

Construct Validity in the Academic Reading Test: a Comparison of Reading Requirements Between China's College English Test CET 4/6 and American TOEFL Test Items and in a China University Study

--A Discussion Based on a Systematic Literature Review

Name: Zhenyu Chen (Vivi)

Student ID: 2300067

Module code: 7016

27/02/2024

Content

1. Aim of Study
2. Background & Context & Rationale
3. Search Process
4. Definition of Literature Review
5. How to Critique the Literature
6. Current Definitions
7. Summary of Key Ideas
8. Research Gaps
9. Research Questions
10. Resources and References

Aim of Study

- **Positionality (Guangzhou Xinhua College in China since 2018.)**
 - supervisor of the program of IELTS and TOEFL training;
 - participation for the course design and textbook of CET4/6 and TOEFL;
- **Research Significance (Research gap in red)**
- The purpose of this mixed methods study is to evaluate the extent to which these assessments measure the same underlying construct of academic reading proficiency by comparing the reading tasks, question types, and cognitive demands of CET 4/6 and TOEFL. One of the assumptions of this research is that there are significant differences in the subject-specific lexical and thematic patterns and cognitive level pattern between the two reading textbooks in the Chinese university context.
- These prior studies reviewed here support the hypothesis that there may be a relationship between the text type and the sort of task or technique that can be used with it (Alderson, 2000, p. 203). The literature on Bachman and Palmer (1996) has highlighted the structure as well as the skills of target language use tasks, which might lead to the development of more authentic test tasks (p.147). Given the evolving nature of language use tasks identified in the literature, it becomes crucial to assess how well the updated textbooks for Chinese CET 4/6 and TOEFL examinations align with these advancements and effectively prepare students. While existing studies have elucidated the relationship between text types and task effectiveness, **there remains a theoretical gap in applying these insights to evaluate and enhance the design of language examination textbooks, particularly in the context of recent reforms.** Bridging this theoretical gap requires a focused investigation into how contemporary textbooks incorporate these insights into task design, thereby offering a pathway to enhance language learning outcomes. Addressing this gap not only enriches our theoretical understanding but also ensures that language education remains responsive to the evolving demands of language proficiency assessments.

Background & Context & Rationale

Construct validity is a fundamental concept in educational assessment, ensuring that a test measures the intended construct accurately and reliably. In the context of academic reading tests, the comparison of reading requirements between different assessments is crucial for understanding the alignment of test items with the targeted construct. **This study focuses on** comparing the reading requirements between China's College English Test (CET 4/6) and the American Test of English as a Foreign Language (TOEFL) in the context of university study in China.

Document analysis requires that data be examined and interpreted in order to **elicit meaning, gain understanding, and develop empirical knowledge** (Corbin & Strauss, 2008; Rapley, 2007). Drawing on the principles of construct validity and educational assessment, this study seeks to contribute to the literature on reading assessment by providing a comprehensive comparison of reading requirements in CET 4/6 and TOEFL. By examining the alignment of test items with academic reading skills and cognitive processes, this research aims to enhance our understanding of the construct validity of these assessments **in the Chinese university context.** The construct validity of such **tests thus has implications for curriculum and classroom practice through the so-called “test washback”** (Alderson and Wall, 1993).

Search Process

The sources we employed for the review were as follows:

1. Electronic databases: Google Scholar, Semantic Scholar, Emerald Insight, IEEE Access, Academic Search Elite/EBSCOhost, Web of Science, Educational Resources Information Center (ERIC), PubMed, JSTOR, and Scopus (databases were last accessed in February 2024).

2. Hand (online) search of journals: Cambridge University Press; Theory and Practice in Language Studies; Saudi Journal of Language Studies; Qualitative Research Journal; International Journal of Qualitative Methods; The Qualitative Report (hand searches completed by February 2024).

Table 1. Search terms derived from initial reviews of the literature.

No.	Search terms
1	Qualitative Document Analysis
2	Theme Analysis
3	Question Tasks Analysis
4	Topic Mining
5	Cognitive Level
6	English Reading Exam Textbooks
7	IELTS & TOEFL
8	Chinese CET 4/6

