RESEARCH SYNTHESIS IN APPLIED LINGUISTICS
https://doi.org/10.1080/29984475.2025.2456880

Routledge
Taylor & Francis Group

METHODOLOGICAL ARTICLE

OPEN ACCESS Check for updates

Synthesis Methods and Reporting Tool (SMART) for Research Syntheses in Applied Linguistics

Sin Wang Chong

International Education and Lifelong Learning Institute, University of St Andrews, St Andrews, UK

ABSTRACT

This article presents the Synthesis Methods and Reporting Tool (SMART), a guide on methodological and reporting practices for all types of research synthesis (RS) in applied linguistics. SMART is developed based on published RSs and methodological publications on RS in applied linguistics and RS checklists in other disciplines. To ensure SMART is field-specific, a group of RS specialists in applied linguistics provided input and feedback. SMART is underpinned by four principles (STAR): *systematicity*, *transparency*, *accessibility*, and *reflexivity,* and comprises three key stages (*preparation*, *method*, and *reporting*); it aims to cover all procedures and practices that need to be considered before, during, and after conducting an RS. A unique feature of SMART is the inclusion of baseline and preferred practices: 'baseline practices' refer to minimum requirements that an RS needs to meet while 'preferred practices' provide suggestions that researchers are encouraged to implement. The introduction of SMART provides the applied linguistics community, especially those new to RS, with an easy-to-use tool to guide RS methodologies. While the intention is not to replace other RS tools, SMART is developed to be responsive to the development, standard, and characteristics of RS in applied linguistics.

ARTICLE HISTORY
Received 18 December 2024
Accepted 8 January 2025

KEYWORDS
Research synthesis; applied linguistics; methodology; secondary research

Research Syntheses in Applied Linguistics: Major Milestones

Research synthesis (RS) refers to the review of literature that uses transparent, systematic, and often pre-registered methodology to identify, extract, and synthesise information from literature, guided by research questions (Chong and Plonsky 2024b). In applied linguistics, RS has been around for longer than most think. One of the earliest RSs was published 40 years ago (Willig 1985), which is a meta-analysis on the effectiveness of bilingual education, published in *Review of Educational Research*. An examination of the two largest bibliographies of RS in applied linguistics hosted by Luke Plonsky and the Research Synthesis in Applied Linguistics Special Interest Group of the British Association for Applied Linguistics (BAAL), a total of 754[1] RS have been published (Figure 1). This number excludes works in progress including RS protocols, which are study plans for ongoing RSs. It is without

CONTACT Sin Wang Chong swc5@st-andrews.ac.uk

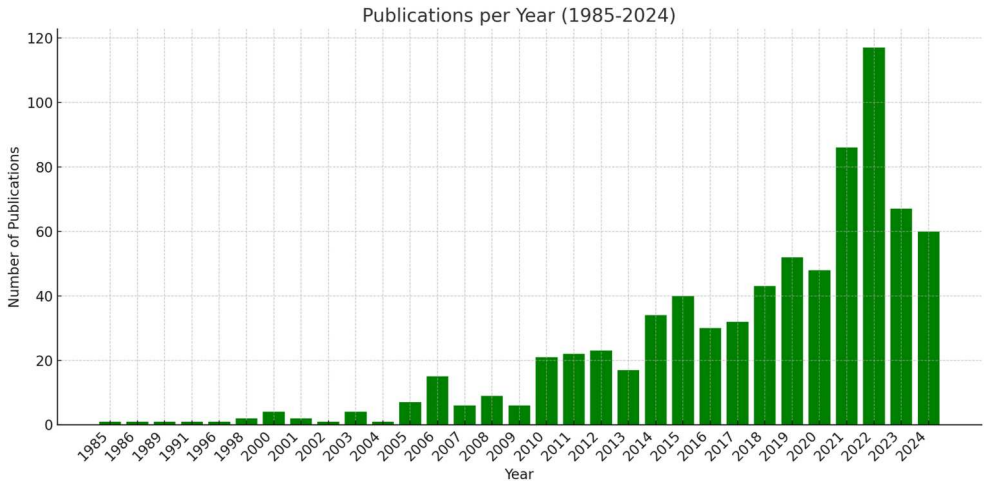

Figure 1. Number of RS publications in applied linguistics from 1985 to October 2024.

exaggeration to say that RS has become integral to the applied linguistics research landscape, especially in the past decade or so, and is usually highly cited (Amini Farsani et al. 2021). Apart from quantity, different types of RS have been published in recent years. While meta-analysis constitutes the majority of the RS published, other types of RS such as scoping reviews, systematic reviews, bibliometric reviews, qualitative research syntheses, and methodological syntheses have started to emerge since 2010.

Some milestones have contributed to the flourishing of RS in applied linguistics. As early as the 1980s, Chaudron (2012) referred briefly to the usefulness of meta-analysis in his book on L2 classroom research. The first systematic exploration of RS in the field was by Norris and Ortega (2006) in their seminal text 'Synthesising Research on Language Learning and Teaching'. This was a collection of synthetic studies and commentaries focusing mostly on meta-analysis. In the following year, Norris and Ortega (2007) published an opinion piece in *TESOL Quarterly* about the values of RS in applied linguistics, arguing that well-conducted RS can inform future research directions, research practices, and broaden the spectrum of primary studies in applied linguistics.

Language Learning, *Language Teaching*, and *Applied Linguistics* are the first applied linguistics journals to publish RS, with some of their meta-analyses published as early as early 2000s (e.g., Masgoret and Gardner 2003; Ortega 2003). In the past few years, more and more journals have regarded RS as a distinct submission category. Some examples include *Language Learning*, *Language Teaching*, *Innovation in Language Learning and Teaching*, *Language Testing*, and *Studies in Second Language Acquisition*. Recently, several special issues have been published in journals that focus on RS. The first special issue on RS was co-guest edited by Rod Ellis, Shaofeng Li, and Natsuko Shintani *for Applied Linguistics* in 2015, entitled 'Synthesizing Research on Form-Focused Instruction: The Complementary Contributions of Narrative Review and Meta-analysis' (Ellis 2015; Li 2015; Shintani 2015); Plonsky (2023) edited a special issue on 'Bibliometrics in Applied linguistics' for *Studies in Second Language Learning and Teaching*; Chong, Bond, and Chalmers (2024a) edited a special issue on 'Research Synthesis in Language Learning and Teaching' for *Applied Linguistics Review*.

The vibrant development of RS in applied linguistics has been propelled by the proliferation of publications on RS methodologies, distinguishing RS from traditional literature reviews (Li and Wang 2018). These publications introduce, survey, and/or critique methodological practices of RS including meta-analysis (Plonsky et al. 2023; Plonsky and Oswald 2012a, 2012b), systematic review (Macaro 2019), qualitative research synthesis (Chong and Plonsky 2021; Chong and Reinders 2021), and innovative types of RS such as meta-reviews (Chong et al., 2025), practice review (Chong and Plonsky 2024a), and mixed methods review (Riazi and Amini Farsani 2024). There are applied linguistics journals that have developed author and/or peer reviewer guidelines for RS such as those by *Innovation in Language Learning and Teaching* and *Language Testing*. There is also the first applied linguistics journal that focuses on research methodologies, *Research Methods in Applied Linguistics* and the first journal that focuses specifically on RS, *Research Synthesis in Applied Linguistics*.

Other more recent initiatives on RS in applied linguistics focus on community building, such as the establishment of the Research Synthesis in Applied Linguistics Special Interest Group (researchsynthesis.weebly.com) by the British Association for Applied Linguistics (BAAL) in 2022, the first and only research network on RS in applied linguistics. The group, at the time of drafting the current article, has over 300 members (mostly UK-based) and organises regular seminars and workshops on conducting RS. Their website hosts a repository of and resources for conducting RS in applied linguistics. The group sends out monthly newsletters to their members that include a monthly pulse survey, the most recently published RS, and information about RS software and tools. Focusing on English teachers, *TESOLgraphics* (www.tesolgraphics.com), which was founded in 2021, summarises RSs in TESOL in one-page infographics to make synthesised findings accessible and usable to teachers (Sato et al. 2024). OASIS (https://oasis-database.org/), which produces accessible summaries of research articles, also publishes lay summaries by authors of RS.

Types of Research Syntheses in Applied Linguistics

While meta-analysis remains the dominant type of RS in applied linguistics (e.g., Ren, Li, and Lu 2023), the diversification of topics and various 'turns' in the field has contributed to the growth of different types of RS. These relatively new types of RS, including methodological synthesis, qualitative research synthesis, systematic review, bibliometric review, mixed review, and scoping review, have been documented in a typology of secondary research in applied linguistics (Chong and Plonsky 2024b).

