# A Comparative Genre Analysis of AI-generated and Scholar-written Abstracts for English Review Articles in International Journals

**Abstract**: There has been growing interest in the performance and efficiency of ChatGPT in generating academic texts. However, little empirical research has been conducted on its performance in producing review article abstracts. This study adopts the genre analysis approach to investigate the rhetorical moves of review article abstracts in hard and soft science disciplines based on two self-complied corpora, respectively including 160 scholar-written abstracts from four high-impact international journals, and 160 abstracts generated by ChatGPT, with an aim to reveal the similarities and differences between human-written and AI-generated English review article abstracts. The results show significant differences between human-written and ChatGPT-generated abstracts, first in the frequency of 3 out of the 5 moves, and then in the sequential order of moves, with each type of abstracts demonstrating a preference for move sequence patterns as well as obligatory and optional elements. The two types of abstracts differ significantly in the frequency of move embedding, but share the same embedding combination patterns. These findings may deepen our understanding of ChatGPT's capabilities and limitations in generating academic texts across different disciplines, help improve the generative AI system, and then highlight the complex relationship among the structure of academic abstracts, discipline cultures and genre knowledge.
**Keywords**: ChatGPT, rhetorical move, review article abstracts, genre analysis, disciplinary variation

## 1. Introduction

Since OpenAI's release of ChatGPT in November, 2023, an increasing literature has been produced to evaluate the performance of ChatGPT on scientific writing (e.g. Benichou, 2023; Dergaa, et al., 2023; Gao et al., 2023; Huang & Tan, 2023; Salvagno et al., 2023). The evaluation is crucial as it might allow researchers to determine the effectiveness of the large language model in specific research practices (van Dis et al., 2023). Extant studies have indicated that ChatGPT might offer benefits for academic article writing (e.g. Benichou, 2023; Gao et al., 2023; Májovský, et al., 2023; Macdonald, et al., 2023).

However, the performance of ChatGPT in generating review article abstracts remains unclear. Review article abstract may constitute an essential component of academic genre (Hyland, 2000; Hyland & Diani, 2009). Review article abstract writing seems to be a demanding task. Some have argued that abstract writing could be complicated for human writers due to potential time constraints (El-Kassas, 2023) and difficulties faced by language learners (Al Fadda, 2012). Therefore, generating review article abstracts could pose a significant challenge for ChatGPT since the composition of review article abstracts might be influenced by factors such as knowledge of abstract move structure and disciplinary cultures.

This study aims to assess ChatGPT's efficacy in composing abstracts for review articles by comparing the rhetorical moves of review article abstracts respectively generated by ChatGPT and composed by human writers across four social and natural science disciplines, with an aim to deepen our understanding of ChatGPT's performance in academic writing and provide valuable insights for refining and optimizing the generative AI technology by illuminating the intricate associations between rhetorical moves and disciplinary cultures in academic text generation.

## 2. Literature Review

### 2.1 Review Article Abstracts Writing

An abstract refers to "a brief, comprehensive summary of the contents of the paper" (American Psychological Association, 2020, p.38). Review article abstracts are an essential part of review articles, which have the explicit purpose of evaluating the research and benefiting the construction of knowledge of various disciplines as well as facilitating the social cohesiveness of scholarly communities (Hyland, 2000; Hyland & Diani, 2009). High-quality abstracts should be concise, self-sufficient, unbiased, and provide complete and accurate information (Bavdekar & Gogtay, 2015).

Abstract writing would be highly important, for well-written abstracts could help researchers to appeal to readers and promote their studies. Abstracts could shape readers' decisions on whether the accompanying article merits further reading (Hyland, 2000). Abstract writing may also be challenging and time-consuming. Academic writing could be a sophisticated and effortful skill, especially for language learners (Al Fadda, 2012; Al-Mukdad, 2019; Ma, 2021; Biber & Gray, 2010; Evans & Green, 2007). For example, English academic writing may be inherently difficult for most writers, particularly for English as a second language (L2) and English as foreign language (EFL) students (Lin & Morrison, 2021). Similarly, acquiring proficiency in academic writing could be a significant hurdle for freshmen (Horstmanshof & Brownie, 2013). Furthermore, non-native English students might perceive academic writing as "a formidable and daunting academic undertaking" mainly because of sociocultural challenges and linguistic difficulties (Bakhou & Bouhania 2020, p. 254). Abstract writing may be influenced by disciplinary cultures. Academic writing, the formation of knowledge, the cultures of various disciplines and social interactions are all closely interlinked (Hyland, 2000).

### 2.2 Genre and Move Analysis

Genre is generally accepted as a rhetorical, socially contextual category of discourse in the field of discourse analysis (Bhatia, 1993; Martin, 2009; Swales, 1990, 2004; Hyland, 2015). The defining characteristic of a genre was its communicative purpose (Swales, 1990). Genres processed recurrent configurations of cultural meanings with social purposes that would be achieved progressively (Martin, 2009). Moreover, genres, as community-based actions, might shape persuasive arguments and reinforce group identity and disciplinary practices (Hyland, 2015).

Genre analysis is generally recognized as a widely used method for examining the structure and purpose of non-literary genres. It has a set of analytic methods used for examining various genres within situated contexts (Tardy, 2011). Additionally, it could serve as a tool for improving educational methodologies in areas like rhetoric, composition studies, professional writing, English for specific purposes (ESP), etc. (Hyon, 1996).

John Swales has made a significant contribution to genre analysis of academic discourses. In his landmark work *Genre Analysis*, Swales (1990) introduced the CARS (Create a Research Space) model for investigating the move structure of article introductions, and the IMRD (i.e. Introduction–Method–Results–Discussion) model for scrutinizing the move structure of full-length empirical research articles (RA). He advocated for a genre-based approach to English for Academic Purposes (EAP) and showed how research and pedagogical application can be intertwined (Aull & Swales, 2015).

Move analysis is a widely used genre analysis approach. It primarily focuses on the rhetorical moves of a genre. A move usually refers to a coherent unit of speech or writing that serves a specific communicative purpose in a conversation or discourse (Swales, 2004). When conducting move analysis, researchers could first identify general structure patterns and initial move categories from a representative text corpus, then determine individual rhetorical moves, distinguishing between obligatory and optional moves, and the possibilities for the sequence of moves (Tardy & Swales, 2014). This analysis, either qualitative or quantitative, may extend to the analysis of the "steps" or sub-moves of each move. In addition, it is crucial to note that moves are basically considered rhetorical categories, not grammatical ones and a move may vary in size and be potentially realized non-verbally.

**2.3 Move Structure of Academic Paper Abstracts**

The move structure of academic paper abstracts refers to the rhetorical discourse units of the abstracts. The structure may be diversified. Different move structure of academic paper abstracts have been found, such as 4-move pattern (i.e. [Introducing purpose], [Describing methodology], [Summarizing results], [Presenting conclusions]) (Bhatia, 1993), 5-move pattern (i.e. [Situating the research], [Presenting the research], [Describing the methodology], [Summarizing the results], [Discussing the research] (Santos, 1996), and 6-move pattern (i.e. [Setting the background for the study and its importance], [Mentioning gaps addressed by the study], [Stating the purpose, question, hypothesis of study], [Identifying the methodology], [Stating the findings or argument], [Offering concluding remarks]) (Tessuto, 2015). Besides, previous studies have shown that moves of academic article abstracts could be classified as conventional and optional (Tanko, 2017). Furthermore, move embedding and the reversal of the sequential order of moves might be two typical characteristics of academic paper abstracts (Santos, 1993; Bhatia, 1993).

**2.4 The Application of ChatGPT in Academic Writing**

ChatGPT, referring to Chat Generative Pre-Trained Transformer, is a large language model (LLM). It allows users to interact with it using prompts and is capable of generating high-quality texts. ChatGPT is built upon GPT-3.5, one of the most extensive large language models to date, with over 175 billion parameters (Shen et al., 2023). Its training data consist of a wide variety of internet texts, such as books, articles, websites, etc. Particularly, it utilizes reinforcement learning from human feedback for training. This method helps ChatGPT to fine-tune its responses based on human evaluators' inputs. Consequently, ChatGPT has greatly improved in understanding users' prompts and generating human-like texts.

Previous studies have shown that ChatGPT could be used in various academic writing tasks, such as drafting research articles (e.g. Benichou, 2023; Májovský, et al., 2023; Macdonald, et al., 2023), generating academic article and conference abstracts (e.g. Babl & Babl, 2023; Gao et al., 2023), writing peer review reports (Hosseini & Horbach, 2023), and producing scientific case reports (Buholayka, et al., 2023). For instance, evaluating the article abstracts generated by ChatGPT for 50 scientific medical papers suggested that the lexicogrammatical expressions of the abstracts generated by ChatGPT seemed to be more formulaic vis-à-vis those of human-written abstracts (Gao et al., 2023). In addition, the text generated by ChatGPT seems to be more generic and less tailored to a specific context. The output from ChatGPT may not fully emulate human-like text, occasionally displaying repetitive phrases or words, along with unusual or inconsistent language usage (Cotton et al., 2023). Furthermore, ChatGPT can be used to help editors with repetitive and tedious tasks, promote the spread of new research ideas, and save time and resources by drafting findings and formatting papers according to journal standards. On the other hand, the ethical issues with ChatGPT cannot be ignored, including authorship ambiguity, plagiarism, and missing references in its outputs (Lund et al., 2023).