Definition of Literature Review

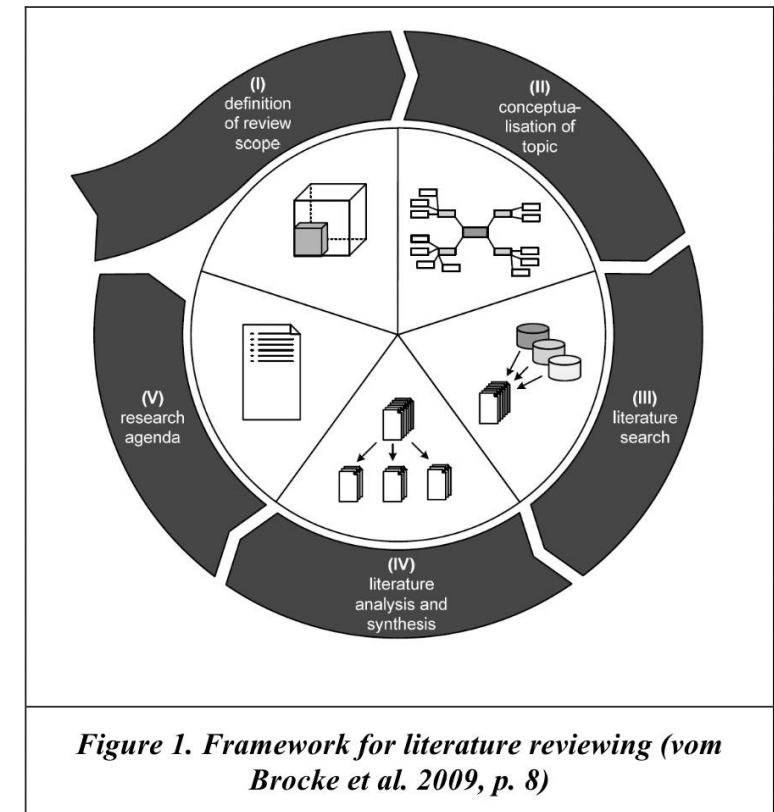
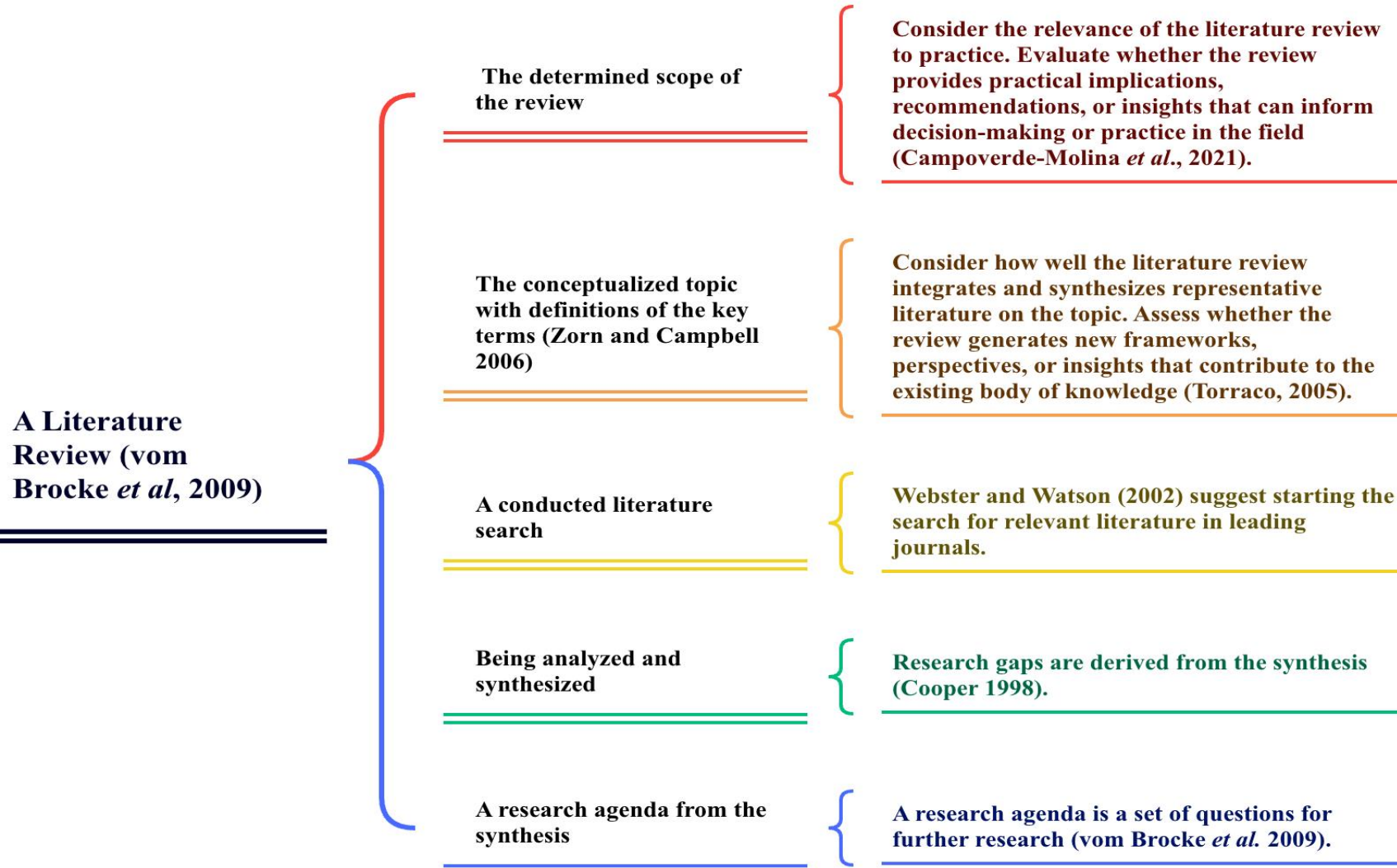
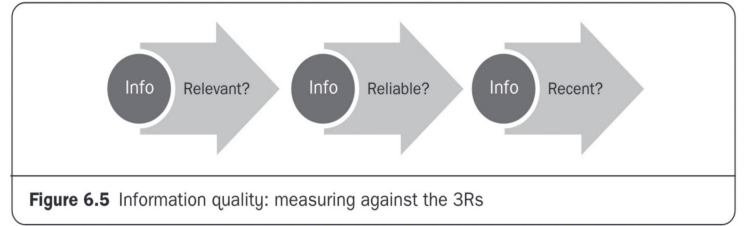
- Snyder, H. (2019) wrote, “Literature review as a research methodology: an **overview and guidelines**”.
- Literature Reviews are also discussed: **Systematic Reviews** (mainly used by health-related disciplines) and **Contextual Reviews** (for those mainly reviewing things, e.g. paintings, more than literature). (John Biggam, 2021, p.10)
- Some studies regard “meta analysis” as a type of **systematic review**, which is primarily a **statistical method** that entails evaluating the research results among many studies on the very **same topic** **using standardized statistical tests** in drawing conclusions and **identify patterns and trends between research results** (Almelhes, 2020).
- **Conceptual model** or **categorization** (Snyder *et al.*, 2016; Witell *et al.*, 2016),
- Nonetheless, it is essential for **each systematic literature review** to state its **intended outcomes and provide a transparent protocol** of how the research has been undertaken to address its research question(s) of interest (Badger, Nursten, Williams & Woodward, 2000).

