



Chinese EFL Learners' Decision-Making while Evaluating Peers' Texts

JINGJING MA

*University of Hong Kong/Fudan University**

Received: 3 July 2010 / Accepted: 29 September 2010

ABSTRACT

Approaching peer review from a process and contextualized perspective, this exploratory case study investigates two Chinese EFL learners' decision-making patterns while evaluating peers' texts in an online peer review and factors influencing these patterns. Detailed qualitative case study data were collected through think-aloud protocols, stimulated recall, semi-structured interviews, classroom observation and document analysis. Analyses indicate that the two learners with higher level of English writing proficiency to a certain extent illustrated contrasting patterns of decision-making, and yet both prioritized specific aspects of peers' texts. Student-related factors such as perceptions of good English expository writing shaped by previous learning and assessment experiences of English (or Chinese) writing, type of writing task and weaknesses of student text interacted with one another to influence the participants' decision-making patterns. Pedagogical implications for the findings are discussed.

KEYWORDS:

Chinese EFL students, peer review, decision-making patterns, perceptions of English writing

RESUMEN

Con un acercamiento a la revisión por pares (*peer review*) en términos de proceso y desde una perspectiva contextualizada, este estudio de caso investiga (i) los procesos de toma de decisiones de dos estudiantes Chinos de inglés como lengua extranjera (EFL) al evaluar los textos producidos por sus compañeros, y (ii) los factores que influyeron en la activación de dichos procesos. Los datos cualitativos del estudio se recogieron mediante protocolos de pensamiento en voz alta, entrevistas de recuerdo estimulado, entrevistas semi-estructuradas, observación en el aula y análisis de textos. Los análisis llevados a cabo indican que los dos aprendices con un nivel alto de inglés escrito presentaban patrones parcialmente distintos en sus procesos de toma de decisiones, y que priorizaban aspectos concretos de los textos de sus compañeros. A su vez, factores tales como (i) las percepciones de los dos estudiantes sobre las características de un buen texto expositivo en inglés (derivadas, a su vez, de sus experiencias previas de aprendizaje y evaluación de la escritura en inglés y en chino), (ii) el tipo de tarea de escritura, y (iii) los puntos débiles de los textos evaluados, interactuaban entre sí, influyendo en sus procesos de toma de decisiones. Se discuten las implicaciones pedagógicas derivadas de estos resultados.

PALABRAS CLAVE:

estudiantes chinos de inglés como lengua extranjera (ILE), revisión por pares, patrones de toma de decisiones, percepciones de la escritura en inglés

**Address for correspondence:* Jingjing Ma., Faculty of Education, University of Hong Kong, Rm. 112, HOC Building, Pokfulam Road, Hong Kong, Tel.: 00852-22415467. E-mail: majingjing79@163.com

1. INTRODUCTION

In recent years peer response has become an important component of both L1 and L2 writing class due to its cognitive, affective, social and methodological benefits (Rollinson, 2005). Peer review in particular, in which students need to give written peer comments, has gained popularity in EFL contexts, given its potential to mitigate time constraint many EFL writing instructors face and to enable student reviewers to think deeply about how to phrase their comments and student writers to revisit written comments as many times as they want for revision (Min, 2005, Rollinson, 2005). The development of technology has made online peer review possible and this innovative mode has been found to further facilitate the peer review process (e.g. Guardado & Shi, 2007).

Particularly concerned with effectiveness of students as responders (Rinnert & Kobayashi, 2001), researchers have explored the characteristics of peer comments in written or online peer review to understand whether peer feedback thus provided was beneficial for student revision (e.g. Caulk, 1994). Factors likely to affect these characteristics have also been proposed (e.g. Liu & Hansen, 2002). However, little attention has been paid to the cognitive processes of peer review and factors affecting them, information that L2 writing instructors needed to prepare their students to become more effective peer reviewers.

This paper reports an exploratory case study conducted in the context of an English expository writing course for non English majors at a university in Mainland China. It aims to investigate Chinese EFL students' cognitive processes in evaluating peers' texts in an online peer review and factors that influenced such processes.

Given the popularity of written or online peer review in EFL contexts, many studies have investigated the characteristics of peer comments, reflecting a product view of peer review. Accordingly, textual analysis has been employed to investigate stance (e.g. Min, 2008), focus (e.g. Caulk, 1994) and quality (e.g. Hu & Lam, 2009) of written peer review, and discourse pattern (e.g. Guardado & Shi, 2007), type, area and nature (e.g. Liu & Sadler, 2003) of online peer review. Findings of these studies have further shed light on training L2 learners to provide quality comments beneficial for student revision (e.g. Min, 2005). However, much still remains to be known about L2 learners' cognitive processes in responding to peers' texts (Hu & Lam, 2009). Such information is important because it serves as direct evidence of (in)effectiveness of students as responders and may further guide peer review training. A process perspective of peer review has thus been suggested by Hu and Lam (2009) to explore the peer review processes that would otherwise remain hidden, particularly with the help of think-aloud protocols.

Previous studies on EMT (English-mother-tongue) and ESL/EFL raters' rating process (e.g. Cumming, Kantor, & Powers, 2002; Vaughan, 1991) have laid the groundwork for an analysis of peer review from a process perspective. For example, Cumming et al. (2002) identified the emergence of various decision-making behaviors with different foci when

experienced raters assessed ESL/EFL writing. Furthermore, individual raters possessed different patterns of rating by focusing more on particular aspect (s) of student writing (Cumming et al., 2002; Vaughan, 1991). Though the cognitive processes of assessing ESL/EFL writing to provide a grade are different from those of responding to peers' texts to provide comments, the afore-mentioned studies become particularly relevant to an investigation of L2 students' decision-making processes in a peer review that requires them to evaluate peers' texts based on course assessment criteria (e.g. Ferris, 2001).

