



“Commenting on your work is a waste of time only!”: An appraisal-based study of evaluative language in supervisory feedback

M. Neupane Bastola^a, Guangwei Hu^{b,*}

^a The Hong Kong Polytechnic University, FG325, Department of English, The Hong Kong Polytechnic University, Hunghom, Kowloon, Hong Kong

^b The Hong Kong Polytechnic University, AG430, Department of English, The Hong Kong Polytechnic University, Hunghom, Kowloon, Hong Kong

ARTICLE INFO

Keywords:

Master's thesis
Supervisory feedback
Supervisors' attitude
Appraisal framework
Nepalese higher education

ABSTRACT

Research on supervisory feedback on master's theses, especially attitudinal stances conveyed in such feedback, is thin on the ground. Students' construal of their supervisors' attitudes, however, can have a profound impact on their engagement with supervisory feedback. Drawing on the appraisal framework, which characterizes attitudinal meanings in terms of affect (i.e., emotional responses), judgement (i.e., normative evaluation of human behaviors) and appreciation (i.e., aesthetically-/socially-based evaluation of objects and products), this study examined Nepalese supervisors' attitudinal stances communicated in written comments on master's thesis drafts ($n = 76$) submitted by English-as-a-foreign-language students and oral feedback on proposal and thesis defences ($n = 89$). Quantitative analyses revealed that while instances of appreciation dominated in the supervisors' use of evaluative language, judgements were also frequent, with affective responses trailing far behind. In both the oral feedback and written comments, significant disciplinary variations were observed for certain types of judgment and appreciation. These findings are discussed in terms of disciplinary culture and the potential impact of the attitudinal stances on students' learning. Implications are derived for the productive framing of supervisory feedback to facilitate students' feedback uptake.

1. Introduction

Writing a master's thesis is a demanding task (Paran, Hyland, & Bentall, 2017). Supervisory feedback is a paramount source of support for students to accomplish the task successfully and learn from the process (Basturkmen, East, & Bitchener, 2014; Bitchener, Basturkmen, & East, 2010). The extent to which supervisory feedback can achieve its intended purposes, however, depends on its quality (Hattie & Timperley, 2007) and students' active engagement with it (Neupane Bastola and Hu, 2020a). Learning-rich feedback makes students aware of expected standards (i.e., goals), informs them of how their performance fares (i.e., feedback), and provides guidance for improvement (feed-forward) (Hattie & Timperley, 2007). Feedback is useful, meaningful and sustainable if it can enhance students' autonomy, self-regulation, metacognition, and zeal for lifelong learning, bringing about a change in students' "thinking processes, emotions, relationships, work strategies, identity and more" (Henderson, Ajjawi, Boud, & Molloy, 2019) and making the feedback provider redundant in the long run (Carless, Salter, Yang, & Lam, 2011; Yang & Carless, 2013). Previous research has revealed that PhD students appreciated feedback that challenged them

to expand their intellectual independence, enhance their self-regulation capacity, and promote critical thinking (East, Bitchener, & Basturkmen, 2012), provided clear and detailed guidelines (Can & Walker, 2011), and had a suggestive rather than a directive tone (Can & Walker, 2011).

Extant research has also demonstrated that the quality of feedback is mediated by the nature of the feedback process per se (i.e., authoritative or dialogic) (Hyland & Hyland, 2019a), students' competence and confidence (Wang & Li, 2011), and the relationship between supervisor and student (Katikireddi & Reilly, 2017; Sutton, 2012). In this regard, the language used to communicate feedback has been recognized as affecting the quality and impact of feedback because students may simply dismiss the feedback received if the language is authoritative, demeaning, or destructive to mutuality (Sopina & McNeill, 2015; Starfield et al., 2015). Thus, a feedback provider intending to encourage and motivate students "could, unwittingly, employ language and tone that undermines [his/her] purpose" (Price, Handley, & Millar, 2011). However, little is known about the actual language features of supervisory feedback, especially the use of evaluative language to convey attitudinal meanings. Furthermore, thesis supervision involves providing feedback on thesis drafts as well as oral defences. These two

* Corresponding author.

E-mail addresses: madhu.neupanebastola@connect.polyu.hk (M. Neupane Bastola), guangwei.hu@polyu.edu.hk (G. Hu).

types of feedback situation create distinctly different interpersonal conditions, may require diverse evaluative resources, and can have varying impacts on students (Don & Izadi, 2011; Lin, 2017; Riekkinen, 2010). However, there is a paucity of research that explores possible differences in the use of evaluative language in supervisory feedback on thesis drafts and oral defences.

It is also important to note that the feedback process for a master's thesis occurs in a disciplinary culture (Hyland & Hyland, 2019a; Hyland, 2008; Winstone & Carless, 2020), which is constituted by "sets of taken-for-granted values, attitudes and ways of behaving" (Becher and Trowler, 2001). Disciplinary culture profoundly influences ontological, epistemological, and pedagogical beliefs and practices of academic communities (Becher & Trowler, 2001; Hu, 2018). Therefore, socializing students into legitimate ways of writing and meaning making in their discipline is one of the fundamental aims of graduate education (van Heerden, Clarence, & Bharuthram, 2017). A lack of such socialization on the part of graduate students can frustrate and alienate them (MacKay, Hughes, Marzetti, Lent, & Rhind, 2019), leading to disciplinary incompetence (Ylijoki, 2000). However, more often than not, research on supervisors' feedback does not "fully consider the concerns of the disciplines in terms of who they want their students to become, what they want their students to know or how they want their students to construct knowledge" (van Heerden et al., 2017). This study set out to bridge this gap by comparing the use of evaluative language to convey attitudinal stances in the oral and written feedback provided by supervisors from four disciplines (i.e., English Education, English Studies, Physics, and Engineering).

1.1. Disciplinary discourse and supervisory feedback

Authentic and effective feedback reflects disciplinary norms and values (Winstone & Carless, 2020). Based on their fundamental beliefs and practices, disciplines are broadly divided into four groups (Becher & Trowler, 2001). The first group comprises hard-pure disciplines (e.g., Physics) that are characterized by the pursuit of cumulative, atomistic, value-free, and universal knowledge and the adoption of rigorous criteria for knowledge creation and verification (Hu, 2018). By contrast, soft-pure disciplines (e.g., English Studies) are epitomized by value-laden and reiterative knowledge, holistic perspectives, and a general lack of universally agreed-upon standards for knowledge creation and verification (Hu and Choo, 2016). The third group, hard-applied disciplines (e.g., Engineering), put a premium on practical knowledge and application, whereas the last group, soft-applied disciplines (e.g., Education), distinguish themselves by focusing on the application of theoretical knowledge for the enhancement of professional practice (Becher & Trowler, 2001). In a widely cited discussion on disciplinary knowledge structures, Bernstein, 1999 posits that hard disciplines are dominated by a hierarchical knowledge structure (i.e., oriented to cumulative, objective, abstract, and general knowledge), whereas soft disciplines operate by a horizontal knowledge structure (i.e., oriented to knowers and specialized language for the creation and integration of knowledge). Developing the idea of disciplinary knowledge structures further, Maton, 2014 legitimation code theory proposes that "for every knowledge structure there is also a knower structure; that is, fields are *knowledge-knower structures*" (p.161; emphasis in original). According to legitimation code theory, disciplines with a hierarchical knowledge structure but a horizontal knower structure (e.g., Physics) operate with a knowledge code, embodied by "a more structured hierarchical body of knowledge that is verified against established scientific principles and procedures" (Hu, 2018). By contrast, disciplines with a horizontal knowledge structure and a hierarchical knower structure (e.g., Humanities) subscribe to a knower code, which legitimates disciplinary knowledge by "appealing to knowers' personal voice, expertise, experience, and authority" (Hu, 2018).

The disciplinary characteristics outlined above have been found to influence academic discourse in various ways. For example, corpus-

based studies (e.g., Cao and Hu, 2014; Hu and Cao, 2015; Hyland, 2005) revealed markedly greater frequencies in soft disciplines of linguistic features (e.g., hedges, boosters, and reader pronouns) constitutive of "a dialogic engagement and more explicit recognition of alternative voices" (Hyland, 2008) or reflecting writers' intention of "increasing commitment to their knowledge claims, asserting their authority, and positioning themselves as privileged knowers in their disciplinary communities" (Hu and Cao, 2015; Cao and Hu, 2014) also found disciplinary variations in the use of such linguistic features as exemplifiers (e.g., *for instance*), comparative transitions (e.g., *however, but*), integral citations (e.g., *as XXX argued*), and linear references (e.g., *as noted earlier*) that are interpretable in terms of Maton, 2014 knowledge-knower structures. Corpus-based studies of students' written discourse also showed that writing in soft disciplines (e.g., Philosophy, Education, Economics, and English Studies) was highly involved, narrative, and characterized by frequent use of stance verbs (e.g., *believe*), personal pronouns, hedges, boosters, active sentences, and past tenses (Hardy & Römer, 2013; Lancaster, 2016), while writing in physical sciences was descriptive, informational, and made more frequent use of nominalizations, adjectives, passive structures, and present tenses (Hardy & Römer, 2013).

