International Journal of Language and Literary Studies
Volume 5, Issue 1, 2023
Homepage : http://ijlls.org/index.php/ijlls

Qualitative Systematic Review in Applied Linguistics: A Synthesis

Jinpei Zhang
School of Foreign Languages, Shenzhen Institute of Information and Technology, Shenzhen, P.R. China
zhangjp@sziit.edu.cn

DOI: http://doi.org/10.36892/ijlls.v5i1.1208

APA Citation: Zhang, J. (2022). Qualitative Systematic Review in Applied Linguistics: A Synthesis. *International Journal of Language and Literary Studies.* 5(1).156-174. http://doi.org/10.36892/ijlls.v5i1.1208

Received:
10/03/2023

Accepted:
29/03/2023

Keywords:
Qualitative Systematic Review; Qualitative Evidence Synthesis; Qualitative Research Synthesis; Synthesis Methods; Applied Linguistics.

Abstract

Qualitative Systematic Review (QSR) is the second-generation literature review that aims to synthesize primary qualitative studies. Its use is growing rather rapidly in medicine, nursing, education and other fields, but it remains underutilized in applied linguistics. In order to invite more qualitative systematic reviews in applied linguistics, this paper first provides the definition of QSR, outlines the key steps for its conduct, elaborates on the core procedure, i.e., synthesis methods, then illustrates with published examples of qualitative systematic reviews in applied linguistics, and finally explores the future directions.

1. INTRODUCTION

In academic research, an author must first identify what has been done by other scholars in the field, and find a research gap and then decide what questions need to be addressed. Literature review is a written appraisal of research findings related to a topic of interest, usually with no prescribed methodology (Jesson et al., 2011). This kind of review, sometimes termed as traditional narrative review, is the first-generation literature review (Pope et al., 2007). It has received a lot of criticism in the research community due to its potential bias and lack of transparency (Ellis, 2015). In order to reduce subjectivity in the review process and improve rigor and replicability, many researchers have begun to conduct literature review in a systematic way, which Pope et al. (2007) labeled as the second-generation review. Systematicity of a systematic review is reflected in the whole review process: A systematic reviewer first formulates research questions, searches the literature in a systematic way,

retrieves studies that meet the prescribed criteria, conducts a critical assessment of their quality, extracts and analyzes data from the included literature, and finally reports the results in accordance with some guidelines. In general, there are two types of systematic reviews: quantitative systematic reviews and qualitative systematic reviews (Booth et al., 2022; Boland et al., 2017), derived from which we may have mixed method reviews (e.g., EPPI-center review by Macaro et al., 2018). In the field of applied linguistics, many quantitative systematic reviews have been done, with meta-analysis being the most frequently used synthesis method to combine the primary research data (Chong & Plonsky, 2021). On the contrary, however, as Ortega (2011) pointed out, little attention has been paid to the synthesis of qualitative data. Similarly, Chong and Plonsky (2021) observed that qualitative research synthesis is underutilized in the TESOL domain. The book edited by Norris and Ortega (2006), consists of a total of eight secondary research papers, seven of which are quantitative meta-analysis, and only one is qualitative synthesis. Khany and Tazik (2017) examined the research designs of papers published in applied linguistic journals between 1976 and 2015, and tallied them in 10-year cycles. The results show that numbers of quantitative and qualitative designs are both growing rapidly and in the same cycle, the total number of quantitative research and qualitative research is not much different. For example, from 2006 to 2015, the proportion of quantitative research and qualitative research are 46.65% and 45.84% respectively. Currently, many applied linguistic studies are qualitative in nature and can be synthesized to inform policy makers, practitioners and language learners.

Qualitative systematic review can either be published as standalone research work, examples of which will be provided in the subsequent sections, or embedded as an alternative to traditional literature review in a study. In the following sections, we will first give a detailed introduction to QSR, including its definition and steps of its conduct, and then describe its core procedure. On this basis, some qualitative systematic reviews conducted in applied linguistics are presented, and future research directions explored.

.

