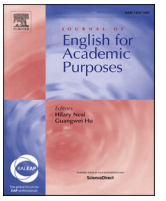
Journal of English for Academic Purposes 69 (2024) 101383

Contents lists available at ScienceDirect

Journal of English for Academic Purposes

journal homepage: www.elsevier.com/locate/jeap



A comparative study of research questions written by L1 English authors and Chinese EFL scholars

Ziqing Gong [a], Yonghou Liu [a,*], Ying Liu [b]

[a] School of Foreign Languages and Literature, Beijing Normal University, Beijing, 100875, People's Republic of China
[b] Department of Foreign Languages, Lyu Liang University, Lyu Liang, Shanxi Province, 033001, People's Republic of China

ARTICLE INFO

Handling Editor: Dr Hilary Nesi

Keywords:
English for academic purposes
Research questions
Applied linguistics
Genre analysis

ABSTRACT

Research questions (RQs) function as an important basis for entire research projects, but scant attention has been paid to their formulation. The current study compares the types and structures of RQs and inter-step shifts involving RQs in English research articles (RAs) written by L1 English authors and Chinese EFL scholars. Our data consisted of 300 English RAs from highly ranked journals in the field of applied linguistics, comprising 150 articles by L1 English authors and 150 articles by Chinese EFL scholars. The findings reveal that RQ types are used by both author groups in the following decreasing order: descriptive questions > contingent questions > comparative questions > explanatory questions > normative questions. Both Chinese and L1 English writers exhibit sophisticated competence in constructing RQs in appropriate hierarchical orders, characterized by patterns of parallel structure, progressive structure, and parallel-progressive structure. However, English L1 scholars outperform their Chinese counterparts in the use of inter-step shifts that integrate RQs into a broader text. Our findings can help writers understand the internal logic of RQs, guide them to formulate hierarchically appropriate RQs and integrate them into the entire research context.

1. Introduction

Academic research begins with research questions. Formulating appropriate research questions is a prerequisite for a successful research project (Alvesson & Sandberg, 2013; Goldschmidt & Matthews, 2022; Lim, 2014). Good research questions not only trigger interesting knowledge, but also function as an important basis for the author(s) to choose the research scope, research design and main content. Moreover, the quality and significance of a study can be first examined by considering its research questions (Agee, 2009; Lipowski, 2008; Zheng & Ruan, 2018). Therefore, research questions serve as the foundation of an investigation.

Given the important role of research questions in academic writing, it is surprising that this topic has received so little attention. In the field of genre studies, scholars have dedicated substantial effort to the analysis of research abstracts (El-Dakhs, 2018; Hyland, 2004a), acknowledgements (An et al., 2017; Hyland, 2004b), introductions (Feak & Swales, 2011; Luo & Lim, 2022), methodology (Cotos et al., 2017; Lim, 2019, 2021), results (Lim, 2010; Yang & Allison, 2003), and discussion sections (Moreno, 2021; Moyetta, 2016; Peacock, 2002). By contrast, research questions are addressed much less frequently and less profoundly. Such neglect is not limited to empirical studies, guidebooks on academic writing also provide little information on research questions (White, 2013). In

* Corresponding author. School of Foreign Languages and Literature, Beijing Normal University, Beijing, 100875, People's Republic of China.
E-mail addresses: freya_gong@163.com (Z. Gong), liuyonghou@bnu.edu.cn (Y. Liu), llgzliuying@163.com (Y. Liu).

https://doi.org/10.1016/j.jeap.2024.101383
Received 3 October 2023; Received in revised form 10 April 2024; Accepted 15 April 2024
Available online 16 April 2024

many cases, research questions are merely discussed in general terms with few detailed suggestions. Consequently, question formulation remains a challenge for many students and even researchers.

While there is a growing number of empirical studies that attempt to contrast research questions used by L1 English writers with those used by EFL scholars (Loi & Evans, 2010; Sheldon, 2011), they have been limited to the frequencies of using research questions, and have not examined the detailed patterns or features involved. Arguably, a critical gap in the literature remains, which underscores the need for a more thorough investigation into the detailed differences in research questions written by L1 English scholars and Chinese EFL scholars.

The present study conducts a detailed comparative analysis of English research articles authored by L1 English scholars and Chinese scholars. We aim to investigate the differences in research question types, research question structures, and the inter-step shifts involving research questions used by the two author groups. Accordingly, our main research question is as follows:

What are the similarities and differences between L1 English and Chinese authors in their research question types, hierarchical orders, and inter-step shifts involving research questions?

2. Literature review

2.1. Previous studies on research questions

The lack of literature on research questions has been a concern for many years. As far back as the 1970s, Lundberg (1976) noted that studies on research questions were "meagre and uneven" (p. 6). Until 2013, the scarcity of empirical studies on research questions remained evident (White, 2013). This issue was also raised by Flick (1998), who noted that many textbooks overlooked the need to introduce research questions. This deficiency of scholarly attention leaves our understanding of research questions incomplete and underscores the need for further investigation.

Fortunately, an increasing number of scholars are calling for attention to the role of research questions. Based on interviews with scholars in social science, Taylor (2002) concluded that even professional researchers paid too little attention to research questions in writing. Bordage (2001) found that a lack of research questions is the second most common reason for paper submissions to be rejected.

At the same time, empirical research has been conducted to develop our knowledge about writing quality research questions in English for academic purposes. Some scholars have examined research articles written by L1 English researchers and EFL researchers, and similar conclusions have been drawn. For instance, Sheldon (2011) conducted a study of applied linguistics research articles and found that 33.3% of L1 English writers incorporated research questions in their articles, while only 5.6% of Spanish writers did the same. Similarly, Loi and Evans (2010) investigated research articles in the field of educational psychology and found that research questions appeared in 35% of English research articles but only in 10% of Chinese research articles in the same discipline. Both studies reveal that L1 English researchers tend to include research questions more often than EFL researchers.

However, merely comparing the frequencies of using research questions is not sufficient, further and deeper analysis is still needed. Zheng and Zuo (2014) analysed Chinese research articles on foreign language teacher research and claimed that many Chinese writers fail to include research questions; moreover, the questions they do formulate are monotonous in type and lack clear logic. Later, Liu and Gong (2023) compared research questions written by Chinese expert writers with those written by Chinese novice writers and found that the questions formulated by the latter had relatively poor quality. Hence, a closer exploration of research questions is greatly needed.

2.2. Classification of research questions

For centuries, there was no definite taxonomy of research questions in social science and natural science (Dillon, 1983; Laudan, 1977), which led to a lack of clear guidance for research question formulation. It was not until the 1980s that notable progress was made, marked by the efforts of several scholars who attempted to classify research questions (de Vaus, 2001; Dillon, 1984; Lim, 2014; Lim & Luo, 2020; Van de Ven, 2007; White, 2009).