Other types of RS that is prominent in disciplines outside of applied linguistics have started to emerge in our field, as shown in some works-in-progress. These include practice review, meta-review, meta-ethnography, framework synthesis, and mixed methods RS. *Practice review* focuses on practice-related information (e.g., pedagogical approaches) through a synthesis of academic and non-academic sources. Kong, Hopkyns, and Chong (2024) published a protocol on an ongoing practice review on translanguaging pedagogies in Anglophone countries. *Meta-review*, or sometimes called an 'umbrella review', 'review of reviews', or 'meta systematic review', synthesises published RS in lieu of primary studies. The first meta-review in applied linguistics has been Plonsky and Ziegler (2016), which is a second-order synthesis on the interface between

computer-assisted language learning and second language acquisition. More recent examples include a meta-review on meta-analyses in L2 research (Vuogan and Li 2024) and an ongoing meta-review commissioned by BAAL on the state-of-play of applied linguistics (Chong, Nie, and Liu, forthcoming). *Meta-ethnography* focuses on extracting, comparing, and reinterpreting metaphors and concepts in primary studies (Atkins et al. 2008). Although I am not aware of any meta-ethnography published in applied linguistics, the rise of ethnographic (including autoethnographies) and narrative inquiry studies makes meta-ethnography a potentially relevant addition to the RS repertoire. *Framework synthesis*, as the name suggests, employs an established framework as an analytical lens to synthesise research insights (Brunton, Oliver, and Thomas 2020). Shen and Chong (2023) employs Ellis' (2010) feedback engagement framework when synthesising evidence on written corrective feedback. Mixed methods RS are syntheses that combine quantitative and qualitative in a single review or conducting multiple reviews focusing on quantitative and qualitative evidence respectively, with insights from these reviews combined into a final review (The Joanna Briggs Institute 2014).

Rapid review and *evidence gap map* are two other kinds of RS that are yet to be featured in applied linguistics but are potentially relevant to the field. Rapid reviews are RS that employ a streamlined process in identifying and analysing literature (Garritty et al. 2024); the urgency of the topic (e.g., the shift to online language teaching during COVID) outweighs the limitations of employing a less thorough methodology. Sometimes, rapid reviews are conducted to accommodate other constraints such as the lack of resources and time for conducting an RS in a funded project. While I have not come across a rapid review in applied linguistics, there are RS that use 'semi-systematic reviews' to refer to this kind of streamlined RS (e.g., Rose et al. 2018). Evidence gap maps, which can be used as part of a living systematic review (Elliott et al. 2017), are visualisations that accompany an RS; their aim is to provide an overview of 'hot spots' and 'cold spots' in a research topic, exemplified using circles of different sizes and colours such as those generated by EPPI-Reviewer. I do not advocate for the compulsory use of research software in RS as technological use needs to be fit-for-purpose, but I support the use of visual aids including tables, diagrams, figures, and maps to present an overview and trends of synthesised data. Although it is irrelevant to applied linguistics and not an RS, one interesting example is the representation of the state-of-play of higher education research as 'research archipelago' by Macfarlane (2012), with each 'island' representing an area of research in higher education (e.g., teaching and learning island) and its size showing its vibrancy.

Table 1 summarises the types of RS in applied linguistics, building on Chong and Plonsky (2024b). The table only includes types of published RS or RS protocols. Different from Chong and Plonsky (2024b), Table 1 focuses on RS but not secondary research in general. I am often asked about methodological frameworks for different types of RS. A closer look at the various types of RS shows that RS types differ from one another in terms of their purposes but not so much their methodologies. When making a decision about which type of RS to conduct, the dealbreaker is almost always the purpose of the review. I categorise the eight types of RS in applied linguistics into three review families based on their purposes: configurative review, exploratory review, and explanatory review. It is important to note that these categories can be conflated.

Table 1. Types of RS in applied linguistics.

Review family	Review type	Purpose	Example
Configurative review	Methodological synthesis	To review how a specific research method or methodology is employed in published studies	Ghanbar et al. (2024)
	Meta-review	To review evidence presented in RS and the quality of RS	Chong, Nie, and Liu (forthcoming)
	Practice review	To review information related to practice e.g., translanguaging pedagogies from academic and non-academic sources	Kong, Hopkyns, and Chong (2024)
	Qualitative research synthesis	To review qualitative evidence presented in primary studies	Chong and Reinders (2020)
	Systematic review	To review qualitative and quantitative evidence presented in primary studies comprehensively	Macaro et al. (2018)
Exploratory review	Bibliometric review	To review quantitative information and patterns of publications e.g., topics, authors, sources, citations	Zakaria and Aryadoust (2024)
	Scoping review	To preliminarily survey qualitative and quantitative evidence on a niche or emergent topic	Chong and Reinders (2022)
Explanatory review	Meta-analysis	To review overall effectiveness of interventions in relation to effect sizes and variance/correlational measures	Teimouri, Goetze, and Plonsky (2019)

Configurative reviews focus on meaning-making through the interpretation of synthesised information. Unlike primary qualitative studies that also fall under the interpretivist paradigm, this meaning-making process usually focuses on the analytic rather than the evocative, with an aim to produce a meta-framework (e.g., development of a methodological or conceptual framework) or meta-narrative on the topic. For example, Çiftçi and Savaş (2018) conducted a qualitative research synthesis to develop a conceptual framework for understanding telecollaboration in language and intercultural learning. Morea and Ghanbar's (2024) methodological synthesis on Q methodology in applied linguistics presents a meta-narrative about research areas that Q methodology has been applied to and characteristics of these studies pertaining to methods and analytical techniques.

Exploratory reviews are reviews that ask broad research questions and are usually for topics that are emergent, niche, or rapidly expanding. Given the preliminary nature of exploratory reviews, they are usually followed by a configurative or confirmatory review. For instance, one purpose of a scoping review, a type of exploratory review, is to identify potential topics for other RS (Pham et al. 2014). *Explanatory reviews* such as meta-analyses employ quantitative analysis techniques to describe the relation between

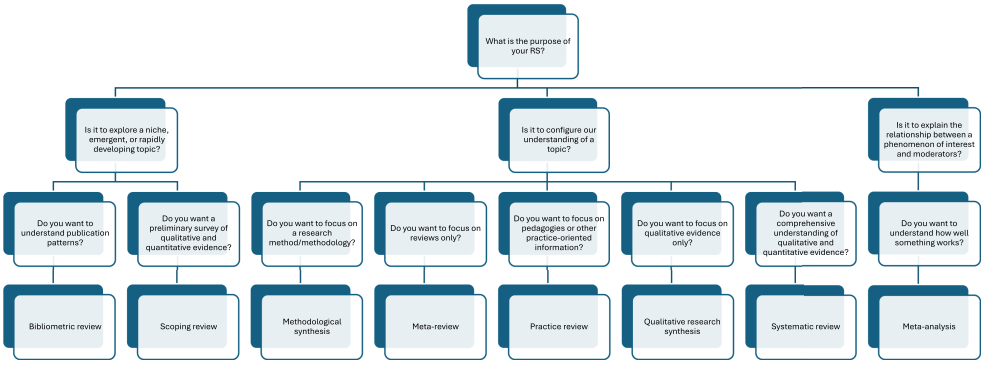

Figure 2. Decision-making for determining the type of RS.

a phenomenon of interest (e.g., the effect of feedback on grammar) and one or more theoretically relevant moderator, one relation at a time (Norouzian and Bui 2024).

When making a decision on the type of RS to conduct, follow the steps in Figure 2.

Synthesis Methods and Reporting: A Tug of War between Standardisation and Diversity

Appraisal of quality is of utmost importance because RSs are often perceived as authoritative publications, and they can be influential in terms of shaping future research directions of a field and determining citations of specific publications to an extent. Therefore, it is crucial that RS are conducted in the most rigorous, systematic, and transparent manner possible, following methodological and quality-assurance guidelines such as SMART. These guidelines can provide useful a practical blueprint for researchers when conducting RS and evaluating the quality of RS.

In disciplines where RS has been more well-established, RS appraisal guidelines have been developed. The one that is most widely known to the applied linguistics community is probably the Preferred Reporting Items for Systematic Review and Meta-Analysis Protocol (PRISMA-P) (Moher et al. 2015) and the Preferred Reporting Items for Systematic Review and Meta-Analysis (PRISMA) (Page et al. 2021). PRISMA-P is a guide for writing an RS protocol while PRISMA focuses on reporting and includes a diagram template for documenting study selection process and outcomes. In addition to PRISMA, there are other guidelines such as the Database of Abstracts and Reviews of Effects (DARE) tool (Centre for Reviews and Dissemination 1995), focusing mainly on inclusion and exclusion criteria, representativeness of the included literature, study appraisal, and description of individual studies in computer science and social sciences. Developed for RS in healthcare, AMSTAR and the revised version, AMSTAR 2 (Shea et al. 2017), is an appraisal tool that allows researchers to score RS based on critical and non-critical methodological weaknesses, although their focus is on RS on intervention studies. Designed for systematic reviews on higher education research, Chong, Lin, and Chen (2022) developed a bottom-up, 20-item methodological checklist based on their analysis of the method section of 160 systematic reviews.