Given these and other discursive and linguistic differences across disciplines, it is crucial to socialize students into disciplinary discourse if they are to become successful members of their community (Gunn, 2014). A recent study (Ashwin, Abbas, & McLean, 2017) at a British university showed that students' understanding of their discipline greatly influenced their learning from the research process because a change in their understanding occurred "only when they saw their discipline as providing a framework that allowed them to answer the questions they were investigating in their research" p. 527). In the context of supervisory feedback, the disciplinary differences discussed above would mean that supervisors in different disciplines are likely to look for and reward different qualities in their students' work. However, only a small body of research on supervisory feedback explored disciplinary variations and produced inconclusive findings. For example, two studies (Basturkmen et al., 2014; Bitchener et al., 2010) found little disciplinary variation in the foci of supervisory feedback. In contrast, more recent studies observed significant disciplinary effects on the development of feedback practices (Winstone & Boud, 2019) as well as the foci and pragmatic functions of supervisors' in-text feedback on master's theses (Neupane Bastola and Hu, 2020b). Notably, no study has examined disciplinary differences in the expression of attitudinal meanings in supervisory feedback. The limited extant research and the inconclusive findings point to the need for further inquiry into the what and the how of potential disciplinary influences on supervisory feedback in general and attitudinal stances in such feedback in particular.

1.2. Evaluative language in feedback

As recognized in the feedback literature (Hyland & Hyland, 2019a), the language used to communicate supervisory feedback can mediate its effectiveness. Hyland and Hyland, 2019b, for example, have noted that feedback providers' self-representation as "impersonal, critical and autocratic, or informed, sympathetic and helpful, and controlling" (p. 166) determines feedback interaction and students' subsequent action. As it is likely for the language of feedback to produce, in some cases, an unintended impact on students, extant research suggests that feedback providers reflect on the language they use and refrain from assuming that "the language they use is inherently meaningful to students" (Higgins, Hartley, & Skelton, 2002). Academics are also advised to make feedback language non-judgmental (Schartel, 2012) and less authoritative (Jonsson, 2013), assuming "a tone that students will read, understand and think about" (p. 237). Too much judgment, excessive criticism, and ill-conceived feedback may make students feel vulnerable and powerless (Boud, 1995), undermine the sense of cordiality (Basturkmen et al., 2014), and have a damaging effect on students (Hyatt,

2005). In contrast, the use of informal language, inclusive pronouns, and hedging may soften the threat to face and promote a common academic identity (Ajjawi and Boud, 2018).

Recommendations of this nature are largely consistent with previous research that examined students' perceptions of feedback language. This line of research found that students valued feedback that was positive, focused on strengths, had an encouraging tone, and showed a sense of caring (Hyland & Hyland, 2019b; Sopina & McNeill, 2015; Sutton, 2012). Empirical investigations into the pragmatic functions of feedback comments (e.g., Basturkmen et al., 2014; Bitchener et al., 2010; East et al., 2012; Hyland & Hyland, 2001, [Hyland and Hyland, 2019b]; Xu, 2017) also revealed that teachers aware of the impact of feedback language employed different strategies, for example, using praise to soften criticisms, enhance students' confidence, and maintain good interpersonal relations; utilizing hedging devices, personal attribution, and interrogative forms to tone down potentially damaging effects of feedback; and avoiding overly negative, uninformative comments. In contrast, teachers lacking such an awareness were found to use imperatives and obligation modality frequently, presenting themselves as an incontestable authority (Hyatt, 2005).

However, little research has focused squarely on the socio-affective features of evaluative language used in supervisory feedback. Two studies (Hu and Choo, 2016; Starfield et al., 2015) that employed the appraisal framework (Martin and White, 2005) to examine evaluative language used in examiners' reports are relevant to the present study. Starfield et al., 2015 explored the use of evaluative language in 142 PhD examination reports written by international, national, and internal examiners for 50 theses across four disciplines (i.e., Health and Health Sciences; Science; Business, Economics, and Accounting; and Humanities) at a New Zealand university. Although the university's assessment criteria (i.e., "official" evaluation) focused on the thesis per se, the examiners frequently slipped into "unofficial" evaluation by passing judgments on the candidates and expressing their affect. Hu and Choo, 2016 study of 84 project assessment reports revealed cross-disciplinary differences in schoolteachers' attitudes conveyed in their written evaluative language. While teachers from soft disciplines expressed satisfaction significantly more frequently, teachers from hard disciplines marked positive tenacity (i.e., dependability/resoluteness/perseverance) more often.

Although providing oral feedback on proposal and thesis defences is common practice in master's thesis supervision, "less research has focused on the defence/viva" (Mežek and Swales, 2016) partly because of difficulty in gaining access to the needed data (Lau, Lin, & Odle, 2020). A few extant studies described the structure of PhD defences (Mežek & Swales, 2016; van der Heide, Rufas, & Supper, 2016), supervisors' and students' anecdotal experiences (Djebali, 2020; Duke, 2020), and metadiscoursal features such as hedging (Don & Izadi, 2011; Riekkinen, 2010) and modality (Recski, 2005). Our extensive searches of the literature have located only three studies that examined master's thesis defences. Two of these studies (i.e., Lin, 2017, [Lin, 2020]) looked into pragmatic force modifiers, while the third one (Lau et al., 2020) examined the use of engagement resources. These studies demonstrated that participants used various interactional devices to make their interaction less face-threatening (Riekkinen, 2010), form dynamic relationships between interactants (Lin, 2017), and modulate their degree of commitment to the propositions made (Lau et al., 2020; Recski, 2005). However, no study has examined the use of evaluative language to convey attitudinal meanings in supervisory feedback.

As argued by Starfield et al., 2015, feedback should evaluate the thesis, not the student, so that they can bridge the gap between their current and expected standards. Therefore, it is essential to examine the evaluative language of oral and written supervisory feedback to "determine whether comments are appropriate, helpful, insightful or not" (Starfield et al., 2015). Notably, commenting on thesis drafts and providing feedback on oral defences constitute distinct interactional conditions for utilizing interpersonal language resources. For one thing,

oral defences tend to "foreground the attitudinally salient information and background ideational content" (Recski, 2005). For another, oral defences are also events where a candidate's personal qualities are assessed (van der Heide et al., 2016). Therefore, by their very nature, such events involve face-threatening acts (Don & Izadi, 2011; Riekkinen, 2010) and may necessitate the use of discourse strategies for maintaining and preserving face (Riekkinen, 2010). For these reasons, evaluative language used in oral feedback on defences warrants independent scrutiny from, and comparison with, that in written comments on theses. Building on Starfield et al., 2015, this study examined evaluative language in both oral and written supervisory feedback within the appraisal framework (Martin & White, 2005) to answer the following research questions:

- (1) What attitudes do supervisors express in their feedback on thesis drafts and oral defences?
- (2) Does the use of evaluative language in supervisory feedback vary across disciplines?

2. Method

2.1. Participants and data collection

The study was conducted at a comprehensive public university in Nepal and involved four disciplines (i.e., English Studies, English Education, Physics, and Engineering) that represented the soft-pure, soft-applied, hard-pure, and hard-applied classifications developed by Becher and Trowler, 2001. In the two-year master's degree programs offered by the focal university, thesis writing is compulsory in English Education and Engineering but optional in English Studies and Physics. In all these disciplines, students have to write their thesis in English, a foreign language in Nepal. At the university concerned, a single faculty member typically supervises a master's thesis. A student must defend his/her proposal and thesis before a research committee usually comprising the Head of the Department, an internal evaluator, an external evaluator, and the thesis supervisor. Unlike in some other contexts where "master's dissertations are marked entirely summatively and the student is not expected to return to them to improve them unless they fail" (Wisker, 2012), proposal and thesis defences are formative at the Nepalese university. Proposal defences aim to help students design more rigorous and robust research with feedback from both their supervisors and other academics. Thesis defences are also formative because they require minor/major revisions or even re-defences, though students are rarely awarded a fail grade. Supervisors are always involved in the final evaluation and commonly comment on their students' work during the defences. In disciplines with infrequent meetings between supervisors and students, such defences are a major source of supervisory feedback. In thesis defences, even external supervisors (as they are usually called) assume a supervisory role in providing suggestions on language and content issues (Wisker, 2012). As a professor in this study put it, a master's thesis defence is more like a teaching event than an examination.

This study drew on two data sets collected as part of a larger project: supervisory feedback on thesis drafts and oral defences. To collect the data needed, supervisors were selected from the four focal disciplines by means of maximum variation sampling based on the criteria of supervisory experience and research output (Patton, 2015). The same sampling method was then used to select students supervised by the sampled supervisors according to supervisor-assessed progress with their thesis writing (i.e., making smooth progress or struggling). Altogether, 76 thesis drafts with supervisory comments were included in the analysis. The second set of data comprised oral feedback on 89 audio-recorded proposal/thesis defences. A great majority of the supervisors involved in the oral defences were different from those involved in the written data set. Before the data collection was undertaken, an a priori power analysis was run using G*Power for the planned statistical tests (i.e.,

one-way between-subjects ANOVAs), with a power level of $1 - \beta = .80$, a pre-specified medium effect size of $f = .436$, and a significance level of $\alpha = .05$. The pre-specified effect size was based on Plonsky and Oswald, 2014 recommendation on a field-specific medium effect size (i.e., $r = .40$). The required sample size was 64, markedly smaller than our sample sizes for both the oral and the written dataset. While the sample sizes were adequate, it is also important to note that the thesis drafts and oral defences were unevenly distributed across the four disciplines, with one or two disciplines under-represented. Such uneven distributions could weaken the robustness of the statistical results obtained. Table 1 summarizes the distributions of the thesis drafts and oral defences across the disciplines. Ethical clearance for the study was obtained from the authors' university before data collection. The participants were duly informed of the purpose of the study, and informed consent was obtained from them.

