a rationale (the examiner needs to fill the missing information)*	`We gave the 'explanation' mark here`
Rationalising credit	*A rationale for a mark decision is given, making the team leader thinking explicit*	`We gave the 'explanation' mark here because...`
General	*Non-specific/non-concept related information*	`Your marking is good`
Technical	*Conveying system level information*	`Annotations: remember to put N/R to all questions not answered.`
State mark	*Number only statements*	`Q1 is 1 mark`

These analyses show that the basis of common ground between examiners was focused on where and why marks were allocated for specific performance features. There was also communication around the general character of an examiner's marking performance, as well as some technical information that was specific to the marking software context in which the examiners worked.

Disagreement

Whilst the findings around content allow insight into what information contributes to the alignment of examiners' thinking (e.g. shared information about where and why marking credit is found in a performance, or how examiners are expected to use the specific marking software), these data are only one part of the story.

An important point to consider is that feedback information was generally shared within the broader context of explicitly or implicitly stated disagreement between the participants. Analysis was carried out to measure the relative balance of agreement or disagreement information within each feedback message. This analysis showed that disagreement was conveyed in around 88% of all of the feedback messages, and that it accounted for the major part of around 70% of the messages in total (Figure 4).

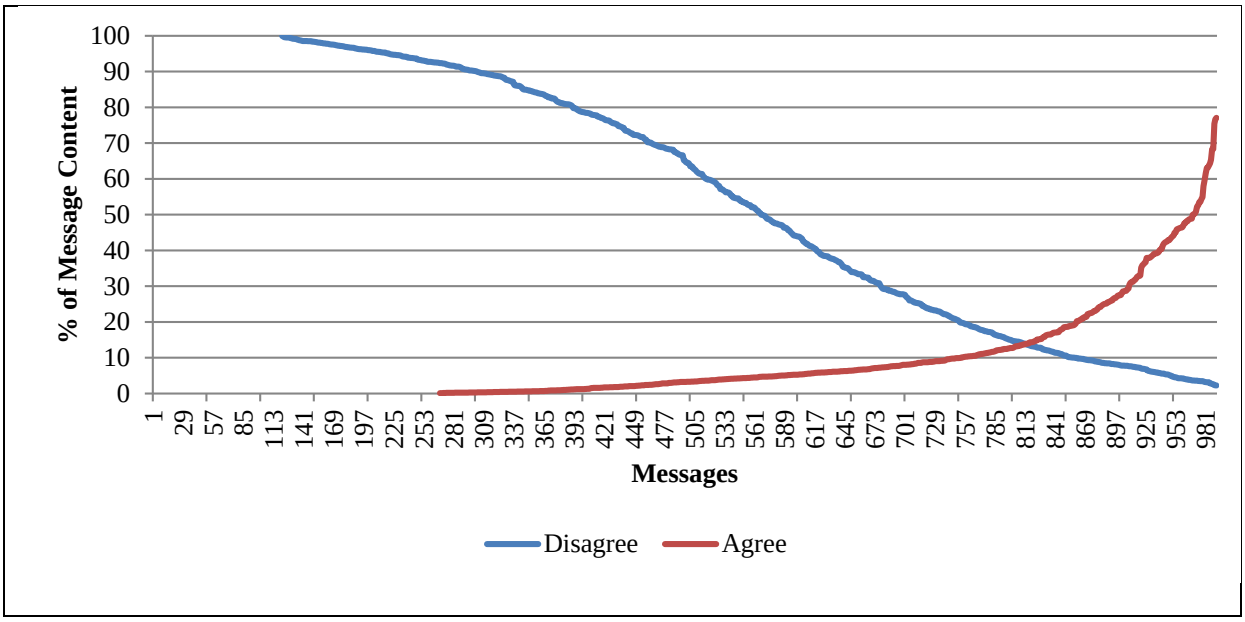

Figure 4: Relative amounts of disagreement and agreement content in feedback messages

Time and resolution

The observation that examiners tended to focus on disagreement may not be surprising (given that the aim of their work was to establish agreement). Analyses were also able to give insight into how the participants interacted to attain a movement towards the resolution of disagreement.

Keyword analysis identified those words that were used with an 'unusual frequency' (Scott, 1997, p. 236) by particular participants. This analysis showed that word use changed over the course of feedback interactions. Table 3 shows those words that were used more frequently by team leaders in the first chronological quarter of feedback discourse compared with the final quarter of discourse. For ease of analysis, the table is also organised to show the words that are generally indicative of *exploratory* discourse (Mercer, 2000), or that are indicative of referencing links between things (a feature of common ground building).