2. QUALITATIVE SYSTEMATIC REVIEW

2.1.Definition

Qualitative systematic review, interchangeably qualitative evidence synthesis (QES), or qualitative research synthesis (QRS), is an umbrella term for the systematic review of qualitative primary research, conducted either as a stand-alone review or embedded as part of a research work (Booth et al., 2022; Fleming & Noyes, 2021; Boland et al.,2017). Researchers from different backgrounds and fields of study might favor different terms. For example, the term QES is the preferred term of Cochrane Qualitative and Implementation Methods Group, representing the field of medicine and healthcare. In applied linguistics, Avgousti and Hadjistassou (2022) opted for QSR, while Tellez and Waxman (2006), Chong and Plonsky (2021), and Chong and Plonsky (2023) preferred the term QRS. This paper endorses qualitative systematic review as the cover term since it not only indicates qualitative data as its sources but also highlights systematicity in the review.

2.2. Major steps

QSR is to be undertaken step by step, and each step has to follow strict standards or guidelines. Based on the guidelines compiled by Boland et al. (2017), Hannes (2019), and Chong and Plonsky (2021), a typical QSR is conducted in the following six steps:

Step 1: Formulate review questions

A qualitative systematic reviewer should first define the review questions. Some scholars pointed out that questions can be designed by identifying the population, specifying phenomenon of interest, and setting the context (Booth et al., 2022). For example, Macaro et al. (2018) identified university teachers and students as population. Their research interest lies in English as the medium of instruction (EMI), and the context is higher education. Therefore, his review questions include beliefs of university EMI teachers and students towards EMI. If we say quantitative meta-analysis answers the question regarding the effect size of intervention, QSR mainly asks why there is such an effect, or synthesizes the views and attitudes of certain groups.

Step 2: Search strategy

After determining the purpose, next step is to search for relevant literature, usually through the internet (such as search engine or database) and/or by manual search. Stevenson (2016) used keywords or search terms such as automated writing evaluation (AWE) and

automatic writing feedback, to search for relevant literature on search engines (such as Google Scholar, Google), databases (such as ERIC, MLA) and websites (such as ETS official website, AWE software official website), and checked reference lists and bibliographies for additional publications. The common databases in applied linguistics and their providers are shown in Table 1.

Table 1 Common Database for Applied Linguistics

	Database	Provider
1	APA PsycARTICLES includes Journal of Applied Psychology, Journal of Educational Psychology, etc.	EBSCO or Proquest
2	APA PyscINFO includes 2285 journals, such as Applied Linguistics, Applied Psycholinguistics, Studies in Second Language Acquisition	EBSCO or Proquest
3	Education Resources Information Center (ERIC)	EBSCO or Proquest
4	Language and Linguistics Behavior Abstracts(LLBA) such as The Modern Language Journal, Language Learning, Foreign Language Annals, EUROSLA Yearbook	Proquest
5	Modern Language Association (MLA) International Bibliography includes Modern Language Journal, Language Learning, etc.	EBSCO
6	Proquest Dissertation and Theses Global	Proquest
7	Social Sciences Citation Index (SSCI) includes 1,700 journals	http://isiknowledge.com

Step 3: Describe the inclusion and exclusion criteria

The number of studies obtained in the initial search is generally rather large, and inclusion and exclusion criteria need to be established to screen out ineligible ones. For example, Mendoza and Phung (2019) proposed three inclusion criteria and two exclusion criteria. Inclusion criteria include: research questions must contain terms such as ideal L2 self, ought-to L2 self, and L2 learning experience; study at least one language other than English; and should be published in a peer-reviewed journal. Exclusion criteria include: articles published in non-English language; articles not retrieved from the four designated databases.

Step 4: Appraise the quality

After literature has been retrieved, its quality needs to be assessed. Macaro et al. (2018) examined the quality of the included literature from the four aspects: relevance to the research question, contribution, appropriateness of the research design, and reliability of the methodology. Only medium and high-level papers are allowed to proceed to the in-depth review stage. Other standards that can be used to assess the quality of literature include Critical Appraisal Skills Program (CASP), etc., (for more such standards, cf. Booth et al., 2022).

Step 5: Extract and synthesize data

Important information of the included literature is extracted and usually presented in the form of a table, including author, year of publication, context, research method, main findings, etc. After data extraction has been completed, they are ready to be synthesized. As mentioned above, there are as many as 20 synthesis methods for QSR, from which reviewers should choose the most appropriate one.