One of the most comprehensible typologies of research questions is provided by de Vaus (2001). He stated that two fundamental questions in social research are "What is going on?" and "Why is it going on?" (p. 1), thereby classifying research questions into "descriptive questions" and "explanatory questions". De Vaus also summarized the linguistic realization of the two types, stating that questions starting with "what", "who", "when" and "where" are descriptive, and questions starting with "how" and "why" are explanatory. Although this classification provides an easy way to determine a question type, the binary classification has its limitations. First, it fails to cover all cases (White, 2009), omitting question types such as comparison questions (e.g., "What similarities and differences, if any, exist between English and non-English majors in the errors of phrasal verb usage?"), as stressed by Gorard (2003). In addition, de Vaus's statement about the linguistic realization of research questions can be oversimplified. Questions starting with "what" can be explanatory questions (e.g., "What is the reason for this situation?"), while questions beginning with "how" can also be descriptive questions (e.g., "How does student learning take place in an online environment?"). In other words, defining question types based on wh-words may not lead to accurate classification.

Compared with de Vaus (2001), Lim's (2014) study offers a more comprehensive classification. Lim categorizes questions into polar and wh-questions, with wh-questions further divided into three subtypes: factual, evaluative, and circumstantial. Polar questions, which address routine events, typically follow an "auxiliary verb + subject + main verb" (V_aSV_m) structure or an existential interrogative structure. In non-polar questions, factual wh-questions seek objective information and occur in a "subject + copular verb + complement" (SVC) interrogative structure. Evaluative wh-questions, which are designed to ascertain the effects of treatment

Table 1
The classification and hierarchy of research questions.

Order	Question Types	Functions	Prototypical Syntactical Choices		Examples
First-Order	Descriptive Questions	To generate knowledge about the properties, characteristics, functions of a phenomenon.	SVC		*What are the characteristics common to the learners in an EAP thesis-writing course?*
			How + (adj/adv) + V_aSV_m		*How do advanced EFL readers make use of multi-word expressions in reading?*
Second-Order	Contingent Questions	To generate knowledge about the contingent relationship between phenomena.	Polar	VSC*/V_aSV_m*	*Are there relationships between individual difference factors and word frequency production in L2 learners?*
			Non-polar	SVC*	*What is the relationship between priming and L2 attainment?*
				How/To what extent + V_aSV_m*	*How is this metacognitive awareness related to L1-to-L2 rhetorical transfer?*
	Comparative Questions	To generate knowledge about the comparison between phenomena.	Polar	VSC**/V_aSV_m**	*Are there significant differences in the frequency of the strategies used between the two time periods?*
			Non-polar	SVC**	*What are the differences in the errors of phrasal verbs usage between English and non-English majors?*
				How/To what extent + V_aSV_m**	*How do the writing behaviours differ between a successful and an unsuccessful novice academic writer?*
Third-Order	Explanatory Questions	To generate knowledge about the reasons for and/or factors causing a phenomenon.	SVO		*What factors may explain such change or lack of change?*
			Why + V_aSV_m		*Why do students use proofreading services?*
Fourth-Order	Normative Questions	To generate knowledge about the implication of a phenomenon.	What/How + V_aSV_m		*How can this analysis inform ESP research and teaching framework?*

Note: S=Subject, C=Complement, O=Object, V=Verb (V_a = Auxiliary verb, V_m = Main verb).
C*: *correlation/relationship/…* V_m*: *correlate/relate/…* C**: *difference/similarity/…* V_m**: *differ/vary/…*

variables, start with a prepositional phrase ("to what extent"), followed by a V_aSV_m structure. Circumstantial wh-questions, aiming to explore how a dependent variable varies due to the influence of an independent variable, start with "how" and a following V_aSV_m structure. Later, Lim and Luo (2020) extended this classification by introducing three additional question types: determinative wh-questions, causative wh-questions, and personal identification wh-questions, but these newly identified types were observed to be less prevalent compared to the previously established categories. Lim's systematic classification offers a valuable framework for understanding the nuanced linguistic expressions inherent in the formulation of research questions.

Nevertheless, neither the work by de Vaus (2001) nor the studies by Lim and his associate (Lim, 2014; Lim & Luo, 2020) explore the logical relations between question types. Their classification framework treats different types of research questions in isolation, with little attention given to their interconnections. However, it is crucial to investigate the hierarchical relationships among these question types because, in this way, the degree to which research questions are logically organized can be revealed, thereby enhancing the overall clarity of the research.

Regarding the logical relations between research questions, Dillon (1984) presents a comprehensive classification framework that encompasses properties, comparisons, contingencies and other types of questions. The study underscores the hierarchical relationship between different types of research questions, asserting that research questions are hierarchically related to each other by the principle of information containment. This means that lower-order questions are entailed in higher-order questions. Higher-order questions can only be addressed if lower-order questions are answered. Therefore, the construction and formulation of research questions should follow the hierarchical order.

Although Dillon's framework has merit, it does not consider explanatory questions as an independent question type, but blurs them into descriptive questions. However, these two types of questions are markedly different, as they belong to different hierarchical orders. The knowledge contained in descriptive questions is more basic than that contained in explanatory questions, which means that explanatory questions can only be addressed if descriptive questions are answered. In this sense, these two question types should not be deemed identical. This framework was later integrated into four types, namely, descriptive questions, comparative questions, contingent questions, and normative questions (Alvesson & Sandberg, 2013), but explanatory questions are still missing.

In response to the identified limitations, we integrated the classification frameworks proposed by de Vaus (2001), Dillon (1984), and Alvesson and Sandberg (2013) into a novel system (Liu & Gong, 2023). In this framework, research questions are categorized into five types within a hierarchy from the first-order to fourth-order questions. However, while comprehensive in its function-oriented structure, this framework lacks an exploration of the linguistic realization inherent in each question type.

Recognizing the importance of bridging the gap between form and function, as underscored by Moreno and Swales (2018), we update our classification in the current study. Inspired by the syntactic forms of research questions outlined by Lim and his associate (Lim, 2014; Lim & Luo, 2020), we incorporate the prototypical linguistic resources associated with each question type and thereby enhance our framework by providing insights into the intersection of form and function. The updated classification framework is illustrated in Table 1.

In this classification framework, research questions are hierarchically related to each other. First-order questions (i.e., descriptive questions) aim to describe what constitutes or characterizes a phenomenon. They are prototypically realized in the form of a non-polar question starting with the question words "what" or "how". When starting with "what", a descriptive question follows a "subject 'what' + auxiliary verb 'be' + complement" (SVC) structure. If it begins with "how", a descriptive question has the following structure: "how + (optional) adjective or adverb + auxiliary verb + subject + main verb" (i.e., how + (adj/adv) + V_aSV_m). Descriptive questions are the most basic question type, establishing the foundation for the following higher-order questions.

Contingent questions and comparative questions are second-order questions. They aim to generate knowledge about the contingent relationship and comparison between phenomena. Syntactically, these questions are prototypically realized by both polar and non-polar questions. Polar questions follow the structure of "auxiliary verb 'be' + subject 'there' + complement" (VSC) or "auxiliary 'do' + subject + main verb" (V_aSV_m). Non-polar questions mirror descriptive questions by employing the SVC and "how (or 'to what extent') + V_aSV_m" structure. However, the lexicogrammatical choices of the "complement" and "main verb" should specifically address concepts of "relationship" and "comparison". Table 1 provides some typical wordings associated with these choices. Second-order questions entail first-order questions and presuppose that descriptive questions can be answered first.