### 2.5 Existing Gaps and the Present Study

Extant literature has recognized that review article appears be a key genre in academic fields (Hyland, 2000; Hyland & Diani, 2009) and abstract writing was considered important (Hyland (2000) and challenging for most writers, particularly for L2 and EFL learners (Al Fadda, 2012; Al-Mukdad, 2019; Ma, 2021; Biber & Gray, 2010; Evans & Green, 2007). Besides, previous studies have shown that ChatGPT could be used in many academic writing tasks (e.g. Benichou, 2023; Májovský, et al., 2023; Macdonald, et al., 2023) and be conducive to the abstract writing of research papers (Gao et al. 2023). However, what is not yet clearly understood is the actual performance of ChatGPT in generating review article abstracts.

In order to address this research gap, we conducted a comparative genre analysis using two self-compiled corpora—the Corpus of Human-Generated Review Article Abstracts and the Corpus of ChatGPT-Generated Review Article Abstracts. The following four questions were attempted to be answered:

1) What is the overall distribution of moves across the two corpora?
2) What is the distribution of moves across the four disciplines across the two corpora?
3) What is the sequential order of the moves across the two corpora?
4) What is the move embedding across the two corpora?

## 3. Research Methodology

### 3.1 Data Collection and Corpus Compilation

The data used in the current study came from two self-complied corpora: the Corpus of Human-Generated Review Article Abstracts (henceforth CHGRAA) and the Corpus of ChatGPT-Generated Review Article Abstracts (henceforth CCGRAA). CHGRAA consisted of the abstracts of 160 review articles published in major journals of the following four social and natural science disciplines: Anthropology, Education, Engineering, and Physics (Becher, 1994). One journal was selected in each discipline based on Scimago Journal Rank (SJR) indicator in 2022. Forty review articles were selected in each journal in the period of 2018-2022. CCGRAA included 160 review article abstracts, which were generated by ChatGPT merely based on the content (without pictures, tables and figures)[1] of the 160 human-written review articles. The corpus information was provided in Table 1. As shown in Table 1, the number of tokens in each discipline was largely comparable.

**Table 1 The information of CHGRAA and CCGRAA**

| Discipline | Journal | Corpus | Texts | Tokens | MW | TTR | AWL | ASL |
|---|---|---|---|---|---|---|---|---|
| Anthropology | Annual Review of Anthropology | CHGRAA | 40 | 6067 | 151 | 31.01% | 5.84 | 27.40 |
| | | CCGRAA | 40 | 5592 | 139 | 26.91% | 6.44 | ·21.29 |
| Education | Educational Research Review | CHGRAA | 40 | 7211 | 177 | 22.32% | 5.83 | 23.03 |
| | | CCGRAA | 40 | 6635 | 165 | 20.77% | 6.39 | 19.82 |
| Engineering | Materials Science and Engineering: C | CHGRAA | 40 | 6719 | 167 | 25.41% | 5.98 | 25.03 |
| | | CCGRAA | 40 | 6328 | 157 | 22.35% | 6.54 | 18.37 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Physics | Nature Reviews Physics | CHGRAA | 40 | 6574 | 163 | 27.88% | 5.82 | 25.85 |
| | | CCGRAA | 40 | 6053 | 150 | 24.98% | 6.43 | 21.10 |
| Total | | CHGRAA | **160** | **26571** | **165** | **17.81%** | **5.87** | **25.14** |
| | | CCGRAA | **160** | **24608** | **153** | **16.13%** | **6.45** | **20.03** |

**Note**: CHGRAA refers to the Corpus of Human-Generated Review Article Abstracts. CCGRAA refers to the Corpus of ChatGPT-Generated Review Article Abstracts. MW, TTR, AWL and MWL refer to the average words of abstracts, type/token ratio, and average word length and average sentence length, respectively.

In the generation process, the review article was divided into several sections. This was mainly due to the character input limit for each interaction with the ChatGPT 3.5 version. As for the abstract generation, the following prompts were fed to the 3.5 model, the details of which were as follows:

(1) 我接下来要在多个对话中，提供给你用「@」编号的文章内容，格式为@i=<内容>，请先记住它们，不要提供各自的摘要，可以吗？(English translation: I will provide article content marked with「@」numbers in the format of @i=<the content> in the following conversations. Please remember them first without providing summaries for each, okay?)

(2) 请记住@i，不要提供摘要: (English translation: Please remember @i and do not provide a summary:)(i=1, 2, 3, …n-1)

(3) 请为@n 到@m 的内容生成一份符合期刊综述论文格式的英文摘要，字数不要超过……个单词。(English translation: Please generate an English abstract for the content from @n to @m in the format of a journal review paper, and the word count should not exceed ... words.[1])

**3.2 Approach to Move Analysis**

The initial step in our research was to reliably annotate the sub-moves of the review article abstracts. To ensure the reliability and validity issues associated with the form-oriented methodology (Lu, et al., 2021), a function-oriented approach was implemented, which implied the annotation was performed based solely on the function or content of the abstracts in this study. The entire corpus was annotated for moves and sub-moves, and then the frequency and sequential pattern of moves, and move embedding were examined. In addition, when the frequency of moves was calculated, consideration was given to whether a move was obligatory or optional if it appeared in 90% of 160 review article abstracts (Santos, 1996). Specifically, a move that occurs in 90% of the 160 abstracts in the corpus was labeled as obligatory while a move that occurs below 90% was counted as being optional.

**3.3 The Adapted Framework for Move Analysis of Review Article Abstracts**

The objectives of the analysis were twofold: to establish a comprehensive analytical framework, and then to use the framework to conduct reliable annotation of moves within our corpora. Santos' (1996) five-move structure model was adopted as the base framework (as shown

---

[1] The pictures, tables and figures in original human-written review article abstracts cannot be fed to ChatGPT 3.5 version, so those data have been deleted when we complied the Corpus of ChatGPT-Generated Review Article Abstracts.

[1] The maximum word limit for each abstract generated by ChatGPT was the same as the maximum word limit for human-written abstracts. For example, if one human-written abstract had 200 words, the maximum word limit for ChatGPT-generated abstracts was also 200 words.

in Table 2).

**Table 2 Santos' (1996) five-move structure model**

| Move | Move structure |
|---|---|
| **Move 1** | **Situating the research**<br>Sub-move 1A: Stating current knowledge<br>Sub-move 1B: Citing previous research<br>Sub-move 1C: Extended previous research<br>Sub-move 2: Stating a problem |
| **Move 2** | **Presenting the research**<br>Sub-move 1A: Indicating main features<br>Sub-move 1B: Indicating main purpose<br>Sub-move 2: Hypothesis raising |
| **Move 3** | **Describing the methodology** |
| **Move 4** | **Summarizing the results** |
| **Move 5** | **Discussing the research**<br>Sub-move 1: Drawing conclusions<br>Sub-move 2: Giving recommendations |

The model has been widely acknowledged within the EAP community (Biber & Conrad, 2019; Jiang & Hyland, 2017; Omidian, Shahriari & Siyanova-Chanturia, 2018; Pho, 2008) as being adept at covering the move structure of both structured and unstructured abstracts (Hartley, Sydes & Blurton, 1996). However, we anticipated that some adjustments to this model would be necessary due to the interdisciplinary variations present in the review article abstract writing practices found in our data. The final analytical framework was provided in Table 3. In order to ensure the reliability of the adapted framework and its application in the annotation, our analysis was carried out in two stages.

**Table 3 A modified move structure analysis model for review article abstracts**

| Move/ Move structure | Discourse Function |
|---|---|
| **Move 1 (M1): Situating the research** | |
| Sub-move 1A (SM1A): Stating current knowledge | Describing the current state or the background information of the research topic |
| Sub-move 1B (SM1B): Stating a problem | Indicating a knowledge gap, making counter-claims, or presenting controversial arguments |
| **Move 2 (M2): Presenting the research** | |
| Sub-move 2A SM2A): Indicating main features | Describing the scope of the paper, summarizing the content, or discussing the interrelated subtopics mentioned in the article. |
| Sub-move 2B SM2B): Indicating main purpose | Indicating the purpose of the research |
| **Move 3 (M3): Describing the methodology** | Describing the theories, analytical frameworks, methods, participants, corpus data, software, tools, etc. used in the research |
| **Move 4 (M4): Summarizing the results** | Describing the research results or findings objectively, which may include the data, the author's subjective opinions or interpretations based on the research |

| | |
|---|---|
| **Move 5 (M5): Discussing the research** | |
| Sub-move 5A (SM5A): Drawing conclusions | Stating the value of the current research, as well as the practical and theoretical significance of the research |
| Sub-move 5B (SM5B): Giving recommendations | Suggesting the potential applications of the research in the future or providing recommendations for future research |

During the initial phase, three annotators[1] independently analyzed a random selection of 32 abstracts, which represented 10% of the total 320 abstracts collected. They tried to apply Santos' (1996) move structure model to the texts, and held several extensive meetings to discuss changes to the move structure model and to reconcile any differences in their individual analyses of the 32 abstracts. It was found that minor modifications were required for Move 1 (**Contextualizing the research**) and Move 2 (**Introducing the research**), with the removal of three sub-moves since they were not present in the corpus data. In the subsequent phase, inter-rater reliability at the sub-move level was assessed to determine the level of subjectivity in the annotation process. The three pairs achieved an average **Cohen's Kappa score of 0.978 (SPSS 23)**, indicating a high level of agreement. The remaining texts were then analyzed by one author and carefully reviewed by another author. All differences identified were collaboratively resolved by the four researchers in a series of extensive team discussions.