2.2. Analytical framework

Evaluative language in the supervisors' oral feedback on proposal/thesis defences and written comments on thesis drafts was analysed according to the attitude system of the appraisal framework (Martin & White, 2005). The framework focuses on lexico-grammatical resources that can mark a speaker's/writer's intersubjective stances in the negotiation of social relationships, and consists of three main systems – attitude, graduation, and engagement (Martin & White, 2005). According to Martin and White, 2005, “[a]ttitude is concerned with our feelings, including emotional reactions, judgments of behavior and evaluation of things” (e.g., *I am satisfied with your work*; *You have not-written the objectives properly*; *This is a good piece of work*). Engagement involves the positioning of the speaker/writer with respect to potential responses to their opinions (e.g., *Your study seems to be a duplication of previous studies*), and graduation modulates the strength of evaluation (e.g., *We are not satisfied at all with your work*).

Because this study focuses on attitudinal stances conveyed in supervisory feedback, the rest of this section outlines the attitude system of appraisal theory. Attitudinal meanings, distinguished in terms of polarity (i.e., positive vs. negative), comprise *affect* (emotional reactions), *judgment* (assessment of human behaviour and characters according to normative principles), and *appreciation* (assessment of objects, artefacts, texts, state of affairs, and processes according to aesthetic principles and systems of social values) (Martin & White, 2005). Affect is further divided into four subcategories: *dis/inclination* (e.g., *I would like to see that a little elaborated, maybe in few pages.*), *un/happiness* (e.g., *I do not mind if you feel bad*), *in/security* (e.g., *I am worried about you all*), and *dis/satisfaction* (e.g., *You did not meet my expectations*). Unlike the individualised nature of affect, both judgment and appreciation are institutionalised feelings concerned with shared community values (Martin & White, 2005). Judgments are broadly categorized into social esteem and social sanction. Judgment about social esteem has to do with normality (e.g., *You are in your own way*), capacity (e.g., *You are not clear about what you are going to do*), and tenacity (e.g., *You worked hard*). Judgments of social sanction are concerned with veracity (e.g., *You have cheated us*) and propriety (e.g., *You have not-written the objectives properly*) (Martin & White, 2005). Appreciation deals with positive and negative assessments of inanimate objects in terms of reactions to their impact and quality (e.g., *Your work is interesting*; *This is a good piece of work*), compositional qualities such as balance and complexity (e.g., *There is a*

Table 1
Distributions of thesis drafts and oral defences by discipline.

Discipline	Oral defence	Draft
English Education	11	19
English Studies	20	10
Physics	13	22
Engineering	45	25

lack of connection between sentences in your writing; *There are so many misleading headings*), and valuation based on a set of institutionalised norms (e.g., *Commenting on your work is a waste of time only!*). Starfield et al., 2015 have added a new subcategory to valuation (i.e., standard) which concerns sufficiency, relevance, authenticity, and timeliness of content (e.g., *The analysis is superficial*). The other subcategories of valuation include normality, capacity, tenacity, veracity, and propriety construed with respect to objects and products. Fig. 1 presents the analytical framework adopted in this study.

2.3. Data coding and analysis

Data analysis involved several steps. First, discipline-based sub-corpora of supervisory feedback on oral defences and written comments on thesis drafts were created and imported to UAM CorpusTool (3.3 version), a free software package for annotating corpora at multiple levels (O'Donnell, 2011). Then, the sub-corpora were coded manually according to the analytical framework presented above. By its very nature, supervisory feedback makes use of evaluative language, and attitudinal meanings may vary across contextual and textual settings (Martin & White, 2005). Therefore, the coding of the feedback

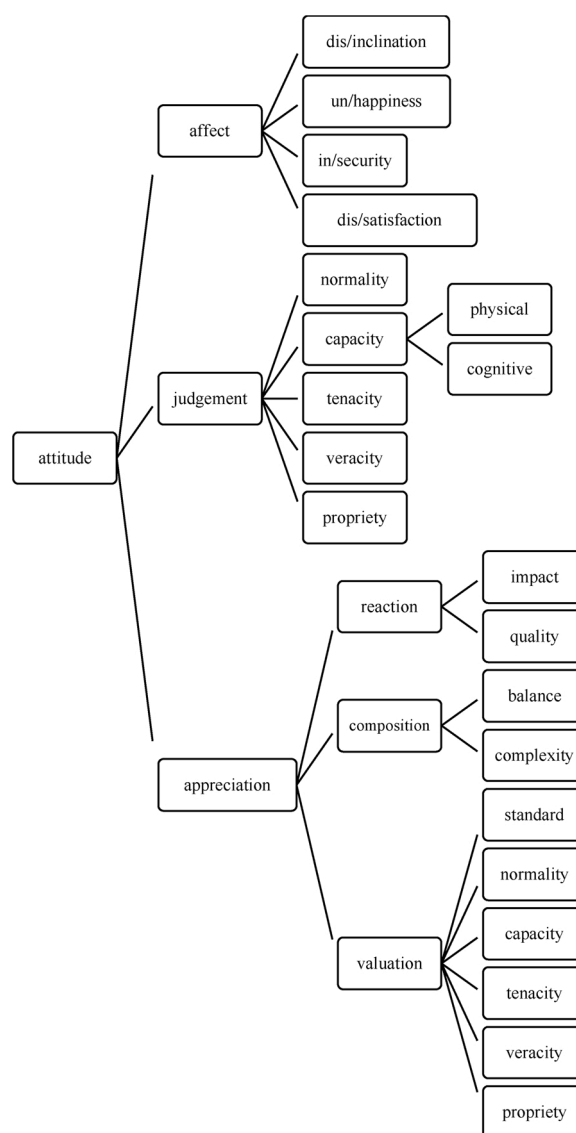


Fig. 1. Categories of attitude (based on Martin & White, 2005; Starfield et al., 2015).

comments involved paying close attention to “the context of use as well as the object of appraisal” (Hu and Choo, 2016). About 12% of the thesis data set (i.e., 16 drafts) were coded independently by the first author of this paper and a PhD student of Applied Linguistics to establish inter-coder reliability. The obtained Cohen’s kappa value of .87 indicated excellent inter-coder reliability. All inter-coder disagreements were resolved through discussion. The first author then coded all the remaining data.

The frequencies of comments falling into the different categories of the analytical framework were normalized by 1,000 words to control for the varying lengths of the supervisors’ commentary. Descriptive statistics for the normalized frequencies were computed to map the distributions of attitudinal meanings. To ascertain disciplinary variations in the incidence of various attitudinal stances, one-way between-subjects ANOVAs were run on the normalized frequencies. Following Hu and Choo, 2016, only those (sub)categories for which at least one disciplinary group had a mean score of 1.00 were subjected to the inferential statistical analyses. As a result, ANOVAs were run only for 19 of the 50 (sub)categories listed in Table 3. The alpha was set at .05 (2-tailed) for all the statistical tests.

3. Results

3.1. Attitudes expressed in the supervisory feedback

Table 2 presents the normalized frequencies of affect, judgment, and appreciation, along with their polarity and other relevant descriptive statistics. A comparison of the oral and written feedback revealed both similarities and differences. First, appreciation was most frequently expressed in both the oral feedback (44.78 per 1000) and the written comments (86.15 per 1000 words), followed in frequency by judgement. Second, the incidence of affect was very low, and instances were found only in the oral feedback. Third, attitudinal meanings were more frequent in the written feedback (151.96 per 1000 words) than in the oral feedback (79.53 per 1000 words). Fourth, instances of both judgement and appreciation occurred twice as frequently in the written comments as in the oral feedback. Finally, the attitudinal meanings were overwhelmingly negative in both the oral feedback (96%) and the written comments (99%). In the following subsections, examples of different attitudinal meanings communicated in the supervisory feedback are presented in the order of relative prominence to give a more detailed view.

3.1.1. Appreciation

The most common type of appreciation found in the supervisory feedback was social valuation. This type of attitudinal meaning was dominated by instances of valuation:standard, which were preponderantly negative.¹ These instances were mainly concerned with the adequacy of information (Example 1); data analysis, interpretation, and explanation (Example 2); and the quality of material presented (Example 3). Most of them were formulated as questions targeting missing information and under-developed arguments (Examples 4–6). Such questions conveyed negative evaluations because they indicated problems in students’ writing or presentations (Starfield et al., 2015). The most frequent question words were *what* (476 times), *how* (223 times), *why* (104 times), *which* (54 times), and *where* (41 times).

- (1) How [come] aspect is *missing* in your analysis. (-ve appreciation: valuation:standard)
- (2) Too much data *without interpretation*. (-ve appreciation:valuation: standard)

- (3) It seems to be *superficial*. (-ve appreciation:valuation:standard)
- (4) *What* is the new contribution of your study? (-ve appreciation: valuation:standard)
- (5) *Which analysis* will you use for optimization? (-ve appreciation: valuation:standard)
- (6) *Where* does the trauma theory come in your analysis? (-ve appreciation:valuation:standard):

The second most frequent type of social valuation was valuation:propriety, which mostly concerned issues of potential academic misconduct. These issues “include, but are not limited to, the inclusion and exclusion of relevant and current references, claims that are (un)justified, (un)substantiated and (un)referenced, the appropriate number of quotes and amount of paraphrasing as well as proper formatting of in-text references, reference lists and bibliographies” (Starfield et al., 2015). Common issues identified by the supervisors concerned in-text citations (Example 7) and plagiarism (Examples 8–9). Terms commonly used to express valuation:propriety were *reference* (200 times), *citation* (61 times), *source* (67 times), *plagiarized* (4 times), and *lifted* (3 times).