Table 3: Keyword use and discourse time

Quartile 1			Quartile 4		
Team Leader	*Exploratory*	*Reference*	*Team Leader*	*Exploratory*	*Reference*
mark		here	bit	because	again
here			many	thought	they
correct				if	you
only				think	your
answer					me
some					my
gave					
awarded					
need					
please					

This analysis suggests that the earliest quartile was characterised by the team leaders engaging in more directive activity than in the final quartile (e.g. ‘we gave a mark here’, ‘we only awarded this’, ‘the correct mark here is’). In the final quartile of the discourse there is a greater presence of terms that are potentially associated with exploratory discourse (e.g. ‘we gave this because’, ‘I thought that it’, ‘I think this’) and personalised connections with actions (‘you’, ‘your’, ‘me’, ‘my’).

Another indication of time was considered through the analysis of the movement towards the attainment of resolution across a series of linked messages (i.e. a public acknowledgement that any differences in professional judgements had been resolved). Figure 5 shows the types of discourse codes that contributed to the attainment of resolution where examiners disagreed around a mark. The light blue codes are team leader discourse codes, the orange code is an examiner only code, and the dark blue codes in bold font are codes that are used by all participants.

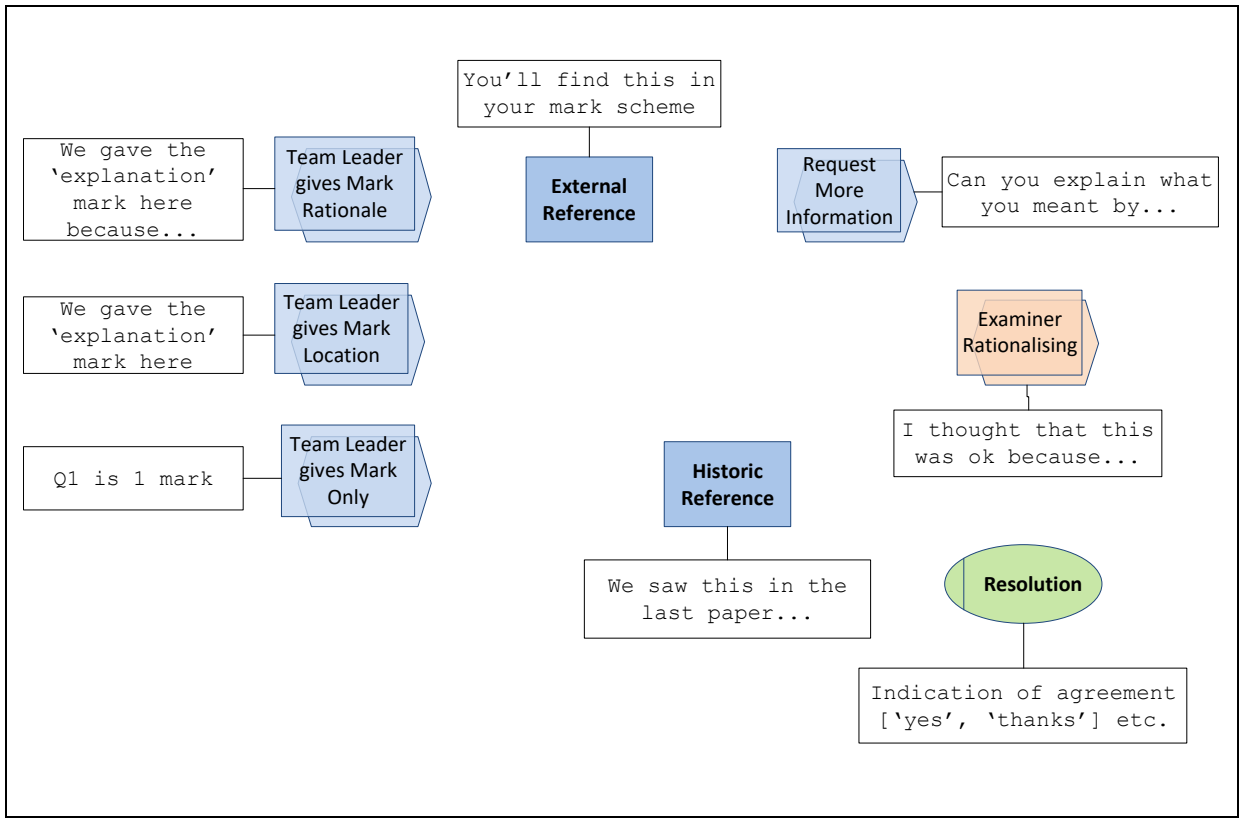

Figure 5: Discourse codes that interacted around resolution

Figure 6 shows the scale of the discourse code interaction around resolution. In this case the most frequently associated codes are represented by the thickest blue lines.