Step 6: Report

The rigor of the systematic review is also manifest in the fact that each process must be reported in accordance with certain norms. There are reporting guidance proposed specifically for a certain synthesis method, for example, France et al. (2019) proposed eMERGe reporting guidance for meta-ethnographic synthesis. There are also more general reporting guidance applicable to different synthesis methods, such as ENTREQ, which intends to enhance transparency in reporting QSR.

2.3. Synthesis: the core step

Synthesis, the major step in the review process, is defined as "making a new whole out of the parts: individual studies or pieces of evidence are somehow combined to produce a

coherent whole, in the form of an argument, theory or conclusions" (Pope et al., 2007:15). At this step, researchers synthesize the evidence extracted from the included literature in a certain way either to form new concepts, or to develop new themes, or to generate theories. According to Barnett-Page & Thomas (2009), there are 12 commonly used qualitative research synthesis methods, including meta-ethnography, grounded theory synthesis (GTS), thematic synthesis, critical interpretative synthesis (CIS), and framework synthesis, etc. Hannes (2019) mentioned more than 20 synthesis methods for qualitative research. QSR is still in the process of rapid development, with the extant synthesis methods being adapted, and new synthesis strategies constantly emerging.

Barnett-Page and Thomas (2009) provided an overview of 12 qualitative synthesis methods and argued that these methods are distinct from each other across a range of dimensions, which include reviewers' epistemological assumptions, iterations in the process, approach to quality assessment, extent of literature problematizing, similarities and differences between primary studies, etc. In terms of reviewer's epistemological positions (see Table 2), these synthesis methods formed a continuum, with subjective idealism at the leftmost end, objective idealism, critical realism and scientific realism in the middle, and naive realism at the rightmost end (Barnett-Page & Thomas, 2009). Subjective idealism holds a rather constructionist view of knowledge and assumes that there are many realities. Reviews informed by such epistemological stance tend to interpret differences in research findings across studies from the assumptions underlying each study. Objective idealism reviewers assert that there are shared realities or objective commonalities across different studies. Critical realism contends that the way we get to know reality depends on our perceptions and beliefs. A review endorsing such epistemology usually problematizes the literature and configures views of different people on issues of interest into one framework. Scientific realism argues that it is possible to obtain the same outcomes or reality as long as scientific methods are applied and the situations remain the same.

Table 2 Epistemological Difference Between Methods (Barnett-Page & Thomas, 2009)

Subjective idealism			Objective idealism		Critical realism			Scientific realism
Meta-narrative	IS	Meta-study	Meta-ethnography	GTS	Thematic synthesis	Textual narrative synthesis	Framework synthesis	Ecological triangulation

The choice of review method is constrained by many factors but to Barnett-Page & Thomas' (2009) epistemological position is probably the single most important one. Even though reviewers' views of knowledge do help explain why there are such an array of different synthesis methods, fitting them into specific epistemological stances can be a hassle (Barnett-Page & Thomas, 2009).

A more comprehensive guidance for choosing an appropriate synthesis method was Booth et al.'s (2018) RETREAT framework (these initial letters are underlined and in bold in the forthcoming sentences). They mapped 15 synthesis methods against the seven RETREAT criteria and worked out a matrix table detailing subdomains of each criterion to represent more nuanced considerations for choice of qualitative synthesis method. **R**eview question is considered as a critical factor when choosing QSR methods. Some review questions are fixed and some emergent. For example, framework synthesis and thematic synthesis usually deal with fixed questions that are formulated even before the review process begins, while review questions addressed by GTS and meta-ethnographic synthesis tend to be adjustable in the review process and thus "emergent". Echoing Barnett-Page and Thomas (2009), Booth et al. (2018) observed that **E**pistemology is a further key consideration in method choice. Different synthesis methods fall into different slots on the idealism-realism continuum. Idealists tend to endorse a configurative review approach (which configures findings across studies) while realists are more likely to adopt an aggregative approach (which adds up findings across studies). Furthermore, some types of synthesis require methods should be compatible with the epistemology of the primary studies (e.g., meta-ethnography and GTS), while other methods (e.g., thematic synthesis) do not demand consistency of epistemological positions. In addition, methods can be distinguished from each other according to the extent to which they are concerned with theory generating, theory exploring, or theory testing. A framework synthesis

