University students' perceptions towards using exemplars dialogically to develop evaluative judgement: the case of a high-stakes language test

Sin Wang Chong

Correspondence: S.Chong@qub.ac.
uk
Queen's University Belfast, Belfast,
UK

Abstract

The revived interest in the notion of 'evaluative judgement' in higher education is motivated by commitment of researchers and practitioners to effectively implement learning-oriented assessment and cultivate this high-order cognitive ability to develop students' capacity for self-regulated learning. Recent studies have examined the affordances and constraints of using exemplars to develop students' evaluative judgement in the fields of Education, Nutrition, Biology, and English for Academic Purposes. The present study, which focuses on use of exemplars in IELTS, analyzes patterns of teacher-students dialogues and 129 university students' perceptions of using exemplars to develop their understanding of assessment standards of IELTS academic writing tasks. Qualitative data were collected through an online questionnaire, individual semi-structured interviews, and workshop observations. Findings suggest that the IELTS instructor/researcher utilized various interactive strategies to develop students' hard, soft, and dynamic dimensions of evaluative judgement. Students identified affordances and limitations of using exemplars for language test preparation. Implications related to dialogic exemplar use to develop students' evaluative judgement are discussed in light of the findings.

Keywords: Exemplars, Evaluative judgement, Language test, Higher education

Introduction

Assessment in higher education is broadly categorized into formative and summative (Carter, Salamonson, Ramjan, & Halcomb, 2018); others have attempted to conceptualize assessment in a more fine-grained manner to encompass three forms: assessment *of* learning, assessment *for* learning, and assessment *as* learning (Chong, 2018a; Lee, 2017). Formative assessment, or learning-oriented assessment (LOA) (i.e. assessment *for* and *as* learning), has been attributed an important role in the learning process of university students because it enhances students' learning through their engagement with and reflection on assessment tasks . While both assessment *for* and *as* learning are regarded as a type of LOA, the former capitalizes the mediating role of

teachers to provide feedback (Carless, 2007) and the latter acknowledges the pivotal role of students as 'active agents in the assessment process' (Earl, 2013, Loc 553). Specifically, the dominant role of the students is exemplified in their involvement in peer (Liu & Carless, 2006) and self-assessment (Panadero, Alonso-Tapia, & Reche, 2013) tasks. With the idea that students learn best when serving as assessors, assessment literature in higher education has documented an array of assessment activities following the principles of assessment *as* learning; in particular, the involvement of students in the assessment of quality of exemplars has gained much attention (Carless & Chan, 2017; Carless, Chan, To, Lo, & Barrett, 2018; Chong, 2018b; Dawson, 2018). Using exemplars as a student-centered assessment activity, teachers develop students' feedback literacy (Carless & Boud, 2018) and refine their understanding about what 'good work' looks like with reference to a set of assessment rubrics (Dawson, 2015), a 'high-level cognitive ability' known as evaluative judgement (EJ) (Tai, Ajjawi, Boud, Dawson, & Panadero, 2018, p. 470).

EJ, which is defined as 'the capacity to make decisions about the quality of work of oneself and others' (Tai et al., 2018, p. 467), has received a revived attention in higher education literature. Originally developed by Sadler (1989) and known as 'evaluative knowledge' (p. 135) and 'evaluative expertise' (p. 138), EJ not only benefits university students' current studies but also their lifelong learning because they can make informed decisions about the quality of work they undertake in a self-directed manner (Cowan, 2010). Despite the potential development of students' EJ through the use of exemplars, there is a dearth of research which examines students' perceptions and experience of using exemplars to develop their EJ in the higher education context. Recent empirical evidence is provided by Carless and Chan (2017) (in the field of Education), Hendry and Jukic (2014) (in the field of Nutrition), Orsmond, Merry, and Reiling (2002) (in the field of Biology), and To and Carless (2016) (in the field of English for Academic Purposes). Nevertheless, there is no attempt to investigate the affordances and limitations of exemplars in developing students' EJ in high-stakes examinations. The present study aims to address this research gap by delving into 129 university students' perceptions of using writing exemplars to develop their EJ of IELTS Academic Writing Tasks 1 (data analysis) and 2 (essay writing) through triangulating qualitative data collected from student-teacher dialogues in lesson videos, an online post-workshop questionnaire, and individual semi-structured interviews.

Literature review

Theoretical underpinning of exemplar use: learning-oriented assessment

The use of exemplar is a form of LOA. Bloxham (2015) defines LOA as assessment tasks which 'lead to learning – when students are involved in evaluating their own work and when feedback is forward-looking so that students can act upon it' (p. 109). Multiple LOA models have been put forward by assessment researchers in the context of higher education which comprise three cornerstones: *tasks, students, feedback.* Carless (2007) puts forth a model which highlights the importance of designing assessment tasks as learning tasks which provide opportunities for students to serve as peer or self-evaluators so that feedback can be useful to students in future tasks (feedforward). Sambell, McDowell, and Montgomery's (2013) model has a slightly different take on

the three components. For assessment tasks, Sambell et al. (2013) contend that tasks need to be authentic, complex, and low-stakes. As for student involvement, these tasks should develop students' self-directed abilities and metacognitive knowledge. Regarding the nature of feedback, feedback should be rich and should be given formally (e.g. teacher feedback) and informally (e.g. peer discussion). In his revised LOA model, Carless (2015) includes the development of students' EJ as one of the components of his framework. Bloxham's (2015) book chapter introduces her LOA models which highlights task design and student engagement: LOA tasks should be authentic, challenging, integrated, purposeful, and relevant to students' learning. In addition, there should be active engagement of students to 'develop the capacity to find things out for themselves and learn independently' (p. 110). Comparing the LOA models of Bloxham (2015), Carless (2007, 2015), and Sambell et al. (2013), despite slight differences in terminologies, all the three models point to the importance of designing student-centered assessment tasks which act as learning opportunities for students to develop self-regulated skills (e.g., evaluating their own work or others' work) and metacognitive knowledge (e.g., developing a more solid understanding of assessment standards, Chong, 2020). Furthermore, all these models seem to hold a positive view towards the usefulness of feedback in these student-centered assessment tasks. Bloxhem (2015) and Sambell et al. (2013) added that tasks need to be authentic and low-stakes in order to facilitate a feedback-rich culture.

The use of exemplars is regarded as a form of LOA because its task, student, and feedback dimensions align with the aforementioned principles. In terms of task, the task of analyzing exemplars is purposeful in a sense that it aims to develop students' understanding of assessment requirements, which are often described by some opaque statements in rubrics. At the same time, the task is authentic and low-stakes. It is authentic because assessing the quality of a product is an indispensable task students will undertake in their professional life (e.g. a teacher will evaluate the quality of a student's work). As far as the role of students is concerned, students are placed at the center of the assessment task because they are the ones who not only are assessed but also assess others. Through repeated analyses of the quality of exemplars, students become independent learners who are readier to make judgements about the quality of a finished product with reference to a set of externally imposed criteria. Lastly, feedback plays a vital role in exemplar analysis activities. When analyzing exemplars, students engage in verbal feedback with peers and the teacher which encompasses four dimensions: *intentionality*, *transcendence*, *meaning*, and *reciprocity* (Lee, 2014, 2017). Intentional feedback refers to focused feedback which targets a particular strength or weakness of the exemplar with reference to a rubric; Transcendental feedback means feedback which functions as 'feed-up' or 'feed-forward' to enable students to improve their performance in future learning tasks (Hattie & Timperley, 2007); meaningful feedback entails feedback which allows students to compare the quality of the exemplar with their own performance. Lastly, reciprocal feedback, which is also known as dialogic feedback, is informed by sociocultural theory (Vygotsky, 1987) and the notion of mediation (Inaba, 2018). Dialogic feedback defines feedback as 'a process through which learners make sense of information from various sources and use it to enhance their work or learning strategies' (Carless et al., 2018, p. 1315) and includes not only cognitive but socio-emotional dimensions (Chong, 2018c).

Dialogic use of exemplars

Having established that use of exemplars constitutes a form of LOA, recent assessment literature in higher education suggests that the use of exemplars is most effective when accompanied with dialogues. In their recent study, Carless and Chan (2017) summarize 16 dialogic moves by a faculty member in a university in Hong Kong who teaches teacher education courses to students of science specialism. These 16 moves can be categorized into three functions: scaffolding, eliciting, and giving (Table 1).

From the 16 moves, Carless and Chan (2017) provide four principles to guide the practice of dialogic use of exemplars:

- Students' opinions, albeit divergent and developmental, should be valued.
- Students should be given the opportunities to verbalize and make explicit their thinking and reasoning.
- Students should engage in dialogues with their peers and the teacher through small-group and whole-class discussions.
- Students should be given scaffolding prior to analyzing exemplars.