Third-order questions (i.e., explanatory questions) yield knowledge about the cause of a phenomenon. These questions are prototypically structured as SVO or "why + auxiliary verb + subject + main verb" (why + V_aSV_m). Explanatory questions presuppose knowledge of lower-order questions.

The last question type is that of normative questions, which belong to the fourth-order. These questions typically follow the form of a "what/how + auxiliary verb + subject + main verb" (what/how + V_aSV_m) structure. Normative questions generate knowledge about the implications of a phenomenon, and their solutions presuppose knowledge of the aforementioned questions.

2.3. Research question as a step in the CaRS model

Swales' (1990, 2004) "Create a Research Space" (CaRS) model has been widely recognized as a valid framework for analysing research article introductions (Ahmed, 2004; Taylor & Chen, 1991) and forms the basis for the genre-based analysis in this study. Swales first proposed the CaRS model in 1990 to predict the organization of research introductions and later modified it in 2004. In both versions, the CaRS model consists of three moves, which are labelled below.

i. Move 1—establishing a research territory
ii. Move 2—establishing a niche

Table 2
Move 3 in the CaRS model (Swales, 2004).

Move 3: Presenting the present work
Step 1 (obligatory) Announcing present research descriptively or purposively
Step 2 (Optional*) Presenting research questions or hypotheses
Step 3 (Optional) Definitional clarifications
Step 4 (Optional) Summarizing methods
Step 5 (PISF**) Announcing principal outcomes
Step 6 (PISF) Stating the value of the present research
Step 7 (PISF) Outlining the structure of the paper

Note: * Step 2-4 are not only optional but less fixed in their order of occurrence than the others.
**PISF: Probable in some fields, but unlikely in others.

iii. Move 3—occupying the niche (Swales, 1990)/
Move 3—presenting the present work (Swales, 2004)

As noted, Move 3 was revised in the later version. Swales (2004) acknowledges that the third move is much more complex than his original envisions and thus proposes a more explicit functional label, specifying "presenting research questions or hypotheses" as one of the rhetorical steps in Move 3. The full version of Move 3 is shown in Table 2.

Specifying research questions as a step in Move 3 in the revised CaRS model has highlighted its significance, promoting increasing attention to the formulation of research questions. Lim (2014) investigated the rhetorical shift from a purpose statement (M3S1) to the presentation of a research question (M3S2) and found that research questions are typically preceded by purpose statements. Therefore, he concluded that the construction of meaningful research questions depends on the clarity of the purpose statements. This finding highlights the rhetorical relationship between research questions and their surrounding steps.

While Lim's (2014) finding is illuminating, it is essential to note that only M3S1 was given the focus in his study. Therefore, a more thorough analysis that encompasses all steps in Move 3 is necessary. This investigation can elucidate how L1 English and Chinese EFL scholars use these rhetorical steps to integrate research questions into their broader texts.

3. Methodology

3.1. Data collection

For the comparative analysis, we built two corpora consisting of 300 research articles written by L1 English scholars and Chinese scholars, i.e., the English Scholar Corpus (ESC) and the Chinese Scholar Corpus (CSC). Each corpus contains 150 research articles.

The data collection followed the following requirements. First, all the articles were randomly drawn from 5 prestigious journals in the area of applied linguistics between 2018 and 2022 based on the impact factor (*Applied Linguistics, Language Teaching Research, Studies in Second Language Acquisition, The Modern Language Journal, English for Specific Purposes and Journal of English for Academic Purposes*). Second, only research articles with specific research questions were collected. Third, as the present study aims to compare research questions written by L1 English authors and Chinese EFL scholars, articles with co-authors from different language backgrounds were excluded. The classification of an author as either an L1 Chinese or L1 English speaker was determined based on their name, institution, and other background information (e.g., previous institutions, education). Specifically, authors with typical English names who are affiliated with institutions in the major English-speaking countries—Britain, the United States, Canada, Australia, and New Zealand—were taken to be L1 English speakers. We also checked authors' educational information and previous institutions from their personal profiles. Articles authored by individuals who obtained their PhDs or have previously worked in non-English-speaking countries are excluded from the corpora. Even so, we acknowledge that these methods may eliminate many L1 English authors working abroad and/or having a non-English name. Nonetheless, we assume that authors with typical English names, who have completed their PhDs and are associated with Anglophone institutions, likely possess near-native English proficiency. On the other hand, authors using Pinyin to Romanise their Chinese names, along with other Chinese background information, were identified as L1 Chinese speakers. This criterion has also been adopted in similar studies (Gao, 2016; Geng & Wharton, 2016).

Accordingly, L1 English scholars and Chinese authors can be considered as being in the same discourse community while having "comparable disciplinary expertise" and "pedagogically relevant expert performances" (Tribble, 2017, p. 34).

3.2. Data coding

The current study involved coding of question types and rhetorical steps. In terms of research question types, the coding was based on the classification framework shown in Table 1. Two steps were followed for the coding. The first step involved classifying research questions into different types. The second step involved examining the hierarchical order of research questions in each article. In this respect, each type was given a number (e.g., 1 for descriptive questions, 2 for contingent questions) and the order of these types was inserted into an Excel file for statistical examination. For example, in one article, the research questions might be coded as 1-2-3-5, while another might be coded as 3-3-1-1. The coded data were then run through the chi-square test to determine whether there were significant differences.

Swales' (2004) CaRS model was chosen for coding the inter-step shifts involving research questions. This model was selected due to its efficiency in implementation and popularity. While the model consists of 3 moves, the current study only focuses on the formulation of research questions, a component within Move 3. Therefore, only the third move was analysed. The coding of inter-step shifts involved two steps. In the first step, the authors read through the introduction (and method, if necessary), highlighting each step with different colours. In the second step, the employment of each step was inserted into an Excel file. Afterwards, both statistical examinations and an in-depth analysis were conducted for comparisons.

4. Results and discussion

4.1. Distribution of research question types

Based on the data collected, the total number of research questions in the two corpora is 796. The ESC accounts for 407, with an average of 2.71 RQs per article and a range from 1 to 7. The CSC accounts for 389, with an average of 2.59 RQs per article and a range from 1 to 6. The findings suggest the similarity in the quantity of research questions per article used by L1 English scholars and Chinese scholars. Overall, most articles from both groups contain 2-3 research questions. This trend may be attributed to the characteristics of the RA genre, as a journal article tends to focus on a narrow topic and has a limited length. Excessive research questions in a single article may impede the depth of exploration, potentially compromising the overall quality of the work.

Our research question first focuses on exploring the frequencies of occurrence of the five question types in the ESC and the CSC. As illustrated in Table 3, the distribution of research question types appears to conform to the hierarchical order, with first-order questions exhibiting the highest frequency, followed by second-order, third-order, and fourth-order questions in a decreasing manner.