During the whole process, we adhered to Swales' (2004) definition of moves as "discoursal or rhetorical units performing coherent communicative functions in texts" (pp. 228–229), referring to moves as functional rather than formal units. Also, the analysis of move structure has been utilized to examine the significance of specific moves (whether they were obligatory or optional), the sequential positioning of these moves, and their embedding (Samraj, 2014). Therefore, these three characteristics of moves have been integrated into our adapted framework for the analysis of move structure of review article abstracts  (see Figure 1 ).

| Dimension 1 | Dimension 2 | Dimension 3 |
|---|---|---|
| • the distribution of moves<br>• the distribution of sub-moves | • the sequential order of moves<br>• the sequential order of sunb-moves | • the embedding of moves |

**Figure 1 The Adapted Framework for Move Analysis of Review Article Abstracts**

## 4. Results and Analysis

### 4.1 The Overall Distribution of Moves across the Two Corpora

Table 4 summarizes the distribution of moves and sub-moves across the two corpora, i.e. the ChatGPT-generated abstracts and the human-written abstracts.

[1] In the project, apart from the two authors, our team also included two doctoral students, Zhao Wen and Xu Wei, specializing in education and applied linguistics respectively. Zhao and Xu independently analyzed a random selection of 32 abstracts.

**Table 4 The overall distribution of moves across the two corpora**

| Move | F(CC) | F(CH) | P(CC) | P(CH) | χ2 | S (p) | | |
|---|---|---|---|---|---|---|---|---|
| M1 | **73** | **170** | **14.87%** | **33.73%** | **46.92** | **0.000** | *** | - |
| M2 | 172 | 171 | 35.03% | 33.73% | 0.13 | 0.715 | | |
| M3 | **24** | **45** | **4.89%** | **8.93%** | **5.68** | **0.017** | * | - |
| M4 | 68 | 60 | 13.85% | 11.90% | 0.67 | 0.412 | | |
| M5 | **154** | **58** | **31.36%** | **11.51%** | **57.31** | **0.000** | *** | + |
| Total | 491 | 504 | | | | | | |

***Note***: F(CC) and F(CH) refer to the move frequency in the Corpus of ChatGPT-Generated Review Article Abstracts (CCGRAA) and the Corpus of Human-Generated Review Article Abstracts (CHGRAA), respectively. P(CC) and P(CH) represent the percentages of move usage in two different corpora, CCGRAA and CHGRAA, respectively. χ2 refers to the Chi-square test value. S(p) stands for the p value. The asterisks (*) indicate significance level: (*) means statistically significant at the 0.05 level; (**) means statistically significant at the 0.01 level; (***) means statistically significant at the 0.001 level. The "+" and "-" signs on the right side indicate "overuse" and "underuse".

As evident from Table 4, the Chi-square test[1] showed statistically significant differences in the frequency of occurrence between Move 1, Move 3, and Move 5 across the Corpus of ChatGPT-Generated Review Article Abstracts (CCGRAA) and the Corpus of Human-Generated Review Article Abstracts (CHGRAA). This finding indicates that the overall distribution of the moves in review article abstract across the two corpora is fairly different. Specifically, Move 1 (**Situating the research**), Move 3 (**Describing the methodology**), and Move 5 (**Discussing the research)** occurred with significantly different frequencies in the two types of abstracts. For example:

(1) *Drawing on various scholarly perspectives, including linguistic anthropology and related fields, the review explores how individuals' recognition of vulnerability and their emotional responses shaped new forms of social engagement.* **[Move 3: Describing the methodology**] (Anthropology by ChatGPT)

The findings suggest that there are discernible and statistically significant patterns in the ways that ChatGPT and humans write abstracts. In other words, even though both humans and ChatGPT were capable of writing review article abstracts, their respective methodologies were not identical. There existed substantial and statistically significant differences in their writing styles or strategies, as evidenced by the different frequencies of usage of Move 1, Move 3, and Move 5.

However, no significant differences were found in Move 2 (**Presenting the research**) and Move 4 (**Summarizing the results) across the two corpora.** This indicates both humans and ChatGPT might be more likely to use similar move structures when presenting research and summarizing results in review articles abstracts. Table 5 provides an overview of the distribution of sub-moves across the two corpora.

[1] The Chi-square test was run by Chi-square and Log-Likelihood Calculator. This tool was developed by Liang Maocheng, National Research Center for Foreign Language Education, Beijing Foreign Studies University, People's Republic of China. The chi-square test used in this program is the classic chi-square test of significance with Yates correction for 2 X 2 tables. The tool is available online and the downloading link is http://corpus.bfsu.edu.cn/LLX2.zip.

**Table 5 The distribution of sub-moves across the two corpora**

| Move | SM | F(CC) | F(CH) | P(CC) | P(CH) | χ2 | S (p) | | |
|---|---|---|---|---|---|---|---|---|---|
| **M1** | SM1A | **73** | **139** | **14.87%** | **27.58%** | **23.22** | **0.000** | *** | - |
| | SM1B | **0** | **31** | **0.00%** | **6.15%** | **29.17** | **0.000** | *** | - |
| **M2** | SM2A | **159** | **131** | **32.38%** | **25.99%** | **4.61** | **0.032** | * | + |
| | SM2B | **13** | **40** | **2.65%** | **7.94%** | **12.77** | **0.000** | *** | - |
| **M3** | | **24** | **45** | **4.89%** | **8.93%** | **5.68** | **0.017** | * | - |
| **M4** | | 68 | 60 | 13.85% | 11.90% | 0.67 | 0.412 | | |
| **M5** | SM5A | **81** | **20** | **16.50%** | **3.97%** | **41.44** | **0.000** | *** | + |
| | SM5B | **73** | **38** | **14.87%** | **7.54%** | **12.75** | **0.000** | *** | + |

***Note***: SM refers to the sub-moves. F(CC) and F(CH) refer to the move frequency in the Corpus of ChatGPT-Generated Review Article Abstracts (CCGRAA) and the Corpus of Human-Generated Review Article Abstracts (CHGRAA), respectively. P(CC) and P(CH) represent the percentages of move usage in two different corpora, CCGRAA and CHGRAA, respectively. χ2 refers to the Chi-square test value. S(p) stands for the p value. The asterisks (*) indicate significance level: (*) means statistically significant at the 0.05 level; (**) means statistically significant at the 0.01 level; (***) means statistically significant at the 0.001 level. The "+" and "-" signs on the right side indicate "overuse" and "underuse".

Considerable variations were found in the use of sub-moves to fulfill the rhetorical functions in the two corpora. Specifically, in terms of the use of sub-move 1A (**Stating current knowledge**), sub-move 1B (**Stating a problem**) and sub-move 2B (**Indicating main purpose**), there was a significantly lower occurrence in the ChatGPT-generated abstracts, while there was a notably higher occurrence in the human-written abstracts. On the contrary, in terms of the usage of sub-move 2A (**Indicating main features**), sub-move 5A (**Drawing conclusions**) and sub-move 5B (**Giving recommendations**), there was a significantly higher occurrence in the ChatGPT-generated abstracts, while there was a notably lower occurrence in the human-written abstracts. For example:

(2) *Since the adoption of the UN Convention on the Rights of Persons with Disabilities in 2006, the educational trajectories of children with disabilities have increasingly become the focus of public interest and the subject of empirical research. The transition to school has received special attention in this research.* **[Sub-move 1A: Stating current knowledge)]** *Yet a systematic overview of current research in this area is not available.* [**Sub-move 1B : Stating a problem)**] (Education by Humans)

(3) *This review aims to give an overview on the latest and most promising strategies to improve the bioactivity of lactic acid-based materials, especially focusing on biomolecule-free bulk approaches such as blending, copolymerization or composite fabrication.* [**Sub-move 2B: Indicating main purpose)**] (Engineering by Humans)

(4) *Here we review the theory of quantum systems with ultrastrong coupling, discussing entangled ground states.... We also overview the multitude of experimental setups, including superconducting circuits... We conclude by discussing the many potential applications that these achievements enable in physics and chemistry.* [**Sub-move 2A: Indicating main features)**] (Physics by Humans)

(5) *This comprehensive review examines the relationship between local food movements and the construction of identity. The analysis encompasses various scholarly perspectives .... The discussion begins with... The review also delves into, ... Additionally, the paper examines the role of local food in addressing social issues, ... The conclusion emphasizes the complexity of defining local food ....* [**Sub-move 2A: Indicating main features)**] (Anthropology by ChatGPT)

This indicates that the algorithm used to generate review article abstracts by ChatGPT may not fully replicate human writing patterns. Specifically, ChatGPT may fail to adequately describe the background information of the review topic, to acknowledge knowledge gaps or controversial arguments, or to state the aims of the study. It may also overstate or overstress the summary or main content of the article, the value of the research and the potential applications and recommendations for future research.