On the other hand, to ensure the effectiveness of students as peer reviewers, researchers have proposed various factors likely to affect L2 learners' feedback giving. These factors are often examined separately in a decontextualized manner. For example, since L2 learners tended to focus on surface errors of student texts (Leki, 1990) instead of global issues (i.e. content and organization) potentially beneficial for student revision (Liu & Hansen, 2002), possible factors to impact on the focus of peer comments were discussed one by one in the hope of enabling writing instructors to guide learners to focus on global issues while commenting (e.g. Liu & Hansen, 2002). However, since peer review does not take place in a vacuum, more studies are needed to explore factors that bear on peer feedback giving in different contexts through a contextualized approach, especially various factors to influence L2 learners' cognitive processes of peer review and possible interaction among them.

The afore-mentioned possible factors to influence the focus of peer comments are likely to be important for the current study. They include student-related factors, such as L2 student reviewers' perceptions of English writing shaped by their experience with English writing instruction (Mangelsdorf & Schlumberger, 1992); the type of writing task (Liu & Hansen, 2002) and strengths and weaknesses of student writing (Liu & Hansen, 2002).

For example, influenced by the emphasis of (previous) English writing instruction, L2 learners might perceive correctness in writing as being more important than expression of meaning, and their comments would focus more on form than content (Mangelsdorf & Schlumberger, 1992). Research shows that both previous L1 and L2 writing instruction may actually affect students' perceptions of L2 writing (Rinnert & Kobayashi, 2009), which may impact on the focus of peer comments. The influence of previous L1 and L2 writing instruction can be explained by a socio-cognitive view of writing which acknowledges that writers' knowledge and practice of writing (seen as a primarily mental activity) is shaped by the social contexts, one of the most important being educational context (Roca & Murphy, 2001). Moreover, research on L1 writing found that L1 students' perceptions of good English writing differed significantly and consistently from those of instructors (Newkirk, 1984).

The type of writing task or weaknesses of student writing may also cause student reviewers to focus on different aspects of their peers' writing (Liu & Hansen, 2002). For instance, they may find it easier to concentrate on content and organization while responding to narration compared with persuasion, a different rhetorical mode. The strengths and weaknesses of

student texts are likely to direct student reviewers' attention to particular aspects of peers' writing as well.

2. THIS STUDY

2.1. Research questions

Based on the research gap identified from a review of the literature, this paper adopts a process and contextualized perspective of peer review to pursue the following research questions:

RQ1. How did the Chinese EFL students in the current study evaluate peers' writing in an online peer review, as reflected by their decision-making processes, especially decision-making patterns?

RQ2. What factors seemed to influence their decision-making patterns?

For the purpose of the paper, decision-making patterns are defined as student reviewers' tendency to decide to focus on particular aspects of student writing.

2.2. Research methodology

Since very few studies have examined L2 learners' decision-making processes while evaluating student writing and factors affecting those processes, exploratory case studies of a group of Chinese EFL students were carried out to gain an in-depth understanding. 6 students were purposefully selected from 18 students in the only English expository writing class that incorporated online peer review in a prestigious university in Mainland China. The selection was based on the following criteria to offer the "richest possible data" (Lofland & Lofland, 1995: 16): different levels of English writing proficiency, different focus and amount of peer comments provided and enthusiasm about and willingness to participate in the study. English writing proficiency was included as one sampling criterion to explore its possible influence on participants' decision-making patterns.

It turned out that rich think-aloud data were elicited from two of the students who enjoyed higher English writing proficiency compared with the other four but who were equally enthusiastic about participating in the study. One of them provided a large number of peer comments focusing on language while the other tended to give relatively fewer comments with a focus on content and organization. As a result, the findings about these two are presented here. Such a focus also allows for an in-depth analysis of their decision-making patterns, which to a certain extent illustrated contrasting patterns and yet shared certain similarities.

2.3. Research context

The investigation took place in the context of a 16-week credit-bearing elective course of English expository writing. Online peer review was implemented as an integral part of a process-approach to writing adopted in the course. At the beginning of the course the teacher briefly introduced the concepts in the peer feedback form (Appendix 1), a simplified version of the course assessment criteria, and reminded students about employing the form to attend to different aspects of peers' writing. The students were then assigned to read their classmates' first drafts downloaded from the Internet and to insert comments. Then they had to revise their first drafts based on peer comments received. The second drafts were for teacher feedback and evaluation. The students might choose to revise the second drafts further based on teacher feedback, but this was not compulsory. After the first round of peer review, the teacher also referred the students to the quality peer comments provided by one student in the class. From the second peer review task on, he did not seem to monitor the peer review process any more and the students were left to themselves to carry out the peer review tasks.

During the 16 weeks, a total of four writing tasks were assigned including two illustration essays, one cause-and-effect essay and one comparison-and-contrast essay. However, only the second and third writing tasks (i.e. the second illustration essay and cause-and-effect essay) were chosen in this research based on my observation in the pilot study that these two tasks could generate the richest information in relation to the research questions pursued. The illustration essay required students to use specific examples to illustrate that modern people are more gullible and superstitious than people were in the Middle Ages while the cause-and-effect essay required them to explore the causes of why no single Chinese has won any Nobel Prizes. Each essay should not exceed 400 words and student writing usually ranged from 350 to 400 words.

2.4. Participants

Background information on the two participants is summarized in Table 1. Both Hyshan and Fiona (their pseudonyms) had studied English for 13 years starting from primary school, mainly through formal English education in mainland China. However, they had not been formally taught English expository writing before entering the university.

Student	Age	Gender	Year of study	Major	Writing proficiency	TOEFL writing score
Hyshan	20	Female	3 rd year	Biology	High	30 out of 30 taken in high school
Fiona	20	Female	3 rd year	Journalism	High	21 out of 30 taken in year 1

Table 1. Background information on case study students

Prior to entering the English expository writing course in their university study, Hyshan had experiences with writing courses like Intermediate English Writing while Fiona had mainly taken English speaking courses. Both students had the same plan of studying abroad after graduation.