As observed, all types of research questions are found in both the ESC and the CSC, but in varying proportions, as follows: descriptive questions (50.01% vs. 46.27%), contingent questions (24.57% vs. 27.25%), comparative questions (17.44% vs. 20.31%), explanatory questions (4.18% vs. 3.34), and normative questions (3.70% vs. 2.83%). According to the chi-square test, there is no significant difference in the frequencies of research questions across the two corpora, both when considering all question types collectively and at each question type individually ($p > 0.05$). This suggests an overall commonality in the usage of research question types by both groups.

Additionally, similarities in the order of frequencies between the ESC and the CSC can be drawn. First, in both corpora, research question types are used in a decreasing order of descriptive questions > contingent questions > comparative questions > explanatory questions > normative questions. Second, questions from different orders are incorporated in a decreasing order of first-order questions > second-order questions > third-order questions > fourth-order questions. Such a descending order aligns with our earlier comparative studies between the research questions written by Chinese expert and novice writers (Liu & Gong, 2023) and echoes Dillon's (1984) view that there is an implicit hierarchy of research questions. Lower-order questions generate more fundamental knowledge than higher-order ones; thus, they occur more frequently than the less fundamental higher-order questions.

Descriptive questions are the first-order questions and the most basic question type. Due to their fundamentality, descriptive questions constitute the majority of inquiries in both the ESC and the CSC. By answering descriptive questions, we can gain a clearer understanding of what characterizes a phenomenon and build the basis for higher-order questions.

Explanatory questions are not substantially represented in either the ESC or the CSC, ranking after descriptive, contingent, and comparative questions. The low proportion of explanatory questions is also noted by Zheng and Zuo (2014), who considered it a deficiency that could diminish the quality and significance of academic research. In our view, however, the low proportion of explanatory questions is reasonable since it conforms to the inherent hierarchical order of research questions. Explanatory questions are third-order questions, which means they are not as fundamental as the lower-order questions, and this can explain their lower frequency in research articles.

As evidenced by the similar decreasing frequency order of research question types in both corpora, we can tell that there is neither superiority nor inferiority in the types of research questions, given that different types are only discriminated by the degree of

Table 3
Distribution of research question types in the ESC and the CSC.

Type of RQ	ESC		CSC		Chi-square
	Frequency	Percentage (%)	Frequency	Percentage (%)	
Descriptive Question	204	50.01	180	46.27	$\chi^2 = 1.181$ P = 0.277
Contingent Question	100	24.57	106	27.25	$\chi^2 = 0.744$ P = 0.388
Comparative Question	71	17.44	79	20.31	$\chi^2 = 1.067$ P = 0.302
Explanatory Question	17	4.18	13	3.34	$\chi^2 = 0.382$ P = 0.536
Normative Question	15	3.70	11	2.83	$\chi^2 = 0.463$ P = 0.496
Total	407	100	389	100	$\chi^2 = 2.845$ P = 0.584

fundamentality in the knowledge they generate. Thus, it is unwise to judge the quality of research questions solely according to question types.

4.2. Hierarchical orders of research questions

As previously stated, the quality of research questions should not only be evaluated by their types but also by whether they conform to the hierarchical order. The formulation of the research question is supposed to follow the hierarchical order, or a study may turn out to be incoherent and vague (White, 2009). In terms of logical appropriateness, the orders of the research questions are shown in Table 4.

The data analysis reveals that there is little difference in the appropriateness of the hierarchical orders of research questions in the ESC and the CSC. Specifically, 95.33% (143/150) of RAs in the ESC and 94.67% (142/150) of RAs in the CSC construct their research questions in an appropriate hierarchical order. This challenges Zheng and Zuo's (2014) assertion that many Chinese authors struggle with formulating logical questions. The high proportion of logically structured questions in both groups suggests that Chinese scholars and L1 English authors possess an equally high ability in constructing research questions in the appropriate hierarchical order.

Through the above comparison, we can gain a general understanding of writers' competence in constructing research questions. However, besides the high proportions of logically constructed research questions in both corpora, some readers may wonder what makes them logical and appropriate. To gain deeper insights into this aspect, it is essential to move beyond superficial quantitative comparisons and delve into the analysis of structural patterns within research questions.

According to the dataset, research questions are either structured in the same hierarchical order, progress from lower-order to higher-order, or present a combination of both same-order and progressive-order questions. The former two patterns echo White's (2009) statement that writers should order their research questions "either into a sequence or hierarchy" (p. 50), yet he did not clearly define what sequence and hierarchy are. Expanding on White's concepts, we define "sequence" and "hierarchy" as the parallel structure and the progressive structure, respectively, and introduce a new parallel-progressive structure based on our dataset. Consequently, three patterns are proposed, they are the parallel structure, the progressive structure, and the parallel-progressive structure. These three patterns are in accordance to our previous research (Liu & Gong, 2023) and will be introduced below. By providing examples of each pattern, we aim to develop a more nuanced understanding of what constitutes appropriately ordered research questions.

4.2.1. The parallel structure

Our analysis has revealed that writers may present research questions in a parallel structure, in which all of the questions belong to the same hierarchical order. A typical example of the parallel structure can be seen in Fig. 1.

In this example, the authors target the differences of the "(un) attended this/these" between high-rated and low-rated ESL argumentative essays and pose four comparative questions in a row. The four questions all belong to the second-order but focus on different details. The differences in "frequency", subsequent "verb types and patterns", "the structural distribution of the antecedents" and "the stance options of attending nouns" are questioned respectively to gain a deep understanding about the usage of "(un) attended this/these" in ESL argumentative essays.

Table 4
Appropriateness in the order of research questions in the ESC and the CSC.

Hierarchical Order	ESC		CSC	
	Frequency	Percentage (%)	Frequency	Percentage (%)
Appropriate Order	143	95.33	142	94.67
Inappropriate Order	7	4.67	8	5.33
Total	150	100	150	100

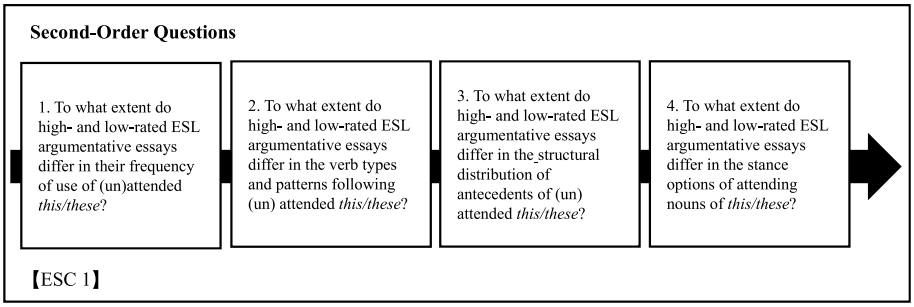


Fig. 1. An example of the parallel structure.

When research questions exhibit a parallel structure, they typically involve detailed inquiry into specific aspects of a phenomenon or an attempt to contextualize the phenomenon in various situations for a more elaborate understanding. Ultimately, a parallel structure can lead to more thorough knowledge, providing a deep understanding of the target phenomenon. Thus, we can conclude that the parallel structure is characterized as being "deep".