is more likely to test theory, while a GTS usually intends to generate theory. In terms of **T**ime, some synthesis involves several iterations, for example, GTS requires constant comparison, open coding and axial coding, and as a result, it can be quite time-consuming. In addition to time, the availability of other **R**esources has major impact on method choices. People (in terms of efforts of reviewer team, besides their expertise) and funding (e.g., cost of meetings or budget for software) are two important factors to consider. When it comes to **E**xpertise, most stages of a systematic review require only general research skills, while other processes (e.g., synthesis) may require specialist input. Some methods, such as GTS and thematic analysis place heavy requirements for expertise in primary qualitative techniques (e.g., constant comparison and axial coding). Systematic reviews cater to the needs of different **A**udiences, including policy makers, practitioners and language learners. Practitioners, for example, desire results that are ready for direct and immediate use. In this case, framework synthesis, thematic synthesis, or meta-aggregation, may be the right synthesis methods. On the contrary, users of the results obtained from meta-ethnographic synthesis and GTS have to consider their applicability in the contexts and make necessary adaptations. Finally, data can be divided into "thick" vs. "thin" and "rich" vs. "poor" **T**ypes. Contextually thin data, i.e., primary studies that give little clue to situational contexts, are only suitable for meta-aggregation, thematic synthesis, or narrative synthesis. Data are divided into "rich" and "poor" according to the richness of conceptual content and the degree to which the studies contribute to theory development. For example, primary studies synthesized with GTS generally contain rich concepts: rich codes are generated from constant comparison process, which are then further categorized or combined through axial coding to develop theory or framework.

2.4. Published Examples of QSR in Applied Linguistics

Table 3 provides some examples of QSR studies undertaken by scholars in the past few years, featuring different synthesis methods. In subsequent paragraphs the review process of each method is briefly discussed and synthesis results reported.

Tellez and Waxman (2006) synthesized the results of 25 included studies using meta-synthesis method. They engaged the open coding and axial coding procedure (Strauss & Corbin, 1990). Four effective teaching practices emerged from the meta-synthesis: communitarian teaching practices, protracted language events, multiple representations designed for understanding target language, and building on prior knowledge.

Based on Noblit and Hare's (1988) meta-ethnographic synthesis method, Badjadi (2015) identified six primary studies and translated the concepts into one another. The term "translate" in this context refers to the process of taking concepts from one study and recognizing them in another, though they may not be expressed exactly in the same words (Thomas & Harden, 2008). Badjadi developed four key concepts from the primary studies: rhetorical awareness, task complexity, literacy, and source writing experience. Then 11 second-order interpretations (derived from the primary sources) falling under the four conceptual categories were identified, on the basis of which 12 third-order interpretations (based on the key concepts and second-order interpretations) were constructed. Finally these were all linked together in a line of argument that accounts for the significant factors which affect the way task requirements and the assigned readings are interpreted, and the usefulness of interventions that help pave the way for a proper understanding of the demands of similar writing tasks.

Avgousti's (2018) systematic review is composed of a mapping review stage and an in-depth review stage. At the mapping review stage, Avgousti described the characteristics of included studies (e.g., author, journal, year of publication, participants, data collection, findings, etc.) and showed the distribution of intercultural communicative competence studies in terms of publication venue, technology mode (e.g., asynchronous, synchronous), technology modality (e.g., text, video, voice), Web 2.0 tools (e.g., email, Skype, social networking tools). The in-depth review focuses on the significant role of online technological tools in successful intercultural exchanges, and the affordances of the tools in the development of intercultural competences.

Inspired by Fleming's (2009) CIS method, Stevenson (2016) first extracted key information to provide a context within which AWE classroom use could be interpreted, then identified 203 "translations" (i.e., concepts, themes and ideas) in the findings, discussion and conclusions sections of the primary studies and synthesized them into three overarching synthetic constructs: Purpose, Action, and Use. "Purpose" discusses the reasons why AWE is applied in the classroom, "action" describes how the teachers and learners actually utilize AWE in the classroom, and "use" points out the issues encountered in the use of AWE in the classroom. Finally one synthetic argument per construct was developed: there are numerous

purposes for using AWE as discussed in primary studies; some teachers have creative ways of integrating AWE in their classrooms; some of the feedback provided by AWE systems is fraught with limitations.