While Carless and Chan (2017) categorize different dialogic moves used by a teacher, O'Donovan, Price, and Rust's (2008) framework differentiates the various degrees of 'dialogic-ness' when teachers use exemplars. The framework comprises four approaches to developing students' understanding of assessment standards: (1) a 'laissez faire' approach, (2) an 'explicit' approach, (3) a 'social constructivist' approach, and (4) a 'community of practice approach' (Fig. 1). In this framework, assessment dialogues between the teacher and students about the writing exemplars (as indicated by dotted and solid arrows in Fig. 1) is the key to facilitate students' understanding of assessment standards of assessment tasks (Ajjawi & Boud, 2018; Chong, 2018b). In the 'laissez fair approach', teachers do not discuss exemplars with students in lessons but distribute exemplars (often as unannotated 'model essays') for students' self-study before or after an assessment task. In this case, only internal dialogues take place within students because there
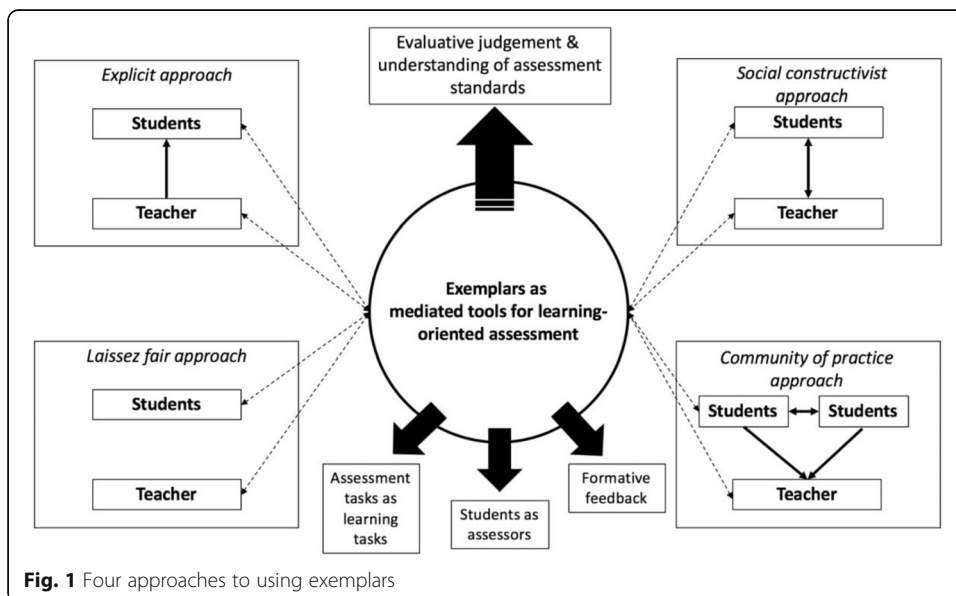

Fig. 1 Four approaches to using exemplars

is no opportunity for students to engage in discussions with the teacher regarding the quality of the exemplars given. Since students study the exemplars on their own without teachers' inputs, it is more likely for internal dialogues to take place only if students are motivated, of higher ability, and annotations are provided on to the exemplar. An 'explicit approach' to using exemplars refers to a more teacher-fronted approach to analyzing exemplars. In such cases, the teacher explains his or her judgement of the exemplar without eliciting students' opinions. A 'social constructivist' approach, on the other hand, resembles the dialogic use of exemplars documented in Carless and Chan (2017) in which students' divergent opinions about an exemplar is prioritized over the teacher's. Last but not least, a 'community of practice' approach is said to be the 'strongest' form of dialogic use of exemplars because, instead of the teacher, discussions on the exemplar are guided by students. A typical scenario of this approach includes students discussing an exemplar with a rubric in small groups and they give verbal comments on areas of the exemplar they choose to focus on. In the meantime, the teacher roams around the room facilitating the discussion of each group. In a 'community of practice' approach to using exemplars, teachers' judgement of an exemplar is usually withheld until the end, that is, after each group has reported their own evaluation, to encourage critical analysis and original thinking from students. While this study focuses on use of exemplars in language test preparation, which is the first of its kind, findings from a study conducted in the same context (Hong Kong) in a related field (English for Academic Purposes) by To and Carless (2016) suggested that students valued the opportunities to discuss exemplars with their teachers and recommended a 'social constructivist' or a dialogic approach to exemplar use to be adopted in language classrooms. Nevertheless, the authors maintained that occasional teacher-fronted instruction is essential to make students elucidate main features of the exemplars.

Conceptual framework of evaluative judgement

Nelson's (2018) framework of EJ is adopted as the conceptual framework of the current study. According to Nelson, a taxonomy of EJ comprises three types: *hard EJ*, *soft EJ*, and *dynamic EJ*. Hard EJ is the objective evaluation concerning the correctness of features of an exemplar; in Nelson's own words, opportunities of hard EJ come to light 'when we discriminate on the basis of truth' (p. 51). In the context of IELTS writing, hard EJ concerns areas involving grammatical (e.g. the use of correct verb tenses) and mechanical accuracy (e.g. the use of correct punctuations). Soft EJ, on the other hand, is judgement based on 'values and quality where the issue is not so much about right and wrong but how important things are' (p. 52). For instance, in IELTS writing, soft EJ is demonstrated through students' use of a variety of vocabulary and sentence patterns. While it is not incorrect to use the same expressions or sentence patterns repeatedly, 'variety of expressions' is regarded as an important area of assessment in the IELTS writing rubrics. Lastly, dynamic EJ concerns 'how to manage content' (Nelson, 2018, p. 53). Nelson admitted that this dimension of EJ is the most 'esoteric' (p. 53) because it is related to 'writing, organizing ideas in a creative or persuasive way'. In other words, dynamic EJ is about how one communicates with the audience of the work through presentation of logical and structured ideas.

Objective of the study

Although EJ has been receiving more attention in assessment literature in higher education, there is still a 'paucity of available research to provide evidence of the effectiveness of … exemplars' (Carter et al., 2018, p. 93). On the other hand, much of the literature on EJ has focused on its historical development (Tai et al., 2018), epistemological underpinnings (Goodyear & Markauskaite, 2018), and practical strategies to develop students' understanding (Thompson & Lawson, 2018) of the concept. It is, therefore, the objective of the current study to add empirical evidence to the research base, especially in an under-explored context which involves a high-stakes language test, IELTS. Specifically, the study is guided by the following research questions:

> RQ1: What dialogic moves were adopted by the teacher and how did these moves facilitate the development of students' soft, hard, and dynamic EJ?

> RQ2: What are the affordances and limitations of adopting a dialogic approach to using exemplars to develop students' EJ?

The study

IELTS writing workshops

The present study was conducted between May 2018 and December 2018 as a part of a Teaching Development Grant project led by the author in a teacher-training university in Hong Kong. The goal of this grant project is to develop an IELTS writing textbook to be used by university English teachers to teach the academic module of IELTS writing using an exemplar-based instructional approach (Chong & Ye, 2020). When the study was conducted, IELTS was the international English proficiency test adopted by the university as a Language Exit Requirement (LER) for all full-time undergraduates. Depending on individual program requirements, most students are expected to obtain at least a Band 6.5 in IELTS. In order to obtain student feedback on the teaching materials, a total of 16 IELTS writing workshops were conducted by the author between October and November 2018. The workshop sessions involved a total of 148 undergraduate and postgraduate students. The design of the workshops is shown in Figs. 2 and 3.

Two types of IELTS writing workshops were organized. While the design of these two aligns with Fig. 2, one type of workshops focused on IELTS Writing Task 1 which is data description; the other was about IELTS Writing Task 2, which is essay writing. In each 3-h workshop session, 10 to 15 students first took a pre-workshop quiz using Google Form which consists of 10 multiple-choice questions to test students' prior knowledge of IELTS writing assessment standards (Additional file 1). Then, students
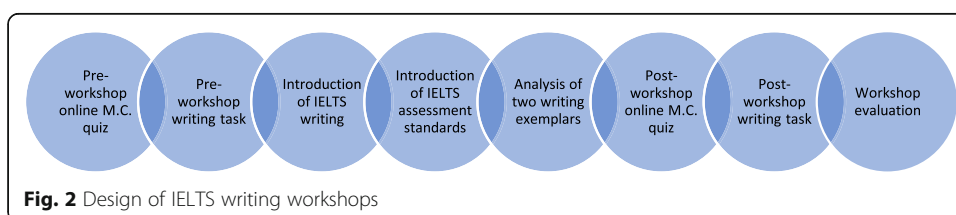

Fig. 2 Design of IELTS writing workshops

were given either a Task 1 or Task 2 question to complete to familiarize themselves with the IELTS writing test format (Additional file 2).[1] After these pre-workshop tasks, students were briefly introduced to the test format and question types of IELTS Writing Task 1 or Task 2. The next part of the workshop involves the introduction of the four domains of IELTS writing assessment standards: task achievement, coherence and cohesion, lexical resource, and grammatical range and accuracy. The introduction of the assessment standards was done using (1) a simplified version of the marking scheme containing only keywords from the original rubric and (2) two writing exemplars (one good and one poor) to illustrate the dimensions of quality. These exemplars were selected from over 100 exemplars collected from undergraduates and postgraduates at the author's university. Written consent was obtained from these students to use, adapt, and publish their writing as teaching materials. These two exemplars were selected by the author, who is an experienced IELTS writing instructor, because they illustrate typical writing features of low- and high-achieving IELTS candidates. When analyzing the exemplars, a social constructivist (dialogic) approach (O'Donovan et al., 2008) (Fig. 1) was adopted to maximize teacher-student interactions and elicit students' opinions. Moreover, such dialogic approach to using exemplars was implemented because research on exemplar use on a closely related field, English for Academic Purposes (To & Carless, 2016), made such recommendation. Teacher-led dialogues with students comprise a range of questioning techniques documented in Carless and Chan (2017) (Table 1). After teacher-led discussions on the two writing exemplars and the assessment standards, students completed two post-workshop tasks to check their understanding: (1) the same online quiz to check whether their declarative knowledge of the assessment standards had improved and (2) the same writing task. In order to avoid the problem of students imitating the surface features of the exemplars (Handley & Williams, 2011), the questions of the writing tasks and the exemplars were different.