How to Critique the Literature

Step 1: Preliminary preparation: gaining relative information sources

According to John Biggam (2021, p.116-117), “Good information has the following characteristics: it is relevant, reliable and recent.”

Step 2: Criticism process (Side-by-Side Reasoning: The Convergent Map)



Current Definitions

Theoretical Framework: definition

Morgan, H. (2022)'s ideas: Document analysis has been an underused approach to qualitative research. Document analysis is a valuable research method that has been used for many years. This method consists of analyzing various types of documents including books, newspaper articles, academic journal articles, and institutional reports.

Any document containing text is a potential source for qualitative analysis (Patton, 2015).

A good textbook should contain clear instruction, systematic coverage of syllabus, and adequate guidance for the teacher (Ur, 1996). Clear instruction plays a significant role since it also helps students to achieve the learning objectives based on the curriculum (Tyler, 2013).

Evaluation and Analysis & Identification of Methodological Issues:

As defined by Beaugrande and Dressler (1981), text performs communicative functions by meeting specific selection criteria to fulfil the needs of a particular discourse community. Based on pre-set criteria, **different textual dimensions are assessed regarding the efficacy of particular strategies by conducting textual analysis.**

Identification of Patterns and Trends

Studies that may be practical to complete only through **a document analysis include those that involve cross-cultural research** (Bowen, 2009). E.g.: Between Chinese CET 4/6 and American TOEFL (in a university in China)

Contextualization

Parallel Case-- Many research inquiries ask for comparisons about two identified cases. This argument pattern type works to make the **critique argument when the two cases are similar**. The parallel case argument pattern type **uses the comparison of two like cases for its logic**. The rule for the parallel case is where two cases are similar, what is true for the first case is also true for the second. Lawrence & Brenda (2022, p160)

Analogy--Use an analogy argument pattern type when a research question seeks to **clarify or expose the qualities of a particular case by comparing it to an archetype**. The analogy argument pattern type also **uses the logic of comparison**. It compares like parts within a defined case to parts within a **prototypical case for purposes of explanation or clarification**. The analogy argues that qualities or parts of the prototype, archetype, or stereotype explain the qualities or parts described by the research question. Lawrence & Brenda (2022, p160)

Synthesis of Findings

1) Theme Analysis: Traditional vs Digital Topic Mining 2) Lexical Patterns 3) Cognitive Level Analysis

Implications for Future Research

Though limited in scope and validity, these findings can **provide a framework of combined research between text analysis and teaching approaches**. More empirical experiments can be followed up to under-gird the efficacy of the suggested lexical syllabus and provide more detailed pedagogical insights. (Yi Wang, 2022).

Conducting a Qualitative Document Analysis

A conducted literature search

Questions:

Another common issue might be that prior document analysis strategies **has only been carried out in a particular cultural context**.

≡ Could such strategies **be applied into other cultural text?**
≡ Could the **context thus have influenced the findings?**

What prior document analysis strategies employed in selected reading exam textbooks **can be used by other reading exam textbooks to achieve their instruction function and students' learning objectives based on the curriculum?**

Summary of Key Ideas

(Side-by-Side Reasoning:
The Convergent Map)

Key Ideas

The Contextual Gap

Parallel Case. & Analogy.