There have been attempts to modify existing RS guidelines and tools to account for differences in evidence synthesis methods. For example, Bond et al. (2024a) combined DARE and AMSTAR 2 to develop a checklist for appraising the quality of RS included in their meta-review on artificial intelligence in higher education. This is a case in point that RS conducted in various fields or disciplines may differ due to the influence of dominant research paradigms, research designs of primary studies, and research questions. The benefits of developing a set of discipline-specific RS guidelines are clear but we need to consider the readiness of researchers in a field; while it is crucial to avoid disciplinary silos for methodological advancement, it would be unwise and impractical to impose a new set of standards rigidly and expect everyone to adopt it without being critical about its applicability, while criticising previous work as sub-par. At the same time, it is my aspiration that, through starting this journal as a dedicated platform for RS in applied linguistics and developing SMART, applied linguistics researchers will become more cognisant of RS methodologies and gradually come to a consensus of a minimum of scientific and reporting rigour of RS in our field.

In fact, while the development of these tools and guidelines is well-intended, there are issues related to how these standards are understood and used. Rader et al. (2014) examined how PRISMA was used in medicine through a survey. Findings suggest that even within the same discipline, there are challenges in understanding the meaning of some items in PRISMA as well as other practical constraints such as the lack of time and resources. In applied linguistics, Chong and Reinders (2021) conducted a methodological synthesis on how qualitative research syntheses on CALL (computer-assisted language learning) were conducted using a field-specific framework (Chong and Plonsky 2021); they identified areas that need to be improved on such as distinguishing between data extraction and synthesis. Chalmers et al. (2024) employed PRISMA to analyse the methodological rigour of published RS in applied linguistics and found that only about half of the RS included essential information such as search terms used, inclusion/exclusion criteria, and data collection.

I argued in Chong, Lin, and Chen (2022) that, while standardisation of how RS is conducted and reported is necessary for methodological discussion and improvement, it is important to recognise the double duty of such tools to encourage both heterogeneity and homogeneity. In other words, RS tools are most useful when they can illustrate a baseline standard while acknowledging that there are variations in how RS is conducted depending on their purposes and research questions. When developing Synthesis Methods and Reporting Tool (SMART) for applied linguistics RS, I am aware that RS is a burgeoning type of research with much of our earlier effort dedicated to meta-analyses (see Figure 1). That is why I decided to write this paper to accompany SMART, starting with the origin and current state-of-play of RS in applied linguistics to make a case for how I envisage SMART to be used: one that applied linguistics researchers can elect to use alongside other established guidelines like PRISMA, one that is user-friendly to researchers new to RS, and one that establishes baseline methodological practices while encouraging researchers to be flexible and innovative depending on the purpose of the RS. Another goal I want to achieve with SMART is to identify dimensions of quality that are core to RS in applied linguistics, similar to how we evaluate quantitative studies using validity and reliability and qualitative studies through trustworthiness and dependability etc.

Four Dimensions of RS Quality: STAR

I developed SMART based on four dimensions of RS quality: *systematicity*, *transparency*, *accessibility*, and *reflexivity* (STAR). STAR aligns with what Plonsky (2024) suggested in his four-part framework of study quality: methodologically rigorous, transparent, ethical, and of value to society but focuses on secondary research. An RS is systematic when it adheres to principles of best practice in secondary research (such as SMART), and it is theoretically grounded; it is transparent when an RS reports in detail its methodological steps and decisions, and is pre-registered (i.e., publishing information about an RS' focus and methodology in advance) (Chong et al. 2024b). I will explain later that there are different ways to pre-register an RS, with protocol (a study plan for an RS) being one of them. An RS is deemed accessible and inclusive when there is evidence to suggest researchers and practitioners (broadly defined as those whose main responsibilities in their vocations are not doing research) collaborate in the review process and/or when synthesised evidence is presented in a reader-friendly way for a non-academic audience.

Accessibility is especially important for applied linguistics RS because a lot of the topics in our field are applied in nature, and it is an ethical imperative that evidence can be usable by practitioners. For an interesting example of the accessibility and inclusivity of RS practices, see Cochrane's *The People's Review* project (www.thepeoplesreview.ie). Finally, reflexivity, which refers to the researchers' awareness of how methodological decisions are shaped by their own experiences, backgrounds, as well as the contexts they are in, is key to high-quality RS because it captures not only methodological actions but the process through which methodological decisions are made; this echoes what I said earlier about encouraging heterogeneity in RS practices. What I wish to encourage through SMART and STAR is that researchers in applied linguistics can be innovative in RS methodologies, not at the expense of rigour, but to answer questions that are most relevant to our community.

Items of SMART

SMART comprises three components and 22 items: preparation (four items), method (seven items), and reporting (11 items) that seek to cover every stage and decision you need to consider prior to, during, and after you conduct an RS (Tables 2–4). A checklist version of SMART is in Appendix A for authors and reviewers to use. Different from other RS tools, SMART is designed to include both 'baseline' and 'preferred' practices. 'Baseline practices' are those that are essential to conducting a high-quality RS while 'preferred practices' are recommendations for researchers to consider. The decision to include both baseline and preferred practices is to make SMART a tool for appraisal and professional development. Researchers can use SMART to self-assess the quality of RS they are conducting, and journal editors or reviewers can use it to determine the rigour of RS and offer suggestions. At the same time, the preferred items raise awareness about good RS practices from other disciplines but which may not be immediately applicable to our field.

Items in SMART are mostly informed by existing tools such as DARE, AMSTAR 2, PRISMA, and PRISMA-P, while some are added based on my experience conducting, reviewing, and analysing RS in applied linguistics. While SMART is developed as part of the conceptualisation and launching of, *Research Synthesis in Applied Linguistics*, the first journal dedicated to RS in our field, I want it to be as relevant and useful to applied linguists as a whole. Therefore, I followed a Delphi process to develop SMART. An open call was issued to invite applied linguists at different career stages to contribute to the development of SMART. I have received a total of 15 expressions of interest, including doctoral researchers, early career researchers, senior academics, journal peer reviewers, and editors in applied linguistics. They represent diverse research interests in applied linguistics including language education, language testing, and second language acquisition. These 15 researchers formed a working group to provide input and feedback on SMART; the editorial board of the *Research Synthesis in Applied Linguistics* journal was also invited to provide input and comments. Before development, they were asked to indicate what they perceive as important elements in SMART. Considering all their input, I developed a draft of SMART, which was shared with the working group and editorial team of *Research Synthesis in Applied Linguistics* for written comments. I then made revisions based on their feedback. The names of

Table 2. 'Preparation' items in SMART.

PREPARATION			
	Item	Baseline Practice (BP)	Preferred Practice (PP)
P1	Originality of RS	Conduct **a preliminary search** on published RSs to ensure that your RS idea is not a duplicated effort.	BP + Reflect on the **original contributions** of your RS, especially in light of recent RSs on the same/similar topics e.g., your RS focuses on a different educational context, the types of evidence you plan to synthesise are different, the rapid development of a topic warrants a new RS (e.g., new uses of AI).
P2	RS team	Assemble **a team of researchers** (at least 3) to work on the RS. One should be an expert on the substantive topic of the RS and the other should have experience with conducting RSs.	BP + Involve **stakeholders** who are related to the topic of your RS such as teachers, school leaders, policymakers. Their role is to review the synthesised findings to ensure they address their needs and are presented in an accessible way to a non-academic audience.
P3	Type of RS	Discuss the purpose of your RS idea with the team and decide on the **type of RS**. Refer to **typologies of RS**. Sutton et al. (2019), for example, identified 48 types of review. For a typology that is specific to applied linguistics, see Chong and Plonsky (2024b). For emergent types of RS in applied linguistics, see Chong and Plonsky (2024a).	BP + It is essential to **maintain consistency when naming the types of research synthesis (RS)** e.g., avoid using 'meta-analysis' when referring to 'meta-synthesis,' which is a specific type of qualitative RS. Additionally, prefer the term 'qualitative research synthesis' over 'qualitative evidence synthesis,' as the former is more widely used in applied linguistics. Consistent naming helps researchers search for and distinguish between different types of RS.
P4	RS registration	Complete a **title registration** form with information about the planned RS: tentative title, type of RS, research questions, personnel, and timeline. The information of the registered RS will be published on a sister website of *Research Synthesis in Applied Linguistics* (T&F). This ensures that there is no duplicated RS and allows researchers with similar interests to explore opportunities to collaborate.	BP + Publish an open-access and peer-reviewed **protocol** on platforms such as *Research Synthesis in Applied Linguistics* (T&F)and *International Database of Education Systematic Reviews* (https://idesr.org/). The protocol should be structured and contain the following information: background and review rationale; review questions; inclusion and exclusion criteria; search strategy; data extraction and management; study appraisal (if applicable); data synthesis; reporting; personnel; conflict of interest; timeline; references; supplementary materials (if applicable). If the protocol cannot be published open access, a pre-print or post-print should be made available on research databases/institutional repositories.

researchers who provided feedback are listed in the acknowledgement section of this paper unless they have indicated otherwise.

