Language Teaching (2024), **57**, 143–182
doi:10.1017/S0261444823000332

CAMBRIDGE UNIVERSITY PRESS

REVIEW OF RECENT SCHOLARSHIP

Mixed-methods research in applied linguistics: Charting the progress through the second decade of the twenty-first century

A. Mehdi Riazi[1] and Mohammad Amini Farsani[2]

[1]Hamad Bin Khalifa University, Doha, Qatar and [2]Iran University of Science and Technology, Tehran, Iran
Corresponding author: A. Mehdi Riazi; Email: ariazi@hbku.edu.qa

(Received 4 July 2023; accepted 20 July 2023)

Abstract

This review of recent scholarship (RRS) paper is a follow-up of the first, published in this journal in 2014. For this RRS paper, we identified and included 304 mixed-methods research (MMR) papers published in 20 top-tier applied linguistics (AL) journals. We used a six-pronged quality and transparency framework to review and analyze the MMR studies, drawing on six quality frameworks and transparency discussions in the MMR literature. Using the quality and transparency framework, we report on: (1) which sources AL MMR researchers use to frame their studies, (2) how explicitly they explain the purpose and design structure of the MMR studies, (3) how transparently they describe method features (sampling procedures, data sources, and data analysis), and (4) how they integrate quantitative and qualitative data and analyses and construct meta-inferences. The results of the analyses will be reported and will show how MMR has developed and is represented in the published articles in the second decade of the twenty-first century. The discussion of the results will also highlight the areas future AL MMR researchers need to consider to make their studies and reports more rigorous and transparent.

1. Introduction

Mixed-methods research (MMR) is a methodological approach that intends to benefit from the affordances of quantitative and qualitative methodologies in a single study. Both quantitative and qualitative data and analysis are mixed in principled ways (a) to address the adversarial incompatibility of quantitative and qualitative approaches and (b) to help researchers produce more comprehensive inferences (Riazi & Candlin, 2014). In Riazi and Candlin's (2014) state-of-the-art review, the authors reviewed 40 MMR studies published in 30 AL journals over the first decade of the twenty-first century. Since the publication of that paper in *Language Teaching*, several important events regarding MMR have taken place. The first has been a more substantive discussion of MMR more broadly and in applied linguistics (AL) more specifically. The discussions are presented in journals, books, edited volumes, conferences, and dissertations. This wealth of updated literature on MMR is valuable since it can inform different disciplines on how best to design and implement MMR studies. The second event relates to the dissemination of MMR studies in the AL journals from 2010–2020. Amini Farsani et al. (2021) report a second-highest frequency for MMR studies (25.9%) after quantitative studies (42.6%) in a review of 3,814 empirical articles published in 18 leading AL journals from 2009–2018. We also identified 304 published articles in 20 top-tier AL journals that have explicitly stated the use of MMR. This trend is in sharp contrast to the number of MMR studies reported in previous MMR reviews.

It is, therefore, timely to update the original review and demonstrate how the use of MMR has developed in AL studies in the period reviewed. To guide our inspection, we posed the following

research questions (RQ) derived from the quality and transparency framework we will discuss later in the article:

RQ1: What are the key areas AL researchers have used MMR to investigate language-related issues?
RQ2: Which MMR references do AL researchers use to document and frame their MMR methodology?
RQ3: How transparently do AL MMR researchers describe and explain their study's purpose and design structure?
RQ4: How transparently do AL MMR researchers describe and explain the general issues or method features (sampling, data sources, and data analysis) of the MMR studies?
RQ5: How do AL researchers integrate quantitative and qualitative data and analyses in their studies and make final conclusions or produce meta-inferences?

Our corpus included 304 MMR studies published in 20 top-tier AL journals over the second decade of the twenty-first century (2011–2020 inclusive). We believe the review is timely, given the increasing trend in using MMR in AL and the need for rigorous reviews and analyses to inform different stakeholders. Our review goes beyond most MMR review papers that follow a prevalence orientation. That is, they have exclusively focused on and sought empirical evidence for the prevalence of MMR in different disciplines, including AL. While those reviews have provided valuable information and knowledge to various stakeholders, they have rarely attended to the quality and transparency of MMR studies. In the inter- and multidisciplinary field of AL, it is time to explore how MMR studies are designed, implemented, and reported. While we acknowledge the contribution of previous studies on the prevalence rates of methodological approaches in general and mixed methods research in particular, the current study is significantly different from earlier studies in focusing on the quality and transparency criteria, as will be discussed later. The outcomes of this and similar reviews can respond to and contribute to what Plonsky (2017) named methodological awareness or what Byrnes (2013, p. 825) has called the rise of "methodological turn". Byrnes referred to methodological turn as professional scrutiny of the methodologies used by researchers so that the discussion can shed light on the relationship between theory, practice, and knowledge.

2. Background to the study

Since the review paper in *Language Teaching* in 2014, MMR has received increasing attention from scholars and researchers in AL. This growing trend corresponds to the broader context of MMR adoption and discussion in other disciplines. Two preliminary points are, however, in order. As Johnson (2006, p. v) stated, "it is important to recognize that the purpose of mixed research is not to replace qualitative or quantitative research". There is no intention to convey chronological supremacy by assuming that MMR is the trend of the time, surpassing quantitative and qualitative approaches to research.

On the contrary, quantitative, qualitative, and MMR are potential research approaches that can help AL researchers investigate a variety of research problems and research questions. There are many topics and research questions that lend themselves to quantitative research. There are also situations where qualitative research will be a better methodological choice than quantitative or MMR. The fact is that quantitative research, qualitative research, MMR, and even secondary research (research synthesis) all have much to offer to the AL community, depending on the conceptualization of research problems and research questions.

The second point regards the rigor of mixing quantitative and qualitative data and analysis. As Hashemi and Babaii (2013) and Riazi and Candlin (2014) indicated through analysis of published journal articles, a considerable number of seemingly MMR studies have not been able to do so rigorously. These studies may thus be named eclectic rather than principled or innovative MMR studies (Riazi, 2016). As such, there is more opportunity for using MMR in more principled and creative ways compared with the eclectic use of quantitative and qualitative data and analysis. Eclectic

https://doi.org/10.1017/S0261444823000332 Published online by Cambridge University Press

MMR studies usually fall short of using MMR jargon and procedures to frame their studies. They are also less innovative in the sense of integrating the two methods creatively. As such, reviews like the current one can help future researchers develop more awareness about the affordances of MMR and thus use the methodology in more principled and innovative ways.

Since 2014, the discussion of MMR has expanded parallel to its using in AL. There have been book chapters (Hashemi, 2020; Ivankova & Greer, 2015; Mackey & Bryfonski, 2018), books (Brown, 2014; Riazi, 2017), and research commentaries (Mirhosseini, 2018; Riazi, 2018). In addition, a separate section is now devoted to MMR in the second edition of *The encyclopedia of applied linguistics* (Chapelle, forthcoming). Moreover, methodological reviews have been the direct focus or part of the objectives of some studies published after the first MMR review paper in *Language Teaching*. Table 1 provides a list of the recently published empirical and review studies that address MMR directly or indirectly. That is, the source might directly discuss MMR, or MMR might have been discussed indirectly as part of the study's focus. The criterion for choosing empirical studies has been their discussion of MMR and not necessarily using MMR to investigate a problem.

Table 1. List of sources discussing methodological orientations in AL since the first state-of-the-art review paper in LT

Empirical & review studies	Domain	Subdomains	Focus/ scope	Orientation
Riazi (2016)	Discussion of MMR in AL	Typology of MMR	Direct	Towards innovativeness in MMR
Tsushima (2015)	Discussion of MMR in AL	The use of MMR in language testing and assessment	Direct	MMR implications for language testing and assessment
Riazi et al. (2018)	Review of the papers published in the *JSLW*	L2 writing strand	Indirect	A contextual, theoretical, and methodological review of the papers in L2 writing
Hashemi and Gohari Moghaddam (2019)	Rhetorical use of MMR	Writing up discussion sections	Direct	Rhetorical implications for writing MMR discussion sections
Riazi et al. (2020)	Review of the papers published in the *JEAP*	English for academic purposes	Indirect	A contextual, theoretical, and methodological review of the papers in EAP
Tazik et al. (2020)	Review of the papers published in AL	AL	Indirect	A methodological review of the papers in AL
Gilmore and Ganem-Gutierrez (2020)	Research of L2 writing using MMR	Complex Dynamic System	Direct	The benefits and challenges of employing mixed methods approaches from a CDST perspective
Amini Farsani et al. (2021)	Review of the papers published in AL journals	AL	Indirect	A methodological and bibliometric review of the papers in AL
Riazi et al. (in press)	Review of the papers published in the *Tesol Quarterly* journal	TESOL and AL	Indirect	A contextual, theoretical, and methodological review of the papers in Tesol and AL
Amini Farsani and Jamali (in press)	A bibliometric review of the papers published in AL journals	AL	Indirect	Collaborative research network represented in AL
Amini Farsani and Mohammadi (2022)	Review of MA theses	The quality appraisal of MMR theses	Indirect	MMR implications for MA students in TEFL research
Amini Farsani et al. (2022)	The pedagogical discussion of MMR in AL	MMR proficiency	Direct	MMR pedagogical implications for AL

https://doi.org/10.1017/S0261444823000332 Published online by Cambridge University Press

Merging the pluralistic nature of MMR with the interdisciplinary function of AL, Riazi (2016) argued that it is possible to design MMR studies that can be innovative. The point was raised that AL researchers might consider MMR as a coherent methodology rather than merely an option for combining quantitative and qualitative methods. To that end, he elaborated on the typology of MMR studies as "eclectic", "principled", and "innovative" initially presented in the 2014 review paper. The eclectic MMR applies to those studies that use quantitative and qualitative data and analyses without necessarily integrating them (see Riazi & Candlin, 2014, for further discussion on eclectic, principled, and innovative MMR). On the other hand, principled MMR studies use relevant MMR literature to frame their study with an explanation of how the two data sources and analyses are integrated. This is usually done by identifying a purpose (e.g., triangulation, expansion, development, etc.) and explicitly stating the design (e.g., concurrent convergent, sequential explanatory, etc.) of the study. Innovative MMR applies to those studies in which researchers go beyond routine MMR purposes and designs (principled MMR) to conceptualize the research problem as multi-dimensional, with a need to address each dimension with relevant methods and ultimately integrating them to provide a coherent and comprehensive understanding of the research problem.

In addition to review papers, some interesting bibliometric studies have also been conducted that address MMR indirectly. For example, Lei and Liu (2019) and Zhang (2020) examined the research trends and topical foci of empirical studies published in AL journals. Taking a bibliometric approach, Lei and Liu (2019) examined 42 journals from 2006–2015, tracing research trends (i.e., the most frequent keywords, the most highly cited publications, and authors, along with the most productive countries). In addition to substantive keywords (e.g., identity, technological issues), the findings highlighted the prevalence of methodological-based keywords such as mixed methods research, meta-analysis, and structural equation modeling. On the other hand, Zhang (2020) conducted a bibliometric review of 16 second language acquisition (SLA)-based journals from 1997–2018. He focused on co-citation, highly cited regions, keyword analysis, and sources of publications. Considering the topical issues and other substantive topics, he found the keywords related to methodological problems (e.g., mixed methods, meta-analysis, eye tracking, and semi-instructed interviews). Both studies (Lei & Liu, 2019; Zhang, 2020) unanimously demonstrated that the representation of methodological-based topics, notably mixed methods research, was one of the most prevalent critical terms in AL articles. As Zhang (2020) put it, "while both quantitative and qualitative have remained relatively stable, the keyword mixed method has made the biggest jump among all keywords" (p. 216).