	TITLE	AUTHOR/YEAR	SUMMARY KEY IDEAS	CONTEXT EXAMPLE
Paper 1	Conducting a Qualitative Document Analysis	Hani Morgan (2022)	close the gap in the literature on conducting a qualitative document analysis; using reflexive thematic analysis	USA
Paper 2	Applying Thematic Analysis to Education: A Hybrid Approach to Interpreting Data in Practitioner Research	Wen Xu and Katina Zammit (2020)	articulating and demonstrating that Thematic Analysis is an accessible and flexible method to analyze qualitative data collected in the natural classroom setting	Australian
Paper 3	Digital Supply Chain Resilience: Analyzing the Literature Using a Topic Modeling Approach	Atiyeh Kazerooni Monfared, Ravi Vatrappu (2024)	Through text mining and topic modeling, we identified the research topics and trends.	Australian
Paper 4	Lexical Cohesion Patterns of Natural Science Academic Text and Implications for Lexical Teaching: Textual Analysis of IELTS Reading Passages	Yi Wang (2022)	to manifest its lexical cohesion patterns and pedagogical implications, and articles' theme	UK
Paper 5	An Analysis Of Reading Comprehension Questions In English Textbook "English For Iraq" According To Revised Bloom's Taxonomy	Maha et.al. (2021)	identify the existence of the cognitive domain levels, in English textbook "English for Iraq" for 2nd intermediate grade based on Bloom's taxonomy.	Iraq
Paper 6	A Comparative Washback Study of IELTS and TOEFL iBT on Teaching and Learning Activities in Preparation Courses in the Iranian Context	Shiva Seyed Erfani (2012)	1.investigate the washback of the two tests to see whether IELTS and iBT functioned in the preparation classes the way the designers anticipated. 2.explore whether the same format preparation took place	Iranian

The Contextual Gap

Parallel Case. & Analogy.

Merriam (1988) emphasizes that documents can help researchers uncover meaning, develop understanding, and discover insights relevant to the research problem. Fereday and Muir-Cochrane (2006) discuss the rigor of thematic analysis, which involves inductive and deductive coding and theme development, highlighting the value of this hybrid approach.

**reflexive
thematic
analysis**

Triangulation (Bowen, 2009):
use data from interviews and observations

Some key unresolved issues regarding thematic analysis include the necessity for further exploration on enhancing the rigor and consistency of thematic analysis methods (Brunsdon et al., 2019).

Contrast

Thematic Analysis

Agreement

4



Agreement

Through **text mining** and **LDA topic modeling**, we identified the **research topics** and **trends**.

Debate

highlighting the importance of digitalization and identifying key research topics and trends

Limitation

One limitation mentioned in the document is that while previous studies have made efforts to use a more systematic method for labeling topics in text mining and topic modeling, the labeling process is still influenced by researchers' judgment.

Unanswered questions : The efficacy of LDA in extracting meaningful topics from chat transcripts, microblogs, and other forms of unstructured text data remains an area of interest for further exploration (Chen & Wang, 2019)

The authors suggest that their contributions could be extended by **conducting a more in-depth review of the dominant topics** to address this limitation

Conducting a Qualitative Document Analysis - - Hani Morgan (2022)

Paper 1

Agreement

Hani Morgan (2022) agrees on the importance and utility of using document analysis as a qualitative research method.

Debate

the debates in the article center around the **proper use** of document analysis in qualitative research, the **need for detailed reporting of procedures**, and the considerations for **integrating document analysis with other research methods for robust data collection and an**