- (7) *Citation?* (-ve appreciation:valuation:propriety)
- (8) This section is *plagiarized*. (-ve appreciation:valuation:propriety)
- (9) Most of the things are *lifted*. (-ve appreciation:valuation:propriety):

There were also a few instances of valuation:capacity (Example 10), valuation:tenacity (Example 11), and valuation:veracity (Example 12):

- (10) If you take companies from different sectors, it is *difficult* to set the parameters. (-ve appreciation:valuation:capacity)
- (11) *Lack of reading* created a big problem. (-ve appreciation:valuation:tenacity)
- (12) Your analysis is totally *wrong*. (-ve appreciation:valuation:veracity):

Another type of appreciation, compositional evaluation, was also quite frequent in our data sets. Composition is concerned with balance (i.e., coherence and logical connection, structure of the thesis at all levels, relationships between text and tables/figures, and formatting issues like font, spacing, labelling, and capitalization) and complexity (i.e., clarity and comprehensibility of information presented). Examples 13–14 illustrate composition:balance, whereas Examples 15–16 illustrate composition:complexity. Notably, instances of composition:balance were far more frequent in the written comments than in the oral feedback. The five most frequently used expressions of appreciation:composition were *not clear* (25 times), *problem with language* (33 times), *space* (24 times), *bold* (21 times), *font* (20 times), and *italics* (20 times).

- (13) This goes into abstract *not here*. (-ve appreciation:composition:balance)
- (14) Your writing *needs thorough editing*. (-ve appreciation:composition:balance)
- (15) It is *difficult to know* whether the terms you have selected are cultural or not. (-ve appreciation:composition:complexity)
- (16) Your methodology is *not clear*. (-ve appreciation:composition:complexity):

The least common type of appreciation, reaction, expressed the supervisors’ personalized and subjectively determined positions regarding a thesis. The supervisors expressed reactions in terms of impact (Example 17) and quality (Example 18). Instances of reaction:impact and reaction:quality were notably more frequent in the oral feedback than in the written comments. The frequently used terms to express reactions were *fine* (40 times), *good* (21 times), *interesting* (20 times), *nice* (8 times), and *poor* (8 times). Unlike the other types of appreciation found in our data sets, positive comments were more frequent for both composition:impact and composition:quality, as can be seen in the fact that

¹ The sign “:” is used to indicate subsumption, with the entity following it being a subcategory of the entity preceding it. Negative and positive attitudes are notated as “-ve” and “+ve”, respectively.

four of the five most frequently used lexical items for expressing reactions were positive in polarity. A closer examination of the word *fine* indicated supervisors' agreement with students' ideas rather than positive appreciation of the text. However, *good* and *interesting* mostly expressed true appreciation of students' work, although, in a few cases, they were employed to preface critical comments (Examples 19–20).

- (17) I found your stories *interesting*. (+ve appreciation:reaction:impact)
- (18) The second paragraph is *fine*. (+ve appreciation:reaction:quality)
- (19) *Good* argument but no research credibility. (+ve appreciation:reaction:quality)
- (20) This is an *interesting* project; we are only worried about the direction it will take. (+ve appreciation:reaction:impact)

3.1.2. Judgment

Although judgement was found by Starfield et al., 2015 to fall outside of officially recommended examination criteria for doctoral theses, judgment:propriety was very prominent in our data sets, especially in the supervisors' written comments. This type of judgment communicated what students should and should not do with respect to different components of their thesis projects (e.g., research topic, abstract, introduction, literature, methodology, analysis and discussion, and conclusion), language, mechanics, and academic writing conventions (Examples 21–23). Expressions frequently used to convey judgments included *do not* (271 times), *need to* (231 times), *have to* (197 times), and *should* (138 times).

- (21) Now in qualitative data, you *do not* impose your themes in your writing. (-ve judgment:propriety)
- (22) You *need to* rewrite to make it readable. (-ve judgment:propriety)
- (23) You *should* not start directly like this. (-ve judgment:propriety):

The oral feedback also contained some instances of judgment:capacity:physical (i.e., ability to do something in physical terms) and judgment:capacity:cognitive (i.e., knowledge and clarity of understanding) (see Examples 24–26). The top five words used to

make such judgments were *can* (71 times), *know* (23 times), *clear* (16 times), *understand* (9), and *could* (7 times).

- (24) *Can* you study 100 buildings? (-ve judgment:capacity:physical)
- (25) It looks like you *do not understand* what you are trying to tell. (-ve judgment:capacity:cognitive)
- (26) Because of lack of reading you *could* not come up with the statement of the problem. (-ve judgment:capacity:cognitive):
Low in frequency, instances of judgment:tenacity (i.e., disposition to work) evaluated whether students were serious, dedicated, committed, meticulous, or patient (Examples 27–29). Expressions most frequently used to construe tenacity were *have not* (47 times), *work hard* (11 times), *careful* (10 times), *serious* (7 times), and *whatever* (4 times).
- (27) You have *not done* enough literature review. (-ve judgment:tenacity)
- (28) You *should be careful* in maintaining links between paragraphs. (-ve judgment:tenacity)
- (29) If you are *not serious*, who will think about you? (-ve judgment:tenacity):

Only a few occurrences of judgment:veracity (i.e., truthfulness and honesty) were found in the oral feedback (Examples 30–31) and mostly contained one of two words: *wrong* (7 times) and *trust* (5 times). Similarly, instances of judgment:normality were few and far between, appearing in the oral feedback only. They addressed students' level of performance (Example 32) or receptivity to feedback (Example 33).

- (30) You have provided *wrong* information. (-ve judgment:veracity)
- (31) We do not have much *trust* in survey questionnaire. (-ve judgment:veracity)
- (32) You were a *good* student. (+ve judgment:normality)
- (33) Learning starts only when you are *open*. (-ve judgment:normality)

3.1.3. Affect

Like judgment, affect is not expected to appear in “official” evaluation (Starfield et al., 2015). In this study, only a few instances of affect were observed in the supervisors' oral feedback, expressing dis/satisfaction (Example 34), in/security (Example 35), and dis/inclination

Table 2
Descriptive statistics for attitudinal meanings aggregated over disciplines (per 1000 words).

Attitude	Oral Feedback					Written feedback				
	+ve	-ve	+&-	M	SD	+ve	-ve	+&-	M	SD
Affect	0.27	0.50	0.77	0.39	1.03	0.00	0.00	0.00	0.00	0.00
Dis/satisfaction	0.16	0.23	0.39	0.26	0.90	0.00	0.00	0.00	0.00	0.00
In/security	0.00	0.16	0.16	0.06	0.33	0.00	0.00	0.00	0.00	0.00
Dis/inclination	0.11	0.11	0.23	0.06	0.28	0.00	0.00	0.00	0.00	0.00
Judgement	0.86	39.40	40.26	33.32	13.32	0.00	83.63	83.63	65.80	62.26
Tenacity	0.16	1.43	1.59	1.51	2.99	0.00	0.10	0.10	0.01	0.08
Propriety	0.32	33.34	33.66	27.05	12.73	0.00	83.44	83.44	65.79	62.26
Veracity	0.00	0.82	0.82	0.80	2.26	0.00	0.10	0.10	0.01	0.08
Normality	0.11	0.36	0.48	0.36	1.36	0.00	0.00	0.00	0.00	0.00
Capacity	3.45	3.45	3.72	3.60	4.20	0.00	0.00	0.00	0.00	0.00
Capacity:physical	0.23	1.73	1.95	2.09	3.14	0.00	0.00	0.00	0.00	0.00
Capacity:cognitive	0.05	1.73	1.77	1.51	2.49	0.00	0.00	0.00	0.00	0.00
Appreciation	2.50	35.99	38.49	44.78	23.90	0.49	115.37	115.86	86.15	104.31
Reaction	2.23	0.20	2.43	1.20	2.19	0.49	0.39	0.88	0.23	1.13
Impact	0.55	0.02	0.57	0.44	1.24	0.10	0.00	0.10	0.01	0.08
Quality	1.68	0.18	1.86	0.76	1.74	0.39	0.39	0.78	0.22	1.12
Composition	0.00	2.86	2.86	3.57	6.97	0.00	33.01	33.01	16.23	20.21
Balance	0.00	2.02	2.02	2.76	6.84	0.00	31.35	31.35	15.41	20.05
Complexity	0.00	0.84	0.84	0.81	1.53	0.00	1.66	1.66	0.82	2.42
Social valuation	0.27	32.93	33.20	40.00	22.18	0.00	81.98	81.98	69.69	100.94
Valuation:standard	0.20	26.43	26.64	33.28	20.76	0.00	51.21	51.21	38.50	73.73
Valuation:capacity	0.05	1.52	1.57	1.46	2.76	0.00	0.00	0.00	0.00	0.00
Valuation:tenacity	0.00	0.16	0.16	0.10	0.41	0.00	0.29	0.29	0.10	0.51
Valuation:veracity	0.00	1.20	1.20	1.33	3.12	0.00	2.34	2.34	1.50	4.68
Valuation:propriety	0.02	3.61	3.63	3.84	5.57	0.00	28.14	28.14	29.59	51.82
Total	3.63	75.90	79.53	78.48	23.97	0.49	199.01	199.49	151.96	120

Note. +ve = positive evaluation, -ve = negative evaluation, +&- = positive and negative evaluation combined, M = mean, SD = standard deviation.