Students' essays were sent to an external rater who is an experienced IELTS writing instructor for feedback and grading. Based on descriptive statistical analysis on SPSS, students ($n = 107$)[2] performed better in the post-workshop writing task in both IELTS Writing Tasks 1 and 2. When looking into students' performance in each of the four IELTS assessment domains in the pre-workshop and post-workshop tests, students' performance in Task 1 in 'task achievement' (TA) and 'coherence and cohesion' (CC) improved while their performance in 'lexical resource' (LR) and 'grammatical range and accuracy' (GRA) did not. In Task 2, students' performance in all the four assessment domains improved in the post-workshop test (Tables 2 and 3).

The pre-test and post-test results are provided in this section but not in the findings to provide the reader with a context and preliminary statistical evidence that exemplars are conducive to developing students' EJ, which was exemplified by their improved writing performance after the workshops. These statistical results are not included as the results of this study because of three reasons. First, this qualitative study focuses on

---

[1]This step was deemed crucial because most of the students at the author's university have not taken IELTS before. In Hong Kong, most students are admitted to university based on their results in the Hong Kong Diploma of Secondary Education (HKDSE), a territory-wide public examination for all secondary school students. The minimal English requirement for a student to be admitted to a university is to obtain a Level 3 in English Language in HKDSE, equivalent to an IELTS Band 5.

[2]Results of only 107 students were included ($n = 60$ for Task 1 and $n = 47$ for Task 2) because not all the students who participated in the workshops completed both the pre-test and post-test.

Table 1 Dialogue features and dialogic moves in Carless and Chan (2017)

Dialogue functions	Dialogic moves
Scaffolding	1. Providing a structural scaffolding to the dialogue
Eliciting	2. Eliciting views from students as a group 3. Eliciting views from an individual student 4. Eliciting questions from students 5. Eliciting explanations from students 6. Eliciting examples from students
Giving	7. Giving a summary of students' views 8. Giving a paraphrased version of students' views 9. Giving an elaboration on students' views 10. Giving the teacher's own views 11. Giving wait time 12. Giving instructions to refer to rubrics 13. Giving instructions to refer to exemplars 14. Giving examples 15. Giving the teacher's own interpretation of exemplars 16. Giving compliments or acknowledgements

dialogic moves of the teacher and students' perception of exemplar use. Second, the statistics were originally collected for internal evaluation purpose and they were not publication-ready (e.g. the absence of another rater to calculate inter-rater reliability due to budget constraint). Third, because of the one-off design of the project, it is difficult to make claims based on statistical comparison because longitudinal observation (e.g., with a delayed post-test) was not possible.

Participants
Amongst the 148 students who participated in the IELTS writing workshops, 129 agreed to participate in the study by indication on the informed consent form. Among the 129 participants, 112 of them (87%) were undergraduate students from a wide range of programmes, mostly from teacher education programmes (e.g. Arts Education, Chinese Language Education, Early Childhood Education, English Language Education, History Education, Science Education, Special Education); 17 (13%) of them were from the Graduate School of the university (Master of Arts, Master of Education, and Master of Teaching). In terms of genders, 102 (79%) of the participants were females while the rest were males (21%). Students were given the chance to indicate whether they would like to participate in a post-workshop individual interview on the consent form. Amongst the 89 students who originally indicated that they were willing to participate in a student interview, only 16 students attended one because the interviews were scheduled in late November and early December during which students were busy with final assignments and exam preparation. Due to the stimulated recall nature of the interviews, interviews could only be conducted immediately after the workshops when students were still able to recall their learning experience.

Data collection and analysis
Data were collected from three sources: (1) online post-workshop questionnaire (Additional file 3), (2) individual student interviews, and (3) student-teacher dialogues during workshops. RQ1 was answered by triangulating data from (2) and (3) whilst RQ2 was answered by triangulating data from (1) and (2).

Table 2 Descriptive statistics of pre-test and post-test results (Task 1)

	N	Range	Minimum	Maximum	Mean	Std. Deviation
TA_Pre	60	4	4	8	5.32	1.033
TA_Post	60	4	3	7	6.03	0.610
CC_Pre	60	5	3	8	5.12	0.904
CC_Post	60	4	3	7	5.53	0.724
LR_Pre	60	4	4	8	5.38	0.846
LR_Post	60	3	4	7	5.17	0.785
GRA_Pre	60	3	4	7	5.07	0.710
GRA_Post	60	2	4	6	5.05	0.675
Total_Pre	60	16	15	31	20.88	2.992
Total_Post	60	13	14	27	21.78	2.218

(1) Online post-workship questionnaire

One hundred twenty-nine students participants completed an online post-workship questionnaire form on Google Form. This evaluation form consists of 21 items. Items 1 to 5 are related to students' personal information; items 6 to 16 are Likert-scale items with four points (strongly disagree, disagree, agree, strongly agree) based on the template of the university's Student Evaluation of Teaching (SET). Finally, items 17 to 21 are open-ended questions. Items on Likert-scales were analysed using descriptive statistical methods to identify the distributions of students' perceptions and experiences of using exemplars to prepare for IELTS. Open-ended items were downloaded to a Word document and inputted into *NVivo* for coding. For the current study, only findings from open-ended responses (items 17–21) are reported.

(2) Individual student interviews

Sixteen students participated in an individual semi-structured individual interview conducted either face-to-face or on the phone by a research assistant which lasted for approximately 30 min. These interviews, which were audio recorded, focused on students' experience of using exemplars to understand the assessment standards of IELTS writing and their perceived usefulness and limitations of using exemplars to develop such understanding and their EJ. The interviews were conducted in either students' first language (Cantonese or Mandarin) or English, based on the

Table 3 Descriptive statistics of pre-test and post-test results (Task 2)

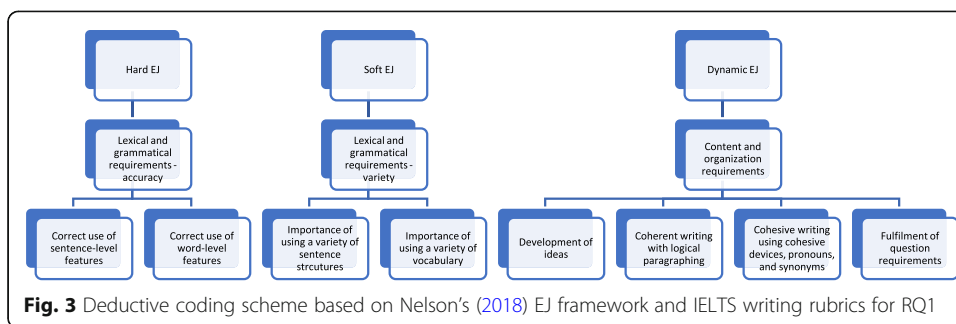|	N	Range	Minimum	Maximum	Mean	Std. Deviation
TA_Pre	47	3	4	7	5.45	0.928
TA_Post	47	3	4	7	5.87	0.647
CC_Pre	47	3	4	7	5.53	0.881
CC_Post	47	3	4	7	5.77	0.666
LR_Pre	47	2	4	6	5.11	0.667
LR_Post	47	2	4	6	5.36	0.568
GRA_Pre	47	2	4	6	4.83	0.637
GRA_Post	47	2	4	6	5.26	0.642
Total_Pre	47	10	16	26	20.91	2.741
Total_Post	47	9	16	25	22.26	2.027

Fig. 3 Deductive coding scheme based on Nelson's (2018) EJ framework and IELTS writing rubrics for RQ1

students' preference to enable effective communication of ideas. All the 16 interviews were (translated and) transcribed by a research assistant who holds a master's degree in education. Accuracy of transcription and translation was checked by the author and another research assistant who holds a master's degree in English language studies by listening to the recordings, comparing the transcriptions with the recordings, and making changes to words or phrases which are grammatically and semantically inaccurate. The transcribed data were inputted into *NVivo* for coding.