Besides the above two studies, which projected MMR from a bibliometric perspective, Amini Farsani et al. (2021) and Amini Farsani and Mohammadi (2022) also used methodological synthesis and bibliometric techniques. They synthesized the methodological orientations, academic citation, and scientific collaboration in 3,992 articles published in 18 leading AL journals from 2009–2018. In the first study (Amini Farsani et al., 2021), their findings indicated that the most frequent research approach was quantitative ($n = 1{,}700$, 42.6%), followed by MMR studies ($n = 1{,}034$, 25.9%) and qualitative studies ($n = 993$, 24.9%). The least represented research approach was systematic reviews ($n = 87$, 2.2%). As for the chronological representation of MMR, their findings showed that mixed methods research studies gained considerable attention from 2014–2018. However, they highlighted the dominant practice in MMR studies to be "eclectic" (see Riazi & Candlin, 2014; Riazi, 2016) or "quasi-mixed" (Teddlie & Tashakkori, 2006). In terms of bibliometrics, they found that systematic review papers had significantly higher citation evidence than quantitative, qualitative, and MMR. This might be because systematic reviews provide evidence from an array of empirical studies and thus attract researchers, practitioners, and policymakers' attention (see, e.g., Tahamtan et al., 2016).

In the second study, Amini Farsani and Jamali (in press) explored the collaborative network of AL studies (i.e., country, author, topic) considering different methodological orientations (i.e., quantitative, qualitative, mixed-methods, and systematic review papers). They reviewed 3,992 articles published in 18 journals of AL from 2009–2018. They found several patterns of collaboration regarding different research methodologies – collaboration as in co-authors on a published paper. The highest number of co-authors was found in quantitative studies ($n = 66.8\%$), followed by systematic reviews

https://doi.org/10.1017/S0261444823000332 Published online by Cambridge University Press

(60.9%) and mixed-methods studies (55.7%). The lowest number of co-authorships was found in qualitative studies (45.5%). Regarding researchers' countries, the USA and UK were the two top collaborative countries regarding the MMR approach. Topic-wise, mixed-methods collaborators were predominantly interested in collaboration on language testing and assessment and educational technology as the two most frequently represented strands in AL.

Apart from the above bibliometric review papers, in a set of studies, Riazi and his collaborators synthesized the substantive content and methodological aspects of studies published in some AL journals, including the *Journal of Second Language Writing (JSLW)*, *English for Academic Purposes (EAP)*, and *TESOL Quarterly (TQ)*. The journals were chosen for no particular reason but to align with the reviewers' interests. Riazi et al. (2018) examined 272 empirical studies published in the *JSLW* from 1992–2016. Besides the content areas (i.e., theoretical and topical issues), they surveyed the prevalence rate of quantitative, qualitative, and MMR overall and chronologically. Overall, the results revealed that qualitative studies ($n = 106$, 39%) were the most prevalent research approach, followed by the eclectic MMR ($n = 85$, 31.3%) and quantitative studies ($n = 76$, 27.9%). Chronologically, there were fewer eclectic MMR in the first period (1992–1999) compared with the following two periods (2000–2010 and 2011–2016), which is not beyond expectation given the recency of MMR adoption in AL. These findings are indicative that second language (L2) writing researchers have welcomed qualitative approaches more than quantitative approaches. One of the reasons could be that most of the L2 writing issues related to writing processes, writing feedback, and so forth, lend themselves more to qualitative data collection and analysis. On the other hand, L2 writing researchers have started to use quantitative and qualitative data analysis in their studies. However, most of the attempts are at the early stages of being eclectic and moving toward principled MMR designs.

Riazi et al. (2020) also systematically reviewed 416 empirical articles published in *EAP* from 2002–2019. They investigated four broad themes: contexts and participants, theoretical underpinnings and research orientations, research methodology and data sources, and pedagogical implications. Methodologically, the results revealed that overall eclectic MMR was the most frequent research approach ($n = 214$, 51%). This was followed by qualitative ($n = 146$, 35%), quantitative (n = 45, 11%), and principled and innovative mixed-methods studies ($n = 11$, 3%), respectively. Chronologically, they found almost the same trend across periods. That is, eclectic MMR was the most pervasive research approach, and principled mixed-methods studies were the least frequent research approach.

Finally, Riazi et al. (in press) reported the results of a bibliometric study through which they reviewed 696 empirical articles published in the *TQ* journal over its lifespan (1967–2019). They provided overall and periodic reviews (1967–1979, 1980–1989, 1990–1999, 2000–2009, 2010–2019) on three themes. The themes were: (1) contexts and participants, (2) research foci and theoretical orientations, and (3) research methodology and data sources. Regarding the third theme, they reported quantitative, qualitative, and eclectic methods as the most frequently used research methodologies, and elicitation, multiple sources, and observation as the top three data sources used by researchers. Quantitative methods were found to be predominant ($n = 286$, 41%) in TESOL studies, followed by qualitative methods ($n = 254$, 36.5%) and eclectic MMR ($n = 143$, 20.5%). Only 1.9% ($n = 13$) of the articles could be coded as principled MMR. These were articles in which researchers relied on the MMR literature, used relevant jargon to explain how they mixed quantitative and qualitative data and analysis, and identified a purpose for mixing methods. As discussed earlier, L2 researchers have started using quantitative and qualitative data and analyses in their studies. However, most of these studies do not provide a clear discussion of how they integrated the two data sources to reach more comprehensive conclusions. These studies are called eclectic (Riazi & Candlin, 2014). As we move forward, more and more L2 researchers are appreciating principled MMR designs. That is, they use relevant MMR sources to frame their studies.

There were also some empirical studies that addressed the role of MMR in investigating some L2 issues. For example, Gilmore and Ganem-Gutierrez (2020) considered the use of MMR in researching L2 writing from the complex dynamic system theory perspective. Drawing on an ongoing research project on Japanese learners of English writing processes, Gilmore and Ganem-Gutierrez discuss

https://doi.org/10.1017/S0261444823000332 Published online by Cambridge University Press

the benefits and challenges of using MMR to investigate L2 writing from a Complex Dynamic Systems Theory (CDST) perspective. Tsushima (2015) also examined the considerable role of MMR in the language assessment strand, notably in classroom-based language assessment (CBLA). He investigated the Japanese English as a foreign language (EFL) teachers' perspectives toward the alignment of teaching and assessment practices through the lens of MMR. More specifically, Tsushima illustrated how MMR could comprehensively depict the complicated nature of teaching and assessment in L2 classes. As such, he used a three-phase sequential explanatory design. Phase one included administering a questionnaire among teachers, phase two was observing classrooms and analyzing term examination samples, and phase 3 was conducting interviews. He highlighted the inadequacies of quantitative or qualitative research designs with detailed examples and instances, as they might not be able to represent the dynamicity of CBLA. He thus recommended a mixed-methods approach to attend to the different layers of CBLA, which might not be comprehensively examined with quantitative or qualitative research approaches alone. He also highlighted language assessment researchers' theoretical and practical challenges when implementing MMR studies.

Tazik et al. (2020) surveyed 7,525 articles published in 10 AL journals from 1986–2015. They followed a methodological prevalence-rate orientation and examined the articles both cumulatively and chronologically. The overall findings revealed that quantitative studies ($n = 3{,}612$, 48%) overtook the qualitative ($n = 2{,}184$, 29.02%) and MMR studies ($n = 467$, 6.21%), respectively. Chronologically, the same trend was also notable with the quantitative approach as frequently used from 1986–2005 ($n = 2{,}530$, 64.23%). However, qualitative studies received significant momentum in AL from 2006–2015 ($n = 1{,}243$, 41.75%). The substantial rise of MMR was also notable, from 0.55% in the first decade to 13.5% in the last decade. The authors further described the most frequently represented designs in each research paradigm. Among the 467 articles with MMR designs, convergent MMR design ($n = 252$, 54%) was more commonly used than sequential designs ($n = 215$, 46%).

All the above studies highlight that since the first state-of-the-art paper in 2014, the appeal for using MMR in AL has been expanding. However, an important question is how MMR researchers strive for quality, transparency, and rigor in designing, conducting, and reporting MMR studies. In the next part, we will review and discuss the quality and transparency discussion in the MMR literature.

2.1 Quality and transparency in MMR

As one of the hotly debated notions in academia, more broadly, quality is characterized as a fluid, slippery, and value-laden undertaking. It means different things to different stakeholders and is subjectively attached to what is worthwhile, transparent, sound, and acceptable (Harvey & Green, 1993; Van Kemenade et al., 2008). In Harvey and Green's (1993) terms, quality is not a "unitary concept"; instead, it must be characterized as a "range of qualities" or constituents (p.28). Hence, as Van Kemenade et al. (2008) put it, quality can be conceptualized and operationalized, and translated into four components: OBJECT (i.e., the quality of what?), STANDARD (i.e., metrics, benchmarks, or criteria used to gauge quality), SUBJECT (i.e., who sets the criteria or standard and for who?), and VALUES (that the object is proper, worthwhile, and acceptable). Such all-encompassing conceptualization and operationalizations are in line with what Crosby (1986) referred to as a quality culture through which everybody is responsible for creating and observing standards.

The criteria or standards for judging quality can provide an evidence-based and pragmatic solution to the complex questions of quality posed in different disciplines. Accordingly, if we want to find an agreed-upon list of criteria for assessing the quality of an object (in this case, MMR), it is axiomatic that we should understand how the quality standards or criteria are conceptualized and operationalized by subjects (MMR scholars and researchers) to determine the value (properness) of the object (MMR). In the following paragraphs, we present how the quality of MMR regarding the four constituents is conceived and operationalized in MMR literature.

From 2004 onwards, an era coinciding with the rise of the expanded procedural developmental period in general MMR research (Creswell & Plano Clark, 2018), MMR researchers and methodologists

https://doi.org/10.1017/S0261444823000332 Published online by Cambridge University Press

have widely addressed the notion of quality. This trend is evidenced by an increasing number of empirical and theoretical publications on this topic. For example, just in the *Journal of Mixed Methods Research*, five editorials (Creswell & Tashakkori, 2007; Fetters & Freshwater, 2015; Fetters & Molina-Azorin, 2019a, 2019b; Mertens, 2011) and one special issue (Fàbregues et al., 2021) addressed MMR quality.

Two systematic reviews (Heyvaert et al., 2013; Fàbregues & Molina-Azorin, 2017) were also conducted to review the published MMR literature on quality from the mid-2000s. First, in response to the lack of quality benchmarks for judging primary MMR studies, Heyvaert et al. (2013) systematically reviewed appraisal frameworks in the MMR literature published up to 2009. Taking an inclusive search strategy, they identified 13 frameworks proposed to evaluate the methodological quality of primary MMR studies. Then, adopting a grounded theory approach, they coded and thematized the 13 identified themes into three broad categories of: (1) generic evaluation benchmarks with nine subcategories, addressing the methodological and epistemological rigor of a study in a general way, (2) specific benchmarks, with two sub-categories, addressing the quality of the qualitative and the quality of the quantitative methods, and (3) MMR-specific benchmarks, with two sub-categories of the necessity and rationale for the use of MMR and the integration quality.

Second, Fàbregues and Molina-Azorin (2017) replicated Heyvaert et al.'s (2013) study by systematically reviewing the quality of MMR studies up to the end of February 2016. They followed three purposes in their review: (1) to examine the trends of MMR literature, (2) to provide an evidence-based summary of the most frequent quality benchmarks, and (3) to suggest quality recommendations for the MMR field. This retrospective-prospective review employed 14 electronic databases, which were then complemented by reviewing methodologically-based journals. The results indicated a sizeable body of research that has addressed the nature of quality (23 research outputs between 2005–2010 and 38 between 2011–2016). As for the criteria set for assessing the quality of MMR studies, they identified 19 criteria, which were then applied to four phases of planning, undertaking, interpreting, and disseminating. The findings, supporting Heyvaert's results, indicated that the identified criteria did not specifically pertain to MMR studies. Instead, they embraced "aspects concerning quantitative and qualitative parts of the study as well as generic research criteria" (Fàbregues & Molina-Azorin, 2017, p. 2856).

As for the MMR quality criteria, Fàbregues and Molina-Azorin found two highly ranked MMR quality criteria, which were: (1) effective integration of the quantitative and qualitative data and analyses and (2) provision of a rationale for using MMR design. On the other hand, they identified other quality criteria that were not highly ranked and included: (1) a transparent report of MMR designs in terms of purpose and emphasis, (2) the linkage of research questions to research designs, (3) interpretive rigor and comprehensiveness, (4) the congruence between rationale and design, and (5) the documentation of crucial MMR literature. They invited prospective researchers to continue working on and publishing frameworks related to MMR quality, working on agreed-upon MMR terminology, and reaching an agreement on a core quality benchmark.