The main purpose of Mendoza and Phung (2019) is to pilot critical research synthesis (CRS), a qualitative synthesis method, which they tout as an alternative for quantitative meta-analysis. To this end, they retrieved 30 studies that applied L2 Motivational Self System (L2MSS) framework to investigate learners of languages other than English (LOTEs). Then the studies were coded for geographical area, educational background, target language, research methods, research questions and findings. The extracted data were then further analyzed by reporting study demographics, synthesizing the substantive findings, and discussing methodological issues. Critiquing penetrated the whole data analysis process. For demographic characteristics, under-researched contexts and population were identified. For substantive findings, issues such as the competition between English and LOTEs, ways to motivate students to learn LOTEs in formal educational settings, applicability of different L2 motivation theories to different types of learners were elaborated on. For methodological issues, strengths and weaknesses of different categories of research methods were examined.

Grounded theory, which used to be a methodology to generate theory out of first-hand qualitative data, has been employed by Chen (2016) as a synthesis method to combine research findings of primary studies in applied linguistics. Chong and Reinders (2020) followed Chen's (2016) example and used GTS to synthesize results from 16 technology-mediated Task-Based Language Teaching (TBLT) studies. In their review, Chong and Reinders (2020) used NVivo to create a total of 332 initial codes, yielding four conceptual categories (composed of 10 descriptive categories, which were themselves made up of 31 sub-categories). The four

conceptual categories include: the characteristics, the affordances, and the limitations of technology-mediated tasks, and the factors determining the effectiveness of TBLT. Following the synthesis, future research directions are proposed and practical implications suggested.

In the above section, QSR examples conducted in applied linguistics are provided to illustrate the whole review process. Compared to fields such as medicine, nursing and education, the number of QSR studies in applied linguistics is still limited. Moreover, there are some inadequacies that need to be addressed. For example, Badjadi (2015) identified only six primary sources in his review, which we find insufficient to derive so many interpretations as he did. At the same time, his exclusion criteria may not be stringent enough so that in one included literature the participants were even university professors who wrote academic papers, while his purpose was to investigate how English learners represented reading-to-write tasks. Another example is the review conducted by Avgousti (2018), in which she claimed to conduct a thematic synthesis review, but did not follow the three-stage protocol devised by Thomas and Harden (2008). Since no explanation was offered why there is such deviation, the transparency and systematicity of the review may be undermined. Such issues will be addressed in the future directions section.

Table 3 QSR Examples in Applied Linguistics

Authors（Year）	Review aim	Process (Search terms, Primary studies, Inclusion/Exclusion criteria)	Synthesis Method
Tellez and Waxman (2006)	To synthesize research on effective teaching practice to English learners.	Five databases. Search terms (English, English Language Learners, Instruction, Effective, etc.). Twenty-five qualitative primary studies. Four inclusion criteria: the study should provide a rationale for choosing its participants and context, etc.	Meta-synthesis
Badjadi (2015)	To synthesize research on task representation of reading-to-write.	Two EBSCO databases, LLBA. Search term, e.g., L2/SL reading-to-write perceptions OR task representations. Six qualitative primary studies. Inclusion criteria: qualitative research papers published between 2003 and 2014 on task representations.	Meta-ethnography
Stevenson (2016)	To synthesize existing research on the integration of AWE into classroom writing instruction.	Twelve databases (ERIC, MLA), search engines (Google scholar), websites (ETS), 22 journals from 1990-2013. Search terms (e.g., AWE). Twenty-one primary studies. Inclusion criteria: The conclusion and discussion parts should mention the application of AWE in classroom teaching.	Critical interpretative synthesis

Avgousti (2018)	To synthesize the influence of Web 2.0 tools on intercultural communicative competence.	Six databases, 37 journals. Search terms use a truncation strategy (e.g. intercult*). Fifty-seven primary studies. Nine inclusion criteria: e.g., the telecollaborative partnership should entail some form of Web 2.0 tool and application. Nine exclusion criteria: e.g., There is no use of Web 2.0 tools and applications.	Thematic synthesis
Mendoza and Phung (2019)	To synthesize research on motivation to learn language other than English on the basis of L2 Self System framework.	Four databases: ERIC, LLBA, PsycInfo, Web of Science. Seven search terms. Thirty primary studies (9 qualitative, 18 quantitative, 3 mixed methods). Three inclusion criteria: research questions should contain ideal self, ought to self; second language learning; language other than English; peer-reviewed journals.	Critical research synthesis
Chong and Reinders (2020)	To synthesize qualitative findings from 16 technology-mediated TBLT studies published between 2002 and 2017 in second and foreign language contexts.	Ten digital libraries or databases, 25 journals. Search terms: e.g., task-based language teaching. Sixteen primary studies (6 qualitative, 10 mixed methods). Three inclusion criteria: e.g., the articles adopted either a qualitative or mixed-methods research design with a significant qualitative component to the research.	Grounded theory