(3) Student-teacher dialogues

Interactions between the teacher and students in the workshops were video-recorded, transcription and counter-checking were done in the same fashion as the interview data. All the data were imported to *NVivo* to ensure consistency and accuracy of coding.

To analyse the qualitative data, both deductive and inductive coding approaches were used (Saldaña, 2016). A deductive coding approach was adopted for RQ1 to determine 'a provisional list of codes' with reference to the study's conceptual frameworks (Nelson's framework of EJ, 2018) (p. 74) (Fig. 3). At the same time, to ensure the data analysis process embraced the complexities of newly emergent themes (Fig. 4), an inductive approach was adopted for RQ2 using the constant comparison method of grounded theory (Charmaz, 2006). According to Charmaz (2014), grounded theory is a set of:

Systematic, yet flexible guidelines for collecting and analysing qualitative data to construct theories from the data themselves … Grounded theory beings with inductive data, invokes iterative strategies of going back and forth between data and analysis, uses comparative methods, and keeps you interacting and involved with your data and emerging analysis. (pp. 1)
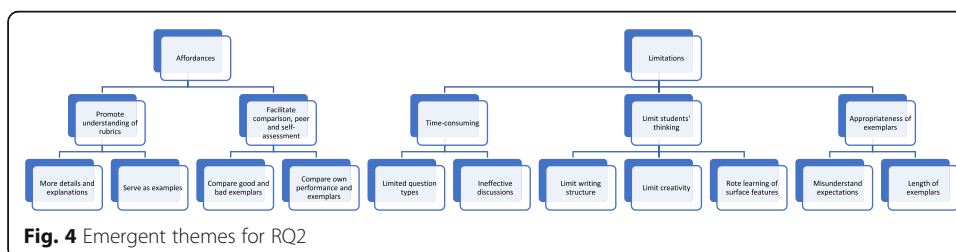

Fig. 4 Emergent themes for RQ2

Such a 'hybrid approach' to coding was encouraged by Saldaña (2016) because 'each qualitative study is unique'; thus, the coding approach must be 'customized to suit the unique needs and disciplinary concerns of your study' (p. 74).

To enhance the credibility and trustworthiness of the qualitative data reported, all translations and transcriptions were done after repeated listening and viewing of the recordings by a research assistant. The translations and transcriptions were cross-checked by the author and another research assistant regarding accuracy of content and grammar. Moreover, triangulation, a 'cross-verification of data from a range of sources', was carried out to compare insights from the interview and questionnaire data with the observations from lesson videos (Lambert, 2019, p. 128). These three mechanisms are adopted because they are regarded as examples of good practices of conducting fine qualitative studies (Savin-Baden & Major, 2012).

Findings

Dialogic patterns to develop students' hard EJ

With reference to Table 1 (Carless & Chan, 2017), the analysis of the dialogic patterns used to develop students' hard EJ indicates that the teacher employed only one dialogic move, that is, giving compliments or acknowledgement. The lack of interactions between the teacher and students when discussing grammatical accuracy in the exemplars is possibly due to the teacher's focus on 'correctness'. Since the grammatical aspect of writing concerns 'absolute truth' (e.g., the past tense should be used in recounting) instead of 'interpretive truth' (e.g., the word used here is not advanced enough), the teacher paid less attention to negotiating understanding with his students. The dialogic pattern employed by the teacher below is accompanied with two examples:

$$\boxed{\text{S's opinion}} \rightarrow \boxed{\text{T's compliment or acknowledgement}}$$

Key:

$\boxed{\text{XXXXX}}$ = dialogic move

Example 1

Speaker	Transcription
Student:	In my opinion, the weakness of this passage is the tenses of some verbs. For example, in paragraph 2, the third sentence, "City A remains …", "remains" is in the present tense; however, we should use the past tense. So, I would suggest him change it to the past tense. For example, I will use the previous sentence as an example. It should be: 'however, the housing price of City B dramatically *increased* from 2008 to 2016'.
Teacher:	Yeah, very clear, thank you so much.

Example 2

Speaker	Transcription
Student:	Our group would like to comment on a grammar problem … We can refer to the third paragraph. He said, 'the society is needing more and more …', we think he could use 'needs' instead of 'is needing', to make it clearer.
Teacher:	Okay, thank you very much.

Regarding hard EJ which focuses on correctness of lexical and grammatical features in writing, students seemed to focus more on the correct use of sentence-level features (10 references) than word-level features (5 references). In alignment with the assessment rubrics of IELTS writing, students analyzed the accurate use of sentence patterns and punctuations in the exemplars given. Among the 10 references, seven references were related to grammatical accuracy whilst three references were associated with the correct use of punctuations. Different from the rubrics' focus on sentence pattern accuracy, students also focused on other linguistic accuracy features they identified during workshop discussions. For example, one student focused on the inaccurate use of adjectives ending with 'ed' to describe a phenomenon:

> *Our group focus on grammatical errors. For the second paragraph, we can see there are a lot of confusing grammar, like here, … "satisfied" is wrong.*

It was only during post-workshop interviews that they mentioned how they could use complex sentences correctly:

> *The teacher emphasized those complex sentences, like some subordinate sentences. All in all, he was more specific than the official one.*

In the interviews, one student was able to point out that the IELTS rubrics focus on the accurate use of simple and complex sentence patterns instead of general grammatical accuracy:

> *I learned something from the class that it is not necessary to pay full attention to all the details of the grammar; instead I could express my ideas with various sentence patterns.*

In contrast, no students focused on the use of punctuations when analyzing the exemplars during workshops; the three instances when students mentioned correct use of punctuations were in the post-workshop interviews when they were asked what assessment standards they have learned in the workshops:

> *Student: Maybe different kinds of punctuation?*

> *Interviewer: okay, different kinds of punctuation.*

> *Student: Yeah, some of them are very specific, but I really learned something.*

> *Interviewer: What did you learn?*

> *Student: double quotation marks, commas, parentheses …*

Endorsed by only 5 references, very few students focused on word-level accuracy. Despite analyzing the exemplars with an emphasis on accurate use of words, these students did not focus on the three word-level linguistic features highlighted in the

rubrics, namely collocations, parts of speech, and spelling. In one of the workshops, a student pointed out a word-level accuracy problem related to subject-verb agreement:

> *I would like to focus on the grammar, for example, in the first paragraph, "appear", it should be "appears", I think.*

In a similar vein, in the interviews, students were concerned with how they should use correct verb tense to describe data which are related to the past for IELTS Writing Task 1:

> *If it requires us to organize some data, we should use the past tense. If it requires us to describe some facts, we should use the present tense.*

Referring to the data from workshop observations and student interviews, students' focus on grammatical accuracy seemed to differ from the IELTS marking scheme. While the rubrics' emphasis was on types of sentence structures (e.g. simple, compound, complex), punctuations, collocations, word forms, and spelling, data from workshop observations suggested that the students commented on other grammatical problems in the exemplars (i.e. inaccurate use of adjectives, verb tenses, and subject-verb disagreement). The students were only able to accentuate the grammatical areas in the rubrics during the post-workshop interviews. One reason why students failed to align their discussions on exemplars with the rubrics' focus is that the teacher did not utilize any elicitation techniques to guide students to focus on rubrics' descriptors. From the workshop observation data, the teacher acknowledged and/or complimented the students' comments on the grammatical inaccuracy of the exemplars without asking them to evaluate the exemplars with reference to those areas underscored in the rubrics. Second, students did not give comments on the word-level and sentence-level features mentioned in the rubrics because they were unfamiliar with those relatively more advanced features. For instance, despite understanding the concept of complex sentences (the combination of an independent clause with at least one dependent clause) from the 3-h workshop, their understanding was still preliminary which limited them from identifying sentences in the exemplars to be rewritten into complex ones.

Dialogic patterns to develop students' soft EJ

Amongst the three dialogue functions suggested by Carless and Chan (2017), the teacher mainly employed the 'giving' and 'eliciting' moves to develop students' soft EJ. Two typical patterns employed by the teacher are as follows:

Pattern 1: Responding to students' opinions

S's opinion A → T's summary/paraphrasing of A → T's opinion on A → (T's reference to rubric)

Pattern 2: Guiding students to elaborate on their opinions

S's opinion A → T's elicitation of view/example → S's response → (T's compliment) → S's opinion B → T's elicitation of view/example → S's response → (T's compliment)

Key:

XXXXX = dialogic move

(XXXXX) = optional dialogic move

In the following excerpt, what the teacher did is responding to the students' elaborated opinions in two ways. First, he used factual information as evidence to support his own viewpoint. For instance, the teacher reckoned the cohesive devices used by the writer were not '*good*' because they were '*not difficult enough*'. He then asked the students to refer to the marking scheme to show them 'difficulty' of vocabulary is one of the assessment standards. Second, the teacher's comments were responsive to the opinion presented by the student. At the end of the dialogue, the teacher summarized what the student said regarding the mechanical use of the term 'virtual career' before commenting that it is not an effective way to write using the same words repeatedly. This approach of dialogic moves was used with more competent students who were able to give a thorough evaluation of the quality of the exemplar regarding its word and sentence choices.