So far, it appears that the two systematic reviews operationalized MMR quality. However, given that MMR quality is highly contingent on context, culture, and discipline, these reviews seem to prescribe a one-size-fits-all recipe to all MMR contexts. As a result, Fàbregues et al. (2019) took the initiative to examine MMR researchers' views on the deployment of quality and how they conceptualized and operationalized quality features in specific disciplines. Adopting a multiple case study design, they examined how 44 international researchers in education, nursing, psychology, and sociology conceptualized and operationalized MMR quality. The 44 international researchers' views on quality resulted in 14 quality criteria. These criteria were not specific to MMR; instead, they embraced quantitative, qualitative, and generic issues.

Furthermore, like the previous studies (e.g., Fàbregues & Molina-Azorin, 2017; Heyvaert et al., 2013), the quality criteria specific to quantitative or qualitative and generic issues outweighed those of MMR. As for MMR-specific criteria, the findings relied on the provision of MMR rationale, the effective integration of quantitative and qualitative methods, and the transparent and accurate description of MMR design and its implementation. Three other criteria with less emphasis were also

https://doi.org/10.1017/S0261444823000332 Published online by Cambridge University Press

identified. These included: (1) an accurate statement of the planned MMR design, (2) a description of the unique inferences drawn from the quantitative and qualitative phases, and (3) congruence between the quantitative and qualitative phases. Regarding views across disciplines, the researchers in nursing and psychology adhered to a fixed perspective (a set of criteria that could be applied to all studies), while those in education and sociology adopted a flexible perspective (that different criteria may apply to different studies). Such findings signify the context-dependency of quality in MMR studies (Creswell, 2015; Plano-Clark & Ivankova, 2015).

Situating the above findings within quality constituents noted earlier, the implication is that MMR researchers and scholars (SUBJECTS) evaluated the MMR studies (OBJECT) using comprehensive benchmarks rather than specific ones (STANDARD) or, in Bryman's (2006) terms, BESPOKE criteria. However, as Heyvaert et al. (2013, p. 321) put it, "[t]here is not yet a consensus on the criteria that should be used to evaluate the quality of primary MMR studies, or on the form in which these criteria should be grouped" (VALUE).

Quite recently, and in response to this debate, three studies have been published. These studies are: Creamer (2018), Harrison et al. (2020), and Hirose and Creswell (2023). Creamer (2018) developed a score-based quality benchmark, Mixed-Methods Evaluation Rubric (MMER). This rubric embraces four mutually exclusive criteria comprising: (1) transparency about the rationale and reasons for mixing methods, (2) level of mixing or integration at different stages, (3) interpretive comprehensiveness, and (4) methodological foundation (i.e., references made to the MMR literature). She then field-tested the newly-developed rubric to address quality in exemplar studies. As Creamer (2018) observed, MMR methodological transparency can promote research replication and "substantiate the results" (p. 154). She conceived of methodological transparency in MMR as "promoting replication by providing explicit detail about the steps taken to complete data collection and analysis as well as by delineating the link between results and the source of data" (p. 154). The design transparency, as Creamer argued, includes the priority or emphasis put on quantitative or qualitative strands (i.e., quantitatively-led, qualitatively-led, or equal status), MMR purpose (i.e., triangulation, complementarity, et.), the nature of MMR designs (i.e., concurrent or sequential or both), and integration of quantitative and qualitative phases in data collection, analysis, and interpretation of the methods. Such transparency conceptualization aligns with the definition proposed by the American Educational Research Association (AERA, 2006, p. 33):

> Reports of empirical research should be transparent; that is, reporting should make explicit the logic of inquiry and the activities that led from the development of the initial interest, topic, problem, or research questions; through the definition, collection, and analysis of data or empirical evidence; to the articulate outcomes of the study.

The main motive behind transparency discussion is to promote research transparency that can facilitate reproducibility and replication in AL research (Marsden et al., 2016). In our review and analysis, we will look for the extent to which AL MMR researchers have transparently explained their study's purpose, design, and methods (sampling, data collection, and data analysis).

Harrison et al. (2020), on the other hand, highlighted the salient role of rigor "as an interdisciplinary baseline for quality evaluation" (p.473). They aimed to develop a rigorous mixed-methods quality framework in light of the related literature and put into practice the newly developed rubric in the management discipline. They developed a two-pronged framework with core and advanced criteria. The core criteria include four components addressing: (1) rigorous data collection of each method, (2) rigorous data analysis of each method, (3) the integration and mixing of the two strands, and (4) the use of a specific mixed methods design type. As for the advanced criteria, the authors referred to the centrality of MMR rationale and rhetorical elements (i.e., referencing mixed methods literature, using joint display (i.e., providing visual displays to show how quantitative and qualitative data and analysis are integrated), and including mixed methods in the title). Harrison et al. believed that these two sets of criteria highlight the methodological rigor associated with various sections within

https://doi.org/10.1017/S0261444823000332 Published online by Cambridge University Press

a manuscript. They used a continuum with three scales of high, medium, and low rigor for each category in their framework. The authors field-tested the framework on 195 articles published in representative journals in management. The results revealed that most studies were characterized as low-rigor ($n = 128$, 65.6%) or low-medium-rigor ($n = 20$, 10.25%). Of the remaining articles, 28 were qualified as medium rigor (14.3%), and 19 (less than 10%) were classified as either medium-high or high rigor MMR studies.

Finally, Hirose and Creswell (2023) synthesized the four crucial sources in the MMR literature on quality criteria. The four sources included Fàbregues and Molina-Azorin (2017), Fàbregues et al. (2019), a United States federal health agency perspective (NIH, OBSS, 2011, 2018), and American Psychological Association (2019). Hirose and Creswell synthesized the four sources and advanced six core criteria. The six core criteria include: (1) providing MMR rationale, (2) posing quantitative, qualitative, and mixed methods questions, (3) reporting the qualitative and quantitative data separately, (4) identifying the type of mixed methods design and depicting a diagram of it, (5) delineating the deployment of integration in a joint display, and finally (6) discussing how meta-inferences and value resulted from the integrated analysis. The authors then implemented the core criteria through a case study. Hirose's (2018) study was used as a case to apply the six quality criteria. The authors found that the six criteria were functional in evaluating the MMR quality. So, more applications of the criteria are needed to evaluate its functionality with a variety of MMR studies.

Overall, the literature on MMR quality can be categorized into two distinctive periods. The first, covering from 2004–2017, represents general research quality benchmarks and portrays comprehensive quantitative, qualitative, and MMR criteria. From 2018 to the present, MMR researchers and methodologists have attended to MMR-specific and more parsimonious criteria. Table 2 summarizes and presents the ranking of the quality criteria in the six frameworks we reviewed above. We used asterisks to represent the level of emphasis put on each quality feature or criterion. Three asterisks show that the criterion was highly emphasized, and three hyphens (---) indicate that the criterion was not discussed in the source.

Based on the above review and the summary presented in Table 2, we used the following six quality criteria to review and analyze the 304 MMR articles published in the 20 AL journals. While acknowledging the additional MMR quality indicators, we uphold the principle of high consensus, as operationalized by those MMR quality elements characterized by the highly emphasized (***) among at least two sources, as outlined in Table 2. The research questions criterion was left out since it received

Table 2. The MMR quality criteria in the six frameworks reviewed

	Framework					
Quality feature	F1	F2	F3	F4	F5	F6
General issues	***	***	***	---	---	---
Quan.–Qual. specific items (data collection and data analysis mainly)	***	***	***	---	***	***
MMR integration	***	***	***	***	***	***
MMR rationale or purpose	***	***	***	***	***	***
MMR design transparency	---	*	***	***	***	***
MMR research question	---	*	---	---	---	***
Meta-inferences/interpretive rigor	---	*	*	***	***	***
MMR rhetoric (documenting MMR with relevant references)	---	*	---	***	***	---

***Highly emphasized; *Less emphasized or missing; ---Not discussed.
F1: Heyvaert et al. (2013); F2: Fàbregues and Molina-Azorin (2017); F3: Fàbregues et al. (2019); F4: Creamer (2018); F5: Harrison et al. (2020); F6: Hirose and Creswell (2023).

https://doi.org/10.1017/S0261444823000332 Published online by Cambridge University Press

highly emphasized from only one source (Hirose & Creswell, 2023). Accordingly, the remaining criteria were used to form the six MMR quality categories:

1. MMR rhetoric (documenting MMR study with relevant references).
2. MMR purpose (identifying a purpose for which the methods were mixed).
3. MMR design transparency (identifying the type of mixed methods design).
4. General issues or method features (including specific items like sampling, data sources, and data analyses of quantitative and qualitative phases).
5. MMR integration (delineating the deployment of integration).
6. Meta-inferences (showing how a meta-inference was developed).

Our six-pronged quality criteria are the same as that of Hirose and Creswell (2023), except that, as stated before, research questions are left out in favour of MMR rhetoric. The designation of MMR rhetoric as the first quality indicator in our framework is warranted, given its central role in shaping and boosting MMR literacy and proficiency (Amini Farsani et al., 2022; Guetterman, 2015, 2017; Riazi, 2018). The second and third criteria highlight the centrality of MMR purpose and design and their transparency in MMR studies. Transparency is discussed and recommended in the recent MMR literature (e.g., Creamer, 2018; Creswell & Plano Clark, 2018), beyond considerations of practicality and functionality of MMR purpose and design. The fourth quality criterion, general issues, reflects the role of quantitative and qualitative rigor in shaping good enough MMR. The fifth criterion is related to how MMR researchers integrate quantitative and qualitative data and analysis (see, e.g., Fetters et al., 2013). Finally, developing meta-inferences (i.e., 1 + 1 = 3) is at the hub of quality-assurance benchmarking (see Riazi, 2017), that is, how meta-inferences and added value resulted from the integrated analysis.

In addition to quality criteria, we thought it might benefit readers to report the topics covered in these 304 articles. As explained in the next section, we followed a research synthesis approach and sought answers to the research questions stated in the introduction using the six quality and transparency criteria.

3. Method and procedures

In this review paper, we followed a synthetic approach (Cooper, 2016; Plonsky & Oswald, 2015) to provide a transparent and objective picture of the status quo of MMR studies published in top-tier AL journals. Employing Plonsky and Oswald's (2015) framework, we pursued four phases to review and analyze published MMR articles. The four phases are:

1. Study identification and retrieval.
2. Developing a coding sheet.
3. Delineating the coding process.
4. Analyzing and interpreting the results.

3.1 MMR study identification and retrieval

Study identification and retrieval refer to defining and characterizing the domain of empirical studies within an AL context (Plonsky, 2013) and include three components:

1. Location or sources of empirical research (e.g., published articles, unpublished M.A. theses or Ph.D. dissertations, and any forms of fugitive literature).
2. Time or temporal scope of the synthesis (e.g., five years, a decade, etc.).
3. Content or the scope of primary studies.

Our sources of empirical research in this study include published MMR studies in top-tier AL journals. The temporal scope covers 2011–2020 (inclusive), and we selected those articles that

https://doi.org/10.1017/S0261444823000332 Published online by Cambridge University Press

explicitly stated the use of MMR in their research report. We relied on two sources to choose the journals and used a criterion to select the top-tier journals. The two sources included: (1) the list of 52 journals related to Applied Linguistics/TESOL, presented at the TESOL Convention and Exhibit in 2015 in Toronto, and (2) the list of journals used by Lei and Liu (2019) and Zhang (2020). Our criterion was to use those AL journals indexed in *Social Science Citation Index* (SSCI) with an impact factor (IF) equal to or higher than one as well as corresponding to the newly developed elite benchmark index projected by Alise and Teddlie (2010). As Alise and Teddlie (2010) put it, the elite benchmark index consists of the Impact Factor[1] (representing a journal's relative significance to the field) multiplied by the Cite Score (indicating the popularity of the journals to create an index of top-tier journals). Using the above two sources of journal lists and the IF index of one or more, we shortlisted 30 journals, as presented in Table 3. We did not include the *Journal of Mixed Methods Research* since our focus was on content-based journals in AL.