3. FUTURE DIRECTIONS

In recent years, QSR has attracted more and more attention from applied linguistic scholars and some exciting developments have been achieved. For example, inspired by grounded theory, a method originally mainly used for primary data analysis, Chen (2016) pioneered the use of GTS in QSR. Mendoza and Phung (2019) were able to put forward CRS, as a qualitative alternative to quantitative meta-analysis. As mentioned above, QSR is still in the initial stage in the field of applied linguistics. Recent years have witnessed many qualitative research conducted on effective teaching, automatic writing evaluation, peer feedback in writing, English as a medium of instruction, and learner motivation, etc. Scholars can try to synthesize existing research to provide insights for policy makers, teachers, students and many other stakeholders alike. Especially, reviewers in applied linguistics can:

3.1. Improve review rigor and transparency

As mentioned in the previous section, some reviews conducted in applied linguistics lack rigor and transparency. Badjadi (2015) did not retrieve enough literature to include in the review and failed to exclude literature that were dissimilar to other primary research. Standards for literature search (e.g., STARLITE, see Booth et al., 2022) might be of help in this regard. Conducting systematic reviews in accordance with standards proposed for different stages of review will undoubtedly improve the quality of reviews. For example, CASP can be resorted to in the quality assessment stage. At the review reporting phase, France et al. (2019) published reporting guidance specifically for meta-ethnographic synthesis. ENTREQ, on the other hand, is more general and applicable to the reporting of a variety of synthesis methods.

Synthesis is the core step of QSR. If the synthesis process is not rigorous, it will inevitably affect the systematicity, reliability and replicability of the review. Some worked examples prepared by specialists, e.g., Britten et al. (2002) on meta-ethnography, Thomas and Harden (2008) on thematic synthesis, and Fleming (2009) on CIS, can provide guidance for novice reviewers and help improve the rigor of the review process. Although Avgousti (2018) claimed to have applied Thomas and Harden' s (2008) thematic synthesis method, she did not follow exactly the three-step thematic synthesis scheme (i.e., line by line coding, descriptive theme, and analytical theme). She simply described the important themes in detail but left out the procedures of line by line coding and the analytical theme, and failed to explain the reasons

for making such adaptations. Most probably, what Avgousti (2018) conducted was an EPPI-center review (see Macaro, 2020), rather than a thematic synthesis.

3.2. Regularly update reviews

Systematic reviews may be designed for periodic update to accommodate the emergence of new evidence. In order to maintain the up-to-dateness of synthesized results, systematic reviews should be renewed regularly. For example, the literature included by Chen (2016) is from 1990 to 2010, but many similar studies have been undertaken in the field of applied linguistics since 2011. Therefore, the systematic review of peer feedback needs to be updated.

3.3.Explore the complementarity of quantitative and qualitative systematic reviews

Ellis (2015, 2018) argued that meta-analysis and traditional narrative review should complement each other to provide a more complete answer: meta-analysis can be used to answer questions such as how effective an intervention is and which variables have mediating and moderating effects, while a traditional narrative review can provide a general picture of the key themes and patterns. However, he cautioned that traditional narrative reviews fail to search systematically for relevant literature and there is often selection bias. Ellis observed that these deficiencies could be remedied with the help of an exhaustive literature search so as to make narrative reviews systematic. Then it can be argued that QSR may be a better candidate to work in tandem with meta-analysis in that thematic synthesis and narrative synthesis are two commonly used qualitative synthesis methods for describing themes and discovering patterns. Moreover, QSR has more advantages than traditional narrative review in terms of systematic retrieval and reduction of subjectivity in the literature selection process. In short, the academic circle should explore the complementarity of meta-analysis and QSR rather than traditional literature review. Badjadi (2015) examined the obstacles of second language learners' reading-based writing tasks, such as the writer's problems in quoting the original text. To take this topic one step further, we can ask: What the obstacles are and how effective teaching intervention is. By combining QSR with meta-analysis, researchers can investigate related issues in a more comprehensive and in-depth manner.