Paper 2

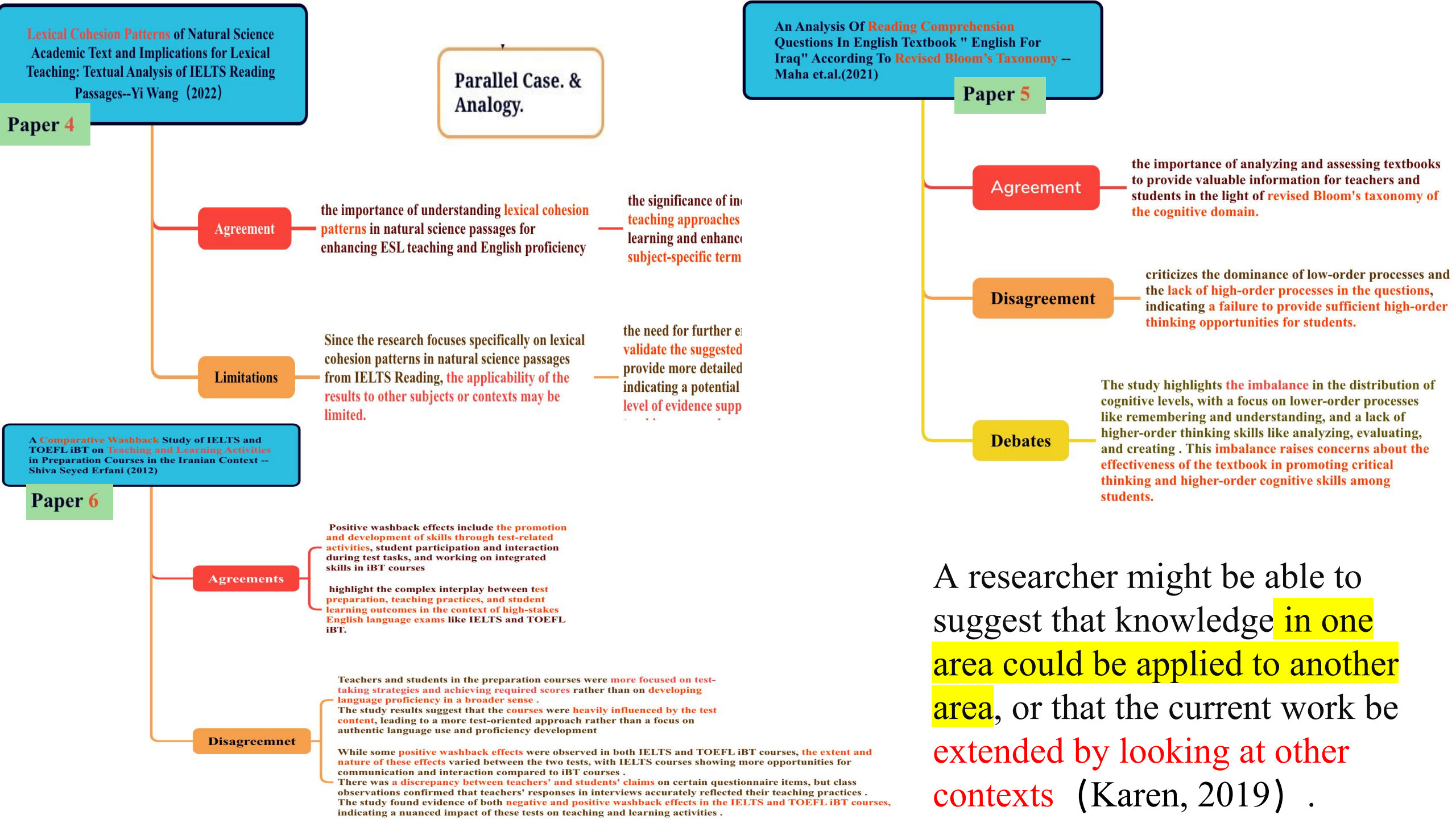
Applying Thematic Analysis to Education: A Hybrid Approach to Interpreting Data in Practitioner Research- Wen Xu and Katina Zammit (2020)

**traditional way:
manual**

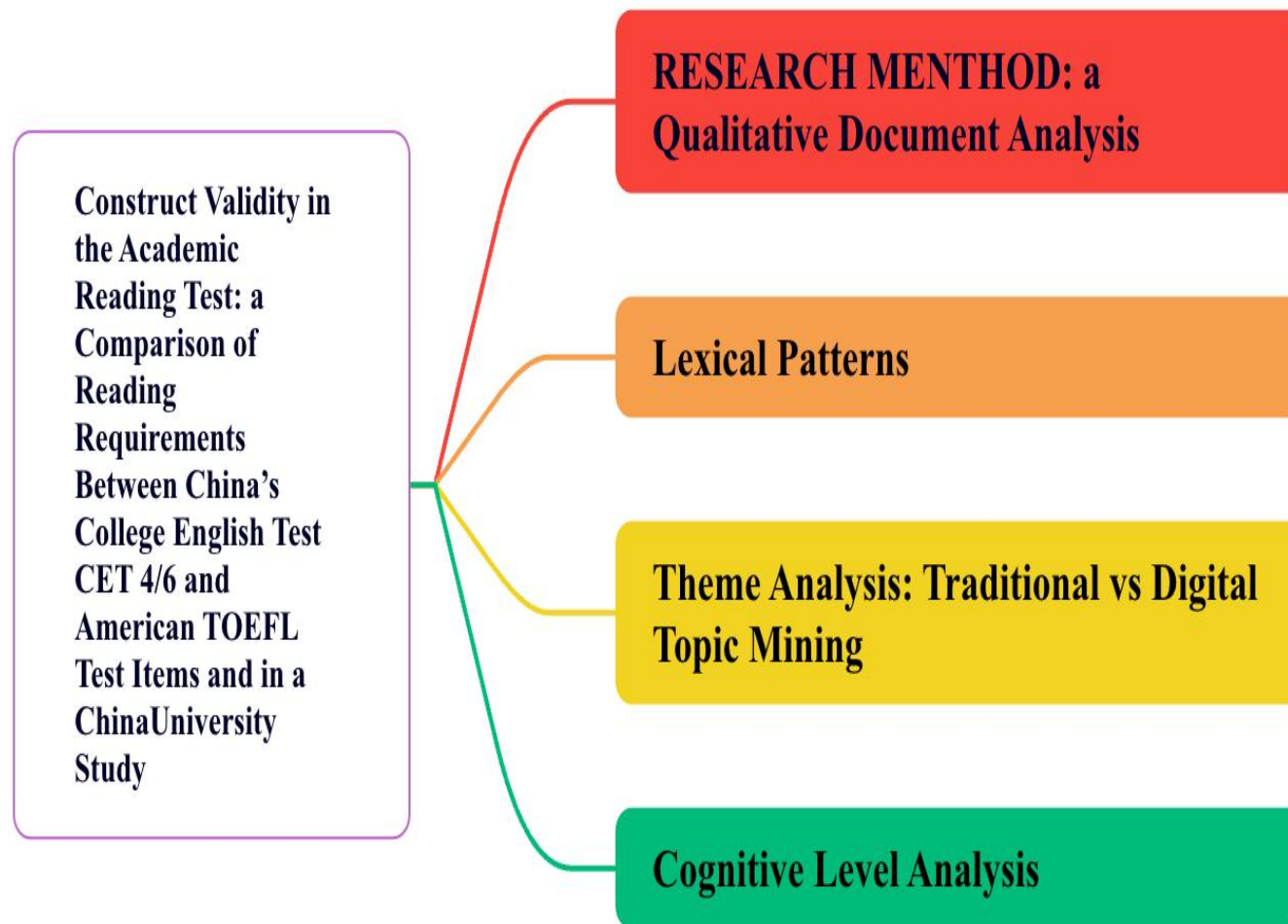
In this respect, data saturation also depends on **researchers' experience** and the number of analysts involved in processing data.