4.2.2. The progressive structure

Our findings also show that writers may use a progressive structure, in which the sequence of research questions progresses from lower-order questions to higher-order questions. It can either progress consecutively (e.g., first-order→second-order→third-order) or proceed in a skipped manner (e.g., first-order→third-order). Fig. 2 illustrates an example of the progressive structure.

In this example, the authors follow the sequence of "what is it", "what are the differences", "why are there differences" and "how should these differences inform us", generating comprehensive knowledge about "lexical bundles".

Such a structure aims to conduct a systematic investigation and generate broad knowledge about a certain phenomenon. Compared with a parallel structure, which targets at depth, a progressive structure is more focused on width. By exploring the characteristics of the phenomenon, comparing its differences when positioned in different situations, and seeking explanations and solutions, it presents readers with extensive knowledge about the focal phenomenon. Therefore, the progressive structure can be characterized as being "broad".

4.2.3. The parallel-progressive structure

Besides parallel and progressive structures, we have also found that writers may use a parallel-progressive structure, in which research questions exhibit characteristics of both structures. Fig. 3 provides an example of the parallel-progressive structure.

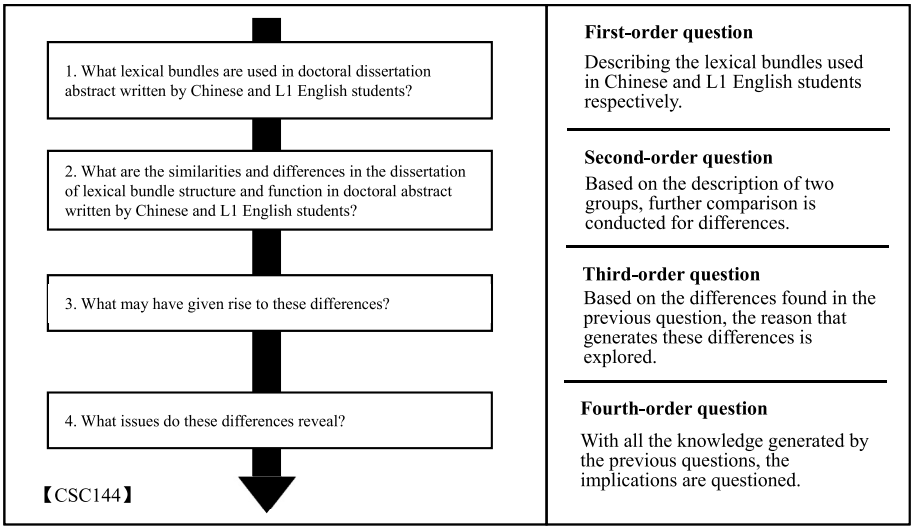


Fig. 2. An example of the progressive structure.

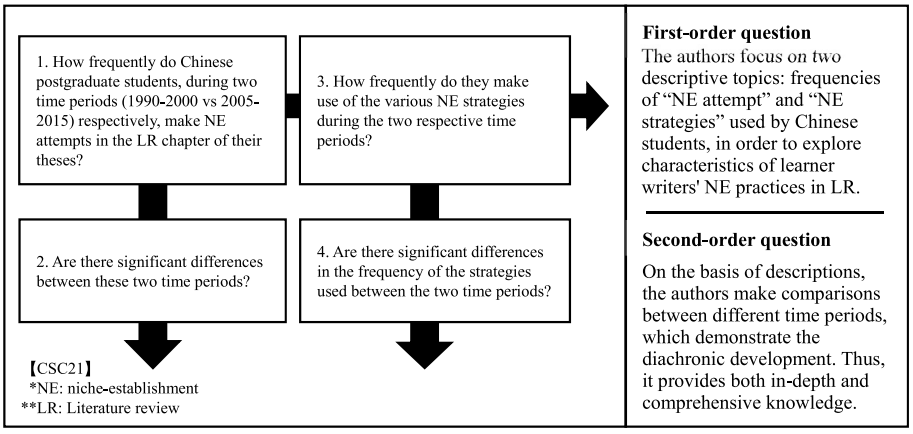


Fig. 3. An example of the parallel-progressive structure.

In this example, the authors not only concentrate on diverse aspects of niche-establishment — "attempt" and "strategies" — which constitute a parallel line of inquiry, but also adhere to a sequence that progresses from the inquiry of "what it is" to "what the differences are", thereby yielding a comprehensive understanding.

4.2.4. Comparison of the research question patterns

Based on the aforementioned patterns in formulating research questions, the frequency of each pattern is shown in Table 5.

As presented, the two predominant structural patterns in both the ESC and the CSC are the parallel and progressive structures. In both corpora, the parallel structure is the most prevailing choice, representing 57.63% in the ESC and 54.93% in the CSC. The use of the progressive structure follows closely, accounting for 38.19% in the ESC and 42.25% in the CSC. The parallel-progressive structure is notably least commonly used in both the ESC (2.78%) and the CSC (2.82%). This observed preference for either a horizontal (i.e., the parallel structure) or vertical (i.e., the progressive structure) thinking process might be attributed to the inherent inclination towards linear thinking.

These three patterns can effectively encompass common structures of logically appropriate research questions. Regardless of which pattern is employed, as long as the questions are arranged in hierarchically appropriate order, the logic of the research can be predicted.

4.3. Inter-step shifts integrating research questions into a broader text

In the aforementioned sections, we have discussed how to formulate research questions in a hierarchically logical order. Nevertheless, research questions should not be considered in isolation. Good research questions also need to be effectively integrated into full articles. In this section, we will explore inter-step shifts that act as a bridge between the research questions and the surrounding steps, and illustrate the strategies employed by L1 English scholars and Chinese scholars in integrating their research questions into the broader context.

We referred to Swales's (2004) CaRS model for the analysis. Typically, research questions are not presented directly at the beginning of Move 3. Instead, writers may utilize rhetorical steps such as concise statements of purposes, methods, findings, and values to lay the groundwork for a comprehensive presentation of research questions. Table 6 presents the frequencies of these steps in the ESC and the CSC.

Similarly, in both corpora, M3S1 stands out as the most frequently used step, while M3S5 and M3S7 consistently rank as the least utilized steps. However, there are differences in the frequencies of specific steps. The frequencies of M3S1, M3S4, and M3S6 in the ESC are notably higher than those in the CSC. These steps contribute to a clearer presentation and enhance the value of the present study. Conversely, the underemployment of these steps may diminish readers' efficiency and lead to an underestimation of the study's significance.

Typically, research questions should be supplemented with a detailed and precise statement of purpose (Lim, 2014; White, 2009). As illustrated in Table 6, this step is included in 97.33% (146/150) of the ESC and 90% (135/150) of the CSC. Notably, the employment

Table 5
Employment of research question patterns.

Pattern	Appropriate Order		
	Parallel structure	Progressive structure	Parallel-progressive structure
ESC	57.63%	38.19%	2.78%
CSC	54.93%	42.25%	2.82%

Table 6
Frequency of each step in Move 3.