Moreover, we calculated the percentage of each move of 160 abstracts in the Corpus of

ChatGPT-Generated Review Article Abstracts (CCGRAA) to determine the obligatory and optional moves. As mentioned above, for a move to be classified as obligatory, it needs to be present in 90% of the 160 abstracts in the corpus. A move might be deemed optional if its occurrence may be less than 90%. The results showed that the obligatory moves in ChatGPT-generated abstracts were Move 1 (107.50%) and Move 5 (96.25%), while the optional moves were Move 1(45.63%), Move 3 (15.00%), and Move 4 (42.50%). However, the obligatory moves in human-written abstracts were Move 1 (106.25%) and Move 2 (106.88%), while the optional moves were Move 3 (28.13%), Move 4 (37.50%), and Move 45(36.25%). This suggests that the typical order of the moves in a ChatGPT-generated or a human-written review article abstract based on the move sequence patterns identified in the corpus can be described as follows, respectively (the bracketed elements and plus signs between them describe the typical order of the functional units, the parentheses indicate optional moves):

ChatGPT-generated Abstracts: M1 [(Situating the research)] + **M2 [Presenting the research]** + M3 [(Describing the methodology)] + M4 [(Summarizing the results)] + **M5 [Discussing the research]**

Humans-generated Abstracts: **M1 [Situating the research]** + **M2 [Presenting the research]** + M3 [(Describing the methodology)] + M4 [(Summarizing the results)] + M5 [(Discussing the research)]

**4.2 The Distribution of Moves across the Four Disciplines in the Two Corpora**

Table 6 provides a synopsis of the use of moves and sub-moves in anthropology throughout the two corpora.

**Table 6 The distribution of sub-moves in anthropology across the two corpora**

| Move | SM | F(CCATH) | F(CHATH) | P(CCATH) | P(CHATH) | χ2 | S (p) | | |
|---|---|---|---|---|---|---|---|---|---|
| **M1** | SM1A | **6** | **29** | **1.22%** | **5.75%** | **13.75** | **0.000** | *** | - |
| | SM1B | **0** | **7** | **0.00%** | **1.39%** | **5.02** | **0.025** | * | - |
| **M2** | SM2A | 40 | 32 | 8.15% | 6.35% | 0.94 | 0.331 | | |
| | SM2B | 2 | 6 | 0.41% | 1.19% | 1.06 | 0.304 | | |
| **M3** | | 3 | 6 | 0.61% | 1.19% | 0.40 | 0.528 | | |
| **M4** | | **5** | **17** | **1.02%** | **3.37%** | **5.34** | **0.021** | * | - |
| **M5** | SM5A | **18** | **5** | **3.67%** | **0.99%** | **6.74** | **0.009** | ** | + |
| | SM5B | 10 | 4 | 2.04% | 0.79% | 1.95 | 0.163 | | |

***Note***: SM refers to the sub-moves. F(CCATH) and F(CHATH) refer to the move frequency in anthropology of the Corpus of ChatGPT-Generated Review Article Abstracts (CCGRAA) and the Corpus of Human-Generated Review Article Abstracts (CHGRAA), respectively. P(CCATH) and P(CHATH) represent the percentages of move usage in anthropology of two different corpora, CCGRAA and CHGRAA, respectively.

As demonstrated in Table 6, there were substantial differences in the application of moves and sub-moves within the discipline of anthropology. In specific terms, the use of sub-move 1A (**Stating current knowledge**), sub-move 1B (**Stating a problem**), and Move 4 (**Summarizing the results**), was noticeably less frequent in the ChatGPT-created abstracts compared to the human-written ones, which showed a substantially higher occurrence of these elements. Conversely, the utilization of sub-move 5A (**Drawing conclusions**) was distinctly more prevalent in the ChatGPT-produced abstracts, while it was considerably less common in the human-composed ones. This implies that in generating review article abstracts in the field of anthropology, ChatGPT may not be sufficient in accurately depicting the research topic's background information, acknowledging contentious arguments or gaps in knowledge, or in summarizing the results. Additionally, it might over-highlight the significance of the research.

A summary of how sub-moves were distributed within the two kinds of review article abstract corpora in education is provided in Table 7.

**Table 7 The distribution of sub-moves in education across the two corpora**

| Move | SM | F(CCEDU) | F(CHEDU) | P(CC) | P(CH) | χ2 | S (p) | | |
|---|---|---|---|---|---|---|---|---|---|
| **M1** | SM1A | **4** | **30** | **0.81%** | **5.95%** | **8.37** | **0.000** | *** | - |
| | SM1B | **0** | **19** | **0.00%** | **3.77%** | **16.91** | **0.000** | *** | - |
| **M2** | SM2A | 39 | 24 | 7.94% | 4.76% | 3.72 | 0.054 | | |
| | SM2B | **9** | **25** | **1.83%** | **4.96%** | **6.45** | **0.011** | * | - |
| **M3** | | **20** | **39** | **4.07%** | **7.74%** | **5.35** | **0.021** | * | - |
| **M4** | | 34 | 32 | 6.92% | 6.35% | 0.06 | 0.812 | | |
| **M5** | SM5A | 19 | 13 | 3.87% | 2.58% | 0.95 | 0.330 | | |
| | SM5B | 36 | 24 | 7.33% | 4.76% | 2.46 | 0.117 | | |

***Note***: SM refers to the sub-moves. F(CCEDU) and F(CHEDU) refer to the move frequency in education of the Corpus of ChatGPT-Generated Review Article Abstracts (CCGRAA) and the Corpus of Human-Generated Review Article Abstracts (CHGRAA), respectively. P(CCATH) and P(CHATH) represent the percentages of move usage in education of two different corpora, CCGRAA and CHGRAA, respectively.

As depicted in Table 7, there were significant differences in the utilization of moves and sub-moves in the field of education. In specific terms, the application of sub-move 1A (**Stating current knowledge**), sub-move 1B (**Stating a problem**), sub-move 2B (**Indicating main purpose**), and Move 3 (**Describing the methodology**) was noticeably less frequent in the abstracts generated by ChatGPT, while these elements were considerably more common in the abstracts created by humans. This suggests that in the creation of review article abstracts in education, ChatGPT may face difficulties in offering relevant background information on the research subject, recognizing knowledge deficiencies, efficiently summarizing the results, or possibly neglecting the methodology.

Table 8 provides a summary of how sub-moves were distributed within the two corpora of physics review article abstracts.

**Table 8 The distribution of sub-moves in physics across the two corpora**

| Move | SM | F(CCPHY) | F(CHPHY) | P(CC) | P(CH) | χ2 | S (p) | | |
|---|---|---|---|---|---|---|---|---|---|
| **M1** | SM1A | 30 | 40 | 6.11% | 7.94% | 1.00 | 0.316 | | |
| | SM1B | 0 | 1 | 0.00% | 0.20% | 0.00 | 0.990 | | |
| **M2** | SM2A | 40 | 37 | 8.15% | 7.34% | 0.13 | 0.721 | | |
| | SM2B | 1 | 2 | 0.20% | 0.40% | 0.00 | 0.982 | | |
| **M3** | | 0 | 0 | 0.00% | 0.00% | 0.00 | 1.000 | | |
| **M4** | | 10 | 5 | 2.04% | 0.99% | 1.19 | 0.275 | | |
| **M5** | SM5A | **25** | **1** | **5.09%** | **0.20%** | **21.52** | **0.000** | *** | + |
| | SM5B | 12 | 5 | 2.44% | 0.99% | 2.32 | 0.128 | | |

***Note***: SM refers to the sub-moves. F(CCPHY) and F(CHPHY) refer to the move frequency in physics of the Corpus of ChatGPT-Generated Review Article Abstracts (CCGRAA) and the Corpus of Human-Generated Review Article Abstracts (CHGRAA), respectively. P(CCATH) and P(CHATH) represent the percentages of move usage in physics of two different corpora, CCGRAA and CHGRAA, respectively.

As illustrated in Table 8, there were significant disparities merely in the employment of sub-moves in the field of physics. Specifically, with regards to the application of sub-move 5A (**Drawing conclusions**), a significantly greater frequency was observed in the ChatGPT-generated abstracts, in contrast to a notably lesser frequency in abstracts written by humans. This implies that the large language model (ChatGPT) may tend to focus more on drawing conclusions compared to human-written abstracts when generating physics review article abstracts. For example:

(6) *Overall, nanotechnology holds significant potential for revolutionizing orthopedic implant technology in the near future.* [**Sub-move 5A: Drawing conclusions)**] (Engineering by ChatGPT)

Table 9 presents an overview of the distribution of moves and sub-moves between the two kinds of engineering review article abstract corpora.