Table 3
Descriptive statistics for normalized frequencies of attitudinal meanings in supervisory feedback per 1000 words.

Attitude	English Education				English Studies				Physics				Engineering			
	Oral		Written		Oral		Written		Oral		Written		Oral		Written	
	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
Affect+	0.49	0.79	0.00	0.00	0.00	0.00	0.00	0.00	0.19	0.68	0.00	0.00	0.00	0.00	0.00	0.00
Affect-	0.29	0.59	0.00	0.00	0.38	0.99	0.00	0.00	0.00	0.00	0.00	0.00	0.35	1.11	0.00	0.00
Satisfaction+	0.31	0.46	0.00	0.00	0.00	0.00	0.00	0.00	0.19	0.68	0.00	0.00	0.00	0.00	0.00	0.00
Dissatisfaction	0.02	0.08	0.00	0.00	0.24	0.94	0.00	0.00	0.00	0.00	0.00	0.00	0.27	1.02	0.00	0.00
Security+	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Insecurity	0.00	0.00	0.00	0.00	0.10	0.25	0.00	0.00	0.00	0.00	0.00	0.00	0.08	0.43	0.00	0.00
Inclination	0.18	0.41	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Disinclination-	0.26	0.58	0.00	0.00	0.03	0.15	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Judgement+	0.99	1.55	0.00	0.00	0.90	2.00	0.00	0.00	0.00	0.00	0.00	0.00	0.77	1.66	0.00	0.00
Judgement-	35.71	8.75	57.71	41.79	29.70	16.66	51.24	52.35	22.49	12.05	83.68	92.85	36.06	10.98	62.05	42.27
Tenacity+	0.30	0.58	0.00	0.00	0.18	0.66	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Tenacity-	0.31	0.47	0.04	0.15	1.93	3.55	0.00	0.00	0.00	0.00	0.00	0.00	1.89	3.33	0.00	0.00
Propriety+	0.39	1.05	0.00	0.00	0.30	0.82	0.00	0.00	0.00	0.00	0.00	0.00	0.17	0.68	0.00	0.00
Propriety-	32.08	9.33	57.64	41.76	24.06	16.10	51.24	52.35	18.38	12.22	83.68	92.85	29.26	10.59	62.05	42.27
Veracity+	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Veracity-	0.67	1.53	0.04	0.15	0.90	3.19	0.00	0.00	0.36	1.31	0.00	0.00	0.91	2.19	0.00	0.00
Normality+	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.27	0.86	0.00	0.00
Normality-	0.13	0.32	0.00	0.00	0.14	0.63	0.00	0.00	0.00	0.00	0.00	0.00	0.35	1.35	0.00	0.00
Capacity+	0.30	0.42	0.00	0.00	0.43	1.84	0.00	0.00	0.00	0.00	0.00	0.00	0.33	1.18	0.00	0.00
Capacity-	2.52	2.87	0.00	0.00	2.67	3.30	0.00	0.00	3.74	6.45	0.00	0.00	3.65	3.76	0.00	0.00
Capacity:physical+	0.30	0.42	0.00	0.00	0.41	1.85	0.00	0.00	0.00	0.00	0.00	0.00	0.30	1.17	0.00	0.00
Capacity:physical-	0.80	1.11	0.00	0.00	1.35	2.22	0.00	0.00	2.05	5.00	0.00	0.00	2.19	2.63	0.00	0.00
Capacity:cognitive+	0.00	0.00	0.00	0.00	0.02	0.08	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.15	0.00	0.00
Capacity:cognitive-	1.72	2.48	0.00	0.00	1.33	2.25	0.00	0.00	1.69	3.76	0.00	0.00	1.46	2.22	0.00	0.00
Appreciation+	3.88	3.11	0.14	0.47	1.00	2.86	0.28	0.89	0.82	2.34	0.00	0.00	0.87	1.78	0.00	0.00
Appreciation-	20.27	9.28	66.97	62.06	35.07	21.89	84.00	92.53	81.80	24.00	146.09	154.95	41.88	14.22	48.63	42.87
Reaction+	3.48	2.65	0.14	0.47	0.56	1.43	0.28	0.89	0.19	0.68	0.00	0.00	0.79	1.73	0.00	0.00
Reaction-	0.10	0.23	0.00	0.00	0.66	1.62	0.56	1.78	0.00	0.00	0.00	0.00	0.12	0.77	0.27	0.97
Impact+	0.92	1.55	0.04	0.15	0.07	0.30	0.00	0.00	0.00	0.00	0.00	0.00	0.61	1.52	0.00	0.00
Impact-	0.00	0.00	0.00	0.00	0.02	0.08	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Quality+	2.57	2.11	0.10	0.33	0.49	1.42	0.28	0.89	0.19	0.68	0.00	0.00	0.18	0.68	0.00	0.00
Quality-	0.10	0.23	0.00	0.00	0.64	1.62	0.56	1.78	0.00	0.00	0.00	0.00	0.12	0.77	0.27	0.97
Composition+	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Composition-	3.22	2.85	17.08	18.22	2.47	4.10	18.38	23.99	11.12	15.06	22.90	25.35	1.96	2.50	8.85	12.25
Balance+	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Balance-	2.46	2.00	16.20	18.00	1.54	2.61	17.54	23.53	11.12	15.06	21.68	25.08	0.96	1.80	7.76	12.29
Complexity+	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Complexity-	0.77	1.18	0.79	1.48	0.94	2.12	0.00	0.00	0.00	0.00	1.21	3.44	1.00	1.48	0.83	2.42
Social valuation+	0.40	0.77	0.00	0.00	0.44	1.98	0.00	0.00	0.64	2.29	0.00	0.00	0.08	0.55	0.00	0.00
Social valuation-	16.95	8.93	49.89	57.94	31.94	22.09	65.06	79.66	70.69	19.11	123.19	156.10	39.80	14.58	39.51	41.56
Valuation:standard+	0.32	0.76	0.00	0.00	0.30	1.32	0.00	0.00	0.64	2.29	0.00	0.00	0.00	0.00	0.00	0.00
Valuation:standard-	11.74	5.38	36.77	41.51	28.77	21.31	31.12	34.48	61.64	16.30	52.89	124.96	31.96	15.04	30.10	36.02
Valuation:capacity+	0.09	0.28	0.00	0.00	0.15	0.66	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Valuation:capacity-	0.26	0.59	0.00	0.00	0.83	1.91	0.00	0.00	0.00	0.00	0.00	0.00	2.37	3.40	0.00	0.00
Valuation:tenacity+	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Valuation:tenacity-	0.00	0.00	0.26	0.80	0.02	0.08	0.28	0.89	0.00	0.00	0.00	0.00	0.18	0.57	0.00	0.00
Valuation:veracity+	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Valuation:veracity-	1.41	1.42	0.61	1.10	0.51	1.22	0.62	1.97	3.63	6.63	2.67	7.32	1.02	2.09	1.50	4.14
Valuation:propriety+	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.08	0.55	0.00	0.00
Valuation:propriety-	3.54	3.70	12.24	21.33	1.82	3.54	33.03	53.93	5.42	11.10	67.64	73.75	4.27	4.19	7.91	11.87

Note. “+” and “-” indicate positive and negative polarity, respectively.

(Example 36). A close examination of the examples in context showed that the supervisors were dissatisfied when they perceived that their students lacked seriousness and dedication to the research that they were undertaking. They also had doubts regarding whether the students would be able to accomplish their thesis projects.

- (34) Today we are *not satisfied* with your work at all. (-ve affect:dis/satisfaction)
- (35) I am *not sure* if you can do that. (-ve affect:in/security)
- (36) I am *worried* about you. (-ve affect:in/security)

3.2. Disciplinary similarities and differences in the use of evaluative language

To determine if there were disciplinary variations in the use of evaluative language, one-way ANOVAs were run on the normalized

frequencies of different types of attitudinal meaning in the oral and written feedback separately. Table 3 presents descriptive statistics for the normalized frequencies by discipline, data set, attitudinal category, and polarity.

3.3.1. Oral feedback

Judgment. A significant main effect of discipline was found for the supervisors’ expression of negative propriety in the oral feedback, $F(3, 85) = 3.73, p = .01, \eta^2 = .12$. The obtained effect size indicated that disciplinary background accounted for 12% of the variance in the incidence of negative propriety. A Bonferroni post hoc test located significant differences with large effect sizes between the English Education and Physics supervisors ($p = .04, d = 1.30$) and between the Engineering and Physics supervisors ($p = .03, d = 0.95$). No significant effect of discipline was observed for negative tenacity, $F(3, 85) = 2.14, p = .10, \eta^2 = .07$; negative veracity, $F(3, 85) = 0.22, p = .89, \eta^2 = .01$; or negative

capacity, $F(3, 85) = 0.45, p = .72, \eta^2 = .02$.

Appreciation. A significant disciplinary effect was found on the frequencies of positive reaction:quality, $F(3, 85) = 13.72, p = .001, \eta^2 = .33$. Disciplinary background explained a sizable 33% of the variance. A Bonferroni post hoc test found that the English Education supervisors expressed positive reactions to quality more frequently than the English Studies supervisors ($p = .001, d = 1.16$), the Physics supervisors ($p = .001, d = 1.52$), and the Engineering supervisors ($p = .001, d = 1.52$), with large effect sizes in all cases.