We excluded journals of *Bilingualism: Language and Cognition, Applied Psycholinguistics*, *Cognitive Linguistics*, *International Journal of Bilingual Education and Bilingualism,* and *The International Journal of Multilingualism* from Table 3. They were excluded from our corpus since searching those journals with the queries of MMR keywords over 2011–2020 did not return any studies. Furthermore, using the elite journal index ($IF^{*}CS \geq 3.5$) suggested by Alise and Teddlie (2010), we left out another five journals, including *Pragmatics*, *Language and Education*, *Language Acquisition*, *Applied Linguistics Review*, and *Language Policy*. Finally, two more journals, *Language Teaching* and *Annual Review of Applied Linguistics,* were also excluded as their aim and scope differed from other journals and because most of the articles are commissioned. The point should be made here that commissioned articles and review studies was also excluded from other journals. Only empirical articles (those that collected first-hand data) were included in our review.

Overall, we could identify 5,093 articles published in 20 top tier journals with quantitative, qualitative, mixed-methods, and review approaches. We adhered to the MMR transparency mantra (Creswell & Plano Clark, 2018) to select articles with only an MMR approach. This approach helped us set stringent rather than lenient criteria for locating the candidate MMR studies. We think it is essential given the recent movement of methodological awareness or, in McKinley's (2020) terms, the "golden age of research in applied linguistics" (p. 2). Thus, we manually skimmed and scanned all the articles and targeted those that transparently used mixed-methods collocations (e.g., mixed methodology, sequential mixed methods research, etc.). This was done by searching for the terms and collocations of "mix*, mixed-methods, mixed-methods design, mixed-methodology, mixed research, and mixed approach".

The reason for using the exact MMR terms instead of "quantitative AND qualitative" strings was that from the first review (Riazi & Candlin, 2014), there has been a substantive number of studies in AL journals with an explicit MMR orientation. This trend might have been partly because some journals like *TESOL Quarterly* have used Riazi and Candlin's (2014) review as a guide for MMR contributors. In addition, there have been many studies that have addressed different aspects of MMR in AL (e.g., Amini Farsani et al., 2021, 2022; Brown, 2014; Hashemi, 2020; Riazi, 2016, 2017; Riazi et al., 2018, 2020). Therefore, the use of strict and transparent queries is warranted given the awareness-raising movement in MMR research in AL since the first state-of-the-art MMR paper was published in *Language Teaching*. However, analysis of studies that used both quantitative and qualitative data and analysis without framing themselves as MMR would be another interesting study. In this review paper, our focus was on those studies that authors explicitly referred to MMR as the methodological approach in their reported studies.

Our search for MMR studies in the 20 AL journals included in our corpus led to 304 articles published over 2011–2020 (inclusive). Figure 1 shows the distribution of these 304 papers over the study period (2011–2020). We included these 304 articles in our review and analysis.

https://doi.org/10.1017/S0261444823000332 Published online by Cambridge University Press

Table 3. The profile of the selected AL journals

Journals	Impact factor	Cite score	IF*CS
Journal of Second Language Writing	4.2	5.6	23.52
Modern Language Journal	3.762	3.39/6	22.57
Applied Linguistics	3.041	6.3	19.16
Bilingualism: Language and Cognition[a]	2.701	4.8	12.96
Language Learning & Technology	2.571	4.6	11.83
Language Teaching Research	2.319	5.1	11.83
TESOL Quarterly	2.718	4.3	11.69
Studies in Second Language Acquisition	2.6	4.1	10.66
Computer-Assisted Language Learning	2.018	4.6	9.28
Language Learning	2	4.5	9
System	1.930	3.3	6.369
English for Specific Purposes	1.704	3.9	6.65
Applied Psycholinguistics[a]	1.76	3.5	6.16
Assessing Writing	1.841	3.2	5.86
Second Language Research	1.750	3.1	5.425
Cognitive Linguistics[a]	1.630	3.3	5.38
Language Culture and Curriculum	1.571	3.4	5.34
ReCALL	1.361	3.5	4.76
Journal of English for Academic Purposes	1.732	2.7	4.68
Language Testing	1.154	3.7	4.27
Journal of Multilingual and Multicultural Development	1.639	2.3	3.77
Foreign Language Annals	1.782	2.1	3.74
The International Journal of Multilingualism[a]	1.375	2.7	3.71
International Journal of Bilingualism[a]	1.259	2.9	3.65
Linguistics and Education	1.516	2.3	3.49
Journal of Pragmatics[a]	1.329	2.6	3.45
Language and Education[a]	1.164	2.8	3.26
Language Acquisition[a]	1.147	2.5	2.87
Applied Linguistics Review[a]	1.098	2.5	2.745
Language Policy[a]	1	2.7	2.7

[a]Excluded journals.

3.2 Developing a coding sheet to code the articles

We developed a coding sheet drawing on the guidelines set in the literature for developing coding sheets in research synthesis (see, e.g., Cooper, 2016; Plonsky, 2013, 2014; Plonsky & Oswald, 2015). We also based our coding scheme on the categories we formulated after we reviewed the different quality and transparency frameworks in Section 2.1. As such, our coding sheet consisted of the following categories:

https://doi.org/10.1017/S0261444823000332 Published online by Cambridge University Press

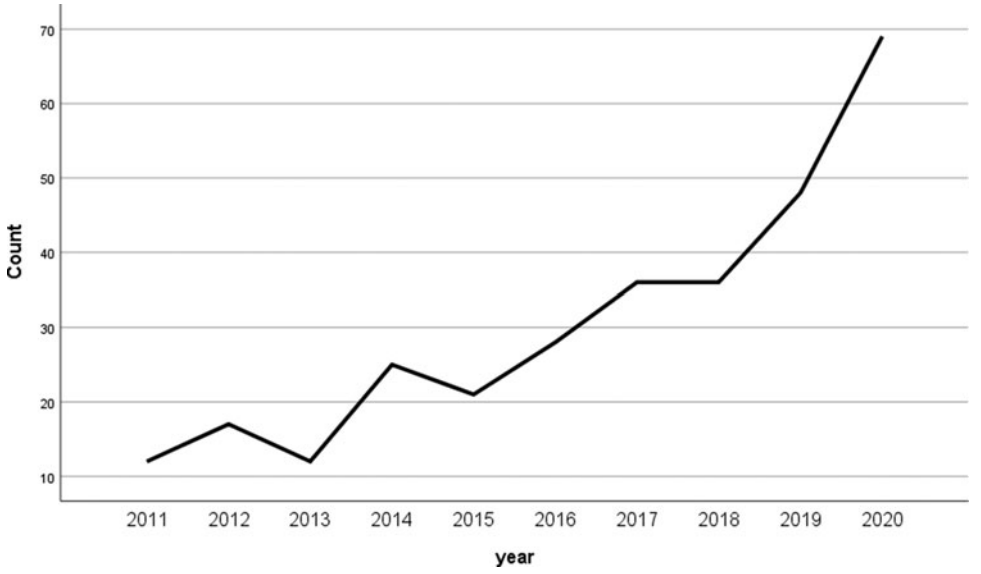

Figure 1. Distribution of MMR studies over the study period (2011–2020)

(a) Bibliographical information (e.g., journal name, year, paper's title, authors, etc.).
(b) MMR term used in the article (e.g., its representation in the title, abstract, keywords, methods, results, and discussion).
(c) MMR rhetoric (using relevant references to document the MMR study).
(d) MMR purpose and design-related features (providing a rationale for mixing methods and delineating the MMR study design).
(e) General issues or method features (sampling, data sources, and data analysis).
(f) Integration of the two strands in the study and deployment of meta-inferences.

To substantiate the above coding categories, we used Greene et al.'s (1989) to code MMR purpose, and we used Creswell and Plano Clark (2018) and Riazi and Candlin (2014), and Riazi (2016) to code design structures and transparency of the designs.

The above coding categories and substantiations helped us examine the level of reporting transparency for each quality category. That is, how explicitly the authors had explained each of the above categories (see online Supplementary Appendix A).

In addition to the coding sheet, we also prepared review notes for each article to supplement the coding with evaluative comments. The review notes for each article included 5–10 pages and delved into various aspects of the MMR approach utilized in the studies. The primary purpose of the notes was to locate, reveal, and illuminate the intricate details pertaining to the strengths and weaknesses of the studies, with a specific concentration on the extent to which they complied with the six quality criteria discussed before. The coding sheet and the review notes helped us achieve what Onwuegbuzie and Corrigan (2014) called for: a transparent, defensible, evaluative, systematic, and comprehensive review. It also enabled us to assess each study's quality regarding "adherence to standards of contextually appropriate, methodological rigor in research practices and transparent and complete reporting of such practices" (Plonsky, 2013, p. 658).

3.3 Delineating the coding process (coding issues and reliability estimates)

Coding the first three categories (a, b, and c in Section 3.2) was straightforward since they covered factual information. As for the reliability of the other categories in the coding sheet, we first coded

https://doi.org/10.1017/S0261444823000332 Published online by Cambridge University Press

Table 4. Reliability indices for the coding sheet categories

Category	Percentage agreement (%)
MMR purpose and design	87
Method features	90
Integration of the two strands and developing meta-inferences	89

a few articles and negotiated the discrepancies and inconsistencies. Each of the two authors then independently coded 20 studies, and the overall percentage agreement between the two coders was about 0.89. In 89% of the cases, we coded the purpose and design, method features, and integration of the two strands in the same way. As Norris and Ortega (2006) put it, "it is essential that reliability be considered and reported not simply overall, but rather for each category under examination" (p. 26). Thus, we present the results of inter-coder reliability for each of the categories in the coding sheet in Table 4 based on the coding of the 20 articles.

Given that many studies did not explicitly state purpose and design (see Table 6), it was normal to have a 13% discrepancy in that category. However, we negotiated these discrepant cases to reach an agreed-upon decision. Following that and considering the high inter-coder reliabilities, the second author rated the whole set of MMR studies. However, to account for the discrepancies in our initial coding, we rechecked the codings on each batch of 50 articles by reviewing and discussing the coding sheets and review notes. This coding procedure was followed throughout the coding process and took 18 months to complete. Once all the 304 articles were coded, we estimated the percentages and frequencies of the coded categories. In addition to the percentages and frequencies, we traced and explored the latent aspects of the data using the review notes we prepared for each article.

4. Results

4.1 RQ1: Topic areas covered in the articles

Using R, Python, and Concordancer software packages, we extracted the keywords represented in the 304 articles. As depicted in the lexical dispersion profile (see online Supplementary Appendix B), we could identify the frequency of use of the key terms. We found that "language", "writing", "online", "foreign", "feedback", "perceptions", and "motivation" and their collocations were the most frequent ones in the dataset, as shown in online Supplementary Appendix C. We then aligned the whole list of keywords with the AL strands disseminated by the American Association for Applied Linguistics (AAAL) and the British Association of Applied Linguistics (BAAL). Table 5 presents the alignment of the keywords with corresponding strands in AAAL and BAAL. The numbers in parentheses show the frequency of the keywords in the dataset. Furthermore, the total number of articles in each strand is provided in the last column of Table 5. The results revealed that almost all the strands of AL used MMR.

Topic-wise, we found that MMR has been employed to examine diverse language-related problems, as outlined in AL strands. It covers all the language-related problems represented in AL strands. Furthermore, as Table 5 presents, the rise and potential of other strands to employ MMR is notable, given the potential of MMR to deal with L2-related problems from different angles. However, the current utilization of MMR is not proportionate and symmetrical, with certain strands, such as SLA, receiving a higher degree of attention. This is fortunate for the field as SLA-oriented researchers are on the verge of digressing from epistemologically one-sided perspectives to multiple perspectives, which consequently affect the quality of inferences and insights in the research projects.