SMART is a 'living' tool. As RS methodologies in applied linguistics continue to develop, the items will be reviewed periodically, by adding new items, modifying existing items, and/or making some preferred practices compulsory. I outline how I intend to make SMART relevant and contemporary in the 'conclusion' section.

"Preparation" Items in SMART

There are four items related to the preparation of an RS (items P1 to P4) (Table 2). These items relate to the originality of RS, the research team conducting the RS, the type of RS to be conducted, and pre-registration of RS. Originality of RS prompts researchers to

Table 3. 'Method' items in SMART.

		METHOD	
	Item	Baseline Practice (BP)	Preferred Practice (PP)
M1	Research questions	Specify the **research questions** that the RS intends to address.	BP + Demonstrate **compatibility between research questions and RS type** e.g., the research questions are broad for a scoping review, which is exploratory in nature.
M2	Inclusion and exclusion criteria	Include **specific criteria** to inform study selection. These criteria should encompass the following: publication years, types of publications, languages. There can be additional criteria based on the focus of the RS.	BP + Specify the **rationale** for having the criteria e.g., reasons for focusing on publications in the past decade.
M3	Search strategy	List out **all the databases** where the search is conducted. Include 3–4 databases (not websites) with additional search strategies (e.g., backward and forward reference searching). Google Scholar is not recommended because it is impossible to screen all search results. List out the **search string(s)** used to conduct the search. Include different versions of the search string if necessary for different databases or for searches in another language. Note that search strings are keywords combined using Boolean operators. When developing a search string, it is recommended to consult a subject librarian.	BP + Include 4 or more databases (not websites) with additional search strategies. Specify the **rationale** for choosing specific databases (and websites), and using particular search string(s).
M4	Screening	Perform **first-level screening** on titles and abstracts based on the inclusion and exclusion criteria. Perform **second-level screening** on full texts based on the inclusion and exclusion criteria. For second-level screening, at least 2 reviewers need to be involved in screening a minimum of 10% of the publications. **Inter-coder reliability** needs to be calculated (e.g., percentage, kappa).	BP + Consider **reflexivity** by documenting the process of discussions and resolving conflicts in second-level screening e.g., by keeping a researcher logbook.
M5	Data extraction	Develop a **data extraction form** based on research questions. **Pilot the data extraction form** on a minimum of 10% of the included publications, with the involvement of at least 2 reviewers. Revise the form if needed based on feedback. **Inter-coder reliability** needs to be calculated (e.g., percentage, kappa).	BP + Consider **reflexivity** by documenting the process of discussions and resolving conflicts in second-level screening e.g., by keeping a researcher logbook.
M6	Study appraisal	Evaluate the methodological rigour of included studies using **existing tools/ checklists** such as Critical Appraisl Skills Programme (CASP; https://casp-uk.net/casp-tools-checklists/). Adaptations may be needed for applied linguistics studies. For a field-specific tool, see Mahboob et al. (2016). Depending on the type and purpose of the RS, methodological rigour may not be an a priori reason for inclusion/exclusion.Study appraisal does not apply to exploratory reviews i.e., scoping review, bibliometric review, and rapid review, or RS that does not examine results.	BP + Be transparent about **how the tools/ checklists are used** to inform decisions about including or excluding studies due to methodological rigour e.g., adopting a scoring system.

(*Continued*)

Table 3. Continued.

		METHOD	
	Item	Baseline Practice (BP)	Preferred Practice (PP)
M7	Data synthesis	Analyse the extracted data qualitatively and/or quantitatively based on RS type and research questions. Use of **software or automation tool** is not required although recommended, especially when conducting configurative and explanatory RSs. If used, it needs to be documented. When coding data qualitatively, at least 2 reviewers need to be involved in a minimum of 10% of the included publications. **Inter-coder reliability** needs to be calculated (e.g., percentage, kappa). Specify how your data are coded e.g., thematic analysis, grounded theory. When analysing data quantitatively (e.g., for meta-analysis or bibliometric review), employ the relevant statistical tests.	BP + Include materials that document the **process of synthesis** such as the coding scheme. Share synthesised findings with **practitioners** to obtain feedback on the usefulness of findings and ways they are represented.

avoid 'research waste' (Isaacs and Chalmers 2023) by conducting a search on published RS on similar topics (P1-BP) and reflecting on the original contributions, the planned RS intends to make (P1-PP). The research team that will undertake the RS needs to be well thought through to include expertise related to the substantive focus of the RS, methodology of RS, information science (P2-BP), and preferably involve practitioners who may have a direct interest in the topic of the RS (P2-PP). As mentioned earlier, I adopt a broad definition of practitioners to refer to anyone whose primary role is not to conduct research in their vocation. In this sense, practitioners can include teachers, translators, policymakers, lawyers, and police; the list goes on depending on the topic of your RS and who the stakeholders are. Having pondered on the originality and team composition of your RS, the next decision to make is the type of RS to conduct. One starting point is to refer to the decision-making tree in this paper (Figure 2) and refer to existing RS typologies for the sake of consistency (P3-BP). Being uniform with naming the type of RS facilitates future searching of relevant RS and discussions on methodologies related to a particular kind of RS (P3-PP). Finally, pre-registration is essential to avoid 'research waste' (Isaacs and Chalmers 2023) and foster collaboration; however, given pre-registration of RS is not a common practice in applied linguistics, I propose title registration as an interim required practice (P4-BP). An online form[2] has been set up for researchers to title register an RS that they plan to conduct, as well as to view RSs that are already registered.[3] The information required will be minimum, including the tentative title, type of RS, research questions, personnel, and timeline. Ideally, researchers will publish a protocol of their RS prior to conducting one (P4-PP).

"Method" Items in SMART

A total of seven items are included to describe the crucial methodological stages in RS: research questions, inclusion and exclusion criteria, search strategy, screening, data extraction, study appraisal, and data synthesis (Table 3). 'Research questions' invite

Table 4. 'Reporting' items in SMART.

	REPORTING		
	Item	Baseline Practice (BP)	Preferred Practice (PP)
R1	Title	Include a title and a subtitle, separated by a colon. The title is the topic of the RS, and the subtitle is the type of RS.	BP + Refer to RS types in Table 1 to ensure consistency.
R2	Abstract	Similar to an abstract for a primary study, include the problem, RS objective(s), **methodological guidelines** that inform the RS, major findings, and implications. The type of RS mentioned in the abstract should be consistent with that in the title.	BP + Create **a graphical abstract** (e.g., infographic) or **lay summary** for a non-academic audience. *Research Synthesis in Applied Linguistics* (Taylor & Francis) collaborates with *TESOLgraphics* (http://tesolgraphics.com/) to create open-access infographics for all published RSs in the journal.
R3	Introduction	Similar to an introduction for a primary study, present the **problem/issue** and **research questions** of the RS.	BP + Mention the **unique contributions** a RS can make to the topic.
R4	Literature review	Similar to a literature review for a primary study, **present and critique relevant publications** on the topic. Present **theoretical and/or conceptual framework** if relevant. Sometimes there are studies that cannot be included in the RS due to inclusion/exclusion criteria or studies are published after data extraction/synthesis is completed. In this case, these studies can be mentioned here (and in the discussion section).	BP + Discuss and critique **published RSs** on the/a cognate topic.
R5	Methodology	Structure the section based on the **Method** component of SMART or other guidelines (Table 4) and include all elements. It should be an **elaboration** of the information in the protocol, focusing on the process and outcome of decision-making. Whenever possible, exemplify your method e.g., include a coding scheme, a funnel plot. Use of **automation tools/software** needs to be reported, explained, and justified. Report and justify any **deviations from the protocol**. Report any **limitations** e.g., selection bias	BP + Add a section on **team credential** to show complementary expertise and professional experience of the team Add a section on **contextual sensitivity**. In applied linguistics, context (e.g., cultural, educational, or policy-related factors) plays an important role. Report how context influenced the methodological choices and outcomes of the RS.
R6	Findings	Summarise **individual study characteristics** e.g., in a table, including quality appraisal (if applicable). Present findings **thematically**, with each theme supported by **illustrations** from representative individual studies.	BP + Whenever possible or appropriate, use **visual aids** e.g., tables, diagrams, figures, evidence gap map to present findings in an accessible manner.
R7	Discussion	Discuss **the meaning and importance** of the findings e.g., using a conceptual/theoretical framework. **Compare and contrast** findings with previous primary studies and RSs.	BP + Discuss relevance of findings to **non-academic audience** such as practitioners in concrete terms.
R8	Conclusion	Reiterate **headlines** of synthesised findings Report future **research directions**	BP + Reflect on how **future RSs** can be done differently to make an original contribution to the topic.
R9	Acknowledgements	Include information about support received from individuals and/or organisations (e.g., funders), if relevant	BP + Include specific contributions of each author in the RS by referring to/modifying the CRediT statement (https://onlinelibrary.wiley.com/doi/full/10.1002/leap.1210).