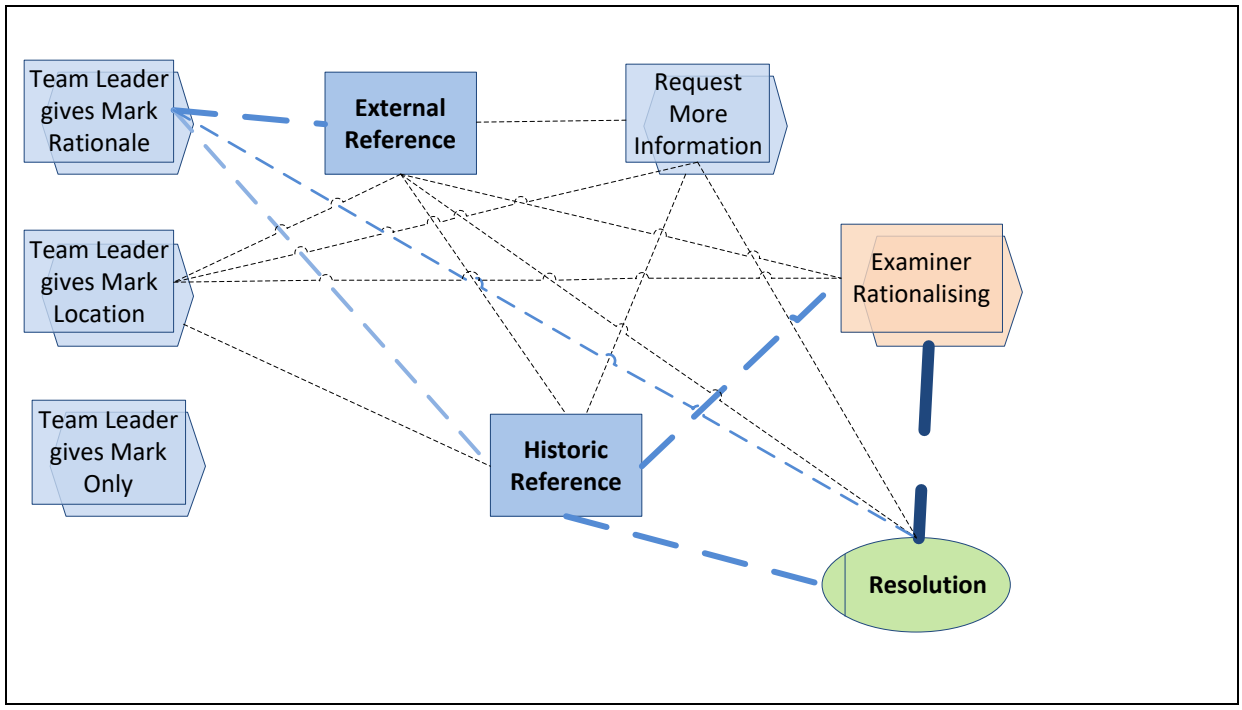

Figure 6: The most common discourse codes that interacted around resolution

This analysis shows that reference making (to other documents or to a shared past experience) was an important mediator of the process of resolution. The analysis also shows that rationalising thinking was also influential in resolution attainment. It is noteworthy that the statement of a mark only was not identified as being useful for mediating the resolution process.

Examiner familiarity and experience

My analyses showed that the characteristics of examiner prior experience and familiarity also had an impact on the types of feedback given. Looking closely at word count indicators, using a Mann-Whitney U nonparametric test, I found that there was significantly more feedback communicated between team leaders and new examiners than between team leaders and experienced examiners (U = 91596.500, p = <.001). Messages to new examiners contained on average 116 words, whilst message to experienced examiners contained on average 75 words.

In addition, and although the groups were highly overlapping, there was significantly more feedback communicated between team leaders and unfamiliar[2] examiners than between team leaders and familiar examiners (U = 100953.000, p = <.001). Messages to unfamiliar

[2] Examiners who had not worked together in the past.

examiners contained on average 101 words, whilst message to familiar examiners contained on average 80 words.

These analyses suggested that team leaders were targeting and adapting their communication to the needs of their examiners in different ways, with new and unfamiliar examiners receiving more feedback than other examiners.