Move	ESC		CSC		Chi-square
	Frequency	Percentage	Frequency	Percentage	
M3S1	146	97.33%	135	90%	**χ^2=6.799 P=0.009**
M3S2	150	100%	150	100%	
M3S3	14	9.33%	18	12%	χ^2=0.560 P=0.454
M3S4	72	47.33%	55	37.33%	**χ^2=3.946 P=0.047**
M3S5	8	5.33%	2	1.33%	χ^2=3.724 P=0.054
M3S6	40	26.67%	15	10%	**χ^2=13.915 P=0.000**
M3S7	8	5.33%	4	26.67%	χ^2=1.389 P=0.239

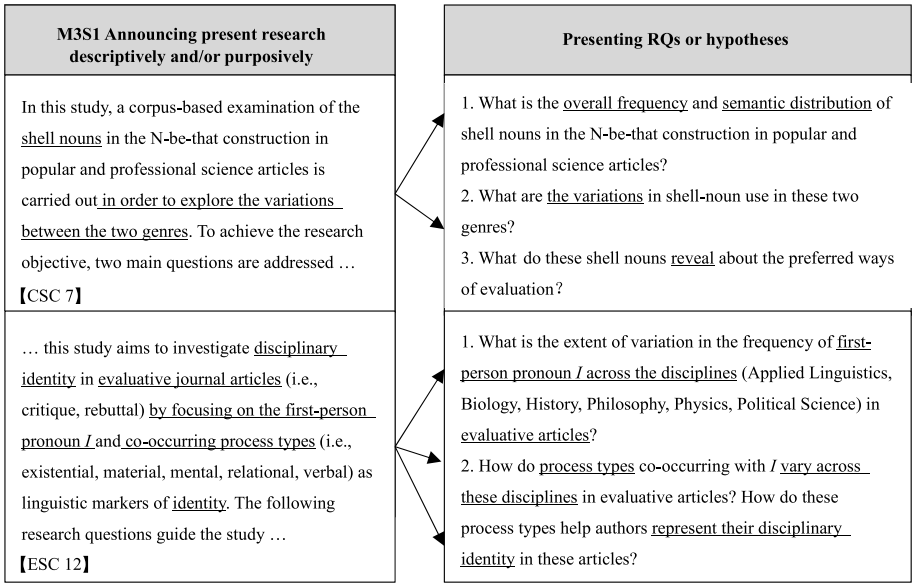


Fig. 4. Inter-step shifts from purpose announcements to research questions.

of the purpose statements in the CSC is significantly lower than that in the ESC ($p = 0.009 < 0.01$), suggesting that Chinese authors may not have given sufficient attention to the lead-in step for research questions. Many Chinese scholars proceed to Move 3 with a direct presentation of research questions, which seems to be a little abrupt. Fig. 4 provides several examples of how a smooth transition from the purpose statement to the presentation of research questions can be achieved.

In the first example, the purpose statement manifests that its objective is to explore the variations of the shell nouns in the N-be-that construction between popular and professional science articles. This objective is then broken down into (i) the overall frequency and semantic distribution of shell nouns (RQ1), (ii) the variations in shell-noun use in these two genres (RQ2), and (iii) implications revealed by these shell nouns (RQ3). Alternatively, the purpose statements may introduce the dependent variables to be measured, with each variable addressed in subsequent research questions. In the second example, the purpose statement specifies three variables: (i) first-person pronoun "I", (ii) co-occurring process, and (iii) disciplinary identity. Subsequently, each of these variables is addressed in three corresponding research questions.

The semantic linkages underscore that a meaningful presentation of research questions relies on clear purpose statements. The lack of emphasis on clear purpose statements may lead to ambiguity or misinterpretation of research questions, potentially hindering communication in both academic and cross-cultural contexts.

In addition to M3S1, the employment of M3S4 in the CSC is also significantly lower than that in the ESC. The method summary serves to provide the audience with a concise overview of the research procedure. Even though this step is not mandatory, it offers a systematic view of the essential information in the research and draws the audience's attention to the full text. Fig. 5 illustrates the rhetorical shift between the presentation of research questions and the method summaries.

In these examples, writers use the "summarizing method" step to elaborate the procedure of addressing the research questions. Writers may introduce their research participants or objectives (e.g., "women scholars", "Chinese EFL learners' writing") and how they collected and analysed their data (e.g., "interview-based study", "corpus-assisted analysis"). Although not obligatory, this step efficiently offers the audience a concise overview, facilitating the comprehension of the current research.

The significantly lower employment of this step by Chinese writers compared with English L1 writers may indicate a lack of awareness among Chinese scholars regarding the importance of presenting concise and informative method summaries. This oversight may hinder readers from quickly grasping the essence of the study.

Moreover, M3S6 is also overlooked by many Chinese scholars, and admittedly, some English L1 authors. This step aims to inform readers about the value of the current study by highlighting its contributions by answering research questions, thereby elevating its status. Compared with L1 English writers, Chinese writers show salient reticence in highlighting the value of their studies. As a step to display the merits of one's work, M3S6 occur in only 10% (15/150) of the RAs in the CSC but 26.67% (40/150) in the ESC, representing a significant difference ($p = 0.000 < 0.01$). Fig. 6 illustrates examples of the inter-step shift from research questions to value statements.

As exemplified in Fig. 6, writers may exhibit their contribution in dealing with the research questions by (i) solving real-world problems (e.g., "lead to a better understanding of …"), (ii) bridging the gap in previous research (e.g., "seldom offer guidance", "often goes unnoticed"), or (iii) laying the groundwork for future investigations (e.g., "be a basis for further research"). Explicitly articulating the potential contributions of the current study enables readers to grasp its significance, potentially captivating their interest and encouraging them to delve into the full text.

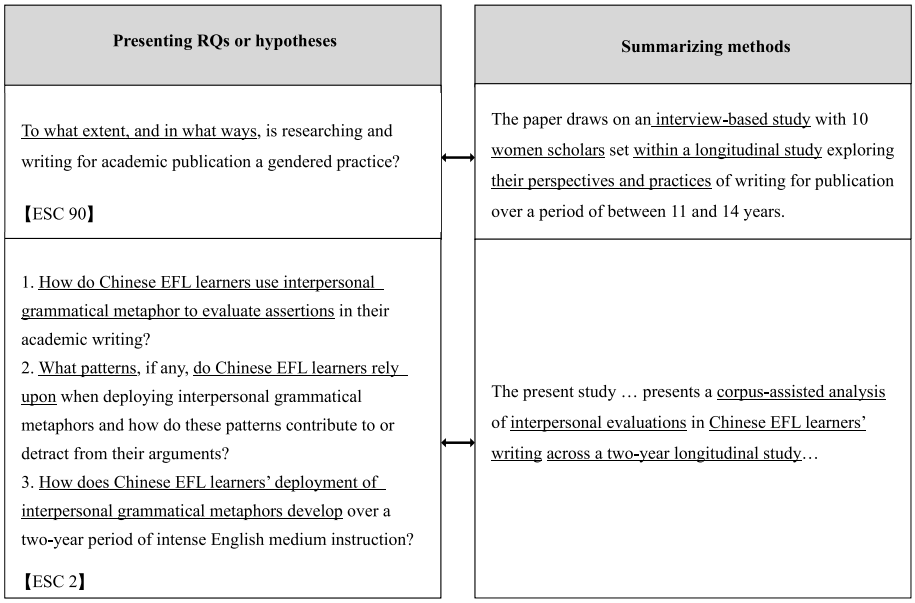


Fig. 5. Inter-step shifts between research questions and method summaries.