Example 3

Speaker	Transcription
Student:	We think in the paragraph, there are not enough cohesive devices. The author always uses "firstly", "secondly". Maybe he should use some other words like "moreover", "furthermore". Second, he always uses the word "virtual career", uses the same words, and he didn't use any synonyms. Okay, these are all of our suggestions.
Teacher:	Okay, thank you so much, let's give them a big hand. I think they are really concerned about vocabulary. They are not happy with the words "firstly", "secondly", and "thirdly". These are "cohesive devices", I understand. These are not good cohesive devices, because they are not difficult enough. Remember the key words in the marking scheme? "difficultly", right? So, we would like to use difficult cohesive devices. Also, they point out that the term "virtual career" is in every paragraph, so there is no attempt by this student to use a synonym; he just copied the word from the question - it is convenient, but it is not very effective, right?

In situations where students seemed to demonstrate a lack of understanding of the quality of the exemplar regarding its use of vocabulary and sentence structures, dialogues were divided into smaller chunks by having students respond to questions the teacher asked to elicit their views or some additional information. To encourage these students to express their thoughts, the teacher occasionally used praises as a means to motivate the students to speak more.

Example 4

Speaker	Transcription
Student:	There could be some change of words, or parts of speech.
Teacher:	But it is pretty difficult to think of a synonym for 'housing price'. What about verbs? Do you see any verb that is used repeatedly?
Student:	'Increase'.
Teacher:	Exactly.
Student:	It is here and, here. *[pointing to the PowerPoint slide showing the exemplar]*
Teacher:	Are you happy with the words being used repeatedly?
Student:	No.
Teacher:	Not really, right? So, what would you suggest as a synonym for the student to use?
Student:	'Rise'.

Findings *(Continued)*

Speaker	Transcription
Teacher:	Yes. Another example? Do you remember what you learnt from the first exemplar, by the way?
Student:	'Surge' and 'climb'.

In terms of students' soft EJ development, more students focused on (1) 'importance of using a variety of vocabulary' (33 references) than (2) 'importance of using a variety of sentence structures' (13 references), which is different from students' hard EJ which focuses on sentence-level accuracy rather than word-level accuracy.

One example to support (1) is how students articulated the importance of using synonyms to replace the verb 'increase' when describing the trend of data for IELTS Writing Task 1 (Example 4). Similarly, in the interviews, students repeated the importance of using synonymous expressions as one of the requirements in the rubrics:

> *Synonyms, I think I have poor vocabulary … … I should really work hard on this part when I write essays. I often repeat the same words because maybe I really don't have very rich vocabulary Sometimes I memorize some words before the test but I forgot the meaning when the test begins, maybe I get nervous, I forgot some of the words. So, I think this is very important.*

Students interviewed also recalled another requirement in the rubrics, the importance of forming new words using suffixes:

> *Interviewer: Okay. So, which aspects do you think you have a better understanding now?*
>
> *Student: All aspects.*
>
> *Interviewer: Like what?*
>
> *Student: Like vocabulary. Teacher taught us what suffix is.*

A number of students focused on evaluating the range of sentences used in the writing exemplars during the workshops:

Example 5

Speaker	Transcription
Teacher	*How about the use of language? what do you think? Or, sentence structure?*
Student	*I think the sentence structure is simpler than the previous one.*
Teacher	*So, you think it is not very good for this paragraph?*
Student	*It is easy to understand, but I think it could include more sentence patterns, or connection words.*

Some students were able to give concrete suggestions on how the writer of the exemplar could make his/her sentences more varied in the workshops:

Example 6

Speaker	Transcription
Student	*The first thing is about the usage of simple sentences. For example, in this essay, the first paragraph, the first sentence, the writer said "I think" here. We think he can use some complex sentences. For instance, use some WH words, use semicolons, etc.*

Student interviewees highlighted a similar focus on the importance of using lengthier sentences, including compound and complex sentences:
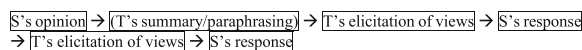
If we want to get a high mark, we need to compare the similarities and differences systematically … to use … some compound sentences or complex sentences.

You learn something good about the exemplar. Under the guidance of the teacher, you know what a better sentence is like, then you can try to learn that sentence or recite that sentence after class.

Compared with the dialogic patterns used by the teacher to develop students' hard EJ, the teacher employed two dialogic patterns when it comes to cultivating students' understanding of the importance of using a variety of vocabulary and sentence patterns. To less capable students, the teacher guided students to elaborate on their opinions by asking them follow-up questions. To the higher-ability students, the teacher summarized the students' opinions before giving his own views, in an attempt to build his opinion on the students' comments. The employment of these two dialogic patterns to cater for learner diversity proved to be quite effective. Triangulating data collected from workshop observations and student interviews, the findings suggest that most of the students were able to identify strengths and weaknesses in the exemplars regarding vocabulary and sentence range, as well as verbalizing the importance of using synonymous words and expressions in IELTS writing. In one of the examples above, however, students were only able to verbalize the significance of forming new words using suffixes in an interview but failed to identify words in the exemplars which could be replaced by a new word formed with a suffix. This setback is believed to be caused by students' lack of domain-specific knowledge on morphology (word formation). Acknowledging students' deficit knowledge on how English words are formed, it can be argued that the two dialogic patterns used by the teacher were constructive because, at the very least, they raised students' awareness of an aspect of English which is previously unknown to students.

Dialogic patterns to develop students' dynamic EJ

Dynamic EJ, in the words of Nelson (2018), is the most 'esoteric' dimension of EJ because it has to do with the logical and coherent presentation of ideas in a written text. Comparatively speaking, dynamic EJ is less definite than hard EJ which is about grammatical accuracy, and it is more abstract than soft EJ which emphasizes the importance of using synonymous words and phrases. With this understanding, it is not surprising to find that the teacher's dialogues with the students focused on clarifying students' understanding of the meaning of idea elaboration and coherent development of ideas through various questioning techniques:

S's opinion → (T's summary/paraphrasing) → T's elicitation of views → S's response
→ T's elicitation of views → S's response

Key:

XXXXX = dialogic move

(XXXXX) = optional dialogic move

In the following excerpts, the teacher asked the students a series of questions in order to guide the students to identify the problematic areas related to content and organization in the exemplars. In the first excerpt, the teacher began by summarizing the comments made by the student. Then, the teacher asked the student a series of yes/no questions to guide her to identify an erroneous area in the introduction. Through these questions, the student reflected on the fact that the author's standpoint should not be placed at the beginning of an introduction but at its end. In the second excerpt, the teacher adopted a similar dialogic pattern. In particular, he used questions to prompt the student to verbalize her comment in a more specific manner. Instead of saying that the overview should be placed in the introductory paragraph, the teacher guided the student to specify the appropriate location of an overview in an introduction.

Example 7

Speaker	Transcription
Student:	He divided the advantages into three parts, "firstly", "secondly", and "thirdly". Also, the ideas are very similar with those in the third paragraph.
Teacher:	Okay, so, the second paragraph is similar with the third paragraph.
Student:	Yes.
Teacher:	And for the background, the position is missing, right? Is there any position? Is there a standpoint?
Student:	Yes, the first sentence.
Teacher:	Yeah, there is a position, but the location of the position, it is very strange. It should be at the end of the paragraph, not the beginning, right?
Student:	Yes.

Example 8

Speaker	Transcription
Student:	First of all, because he puts the overview in the second paragraph …, we suggest that he put the sentence in the first paragraph.
Teacher:	So, where should you put 'overall, the housing price of both the cities and the other period given'?
Student:	The first paragraph.
Teacher:	Where? Before the sentence or …
Student:	The last sentence.
Teacher:	The last sentence of the paragraph?
Student:	Yes.
Teacher:	Because this is the what?
Student:	Because this is the overview.
Teacher:	This is the overview.
Student:	Yes.

Dynamic EJ focuses on students' judgement on the content and organization of a written work. Specifically, four aspects of dynamic EJ were noted based on Nelson's conceptual framework and the deductive coding scheme (Fig. 3): (1) Coherent writing with logical paragraphing (29 references), (2) fulfilment of question requirements (14 references), (3) development of ideas (12 references), and (4) cohesive writing using cohesive devices, pronouns, and synonyms (6 references).