2.5. Data collection

Multiple sources of data were collected through think-aloud protocols, stimulated recall, semi-structured interviews, classroom observation and document analysis. To capture their decision-making patterns, the two participants were invited after a training session to think aloud while evaluating respectively the illustration and cause-and-effect essays written by their peers and displayed on a computer screen. They used Mandarin, the native language shared by them and the researcher for think-aloud to avoid difficulties in verbalizing their thinking (Manchon, Murphy, & Roca, 2005). To counter reactivity issues, the instruction given to them was carefully worded to guide the participants to verbalize everything that came to their mind rather than give explanation or justification of what they were doing while thinking aloud (Ericsson & Simon, 1993). In total, 5.5 hours of think-aloud protocols were recorded with every operation on the computer screen captured during the think-aloud sessions.

For the sake of triangulation, the two participants were also asked to attend a stimulated recall session immediately after the think-aloud whenever possible or within one week of think-aloud sessions, with the peer comments provided on the computer screen serving as stimulus (Gass & Mackey, 2000). To ensure its validity, the instruction of the stimulated recall emphasized that the participants should report what they were thinking at the time of evaluating student texts instead of at the time of the recall. The stimulated recall was conducted in Mandarin and a total of 1.5 hours of data were audio-taped.

To probe into factors likely to influence the participants' decision-making patterns, a series of semi-structured interviews, ranging from 20 minutes to 1 hour, were conducted with the two students in Mandarin to elicit information concerning previous learning and assessment experiences of English and Chinese writing and perceptions of it, understanding of the concepts in the peer feedback form, as well as understanding of the requirement of the illustration and cause-and-effect essays. Follow-up interviews were held to probe into issues arising from ongoing data analysis. 6 hours of interview data were audio-taped.

To gain an understanding of the classroom context which may affect the participants' decision-making patterns, systematic observation of the writing class was carried out. Field notes were taken during observation and were later developed into full notes. For the sake of triangulation, documents were also collected including the course syllabus, student texts both participants evaluated and accompanying peer comments.

2.6. Data analysis

All data were transcribed in full and translated into English. After being compared with the stimulated recall data for triangulation, the think-aloud data were then parsed into meaningful units, that is, “a section of text that states a single idea” (Smagorinsky, 1991: 346), be it a phrase, a clause or a sentence. The coding scheme developed by Cumming et al. (2002) was revised in a recursive manner after its repeated application to the data at hand. Notably, while their coding scheme was originally intended to categorize raters' decision-making behaviors in assessing ESL/EFL compositions, it was found to be highly relevant to this research context.

The original scheme encompassed both interpretation and evaluation behaviors with different foci: self-monitoring, ideational and rhetorical, and linguistic. Due to the focus of the current study, the participants' interpretation behaviors were not included for analysis. Furthermore, for convenience of exposition, the original category of rhetorical and ideational focus was divided into organization and content while the other category of self-monitoring focus was not included. The distinction between content and organization was sometimes difficult, so the current study followed Rinnert and Kobayashi (2001) 's distinction in their study of Japanese EFL students' perceptions of English compositions, that is, “quality and development of ideas belonged to content, whereas ordering and structuring those ideas belonged to organization” (p.204). Following this procedure, the participants' evaluation behaviors were identified, coded into different types and grouped under the foci of language, content or organization (Appendix 2).

To explore the two participants' decision-making patterns, the percentage of decision-making behaviors with the same focus to the total was calculated, followed by a calculation of the percentage of each type to the total. A quantification of data is necessary not only for the purpose of data reduction, but also pattern generation (Miles & Huberman, 1984).

By an iterative process of reviewing the rest of the qualitative data including interview data, field notes and documents, categories were developed and patterns and relations among categories were explored. Emerging themes about factors affecting decision-making patterns were thus identified.

Measures taken to ensure the trustworthiness of the case study included member checks and triangulation. Member checks (Lincoln & Guba, 1985) were carried out by presenting the participants with my interpretations and inviting their opinions 4 months after data collection. Triangulation of methods as mentioned above was also employed to strengthen the credibility of the case study. Though a small-scale exploratory study like this is prone to criticisms of lack of generalizability, the readers can still make naturalistic generalization (Stake, 1995) depending on a comparison of the “sending and receiving contexts” (Lincoln & Guba, 1985: 316).

3. FINDINGS

3.1. Decision-making patterns in evaluating peers' texts

The findings concerning the participants' decision-making patterns were mainly derived from think-aloud and stimulated recall. Altogether 32 distinct decision-making behaviors in evaluating peers' texts (Appendix 2) were identified from the two participants' think-aloud protocols and stimulated recall across tasks. Table 2 summarizes respectively the percentage of Hyshan's and Fiona's decision-making behaviors by foci across tasks. Hyshan's general decision-making pattern across tasks reflected a predominant focus on language, as evidenced by 64.9% and 91.7% of language-related decision-making behaviors for the illustration and cause-and-effect essay respectively. On the other hand, Fiona displayed a contrasting pattern of decision-making for the illustration essay compared with Hyshan, since a total of 52.8% of her decision-making was concerned with organization and content, but she seemed to be far more concerned with language for the cause-and-effect essay, as demonstrated by a surprisingly high percentage of language-related decision-making behaviors (73.5%).