(*Continued*)

Table 4. Continued.

		REPORTING	
	Item	Baseline Practice (BP)	Preferred Practice (PP)
R10	References	Include all cited sources in **APA 7th format**. Indicate publications included in the RS by putting **an asterisk (*)** in front of them.	N/A
R11	Appendices	Include any **tools/instruments** used e.g., study appraisal tool and/or **data** e.g., complete search strategy/Boolean strings used for each database queried, coding scheme, completed extraction forms.	BP + Make the tools, instruments, and/or data **publicly available** on an open science site (e.g., OSF) and/or research repository.

researchers to consider the objectives of an RS (M1-BP) as well as the compatibility between research questions and RS type (M1-PP). 'Inclusion and exclusion criteria' ask researchers to spell out the criteria used to inform decisions of including publications in an RS. These criteria usually cover areas such as publication years, types of publications (including outlet types e.g., journals, books, conference proceedings and nature of publications e.g., editorial, primary study, and opinion piece), and languages. With the primary focus of applied linguistics research on the use of languages, which is highly contextualised, applied linguistics researchers are encouraged to consider including publications in different languages. This decision should be based on the diversity of language users in the research team and the specific topic of the review (M2-BP); preferably, justifications are provided regarding specific criteria. For instance, the reason for focusing on publications in the past decade (M2-PP). Item M3 concerns 'search strategy'; information that needs to be included is: the databases and search string(s) used in the search. To ensure comprehensive coverage of RS, databases should be the primary means through which literature is identified (M3-BP). Websites such as journal or publisher websites can be used in a supplementary manner. RS are perceived as comprehensive and dependable pieces of research, and their coverage needs to be exhaustive. To achieve this, SMART suggests RS in applied linguistics include three to four databases with additional search strategies such as background or forward reference searching (M3-BP). Preferred practices pertaining to search strategy include explaining the reasons for focusing on specific databases (and websites) as well as including four or more databases to ensure comprehensiveness of the search (M3-PP).

Item M4 is about the screening of potentially relevant literature. According to M4-BP, first-level and second-level screenings need to be conducted. First-level screening focuses on titles and abstracts, reviewing these elements based on the inclusion and exclusion criteria to determine a 'shortlist' of publications. Second-level screening further scrutinises the shortlist by reviewing the full texts of these publications against the inclusion and exclusion criteria. A minimum of two reviewers need to be involved in screening at least 10%[4] of the publications in the shortlist (O'Connor and Joffe 2020); disagreements need to be discussed and resolved, and inter-coder reliability calculated (e.g., in the forms of percentage or kappa[5]). Inter-coder reliability needs to be at least 80% (Belur et al. 2021) or 0.75 in kappa value (Lange 2011); if agreement is below this threshold, another 10% of the publications need to be screened and the process continues until the accepted inter-coder reliability is achieved. Researchers are encouraged to consider reflexivity, one of the core dimensions of RS quality (Table 5), in the screening process (M4-PP). In addition to

Table 5. Four dimensions of RS quality.

	Dimension of quality	Definition	SMART items*						
☆	Systematicity	Principles of best practice in secondary research (such as those outlined in SMART and STAR) are adhered to, and the RS is theoretically/conceptually grounded.	M1-BP	M2-BP	M3-BP	M4-BP	M5-BP	M6-BP	M7-BP
			M1-PP	M2-PP	M3-PP	M4-PP	M5-PP	M6-PP	M7-PP
			R1-BP	R2-BP	R3-BP	R4-BP	R5-BP	R6-BP	R7-BP
			R8-BP	R9-BP	R10-BP	R11-BP	R1-PP	R2-PP	R3-PP
			R4-PP	R5-PP	R6-PP	R7-PP	R8-PP	R9-PP	R10-PP
			R11-PP						
☆	Transparency	All methodological steps and decisions are reported in the RS. The RS is pre-registered and has relevant information available before its publication.	P1-BP	P2-BP	P3-BP	P4-BP	R11-PP	P1-PP	P2-PP
			P3-PP	P4-PP					
☆	Accessibility	Stakeholders are involved in the RS. Findings are relevant to stakeholders and presented in an accessible manner.	P2-PP	M7-PP	R2-PP	R6-PP	R7-PP		
☆	Reflexivity	The RS reports not only the outcome but the process of methodological decision-making, as well as factors that influence the decisions.	M4-PP	M5-PP	R5-PP				

* All shaded items are compulsory to qualify an RS to achieve a dimension of quality while those that are not shaded are optional items.

reliability, reflexivity, which refers to researchers' awareness of how their positionality, experience, and background may affect them in the research process, is equally important in RS (Chong, Liu, and Tegama 2025). It is because researchers of RS were not involved in the studies they synthesise and may not be familiar with the contexts of those studies; equally, there are challenges pertaining to the power difference in the review process. Imagine a research assistant discussing whether to include a publication with a principal investigator and they have different opinions. It is possible that the research assistant feels obliged to agree with the principal investigator, resulting in a relatively high inter-coder reliability. But what inter-coder reliability does not tell us is the processes through which agreements are reached and disagreements are resolved. One way to demonstrate reflexivity in RS is through keeping a researcher logbook, which is used to document the process through which agreement is reached in the forms of meeting notes, self-reflection entries, etc. (see examples in Liu and Chong 2024 and Chong, Liu, and Tegama 2025).

As for 'data extraction' (item M5), a data extraction form or coding sheet needs to be developed and piloted based on the research questions (M5-BP). The piloting process resembles the second-level screening process in that at least two reviewers need to be involved in extracting information from at least 10% of the included studies using the form. The purpose of this is twofold: piloting the form may reveal ambiguity in the wordings of extraction items that can lead to revision of the form; divergent understanding can be discussed and resolved between reviewers and consensus between reviewers can be demonstrated through calculating inter-coder reliability. Similar to screening, researchers are encouraged to consider reflexivity alongside reliability (M5-PP). Item M6 concerns 'study appraisal', which does not apply to exploratory reviews such as scoping reviews, bibliometric reviews, and rapid reviews. Researchers are to evaluate the methodological rigour of the included studies (M6-BP), using tools and checklists such as Critical Appraisal Skills Programme (CASP) (Critical Appraisal Skills Programme UK 2024), Mixed Methods Appraisal Tool (MMAT) (Hong, Gonzalez-Reyes, and Pluye 2018) although adaptations may be needed when using them for applied linguistics research. The editorial team of *TESOL Quarterly* developed a

detailed field-specific methodological guide for qualitative, quantitative, and mixed methods research (Mahboob et al. 2016[6]). Researchers are encouraged to be transparent and strive to remain objective when appraising the methodological rigour of studies (M6-PP). A scoring system can be developed using the tools/checklists to inform decisions about which studies to include/exclude based on methodological considerations; for example, a score of '2' can be given to studies that fully meet a criterion, a score of '1' for studies that partially meet a criterion, and a score of '0' for those that do not meet the criterion. Then, the research team can determine a baseline total score that a study needs to obtain to qualify it to be included in the RS. Having said that there remain reliability and validity issues if the study appraisal tools are used in such a way. At the same time, there may be legitimate reasons to adopt a more inclusive approach to including studies (e.g., in a niche research area, of studies that employ innovative methods).

The final item pertains to 'data synthesis' (M7), which is about the analysis of the extracted data (M7-BP). The extracted information can be analysed qualitatively and/or quantitatively based on the type of RS and research questions. While the use of software or automation tools is not required (see a list of digital tools available for conducting RS in Bond et al. 2024a, 2024b), it is important to mention any digital tool(s) used. Experience or training in using the digital tool(s) by the team should be reported to establish credibility and justifications for using the tool(s) need to be provided. When analysing data qualitatively, previous descriptions about inter-coder reliability and preferably reflexivity need to be adhered to. Additionally, the qualitative analytical approach that informs the analysis needs to be reported (e.g., thematic analysis, grounded theory). For quantitative analysis of extracted information, relevant statistical tests need to be employed (e.g., descriptive statistics, correlation analysis, regression analysis, network analysis, hypothesis testing, keyword analysis, and co-occurrence analysis) and justified. Responding to the 'accessibility' dimension of RS quality (Table 5), researchers are encouraged to share synthesised findings with non-academic stakeholders such as practitioners to obtain feedback on the usefulness of findings and ways that they can be effectively represented (M7-PP).