Discipline was also found to have a significant main effect on the frequencies of negative composition:balance, $F(3, 85) = 10.13, p = .001, \eta^2 = 0.26$. The effect size was big, and the post hoc test revealed that the Physics supervisors expressed such attitudinal meanings more frequently than did their counterparts from English Education ($p = .004, d = 0.81$), English Studies ($p = .001, d = 0.89$), and Engineering ($p = .001, d = 0.95$). By contrast, no significant main effect of discipline was found for composition:complexity, $F(3, 85) = 1.52, p = .22, \eta^2 = 0.05$.

A significant main effect of discipline was detected for the frequencies of negative valuation:standard, $F(3, 85) = 20.60, p = .001, \eta^2 = 0.40$. Disciplinary background had a huge effect as it explained 40% of the variance. The Physics supervisors differed greatly from their counterparts in English Education ($p = .001, d = 4.11$), English Studies ($p = .001, d = 1.73$), and Engineering ($p = .001, d = 1.89$). Significant differences were also observed between the English Studies and English Education supervisors ($p = .04, d = 1.10$) and between the Engineering and English Education supervisors ($p = .002, d = 1.79$), with large effect sizes. A significant main effect of discipline was also found in the incidence of negative valuation:capacity, $F(3, 85) = 4.33, p = .01, \eta^2 = 0.13$, with the Engineering supervisors expressing negative capacity more frequently than the Physics supervisors did ($p = .03, d = 0.98$). A disciplinary effect was also detected for negative valuation:veracity, $F(3, 85) = 3.20, p = .03, \eta^2 = 0.10$, with the Physics supervisors expressing such evaluations more frequently than both the English Studies ($p = .03, d = 0.65$) and the Engineering supervisors ($p = .04, d = 0.53$). No significant main effect of discipline was found for negative valuation:propriety, $F(3, 72) = 1.33, p = .27, \eta^2 = .04$.

3.3.2. Written feedback

Judgement. There was no significant main effect of discipline on the incidence of negative propriety, $F(3, 72) = 0.93, p = .43, \eta^2 = .04$.

Appreciation. Disciplinary background had a significant main effect on the incidence of negative valuation:propriety, $F(3, 72) = 7.81, p = .001, \eta^2 = 0.25$, with the Physics supervisors expressing such evaluations more frequently than did the English Education ($p = .002, d = 1.02$) and Engineering supervisors ($p = .001, d = 1.13$). No significant main effect of discipline was observed on the incidence of negative composition:balance, $F(3, 72) = 2.07, p = .11, \eta^2 = .08$; negative composition:complexity, $F(3, 72) = 0.57, p = .64, \eta^2 = .02$; negative valuation:standard, $F(3, 72) = 0.41, p = .74, \eta^2 = .02$; or negative valuation:veracity, $F(3, 72) = 0.80, p = .50, \eta^2 = .03$.

4. Discussion

The results presented in the previous section showed that the supervisors engaged in both “official” evaluation (i.e., appreciation) and “unofficial” evaluation (i.e., judgement and affect) in their supervisory feedback. There were significant differences in the incidence of various categories and subcategories of attitudes across the disciplines and between the modes of feedback. These findings are discussed below in relation to the research questions that guided this study.

4.1. What attitudes do supervisors express in their feedback?

As reported above, of the three main categories of attitude (i.e., affect, judgement, and appreciation) in Martin and White, 2005 appraisal framework, appreciation predominated in both the oral and

written supervisory feedback. This pattern was compliant with the recommendation that the major concern of thesis evaluation is “the quality of the research” (Holbrook, Bourke, Fairbairn, & Lovat, 2014). Appreciation of a thesis in terms of its value, as well as coherence and clarity in presentation, constitutes “official” evaluation, whereas the expression of judgements and affective responses is considered “unofficial” evaluation (Starfield et al., 2015). However, it is important to note that judgments were also highly prominent in our data sets and that affect was occasionally expressed in the oral feedback. These patterns of attitudinal meanings appeared to reflect a general lack of understanding among the supervisors that their feedback should not be directed at the personal attributes of their students or ventilate their own affective responses but should focus on the quality of student work. They also seemed to reveal a lack of awareness that the types of judgment found in our data contained little guidance that would help students understand the standard expected of their thesis (i.e., goals) and scaffold their efforts to improve its quality (i.e., feed-forward) (Hattie & Timperley, 2007). For the same reason, the expression of negative affect (e.g., dissatisfaction and worries) in some supervisors’ feedback on their students’ oral defences did little to scaffold the latter’s further efforts to enhance the quality of their work and might have alienated them emotionally.

The oral and written supervisory feedback differed in the expression of certain attitudinal stances. Specifically, instances of affect and judgment related to students’ tenacity, capacity, veracity, and normality appeared only in the oral feedback. One plausible explanation of these observed differences lay in the immediate impact of the face-to-face communication that occurred in the oral defences. Because of such face-to-face interactions, the presentations might have directed the supervisors’ evaluative attention to the students (i.e., the presenters) and themselves (i.e., their affective responses) more easily than the thesis drafts did. However, the supervisors’ public expression of their negative affect and harsh judgements on their students’ personal attributes could make the latter feel vulnerable and powerless (Boud, 1995). Furthermore, as pointed out earlier, such feedback contained little information to enhance students’ “engagement, commitment to the learning goals, ... self-efficacy, or understanding about the task” (Hattie and Timperley, 2007).

As reported above, the attitudinal meanings expressed in both the oral and written feedback were predominantly negative. This finding was consistent with what was reported in previous studies (Basturkmen et al., 2014; Hyatt, 2005; Starfield et al., 2015). Three factors might have contributed to the predominance of negative attitudinal meanings. First, in its general sense, feedback entails identifying areas that need improvement “to reduce the gap between a current and the desired performance” (Aben, Dingyloudi, Timmermans, & Strijbos, 2019). Thus, it tends to focus on what students have not done well and where further work is needed. Consequently, there is more opportunity to communicate negative appraisal. Second, as our closer examination of the thesis drafts and oral presentations revealed, poor-quality theses and defences tended to elicit more negative comments, a pattern that was also found by previous research (Paltridge & Starfield, 2019; Skinner & Pitzer, 2012). Such unsatisfactory performance could, in turn, be attributed to suboptimal supervision processes characterized by low supervisor-student collaboration, poor supervisor-student relationship, and minimal informative feedback. Third, the overly critical comments targeting students’ social esteem might also indicate the supervisors’ limited feedback literacy, that is, “expertise in knowing how to enhance feedback processes” (Winstone and Carless, 2020).

As our earlier review of previous research has indicated, how feedback is framed can greatly influence student agency and engagement. The supervisory feedback analysed in this study, in many cases, did not seem to be “characterized by an ethos of care” (Sutton, 2012). Rather than motivating, encouraging, and inspiring students to invest efforts in learning, more often than not, it tended to demoralize them. The extremely direct and overly negative judgments threatened students’ face in front of their professors and fellow students (Hyatt, 2005) and

appeared to result in their plummeting self-confidence, demotivation, and negativity toward thesis writing itself (Henderson et al., 2019; Authors, 2020). As a English Studies student shared with us,

Thesis writing has been one of the bitter experiences.... [Once] my supervisor screamed at me, and his angry words still sting me. Those outspoken words can be the best for discouraging students.

Such reactions, common among the student participants in this study, lent support to the observation that “learners’ attitudes and emotions are mediated by encouragement and support from others” (Han and Hyland, 2019). The predominance of negative face-threatening comments was also likely to undermine the development of a cordial and collaborative supervisor-student relationship and ultimately reduce the chances of students’ cognitive and behavioural engagement with supervisory feedback (Wagener, 2018). Such comments could be attributed in part to a lack of awareness that “[t]he ways we convey our praise or criticism, and how we phrase our suggestions, are central to effective feedback” (Hyland and Hyland, 2019b).

4.2. Does the use of evaluative language in supervisory feedback vary across the disciplines?

Our analyses revealed significant disciplinary variations in the incidence of certain types of attitudinal meaning. In the oral feedback, negative judgement:propriety and positive reaction:quality occurred more frequently in English Education; negative composition:balance, negative valuation:standard, and negative valuation:veracity were found more frequently in Physics; and negative valuation:capacity was more common in Engineering. In the written feedback, the Physics supervisors also expressed negative valuation:propriety significantly more frequently than their counterparts in the other disciplines did. These differences can be explained in terms of the knowledge-knower structures underlying the disciplines (Becher & Trowler, 2001; Bernstein, 1999; Maton, 2014) and contextual factors.

Education as a soft applied discipline is concerned with utilitarian knowledge for the “enhancement of [semi-] professional practice” (Becher and Trowler, 2001), has a horizontal knowledge structure but a hierarchical knower structure (Bernstein, 1999), and places an emphasis on personal understanding and the subjectivity of knowledge (Becher & Trowler, 2001; Maton, 2014). These disciplinary characteristics could explain in part the English Education supervisors’ more frequent use of negative judgement:propriety. It would seem that they used such judgments not only to legitimate their own disciplinary knowledge by “appealing to knowers’ personal voice, expertise, experience, and authority” (Hu, 2018) but also to tell their students what they, as would-be legitimate knowers, should and should not do (i.e., person-oriented feedback). In a similar vein, the English Education supervisors’ more frequent expression of positive reaction:quality could be attributed partly to the strong tendency of soft disciplines to be “concerned with particulars, qualities, and complication” (Becher and Trowler, 2001) and to valorize personal interpretations and individual responses (Hu and Choo, 2016), in contrast to the premium placed by hard disciplines on value-free, objective knowledge (Becher & Trowler, 2001).