Focus	Illustration (Hyshan)	Cause and effect (Hyshan)	Illustration (Fiona)	Cause and effect (Fiona)
Language	64.9%	91.7%	47.2%	73.5%
Organization	15.4%	8.3%	17.6%	10.3%
Content	19.7%	0	35.2%	16.2%

Table 2. Percentage of Hyshan's and Fiona's decision-making behaviours by foci across tasks

Table 3 and 4 respectively present Hyshan's and Fiona's top two most evoked types of decision-making behaviors in language, content and organization across tasks. (Considering conciseness was included in Table 3 because it was consistently the third most enacted decision-making behavior within language across tasks.) In this way the participants' decision-making patterns within each category can be obtained. A compromise was also made between the terms in the coding scheme and those used by the participants to refer to aspects of English expository writing. For example, Hyshan and Fiona used different terms (e.g. a "clear line of thinking" in Table 3 and "logic" in Table 4) to refer to the same concept of logical organization, which according to them was represented by the existence of clear thesis statement and topic sentences, link between the two and paragraph unity. Consequently, decision-making behaviors related to the above-mentioned elements were presented separately first (in a smaller font) and then grouped under a "clear line of thinking" or "logic" respectively.

	Illustration essay	Cause-and-effect essay
Language	· consider word and phrase use : 24.7% · evaluate grammar : 18.6% · consider conciseness : 4.1%	· consider word and phrase use : 37.9% · evaluate grammar : 27.7% · consider conciseness : 8.3%
Organization	· assess connection between sentences : 8.2% ·consider connection between body and conclusion: 3.1%	· assess connection between sentences : 2.8% · consider a “clear line of thinking” in Hyshan’s terms : 2.8% -assess connection between topic sentence and thesis statement: 1.9% -evaluate topic sentence: 0.9%
Content	·assess clarity: 6.2% · evaluate appropriateness of examples : 6.2%	·no decision-making for content

Table 3. Hyshan's top two most enacted decision-making behaviors across tasks

	Illustration essay	Cause-and-effect essay
Language	· consider word and phrase use : 24.1%	· evaluate grammar : 50
	· evaluate grammar : 7.4%	· consider word or phrase use : 16.1%
Organization	· evaluate “logic” in Fiona’s terms : 9.4% -evaluate topic sentence: 3.7% -evaluate thesis statement: 1.9% -assess connection between topic sentence and thesis statement: 1.9% -evaluate unity: 1.9%	· evaluate “logic” in Fiona’s terms : 7.4% -assess connection between topic sentence and thesis statement: 7.4%
	·consider support for main points: 2.8%	·consider overall organization: 2.9%
Content	· evaluate appropriateness of examples : 13.9%	· evaluate idea : 5.9%
	· evaluate idea : 4.6%	· evaluate appropriateness of examples : 4.4%
	· <u>consider detailed example</u> : 4.6%	·evaluate appropriateness of supporting evidence (other than examples): 4.4%

Table 4. Fiona's top two most enacted decision-making behaviors across tasks

Items in bold: commonly prioritized aspects of student texts across tasks

Underlined items: examples of prioritized aspects of student texts for different tasks

It is not difficult to discern from Tables 3 and 4 that each participant's most enacted (and sometimes top two or three most enacted) decision-making behaviors in language, organization (and content in Fiona's case) were identical across tasks, indicating a tendency to prioritize specific aspects of peers' texts in these categories (items in bold) regardless of the writing task

type. For example, Hyshan consistently paid great attention to vocabulary, grammar and conciseness within language, and connection between sentences within organization across tasks. Fiona consistently attached great importance to vocabulary and grammar within language, “logic” or logical organization within organization and appropriateness of examples and ideas within content across tasks. Since Hyshan did not evaluate the content of the cause-and-effect essay, she attended most to clarity of content and then appropriateness of examples only for the illustration essay.

Noticeably, both participants were concerned with vocabulary and grammar across tasks. For example, Hyshan consistently prioritized vocabulary and then grammar each time (Table 3), leading to her predominant emphasis on language for both essays. The importance accorded to each by Fiona, however, was not consistent across tasks (Table 4), with a far greater concern for grammar (50%) for the cause-and-effect essay causing her general decision-making pattern to change from being content-and-organization-oriented to language-focused (Table 2).

After consistently prioritizing specific aspects of student texts across tasks, both participants also attended to different elements of writing for different tasks. For example, within organization what Hyshan considered after connection between sentences was connection between body and conclusion for the illustration essay and a “clear line of thinking” for the cause-and-effect essay. Within content what Fiona evaluated next after appropriateness of examples and idea was detailed example for the illustration essay and appropriateness of supporting evidence (other than examples) for the cause-and-effect essay.

3.2. Factors impacting on decision-making patterns

The findings concerning factors influencing decision-making patterns were mainly derived from interview data and document analysis. Student-related factors such as perceptions of good English expository writing largely shaped by previous learning and assessment experiences of English (or Chinese) writing and English extra-curricular activities (in Fiona’s case), task type (e.g. illustration vs cause-and-effect) and weaknesses of student texts were found to mutually influence the participants’ decision-making patterns in this study. Given its key role emerging from data analysis, more emphasis is placed on the student-related factor.

3.2.1. Student-related factors

Both participants had their own idea about what was good English expository writing, represented by the general criteria for it and specific criteria triggered in response to particular type of essay they were evaluating. Such perceptions in turn influenced their decision-making patterns.

Hyshan listed clarity of content, concise sentences and relevance to topic as important qualities of good English expository writing. To achieve clarity of content, connection between

sentences was also crucial. Most importantly, she regarded vocabulary and grammar accuracy as the basic quality.

...If problems concerning tense, sentence structure or inaccurate use of vocabulary appear in writing, these should be attended to first... As an English learner myself, I think the top priority is that there should not exist any wrong usage of words or sentence errors.

In addition, Hyshan especially articulated the specific criteria of appropriate examples and a “clear line of thinking” for the illustration and cause-and-effect essay respectively.

It can be found that all components of Hyshan's general criteria except relevance to topic (i.e. vocabulary and grammar accuracy, concise sentences, clarity of content and connection between sentences) coincided with the most frequently evaluated aspects of student writing in her decision-making across tasks (items in bold in Table 3) and the most evaluated aspect in content for the illustration essay (i.e. clarity), suggesting a strong influence of the former on the latter. In particular, Hyshan's greatest attention paid to vocabulary and grammar accuracy cross tasks, which contributed to her general decision-making pattern for both essays, indicated the great influence of the perceived importance of the basic quality of good English writing.