"Reporting" Items in SMART

The final section of SMART is about ways to structure RS and reporting synthesised findings. It includes 11 items covering all sections of an RS report including title, abstract, introduction, literature review, methodology, findings, discussion, conclusion, acknowledgement, references, and appendices. Details of each item can be found in Table 4.

Using STAR and SMART

STAR will be an RS quality -assurance framework, with a star (☆) awarded to the RS that satisfies each dimension, with a total of four stars (☆☆☆☆) awarded to an RS (Table 5). Authors are encouraged to use this for self-assessment purposes and indicate the number of stars their review has obtained in the method section. The intention is to help researchers and practitioners who want to use the evidence in an RS to approach it critically. Journal editors can also refer reviewers to STAR

when reviewing RSs. It is noted that some elements in the four dimensions of quality overlap as they are interrelated; researchers are encouraged to refer to the definitions of the quality dimensions and specific SMART items for clarity. To make STAR easy to use, I have referred to specific items in SMART (Tables 2–4). I encourage other applied linguistics journals to consider how best to use STAR to assist in the peer-review process. Moreover, STAR can be used for assessing RS conducted as (part of) a postgraduate thesis, although adjustments are needed to make it more suitable for a supervised piece of work (e.g., the number of required databases can be reduced, inter-coder reliability may not be compulsory).

Conclusion

The development of SMART is a timely response to the growing interest in RS in applied linguistics. I am aware SMART will not be a panacea for all methodological challenges in conducting high-quality RS in applied linguistics; in fact, I envisage it will prompt further questions, debates, and refinements of the tool. My humble aspiration for SMART is that it can promote wider and more fruitful methodological discussions about RS in applied linguistics, and become a one-stop shop for those looking for answers related to RS methodologies. As I mentioned, SMART is not only for appraisal of RS; it can be used to develop researchers' capacity to conduct high-quality RS including for training purposes for graduate students. I am also cognisant of how rapid the development of RS methodologies is, especially with new technologies such as Generative AI. It is crucial that SMART is constantly reviewed and revised to better reflect the latest best practices of RS both within and outside of applied linguistics. I welcome adaptations of SMART to serve different purposes such as for teaching and assessing RS at postgraduate levels and for use as part of a journal's submission guidelines or checklist for peer reviewers. I will also be creating supplementary resources to help researchers better understand SMART.

Like this new journal, *Research Synthesis in Applied Linguistics*, SMART is part of a long-term and collaborative project that needs your participation, and I hope I have offered something useful to get our conversations started.

Notes

1. This number is based on the entries in the two RS bibliographies, after duplicates have been removed.
2. https://forms.gle/Y5au1AqnAUnRbMNG6.
3. https://docs.google.com/spreadsheets/d/1VfMjcHiT-mNPGh0s7ts9phXE1sBkJ4yA5bCQ5vxvflc/edit?usp=sharing.
4. This needs to be contextualised in term of the total number of titles and abstracts to be screened. For smaller-scale RSs, researchers are recommended to consider dual-screening more than 10% if not all of them.
5. The use of Kappa is not without limitation, so researchers should exercise judgement to employ it sensibly or alternative measurements.
6. Note that Mahboob et al. (2016) was not designed to be used as a study appraisal tool. Also, there are other appropriate tools for quality appraisal. See for example Gorard's (2014) sieve (https://durham-repository.worktribe.com/output/1432280) for experiments and QEDs, Cochrane RoB2 for RCTs (https://methods.cochrane.org/bias/resources/rob-2-revised-

cochrane-risk-bias-tool-randomized-trials), Cochrane ROBINS-I for non-randomised comparisons (https://methods.cochrane.org/robins-i), Tracy's (2010) Eight Big Tent criteria for qualitative research (10.1177/1077800410383121) and the suite of tools for a variety of different designs are available at https://jbi.global/critical-appraisal-tools.

Acknowledgement

The development of SMART benefits from the advice by the following applied linguists and/or RS specialists: Mohammad Amini Farsani, Alexander Black, Melissa Bond, Serge Bibauw, Hamish Chalmers, Peter Crosthwaite, Mara Fuertes Gutiérrez, Xuesong Gao, Zoe Handley, Tim Hampson, Phil Hiver, Xuechun Huang, Talia Isaacs, Shaofeng Li, Chang Liu, Dongxia Nie, Reza Norouzian, Luke Plonsky, Hayo Reinders, Mehdi Riazi, Masatoshi Sato, Ekaterina Sudina, Jing Yun, and Xuan Zhao. This working group and some members of the RSAL editorial board provided initial input on a Microsoft Form and offered comments on a draft of the tool and this paper. I would also like to thank Sathena Chan for creating the checklist version of SMART (Appendix A).

Disclosure Statement

No potential conflict of interest was reported by the author(s).

Data Availability Statement

Data sharing is not applicable to this article as no new data were created or analysed in this article.

ORCID

Sin Wang Chong http://orcid.org/0000-0002-4519-0544

References

Amini Farsani, M., H. R. Jamali, M. Beikmohammadi, B. D. Ghorbani, and L. Soleimani. 2021. "Methodological Orientations, Academic Citations, and Scientific Collaboration in Applied Linguistics: What do Research Synthesis and Bibliometrics Indicate?" *System* 100: 1–17. https://doi.org/10.1016/j.system.2021.102547.

Atkins, S., S. Lewin, H. Smith, M. Engel, A. Fretheim, and J. Volmink. 2008. "Conducting a Meta-Ethnography of Qualitative Literature: Lessons Learnt." *BMC Medical Research Methodology* 8 (1): 1–10. https://doi.org/10.1186/1471-2288-8-21.

Belur, J., L. Tompson, A. Thornton, and M. Simon. 2021. "Interrater Reliability in Systematic Review Methodology: Exploring Variation in Coder Decision-Making." *Sociological Methods & Research* 50 (2): 837–865. https://doi.org/10.1177/0049124118799372.

Bond, M., H. Khosravi, N. Bergdahl, K. Buntins, M. De Laat, E. Oxley, V. Negrea, S. W. Chong, and M. Händel. 2024b. "Digital Evidence Synthesis Tools in Educational Technology Research: A Systematic Mapping Review." *Pre-print,* 1–26. https://doi.org/10.13140/RG.2.2.30594.25288.

Bond, M., H. Khosravi, M. De Laat, N. Bergdahl, V. Negrea, E. Oxley, P. Pham, S. W. Chong, and G. Siemens. 2024a. "A Meta Systematic Review of Artificial Intelligence in Higher Education: A Call for Increased Ethics, Collaboration, and Rigour." *International Journal of Educational Technology in Higher Education* 21 (1): 4, 1-41. https://doi.org/10.1186/s41239-023-00436-z.

Brunton, G., S. Oliver, and J. Thomas. 2020. "Innovations in Framework Synthesis as a Systematic Review Method." *Research Synthesis Methods* 11 (3): 316–330. https://doi.org/10.1002/jrsm.1399.

Centre for Reviews and Dissemination (UK). 1995. *Database of Abstracts of Reviews of Effects (DARE): Quality-Assessed Reviews*. https://www.ncbi.nlm.nih.gov/books/NBK285222/.

Chalmers, H., J. Brown, and A. Koryakina. 2024. "Topics, Publication Patterns, and Reporting Quality in Systematic Reviews in Language Education. Lessons from the International Database of Education Systematic Reviews (IDESR)." *Applied Linguistics Review* 15 (4): 1645–1669. https://doi.org/10.1515/applirev-2022-0190.

Chaudron, C. 2012. *Second Language Classrooms*. Oxford: Oxford University Press.

Chong, S. W., N. Bergdahl, M. Bond, S. Miller, and A. W. Y. Wong. 2024b. "Re-imagining 'Openness' in *Review of Education*: Methodological Standards, Open Science, and Nurturing the Next Generation of Researchers." *Review of Education* 12 (1): Article e3468. https://doi.org/10.1002/rev3.3468.

Chong, S. W., M. Bond, and H. Chalmers. 2024a. "Opening the Methodological Black Box of Research Synthesis in Language Education: Where Are We Now and Where Are We Heading?" *Applied Linguistics Review* 15 (4): 1557–1568. https://doi.org/10.1515/applirev-2022-0193.

Chong, S. W., T. J. Lin, and Y. Chen. 2022. "A Methodological Review of Systematic Literature Reviews in Higher Education: Heterogeneity and Homogeneity." *Educational Research Review* 35: 100426. https://doi.org/10.1016/j.edurev.2021.100426.

Chong, S. W., Q. Liu, and N. Tegama. 2025. "Re-imagining Quality in Conducting Systematic Literature Reviews in the Context of Dissertation Supervision: From Reliability to Reflexivity." *Research Ethics in Educational Research*.

Chong, S. W., D. Nie, and Q. Liu. Forthcoming. *A Meta-Review on Applied Linguistics Researchscape*. British Association for Applied Linguistics.