3.4.Use software to improve review accuracy and efficiency

Among the examples of applied linguistics introduced in the previous section, Avgousti (2018) mentioned the use of EPPI Reviewer 4 in the process of systematic review. EPPI Reviewer 4 is a software that can be applied in meta-analysis, meta-ethnographic synthesis,

etc. The use of such software would greatly improve efficiency when two or more reviewers collaborate on a review project. Chen (2016) was quite innovative in that she adapted grounded theory to synthesize peer feedback research. However, if software such as NVivo were used at the coding stage and then to discover the connection between codes, it will increase review accuracy and efficiency. Khol et al. (2018) and Dawson (2019) introduced some common software for QSR in detail. Reviewers can choose the appropriate software on the basis of expertise, funding, and synthesis methods.

4. CONCLUSION

This paper first provides a general introduction to QSR, outlines the steps for its conduct, elaborates on its core procedure, and then illustrates the review process with examples conducted in the field of applied linguistics, and finally puts forward future directions. As mentioned earlier, QSR can be published independently as standalone research or embedded as a substitute for traditional literature reviews in research work. As a secondary research, QSR, especially those aggregative in nature, will greatly facilitate the development of cumulative knowledge and provide valuable information for those stakeholders (e.g., policymakers, practitioners). Moreover, a QSR well conducted might help to dismiss the accusations leveled against qualitative research such as its reliability and generalizability. We may not be able to undertake a full systematic review due to considerations of budget, time, resources, etc., but we can make our review more systematic by focusing on the critical steps and following the relevant guidelines (Booth et al., 2022). Above all, there is great potential for QSR to gain prominence in applied linguistics in the not too distant future.

REFERENCES

Avgousti, M. I. (2018). Intercultural communicative competence and online exchanges: a systematic review. *Computer Assisted Language Learning*, 31(8): 819-853.

Avgousti, M. I., & Hadjistassou, S. K. (2022). Exploring the (un-)explored in applied linguistics: Conducting a systematic review in intercultural communicative competence. In K. Dikilitas & K. M. Reynolds (Eds.), *Research methods in language teaching and learning: A practical guide* (pp. 288-310). John Wiley & Sons.

Barnett-Page, E., & J. Thomas. (2009). Methods for the synthesis of qualitative research: a critical review. *BMC Medical Research Methodology*, 9: 59.

Badjadi, N. E. I. (2015). A meta-ethnography of task representation in second language reading-to-write. *The Asian EFL Journal*, (86): 25-44.

Boland, A., Cherry, M. G., & Dickson, R. (2017). *Doing a systematic review: A student's guide.* London: Sage Publications.

Booth, A., Noyes, J., Flemming, K., Gerhardus ,A., Wahlster, I., van der Wilt, G. J., Mozygemba , K., Refolo, P., Sacchini, D., Tummers, M., & Rehfuess, E. (2018). Structured methodology review identified seven (RETREAT) criteria for selecting qualitative evidence synthesis approaches. *Journal of Clinical Epidemiology*, 99: 41-52.

Booth, A., Sutton, A., & Papaioannou, D. (2022). *Systematic approaches to a successful literature review* (3rd ed.). London: Sage Publications.

Britten, N., Campbell, R., Pope, C., Donovan ,J., Morgan, M., Pill, R. (2002).Using meta ethnography to synthesise qualitative research: a worked example. *Journal of Health Services Research & Policy*, 7(4):209-215.

Chen, T. P. (2016). Technology-supported peer feedback in ESL/EFL writing classes: a research synthesis. *Computer Assisted Language Learning*, 29 (2): 265-397.

Chong, S. W., & Plonsky, L. (2021). A primer on qualitative research synthesis in TESOL. *TESOL Quarterly*, 55(3):1024–1034. https://doi.org/10.1002/tesq.3030.