As a typical hard-pure discipline, Physics has a hierarchical knowledge and a horizontal knower structure (Maton, 2014), is concerned with universals, seeks impersonal knowledge, and has rigorous, commonly agreed-upon criteria for generating and verifying knowledge (Becher & Trowler, 2001). These disciplinary characteristics would have inclined the Physics supervisors to attend more to objectivity, criteria, accuracy and order in their supervisory feedback (Hu and Wang, 2014). It would seem that such attention was manifested in their use of attitudinal language to evaluate students’ performance in relation to standards (i.e., valuation:standard), content accuracy (i.e., valuation:veracity) and what Martin and White, 2005 refer to as “view of order” (i.e., composition:balance) in their oral feedback and compliance with

established conventions (i.e., valuation:propriety) in their written feedback. These observed patterns were consistent with Hu and Choo, 2016 observation that “it is natural for teachers from hard disciplines to be impersonal in their feedback and focus on content of work, hard knowledge involved, and criteria of performance” (p.343).

As a hard-applied discipline, Engineering is concerned with the mastery of the physical environment and the creation of new or more advantageous products or techniques (Becher & Trowler, 2001). Given the utilitarian nature of the discipline, Engineering places a premium on efficiency and clarity in design (Vikers, 2007). Such disciplinary priorities could be reasonably expected to find expression in the Engineering supervisors’ more frequent use of valuation:capacity to comment on the feasibility of a study and the utility of a product. Given the observed associations between disciplinary characteristics and the supervisors’ use of particular types of evaluative language, we argue that disciplinary norms and values constitute a major influence on supervisors’ expression of attitudinal stances in their feedback on student work (Becher & Trowler, 2001; Ylijoki, 2000).

5. Conclusion: implications and limitations

Since the primary purpose of feedback is to develop students’ learning capacity so that, in the long run, they become autonomous, independent, and self-regulated (Carless et al., 2011; Henderson et al., 2019; Yang & Carless, 2013), supervisory feedback practice should center around students’ academic growth. In this light, several implications can be derived from the findings of this study. First, although covert expression of emotions (i.e., in the form of appreciation:reaction:impact) is acceptable (Starfield et al., 2015), it might be beneficial to rid supervisory feedback of harsh expressions of negative affect so as to foster a trusting supervisor-student relationship and facilitate students’ engagement with such feedback. Second, supervisors might consider avoiding direct negative judgements on students’ social esteem (e.g., normality, tenacity, and capacity) because such comments rarely contain the guidance (i.e., feedforward) needed to enhance the latter’s learning (Hattie & Timperley, 2007) but are likely to alienate them by threatening their face and demoralizing them. Third, supervisory feedback should be “characterized by an ethos of care” (Sutton, 2012) and be peppered with instances of positive appreciation (e.g., various types of positive reaction and social valuation) of students’ efforts to motivate, encourage and inspire them (de Kleijn, Meijer, Pilot, & Brekelmans, 2014; Wei, Carter, & Laurs, 2019). This is particularly important in oral defences because students may feel more psychologically vulnerable in such face-to-face, high-stakes examinations involving individuals other than their supervisors.

Fourth, although supervisors’ expression of attitudinal stances in their feedback is shaped by disciplinary norms and conventions and, consequently, constitutes implicit disciplinary socialization, such enculturation could be more effective if they explicitly communicate to their students ways of thinking, acting, being, and feeling that are valued by their discipline. For example, they could annotate the attitudinal meanings expressed in their feedback with explicit explanations of why certain ways of reasoning, arguing, and knowing are valued or depreciate in disciplinary practices. Finally, faculty development programs should be offered to help supervisors develop their feedback literacy so that they can “place student progress and improvement at the core of their practice” (Winstone and Carless, 2020). An integral part of feedback literacy is an awareness of both the affordances and potentially detrimental impacts that the expression of attitudinal meanings in supervisory feedback can have. The importance of being aware of the power of evaluative language cannot be overstressed because, as suggested by previous research (Hyland & Hyland, 2019a; Winstone & Carless, 2020), supervisors with a good understanding of attitudinal stances and their effects are able to use evaluative language effectively to facilitate students’ understanding and uptake of supervisory feedback, promote their sense of well-being, develop a productive

supervisor-student relationship characterized by a high degree of mutuality, and inspire independent learning.

The above-discussed implications notwithstanding, it is imperative to recognize several limitations of this study so that future research can address them. First, the robustness of our findings might have been affected by the relatively small number of participants representing English Education and Physics (oral defences) and English Studies (thesis drafts). Although we tried to alleviate this limitation by conducting an a priori power analysis and ensuring that our sample sizes were adequate for the statistical analyses, the uneven distributions of the participants across the disciplines could weaken the generalizability of our findings to the disciplines involved. Researchers interested in comparing discipline-specific supervisory practices might wish to ensure a more balanced representation of the disciplines in their future studies. Second, our data came from a single university and, given the situated nature of supervisory feedback, our findings may not be generalizable to other educational contexts. Therefore, similar studies need to be conducted in other settings to verify our findings. Third, logistical constraints prevented us from collecting all the participating students' emotional responses after they received the written or oral supervisory feedback to gauge its emotional impact, though a small subset of students was interviewed to that end. Future studies along a similar line are well advised to collect such data more systematically to determine, in a more direct manner, the relationship between particular attitudinal meanings conveyed in supervisory feedback and students' emotional temperature and to achieve a more comprehensive understanding of the use of evaluative language in supervisory feedback.

Acknowledgment

We would like to thank the Research Center for Professional Communication in English (RCPCE) at the Department of English of The Hong Kong Polytechnic University for its support.