Furthermore, the specific criteria of appropriate examples and a “clear line of thinking” developed respectively for the illustration and cause-and-effect essay also seemed to exert an influence on Hyshan's decision-making patterns in that these two aspects attracted great attention from her to become the most evaluated areas in content and organization respectively for the two essays (underlined items in Table 3), along with the commonly prioritized elements in these two categories.

Fiona, on the other hand, seemed to have somewhat different perceptions of good English expository writing from those of Hyshan. According to her, it should exhibit the following qualities: “logic” or logical organization, well-organized overall structure, good development of content (i.e. use of appropriate examples and different forms of support), use of accurate but not necessarily big vocabulary and no grammar mistakes. Different from Hyshan, she explicitly mentioned that organization and content were the two aspects she would focus on in providing peer comments. Moreover, Fiona rendered use of detailed examples especially important for the illustration essay but did not raise any other specific criteria for the cause-and-effect essay.

Similarly, Fiona's general criteria seemed to impact on her decision-making patterns, as evidenced by the correspondence between particular elements of her general criteria (i.e. vocabulary and grammar accuracy, “logic” in Fiona's terms and use of appropriate examples) and the aspects she most attended to in language, organization and content across tasks (items in bold in Table 4, with evaluation of idea in content an exception). The perceived importance of organization and content could further explain Fiona's general decision-making pattern for the illustration essay, though it could not account for her general decision making for the cause-and-effect essay, a point to be discussed in 3.2.3. This orientation also seemed to give rise to Fiona's consistent prioritization of particular aspects of content and organization across

tasks. The specific criteria of detailed examples generated for the illustration essay also played a role in impacting on Fiona's decision-making pattern, since detailed example turned out to be the second most evaluated aspect in content (along with idea) (underlined item in Table 4) after the commonly prioritized aspect of appropriateness of examples.

Notably, both participants' general criteria were to a certain extent similar to the course assessment criteria reflected by the peer feedback form (Appendix 1). For example, Hyshan's general criteria matched several (but not all) components of the course assessment criteria, including relevance and clarity in "Content 1", coherence in "Organization 2", and all the elements in "Language 3" while Fiona's general criteria matched development in "Content 2", thesis statement and topic sentence in "Organization 1", unity in "Organization 2" and all the elements in "Language 3". However, the course assessment criteria seemed to be of a broader range because it encompassed more elements of writing than the participants' general criteria. Moreover, although both participants' language-related criteria corresponded with the elements in "Language 3", the weight given to them differed, as evidenced by the perceived priority of grammar and vocabulary accuracy on Hyshan's part and the less importance attached to them by Fiona.

Despite the similarity between the participants' general criteria and the course assessment criteria, the former was not influenced by the latter, because neither did the participants frequently refer to the course assessment criteria (in the form of peer feedback form) while evaluating peers' writing, nor did they specifically mention it as an important source of influence. As a matter of fact, their general and specific criteria for English expository writing were found to be mainly shaped by the previous experiences of learning English (or Chinese) writing and taking high-stakes English test such as TOEFL. In Fiona's case, her past English extra-curricular activities also played a role. Only influences regarded as being identifiable by the participants were reported here.

When asked about the instructional emphasis of English writing in high school, Hyshan still remembered clearly that her English teacher emphasized grammatical correctness as one of the most important qualities to avoid losing marks in National Matriculation English Test. In response to the same question, Fiona also recalled:

Actually the only requirement for English writing in high school is no grammar mistakes...Grammar mistakes are fatal mistakes that can lead to a lowering of grade [in the National Matriculation English Test].

It seems that both Hyshan and Fiona learned the importance of avoiding "fatal mistakes" in English writing from their teachers' emphasis so that they could achieve good exam results. It was likely that they in turn included "no grammar mistakes" as one component of good English writing that they could employ to evaluate peers' writing.

Hyshan had already taken TOEFL test in high school and she stated the following as things she had learned in preparation for TOEFL writing:

From my experience of preparing for the TOEFL writing test, I find making an outline effective...to speed up my writing. What's more, doing so can develop a clear line of thinking.

One of the TOEFL reference books underscores coherence between sentences, saying that it decides your score.

She further elaborated on the function of the connection between sentences:

If the rater of TOEFL writing found that the coherence between sentences were poor..., you were likely to get a low score. As a reader, the rater would also find your writing difficult to understand.

It seems that Hyshan learned the importance of a clear line of thinking (by means of outline-making) and connection between sentences from her preparation for the TOEFL writing test, with the latter being regarded as one way to achieve clarity in content. Judging from the fact that she obtained a full mark in TOEFL writing, such a successful experience probably reinforced the new knowledge about good English writing gained from test preparation. Part of this new knowledge seemed to be incorporated into her specific criteria for the cause-and-effect essay (i.e. a clear line of thinking) and part into her general criteria (i.e. connection between sentences as one strategy to achieve clarity of content). In other words, Hyshan's experiences of TOEFL writing helped to shape particular components of her general and specific criteria.

Whereas Hyshan's successful experience with TOEFL writing test consolidated her knowledge about English writing gained from test preparation, Fiona learned a lesson about the importance of logical organization from her poor performance in the test taken in year 1:

I did not take logical organization so seriously when I took the TOEFL writing test. Influenced by the Chinese way of writing, I thought it was troublesome to pay attention to it, and I could produce an interesting piece even without listing my first, second and third point. Then I didn't do a good job.

Fiona's unsuccessful experience with TOEFL writing test probably made her realize that the difference between good Chinese and English writing in part resides in the way it is organized and caused her to incorporate logical organization into her general criteria for good English expository writing.