Chong, S. W., & Plonsky, L. (2023). A typology of secondary research in Applied Linguistics. *Applied Linguistics Review*. https://doi.org/10.1515/applirev-2022-0189

Chong S.W., & Reinders, H. (2020). Technology-mediated task-based language teaching: A qualitative research synthesis. *Language Learning & Technology*, 24(3), 70–86. http://hdl.handle.net/10125/44739

Dawson, A. J. (2019). Meta-synthesis of qualitative research. In P. Liamputtong (Ed.), *Handbook of research methods in health social sciences* (pp. 785-804). New York: Springer.

Ellis, R. (2015). Introduction: Complementarity in research syntheses. *Applied Linguistics*, 36(3): 258–259.

Ellis, R. (2018). Meta-analysis in second language acquisition research: A critical appraisal. *Journal of Second Language Studies*, 1(2): 231 – 253.

Fleming, K. (2009). Synthesis of quantitative and qualitative research: an example using Critical Interpretative Synthesis. *Journal of Advanced Nursing*, 66(1), 201–217.

Flemming, K., & Noyes, J. (2021). Qualitative evidence synthesis: Where are we at? *International Journal of Qualitative Methods*, 20. https://doi.org/10.1177/1609406921993276

France, E. F., Cunningham, M., &Ring, N., et al. (2019). Improving reporting of meta-ethnography: The eMERGe reporting guidance. *Review of Education*, 7: 430-451.

Hannes, K. (2019). Qualitative evidence synthesis. http://dx.doi.org/10.4135/9781526421036838863.

Jesson, J. K., Matheson, L., & Lacey, F. M. (2011). *Doing your literature review: Traditional and systematic techniques*. London: Sage Publications.

Khany, R., & Tazik, K. (2017). 40 Years of applied linguistics: Investigating content areas, research methods, and statistical techniques. *Chinese Journal of Applied Linguistics*, 40(3): 316-332.

Kohl, C., Mcintosh, E. J., Unger, S., Haddaway, N. R., Keche, S., Schiemann, J., &Wilhelm, R. (2018). Online tools supporting the conduct and reporting of systematic reviews and systematic maps: a case study on CADIMA and review of existing tools. *Environmental Evidence*, 7(8): 1-17.

Macaro, E. (2020). Systematic reviews in applied linguistics. In J. McKinley & H. Rose (Eds.), *The Routledge handbook of research methods in applied linguistics* (pp. 230-252). New York: Routledge.

Macaro, E., Curle, S., Pun, J., An, J., & Dearden, J. (2018). A systematic review of English medium instruction in higher education. *Language Teaching*, 51(1), 36-76.

Mendoza, N., & Phung, H. (2019). Motivation to learn languages other than English: A critical research synthesis. *Foreign Language Annals*, 52 (1): 121–140.

Noblit, G. W., & Hare, R. D. (1988). *Meta-ethnography: Synthesizing qualitative studies*. Newbury Park: Sage Publications.

Norris, J. M., & Ortega, L. (Eds.). *Synthesizing research on language learning and teaching*. Amsterdam: John Benjamins.

Ortega, L. (2011). *Doing synthesis and meta-analysis in applied linguistics.* https://slideplayer.com/slide/7072646/.

Pope, C., Mays, N., & Popay, J. (2007). *Synthesising qualitative and quantitative health research: A guide to methods*. Maidenhead (UK): Open University Press.

Stevenson, M. (2016). A critical interpretative synthesis: The integration of Automated Writing Evaluation into classroom writing instruction. *Computers and Composition*, 42: 1–16.

Strauss, A. L., & Corbin, J. (1990). Basics of qualitative research: Grounded theory procedures and techniques. Newbury Park, CA:Sage Publications.

Tellez, K., & Waxman, H. C. (2006). A meta-synthesis of qualitative research on effective teaching practices for English language learners. In J. M. Norris & L. Ortega (Eds.). *Synthesizing research on language learning and teaching* (pp. 245-278). Amsterdam: John Benjamins.

Thomas, J. & Harden, A. (2008). Methods for the thematic synthesis of qualitative research in systematic reviews. *BMC Medical Research Methodology*, 8(45).

AUTHOR'S BIO

***Jinpei Zhang** is associate professor at Shenzhen Institute of Information Technology, Guangdong, China. He holds a PhD in Applied Linguistics. His research interests include translation, intercultural communication and second language acquisition.*