Students were most competent in assessing the appropriateness of paragraphing of a written work when analyzing the two IELTS writing exemplars. In the workshops, some students were able to identify the essential components of a paragraph in IELTS writing:

Example 9

Speaker	Transcription
Teacher	*How many parts are there in the introduction?*
Student	*Three?*
Teacher	*Good, can you tell me what the three parts are?*
Student	*First one is the definition of language franca, second one is people's opinions of language, last one is the writer's opinion.*
Teacher	*Exactly.*

Paragraphing was also mentioned by students in the interviews as one of the organizational requirements of IELTS writing:

Coherence is paragraph. I should divide the passage into suitable paragraphs, maybe 3 paragraphs or 4 paragraphs, depending on the type of the essay.

In other occasions, students demonstrated a more in-depth understanding of the structural requirements of IELTS writing by commenting on how the students combined unrelated ideas into one paragraph and giving relevant suggestions:

Example 10

Speaker	Transcription
Student	*The most obvious problem of this essay is the structure … the fourth paragraph, "although", and the last paragraph, it begins with "for example", that is very confusing.*
Teacher	*Very confusing, okay.*
Student	*Yeah, because we don't know which paragraph supports which idea. My suggestion is to separate the last two paragraphs into several sentences and divide them into the first and the third paragraphs.*
Teacher	*So, you would group some of the ideas together.*
Student	*Yeah.*
Teacher	*Okay, thank you very much.*

Regarding fulfillment of question requirements, students participating in both Task 1 and Task 2 workshops were able to verbalize the importance of including relevant ideas as required by the writing questions:

Task 1

Student: The other one is task achievement. Wait, let me think about it. For task, for example, can I talk about Task 1?

Interviewer: Sure.

Student: For Task 1, we need to compare and contrast, and describe some data.

Interviewer: Data, okay.

Student: And, … we should cover the whole period.

Task 2

We need to cover all of the questions in our articles, which is the basic requirement.

Another aspect of dynamic EJ students focused on when analyzing the writing exemplars was how ideas should be developed. One specific method frequently mentioned by students was the use of relevant examples:

Example 10

Speaker	Transcription
Teacher	*What do you think about the first body paragraph?*
Student	*Very clear.*
Teacher	*What makes it clear?*
Student	*His examples relate to the key point.*

In addition, with reference to the rubrics and exemplars, students were able to identify irrelevant examples which are detrimental to logical development of ideas:

Example 11

Speaker	Transcription
Student	*And we think the examples are not relevant to his viewpoints. He supported that virtual jobs have advantages, but the examples are not really supporting it.*

Another feature described by the student interviewees as indicative to good idea development is to focus on developing ideas related to the thesis statement of the writing:

When the teacher taught us how to support our ideas, I got to know how to control the length of my article. For example, if I have two different points of view, one is positive and one is negative, I might write more sentences to support my positive one than the negative one, because the positive one is my main idea that I want to emphasize.

Very few students, however, focused on how to structure a piece of writing by using word-level features, including cohesive devices (e.g. in addition, nevertheless) in the workshops.

Unlike the dialogic patterns used by the teacher to develop students' hard and soft EJ, the teacher mainly relied on elicitation of clarification to hone students' dynamic EJ. For instance, as shown in the excerpts, students' comments on the content and organization aspects of the exemplars were generally brief. The students were only able to give more detailed feedback to the exemplars after being prompted by the teacher. In other words, the teacher's dialogic pattern was useful in a sense that it assisted students in deepening their understanding of the rubrics and, in so doing, giving more elaborate feedback. Nonetheless, the dialogic pattern used by the teacher was limited because students were unable to identify from the exemplars more abstract issues such as macro-structure (e.g., the connection between body paragraphs and thesis statement) and issues which require domain-specific (linguistic) knowledge, including the use of different cohesive devices.

Affordances of using exemplars

From the student interviews and open-ended questionnaire responses, students perceived the use of exemplars as an effective means to promote their understanding and EJ of IELTS writing assessment standards because of two reasons: it (1) promotes students' understanding of rubrics (31 references) and (2) facilitates comparison through peer and self-assessment (24 references).

The writing exemplars and discussion promote students' understanding of rubrics by providing more details and explanations about the keywords of the rubrics ("*I learn how to understand the descriptors in a more detailed way, and I now know how to write to gain a higher mark*" – open-ended questionnaire response). The following student explained how the exemplars could enhance students' understanding of a keyword of the rubrics:

> I think it could be better if we talk about these [requirements in rubrics] while reading the exemplars, integrating these two, and no need to explain them separately. It is meaningless to explain them separately. For example, when we talked about pronouns, we might not know how to place the pronouns in a sentence. We need to learn how to use pronouns. (student interview)

The reason why students find the exemplars useful in explaining the keywords in the rubrics is because they are '*concrete example [s]*' and '*vivid illustration [s]*' which make '*the marking scheme more specific*' (open-ended questionnaire response).

Another affordance of exemplars which develops students' EJ is that they enable students to compare their own work with the exemplars and have '*a model sample for reference*' (open-ended questionnaire response). Such opportunities for individual reflections and group discussions on exemplars help students reflect on areas which they can improve in their work:

When we use exemplar, we would automatically compare our own articles with the exemplar, thinking about the differences between our articles and the exemplar, then I know what I should improve on. (student interview)

When our group looked through the bad exemplar, which was of 5.5 points to 6 points, we know that is an article that is at the same level as ours, then we know what our problems are. (student interview)

In addition to self-assessment, peer assessment of different exemplars written by other undergraduate and postgraduate students may promote deep learning of assessment standards of IELTS writing. Students expressed the usefulness of comparing exemplars at different levels because the exemplars enriched their understanding of the quality of writing against a rubric through '*show [ing] me the different levels of writing and giv [ing] me some references*' (open-ended questionnaire response):

You can compare, because he (the teacher) gave us two passages. So, while you are comparing the two passages, you can consider why this one gets higher mark, or somehow, to learn the … how to say … to learn from this comparison, and to improve your writing skills. (student interview)

The exemplars can let me know more about the differences between a higher-scored and lower-scored writing. (student interview)

Specifically, one student demonstrated how she gained a better understanding of 'good' and 'bad' essays for IELTS Writing Task 1 in an interview:

If an article is not that good, it might not be able to describe expressively at the beginning of their article, I mean the topic sentence, and might not be able to use enough data to support, or it might use the tense in a wrong way. (student interview)

Limitations of using exemplars

Although the findings described earlier alleged to the fact that the use of exemplars has the potential to develop students' EJ and students seemed to acknowledge the usefulness of using exemplars in developing their understanding of assessment standards, three limitations have been identified in this study:

- It is difficult to select 'appropriate' exemplars for analysis (22 references).
- It is time-consuming to analyze exemplars (18 references).
- It may limit students' thinking (18 references).

Many students shared the opinion that some of the exemplars discussed in the workshop were not too appropriate for them for a number of reasons. First, they were concerned with the length of the exemplars, commenting that it '*takes too long*' to read (open-ended questionnaire response):

There are many good vocabulary items and sentences in the exemplar, but it is a little bit long. (student interview)

Second, they found the exemplars limiting because they did '*not include all the problems that students usually make*' and '*may not include all aspects of the marking scheme*' (open-ended questionnaire response). One student suggested in the interview that the teacher should make changes to the exemplars to cover more assessment standards in the rubrics:

I am not sure if the exemplars were the original ones from some students or were modified a little bit by the teacher. I hope the teacher could do some modifications on it, to cover as many points of the descriptors as possible. (student interview)

Lastly, a negative remark made by the students towards the use of exemplars to develop their EJ is that the exemplars cannot cover all the question types which may appear in the high-stakes language test:

I may face other types of questions which are not mentioned in the workshop in the IELTS exam later. (open-ended questionnaire response)

I hope there will be different exemplars including line graph, pie chart, and they can be categorized according to their types, especially those mixed types. (student interview)

While students perceived the usefulness of exemplars as effective in developing their EJ, some found the discussions, analyses, and explanations of exemplars too '*time-consuming*' (open-ended questionnaire response). Because of the long duration needed for students to discuss the exemplars and for the teacher to explain the exemplars in relation to the rubrics, only a small number of exemplars could be discussed:

Due to the time limitation, we could only read 2 exemplars. If possible, I really wanted to read more in class so that I could learn more. (student interview)

There might be other bad exemplars in which there might be other types of mistakes that we have not explored . (student interview)

From the perspective of the students, the third limitation of using exemplars to develop EJ is that the exemplars would '*confine thinking*' and '*limit our own thinking*' because students may '*imitate others unconsciously*' (open-ended questionnaire response). Similar views are shared by students who attended the interviews who thought both their ideas and structure of their writing would be limited by the exemplars they have read:

If I read the exemplar first, every idea and logic in every paragraph of the exemplar, it would definitely affect what I am going to write. (student interview)

I would be limited by the structure and cannot think of something else and new. When you attend a test, without any exemplars, you would not know how to structure, how to group ideas. (student interview)

Upon reviewing students' perception of the affordances and limitations of using exemplars to prepare them for a high-stakes language test, it is found that students generally appreciated the use of exemplars to help them better understand the different dimensions of quality in the rubrics. Moreover, the use of exemplars was perceived as a metacognitive experience by students in which they compare their writing performance with the exemplars analysed and reflect on their strengths and weaknesses. Nevertheless, some students were concerned that the exemplars would limit their creativity and thinking because they would copy the surface features of the exemplars. Another worry is related to the amount of time needed to analyse an exemplar. Since the context of this study is a one-off, three-hour workshop, students would like to read more exemplars which cover different facets of the assessment requirements. The opinions of the students suggest that students acknowledged the usefulness of exemplars in developing their EJ and understanding rubrics of IELTS, but their concerns reveal the need for a more thorough and thoughtful planning of how exemplars are to be included in a language classroom. In light of these findings, some suggestions about how exemplars can be incorporated into a language curriculum are offered in the next section.