References

- Ajjawi, R., & Boud, D. (2018). Examining the nature and effects of feedback dialogue. *Assessment & Evaluation in Higher Education*, 43, 1106–1119.
- Aben, J. E. J., Dingyloudi, F., Timmermans, A. C., & Srijbos, J.-W. (2019). Embracing errors for learning: Intrapersonal and interpersonal factors in feedback provision and processing in dyadic interactions. In M. Henderson, R. Ajjawi, D. Boud, & E. Molloy (Eds.), *The impact of feedback in higher education* (pp. 107–125). Cham: Springer.
- Ashwin, P., Abbas, A., & McLean, M. (2017). How does completing a dissertation transform undergraduate students' understandings of disciplinary knowledge? *Assessment and Evaluation in Higher Education*, 42, 517–530.
- Cao, F., & Hu, G. (2014). Interactive metadiscourse in research articles: A comparative study of paradigmatic and disciplinary influences. *Journal of Pragmatics*, 66, 15–31.
- Hu, G. (2018). Disciplinary knowledge making and academic discourse. In Y. Leung, J. Katchen, S. Hwang, & Y. Chen, *Reconceptualizing English language teaching and learning in the 21st century* 553-573. Taipei. Crane Publishing.
- Hu, G., & Cao, F. (2015). Disciplinary and paradigmatic influences on interactional metadiscourse in research articles. *English for Specific Purposes*, 39, 12–25.
- Hu, G., & Choo, L. (2016). The impact of disciplinary background and teaching experience on the use of evaluative language in teacher feedback. *Teachers and Teaching*, 22, 329–349.
- Hu, G., & Wang, G. (2014). Disciplinary and ethnolinguistic influences on citation in research articles. *Journal of English for Academic Purposes*, 14, 14–28.
- Neupane Bastola, M. (2020a). in press b) Formulation of feedback comments: Insights from supervisory feedback on master's theses. *Innovations in Education and Teaching International*, s1. in press.
- Neupane Bastola, M., & Hu, G. (2020b). in press). Supervisory feedback across disciplines: Does it meet students' expectations? *Assessment & Evaluation in Higher Education*. in press).
- Basturkmen, H., East, M., & Bitchener, J. (2014). Supervisors' on-script feedback comments on drafts of dissertations: Socialising students into the academic discourse community. *Teaching in Higher Education*, 19, 432–445.
- Becher, T., & Trowler, P. (2001). *Academic tribes and territories: Intellectual enquiry and the culture of disciplines*. Buckingham: Society for Research into Higher Education & Open University Press.
- Bernstein, B. B. (1999). *Pedagogy, symbolic control, and identity: Theory, research, critique*. London: Rowman & Littlefield.
- Bitchener, J., Basturkmen, H., & East, M. (2010). The focus of supervisor written feedback to thesis/dissertation students. *International Journal of English Studies*, 10 (2), 79–97.
- Boud, D. (1995). *Enhancing learning through self assessment*. London: Kogan Page.
- Can, G., & Walker, A. (2011). A model for doctoral students' perceptions and attitudes toward written feedback for academic writing. *Research in Higher Education*, 52, 508–536.
- Carless, D., Salter, D., Yang, M., & Lam, J. (2011). Developing sustainable feedback practices. *Studies in Higher Education*, 36, 395–407.
- de Kleijn, R. A. M., Meijer, P. C., Pilot, A., & Brekelmans, M. (2014). The relation between feedback perceptions and the supervisor–student relationship in master's thesis projects. *Teaching in Higher Education*, 19, 336–349.
- Djebali, Z. (2020). My viva voce examination: A rollercoaster. In K. Townsend, M. N. K. Saunders, R. Loudoun, & E. A. Morrison (Eds.), *How to keep your doctorate on track: Insights from students' and supervisors' experiences* (pp. 356–364). Cheltenham, UK: Edward Elgar Publishing.
- Don, Z. M., & Izadi, A. (2011). Relational connection and separation in Iranian dissertation defences. *Journal of Pragmatics*, 43, 3782–3792.
- Duke, D. C. (2020). The final mile: Avoiding and overcoming viva trip-ups. In K. Townsend, M. N. K. Saunders, R. Loudoun, & E. A. Morrison (Eds.), *How to keep your doctorate on track: Insights from students' and supervisors' experiences* (pp. 345–355). Cheltenham, UK: Edward Elgar Publishing.
- East, M., Bitchener, J., & Basturkmen, H. (2012). What constitutes effective feedback to postgraduate research students? The students' perspective. *Journal of University Teaching & Learning Practice*, 9(2), 1–16.
- Gunn, V. (2014). Mimetic desire and intersubjectivity in disciplinary cultures: Constraints or enablers to learning in higher education? *Studies in Continuing Education*, 36, 67–82.
- Han, Y., & Hyland, F. (2019). Learner engagement with written feedback: A sociocognitive perspective. In K. Hyland, & F. Hyland (Eds.), *Feedback in second language writing* (2nd ed., pp. 247–264). Cambridge: Cambridge University Press.
- Hardy, J. A., & Römer, U. (2013). Revealing disciplinary variation in student writing: A multi-dimensional analysis of the Michigan Corpus of Upper-level Student Papers (MICUSP). *Corpora*, 8, 183–207.
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77, 81–112.
- Henderson, M., Ajjawi, R., Boud, D., & Molloy, E. (2019). Identifying feedback that has impact. In M. Henderson, R. Ajjawi, D. Boud, & E. Molloy (Eds.), *The impact of feedback in higher education* (pp. 15–34). Cham: Palgrave Macmillan.
- Higgins, R., Hartley, P., & Skelton, A. (2002). The conscientious consumer: Reconsidering the role of assessment feedback in student learning. *Studies in Higher Education*, 27, 53–64.
- Holbrook, A., Bourke, S., Fairbairn, H., & Lovat, T. (2014). The focus and substance of formative comment provided by PhD examiners. *Studies in Higher Education*, 39, 983–1000.
- Hyatt, D. F. (2005). “Yes, a very good point!”: A critical genre analysis of a corpus of feedback commentaries on Master of Education assignments. *Teaching in Higher Education*, 10, 339–353.
- Hyland, K. (2005). Stance and engagement: A model of interaction in academic discourse. *Discourse Studies*, 7(2), 173–192.
- Hyland, K. (2008). Disciplinary voices: Interactions in research writing. *English Text Construction*, 1, 5–22.
- Hyland, F., & Hyland, K. (2001). Sugaring the pill: Praise and criticism in written feedback. *Journal of Second Language Writing*, 10, 185–212.
- Hyland, K., & Hyland, F. (2019a). Contexts and issues in feedback on L2 writing. In K. Hyland, & F. Hyland (Eds.), *Feedback in second language writing* (2nd ed., pp. 1–22). Cambridge: Cambridge University Press.
- Hyland, K., & Hyland, F. (2019b). Interpersonality and teacher-written feedback. In K. Hyland, & F. Hyland (Eds.), *Feedback in second language writing* (2nd ed., pp. 165–183). Cambridge: Cambridge University Press.
- Jonsson, A. (2013). Facilitating productive use of feedback in higher education. *Active Learning in Higher Education*, 14, 63–76.
- Katikireddi, S. V., & Reilly, J. (2017). Characteristics of good supervision: A multi-perspective qualitative exploration of the Masters in Public Health dissertation. *Journal of Public Health*, 39, 625–632.
- Lancaster, Z. (2016). Expressing stance in undergraduate writing: Discipline-specific and general qualities. *Journal of English for Academic Purposes*, 23, 16–30.
- Lau, K., Lin, C.-Y., & Odle, E. (2020). ‘I am just saying maybe ...’: Engagement in dissertation defenses. *Language and Education*. <https://doi.org/10.1080/09500782.2020.1828450>.
- Lin, C.-Y. (2017). “I see absolutely nothing wrong with that in fact I think ...”: Functions of modifiers in shaping dynamic relationships in dissertation defenses. *Journal of English for Academic Purposes*, 28, 14–24.
- Lin, C.-Y. (2020). Pragmatic force modifiers in English-medium master's thesis defenses in Taiwan universities. *English for Specific Purposes*, 53, 16–30.
- MacKay, J. R. D., Hughes, K., Marzetti, H., Lent, N., & Rhind, S. M. (2019). Using National Student Survey (NSS) qualitative data and social identity theory to explore students' experiences of assessment and feedback. *Higher Education Pedagogies*, 4, 315–330.
- Martin, J. R., & White, P. R. R. (2005). *The language of evaluation: Appraisal in English*. New York: Palgrave Macmillan.
- Maton, K. (2014). *Knowledge and knowers: Towards a realist sociology of education*. New York: Routledge.
- Mezek, S., & Swales, J. M. (2016). PhD defences and vivas. In K. Hyland, & P. Shaw (Eds.), *The Routledge handbook of English for academic purposes* (pp. 361–375). London: Routledge.
- O'Donnell, M. (2011). *UAM CorpusTool (Version 3.3.2)*. Retrieved from <http://www.corpusool.com/download.html>.
- Paltridge, B., & Starfield, S. (2019). *Thesis and dissertation writing in a second language: A handbook for supervisors* (2nd ed.). London: Routledge.

- Paran, A., Hyland, F., & Bentall, C. (2017). Managing and mediating the research element on master's courses: The roles of course leaders and supervisors. In R. Breeze, & C. Sancho Guinda (Eds.), *Essential competencies for English-medium university teaching* (pp. 267–280). Cham: Springer.
- Patton, M. Q. (2015). *Qualitative research & evaluation methods: Integrating theory and practice* (4th ed.). Thousand Oaks, CA: SAGE.
- Plonsky, L., & Oswald, F. L. (2014). How big is “big”? Interpreting effect size in L2 research. *Language Learning, 64*, 878–912.
- Price, M., Handley, K., & Millar, J. (2011). Feedback: Focusing attention on engagement. *Studies in Higher Education, 36*, 879–896.
- Recski, L. (2005). Interpersonal engagement in academic spoken discourse: A functional account of dissertation defenses. *English for Specific Purposes, 24*, 5–23.
- Riekkinen, N. (2010). “This is not criticism, but...” softening criticism: The use of lexical hedges in academic spoken interaction. *Helsinki English Studies, 6*, 75–87.
- Schartel, S. A. (2012). Giving feedback – An integral part of education. *Best Practice & Research Clinical Anaesthesiology, 26*, 77–87.
- Skinner, E. A., & Pitzer, J. R. (2012). Developmental dynamics of student engagement, coping, and everyday resilience. In S. L. Christenson, A. L. Reschly, & C. Wylie (Eds.), *Handbook of research on student engagement* (pp. 21–44). Boston, MA: Springer.
- Sopina, E., & McNeill, R. (2015). Investigating the relationship between quality, format and delivery of feedback for written assignments in higher education. *Assessment and Evaluation in Higher Education, 40*, 666–680.
- Starfield, S., Paltridge, B., McMurtrie, R., Holbrook, A., Bourke, S., Fairbairn, H., ... Lovat, T. (2015). Understanding the language of evaluation in examiners' reports on doctoral theses. *Linguistics and Education, 31*, 130–144.
- Sutton, P. (2012). Conceptualizing feedback literacy: Knowing, being, and acting. *Innovations in Education and Teaching International, 49*, 31–40.
- van der Heide, A., Rufas, A., & Supper, A. (2016). Doctoral dissertation defenses: Performing ambiguity between ceremony and assessment. *Science as Culture, 25*, 473–495.
- van Heerden, M., Clarence, S., & Bharuthram, S. (2017). What lies beneath: Exploring the deeper purposes of feedback on student writing through considering disciplinary knowledge and knowers. *Assessment and Evaluation in Higher Education, 42*, 967–977.
- Vikers, C. H. (2007). Second language socialization through team interaction among electrical and computer engineering students. *Modern Language Journal, 91*, 621–640.
- Wagener, B. (2018). The importance of affects, self-regulation and relationships in the writing of a master's thesis. *Teaching in Higher Education, 23*, 227–242.
- Wang, T., & Li, L. Y. (2011). “Tell me what to do” vs. “guide me through it”: Feedback experiences of international doctoral students. *Active Learning in Higher Education, 12*, 101–112.
- Wei, J., Carter, S., & Laurs, D. (2019). Handling the loss of innocence: First-time exchange of writing and feedback in doctoral supervision. *Higher Education Research & Development, 38*, 157–169.
- Winstone, N., & Boud, D. (2019). Exploring cultures of feedback practice: The adoption of learning-focused feedback practices in the UK and Australia. *Higher Education Research & Development, 38*, 411–425.
- Winstone, N. E., & Carless, D. (2020). *Designing effective feedback processes in higher education: A learning-focused approach*. London: Routledge.
- Wisker, G. (2012). *The good supervisor: Supervising postgraduate and undergraduate research for doctoral theses and dissertations*. Basingstoke, England: Palgrave Macmillan.
- Xu, L. (2017). Written feedback in intercultural doctoral supervision: A case study. *Teaching in Higher Education, 22*, 239–255.
- Yang, M., & Carless, D. (2013). The feedback triangle and the enhancement of dialogic feedback processes. *Teaching in Higher Education, 18*, 285–297.
- Ylijoki, O.-H. (2000). Disciplinary cultures and the moral order of studying – A case-study of four Finnish university departments. *Higher Education, 39*, 339–332.