Distancing strategies

To make sense of these discrepancies I was able to consider any differences in the nature of the information that was being conveyed between these different groups of examiners. The clearest difference was in the way that team leaders employed 'distancing strategies' with new and unfamiliar examiners (compared with other examiners).

In my study, distancing strategies describe the deployment of politeness in discourse. It has already been noted that the presence of negative information in social interaction, such as criticism, disagreement, and interruption, embroils issues of face management. Importantly, it has been observed that politeness can minimise face threat (Goffman, 1967; Morand, 2000), and has been found to be used in professional contexts where bad news needs to be delivered (Sussman & Sproull, 1999).

Theorist also observe that politeness may be of a positive or a negative variety, with each affording the user the opportunity to either increase or reduce the perceived social distance in interaction (Brown & Levinson, 1987). Positive politeness reduces the threat to the recipient's positive face by accentuating empathy and common ground between the participants, therefore acting as a kind of social accelerator. These positive politeness tactics include comments on admiration and the use of 'in group' speech forms (e.g. the use of ellipsis and the inclusive pronoun form 'we'). On the other hand, negative politeness avoids imposition on the recipient's negative face (i.e. the desire to act unimpeded) by creating a respectful distance between the participants. Negative politeness tactics act as a form of 'social brake' (Culpeper, 1996) through the judicious use of words to construct messages that include apology, verbal hedging, and honorific term use.

My analyses showed that distancing strategies were used more frequently with new examiners (U = 103452.500, p = <.001) and with unfamiliar examiners (U = 110439.500, p = .008). Over the course of their feedback discourse, on average, new and/or unfamiliar examiners' messages contained around 27-33 distancing strategies, compared with a range of 21-26 distancing strategies for experienced and/or familiar examiners' messages.

Careful analysis of the feedback discourse showed that distancing strategies were deployed in a number of ways across the feedback corpus. These included the personalisation of messages, the use of apology, and the use of modal forms.

Greetings and closings

Nearly all of the feedback messages included a greeting and a farewell statement. Analyses suggest that the participants used these openings and closings in a purpose driven way so as to achieve particular effects. For example, one team leader, (*Roy*)[3] explained how in general he preferred concise, targeted message writing: *"I don't need to waste [words], the potency of the message goes in the more words you use in my opinion."* At the same time he rationalised how the effort expended on personalising messages, including elements such as informal greetings and closings (Figure 7), may have a motivating impact on this particular examiner.

```
Hi Eric, This is out by quite a bit again but I can see the calls you have
made (except for one mistake). Can you please carefully review the below?
…
[A list of five detailed marking points to consider]
…
Thanks, Roy
```

Figure 7: Personalised greeting and closing

When discussing this feedback message the team leader reported: *"I have been at the other end of this and you really just want to know where you have gone wrong… I honestly don't want to dishearten [Eric]. To be fair the message was 'You are out'. There is a human interaction here. Whereas actually when you are giving just very straight forward feedback 'This is right, this is wrong', you don't need as much as that, but you need to be a bit softer [here] I guess."*

Apologies

The use of apology is a strategy for manipulating the perceived social distance in interaction. The feedback extract in Figure 8 shows how apology is used by an examiner (*Teresa*) to preserve the negative face of the team leader (*Serena*) and to reduce intrusion into their professional space.

[3] All names have been changed to preserve participant anonymity.

Sorry Serena another question, Q8 p10 - is this enough for L3B2? - thanks Teresa

Figure 8: Apology use in written feedback

This pattern is also noticeable in the telephone extract in Figure 9, where a team leader (*Ben*) has to inform an examiner (*Gerry*) that he has to send some standardisation scripts back to the examiner for reconsideration.

```
001   Ben   Hi. Gerry?[4]
002   Gerry Yes speaking
003   Ben   Ben. Hi. I'm s:: I'm sorry I had to send you back another set
004         of scripts but erm unfortunately with the first batch being a
005         a little bit over the [limit
006   Gerry                       [I was a bit yeah]
008   Ben                                          Yah], I'm afraid I've got
009         to send them back so you'll have to do another set I'm afraid?
010         And then submit those before we can [be]
011   Gerry                                    [Yeah]
012   Ben   Up and running. Now is there anything in particular?
```

Figure 9: Apology use in telephone feedback

In this feedback Ben manages engagement through underplaying the seriousness of the disagreement (line 005) [little bit]. The use of apology also reinforces the dispreference related to giving bad news (lines 003, 004, 008, 009) [sorry] [unfortunately] [afraid]. Ben then shifts the focus of the conversation (line 012), emphasising [now] before Gerry can dwell on the disagreement. This is a shift towards positive help following the delivery of negative news.