Discussion and implications

Developing students' hard, soft, and dynamic EJ

Referring to Nelson's (2018) conceptual framework of EJ, the findings suggest that the use of exemplars can encourage students to focus on all three dimensions of EJ: soft, hard, and dynamic, as evident by the teacher-student dialogues and individual student interviews. The deepening of students' EJ can be attributed to students' acute understanding of the assessment rubrics and opportunities for students to engage in self-assessment and peer assessment to compare written work against an explicated rubric. Among the three types of EJ, students appeared to understand the dynamic dimension of EJ the most, with 61 references, followed by soft EJ, with 46 references, and hard EJ, with 15 references. This finding is in contrast with Nelson's (2018) argument that dynamic EJ is the most 'esoteric', which makes it the hardest for students to understand.

From the findings presented, students' declarative knowledge of the dynamic aspect of EJ of IELTS writing was strong because students were able to articulate different facets of content and organization requirements of IELTS writing tasks, namely development of ideas, coherent writing with logical paragraphs, cohesive writing using cohesive devices, pronouns, and synonyms, and fulfilment of question requirements. The affordances of exemplars in developing students' understanding of such esoteric requirement of academic writing could be explained by the notion of 'tacit knowledge', which was first discussed by Polanyi (1958). Tacit knowledge is referred to as the 'unarticulated' form of knowledge which 'can be communicated only by example, not by percept' (Polanyi, 1958, p. 56). In other words, tacit knowledge is knowledge which

is 'difficult to transmit through speaking and writing' (Chong, 2018b, p. 751-752) but through the illustration of exemplars. In the case of IELTS writing, among the three domains of EJ, the content and organization requirements under dynamic EJ exhibit higher 'tacitness' than the other two. Unlike the linguistic accuracy requirement under hard EJ, which could be explained clearly through grammatical rules, students brainstorm ideas based on their prior experience and thinking habits, and there are different ways to structure an essay. Unlike the linguistic variety requirement under soft EJ, it is more difficult for students to copy ideas and imitate structures than using words and sentence structures which they have read elsewhere. Despite its ambiguity, students in the present study were able to articulate the dynamic domain of EJ of IELTS writing rubrics most frequently, which confirms the findings from previous studies that tacit knowledge is best developed through the use of exemplars (e.g., Smyth & Carless, 2021).

Additionally, the findings suggest that, with only 15 references, the use of exemplars is less effective for students' development of their understanding of hard EJ, which, in this case, is related to using accurate word-level and sentence-level grammatical features in writing. This result may be due to the fact that students need to possess other types of specialist knowledge (i.e. grammatical knowledge) in order to fully understand this requirement related to linguistic correctness. The student participants of the workshops major in different disciplines in the fields of humanities, social sciences, and science, which explained some students' lack of grammatical knowledge of English language. From the findings, the possession of domain-specific knowledge, be it grammatical, morphological, syntactical, or structural, appears to mediate the positive impact of dialogic use of exemplars on students. For example, even though students became more aware of the various soft EJ requirements of the writing rubrics, students were not able to evaluate the exemplars in terms of the formation of new words through suffixes. This situation was caused by the fact that students did not possess basic knowledge about English word formation. Being limited by this lack of domain-specific knowledge, students' awareness of such assessment standard was not translated into critical evaluation of exemplars.

Maximizing the affordances of using exemplars

Alongside the affordances discussed earlier, the workshop participants pointed out a number of limitations of using exemplars to develop students' EJ. In order to overcome these constraints, I would propose (1) an 'eclectic' approach to using exemplars and (2) the incorporation of exemplar use into course design.

Regarding (1), despite their usefulness in developing students' understanding of assessment standards, exemplars should not be used in isolation because the use of exemplars per se cannot develop domain-specific knowledge (or subject matter knowledge (McNamara, 1991)) of students. For instance, in the present study, although students were able to articulate some aspects of hard EJ, which is associated with the accurate use of word-level and sentence-level grammatical features, this assessment standard was mentioned the least frequently by students and only very general descriptions of such standard were given. As explained, this was due to students' lack of domain-specific knowledge (i.e. English grammatical knowledge). It is, therefore,

important for teachers to develop students' domain-specific knowledge alongside their EJ through other pedagogical tasks. In the case of IELTS writing, it is crucial that teachers provide scaffolding to grammatical items which are commonly used in the two IELTS writing tasks (e.g. the use of verbs to describe trends and data for Task 1). Additionally, more traditional grammar and writing practices prior to students' analysis of grammatical features of writing samples may contribute to students' better understanding of this assessment standard. This finding sheds important light to language teachers who would like to incorporate exemplars into their lessons to consider the timing of introducing exemplars – the use of exemplars would be the most conducive to the development of students' understanding of assessment standards after they have a basic level of understanding of domain-specific knowledge included in the assessment rubrics.

Findings from RQ2 suggest that it is important for assessment researchers in higher education and university language teachers to not only approach the use of exemplars as a 'lesson activity' but consider the place of exemplars in curriculum planning and course design (Hawe, Lightfoot, & Dixon, 2019). As illustrated from the findings reported, albeit useful, the use of exemplars was perceived by some students as time-consuming and incomprehensive. Given the constraint that the context of the present study is a one-off, three-hour workshop, it was not possible for the teacher/researcher to introduce a range of exemplars to illustrate different facets of the rubrics and the range of question types of IELTS. When using exemplars in writing courses, it is important for teachers to plan ahead and identify opportunities for introducing exemplars in different lessons. For instance, short exemplars could be introduced near the end of each lesson after a piece of domain-specific knowledge is taught as a consolidation task. In this way, students are continuously exposed to different exemplars which are selected to illustrate different facets of assessment standards and domain-specific knowledge. Additionally, from the perspective of curriculum planning, teachers should refrain from setting assessment questions which are too similar to the topics illustrated by exemplars to avoid limiting students' creativity and thinking when completing the assessment task (Carless et al., 2018).

Despite providing initial empirical evidence to support the use of exemplars, the current study is limited because the findings, which were collected in an nonformal educational context (workshops offered in a self-access language learning centre), may not be generalizable to other more formal educational environments (e.g. formal classroom setting). Moreover, the workshops were one-off, rendering the duration of the study short and longitudinal observations not possible.

Conclusion

The current study presents qualitative evidence of dialogic moves of a teacher concerning exemplar use as well as the affordances and limitations of using exemplars with a group of undergraduate and postgraduate students in the context of IELTS writing workshops at a university in Hong Kong. The findings indicate that dialogic use of exemplars is conducive to developing students' hard, soft, and dynamic EJ because they develop a more sophisticated understanding of the assessment rubrics and are offered the opportunities to engage in self-assessment and peer assessment. Furthermore, this article suggests the adoption of an eclectic approach to using exemplars and

consideration of incorporating exemplars strategically into course design to maximize its affordances. Future research can consider adopting a quasi-experimental design to investigate the effectiveness of using exemplars in improving students' actual performance. Another research direction can be to conduct ethnographic studies focusing on a smaller number of students to examine how students' EJ develop longitudinally with exemplars as the intervention.

Abbreviations
EJ: Evaluative judgement; IELTS: International English Langauge Testing System; LOA: Learning-oriented assessment

Supplementary Information
The online version contains supplementary material available at https://doi.org/10.1186/s40862-021-00115-4.

Additional file 1.

Additional file 2.

Additional file 3.

Availability of data and materials
The datasets generated and analysed during the current study are not publicly available but are available from the corresponding author on reasonable request.

Declarations

Ethics approval and consent to participate
The Human Research Ethics Committee (HREC) at the Education University of Hong Kong granted approval for this study to be conducted on 20 June 2018 (Ref. no. 2017–2018-0411).

Consent for publication
Consent for publishing data collected from participants has been sought when participants signed the informed consent form electronically.

Competing interests
Not applicable.