Chong, S. W., and L. Plonsky. 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., and L. Plonsky. 2024a. "Research Synthesis in Language Education." In *International Encyclopedia of Language and Linguistics*, edited by H. Nesi, P. Milin, and S. W. Chong, 3rd ed, 1–5. Amsterdam: Elsevier. https://doi.org/10.1016/B978-0-323-95504-1.00010-7.

Chong, S. W., and L. Plonsky. 2024b. "A Typology of Secondary Research in Applied Linguistics." *Applied Linguistics Review* 15 (4): 1569–1594. https://doi.org/10.1515/applirev-2022-0189.

Chong, S. W., and H. Reinders. 2020. "Technology-Mediated Task-Based Language Teaching: A Qualitative Research Synthesis." *Language Learning & Technology* 24 (3): 70–86. http://hdl.handle.net/10125/44739.

Chong, S. W., and H. Reinders. 2021. "A Methodological Review of Qualitative Research Syntheses in CALL: The State-of-the-Art." *System* 103: 102646. https://doi.org/10.1016/j.system.2021.102646.

Chong, S. W., and H. Reinders. 2022. "Autonomy of English Language Learners: A Scoping Review of Research and Practice." *Language Teaching Research*, 1–26. Advanced online publication. https://doi.org/10.1177/13621688221075812.

Çiftçi, E. Y., and P. Savaş. 2018. "The Role of Telecollaboration in Language and Intercultural Learning: A Synthesis of Studies Published Between 2010 and 2015." *ReCALL* 30 (3): 278–298. https://doi.org/10.1017/S0958344017000313.

Critical Appraisal Skills Programme UK. 2024. *CASP Checklists*. https://casp-uk.net/casp-tools-checklists/.

Elliott, J. H., A. Synnot, T. Turner, M. Simmonds, E. A. Akl, S. McDonald, G. Salanti, et al. 2017. "Living systematic review: 1. Introduction—the why, what, when, and how." *Journal of Clinical Epidemiology* 91: 23–30. http://dx.doi.org/10.1016/j.jclinepi.2017.08.010.

Ellis, R. 2010. "Epilogue: A Framework for Investigating Oral and Written Corrective Feedback." *Studies in Second Language Acquisition* 32(2): 335–349. https://doi.org/10.1017/S0272263109990544.

Ellis, R. 2015. "Introduction: Complementarity in Research Syntheses." *Applied Linguistics* 36 (3): 285–289. https://doi.org/10.1093/applin/amv015.

Garritty, C., C. Hamel, M. Trivella, G. Gartlehner, B. Nussbaumer-Streit, D. Devane, C. Kamel, et al. 2024. "Updated recommendations for the Cochrane rapid review methods guidance for

rapid reviews of effectiveness." *BMJ (Clinical research ed.),* 384: e076335. https://doi.org/10.1136/bmj-2023-076335.

Ghanbar, H., C. Cinaglia, R. A. Randez, and P. I. De Costa. 2024. "A Methodological Synthesis of Narrative Inquiry Research in Applied Linguistics: What's the Story?" *International Journal of Applied Linguistics* 34 (4): 1629–1655. Advanced online publication. https://doi.org/10.1111/ijal.12591.

Gorard, S. 2014. *A proposal for judging the trustworthiness of research findings. Radical statistics* 110: 47–59.

Hong, Q. N., A. Gonzalez-Reyes, and P. Pluye. 2018. "Improving the Usefulness of a Tool for Appraising the Quality of Qualitative, Quantitative and Mixed Methods Studies, the Mixed Methods Appraisal Tool (MMAT)." *Journal of Evaluation in Clinical Practice* 24 (3): 459–467. https://doi.org/10.1111/jep.12884.

Isaacs, T., and H. Chalmers. 2023. "Reducing 'Avoidable Research Waste' in Applied Linguistics Research: Insights from Healthcare Research." *Language Teaching*, 1–18. Advanced online publication. https://doi.org/10.1017/S0261444823000411.

The Joanna Briggs Institute. 2014. *Joanna Briggs Institute Reviewers' Manual, 2014 Edition.* http://joannabriggs.org/assets/docs/sumari/ReviewersManual-The-Systematic-Review-of-Economic-Evaluation-Evidence-2014_v2.pdf.

Kong, M. S., S. Hopkyns, and S. W. Chong. 2024. *A Protocol for a Practice Review of Translanguaging in UK Classrooms*. International Database of Education Systematic Reviews. https://idesr.org/article/IDESR000158.

Lange, R. T. 2011. "Inter-Rater Reliability." In *Encyclopedia of Clinical Neuropsychology*, edited by J. S. Kreutzer, J. DeLuca, and B. Caplan, 1348. Springer. https://doi.org/10.1007/978-0-387-79948-3_1203.

Li, S. 2015. "The Associations Between Language Aptitude and Second Language Grammar Acquisition: A Meta-Analytic Review of Five Decades of Research." *Applied Linguistics* 36(3): 385–408. https://doi.org/10.1093/applin/amu054.

Li, S., and H. Wang. 2018. "Traditional Literature Review and Research Synthesis." In *Palgrave Handbook of Applied Linguistics Research Methodology*, edited by A. Phakiti, P. De Costa, L. Plonsky, and S. Starfield, 123–144. New York: Palgrave.

Liu, Q., and S. Chong. 2024. "Bilingual Education in China: A Qualitative Synthesis of Research on Models and Perceptions." *Applied Linguistics Review* 15 (4): 1671–1693. https://doi.org/10.1515/applirev-2022-0194.

Macaro, E. 2019. "Systematic Reviews in Applied Linguistics." In *The Routledge Handbook of Research Methods in Applied Linguistics*, edited by J. McKinley and H. Rose, 230–239. Routledge.

Macaro, E., S. Curle, J. Pun, J. An, and J. Dearden. 2018. "A Systematic Review of English Medium Instruction in Higher Education." *Language Teaching* 51 (1): 36–76. https://doi.org/10.1017/S0261444817000350.

Macfarlane, B. 2012. "The Higher Education Research Archipelago." *Higher Education Research & Development* 31 (1): 129–131. https://doi.org/10.1080/07294360.2012.642846.

Mahboob, A., B. Paltridge, A. Phakiti, E. Wagner, S. Starfield, A. Burns, R. H. Jones, and P. I. De Costa. 2016. "TESOL Quarterly Research Guidelines." *TESOL Quarterly* 50 (1): 42–65. https://doi.org/10.1002/tesq.288.

Masgoret, A.-M., and R. C. Gardner. 2003. "Attitudes, Motivation, and Second Language Learning: A Meta-Analysis of Studies Conducted by Gardner and Associates." *Language Learning* 53 (1): 123–163. https://doi.org/10.1111/1467-9922.00212.

Moher, D., L. Shamseer, M. Clarke, D. Ghersi, A. Liberati, M. Petticrew, P. Shekelle, L. A. Stewart, and PRISMA-P Group. 2015. "Preferred Reporting Items for Systematic Review and Meta-Analysis Protocols (PRISMA-P) 2015 Statement." *Systematic Reviews* 4 (1): 1–9. https://doi.org/10.1186/2046-4053-4-1.

Morea, N., and H. Ghanbar. 2024. "Q Methodology in Applied Linguistics: A Systematic Research Synthesis." *System* 120: 103194. https://doi.org/10.1016/j.system.2023.103194.

Norouzian, R., and G. Bui. 2024. "Meta-analysis of Second Language Research with Complex Research Designs." *Studies in Second Language Acquisition* 46 (1): 251–276. https://doi.org/10.1017/S0272263123000311.

Norris, J. M., and L. Ortega. 2006. *Synthesizing Research on Language Learning and Teaching*. Amsterdam: John Benjamins.

Norris, J. M., and L. Ortega. 2007. "The Future of Research Synthesis in Applied Linguistics: Beyond Art or Science." *TESOL Quarterly* 41 (4): 805–815. https://doi.org/10.1002/j.1545-7249.2007.tb00105.x.

O'Connor, C., and H. Joffe. 2020. "Intercoder Reliability in Qualitative Research: Debates and Practical Guidelines." *International Journal of Qualitative Methods* 19: 1–13. https://doi.org/10.1177/1609406919899220.

Ortega, L. 2003. "Syntactic Complexity Measures and Their Relationship to L2 Proficiency: A Research Synthesis of College-Level L2 Writing." *Applied Linguistics* 24 (4): 492–518. https://doi.org/10.1093/applin/24.4.492.

Page, M. J., J. E. McKenzie, P. M. Bossuyt, I. Boutron, T. C. Hoffmann, C. D. Mulrow, et al. 2021. "The PRISMA 2020 Statement: An Updated Guideline for Reporting Systematic Reviews." *BMJ* 372: n71. https://doi.org/10.1136/bmj.n71.

Pham, M. T., A. Rajić, J. D. Greig, J. M. Sargeant, A. Papadopoulos, and S. A. McEwen. 2014. "A Scoping Review of Scoping Reviews: Advancing the Approach and Enhancing the Consistency." *Research Synthesis Methods* 5 (4): 371–385. https://doi.org/10.1002/jrsm.1123.