A closer look at the use of apology also gives insights into how the participants maintained order and on-going professional interaction in a context where face threat was present. Log likelihood ratio analyses that identify keywords (i.e. words that are used significantly more frequently than others in a discourse) showed that the word 'please' was used more by team leaders than by examiners (*p=<0.05*). In addition, a search of the whole corpus using the search terms [sorry][apol*] located 142 instances of apologetic utterance. Most of these apologies (n=125) took used negative politeness forms (e.g. 'Sorry this feedback is a bit lengthy').

[4] The Conversation Analysis conventions used in this transcription are adapted from Jefferson (2004).

. *short pause*; :: *long pause*; ? *high rise*; [ ] *overlapping talk*; Now *emphasis*

Modals

At times team leaders were seen to soften the definitiveness of their judgement through the use of modal forms (e.g. could, may, might). Phrases that use these types of words are sometimes called 'hedges' and they express tentativeness and avoid strong statements that may be construed as being confrontational (Lakoff, 1973).

In the feedback extract in Figure 10 the team leader (*Ben*) responds to a message from an examiner (*Tony*) by embedding his comments in the original email wording (indicated in red font). Tony has alerted the team leader to an apparent mixed message in the mark scheme (lines 004-008), and then asks for clarification on a marking point (lines 012-013). In his response, Ben's disagreement is weakly stated (line 014). Ben also softens the definitive nature of the responses through the use of modals on lines 010/011 (may, if), which reduces the implication that the examiner is completely incorrect.

```
001   Hi Tony
002   Thank you for the feedback, I have amended the 2 you sent back to me.
003   2 queries:
004   [Script] ID 649581302 – Q1 g ii – MS says 'it' should be assumed to
005   mean cyclohexane. Do they still need to have written cyclohexane
006   somewhere in their answer to get the mark? I accepted 'It burns more
008   effectively'.
009   I cannot find the comment re: 'assumed to be cyclohexane' in the mark
010   scheme – it may have appeared in the practice scripts by the sound of
011   it, and was incorrect if it did.
012   [Script] ID 649661411 – Q2b – do they get the mark even though
013   'curly' arrow is almost straight?
014   I am afraid so – possibly a little generous.
015   I will look through the other 5 and send over.
016   Thanks.
017   Ben
```

Figure 10: Modal use in feedback

Discussion

Through my close analysis of team leader and examiner feedback discourse I have been able to gain insight into the nature of the communication that supports distributed marking processes, with the developmental aim of supporting examiners to think alike. Analysis of feedback content shows that examiners are given important information that steers their practice. This communication content helps examiners to refine their interpretations of mark schemes and helps to reduce any marking discrepancies between examiners and more

senior examiners/team leaders. This content can be interpreted as being a component of an expansive learning environment (Fuller & Unwin, 2003) since it gives new examiners access to the important knowledge that then allows them to go on to become more independent markers.

This analysis also draws attention to the *intermental* nature of professional development, with examiners developing their understanding of a mark scheme through receiving (sometimes contrasting) perspectives from their team leaders on a shared marking performance. Feedback content frequently focused on the location of, and the rationalisations for, marking credit. Drawing on learning communication literature, this content can be interpreted as providing the foundation for the team leaders and the examiners to develop a shared view. According to this perspective, productive learning communication relies on the participants developing and maintaining common ground through their discourse (Beers, Boshuizen, Kirschner, & Gijselaers, 2007; Clark, 1996; Clark & Brennan, 1991; Edwards & Mercer, 1987; Mercer, 2000; Whittaker, 2003).

The acknowledgement of time as a focus for analysis is a key element of my theoretical framework. My analyses of cases where team leaders and examiners worked from disagreement to resolution allowed me to consider how alignment was achieved across the course of feedback discourse. In particular, the content of what the team leaders and examiners communicated gave insight into the common ground that they established across their discourse. My analyses suggested that team leader feedback discourse moved across distinct stages.

The initial stage of discourse centred on a trouble source; an indicator of disagreement that suggested that the team leader and the examiners held different perspectives on a common examination script. In terms of previous work on the links between discourse and learning, this disagreement has *disputational* characteristics that can threaten productive communication (Mercer, 2000). In the first phase language was used by team leaders to locate the position of credit worthy elements of a performance. Although this type of content was not generally correlated with attaining resolution, it was perhaps a logical first step in the process of building and testing the limits of any shared common ground.