4.2 RQ2: MMR documentation patterns

In the introduction section of this review paper, we highlighted some indicators of how MMR is gaining momentum in AL research. One significant indicator is whether researchers document and cite

https://doi.org/10.1017/S0261444823000332 Published online by Cambridge University Press

Table 5. The alignment of the keywords with corresponding strands in AAAL and BAAL

AAAL and BAAL strands	Keywords	No. of articles (%)
Language and technology (TEC)	Online (110); Technology (27); Digital (26); Mobile (17); Synchronous (14); Virtual (10)	57 (18.75)
Second language acquisition, language acquisition, and attrition (SLA)	Feedback (85); Motivation (82); Linguistic (62); Engagement (40); Self-efficacy (38); Communication (36); Grammar (36); Listening (36); Motivational (37); Interaction (34); Communicative (30); Critical (30); Errors (27); Collaborative (29); Instructional (29); Peer (28); WTC (23); Anxiety (22); Engage (22); Acquisition (21); Identity (18); Improvement (17); Pronunciation (17); Sociocultural (17); Accuracy (16); Self-regulation (13); Syntactic (13); Autonomy (10); Efficacy (10)	55 (18.09)
Reading, writing, and literacy (RWL)	Writing (299); Reading (41); Written (37); Literacy (29); Writers (26); Essays (18); Argumentative (17); Plagiarism (17); Voice (10)	33 (10.85)
Assessment and evaluation (ASE)	Test (71); Assessment (62); Rater (34); Assess (23); Evaluation (15); AWE/ automated writing evaluation (12)	29 (9.53)
Second and foreign language pedagogy (PED)/context	Foreign (104); EFL (81); Chinese (78); Spanish (41); Japanese (29); China (28); ESL (18); Pedagogy (18); Flipped (24); French (16); Germany (16); Kanji (15); Korean (14)	28 (9.21)
Teacher education, beliefs, and identities (TED)	Beliefs (56); Identity (18)	22 (7.23)
EAP/ESP	ESP (22); EAP (17)	21 (6.90)
Bilingual, immersion, heritage, and language minority education (BIH)	Multilingual (21); Bilingual (16); Heritage (12)	20 (6.57)
Language, culture, socialization, and pragmatics (LCS)	Cultural (47); International (47); Intercultural (33); Pragmatic (21); Culture (19)	20 (6.57)
Analysis of discourse (DIS)/ Text analysis (written discourse) (TXT)	Discourse (18); Genre (13)	6 (1.97)
Language, planning and policy (LPP/POL)	Curriculum (19); Textbooks (16)	5 (1.64)
Corpus Linguistics (COR)	Corpus (23)	4 (1.31)
Vocabulary and Lexical Studies (VOC)	Vocabulary (52); Lexical (16)	4 (1.31)

Table 6. MMR purposes identified in the 304 articles

	Implicitly stated	Explicitly stated	Total
Purpose	*N*	*N*	*N* (%)
Triangulation	81	61	142 (46.7)
Complementarity	46	36	82 (27)
Expansion	48	0	48 (15.80)
Development	23	4	27 (9)
Initiation	3	2	5 (1.6)
Total	201	103	304 (100)

https://doi.org/10.1017/S0261444823000332 Published online by Cambridge University Press

relevant sources for using MMR and its features (Fàbregues et al., 2021). Also, MMR rhetoric was one of the quality criteria discussed in the MMR quality frameworks. Therefore, we examined the reference sections of all 304 articles to identify the references to MMR literature. Surprisingly, of 304 articles, more than half of them ($n = 179$, 58.9%) did not reference the MMR sources to frame their MMR study, that is, delineating the purpose and design. Only 125 articles (41.1%) referenced MMR literature to frame the MMR studies. We examined the MMR references used in the 125 articles and were able to categorize the MMR sources used into three categories, namely: (1) multidisciplinary sources; that is, MMR sources targeting different disciplines, (2) discipline-specific sources; that is, MMR sources targeting AL discipline, and (c) general sources; that is, general research methodology sources. Overall, there were 226 MMR references cited in the 125 articles. Multidisciplinary MMR sources were the dominant category with 170 out of 226 instances (75.22%), followed by discipline-specific with 40 (17.7%) instances, and general research methodology sources with 16 cases (7%) (see online Supplementary Appendix D for the breakdown of the sources used).

As stated above, the bulk of references to MMR was multidisciplinary. We found that John Creswell and collaborators ($n = 92$) were the most represented MMR figures in AL studies. This was followed by Charles Teddlie and collaborators ($n = 26$), Burk Johnson and collaborators ($n = 13$), and Green and collaborators ($n = 12$).

Regarding discipline-specific categories, we identified two unique patterns. Of 40 studies that used AL-specific MMR sources, 22 adopted general research methodology sources in AL. One of these sources was Dörnyei's (2007) book that addresses and discusses quantitative, qualitative, and MMR research. The second pattern was citations to AL-MMR-specific sources (e.g., Hashemi & Babaii, 2013; Riazi & Candlin, 2014) and included 18 articles. Some articles used field-specific MMR sources such as computer assisted language learning (CALL) (Cerezo, 2016), complex dynamic system (Gilmore & Ganem-Gutierrez, 2020), language testing and assessment (Jang, Wagner, & Park, 2014), SLA (Singleton & Pfenninger, 2015), and needs analysis (Long, 2005). The two books on MMR, Brown (2014) and Riazi (2017), were almost missing in all the articles, with Brown's book cited only in one.

The smallest category was related to those that had concentrated on general aspects of research in different disciplines. In this category, Cohen et al. (2002) was the most frequent source used in the AL MMR studies (see online Supplementary Appendix E). MMR researchers are expected to refer to relevant MMR literature, especially the key ones.

4.3 RQ3: Purpose and design structures

4.3.1 Purpose of mixing quantitative and qualitative methods

As stated in the methods section, we used Greene et al.'s (1989) purposes typology to identify and code the specific purposes in the 304 articles. The purposes of mixing quantitative and qualitative methods identified by Green et al. include "triangulation" (corroborating findings from one method with those from another), "complementarity" (addressing different dimensions of a research problem with different data and analysis), "development" (using the outcomes of one strand to develop another strand), "initiation" (drawing on the conflicting results to initiate a new MMR study), and "expansion" (adding another set of data and analysis to expand the scope of the study). Table 6 summarizes the purposes of mixing quantitative and qualitative data and analysis in the 304 articles.

As seen in Table 6, only 103 articles (34%) explicitly stated a purpose for which the researchers mixed the two methods. However, 201 articles (66%) did not do so. We thus used the information in the articles to code (see Section 3.3 for the coding process) the purpose for mixing the two methods.

Based on the explicit and implicit coding of the purposes, we found that triangulation purpose ($n =$ 142, 46.71%) – which seeks "convergence, corroboration, correspondence of results from different methods" (Greene et al., 1989, p. 259)" – was the most frequent purpose for mixing the two sets of data and analysis. For example, Sato (2013), who explicitly stated triangulation as the purpose, examined L2 language learners' beliefs toward peer interaction and peer corrective feedback in a

https://doi.org/10.1017/S0261444823000332 Published online by Cambridge University Press

collaborative approach intervention for improving speaking fluency. To triangulate the quantitative and qualitative results, the researcher traced the belief profiles before and after instruction using pre-post survey questionnaires and semi-instructed interviews and supported results from one method with those from another. Likewise, Révész et al. (2019), adopting a convergent MMR design, examined the cognitive processes of Chinese L2 writers at various textual locations and different levels of revision. They used stimulated recall, keystroke logging, and eye-tracking to triangulate information about the learners' cognitive processes, text production behaviors, and reading practices during pauses and before revisions. A similar study is Gánem-Gutiérrez and Gilmore (2018).

The second most frequently used purpose was complementarity, which was pursued in almost one-third of the studies ($n = 82$, 27%). Applied linguists pursuing this purpose sought "to measure overlapping but also different facets of a phenomenon, yielding an enriched, elaborated understanding of that phenomenon" (Greene et al., 1989, pp. 258–259). Moeller and Theiler (2014) shed light on different aspects of the development of the spoken Spanish language at the high school level. They quantitively monitored students' progress in speaking over four years of high school Spanish learning. This is followed by a qualitative analysis of the students' speech samples. As they explicitly reported with specific references, the purpose of the study was complementarity, "a design element used to measure overlapping but distinct facets of a phenomenon under investigation" (Caracelli & Greene, 1993, p. 196).

Two other purposes, namely, "expansion" ($n = 48$, 15.8%) and "development" ($n = 27$, 9%), received less attention. "Expansion" purpose – seeking to broaden the breadth and depth of inquiry – was implemented in almost 16% of the studies. "Development" purpose with its informative nature (i.e., the results from one phase inform the development of the other phase) was followed in 9% of the studies. Finally, studies undertaken with an "initiation" purpose ($n = 5$, 1.6%) were the least frequently represented in the MMR studies. The point must be raised here that apart from two purposes (triangulation and development), the other three purposes are challenging to conceive and implement. For example, in the complementarity purpose, it is challenging to conceptualize a research problem as multi-faceted or multi-dimensional, with each facet or dimension lending itself to a certain type of data and analysis. Or, for the expansion purpose, it is not immediately obvious why and when there is a need to expand and how the expansion is to be designed. Similarly, it is said that when researchers face contradictory results from previous research, they may initiate an MMR study to address the conflicting results with different data and analysis. However, little is discussed to unpack the subtilities of the initiation study. There is, therefore, a need for discussing and clarifying these issues and challenges when MMR purposes are discussed.

4.3.2 Design structures

Table 7 presents the MMR designs based on Creswell and Plano Clark's (2018) design structures. The results revealed that AL researchers used both types of MMR designs: core and complex. Creswell and Plano Clark's core designs correspond to the triangulation and development purposes of Greene et al. (1989). The first core design is "convergent design", which is exactly the same as the triangulation purpose. That is, when the outcomes of one data and analysis method converge with those from the other method. Considering development purpose, Creswell and Plano Clark have broken it down into explanatory and exploratory sequential designs. When researchers use the outcome of a quantitative survey to collect more in-depth qualitative data from a sub-sample of the survey study, they are doing a sequential explanatory design. On the other hand, when researchers use the outcome of an initial qualitative study to design and administer a survey to a larger sample, they are using a sequential exploratory design. Our results showed that core designs ($n = 214$, 71.38%) were used more than complex designs ($n = 77$, 25.32%). The results also indicated that "explanatory sequential designs" ($n = 110$, 36%), using qualitative data and analysis to explain the quantitative findings, were more popular than "convergent" ($n = 88$, 28.9%) designs in which quantitative and qualitative results are used for triangulation purpose, and "exploratory sequential" ($n = 16$, 5.3%) designs, using quantitative data and analysis to explore qualitative findings through larger samples. Complex designs include embedding a qualitative phase in the process of either and experiment or a pre-post survey. A multistage design

https://doi.org/10.1017/S0261444823000332 Published online by Cambridge University Press

Table 7. MMR designs in the 304 articles

		Transparency		
		Implicit N (%)	Explicit N (%)	Total N (%)
Core designs	Explanatory sequential	71 (64.54)	39 (35.45)	110 (36)
	Convergent design	64 (72.72)	24 (27.28)	88 (28.9)
	Exploratory sequential	11 (68.75)	5 (31.25)	16 (5.3)
	MMR conversion	3 (100)	0	3 (0.98)
Complex designs	MMR embedded designs	29 (72.5)	11 (27.5)	40 (13.2)
	Hybrid/Multistage	18 (75)	6 (25)	24 (7.89)
	Embedded pre-post survey	13 (100)	0	13 (4.27)
Other	MMR content analysis	4 (66.6)	2 (33.3)	6 (1.97)
	Q-methodology	0	3 (100)	3 (0.98)
	MMR action research	1 (100)	0	1 (0.33)
Total		214 (70.4%)	90 (29.6%)	304 (100)

alludes to different quantitative and qualitative data collection and analysis at different stages in the process of the research.

As Table 7 presents, many articles ($n = 214$, 70.4%) did not explicitly describe the MMR design structure, which we think they must have. Only one-third of the articles did so explicitly.