**Table 9 The distribution of sub-moves in engineering across the two corpora**

| Move | SM | F(CCENG) | F(CHENG) | P(CC) | P(CH) | χ2 | S (p) | | |
|---|---|---|---|---|---|---|---|---|---|
| **M1** | SM1A | 33 | 40 | 6.72% | 7.94% | 0.38 | 0.539 | | |
| | SM1B | 0 | 4 | 0.00% | 0.79% | 2.18 | 0.140 | | |
| **M2** | SM2A | 40 | 38 | 8.15% | 7.54% | 0.06 | 0.812 | | |
| | SM2B | 1 | 7 | 0.20% | 1.39% | 3.02 | 0.082 | | |
| **M3** | | 1 | 0 | 0.20% | 0.00% | 0.00 | 0.990 | | |
| **M4** | | **19** | **6** | **3.87%** | **1.19%** | **6.24** | **0.013** | * | + |
| **M5** | SM5A | **19** | **1** | **3.87%** | **0.20%** | **15.21** | **0.000** | ** | + |
| | SM5B | **15** | **5** | **3.06%** | **0.99%** | **4.38** | **0.036** | * | + |

***Note***: SM refers to the sub-moves. F(CCATH) and F(CHATH) refer to the move frequency in engineering of the Corpus of ChatGPT-Generated Review Article Abstracts (CCGRAA) and the Corpus of Human-Generated Review Article Abstracts (CHGRAA), respectively. P(CCATH) and P(CHATH) represent the percentages of move usage in engineering of two different corpora, CCGRAA and CHGRAA, respectively.

As illustrated in Table 9, there were notable differences in the application of moves and sub-moves in the field of engineering. Specifically, when it comes to the application of Move4 (**Summarizing the results**), sub-move 5A (**Drawing conclusions**), and sub-move 5B (**Giving recommendations**), a significantly greater frequency was observed in the ChatGPT-generated abstracts, while there was a markedly lower frequency in the abstracts written by humans. For example:

(7) *Additionally, the use of predictive systems based on neural networks is recommended for a comprehensive understanding of the SES-academic performance relationship in higher education.* [***Sub-move 5B: Giving recommendations***] (Education by ChatGPT)

This indicates that ChatGPT-generated abstracts might be more likely to summarize results, draw conclusions, and provide recommendations more frequently than human-written abstracts in engineering review article abstracts. This discrepancy could potentially reflect differences in stylistic or structural preferences between the large language model and human writers in producing review article abstracts in the field of engineering.

In conclusion, these findings suggest that the large language model (ChatGPT) may exhibit limitations when generating review article abstracts across various fields. In anthropology and education, ChatGPT might not sufficiently depict the research topic's background, acknowledge contentious arguments or knowledge gaps, and may overstate the research's significance. Meanwhile, in physics and engineering, ChatGPT appears to emphasize drawing conclusions, summarizing results, and offering recommendations over other aspects, potentially indicating stylistic or structural differences between the large language model and human writing.

### 4.3 The Sequential Order of the Moves and the Sub-moves Across CCGRAA and CHGRAA in RA Abstracts

Table 10 provides a summary of move sequences across CCGRAA and CHGRAA in review article abstracts.

**Table 10 The classification of move sequence patterns across the two corpora**

| | CCGRAA | F(CC) | P(CC) | CHGRAA | F(CH) | P(CH) |
|---|---|---|---|---|---|---|
| 1 | M1-M2 | 14 | 8.75% | M2 | 5 | 3.13% |
| 2 | M2 | 8 | 5.00% | M4 | 1 | 0.63% |
| 3 | M1-M2-M3-M4-M5 | 2 | 1.25% | M1-M2 | 69 | **43.13%** |
| 4 | M1-M2-M4 | 1 | 0.63% | M1-M2-M3-M2 | 1 | 0.63% |
| 5 | M1-M2-M4-M5 | 23 | 14.38% | M1- M2-M3-M2-M4 | 1 | 0.63% |
| 6 | M1-M2-M5 | 32 | **20.00%** | M1-M2-M3-M2-M4-M5 | 1 | 0.63% |
| 7 | M1-M3-M2-M4-M5 | 1 | 0.63% | M1-M2-M3-M2-M5 | 3 | 1.88% |
| 8 | M2-M5 | 33 | **20.63%** | M1-M2-M3-M4 | 5 | 3.13% |
| 9 | M2-M3-M2-M4-M5 | 2 | 1.25% | M1-M2-M3-M4-M5 | 18 | **11.25%** |
| 10 | M2-M3-M4-M5 | 14 | 8.75% | M1-M2-M4 | 10 | 6.25% |
| 11 | M2-M4 | 8 | 5.00% | M1-M2-M4-M5 | 2 | 1.25% |
| 12 | M2-M4-M5 | 17 | 10.63% | M1-M2-M5 | 14 | 8.75% |
| 13 | M3-M2-M4 | 1 | 0.63% | M1-M3-M2 | 4 | 2.50% |
| 14 | M3-M2-M5 | 4 | 2.50% | M1-M3-M2-M5 | 2 | 1.25% |
| 15 | | | | M1-M3-M4 | 1 | 0.63% |
| 16 | | | | M1-M3-M4-M5 | 1 | 0.63% |
| 17 | | | | M1-M4 | 7 | 4.38% |
| 18 | | | | M1-M4-M5 | 1 | 0.63% |
| 19 | | | | M2-M3-M2 | 1 | 0.63% |
| 20 | | | | M2-M3-M2-M4-M5 | 1 | 0.63% |
| 21 | | | | M2-M3-M4 | 1 | 0.63% |
| 22 | | | | M2-M3-M4-M2 | 1 | 0.63% |
| 23 | | | | M2-M3-M4-M5 | 5 | 3.13% |
| 24 | | | | M2-M4 | 3 | 1.88% |
| 25 | | | | M2-M4-M5 | 1 | 0.63% |
| 26 | | | | M2-M5 | 1 | 0.63% |
| Total | | 160 | | | 160 | |

***Note***: F(CC) and F(CH) refer to the frequency of move sequence patterns in the Corpus of ChatGPT-Generated Review Article Abstracts (CCGRAA) and the Corpus of Human-Generated Review Article Abstracts (CHGRAA), respectively. P(CC) and P(CH) represent the percentages of move sequence patterns in two different corpora, CCGRAA and CHGRAA, respectively.

It was evident that a typical abstract written by a human may follow a 2-move sequence (M1-M2) or a 5-move sequence (M1-M2-M3-M4-M5). In contrast, a typical abstract generated by ChatGPT may adhere to a 2-move sequence (M2-M5) or a 3-move sequence (M1-M2-M5). For example:

(8) *Despite the buzz around gamification as an exciting new method to engage students, evidence of its ability to enhance learning is mixed. [**M1**] …Therefore, in order to make the case for or against gamification in education, it is important to examine … student learning achievements. [**M2**] This study is a meta-analysis of 30 independent interventions (3,202 participants) drawn from 24 quantitative studies… [**M3**] The results show an overall significant medium effect size … An analysis of 32 qualitative studies reveals four reasons … (b) gamification can cause anxiety or jealousy. [**M4**] We conclude by highlighting two unresolved questions, and suggesting several future research directions concerning gamification in educational contexts.* [**M5**] (Education by Humans)

(9) *This comprehensive review examines … their SJRt on human behavior. The reviewed studies emphasize the significance… Additionally, this review underscores the dynamic nature of social ecologies ….* [**M2**] *Ultimately, this psychoanalytically-informed anthropological research contributes to a more nuanced understanding … emphasizing the transformative power of human actions and imaginative capacities.* **[M5]** (Anthropology by ChatGPT)

This suggests that the machine-generated (ChatGPT) abstracts and human-authored abstracts might exhibit different move structure patterns. The humans were more likely to use a more elaborate sequence, incorporating all five moves in some cases. On the other hand, the large language model seemed to favor simpler, shorter move sequences, possibly omitting or overlooking certain moves or sub-moves. This indicates that the large language model might not fully capture the complexity and nuance often present in human-written review article abstracts.

Furthermore, fourteen distinct move sequence patterns were observed in the ChatGPT-generated abstracts, while the human-written abstracts displayed 26 move sequence patterns. This suggests that the human-written abstracts could demonstrate a greater diversity in their move sequence patterns compared to the ChatGPT-generated abstracts. This might be due to human authors' creative ability to adapt their writing style for different social contexts, or communicative purposes, whereas the machine-generated text might be more rigid or limited due to the programming or the sparsity of training data encountered by the large language model.

Table 11 displays the results obtained from the move sequence pattern analysis in two hard science disciplines, i.e., physics and engineering in CHGRAA.