Following an indication of persistent disagreement, the character of feedback shifted towards having a greater emphasis on the team leader sharing their rationale and elaborating on a marking point. This move represented a growth in the base of potential common ground as the team leader provided the examiner with information about the content that they felt was pertinent to the disagreement. Again, in terms of previous work on the links between

discourse and learning, this shift represents a move towards more *cumulative* and *exploratory* discourse, which has links with productive learning contexts (Mercer, 2000).

As well as providing empirical evidence of the transactional content of feedback information, my analyses also give insight into the allied interactional dimension of communication (Brown & Yule, 1983). My analyses suggest that the common ground that is established through feedback interaction, and the expansiveness that is derived from it, is potentially threatened by the prevalence of negative information (marking disagreement) within the communication. Analysis suggests that the structure of team leader feedback communication is influenced by the nature of the information conveyed within the messages. Moreover, this structuring is to some extent conscious and purpose driven on the part of the team leaders. Feedback information that conveys disagreement is a negative basis for establishing productive, on-going relations. Team leaders appear to structure negative feedback messages in ways that attempted to maintain productive engagement through reinforcing an examiner's sense of professionalism. This is most clearly demonstrated in the prevailing use of negative politeness strategies in such messages. This is particularly the case with new and/or unfamiliar examiners, with whom team leaders would be expected to have the weakest common ground.

Drawing on sociological theory, this form of relationship management through feedback can be interpreted as a form of articulation work (Schmidt, 1994, 2011; Strauss, 1985). Articulation work describes the communication work that is carried out to coordinate individuals whose work is interconnected. It also describes the often unnoticed and taken for granted work that is carried out by managers to ensure that those around them complete their own tasks, and thus ensure that mutually important strategic goals are attained. Team leaders use feedback to communicate important content to examiners whilst also mitigating the threats to common ground building that pertains to the negative information that the messages sometimes need to convey. The use of negative politeness helps the participants to maintain a respectful professional distance, and a corollary of this is that marking work is maintained (and not curtailed prematurely due to a lack of *examiner will* rather than *examiner skill*). Having an on-going feedback interaction over time allows a virtuous cycle of examiner development to be constructed. On-going marking experience leads to attendant feedback, a process of examiner reflection, and the consolidation of examiner thinking that is reinforced by a team leader's perspective.

The insights from this study set out the complexity of the feedback giving task, and how it interacts with the nature of professional examiner development. It also gives insights into the

nature of the relationships that foster professional development, and the importance of the forms of communication that lay the foundations for both examiner learning and the completion of marking tasks to a high standard. My analyses illuminate the way that team leaders manipulate the perceived social distance within their remote feedback communication so as to attend to the dual functions of (a) monitoring the standard of examiner marking, and (b) giving examiners information that supports their on-going development.

References

Ackerman, D. S., & Gross, B. L. (2010). Instructor Feedback: How Much Do Students Really Want? *Journal of Marketing Education*, *32*(2), 172–181. https://doi.org/10.1177/0273475309360159

AlphaPlus. (2014). *Standardisation methods, mark schemes, and their impact on marking reliability* (No. Ofqual/14/5380) (p. 70). Coventry: Office of Qualifications and Examinations Regulation. Retrieved from https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/397383/2014-02-14-standardisation-methods-mark-schemes-and-their-impact-on-marking-reliability.pdf

Beers, P. J., Boshuizen, H. P. A., Kirschner, P. A., & Gijselaers, W. H. (2007). The analysis of negotiation of common ground in CSCL. *Learning and Instruction*, *17*(4), 427–435. https://doi.org/10.1016/j.learninstruc.2007.04.002

Beighton, C., Poma, S., & Leonard, V. (2015). Expansive learning in Firearms training: the HE/professional learning interface. *Police Practice and Research*, *16*(5), 378–390. https://doi.org/10.1080/15614263.2014.951045

Bietz, M. J. (2008). Effects of Communication Media on the Interpretation of Critical Feedback. In *Proceedings of the 2008 ACM Conference on Computer Supported Cooperative Work* (pp. 467–476). New York, NY, USA: ACM. https://doi.org/10.1145/1460563.1460637

Boje, D. M. (2001). *Narrative Methods for Organizational & Communication Research*. London: SAGE Publications Ltd.