The most dominant complex design, as Table 7 shows, was "MMR embedded designs" with two variants. The first variant is "MMR experimental" designs' ($n = 40$, 13.2%), in which a qualitative phase is embedded into an experiment. An example is Alwaleedi et al. (2019), who examined the impact of collaborative writing on students' writing performance in an Arabic context. They used a quasi-experimental design (i.e., pre-and post-test scores of students in experimental and control groups) and a qualitative case study design (classroom observations and audiotapes of verbal interactions). The second variant was named "MMR nested or embedded pre-post survey design" ($n = 13$, 4.27%). For example, Han and Hiver (2018) examined the motivational change processes for 174 middle school language learners due to genre-based writing instruction. There was a pre-post-survey and an intervention (teaching the students genre-based writing). The pre- and post-survey data were quantitatively analyzed to check the changes that occurred to participants' psychological factors due to the intervention (genre-based writing instruction). The researchers collected reflective journals through the intervention process and arranged three interviews with a sub-sample to explain the quantitative survey findings. This is very similar to an experimental design except that instead of pre- and post-test, the researchers used pre- and post-survey. We call this design MMR nested or embedded pre-post survey design.

The second most frequent advanced design is the hybrid design (Johnson & Christensen, 2019; Riazi, 2017), in which one or both strands of the MMR will include more than one phase. For example, in [(quan + QUAL) → QUAL] hybrid design, we have two strands. The first strand consists of two phases: quan and QUAL (capital letters show more emphasis on the strand). The second strand includes only one phase, QUAL. As Table 8 presents, AL researchers predominantly used two-strand designs with a combination of sequential and convergent designs ($n = 15$, 62.5%) and two or more strands with sequential designs ($n = 9$, 37.5%) to address L2-language problems. For example, Copland et al. (2014) used a hybrid design [(QUAN + qual) → QUAL)] to address challenges English teachers encountered when they were working with young learners. The first strand included two phases: one primary quantitative phase, and a second more minor qualitative phase. The second

https://doi.org/10.1017/S0261444823000332 Published online by Cambridge University Press

Table 8. A profile of MMR hybrid designs

Hybrid designs	Notations' variations*	*N* (%)
Combination of convergent and sequential	(quan + QUAL) → QUAL (QUAL + quan) → qual (QUAN + QUAL) → (QUAN + qual) QUAL + (QUAN → QUAL) (quan + QUAL) → QUAN → QUAL	15 (62.5)
Expanded sequential	qual → QUAN → qual QUAN → QUAL → QUAN (QUAL → QUAN) → (QUAL → QUAN)	9 (37.5)

QUAN: major quantitative; quan: minor quantitative; +: convergent.
QUAL: major qualitative; qual: minor qualitative; →sequential.

strand had only one phase: a major qualitative one. These designs depend on the research problem and research questions and researchers' familiarity with MMR designs.

Macaro and Lee (2013), adopting another version of the hybrid MMR design, examined the perceptions of Korean learners of English towards English-only instruction as opposed to the type of instruction switching to the learners' first language. They followed a sequential design. In the first strand, they conducted interviews to construct a questionnaire. In the second strand, they created a questionnaire using the categories extracted from the first phase and distributed the questionnaire to a larger sample. Finally, they conducted another interview with the teachers to get confirmative results. The notation for this hybrid design is QUAL → QUAN → qual.

Besides core and complex designs presented in Table 8, we found some other MMR designs. One conspicuous finding was the use of *Q*-methodology for examining motivation from the lens of CDST. *Q*-methodology is similar to MMR in that the perspectives of participants on an issue are collected by having participants rank and sort a series of statements (Riazi, 2016). Newman and Ramlo (2010) believed that *Q*-methodology lends itself to the framework of mixed methods research. The philosophical underpinnings of the *Q* are a "mixture of qualitative and quantitative ideas" (Newman & Ramlo, 2010, p. 506). *Q*-methodology "neither tests its participants nor imposes meanings a priori. Instead, it asks its participants to decide what is 'meaningful' and hence what does (and what does not) have value and significance from their perspective" (Watts & Stenner, 2005, p. 74). *Q*-methodology was used in three studies addressing students' motivational profiles of 15 Chinese learners engaging in L2 and L3 learning (Zheng et al., 2020), focusing on future language selves and future multilingual self-vision of students of Korean language in an Australian university (Fraschini & Caruso, 2019), and profiling Chinese students' motivation to learn multiple languages (Zheng et al., 2019). The CDST approach helped the researchers to (a) conceptualize motivation as a complex construct and (b) address the dynamics of the construct.

Regarding the principled and innovative nature of the MMR studies (Riazi & Candlin, 2014; Riazi, 2016), we found that more than half of the studies ($n = 196$, 64.1%) could be categorized as principled MMR. The rest were coded to the eclectic ($n = 56$, 18.4%) and innovative ($n = 53$, 17.4%) categories.

We further evaluated the priority, or the drive, as one of the critical features of MMR designs. As noted by Johnson, Onwuegbuzie, and Turner (2007), MMR studies can be represented by three priorities or drives: (1) qualitatively driven or qualitative dominant MMR; (2) quantitatively driven or quantitative dominant MMR; and (3) qualitative-quantitative continuum or equal status MMR. We found that nearly half of the 304 MMR studies ($n = 131$, 43.1%) gave equal weight to the quantitative and qualitative methods considering data and analysis. Ninety-five (31.3%) studies were coded as quantitatively driven, while 78 (25.7%) were coded as qualitatively driven.

4.3.3 Epistemological orientations

AL MMR researchers' epistemological orientation was not included in our review paper. However, one of the reviewers of the paper was concerned about the absence of this important aspect of MMR in this

https://doi.org/10.1017/S0261444823000332 Published online by Cambridge University Press

state-of-the-art paper. To respond to this concern, we surveyed all 304 articles to see if the authors of those articles alluded to any epistemological orientation in their MMR studies. We found that only eight studies (2.63%) explicitly referred to an epistemological orientation in their studies. The stated epistemological orientations included pragmatism ($n = 4$), transformative ($n = 2$), pluralistic ($n = 2$), and critical realism ($n = 1$). For example, Kwan and Dunworth (2016), adhering to pragmatism, investigated the use of English as a lingua franca in domestic workplaces in Hong Kong. In describing their MMR design, they noted: "The study was undertaken within the paradigm of pragmatism and adopted a mixed methods approach. Within pragmatism, knowledge is seen as socially situated, arising from actions and situations (Creswell, 2009) and having multiple facets and dimensions (Tashakkori & Teddlie, 2010)" (Kwan & Dunworth, 2016, p. 15).

Cai and Zhu (2012) and Han and Hiver (2018) framed their study within a transformative epistemology. The former studied the effectiveness of an online learning community project on university students' motivation to learn Chinese as a foreign language, while the latter investigated the psychological profiles of L2 learners in genre-based L2 writing instruction. The only study with a critical realist stance belonged to Razfar and Simon (2011); they conducted a longitudinal study comparing the success of students who mainstreamed into college-level content courses and those who did not. This very low number of studies with a clearly stated epistemological orientation is an issue worth considering in the discussion and recommendations that follow in the later sections of this review paper.

4.4 RQ4: Method features (sampling, data sources, and data analysis)

4.4.1 Sampling procedures

One of the most salient features of MMR studies is the sampling procedure and "how the sample or samples for the qualitative and quantitative data collection are related to each other" (Fetters, 2020, p. 222). Following a bidimensional model of MMR sampling proposed by Onwuegbuzie and Collins (2007), we categorized the MMR sampling procedure and how the quantitative and qualitative samples are related. The four categories suggested by Onwuegbuzie and Collins are identical, parallel, nested, or multilevel, as discussed earlier. Table 9 presents a profile of the MMR sampling procedures found in the 304 studies.

As Table 9 presents, it was found that AL researchers used a variety of MMR sampling procedures in sequential and convergent designs. As for MMR sampling in sequential design, we found that a "sequential design with nested samples" ($n = 79$, 26%) was more represented than identical ($n = 21$, 6.9%), multilevel ($n = 18$, 5.9%), and parallel ($n = 12$, 3.9%) relationships, respectively. Regarding MMR sampling in convergent designs, we found that a "convergent design with identical samples" ($n = 40$, 13.2%) was more popular than the other sampling procedures.

As for methodological transparency for reporting MMR sampling, it seems that MMR sampling transparency was utterly absent in the dataset. That is, while the authors mentioned the samples for the two strands, they did not explain how the two samples were related to each other. Hence, we IMPLICITLY identified MMR sampling in light of the main components of the sampling design reported in each article.

Besides reporting the MMR sampling relationships in the studies, we examined the sampling representations in both quantitative and qualitative phases separately. Table 10 presents the sampling

Table 9. A profile of MMR sampling procedures in the 304 articles

Sequential MMR sampling					Convergent MMR sampling			
Nested	Identical	Parallel	Multilevel	Other	Nested	Identical	Parallel	Multilevel
N (%)	*N* (%)	*N* (%)	*N* (%)	*N* (%)	*N* (%)	*N* (%)	*N* (%)	*N* (%)
79 (26)	21 (6.9)	12 (3.9)	18 (5.9)	9 (3)	16 (5.3)	40 (13.2)	12 (3.9)	18 (5.9)

https://doi.org/10.1017/S0261444823000332 Published online by Cambridge University Press

Table 10. A profile of quantitative and qualitative sampling procedure represented in the MMR studies

Phases	Sampling scheme						
Quantitative	Convenience *N* (%)	Purposive *N* (%)	Stratified *N* (%)	Snowball *N* (%)	Random *N* (%)	Cluster *N* (%)	Other *N* (%)
	213 (70.1)	55 (18.1)	10 (3.3)	6 (2)	5 (1.6)	3 (1)	12(3.9)
Qualitative	Convenience *N* (%)	Criterion-based *N* (%)	Maximum variation *N* (%)	Extreme case *N* (%)	Typical case *N* (%)	Other *N* (%)	
	155 (51)	88 (28.9)	9 (3)	6 (2)	2 (0.7)	44(14.5)	

procedures in the two strands of the MMR studies. Given the dominance of the quantitative method in MMR studies (see Section 4.3), we initially traced the representation of quantitative sampling in MMR studies in terms of the sampling scheme, sample size, and transparency. The results indicated that nonprobability sampling schemes were more employed than probability ones. A significant portion of the MMR studies ($n = 213$, 70.1%) employed convenience sampling in the quantitative phase. This was followed by 18.1% for purposive sampling. The least prevalent sampling scheme was cluster sampling ($n = 3$, 1%). Transparency-wise, we found that explicitly reporting of quantitative sampling method was not followed in almost 69% of the studies ($n = 210$); just one-third of them ($n = 94$, 30.9%) provided a transparent reporting of quantitative sampling.

Furthermore, concerning reporting qualitative sampling procedures as presented in Table 10, we found that half of the studies ($n = 155$, 51%) employed convenience sampling for their qualitative phase. With a gap, the second frequent sampling used in MMR studies was criterion-based. Typical case sampling was used as the least frequent qualitative sampling. Regarding transparency, the results revealed that more than half of the studies ($n = 190$, 62.5%) did not explicitly describe the qualitative sampling procedure.

Table 11 presents the MMR studies' sample size[2] of quantitative and qualitative phases. The results revealed that 50% of the studies had a sample size of 64 in the quantitative phase. The lower and upper 25% and 75% quartiles were 32 and 193, respectively. As for the sample size of the qualitative phase, it was found that 50% of the studies had a sample size of 18, while the lower and upper 25% and 75% quartiles were reported as 8 and 37, respectively. It is indeed understandable that the qualitative samples are smaller than the quantitative samples. This is because of the nature of qualitative studies that require in-depth data collection from smaller samples to shed more light on research issues.

4.4.2 Data sources

Table 12 presents the quantitative and qualitative data collection techniques and data sources AL researchers used in their studies.

Regarding quantitative data collection techniques and data sources, we found that close-ended questionnaires ($n = 197$, 64.8%) were notably more utilized than other techniques and sources such as tests ($n = 91$, 29.9%), coding sheets ($n = 29$, 9.53%), and texts ($n = 25$, 8.22%). We also found that a few AL researchers used the eye-tracking technique to collect data ($n = 6$, 2%). On the other hand, our findings, as represented in Table 12, revealed that the most frequent qualitative data sources were interviews ($n = 199$, 65.46%), open-ended questionnaires ($n = 78$, 25.65%), and observational field notes ($n = 41$, 13.48%), respectively. The least frequent data sources were stimulated recall and think-aloud protocols.

Besides looking at data sources in each of the methods in MMR studies, we also examined the combined use of the data sources in the articles. Table 13 presents a rank-ordered use of combined data sources in the 304 articles.