In the interview Hyshan also talked about how she learned the importance of providing appropriate examples in English expository writing from the emphasis of Chinese writing instruction in her secondary school:

My Chinese teacher taught us that we should use appropriate examples in essay writing, so I remembered this principle ever since. I think this can be applied to both Chinese and English expository writing.

It appears that Hyshan transferred what she learned from her Chinese teacher about good L1 writing to L2 writing, and this source of influence contributed to one element of her specific criteria for the illustration essay –use of appropriate examples.

Regarding the use of appropriate examples, Fiona acknowledged that she learned its importance from her experience of participating in English debating contests:

My idea of the need to provide appropriate examples in English writing might stem from a debater's awareness that seemingly high-sounding examples may only help to use up the time or catch attention, but they won't be very effective unless they can support my own argument.

Characterizing herself as an “English speaking contest maniac”, Fiona took part in English debating contests every semester for the first two years of university study. Such an active participation might reinforce the significance of appropriate examples in English debating and make it possible for her to transfer such knowledge from good English speaking to good English writing, thus helping to shape one component of her general criteria for English writing.

3.2.2. Task type

Since in most of the cases the participants generated specific criteria for the illustration and cause-and-effect essay, and such specific criteria did impact on their decision-making (3.2.1), it is worth noting what led them to do so. Hyshan explicitly mentioned the impact of task type on what kind of specific criteria she would develop:

I think each time the task type will exert an influence. For example, if this task [of cause-and-effect essay] requires use of examples [like the last time], this writing may not be considered good. But this time it requires analysis, so I think the writing will be good as long as it is convincing.

Though Fiona did not explicitly articulate the impact of task type, in response to the illustration essay, she also developed specific criteria that influenced her decision-making.

Whereas the participants' specific criteria seemed to be shaped by their past learning and assessment experiences of English (or Chinese) writing, it was the specific task type (e.g. illustration *vs* cause and effect) that triggered them. In other words, it was the interaction between the two afore-mentioned factors that affected the participants' frequent decision-making on specific aspects of writing required by the specific criteria.

3.2.3. Weaknesses of student texts

Weaknesses of student writing were another factor to consider. The following is the introduction of the cause-and-effect essay Fiona evaluated, with the underlined part indicating her decision-making concerning grammar mistakes:

It is universally acknowledged that knowledge is power, and the number of Nobel Prize winner becomes the measurement of a country's science and technology strength. Since 1949, 16 Japanese scientist have won Nobel Prize, and it is more and more frequently asked that why hasn't a single native Chinese scientists won Nobel Prize. China is an ancient country experienced 5000 years history and is the cradle of the Four Inventions, being excluded from the Nobel Prize family is not only a pity but a shame. There are three reason can account for the phenomenon.

Within the short 94-word introduction, Fiona evoked grammar-related decision-making 6 times. Upon close examination, the rest of the student text was also filled with grammatical mistakes (See Appendix 3 for the second paragraph from the same student text).

Since a far greater amount of attention was paid to grammar for the cause-and-effect essay compared with the illustration essay (i.e. 50% vs 7.4%), this could not simply be explained by Fiona's general requirement of grammatical accuracy. As shown by the introduction of student text above, the requirement of grammatical accuracy seemed to interact with the grammatical errors of student text to jointly influence her grammar-related decision-making for the cause-and-effect essay, so much so that her general orientation to content and organization for the illustration essay was reversed.

It was noticeable that the classroom context or current English instruction did not seem to exert much influence in the current study. This might result from the similarity between both participants' perceptions of good English expository writing and those of their instructor's due to their high level of English writing proficiency. More importantly, it seems that their current English instructor did not make much effort to ensure the students' deep understanding of all the concepts on the course assessment criteria, making it less likely for the participants' perceptions to be reshaped to affect their decision-making. For example, he did not frequently explain the meaning of the concepts in the peer feedback form except for a brief introduction of them in class at the very beginning. Consequently, neither participant acknowledged that they understood all the concepts in the form and they tended to rely on their original perceptions of good English writing for decision-making.

4. DISCUSSION

This paper sought to explore two Chinese EFL students' decision-making patterns in evaluating student texts based on course assessment criteria and the factors affecting such patterns. Regarding the students' decision-making patterns, Hyshan's general decision-making pattern

demonstrated a heavy focus on language across tasks while Fiona's displayed an orientation to content and organization for the illustration essay, but that orientation was reversed for the cause-and-effect essay due to her great concern with grammar. Furthermore, both participants tended to consistently prioritize specific aspects of language, organization (and content in Fiona's case) in their decision-making for both essays.

With reference to the factors influencing students' decision-making patterns, it was found that student-related factors, task type and weaknesses of student text might interact with one another to exert a mutual influence. Of particular importance was the role of the student-related factor, namely, students' perceptions of good English expository writing largely shaped by previous learning and assessment experiences of English (or Chinese) writing and past English extra-curricular activities. As mentioned earlier, the teacher's failure to enhance his students' understanding of the course assessment criteria might be an important explanation for the limited influence of the classroom context, an issue discussed later in relation to pedagogical implications.

Similar to the raters in Vaughan's (1991) studies, the participants in the current study seemed to possess individual patterns of decision-making that predisposed them to focus on particular elements of student writing regardless of task type. However, compared with Hyshan, Fiona's general decision-making pattern for the illustration essay seemed to be more likely to give rise to content-and-organization-related peer comments, regarded as being more beneficial for student revision (Liu & Hansen, 2002).