Digital one with big data analysis

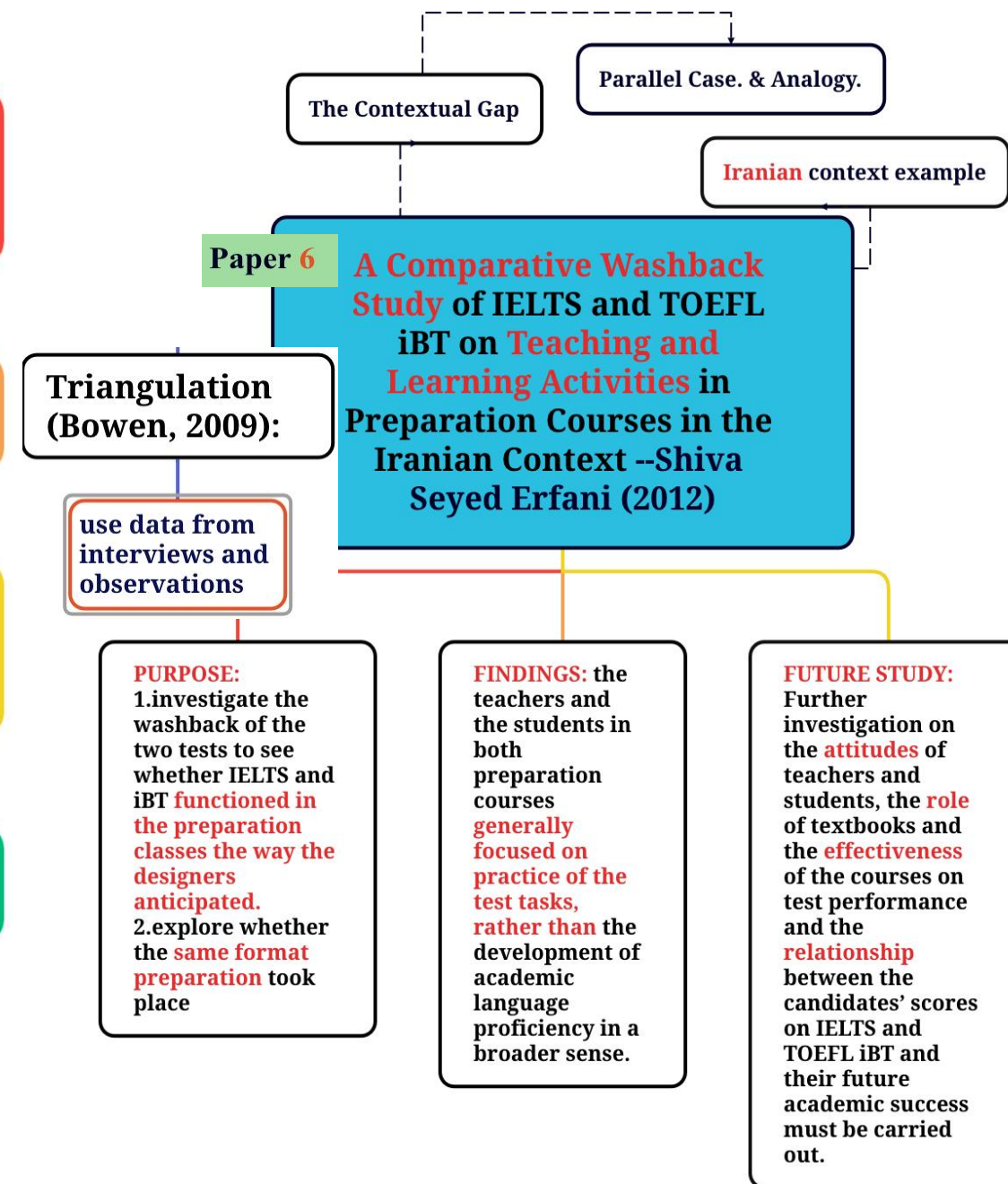
Once the papers have been assigned to their respective topics, the topics need to be labeled. The labeling process requires the domain expert researchers to manually identify the main theme