Bramley, T. (2008). Mark scheme features associated with different levels of marker agreement. Presented at the British Educational Research Association annual conference, Heriot-Watt University, Edinburgh.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Brown, G., & Yule, G. (1983). *Discourse Analysis*. Cambridge: Cambridge University Press.

Brown, P., & Levinson, S. C. (1987). *Politeness: Some Universals in Language Usage*. Cambridge: Cambridge University Press.

Clark, H. H. (1992). *Arenas of Language Use*. Stanford, CA: University of Chicago Press.

Clark, H. H. (1996). *Using Language*. Cambridge: Cambridge University Press.

Clark, H. H., & Brennan, S. E. (1991). Grounding in Communication. In L. B. Resnick, J. Levine, & S. D. Teasley (Eds.), *Perspectives on Socially Shared Meaning* (pp. 127–149). Washington, DC: American Psychological Association.

Culpeper, J. (1996). Towards an anatomy of impoliteness. *Journal of Pragmatics*, *25*(3), 349–367.

Dennen, V. P. (2004). Cognitive apprenticeship in educational practice: Research on scaffolding, modeling, mentoring, and coaching as instructional strategies. In D. H. Jonassen (Ed.), *Handbook of Research on Educational Communications and Technology* (2nd ed., pp. 813–828). Mahwah, NJ: Lawrence Erlbaum Associates.

Edwards, D., & Mercer, N. (1987). *Common Knowledge: The Development of Understanding in the Classroom*. London: Methuen.

Fuller, A., & Unwin, L. (2003). Learning as Apprentices in the Contemporary UK Workplace: creating and managing expansive and restrictive participation. *Journal of Education and Work*, *16*(4), 407–426. https://doi.org/10.1080/1363908032000093012

Gee, J. P., & Green, J. L. (1998). Discourse Analysis, Learning, and Social Practice: A Methodological Study. *Review of Research in Education*, *23*, 119–169. https://doi.org/10.2307/1167289

Goffman, E. (1967). *Interaction Ritual: Essays in Face to Face Behavior*. Chicago: Aldine.

Jefferson, G. (2004). Glossary of transcript symbols with an introduction. In G. H. Lerner (Ed.), *Conversation Analysis: Studies from the First Generation* (pp. 13–31). Philadelphia: John Benjamins Publishing.

Johnson, M. (2015). Articulation work: Insights into examiners' expertise from their remote feedback interactions. *Communication & Language at Work*, *1*(4), 28–52. https://doi.org/http://dx.doi.org/10.7146/claw.v1i4.20771

Johnson, M. (2016a). All in good time: Influences on team leaders' feedback communication choices when giving feedback to examiners. *Research Matters: A Cambridge Assessment Publication*, *21*, 28–33.

Johnson, M. (2016b). "Teachers tell and children show and explain": the complementary dynamics of teacher and peer scaffolding techniques in a mixed ability group.

International Journal of Teaching and Education, *4*(4), 12–29. https://doi.org/10.20472/TE.2016.4.4.002

Johnson, M. (2017). Reading between the Lines: Exploring Methods for Analysing Professional Examiner Feedback Discourse. *International Journal of Research & Method in Education*, *40*(5), 456–470. https://doi.org/10.1080/1743727X.2016.1166484

Johnson, M., & Black, B. (2012). Feedback as scaffolding: Senior examiner monitoring processes and their effects on examiner marking. *Research in Post-Compulsory Education*, *17*(4), 391–407. https://doi.org/10.1080/13596748.2012.738965

Lakoff, G. (1973). Hedges: A study in meaning criteria and the logic of fuzzy concepts. *Journal of Philosophical Logic*, *2*(4), 458–508. https://doi.org/10.1007/BF00262952

Littleton, K., & Mercer, N. (2013). *Interthinking: putting talk to work*. Abingdon, Oxon: Routledge.

Llewellyn, N., & Hindmarsh, J. (Eds.). (2010). *Organisation, Interaction and Practice: Studies of Ethnomethodology and Conversation Analysis*. Cambridge: Cambridge University Press.

Massey, A. J., & Raikes, N. (2006). Item level examiner agreement. Presented at the British Educational Research Association annual conference, University of Warwick, UK.

Mercer, N. (2000). *Words and Minds: How We Use Language to Think Together and Get Things Done*. London: Routledge.