The factors identified by the current study are generally consistent with those proposed in previous research as being likely to affect the focus of peer comments (e.g. Liu & Hansen, 2002; Mangelsdorf & Schlumberger, 1992). More importantly, they were found to interact with one another to mutually influence the participants' decision-making patterns in the context of the present study. Different from previous studies that examined a number of factors separately in a decontextualized manner (e.g. Liu & Hansen, 2002), this study painted a more complex picture of multiple variables interacting with one another to exert an influence in one particular setting, lending support to a contextualized approach to the investigation of peer review. Such a complex picture does not mean that L2 writing teachers could do nothing about their students' decision-making patterns in responding to peers' writing. Instead, it highlighted the importance of exploring key factors to influence their decision-making patterns in feedback giving.

One such key factor identified by the present study was the role of the participants' perceptions of good English expository writing, represented by their general and specific criteria for it. Their general criteria, as mentioned earlier, were similar to the instructor's course assessment criteria but of a narrower range. The Chinese EFL university students in the current study seemed to differ from their English-speaking counterparts who applied consistently and significantly different criteria than those of their instructors in evaluating student texts (Newkirk, 1984). Despite the similarity to the instructor's course assessment criteria, each participant's general criteria still differed, and the weight given to the same language-related

criteria (matching “Language 3” of the peer feedback form) by each was still different, suggesting individual differences in their perceptions, which played a key role in affecting their individual decision-making patterns.

The present study also found that both participants' perceptions of English expository writing were largely shaped by their previous learning and assessment experiences of English (or Chinese) writing. In particular, the identified impact of previous experiences of English writing instruction (e.g. English writing instruction in high school and Chinese writing instruction in Hyshan's case) demonstrated the positive role of L1 and L2 writing instruction in shaping L2 learners' perceptions of English writing (Rinnert & Kobayashi, 2009). However, different from the Japanese EFL students in Rinnert and Kobayashi (2009)'s research whose L1 and L2 writing instruction mainly influenced their perceptions of English writing, the participants' past experiences of high stakes English writing tests such as TOEFL, and past English extra-curricular activities (in Fiona's case) played a part as well. Seen from the lens of socio-cognitive view of writing (Roca & Murphy, 2001), it is through previous experiences of English (or Chinese) writing instruction, high-stakes English writing test such as TOEFL and past English extra-curricular activities (in Fiona's case) that the participants learned about English writing. These experiences have become part of their individual thinking about good English writing that influenced their decision-making patterns and would supposedly influence their writing processes. One limitation, though, was that not every shaping influence on perceptions of English writing could be reported by the participants themselves, such as the influence on both participants' emphasis on vocabulary accuracy.

Findings of the current study suggest several implications for practice, especially in relation to utilizing peer review training to turn L2 students into effective peer reviewers. First, like the raters in Vaughan's (1991) study, L2 learners are not *tabula rasa*, and they have their own perceptions of what makes good English writing, which may influence their decision-making patterns in evaluating peers' texts. In the case of utilizing course assessment criteria for peer review, instead of simply presenting learners with a set of assessment criteria in peer review training, it's recommended that L2 writing teachers first dialogue with them to raise their awareness about the similarities and differences in the two parties' perceptions of components of good English writing and weight given to each component. In addition, it might also be helpful to highlight the weight given to aspects of English writing on the course assessment criteria that were particularly important for specific types of essay (e.g. illustration or cause-and-effect essay), since both participants had realized this by generating specific criteria that corresponded with particular components of the course assessment criteria for each essay type. For example, Fiona's specific criterion for the illustration essay, which caused detailed example to be the second most evaluated area in content, coincided with “detail support” on the course assessment criteria.

Second, instead of carrying out peer review training as a one-off event, as what the teacher did in the present study, L2 writing instructors might consider monitoring peer review

activity and providing constant support to repeatedly reinforce students' understanding of the course assessment criteria to the extent of reshaping their perceptions of good English writing. For instance, instructors may discuss the concepts in the course assessment criteria before each peer review task or comment on students' performance after each one. Without constant support, even students with high writing proficiency-like the two in this study-could not deepen their understanding of good English writing to guide their decision-making. By the aforementioned strategies, L2 instructors can hopefully affect students' perceptions to provide them with a basis for informed decision-making in peer review activity.

5. CONCLUSION

The present study takes a small step towards unveiling L2 learners' cognitive processes in evaluating peers' writing as represented by their individual decision-making patterns. Among the interacting factors affecting the participants' decision-making patterns, what stood out were their individual perceptions of English writing, largely shaped by previous learning and assessment experiences of English (or Chinese) writing as opposed to classroom context. Future research needs to further investigate the role of previous learning and assessment experiences as well as current classroom instruction in affecting the perceptions and decision-making patterns of L2 learners with various levels of English writing proficiency both in the context of the present study and other L2 contexts.

ACKNOWLEDGEMENTS

This study was supported by the Faculty Research Fund (FRF) for PhD. students in the Faculty of Education, University of Hong Kong. I also would like to thank the anonymous reviewers for their comments on an earlier draft of the paper.

REFERENCES

- Caulk, N. (1994). Comparing teacher and student responses to written work. *TESOL Quarterly*, 28:1, 181-187.
- Cumming, A., Kantor, R., & Powers, D. E. (2002). Decision making while rating ESL/EFL writing tasks: A descriptive framework. *Modern Language Journal*, 86: 1, 67-96.
- Ericsson, K. A., & Simon, H. A. (1993). *Protocol analysis: Verbal reports as data*. Cambridge, Mass.: MIT Press.
- Ferris, D. (2001). Teaching "Writing for Proficiency" in Summer School: Lessons from a Foxhole. In J. Murphy & P. Byrd (Eds.), *Understanding the courses we teach: Local perspectives on English language teaching*. Ann Arbor: The University of Michigan Press, pp. 328-345.
- Gass, S. M., & Mackey, A. (2000). *Stimulated methodology in second language research*. Mahwah: Lawrence Erlbaum Associates.