References
Ajjawi, R., & Boud, D. (2018). Examining the nature and effects of feedback dialogue. *Assessment & Evaluation in Higher Education, 43*(7), 1106–1119. https://doi.org/10.1080/02602938.2018.1434128.

Bloxham, S. (2015). Assessing assessment: New developments in assessment design, feedback practices and marking in higher education. In H. Fry, S. Ketteridge, & S. Marshall (Eds.), *A handbook for teaching and learning in higher education: enhancing academic practice*, (4th ed., pp. 107–122). New York: Routledge.

Carless, D. (2015) Exploring learning-oriented assessment processes. *Higher Education, 69*, 963–976. https://doi.org/10.1007/s10734-014-9816-z

Carless, D. (2007). Learning-oriented assessment: conceptual bases and practical implications. *Innovations in Education and Teaching International, 44*(1), 57–66. https://doi.org/10.1080/14703290601081332.

Carless, D., & Boud, D. (2018). The development of student feedback literacy: enabling uptake of feedback. *Assessment & Evaluation in Higher Education, 43*(8), 1315–1325. https://doi.org/10.1080/02602938.2018.1463354.

Carless, D., & Chan, K. K. H. (2017). Managing dialogic use of exemplars. *Assessment & Evaluation in Higher Education, 42*(6), 930–941. https://doi.org/10.1080/02602938.2016.1211246.

Carless, D., Chan, K. K. H., To, J., Lo, M., & Barrett, E. (2018). Developing students' capacities for evaluative judgement through analysing exemplars. In D. Boud, R. Ajjawi, P. Dawson, & J. Tai (Eds.), *Developing evaluative judgement in higher education: assessment for knowing and producing quality work*, (pp. 108–116). London: Routledge.

Carter, R., Salamonson, Y., Ramjan, L. M., & Halcomb, E. (2018). Students use of exemplars to support academic writing in higher education: an integrative review. *Nurse Education Today, 65*, 87–93. https://doi.org/10.1016/j.nedt.2018.02.038.

Charmaz, K. (2006). *Constructing grounded theory: a practical guide through qualitative analysis*. Thousand Oaks: Sage.

Charmaz, K. (2014). *Constructing grounded theory*. Thousand Oaks: Sage.

Chong, S, W, (2018a). Three paradigms of classroom assessment: implications for written feedback research. *Language Assessment Quarterly, 15*(4), 330-347. https://doi.org/10.1080/15434303.2017.1405423.

Chong, S. W. (2018b). The use of exemplars in English writing classrooms: from theory to practice. *Assessment & Evaluation in Higher Education, 44*(5), 748-763. https://doi.org/10.1080/02602938.2018.1535051.

Chong, S. W. (2018c). Interpersonal aspect of written feedback: a community college students' perspective. *Research in Post-Compulsory Education 23*(4), 499-519. https://doi.org/10.1080/13596748.2018.1526906.

Chong, S. W. (2020). *Meetacognitive mindscapes: Understanding secondary EFL writing students' systems of knowledge*. Routledge.

Chong, S. W., & Ye, X. (2020). *Developing writing skills for IELTS: a research-based approach*. Routledge.

Cowan, J. (2010). Developing the ability for making evaluative judgements. *Teaching in Higher Education*, *15*(3), 323–334. https://doi.org/10.1080/13562510903560036.

Dawson, P. (2015). Assessment rubrics: towards clearer and more replicable design, research and practice. *Assessment & Evaluation in Higher Education*, *42*(3), 347–360. https://doi.org/10.1080/02602938.2015.1111294.

Dawson, P. (2018). Exemplars, feedback and bias: How do computers make evaluative judgements? In D. Boud, R. Ajjawi, P. Dawson, & J. Tai (Eds.), *Developing evaluative judgement in higher education: assessment for knowing and producing quality work*, (pp. 99–107). London: Routledge.

Earl, L. (2013). *Assessment as learning: Using classroom assessment to maximize student learning*, (2nd ed.,). [E-reader version]. Thousand Oaks: Corwin Press.

Goodyear, P., & Markauskaite, L. (2018). Epistemic resourcefulness and the development of evaluative judgement. In D. Boud, R. Ajjawi, P. Dawson, & J. Tai (Eds.), *Developing evaluative judgement in higher education: assessment for knowing and producing quality work*, (pp. 28–38). London: Routledge.

Handley, K., & Williams, L. (2011). From copying to learning: using exemplars to engage students with assessment criteria and feedback. *Assessment & Evaluation in Higher Education*, *36*(1), 95–108. https://doi.org/10.1080/02602930903201669.

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research, 77*(1), 81–112. https://doi.org/103102/003465430298487.

Hawe, E., Lightfoot, U., & Dixon, H. (2019). First-year students working with exemplars: promoting self-efficacy, self-monitoring and self-regulation. *Journal of Further and Higher Education*, *43*(1), 30–44. https://doi.org/10.1080/0309877X.2017.1349894.

Hendry, G. D., & Jukic, K. (2014). Learning about the quality of work that teachers expect: students' perceptions of exemplar marking versus teacher explanation. *Journal of University Teaching & Learning Practice*, *11*(2), 5.

Inaba, M. (2018). *Second language literacy practices and language learning outside the classroom*. Bristol: Multilingual Matters.

Lambert, M. (2019). *Practical research methods in education*. New York: Routledge.

Lee, I. (2014). Revisiting teacher feedback in EFL writing from sociocultural perspectives. *TESOL Quarterly*, *48*(1), 201–213. https://doi.org/10.1002/tesq.153.

Lee, I. (2017). *Classroom writing assessment and feedback in L2 school contexts*. Singapore: Springer.

Liu, N.-F., & Carless, D. (2006). Peer feedback: the learning element of peer assessment. *Teaching in Higher Education*, *11*(3), 279–290. https://doi.org/10.1080/13562510600680582.

McNamara, D. (1991). Subject knowledge and its application: problems and possibilities for teacher educators. *Journal of Education for Teaching*, *17*(2), 113–128. https://doi.org/10.1080/0260747910170201.

Nelson, R. (2018). Barriers to the cultivation of evaluative judgement: a critical and historical perspective. In D. Boud, R. Ajjawi, P. Dawson, & J. Tai (Eds.), *Developing evaluative judgement in higher education: assessment for knowing and producing quality work*, (pp. 51–59). London: Routledge.

O'Donovan, B., Price, M., & Rust, C. (2008). Developing student understanding of assessment standards: a nested hierarchy of approaches. *Teaching in Higher Education*, *13*(2), 205–217. https://doi.org/10.1080/13562510801923344.

Orsmond, P., Merry, S., & Reiling, K. (2002). The use of exemplars and formative feedback when using student derived marking criteria in peer and self- assessment. *Assessment & Evaluation in Higher Education*, *27*(4), 309–323. https://doi.org/10.1080/0260293022000001337.

Panadero, E., Alonso-Tapia, J., & Reche, E. (2013). Rubrics vs. self-assessment scripts effect on self-regulation, performance and self-efficacy in pre-service teachers. *Studies in Educational Evaluation*, *39*(3), 125–132. https://doi.org/10.1016/j.stueduc.2013.04.001.

Polanyi, M. (1958). *Personal knowledge: Towards a post-critical philosophy*. London: Routledge.

Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional Science, 18*(2), 119–144. https://doi.org/10.1007/BF00117714.

Saldaña, J. (2016). *The coding manual for qualitative researchers*, (3rd ed.,). Thousand Oaks: Sage.

Sambell, K., McDowell, L., & Montgomery, C. (2013). *Assessment for learning in higher education*. New York: Routledge.

Savin-Baden, M., & Major, C. H. (2012). *Qualitative research: the essential guide to theory and practice*, (1st ed.,). New York: Routledge.

Smyth, P., & Carless, D. (2021). Theorising how teachers manage the use of exemplars: Towards mediated learning from exemplars. *Assessment & Evaluation in Higher Education*, *46*(3), 393–406. https://doi.org/10.1080/02602938.2020.1781785.

Tai, J., Ajjawi, R., Boud, D., Dawson, P., & Panadero, E. (2018). Developing evaluative judgement: enabling students to make decisions about the quality of work. *Higher Education*, *76*(3), 467–481. https://doi.org/10.1007/s10734-017-0220-3.

Thompson, D. G., & Lawson, R. (2018). Strategies for fostering the development of evaluative judgement. In D. Boud, R. Ajjawi, P. Dawson, & J. Tai (Eds.), *Developing evaluative judgement in higher education: Assessment for knowing and producing quality work*, (pp. 136–144). London: Routledge.

To, J., & Carless, D. (2016). Making productive use of exemplars: peer discussion and teacher guidance for positive transfer of strategies. *Journal of Further and Higher Education*, *40*(6), 746–764. https://doi.org/10.1080/0309877X.2015.1014317.

Vygotsky, L. S. (1987). *Thinking and speech*. New York: Plenum.

Publisher's Note