Mercer, N. (2004). Sociocultural discourse analysis: analysing classroom talk as a social mode of thinking. *Journal of Applied Linguistics*, *1*(2), 137–168. https://doi.org/10.1558/japl.v1i2.137

Mercer, N. (2008). The Seeds of Time: Why Classroom Dialogue Needs a Temporal Analysis. *Journal of the Learning Sciences*, *17*(1), 33–59. https://doi.org/10.1080/10508400701793182

Mercer, N., Littleton, K., & Wegerif, R. (2004). Methods for Studying the Processes of Interaction and Collaborative Activity in Computer-Based Educational Activities. *Technology, Pedagogy and Education*, *13*(2), 195–212. https://doi.org/10.1080/14759390400200180

Mercer, N., Wegerif, R., & Dawes, L. (1999). Children's Talk and the Development of Reasoning in the Classroom. *British Educational Research Journal*, *25*(1), 95–111. https://doi.org/10.1080/0141192990250107

Moore, M. G. (1993). Theory of transactional distance. In D. Keegan (Ed.), *Theoretical Principles of Distance Education* (pp. 22–38). London: Routledge.

Morand, D. A. (2000). Language and power: an empirical analysis of linguistic strategies used in superior–subordinate communication. *Journal of Organizational Behavior*, *21*(3), 235–248. https://doi.org/10.1002/(SICI)1099-1379(200005)21:3<235::AID-JOB9>3.0.CO;2-N

Münzer, S., & Holmer, T. (2009). Bridging the Gap Between Media Synchronicity and Task Performance: Effects of Media Characteristics on Process Variables and Task Performance Indicators in an Information Pooling Task. *Communication Research*, *36*(1), 76–103. https://doi.org/10.1177/0093650208326464

Murphy, E. A., & Rodriguez, M. A. (2008). Revisiting Transactional Distance Theory in a Context of Web-Based High-School Distance Education. *The Journal of Distance Education / Revue de l'Éducation à Distance*, *22*(2). Retrieved from http://www.jofde.ca/index.php/jde/article/view/38

Ofqual. (2013). *Review of Quality of Marking in Exams in A levels, GCSEs and Other Academic Qualifications: Interim Report* (No. Ofqual/13/5287). Coventry: Office of Qualifications and Examinations Regulation. Retrieved from http://dera.ioe.ac.uk/17772/1/2013-06-07-review-of-quality-of-marking-in-exams-in-a-levels-gcses-and-other-academic-qualifications-interim-report.pdf

Ofqual. (2016). *General Conditions of Recognition*. Coventry: Office of Qualifications and Examinations Regulation. Retrieved from https://www.gov.uk/government/publications/general-conditions-of-recognition

Ofqual. (2017). *GCE Qualification Level Conditions and Requirements*. Coventry: Office of Qualifications and Examinations Regulation. Retrieved from https://www.gov.uk/government/publications/gce-qualification-level-conditions-and-requirements

Rovai, A. P. (2000). Building and sustaining community in asynchronous learning networks. *The Internet and Higher Education*, *3*(4), 285–297. https://doi.org/10.1016/S1096-7516(01)00037-9

Schmidt, K. (1994). Cooperative Work and its Articulation: Requirements for Computer Support. *Le Travail Humain*, *57*(4), 345–366.

Schmidt, K. (2011). *Cooperative Work and Coordinative Practices: Contributions to the Conceptual Foundations of Computer-Supported Cooperative Work (CSCW)*. London: Springer-Verlag.

Scott, M. (1997). PC analysis of key words — And key key words. *System*, *25*(2), 233–245. https://doi.org/10.1016/S0346-251X(97)00011-0

Stannard, R. (2008). A New Direction in Feedback. *Humanising Language Teaching*, *10*(6). Retrieved from http://www.hltmag.co.uk/dec08/mart04.htm

Strauss, A. (1985). Work and the Division of Labor. *The Sociological Quarterly*, *26*(1), 1–19.

Sussman, S. W., & Sproull, L. (1999). Straight Talk: Delivering Bad News through Electronic Communication. *Information Systems Research*, *10*(2), 150–166. https://doi.org/10.1287/isre.10.2.150

Vonderwell, S. (2003). An examination of asynchronous communication experiences and perspectives of students in an online course: a case study. *The Internet and Higher Education*, *6*(1), 77–90. https://doi.org/10.1016/S1096-7516(02)00164-1

Vygotsky, L. S. (1978). *Mind in Society: The Development of Higher Psychological Processes*. Cambridge, Mass: Harvard University Press.

Whittaker, S. (2003). Theories and methods in mediated communication. In A. C. Graesser, M. A. Gernsbacher, & S. R. Goldman (Eds.), *Handbook of discourse processes* (pp. 243–286). Mahwah, NJ: Lawrence Erlbaum Associates Publishers.