- Guardado, M., & Shi, L. (2007). ESL students' experiences of online peer feedback. *Computers and Composition*, 24, 443-461.
- Hu, G., & Lam, S. (8. Jan., 2009). Issues of cultural appropriateness and pedagogical efficacy: exploring peer review in a second language writing class. *Instructional Science*, 38: 4, 371-394. doi: 10.1007/s11251-008-9086-1.
- Leki, I. (1990). Coaching from the margins: Issues in written response. In B. Kroll (Ed.), *Second Language Writing: Research Insights for the Classroom* (pp. 57-68). Cambridge, UK: Cambridge University Press.
- Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic Inquiry*. Beverly Hills: Sage.
- Liu, J., & Hansen, J. G. (2002). *Peer response in second language writing classrooms*. Ann Arbor, MI: The University of Michigan Press.
- Liu, J., & Sadler, R. (2003). The effect and affect of peer review in electronic versus traditional modes on L2 writing. *Journal of English for Academic Purposes*, 2: 3, 193-227.
- Lofland, J., & Lofland, L. H. (1995). *Analysing social settings: A guide to qualitative observation and analysis* (3rd ed.). Belmont, CA: Wadsworth.
- Manchon, R., Murphy, L., & Roca, J. (2005). Using concurrent protocols to explore L2 writing processes: Methodological issues in the collection and analysis of data. In P. K. Matsuda & T. Silva (Eds.), *Second language writing research: Perspectives on the process of knowledge construction*. Mahwah, NJ: Lawrence Erlbaum Associates, pp. 191-206.
- Mangelsdorf, K., & Schlumberger, A. (1992). ESL student response stances in a peer review task. *Journal of Second Language Writing*, 1: 3, 235-254.
- Miles, M. B., & Huberman, A. M. (1984). *Qualitative data analysis: A sourcebook of new methods*. CA: Beverly Hills: Sage.
- Min, H. T. (2005). Training students to become successful peer reviewers. *System*, 33, 293-308.
- Min, H. T. (2008). Reviewer stances and writer perceptions in EFL peer review training. *English for specific purposes*, 27, 285-305.
- Newkirk, T. (1984). How students read student papers: An exploratory study. *Written Communication*, 1: 3, 283-305.
- Rinnert, C., & Kobayashi, H. (2001). Differing perceptions of EFL writing among readers in Japan. *Modern Language Journal*, 85: 2, 189-209.
- Rinnert, C., & Kobayashi, H. (2009). Situated writing practices in foreign language settings: The role of previous experience and instruction. In R. Manchon (Ed.), *Writing in foreign language contexts: Learning, teaching and research*. Bristol: Multilingual Matters, pp. 23-48.
- Roca, J., & Murphy, J. (2001). Some steps towards a socio-cognitive interpretation of second language composition processes. *International Journal of English Studies*, 1: 2, 25-46.
- Rollinson, P. (2005). Using peer feedback in the ESL writing class. *ELT Journal*, 59, 23-30.
- Smagorinsky, P. (1991). The writer's knowledge and the writing process: A protocol analysis. *Research in the Teaching of English*, 25: 3, 339-364.
- Stake, R. E. (1995). *The art of case study research*. Thousand Oaks, Calif. : Sage Publications.
- Vaughan, C. (1991). Holistic assessment: What goes in the raters' minds? In L. Hamp-Lyons (Ed.), *Assessing second language writing in academic contexts*. Norwood, NJ: Ablex, pp. 181-195.

APPENDIX 1
Peer feedback form

Content 1	relevance	creativity & interest	clarity
	A B C	A B C	A B C
Content 2	development	redundancy	
	A B C	A B C	
Organization1	thesis statement	topic sentence	detail support
	A B C	A B C	A B C
Organization2	paragraphing	coherence	unity
	A B C	A B C	A B C
Language1	vocabulary range	words variety	idiomaticity
	A B C	A B C	A B C
Language2	structure range	sentence variety	
	A B C	A B C	
Language3	word errors	sentence errors	grammar errors
	A B C	A B C	A B C
Language4	spelling	punctuation	
	A B C	A B C	
Language5	formal, informal, colloquial		
Impression	Excellent Very Good Good PassB Weak		

APPENDIX 2

Decision-making behaviors in evaluating peers' texts

a. language focus

- (1) consider word or phrase use
- (2) evaluate grammar
- (3) evaluate sentence structure
- (4) consider sentence length
- (5) evaluate word variety
- (6) assess fluency
- (7) evaluate mechanics
- (8) consider conciseness
- (9) consider idiomaticity
- (10) assess formality
- (11) consider tone
- (12) consider figure of speech

b. organization focus

- (1) assess connection between sentences
- (2) consider connection between introduction and conclusion
- (3) consider connection between body and conclusion
- (4) assess connection between topic sentence and thesis statement
- (5) consider support for main points
- (6) evaluate thesis statement
- (7) evaluate topic sentence
- (8) assess effectiveness of conclusion
- (9) consider transitional marker
- (10) evaluate ordering of ideas
- (11) consider overall organization
- (12) evaluate unity

c. content focus

- (1) assess clarity
- (2) consider detailed example
- (3) evaluate thorough development of thesis
- (4) evaluate idea
- (5) consider relevance
- (6) evaluate redundancy
- (7) evaluate appropriateness of examples
- (8) evaluate appropriateness of supporting evidence (other than examples)

APPENDIX 3

Sample of part of student text

First, some Universities and research institutions lay too much emphasis on winning Nobel Prize, which makes the scientific research a purposeful campaign. The seed of reaching for what is beyond one's grasp rooted when the Chinese kids go to kindergarten. The parents sent their children to piano class, mathematics class, chess class, you name it. When a child is asked, he/she will tell you that he/she want to be scientist, artist or even Nobel Prize winner. The formative study has misled the children to what is not their foremost interest. And it deprives them of a happy childhood and the chance to find their true love. It is the same in the scientific research field. We only see the importance of winning Nobel Prize but ignore the true meaning of research. To some extent, if Chinese want to win Nobel Prize we should first get rid of the high-flown attitude.