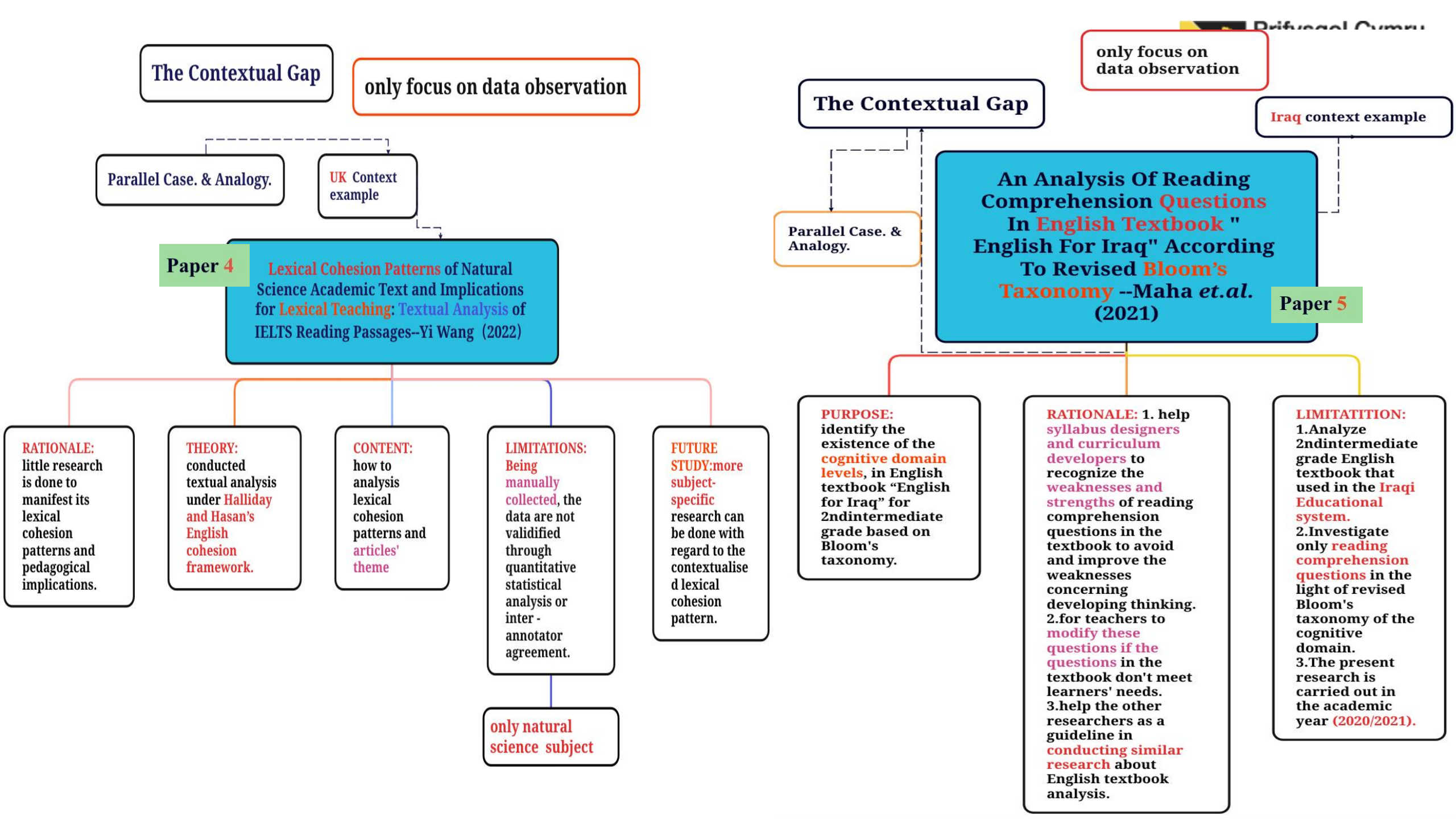


A Research Agenda



A research agenda is a set of questions for further research (vom Brocke et al. 2009).