As presented in Table 13, the use of closed-ended questionnaires and an interview was substantially higher than other combination patterns. This is probably because this combination is the trend and is easy to design and implement.

https://doi.org/10.1017/S0261444823000332 Published online by Cambridge University Press

Table 11. Median and lower and upper quartiles of sample sizes

	Median	Lower 25%	Upper 25%
Quantitative	64	32	193
Qualitative	18	8	37

Table 12. Data sources used for collecting quantitative and qualitative data

	No. of articles[a]	Number (Percentage)
Quantitative data collection techniques	Closed-ended questionnaires	197 (64.8)
	Tests	91(29.9)
	Coding schemes	29 (9.53)
	Texts[b]	25 (8.22)
	Rubrics	13 (4.27)
	Other quantitative data collection techniques, such as eye-tracking	6 (1.97)
Qualitative data collection techniques	Interviews	199 (65.46)
	Open-ended questionnaires	78 (25.65)
	Observational field notes	41 (13.48)
	Texts	32 (10.52)
	Diary/journal logs	32 (10.52)
	Stimulated recall interviews	16 (5.26)
	Think-aloud protocols	9 (2.96)

[a]In some of the studies, we found more than one data source, so the frequency exceeds 304.
[b]Includes documents such as essays and textbooks, which quantitively collected and analyzed.

4.4.3 Data analyses procedures

Table 14 presents the breakdown of the parametric and nonparametric statistical procedures used to analyze the quantitative data in one of the two strands in the MMR studies. The point should be made here that some studies used more than one statistical test of significance; thus, the total in Table 14 is more than 304.

Parametric tests were used more dominantly ($n = 295$, 77.63%) than the nonparametric counterparts ($n = 81$, 21.31%). Considering parametric tests of significance, *t*-tests ($n = 95$, 25%), Pearson correlation ($n = 52$, 13.68%), and one-way ANOVA ($n = 40$, 10.52%) were prevalently used in the MMR studies. Regarding nonparametric tests, Chi-square ($n = 26$, 6.84%), Wilcoxon signed-rank test ($n = 16$, 4.25%), Mann-Whitney U test ($n = 13$, 3.42%), Spearman rho ($n = 10$, 2.65%), Friedman test ($n = 5$, 1.32%); Kruskal Wallis ($n = 4$, 1.06%), and Log-likelihood ($n = 2$, 0.53) were frequently used in the studies.

The qualitative literature elucidates diverse methods employed for analyzing qualitative data, among which three techniques have been singled out for prevalent utilization: "content analysis", "thematic analysis", and "grounded theory" (see Charmaz, 2006; Miles et al., 2014; Patton, 2014). Table 15 presents the breakdown of the qualitative data analysis methods and their reporting transparency.

We found that thematic analysis ($n = 162$, 53.28%) was the most frequent qualitative data analysis method reported in the articles. Across all 162 studies examined, AL authors who utilized thematic analysis provided detailed and thorough information regarding the method, even if the thematic analysis was implicitly identified. Roughly 50% of the studies explicitly indicated their

https://doi.org/10.1017/S0261444823000332 Published online by Cambridge University Press

Table 13. The rank-ordered combination of quantitative and qualitative data sources in MMR studies

Data sources[a]	*N* (%)
"*Close-ended questionnaires*, *interviews*"	66 (21.7)
"*Close-ended questionnaires*, *open-ended questionnaires*"	16 (5.26)
"*Close-ended questionnaires, interviews, open-ended questionnaires*"	16 (5.26)
"*Tests*, *interviews*"	14 (4.60)
"*Texts*, *texts*"	13 (4.27)
"*Close-ended questionnaires*, *tests, interviews*"	12 (3.94)
"*Close-ended questionnaires*, *interviews*, *observational field notes*"	8 (2.63)
"*Close-ended questionnaires*, *tests*, *interviews*, *open-ended questionnaires*"	5 (1.64)
"*Close-ended questionnaires*, *interviews*, *reflective journals*"	4 (1.31)
"*Close-ended questionnaires*, *interviews*, *stimulated recall protocols*"	4 (1.31)
"*Close-ended questionnaires*, *reflective journals*"	4 (1.31)
"*Tests*, *interviews*, *observational field notes*"	4 (1.31)
"*Tests*, *open-ended questionnaires*"	4 (1.31)
"*Texts*, *interviews*"	4 (1.31)

[a]Those data sources with frequencies less than four are not reported here.

Table 14. Statistical tests used in the MMR studies

Statistical tests of significance	*N* (%)
t-tests	95 (25)
Pearson correlation	52 (13.68)
One-way ANOVA	40 (10.52)
Chi-square test	26 (6.84)
Other nonparametric	50 (13.15)
Simple/multiple regression	22 (5.78)
One-way ANCOVA	13 (3.42)
Factorial ANOVA	13 (3.42)
Other types of correlation	3 (0.79)
Factor analysis	21 (5.52)
MANOVA	14 (3.68)
Rash analysis	9 (2.36)
Structural equation modelling	7 (1.84)
Mixed-effect modeling	5 (1.31)
Cluster analysis	5 (1.31)
MANCOVA	4 (1.05)
Discriminant function analysis	1 (0.26)
Total	380 (100)

https://doi.org/10.1017/S0261444823000332 Published online by Cambridge University Press

Table 15. Qualitative data analysis methods used in the MMR studies

	TA[a]	CA	GT	Others
	N (%)	*N* (%)	*N* (%)	*N* (%)
Explicitly stated	73 (45.07)	31 (28.71)	23 (76.66)	4 (100)
Implicitly stated	89 (54.93)	77 (71.29)	7 (23.34)	0 (0)
Total	162 (53.28)	108 (35.52)	30 (9.86)	4 (100)

[a]Thematic analysis (TA); Content analysis (CA); Grounded theory (GT).

utilization of thematic analysis, elaborating on the details of the thematic qualitative analysis method with explicit references. For example, the qualitative analysis method Schwartz and Gorbatt (2017) employed for analyzing teachers' mediation strategies encompasses a thematic analysis of "the observed video-recorded discourses, field notes, and the semi-structured interviews with the teachers" (p. 149). This exemplary study, akin to the majority of the articles, explicitly referenced the key article on thematic analysis – "Using thematic analysis in psychology" – co-authored by Braun and Clarke (2006).

The second most frequent qualitative data analysis method was content analysis ($n = 108$, 35.52%). However, a small portion of the studies ($n = 31$, 28.7%) explicitly indicated content analysis. Since content analysis can be used differently by different researchers, it is critical that researchers explain their approach to content analysis. For example, Sampson (2012) examined how possible self-images, socially constructed possible self-images, and language-learning motivation co-varied in an EFL university context. Qualitatively, he reported the use of qualitative content analysis in this way: "The document was collected before the following class and underwent qualitative content analysis, searching for recurring themes mentioned across the texts" (p. 321).

Grounded theory emerged as the third most frequent method of analysis. Notably, this method stood out for its explicit and transparent method of data analysis. For example, Kim et al. (2019) utilized grounded theory to investigate the performance of L2 learners on a writing test in different test modes. As they highlighted, "Data from observation notes were analyzed by two independent research team members using a grounded theory approach (Creswell & Creswell, 2017) to identify emerging patterns of children's writing behavior across the three test modes" (p. 492). From an innovative MMR perspective, as outlined by Riazi (2016), this qualitative data analysis method is significant as it can pave the way for the rise of new insights in AL through mixed-methods grounded theory (MM-GT). This innovative mixed-methods design, as Johnson et al. (2010, p. 65) argue, "works well in connecting theory generation with theory testing, linking theory and practice, and linking general/nomological description/explanation with idiographic understandings of the human world". As such, when addressing complex and multilayered research questions, MM-GT can potentially facilitate the generation of nuanced and rich insights. Accordingly, it enables the projection of substantive L2 theories that are firmly grounded in empirical data and evidence.

The last category, "others", represents four other techniques for analyzing qualitative data reported in four articles. Two articles (Murray et al., 2020; Omidian et al., 2018) employed discourse analysis as the main method for analyzing qualitative data, and two articles (MacIntyre & Legatto, 2011; Tian & Low, 2012) used vertical analysis. For example, MacIntyre and Legatto (2011), adopting an idiodynamic method, examined learners' fluctuations in willingness to communicate that happen moment-to-moment. Qualitatively, as outlined in the study, they utilized a vertical analysis of learners' response patterns per individual respondent. As they put it, "[i]n the present study, the major research focus is on vertical analysis and therefore is aligned epistemologically with qualitative research" (p. 153).

https://doi.org/10.1017/S0261444823000332 Published online by Cambridge University Press

4.5 RQ5: Integration at the interpretation level and making meta-inferences

To examine the data integration and analysis at the interpretation level, we analyzed each article's discussion and conclusion sections to locate signs of integration and meta-inferences. We followed the model proposed by Fetters et al. (2013), in which they cataloged integration at the interpretation level into three categories. The first category, integration through narrative, deals with a description of the qualitative and quantitative findings in the discussion or conclusion of a study. The second category, integration through data transformation, addresses converting one data set into the other (quantitative into qualitative and vice versa). The third one, integrating through joint displays, refers to when "researchers integrate the data by bringing the data together through visual means to draw new insights beyond the information gained from the separate quantitative and qualitative results" (Fetters et al., 2013, p. 2143). Our analyses revealed that in 250 (82.2%) articles, the integration of quantitative and qualitative data and analysis was reported. However, AL MMR researchers preferred narrative integration ($n = 248$, 99.2%) over joint display ($n = 2$, 0.66%). The data transformation type was not implemented in the corpus.

The final inferences (meta-inference) made in the MMR studies are related to the initial purposes for which the quantitative and qualitative methods were mixed. As such, and to provide an objective presentation of meta-inferences, we examined each study considering the purpose for implementation to trace the actual use of quantitative-qualitative integration. The motive behind such analysis was to see (in)consistency between the initially stated MMR purpose and its final reporting. Hence, we reanalyzed each article's discussion and conclusion sections to ensure consistency. That is the researchers' initial purpose and the inferences they made at the end using quantitative and qualitative data and analysis. Our results revealed that, except for the initiation purpose, there was a systematic consistency between the initially stated MMR purpose and final integration and inference. As Table 16 presents, we found an 88% consistency between the initially stated purpose and final meta-inference of triangulation, 86.6% for complementarity, 74% for development, and 91.6% for expansion.

Table 17 presents some excerpts from the papers and illustrates the congruence between the initial purpose and the conclusions made. The code following the quotes (e.g., JSLW2) shows the journal name and the article number.

Although we found a high consistency between the stated MMR purpose and final inferences, we found some interesting cases. These cases are the articles in which there is a discrepancy between the stated purpose and conclusions. As Table 18 presents, for example, when a study had an initiation purpose, it triangulated quantitative and qualitative data and analysis. That is, the initial initiation purpose ended up with a triangulation inference. This pattern (initial initiation leading to triangulation) was prevalent in five studies. Other patterns are presented in Table 18.