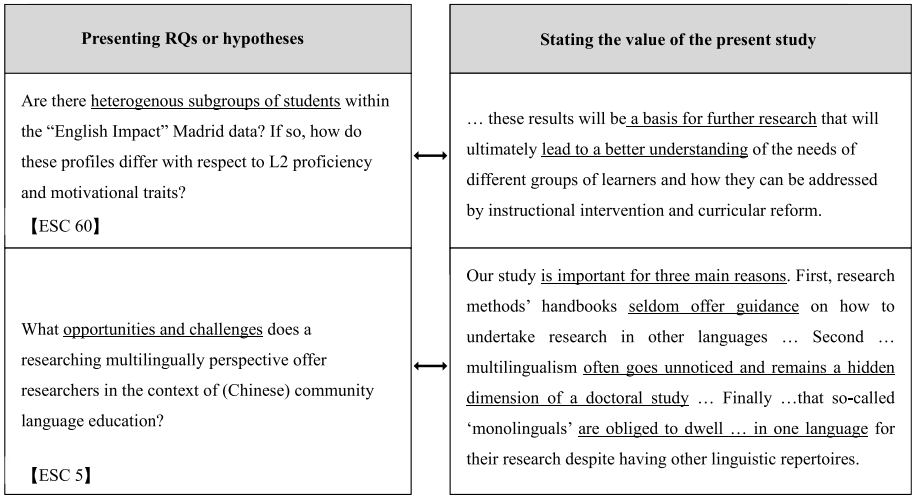


Fig. 6. Inter-step shifts between research questions and value statements.

In contrast, Chinese expert scholars prefer to be reserved in highlighting the value of their own work. Such reticence may result from traditional Chinese ethics, which favour humility rather than self-praise (Lee, 2020; Loi, 2010). However, considering reading efficiency, being taciturn in the value statement step may not be helpful.

5. Conclusion and implications

In this study, we report a comparative analysis of research questions in applied linguistics constructed by L1 English scholars and Chinese scholars. We have updated the classification framework for research questions based on previous studies, and employed Swales' CaRS model (2004) for the rhetorical analysis.

Regarding the distribution of research question types, both Chinese and L1 English scholars incorporate research questions in the descending order of descriptive questions > contingent questions > comparative questions > explanatory questions > normative questions and in a decreasing hierarchical order of first-order questions > second-order questions > third-order questions > fourth-order questions. This finding validates the inherent hierarchical order of research questions and refutes the view that regards question type as the sole criterion for evaluating the quality of research questions.

Regarding the organization of the hierarchical order of research questions, it is observed that in both the ESC and the CSC, most research questions are formulated in an appropriate order, following the inherent hierarchy. This suggests sophistication in question formulation by both author groups. Furthermore, we propose research questions conforming to this hierarchical order follow three patterns: the parallel structure, the progressive structure, and the parallel-progressive structure. Among them, the parallel and progressive structures are the most prevalent in both corpora, while the parallel-progressive structure is used to a lesser extent.

However, despite being evenly matched with L1 English writers in terms of question formulation, Chinese EFL writers face challenges in effectively integrating their research questions into their whole writing. This is largely due to a tendency among some Chinese writers to undervalue the rhetorical steps surrounding research questions, potentially making their studies seem less competitive than those produced by their ESC counterparts. Specifically, steps including purpose announcements, method summaries, and value statements are underused by Chinese writers, so audiences may not be able to efficiently gain an overview of these studies.

Theoretically, based on previous classifications of research questions, the current study proposes a refined classification framework for research questions consisting of four orders and five types. This framework provides readers with a clear illustration of the distinctions between different types of research questions as well as the inherent hierarchy among the different orders.

Pedagogically, this study provides some implications for EAP instruction, particularly concerning the hierarchy and linguistic resources used in formulating research questions. Given the inherent hierarchical nature of research questions, EAP instructors may have to first explain various types of research questions and their hierarchical relationships. Additionally, the introduction of the three identified structural patterns with concrete examples can further facilitate students in constructing hierarchically appropriate research questions. In terms of applying these structural patterns, instructors should assess the focus of the research undertaken by novice writers before providing targeted guidance. If learners aim for a deep and detailed exploration, more emphasis can be placed on guiding them to formulate research questions in a parallel structure. If learners seek a broad understanding, probing not only "what" but also "why" and "how", additional guidance on the progressive structure may be beneficial. In cases where learners aspire to encompass both deep and broad exploration, guidance on the parallel-progressive structure would be appropriate. When writers encounter language-related challenges in formulating research questions, introducing the linguistic resources for each question type (as outlined in Table 1) can be beneficial. As these linguistic resources largely rely on real applied linguistic studies, instructors can use them as a foundation to determine what to emphasize while teaching students to formulate actual research questions in EAP classes. Furthermore, instructors need to highlight the need of rhetorically connecting research questions to surrounding steps, such as purpose statements, value announcements, and other related rhetorical steps. This ensures that the questions are effectively integrated into the broader research context, which can enhance the overall quality and impact of the research.

The present study may have several limitations. First, the research required a substantial amount of manual effort, as close reading of certain sections of research articles was necessary for identifying question types and rhetorical steps. As a result, there might be minor mistakes. Second, aiming to bridge the function-form gap, we tentatively propose a version of linguistic realizations for our function-oriented classification, but this proposal might not be comprehensive. Finally, additional work is needed to test the generalizability of our current findings and determine whether cross-discipline variation exists in the formulation of research questions.

CRediT authorship contribution statement

Ziqing Gong: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Data curation. **Yonghou Liu:** Writing – review & editing, Validation, Supervision, Resources, Project administration, Methodology, Investigation, Conceptualization. **Ying Liu:** Writing – review & editing.

Declaration of competing interest

None.