**Table 11 The classification of move sequence patterns in physics and engineering in CHGRAA**

| | Physics | F(PH) | P(PH) | Engineering | F(EN) | P(EN) |
|---|---|---|---|---|---|---|
| 1 | SM1A- SM1B- SM2A | 1 | 2.50% | SM1A- SM1B- SM2A | 3 | 7.50% |
| 2 | SM1A- SM2A | 26 | **65.00%** | SM1A- SM1B- SM2A-SM5A | 1 | 2.50% |
| 3 | SM1A- SM2A- SM2B | 1 | 2.50% | SM1A- SM2A | 23 | **57.50%** |
| 4 | SM1A- SM2A-M4 | 2 | 5.00% | SM1A- SM2A- SM2B | 1 | 2.50% |
| 5 | SM1A- SM2A- SM5A | 1 | 2.50% | SM1A- SM2A-M4 | 2 | 5.00% |
| 6 | SM1A- SM2A- SM5B | 5 | 12.50% | SM1A- SM2A-M4-SM5B | 2 | 5.00% |
| 7 | SM1A-M3- SM2B- SM2A | 1 | 2.50% | SM1A- SM2A- SM5B | 1 | 2.50% |
| 8 | SM1A-M4 | 3 | 7.50% | SM1A- SM2B- SM2A | 4 | 10.00% |
| 9 | | | | SM1A- SM2B- SM2A-M4 | 1 | 2.50% |
| 10 | | | | SM1A- SM2B- SM5B | 1 | 2.50% |
| 11 | | | | SM1A-M4- SM5B | 1 | 2.50% |
| Total | | 40 | | | 40 | |

***Note***: F(PH) and F(EN) refer to the frequency of move sequence patterns in physics and engineering of the Corpus of Human-Generated Review Article Abstracts (CHGRAA), respectively. P(PH) and P(EN) represent the percentages of move sequence patterns in physics and engineering of CHGRAA.

Eight and eleven different move sequence patterns were identified in these two disciplines, respectively. As can be seen from the table, SM1A-SM2A was the most used move sequence in these two kinds of review article abstracts. SM1A and SM2A represented the sub-move 1A (**Stating current knowledge**) and sub-move 2A (**Indicating main features**), respectively. This suggests that human scholars might prefer to present the background information and demonstrate the main content or the scope of the review articles in those two disciplines. The preference of human scholars to display that information in the abstracts might be due to the influence of disciplinary cultures, or to their knowledge of the genre of the review article. In other words, in physics and engineering, human researchers might prefer a more detailed, context-oriented approach when writing the review article abstracts.

Table 12 showcases the result of an analysis of move sequence patterns within two soft science disciplines, specifically, anthropology and education in CHGRAA.

**Table 12 The classification of move sequence patterns in anthropology and education in CHGRAA**

| | Anthropology | F(AN) | P(AN) | Education | F(ED) | P(ED) |
|---|---|---|---|---|---|---|
| 1 | SM1A, SM1B, SM2A | 3 | 7.50% | SM1A-M3-SM2A | 1 | 2.50% |
| 2 | SM1A, SM1B, SM2A, SM5B | 1 | 2.50% | SM1A-SM1B-M3-M4-SM5B | 1 | 2.50% |
| 3 | SM1A, SM1B, SM2B, M3, SM2A, SM5A | 1 | 2.50% | SM1A-SM1B-M3-SM2A-SM5B | 1 | 2.50% |
| 4 | SM1A, SM1B, SM2B, M3, M4 | 1 | 2.50% | SM1A-SM1B-M3-SM2B-SM5B | 1 | 2.50% |
| 5 | SM1A, SM1B, SM2B, M4 | 1 | 2.50% | SM1A-SM1B-SM2A-M3-M4-SM5A | 1 | 2.50% |
| 6 | **SM1A, SM2A** | **7** | **17.50%** | SM1A-SM1B-SM2A-M3-M4-SM5B | 3 | 7.50% |
| 7 | SM1A, SM2A, M4 | 3 | 7.50% | SM1A-SM1B-SM2A-M3-SM2B-M4-SM5A-SM5B | 1 | 2.50% |
| 8 | SM1A, SM2A, SM5A | 2 | 5.00% | SM1A-SM1B-SM2B-M3-M4 | 2 | 5.00% |
| 9 | SM1A, SM2A, SM5A, SM5B | 1 | 2.50% | **SM1A-SM1B-SM2B-M3-M4-SM5A** | **4** | **10.00%** |
| 10 | SM1A, SM2A, SM5B | 1 | 2.50% | SM1A-SM1B-SM2B-M3-M4-SM5B | 1 | 2.50% |
| 11 | SM1A, SM2B, M4 | 1 | 2.50% | SM1A-SM1B-SM2B-M3-SM2A | 1 | 2.50% |
| 12 | SM1A, M3, SM2A | 2 | 5.00% | SM1A-SM1B-SM2B-M3-SM2A-M4 | 1 | 2.50% |
| 13 | SM1A, M3, M4 | 1 | 2.50% | SM1A-SM1B-SM2B-M3-SM2A-SM5A | 1 | 2.50% |
| 14 | SM1A, M4 | 4 | 10.00% | SM1A-SM2A-M3-M4 | 1 | 2.50% |
| 15 | SM2A | 3 | 7.50% | SM1A-SM2A-M3-M4-SM5A | 1 | 2.50% |
| 16 | SM2A, SM2B | 1 | 2.50% | SM1A-SM2A-M3-M4-SM5B | 3 | 7.50% |
| 17 | SM2A, M3, SM2B, M4, SM5A | 1 | 2.50% | SM1A-SM2A-M3-M4-SM5B-SM5A | 1 | 2.50% |
| 18 | SM2A, M4 | 3 | 7.50% | SM1A-SM2B-M3-M4 | 1 | 2.50% |
| 19 | SM2A, M4, SM5B | 1 | 2.50% | SM1A-SM2B-M3-M4-SM5A-SM5B | 1 | 2.50% |
| 20 | SM2B, SM2A | 1 | 2.50% | SM1A-SM2B-M3-M4-SM5B | 2 | 5.00% |
| 21 | M4 | 1 | 2.50% | SM1A-SM2B-M3-SM2A-SM5B | 1 | 2.50% |
| 22 | | | | SM1B-SM2B-M3-M4-SM5A-SM5B | 1 | 2.50% |
| 23 | | | | SM2A-M3-M4-SM5A-SM5B | 2 | 5.00% |
| 24 | | | | SM2A-M3-M4-SM5B | 2 | 5.00% |
| 25 | | | | SM2B-M3-M4 | 1 | 2.50% |
| 26 | | | | SM2B-M3-SM2A | 1 | 2.50% |
| 27 | | | | SM2B-SM2A-SM5B | 1 | 2.50% |
| 28 | | | | SM2B-SM2B-M3-M4-SM2A | 1 | 2.50% |
| 29 | | | | SM2B-SM2B-M3-M4-SM5A-SM5B | 1 | 2.50% |

| | | |
|---|---|---|
| Total | 40 | 40 |

*Note*: F(AN) and F(ED) refer to the frequency of move sequence patterns in anthropology and education of the Corpus of Human-Generated Review Article Abstracts (CHGRAA), respectively. P(AN) and P(ED) represent the percentages of move sequence patterns in anthropology and education of CHGRAA.

We identified 21 and 29 different move sequence patterns in these two disciplines, respectively. As can be seen from the table, within the abstracts of review articles from anthropology and education, the most commonly employed move sequences were SM1A-SM2A and SM1A-SM1B-SM2B-M3-M4-SM5A, respectively. This implies that human reviewers in anthropology might be more likely to present the background or the context of the topic in question and tend to provide the summary of the content or the overview of topics dealt with in the review. When it comes to education review articles, human writers may prefer to provide a more structured abstract, covering the background, the knowledge gap or the niche, the aims, the methodology, the results and the significance of the review topic.

Furthermore, as can be seen from Table 11 and 12, there were more diversified move sequence patterns in the soft disciplines in comparison with those patterns in the hard disciplines. This indicates that human reviewers in the soft disciplines might be more likely to organize the structure of the review article abstracts in a flexible way.

**4.4 The Move Embedding Across CCGRAA and CHGRAA in RA Abstracts**

The move embedding refers to the blending of moves into the same statement (Santos, 1996). We identified 5 embedded moves in ChatGPT-generated review article abstracts and 22 in human-written ones. In the ChatGPT-generated abstracts, most fully embedded moves (N=5) were composed of Move 2 (**Presenting the research**) and Move 3 (**Describing the methodology**). In the human-written abstracts, most of the fully embedded moves (N=10) also consisted of Move 2 and Move 3. Moreover, Move 3 and Move 4 (**Summarizing the results**) were frequently embedded together (N=6) in the abstracts written by humans. The Chi-square test demonstrated statistically significant differences in the frequency of move embedding between the Corpus of ChatGPT-Generated Review Article Abstracts (CCGRAA) and the Corpus of Human-Generated Review Article Abstracts (CHGRAA) ($\chi 2=10.36$, $p < 0.01$). This implies that move embedding might occur significantly less frequently in the review article abstracts generated by ChatGPT compared to those produced by humans. It also suggests that ChatGPT may generate simpler or more straightforward sentence structures compared to human writers when producing the abstracts. For example:

(10) *Drawing on various scholarly perspectives, including linguistic anthropology and related fields* [**M3**], *the review explores how individuals' recognition of vulnerability and their emotional responses shaped new forms of social engagement.* [**M2**] (Education by ChatGPT)