Plonsky, L. 2023. "Introduction to the Special Issue on Introducing Bibliometrics in Applied Linguistics." *Studies in Second Language Learning and Teaching* 13 (4): 721–726. https://doi.org/10.14746/ssllt.40213.

Plonsky, L. 2024. "Study Quality as an Intellectual and Ethical Imperative: A Proposed Framework." *Annual Review of Applied Linguistics*, 1–15. https://doi.org/10.1017/S0267190524000059.

Plonsky, L., Y. Hu, E. Sudina, and F. L. Oswald. 2023. "Advancing Meta-Analytic Methods in L2 Research." In *Current Approaches in Second Language Acquisition Research: A Practical Guide, I*, edited by A. Mackey and S. M. Gass. Wiley.

Plonsky, L., and F. L. Oswald. 2012a. "How to Do a Meta-Analysis." In *Research Methods in Second Language Acquisition: A Practical Guide*, edited by A. Mackey and S. M. Gass, 275–295. Oxford: Basil Blackwell.

Plonsky, L., and F. L. Oswald. 2012b. "Meta-Analysis." In *The Routledge Encyclopedia of Second Language Acquisition*, edited by P. Robinson, 420–423. London: Routledge.

Plonsky, L., and N. Ziegler. 2016. "The CALL-SLA Interface: Insights from a Second-Order Synthesis." *Language Learning & Technology* 20 (2): 17–37. http://hdl.handle.net/10125/44459.

Rader, T., M. Mann, C. Stansfield, C. Cooper, and M. Sampson. 2014. "Methods for Documenting Systematic Review Searches: A Discussion of Common Issues." *Research Synthesis Methods* 5 (2): 98–115. https://doi.org/10.1002/jrsm.1097.

Ren, W., S. Li, and X. Lu. 2023. "A Meta-Analysis of the Effectiveness of Second Language Pragmatics Instruction." *Applied Linguistics*, 1–21. First view.

Riazi, A. M., and M. Amini Farsani. 2024. "Mixed-methods Research in Applied Linguistics: Charting the Progress Through the Second Decade of the Twenty-First Century." *Language Teaching* 57 (2): 143–182. https://doi.org/10.1017/S0261444823000332.

Rose, H., J. G. Briggs, J. A. Boggs, L. Sergio, and N. Ivanova-Slavianskaia. 2018. "A Systematic Review of Language Learner Strategy Research in the Face of Self-Regulation." *System* 72: 151–163. https://doi.org/10.1016/j.system.2017.12.002.

Sato, M., S. W. Chong, T. Aktar, J. Cowell, M. S. Kong, and M. Shaahdadi. 2024. "Creating and Sustaining a Platform for Researchers and Teachers to Communicate: An Example of *TESOLgraphics*." *Innovation in Language Learning and Teaching*, 1–11. Advanced online publication. https://doi.org/10.1080/17501229.2024.2404613.

Shea, B. J., B. C. Reeves, G. Wells, M. Thuku, C. Hamel, J. Moran, D. Moher, et al. 2017. "AMSTAR 2: A Critical Appraisal Tool for Systematic Reviews That Include Randomised or non-Randomised Studies of Healthcare Interventions, or Both." *BMJ* 358: 1–9. https://doi.org/10.1136/bmj.j4008.

Shen, R., and S. W. Chong. 2023. "Learner Engagement with Written Corrective Feedback in ESL and EFL Contexts: A Qualitative Research Synthesis Using a Perception-Based Framework." *Assessment & Evaluation in Higher Education* 48 (3): 276–290. https://doi.org/10.1080/02602938.2022.2072468.

Shintani, N. 2015. "The Effectiveness of Processing Instruction and Production-Based Instruction on L2 Grammar Acquisition: A Meta-Analysis." *Applied Linguistics* 36 (3): 306–325. https://doi.org/10.1093/applin/amu067.

Sutton, A., M. Clowes, L. Preston, and A. Booth. 2019. "Meeting the review family: exploring review types and associated information retrieval requirements." *Health Information & Libraries Journal* 36 (3): 202–222. https://doi.org/10.1111/hir.12276.

Teimouri, Y., J. Goetze, and L. Plonsky. 2019. "Second Language Anxiety and Achievement: A Meta-Analysis." *Studies in Second Language Acquisition* 41 (2): 363–387. https://doi.org/10.1017/S0272263118000311.

Tracy, S. J. 2010. "Qualitative Quality: Eight "Big-Tent" Criteria for Excellent Qualitative Research." *Qualitative Inquiry* 16 (10): 837–851. https://doi.org/10.1177/1077800410383121.

Vuogan, A., and S. Li. 2024. "A Systematic Review of Meta-Analyses in Second Language Research: Current Practices, Issues, and Recommendations." *Applied Linguistics Review* 15 (4): 1621–1644. https://doi.org/10.1515/applirev-2022-0192.

Willig, A. C. 1985. "A Meta-Analysis of Selected Studies on the Effectiveness of Bilingual Education." *Review of Educational Research* 55 (3): 269–317. https://doi.org/10.3102/00346543055003269.

Zakaria, A., and V. Aryadoust. 2024. "A Scientometric Analysis of Applied Linguistics Research (1970–2022): Methodology and Future Directions." *Applied Linguistics Review* 15 (4): 1595–1619. https://doi.org/10.1515/applirev-2022-0210.

Appendix

Appendix A: SMART checklist

Preparation		
P1	Originality	Have you conducted a preliminary search on published research syntheses (RSs)?
		Have you reflected on original contributions and/or specific contexts of the current RS?
P2	Research team	Have you assembled the research team and discussed roles/specific contributions of team members?
		Have you considered the involvement of non-academic stakeholders?
P3	Type	Have you decided on the type of RS?
		Have you referred to existing typologies (e.g., Chong & Plonsky, 2024)?
P4	Registration	Have you completed title registration?
		Have you published an open-access and/or peer-reviewed protocol?
Method		
M1	Research questions	Have you included the research questions?
		Have you reviewed the alignment between research questions and RS type?
M2	Incl. and excl. criteria	Have you included the specific criteria to inform study selection?
		Have you included the rationale of the criteria?
M3	Search strategies	Have you included all the databases (min. of 3) and the search strings used?
		Have you included the rationale for choosing the databases and search strings?
M4	Screening	Have you conducted first-level screening on titles and abstracts?
		Have you conducted second-level screening on full-texts?
		Have you calculated inter-coder reliability?
		Have you documented the processes and any issues that arose and solutions?
M5	Data extraction	Have you developed and piltoed the data extraction form?
		Have you calculated inter-coder reliability?
		Have you documented the processes and any issues that arose and solutions?
M6	Appraisal	Have you evaluated the methodological rigour of included studies using existing tools/checklists (if applicable)?
		Have you documented the processes and any issues that arose and solutions?
M7	Data synthesis	Have you identified an approach or method for analysising and interpretating the extracted data?

		Have you used any software or automation tools (if any)?
		Have you calculated inter-coder reliability?
		Have you documented the processes and any issues arose and solutions?
Reporting		
R1	Title	Have you included a title and a subtitle? (the title is the focus of the RS and the subtitle is the type of RS)
		Have you checked the RS type is consistent with the RS typologies in applied linguistics (Chong & Plonsky, 2024; Chong, 2025)?
R2	Abstract	Have you included an abstract?
		Have you included an open-access complimentary graphical abstract or lay summary?
R3	Introduction	Have you included an introduction?
		Have you presented the unique contributions and/or specific contexts of the current RS?
R4	Literature review	Have you included a literature review with a critical engagement?
		Have you reviewed other relevant published RSs?
R5	Methodology	Have you described the methodology with all essential elements (see M1-7)?
		Have you reported any deviations from the protocol?
		Have you reported any limitations e.g., selection bias?
		Have you described the team's credential to show members' expertise and roles?
		Have you discussed contextual sensitivity?
R6	Findings	Have you included a thematic presentation of findings with illustrated examples?
		Have you provided a summary of individual study characteristics in a table?
		Have you included appropriate visual aids?
R7	Discussion	Have you included a discussion of the meaning and importance of the findings?
		Have you discussed how current findings relate to previous primary studies and RSs?
		Have you mentioned the relevance of findings to non-academic audience?
R8	Conclusion	Have you summarised headlines of the synthesised findings?
		Have you identified future research directions?
		Have you discussed implications for future RSs?
R9	Acknowledgements	Have you acknowledged support from individuals and/or organisations?
		Have you used a CRediT statement to specify contributions of each author?
R10	References	Have you cited sources in APA 7th format?
		Have you indicated publications included in the RS with an asterisk (*)?
R11	Appendices	Have you included a copy of the tools/instruments used and/or data?
		Have you considered making the research materials publicly available?