Research Gaps

- The text starts by identifying a gap in comprehensive studies on the latest versions of reading textbooks, focusing on lexical, thematic patterns, and cognitive levels. It mentions **a theoretical gap but does not clearly connect this gap to the subsequent discussion on methodological approaches or contextual evaluations.**
- To bridge this theoretical and methodological gap, adopting advanced analytical techniques such as digital LDA topic modelling, **previously applied in the SCM field** (Ali & Kannan, 2022; Sangari & Mashatan, 2022; Duong et al., 2022; Saravanan et al., 2022), **could offer new insights into English textbook analysis.** We can do a more in-depth and large-scale analysis of subject-specific lexical and thematic patterns in English textbooks using digital LDA topic modeling, **a method that hasn't been used before in this field.** This could help us find trends and gaps that we hadn't seen before.
- Knowledge in one area could be applied to another area, or the current work could be extended by looking at other contexts. (Karen, 2019). Given the methodological suggested, **this study will focuses on a close look at cognitive levels in a Chinese setting to fill in gaps and add to what is already known about English reading textbooks.** This will make it possible for more complete and useful learning materials to be made.

Research Questions

In general, this research aims to answer the following research questions:

- 1) What are the **differences** in the **subject-specific lexical and thematic patterns**, of two reading textbooks?
- 2) **How much** do the reading parts include the **lower and higher order cognitive levels** in two reading textbooks respectively outlined in the updated Bloom's taxonomy in China context?
- 3) What is the **weaknesses and strengths** of reading comprehension in two reading textbooks for syllabus designers and curriculum developers to avoid and improve the weaknesses concerning developing thinking in future class design?

Resources and References



Alderson, J.C. (2000) *Assessing Reading*. Cambridge: Cambridge University Press.

Alderson, J.C. and Wall, D. (1993) 'Does Washback Exist?', *Applied Linguistics*, 14(2), pp.115–129. Available at: <https://doi.org/10.1093/applin/14.2.115>.

Almelhes, S. (2020) 'Second language acquisition through the flipped learning paradigm: a systematic literature review', *Journal of the Islamic University for Social and Educational Sciences*, 2nd ed., Almadinah Almonorah, pp. 537-569. Available at: <http://search.mandumah.com/Record/1091211>.

Ali, A., Mahfouz, A. and Arisha, A. (2017) 'Analysing supply chain resilience: integrating the constructs in a concept mapping framework via a systematic literature review', *Supply Chain Management*, 22(1), pp. 16–39. Available at: <https://doi.org/10.1108/SCM-06-2016-0197>.

Ali, I. and Kannan, D. (2022) 'Mapping research on healthcare operations and supply chain management: a topic modelling-based literature review,' *Annals of Operations Research*, 315(1), pp. 29–55. Available at: <https://doi.org/10.1007/s10479-022-04596-5>.

Arksey, H. and O'Malley, L. (2005) 'Scoping studies: towards a methodological framework', *International Journal of Social Research Methodology*, 8(1), pp. 19-32. Available at: <https://doi.org/10.1080/1364557032000119616>

Appleby, A., Wilson, P., & Swinton, J. (2018) 'Spiritual care in general practice: rushing in or fearing to tread? an integrative review of qualitative literature', *Journal of Religion and Health*, 57(3), pp. 1108-1124. Available at: <https://doi.org/10.1007/s10943-018-0581-7>

Bachman, L.F. and Palmer, A.S. (1996) *Language Testing in Practice: Designing and Developing Useful Language Tests*. Oxford: Oxford University Press.

Badger, D., Nursten, J., Williams, P. and Woodward, M. (2000) 'Should All Literature Reviews be Systematic?', *Evaluation and Research in Education*, 14(3-4), pp.220–230. Available at: <https://doi.org/10.1080/095007900008666974>.

Beattie, M., Murphy, D., Atherton, I., & Lauder, W. (2015) 'Instruments to measure patient experience of healthcare quality in hospitals: a systematic review', *Systematic Reviews*, 4(1). Available at: <https://doi.org/10.1186/s13643-015-0089-0>

Braun, V. and Clarke, V. (2006) 'Using thematic analysis in psychology', *Qualitative Research in Psychology*, 3(2), pp. 77-101.

<https://doi.org/10.1191/1478088706qp063oa>

Brunsdon, D., Biesty, L., Brocklehurst, P., Brueton, V., Devane, D., Elliott, J., ... & Gillies, K. (2019) 'What are the most important unanswered research questions in trial retention? a james lind alliance priority setting partnership: the priority ii (prioritising retention in randomised trials) study', *Trials*, 20(1). Available at: <https://doi.org/10.1186/s13063-019-3687-7>

Biggam, J. (2021) *Succeeding with your master's dissertation: a step-by-step handbook*. London: Open University Press, McGraw Hill.

Bowen, G.A. (2009) 'Document Analysis as a Qualitative Research Method', *Qualitative Research Journal*, 9(2), pp. 27–40. Available at: <https://doi.org/10.3316/QRJ0902027>.

Campoverde-Molina, M., Luján-Mora, S., & Valverde, L. (2021) 'Accessibility of university websites worldwide: a systematic literature review', *Universal Access in the Information Society*, 22(1), pp. 133-168. Available at: <https://doi.org/10.1007/s10209-021-00825-z>

Chen, X. and Wang, H. (2019) 'Automated chat transcript analysis using topic modeling for library reference services', *Proceedings of the Association for*

Thank you.

Name: Zhenyu Chen (Vivi)

Student ID: 2300067

Module code: 7016

27/02/2024