(11) *Drawing on the anthropology of violence, genocide, and the state, within the context of anthropology's colonial turn since the 1970s,* **[M3]** *this article describes how states within colonial modernity create and exploit population-differentiated death through practices of social death, slow violence, and slow death.* [**M2**] (Anthropology by ChatGPT)

(12) *This study aimed to investigate the effect of feedback on students' academic achievement in a technology-rich environment* **[M2]** *through a systematic and quantitative synthesis of the studies conducted over several decades.* [**M3**] (Education by Humans)

(13) *Based on the analysis of 66 selected empirical and non-empirical articles* **[M3]**, *the results show that there is a reciprocal relationship between barriers that occur at the macro and micro levels.* [**M4**] (Education by Humans)

## 5. Discussion

Our research primarily focused on comparing the rhetorical moves in review article abstracts generated by ChatGPT versus those written by humans across four soft and hard science disciplines, attempting to identify their similarities and differences, along with an evaluation of the ChatGPT's performance and efficiency in generating these abstracts. One of the important findings reveals that the overall pattern of the moves in review article abstracts varied considerably between the Corpus of ChatGPT-Generated Review Article Abstracts (CCGRAA) and the Corpus of Human-Generated Review Article Abstracts (CHGRAA). Notably, Move 1 (i.e. Situating the research), Move 3 (i.e. Describing the methodology), and Move 5 (i.e. Discussing the research) occurred with markedly different frequencies in the two kinds of abstracts. On the contrary, no notable differences were observed in Move 2 (i.e. Presenting the research) and Move 4 (i.e. Summarizing the results) across the two corpora. This indicates that both humans and the large language model may demonstrate similarities when presenting research and summarizing results in review article abstracts, but might differ in presenting the background, describing the methodology, and discussing the results. Another intriguing finding was that human writers often used a more comprehensive move sequence pattern, at times including all five moves, whereas ChatGPT preferred less complicated and shorter move sequences, possibly omitting some moves or sub-moves. This finding may partially support the assumption that the writing practices in different science disciplines could be influenced by disciplinary cultures, social contexts, recontextualization, the authors' partly language-dependent genre-specific knowledge and language-independent metacognitive genre awareness (Hyland, 2000, 2004; Johns, 2008; Kim & Belcher, 2018; Martin & Rose, 2008; Swales, 1990; Tardy, Sommer-Farias, & Gevers, 2020).

Another important finding was that ChatGPT-generated abstracts demonstrate a less diversity in their move sequence patterns compared to the human-written abstracts. This result is consistent with prior findings on the performance of ChatGPT in generating academic journal articles (Gao, *et al*., 2023; Wang, *et al*., 2023). Specifically, the abstracts produced by ChatGPT appeared to be more ambiguous and standardized (Gao, *et al*., 2023). Similarly, the abstracts generated by ChatGPT showed a high degree of uniformity and strong rationality, leaning towards academic discourse such as summaries (Wang *et al*., 2023). In contrast, abstracts written by human scholars displayed distinct individuality, frequently employing terms specific to certain research contexts and policy-related vocabulary. These variations may be partially explained by the assumption that human authors might possess more creative capacities to modify their writing style according to different social contexts or communicative purposes. On the other hand, the texts generated by ChatGPT could be more inflexible or constrained due to the programming or the scarcity of data the large language model has been trained on. Since ChatGPT was a kind of deep learning model (Baidoo-Anu & Ansah, 2023), and it was trained on a large dataset of text, such as books and articles, and learns to generate text that was similar to the text it was trained on, so it might suffer from some data training problems, such as limited training data, a lack of relevant data, model overfitting, and an unbalanced dataset (Bansal, Sharma, & Kathuria, 2022).

The most interesting finding was that in both ChatGPT-generated and human-written abstracts, the majority of fully embedded moves were comprised of Move 2 (i.e. Presenting the research) and Move 3 (i.e. Describing the methodology). This finding is in agreement with Santos (1996) findings which showed, Move 3 was more likely to merge with Move 2, either partially or totally on most occasions. The embedding of two (or more) moves was a typical phenomenon across all five moves of the abstract.

## 6. Conclusion and Implications

Recently, the role of ChatGPT in text generation has received increased attention across a number of disciplines. Previous research suggests that ChatGPT has the potential to enhance the generation of academic texts. This study annotated a corpus of 320 review article abstracts generated by ChatGPT and human writers in four soft and hard science disciplines, respectively. This study provides an assessment of the performance and efficiency of ChatGPT in generating academic texts. Our results reveal significant differences in the types of rhetorical moves, move patterns, and the frequency of move embedding between abstracts generated by ChatGPT and those authored by humans. This broadens the use of move analysis to the abstracts of review articles generated by ChatGPT and human authors in both soft and hard disciplines, thereby enhancing our understanding of how the abstract writing practices in various disciplines could be affected by factors such as disciplinary cultures, social contexts, the authors' specific genre knowledge, the awareness of genre, etc. The insights gained from this research offer a contribution to the understanding of the review genre in academic writing and underscore the degree to which move analysis provides a comprehensive view of the structural patterns of a review article abstract. In addition to its theoretical contributions to discourse analysis, this research also has practical benefits for those looking to enhance the text generation capabilities of large language models, or those interested in genre-based EAP pedagogy for novices or non-native speakers.

The present study was subject to several limitations. The principal limitation of this study was that we only examined some review articles from two hard and two soft disciplines and our data analysis relied heavily on small sample sizes. Besides, our focus was mainly on the rhetorical moves of the abstracts generated by both ChatGPT and human writers. Finally, another limitation arises from the absence of relevant prior studies directly based on the analytical framework used in this study for these two types of journal review article abstracts. This means that it is unlikely to take into account the effect size of our results. As a result, future studies could cover more disciplines and consider the distinctions in sub-types of review articles and structured and unstructured abstracts. Furthermore, grammatical metaphor and stance markers could be another two crucial factors which might be worthy of being used to evaluate the performance of ChatGPT in generating academic texts.

Our analyses suggest that ChatGPT may not be used to generate review article abstracts, for it cannot fully demonstrate the move structure of the genre and showcases the formulaic patterns of the move structure of the genre. Specifically, ChatGPT cannot generate some sub-moves, such as the Sub-move: Stating a problem, and shows less diversity and flexibility in the move patterns and the sequential orders, this might undermine the capability of ChatGPT in generating the review article abstracts. The move structure of the abstracts may be used to evaluate the performance of the ability of ChatGPT in academic text generation.

**References:**

[1] Al Fadda, H. (2012). Difficulties in academic writing: From the perspective of King Saud university postgraduate students. *English Language Teaching*, *5*(3), 123-130. http://dx.doi.org/10.5539/elt.v5n3p123.

[2] Al-Mukdad, S. (2019). Investigating English academic writing problems encountered by Arab

International University students. *Theory and Practice in Language Studies*, *9*(3), 300-306. http://dx.doi.org/10.17507/tpls.0903.07.

[3] American Psychological Association. (2020). *Publication manual of the American psychological association (*7th ed.) American Psychological Association.

[4] Aull, L. L., & Swales, J. M. (2015). Genre analysis: Considering the initial reviews. *Journal of English for Academic Purposes*, *19*, 6-9. https://doi.org/10.1016/j.jeap.2015.05.007.

[5] Babl, F. E., & Babl, M. P. (2023). Generative artificial intelligence: Can ChatGPT write a quality abstract?. *Emergency Medicine Australasia*, 35, 809–811. https://doi:10.1111/1742-6723.14233.

[6] Baidoo-Anu, D., & Ansah, L. O. (2023). Education in the era of generative artificial intelligence (AI): Understanding the potential benefits of ChatGPT in promoting teaching and learning. *Journal of AI*, *7*(1), 52-62. https://doi.org/10.61969/jai.1337500.

[7] Bakhou, B., & Bouhania, B. (2020). A qualitative inquiry into the difficulties experienced by Algerian EFL master students in thesis writing:'Language is not the only problem'. *Arab World English Journal,11(2),* 243-257. https://dx.doi.org/10.24093/awej/vol11no2.17.

[8] Bavdekar, S. B., & Gogtay, N. J. (2015). Writing an abstract for a research manuscript: Providing an honest, succinct and complete summary. *The Journal of the Association of Physicians of India*, *63*(12), 64-67.

[9] Bansal, M. A., Sharma, D. R., & Kathuria, D. M. (2022). A systematic review on data scarcity problem in deep learning: solution and applications. *ACM Computing Surveys (CSUR)*, *54*(10s), 1-29. https://doi.org/10.1145/3502287.

[10] Becher, T. (1994). The significance of disciplinary differences. *Studies in Higher Education*, *19*(2), 151-161. https://doi.org/10.1080/03075079412331382007.

[11] Benichou, L. (2023). The role of using ChatGPT AI in writing medical scientific articles. *Journal of Stomatology, Oral and Maxillofacial Surgery*, *124*(5), 101456. https://doi.org/10.1016/j.jormas.2023.101456.

[12] Bhatia, V. K. (1993). *Analysing genre: Language use in professional settings*. Routledge.

[13] Biber, D., & Conrad, S. (2019). *Register, genre, and style*. Cambridge: Cambridge University Press.