Funding Acknowledgement

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

References

Agee, J. (2009). Developing qualitative research questions: A reflective process. *International Journal of Qualitative Studies in Education, 22*(4), 431–447.
Ahmed, F. (2004). Rhetorical properties of Arabic research article introductions. *Journal of Pragmatics, 36*(6), 1119–1138.
Alvesson, M., & Sandberg, J. (2013). *Constructing research questions: Doing interesting research*. CA: Sage.
An, J., Jeon, S., Jones, T., & Song, M. (2017). Data-driven pattern analysis of acknowledgments in the biomedical domain. *Data and Information Management, 1*(1), 35–48.
Bordage, G. (2001). Reasons reviewers reject and accept manuscripts: The strengths and weaknesses in medical education reports. *Academic Medicine, 76*(9), 889–896.
Cotos, E., Huffman, S., & Link, S. (2017). A move/step model for methods sections: Demonstrating rigour and credibility. *English for Specific Purposes, 46*, 90–106.
de Vaus, D. A. (2001). *Research design in social research*. London: Sage.
Dillon, J. T. (1983). Use of questions in educational research. *Educational Researcher, 12*(9), 19–24.
Dillon, J. T. (1984). The classification of research questions. *Review of Educational Research, 53*, 327–361.
El-Dakhs, D. (2018). Why are abstracts in PhD theses and research articles different? A genre-specific perspective. *Journal of English for Academic Purposes, 36*, 48–60.
Feak, C. B., & Swales, J. M. (2011). *Creating contexts: Writing introductions across genres*. MI: University of Michigan Press.
Flick, U. (1998). *An introduction to qualitative research*. London: Sage.

Gao, X. (2016). A cross-disciplinary corpus-based study on English and Chinese native speakers' use of linking adverbials in academic writing. *Journal of English for Academic Purposes, 24*, 14–28.
Geng, Y., & Wharton, S. (2016). Evaluative language in discussion sections of doctoral theses: Similarities and differences between L1 Chinese and L1 English writers. *Journal of English for Academic Purposes, 22*, 80–91.
Goldschmidt, G., & Matthews, B. (2022). Formulating design research questions: A framework. *Design Studies, 78*, Article 101062.
Gorard, S. (2003). *Quantitative methods in social science*. London: Continuum.
Hyland, K. (2004a). *Disciplinary discourses: Social interactions in academic writing*. Ann Arbor: University of Michigan Press.
Hyland, K. (2004b). Graduates' gratitude: The generic structure of dissertation acknowledgements. *English for Specific Purposes, 23*, 303–324.
Laudan, L. (1977). *Progress and its problems*. Berkeley: University of California Press.
Lee, W. W. L. (2020). Impression management through hedging and boosting: A cross-cultural investigation of the messages of U.S. And Chinese corporate leaders. *Lingua, 242*, Article 102872.
Lim, J. M. H. (2010). Commenting on research results in applied linguistics and education: A comparative genre-based investigation. *Journal of English for Academic Purposes, 9*, 280–294.
Lim, J. M. H. (2014). Formulating research questions in experimental doctoral dissertations on Applied Linguistics. *English for Specific Purposes, 35*, 66–88.
Lim, J. M. H. (2019). Explicit and implicit justifications of experimental procedures in language education: Pedagogical implications of studying expert writers' communicative resources. *Journal of English for Academic Purposes, 37*, 34–51.
Lim, J. M. H. (2021). *Writing the methods sections of research articles in language education: A genre-based resource book for writers in language education and the teaching of English as a Second Language*. Singapore: Partridge.
Lim, J. M. H., & Luo, X. (2020). Writing research questions and hypotheses: A genre-based investigation into writers' linguistic resources in social science. *ESP Today, 8*(2), 206–226.
Lipowski, E. E. (2008). Developing great research questions. *American Journal of Health-System Pharmacy, 65*(17), 1667–1670.
Liu, Y., & Gong, Z. (2023). Typology and hierarchy of research questions: A corpus-based comparative study between novice and expert writers' academic English papers. *Foreign Language World*, (5), 31–39.
Loi, C. K. (2010). Research article introductions in Chinese and English: A comparative genre-based study. *Journal of English for Academic Purposes, 9*(4), 267–279.
Loi, C. K., & Evans, M. S. (2010). Cultural differences in the organization of research article introductions from the field of educational psychology: English and Chinese. *Journal of Pragmatics, 42*, 2814–2825.
Lundberg, C. C. (1976). Hypothesis creation in organisational behaviour research. *Academy of Management Review, 1*(2), 5–12.
Luo, X., & Lim, J. M. H. (2022). Succinct summaries of research methods: A genre analysis of expert writers' communicative resources in social sciences. *Iberica, 44*, 75–100.
Moreno, A. I. (2021). Selling research in RA discussion sections through English and Spanish: An intercultural rhetoric approach. *English for Specific Purposes, 63*, 1–17.
Moreno, A. I., & Swales, J. M. (2018). Strengthening move analysis methodology towards bridging the function-form gap. *English for Specific Purposes, 50*, 40–63.
Moyetta, D. (2016). The discussion section of English and Spanish research articles in psychology: A contrastive study. *ESP Today, 4*(1), 87–106.
Peacock, M. (2002). Communicative moves in the discussion section of research articles. *System, 30*(4), 479–497.
Sheldon, E. (2011). Rhetorical differences in RA introductions written by English L1 and L2 and Castilian Spanish L1 writers. *Journal of English for Academic Purposes, 10*, 238–251.
Swales, J. M. (1990). *Genre analysis: English in academic and research settings*. Cambridge: Cambridge University Press.
Swales, J. M. (2004). *Research genres: Exploration and applications*. Cambridge: Cambridge University Press.
Taylor, C. (2002). *The RCBN consultation exercise: Stakeholder report*. Cardiff: Cardiff University. *Cardiff University School of social sciences occasional paper 50*.
Taylor, G., & Chen, T. (1991). Linguistic, cultural, and subcultural issues in contrastive discourse analysis: Anglo-American and Chinese scientific texts. *Applied Linguistics, 12*, 319–336.
Tribble, C. (2017). Elfa vs. Genre: A new paradigm war in EAP writing instruction? *Journal of English for Academic Purposes, 25*, 30–44.
Van de Ven, A. H. (2007). *Engaged scholarship: A guide for organizational and social research*. New York: Oxford University Press.
White, P. (2009). *Developing research questions: A guide for social sciences*. New York: Palgrave.
White, P. (2013). Who's afraid of research questions? The neglect of research questions in the methods literature and a call for question-led methods teaching. *International Journal of Research and Method in Education, 36*(3), 213–227.
Yang, R., & Allison, D. (2003). Research articles in applied linguistics: Moving from results to conclusions. *English for Specific Purposes, 22*(4), 365–385.
Zheng, X., & Ruan, X. (2018). An explorative study of paper writing practice in a foreign language teachers' learning community. *Foreign Language World*, (4), 18–26.
Zheng, X., & Zuo, X. (2014). Clear-cut research questions and appropriate research methods are a must in foreign language teacher research: A critical review of the literature between 2004 and 2013 in China. *Foreign Languages and Their Teaching*, (6), 7–12.

ZIQING GONG is a postgraduate student at School of Foreign Languages and Literature, Beijing Normal University. Her main research interests include English for academic purposes and applied linguistics. E-mail: freya_gong@163.com

YONGHOU LIU is a professor at School of Foreign Languages and Literature, Beijing Normal University. His primary research interests are applied linguistics and sociolinguistics. E-mail: liuyonghou@bnu.edu.cn

YING LIU is a lecturer in Department of Foreign Languages, Lyu Liang University. Her main research interests include English for academic purposes and applied linguistics. E-mail:llgzliuying@163.com