[14] Biber, D., & Gray, B. (2010). Challenging stereotypes about academic writing: Complexity, elaboration, explicitness. *Journal of English for Academic Purposes*, *9*(1), 2-20. https://doi.org/10.1016/j.jeap.2010.01.001.

[15] Buholayka, M., Zouabi, R., & Tadinada, A. (2023). The Readiness of ChatGPT to Write Scientific Case Reports Independently: A Comparative Evaluation Between Human and Artificial Intelligence. *Cureus*, *15*(5). https://doi.10.7759/cureus.39386.

[16] Cotton, D. R., Cotton, P. A., & Shipway, J. R. (2023). Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. *Innovations in Education and Teaching International*, 1-12. https://doi.org/10.1080/14703297.2023.2190148.

[17] Dergaa, I., Chamari, K., Zmijewski, P., & Saad, H. B. (2023). From human writing to artificial intelligence generated text: examining the prospects and potential threats of ChatGPT in academic writing. *Biology of Sport*, *40*(2), 615-622. https://doi.org/10.5114/biolsport.2023.125623.

[18] El-Kassas, W. S., Salama, C. R., Rafea, A. A., & Mohamed, H. K. (2021). Automatic text summarization: A comprehensive survey. *Expert Systems with Applications*, *165*, 113679. https://doi.org/10.1016/j.eswa.2020.113679.

[19] Evans, S., & Green, C. (2007). Why EAP is necessary: A survey of Hong Kong tertiary students. *Journal of English for Academic Purposes*, *6*(1), 3-17. https://doi.org/10.1016/j.jeap.2006.11.005.

[20] Gao, C. A., Howard, F. M., Markov, N. S., Dyer, E. C., Ramesh, S., Luo, Y., & Pearson, A. T. (2023). Comparing scientific abstracts generated by ChatGPT to real abstracts with detectors and blinded human reviewers. *NPJ Digital Medicine*, *6*(1), 75. https://doi.org/10.1038/s41746-023-00819-6.

[21] Hartley, J., Sydes, M., & Blurton, A. (1996). Obtaining information accurately and quickly: are structured abstracts more efficient?. *Journal of Information Science*, *22*(5), 349-356. https://doi.org/10.1177/016555159602200503.

[22] Horstmanshof, L., & Brownie, S. (2013). A scaffolded approach to discussion board use for

formative assessment of academic writing skills. *Assessment & Evaluation in Higher Education*, *38*(1), 61-73. https://doi.org/10.1080/02602938.2011.604121.
[23] Hosseini, M., & Horbach, S. P. (2023). Fighting reviewer fatigue or amplifying bias? Considerations and recommendations for use of ChatGPT and other Large Language Models in scholarly peer review. *Research Integrity and Peer Review*, *8*(1), 4. https://doi.org/10.1186/s41073-023-00133-5.
[24] Huang, J., & Tan, M. (2023). The role of ChatGPT in scientific communication: writing better scientific review articles. *American Journal of Cancer Research*, *13*(4), 1148.
[25] Hyland, K. (2000). *Disciplinary discourses: Social interactions in academic writing*. London: Longman.
[26] Hyland, K. (2015). Genre, discipline and identity. *Journal of English for Academic Purposes*, *19*, 32-43. https://doi.org/10.1016/j.jeap.2015.02.005.
[27] Hyland, K., & Diani, G. (Eds.). (2009). *Academic evaluation: Review genres in university settings*. London: Palgrave Macmillan.
[28] Hyon, S. (1996). Genre in three traditions: Implications for ESL. *TESOL quarterly*, *30*(4), 693-722. https://doi.org/10.2307/3587930.
[29] Jiang, F. K., & Hyland, K. (2017). Metadiscursive nouns: Interaction and cohesion in abstract moves. *English for Specific Purposes*, *46*, 1-14. https://doi.org/10.1016/j.esp.2016.11.001.
[30] Johns, A. M. (2008). Genre awareness for the novice academic student: An ongoing quest. *Language teaching*, *41*(2), 237-252. https://doi.org/10.1017/S0261444807004892.
[31] Kim, M., & Belcher, D. D. (2018). Building genre knowledge in second language writers during study abroad in higher education. *Journal of English for Academic Purposes*, *35*, 56-69. https://doi.org/10.1016/j.jeap.2018.06.006.
[32] Lin, L. H., & Morrison, B. (2021). Challenges in academic writing: Perspectives of Engineering faculty and L2 postgraduate research students. *English for Specific Purposes*, *63*, 59-70. https://doi.org/10.1016/j.esp.2021.03.004.
[33] Lund, B. D., Wang, T., Mannuru, N. R., Nie, B., Shimray, S., & Wang, Z. (2023). ChatGPT and a new academic reality: Artificial Intelligence-written research papers and the ethics of the large language models in scholarly publishing. *Journal of the Association for Information Science and Technology*, *74*(5), 570-581. https://doi.org/10.1002/asi.24750.
[34] Macdonald, C., Adeloye, D., Sheikh, A., & Rudan, I. (2023). Can ChatGPT draft a research article? An example of population-level vaccine effectiveness analysis. *Journal of Global Health*, *13*. https://doi:10.7189/jogh.13.01003.
[35] Májovský, M., Černý, M., Kasal, M., Komarc, M., & Netuka, D. (2023). Artificial Intelligence Can Generate Fraudulent but Authentic-Looking Scientific Medical Articles: Pandora's Box Has Been Opened. *Journal of Medical Internet Research*, *25*, e46924. https://doi:10.2196/46924.
[36] Martin, J. R. (2009). Genre and language learning: A social semiotic perspective. *Linguistics and Education*, *20*(1), 10-21. https://doi.org/10.1016/j.linged.2009.01.003.
[37] Martin, J. R., & Rose, D. (2008). *Genre relations: Mapping culture*. London: Equinox Publishing Ltd.
[38] Omidian, T., Shahriari, H., & Siyanova-Chanturia, A. (2018). A cross-disciplinary investigation of multi-word expressions in the moves of research article abstracts. *Journal of English for Academic Purposes*, *36*, 1-14. https://doi.org/10.1016/j.jeap.2018.08.002.
[39] Pho, P. D. (2008). Research article abstracts in applied linguistics and educational technology: A study of linguistic realizations of rhetorical structure and authorial stance. *Discourse Studies*, *10*(2), 231-250. https://doi.org/10.1177/1461445607087010.
[40] Salvagno, M., Taccone, F. S., & Gerli, A. G. (2023). Can artificial intelligence help for scientific writing?. *Critical Care*, *27*(1), 1-5. https://doi.org/10.1186/s13054-023-04380-2.
[41] Samraj, B. (2014). Move structure. In K. P. Schneider & A. Barron (Eds.), *Pragmatics of Discourse* (pp. 385-405). Walter de Gruyter GmbH.
[42] Santos, M. B. Dos (1996). The textual organization of research paper abstracts in applied linguistics. *Text & Talk*, *16*(4), 481-500. https://doi.org/10.1515/text.1.1996.16.4.481.
[43] Shen, Y., Heacock, L., Elias, J., Hentel, K.D., Reig, B., Shih, G., & Moy, L. (2023). ChatGPT and Other Large Language Models Are Double-edged Swords. *Radiology*, 230163. https://doi.org/10.1148/radiol.230163.
[44] Swales, J. M. (1990). *Genre analysis: English in academic and research settings*. Cambridge

university press.
[45] Swales, J. M. (2004). *Research genres: Explorations and applications*. Cambridge University Press.
[46] Tardy, C. M., Sommer-Farias, B., & Gevers, J. (2020). Teaching and researching genre knowledge: Toward an enhanced theoretical framework. *Written Communication*, *37*(3), 287-321.
[47] Tardy, M. C. & Swales, J. M. (2014). Genre analysis. In K. P. Schneider & A. Barron B. (Eds.), *Pragmatics of Discourse*, (pp.165-187). Walter de Gruyter GmbH & Co KG.
[48] Tardy, M. C. (2011). Genre analysis. In K. Hyland & B. Paltridge (Eds.), *Continuum companion to discourse analysis* (pp.54-68). London: Continuum International Publishing Group.
[49] Tessuto, G. (2015). Generic structure and rhetorical moves in English-language empirical law research articles: Sites of interdisciplinary and interdiscursive cross-over. *English for Specific Purposes*, *37*, 13-26. https://doi.org/10.1016/j.esp.2014.06.002.
[50] Van Dis, E. A., Bollen, J., Zuidema, W., van Rooij, R., & Bockting, C. L. (2023). ChatGPT: five priorities for research. *Nature*, *614*(7947), 224-226. https://doi.org/10.1038/d41586-023-00288-7.
[51] Wang, Y., Guo, X., & Liu, Z. (2023). Detection and comparative study of differences between AI-generated and scholar-written Chinese abstracts: A case study in the field of library science. *Journal of Intelligence,* 42(9), 127-134.

Title Page

# A Comparative Genre Analysis of AI-generated and Scholar-written Abstracts for English Review Articles in International Journals

Xinwan Kong[a], Chengyu Liu[a]

[a] College of International Studies, Southwest University, Chongqing, China