5. Discussion

All the evidence attests to the fact that more and more AL researchers use MMR to investigate language-related issues. In the introduction section, we introduced some specific sources that have

Table 16. MMR purpose and its implementation in the MMR studies

MMR initial purpose–Final conclusion	*N* (%)
Triangulation–Triangulation	125 (88.02%)[a]
Complementarity–Complementarity	71 (86.6%)[b]
Expansion–Expansion	44 (91.6%)[c]
Development–Development	20 (74.07%)[d]
Initiation–Initiation	0

Note. We compared the initially stated triangulation purpose with the conclusions made in the studies ($n = 142$). The same holds for all the other purposes (see Table 6).
[a]Out of 142 cases; [b]Out of 82 cases; [c]Out of 27 cases; [d]Out of 48 cases.

https://doi.org/10.1017/S0261444823000332 Published online by Cambridge University Press

Table 17. Excerpts taken from the papers regarding the match between the initial purpose and conclusions (emphasis added)

Initial MMR purpose–Final conclusion	Excerpts taken from the introduction and method sections of the MMR studies	Excerpts taken from the discussion and conclusion sections of the MMR studies
Triangulation–Triangulation	"This MIXED METHODS TRIANGULATION study examines these questions in university L2 academic writing classes through a quantitative text-based analysis of academic essay exams, student questionnaires, and teacher and student interviews."	"The DATA FROM THE STUDENT QUESTIONNAIRES AND INTERVIEW indicate that the students in this study were aware of their teachers' primary focus on accuracy as an assessment criterion. Students could enumerate particular areas of weaknesses that their teachers had previously identified in their writing using the evaluation grid." (JSLW2[a])
	"This corpus-based, mixed-methods study fills in a gap by COMPARING SYNCHRONOUS AND ASYNCHRONOUS TEF as well as uptake."	"TRIANGULATED, the findings of our study suggest TEF that involves the combined use of asynchronous Word comments and synchronous text-based chats is beneficial in ESL composition, and that teachers and students perceive TEF positively." (JSLW8)
	"This combination of quantitative and qualitative methods allowed us, BASED ON A SINGLE DATASET, TO TRIANGULATE INFORMATION about L2 writers' thought processes during pauses and revisions (stimulated recall), real-time text production behaviors (keystroke logging), as well as viewing behaviors including reading during pauses and before revisions (eye-tracking)."	"Importantly, however, THROUGH THE TRIANGULATION of keystroke logging, eye-tracking, and stimulated recall data, we provided evidence for these patterns based on a single dataset in the current study, allowing for drawing more VALID INFERENCES about the processes underlying pausing behaviors." (SSLA2)
	"Each of these data sources allowed me to gain a fuller picture of the participants' experiences through METHODOLOGICAL TRIANGULATION."	"As revealed in the study, the findings are based ON SELF-EFFICACY SCALES, REFLECTIVE DIARIES, AND COURSE EVALUATION FORMS RESULTING IN TRIANGULATION. QUANTITATIVE DATA FROM SELF-EFFICACY SCALES WERE IN AGREEMENT WITH QUALITATIVE DATA FROM REFLECTIVE DIARIES AND COURSE EVALUATION FORMS." (System8)
	"The present study also adopted a mixed-methods approach employing KEYSTROKE LOGGING, EYE-TRACKING, AND STIMULATED RECALL. Specifically, we investigated L2 writing processes across independent versus integrated tasks, hoping that THE TRIANGULATION of methods would afford deeper insights into writing processes across these task types."	"TRIANGULATING THESE FINDINGS, we may infer that, as expected, during stage 1 participants focused on reading the source text and/or notes. In stages 2 to 4, they primarily engaged in text construction involving both higher- and lower-order writing processes." (SLR1)
Complementarity–Complementarity	"… ensuring child English language tests produce reliable and valid scores REQUIRES MORE THAN A QUANTITATIVE EVALUATION of the tests' outcomes; applied linguists and language testers NEED ROBUST QUALITATIVE ANALYSES TO UNDERSTAND WHY certain tasks are appropriate for measuring a child's language skills and why other tasks are not."	"We found OUR QUALITATIVE DATA PROVIDED INSIGHTS THAT OUR QUANTITATIVE DATA could not provide, namely, what the children thought while being tested… why children selected the responses that they did, and whether those selections stemmed from a true measurement of the construct or from construct-irrelevant issues." (TQ10)
	"The study focuses on the efficacy of applying CL analysis of English conditionals to L2 instruction, USING A MIXED-METHODS DESIGN AND SUPPLEMENTING CL APPROACH with task-supported language teaching."	"COMBINED WITH QUALITATIVE FINDINGS (see Jacobsen 2015), these results suggest that the focus on usage-based patterns and meaning indeed made a difference in the subjects' performance." (AL3)
	"Inspired by a CDST perspective, we adopted a longitudinal Q methodological design to HOLISTICALLY CAPTURE the dynamics of the learners' motivational profiles. We intended	"By longitudinally tracking a group of multilingual learners and capturing the crystalized patterns of their motivational trajectories as encapsulated by their

(*Continued*)

https://doi.org/10.1017/S0261444823000332 Published online by Cambridge University Press

Table 17. (*Continued*)

Initial MMR purpose–Final conclusion	Excerpts taken from the introduction and method sections of the MMR studies	Excerpts taken from the discussion and conclusion sections of the MMR studies
	to SHOW THE DIFFERENT ROUTES OF DEVELOPMENT of multilingual motivation."	motivational profiles, the study has contributed to CDST-inspired empirical research on L2 MOTIVATION BY ADDING INSIGHTS ABOUT THE HOLISTIC, dynamic, and relational multilingual self-system." (MLJ17)
	"The present study has thus highlighted the usefulness of employing COMPLEMENTARY quantitative methods for examining connections between L1 and L2 fluency, WHILE ALSO UNDERSCORING THE IMPORTANCE OF COMPLEMENTING QUANTITATIVE METHODS with a qualitative analysis."	"To conclude, THE MIXED METHODS APPROACH, COMPLEMENTING THE QUANTITATIVE APPROACH WITH THE QUALITATIVE ANALYSIS, SUGGESTS THAT THE LACK OF GROUP level connections does not necessarily indicate that L1 and L2 stalling mechanisms are not connected for individual speakers; the connections may simply be idiosyncratic." (MLJ10)
	"In line with this view, data should be collected AT DIFFERENT INTERVALS OF A PROGRAMME because there are ups and downs in students' learning. It would also be interesting to find out if re-motivation is A PHENOMENON, and IDENTIFY POSSIBLE IMPLICATIONS FOR TEACHING AND LEARNING. In response to the call for MORE RESEARCH TO BETTER UNDERSTAND the factors that influence tertiary learners' persistence…"	"Drawing on the results of this study, the students' attendance records and the SURVEY AND INTERVIEW DATA COMPLEMENT ONE ANOTHER. They have identified similar factors that encourage and hinder students' participation, continuation, and persistence in the SDLLS. A majority of the factors are external, and these factors, in turn, determine students' learning processes despite their desire to improve their English." (System33)
	"A mixed-methods approach was then adopted through which the PATTERNS OF VARIATION reflected in the use of the identified sequences were examined both QUANTITATIVELY AND QUALITATIVELY."	"We ADOPTED A MIXED-METHODS APPROACH THROUGH which the identified patterns of variation were examined both quantitatively and qualitatively. THE COMPLEMENTARY STRENGTH of this hybrid methodology enabled us to uncover a series of differences in focus and practice between hard and soft science fields. Below, we first discuss the patterns of variation in light of our quantitative analyses and then specify some of the major differences uncovered in our qualitative inquiry." (JEAP5)
Development–Development	Adopting mixed methods of data collection and analysis, the current study MODELS the "perceived value of compulsory English language education" in a sample of 138 undergraduate non-language majors of Japanese nationality at a national university in Japan."	"The theoretical basis of the tested MODEL WAS THE QUALITATIVE DATA produced in phase one of the current study THAT INFORMED THE CREATION OF THREE THEMATIC CATEGORIES. THE TESTED MODEL SHOWN IN FIGURE 1 SUPPORTS this position as does the factor analysis shown in Table 2 where the international friendship orientation accounted for 41.5% of the total variance explained." (JMMD1)
	"Using mixed methods approach, SURVEY ITEMS WERE GENERATED first using qualitative methods (e.g. expert interviews and document analysis). To do this, WE CREATED THE TPACK-EFL SURVEY to collect data on preservice teachers' self-assessment of the seven types of knowledge represented within TPACK."	"It is possible that the seven-factor structure in this study was due to the use of the INSTRUMENT DEVELOPMENT MODEL proposed by Creswell and Plano Clark (2007). The Creswell and Plano Clark model provides a rigorous and robust process for instrument development through the use of multiple methods. It is possible that DRAWING FROM IN-DEPTH QUALITATIVE DATA TO DEVELOP TPACK ITEMS associated with EFL led to a survey that was designed well conceptually." (CALL7)

[a]AL: *Applied Linguistics*; CALL: *Computer Assisted Language Learning*; JEAP: *Journal of English for Academic Purposes*; JSLW: *Journal of Second Language Writing*; TQ: *TESOL Quarterly*; MLJ: *Modern Language Journal*; JMMD: *Journal of Multilingual and Multicultural Development*

https://doi.org/10.1017/S0261444823000332 Published online by Cambridge University Press

Table 18. Initial MMR purpose and final integration patterns

Variants	Purpose–Practice	*N* (%)	Examples for initial purpose	Examples for practice
Other triangulation variants[a]	Triangulation–Complementarity	13 (8.78)	"The RATIONALE for utilizing a mixed-methodology in this study was to be able to TRIANGULATE THE INFORMATION."	The initial purpose was triangulation; the conclusion was complementarity
	Triangulation–Expansion	10 (6.75)	"Interview data were used to ELABORATE ON TRENDS IN THE QUESTIONNAIRE RESULTS."	The initial purpose was triangulation; the conclusion was an expansion
Other complementarity variants[b]	Complementarity–Triangulation	8 (9.75)	"… FROM DIFFERENT PERSPECTIVES, A SEQUENTIAL MIXED-METHODS DESIGN known as the two-phase design approach (Creswell, 1994) was adopted. It consists of a quantitative phase (Phase I) and a separate qualitative phase (Phase II)."	The initial purpose was complementarity; the conclusion was triangulation
	Complementarity–Expansion	3 (3.65)	The authors sought to analyze the content of each article. This process was assisted by the use of the MAXQDA 12 QUALITATIVE ANALYSIS software package. FINALLY, AND WHERE APPROPRIATE, chi-square analyses were conducted to determine any statistically significant relationships between and within the features and themes of the newspaper articles.	The initial purpose was complementarity; the conclusion was an expansion
Other development variants[c]	Development–Expansion	7 (22.6)	"This study aimed to VALIDATE the effectiveness of the mobile learning platform."	The initial purpose was development; the conclusion was an expansion
	Development–Triangulation	4 (12.9)	"The researchers DEVELOPED a series of questions based on the literature review and the questionnaire results to explore teachers' attitudes towards teaching using an online mode to discover . . ."	The initial purpose was development; the conclusion was triangulation
Other expansion variants[d]	Expansion–Triangulation	2 (4.16)	"To answer the research questions, we conducted A MIXED-METHOD STUDY IN THREE DATA COLLECTION PHASES: (1) collect compositions written by Primary 6 students; (2) recruit teachers to grade the translated scripts and the English compositions; (3) interview teachers and inquire about their general impressions of the scripts and their attitudes towards GT as a pedagogical tool."	The initial purpose was expansion; the conclusion was triangulation
	Expansion–Complementarity	2 (4.16)	"All Zoom sessions were TRANSCRIBED and then coded using researcher-created categories via NVivo 11 software. Zoom recordings as well as pre- and post-surveys before and after the 6-week treatment implementation and weekly surveys with Likert-scale questions WERE ANALYZED through IBMSPSS Statistics 25.0."	The initial purpose was expansion; the conclusion was complementarity

(*Continued*)

https://doi.org/10.1017/S0261444823000332 Published online by Cambridge University Press

Table 18. (*Continued*)

Variants	Purpose–Practice	*N* (%)	Examples for initial purpose	Examples for practice
Other initiation variants[e]	Initiation–Triangulation	5 (100)	"The intervention was NOVEL both for the learners in the present study and researchers. Thus, the nature of the study was EXPLORATORY calling for a design that provides "the most INFORMATIVE, COMPLETE, BALANCED, AND USEFUL RESEARCH RESULTS. Second, research concerning learner beliefs has EMPLOYED DIFFERENT APPROACHES (i.e., quantitative and qualitative) and REPORTED CONFLICTING FINDINGS."	The initial purpose was initiation; the conclusion was triangulation

[a]Out of 148 cases; [b]Out of 82 cases; [c]Out of 48 cases; [d]Out of 31 cases; [e]Out of 5 cases.

discussed the increasing use of MMR in AL. To broaden that perspective, we did a quick search in an online library on 26 September 2022 for January 2011–26 September 2022. The keywords used were MMR AND AL OR language teaching and learning, and the search returned 720,687 sources. The breakdown of the sources was 453,233 articles, 140,511 theses, 55,967 book chapters, 18,135 newspaper articles, and 15,652 reviews. Searching Google Scholar on the same day for the same period but using the keywords of applied linguistics with mixed methods research only returned 17,600 sources. These ad hoc search outcomes align with more bibliometric findings, such as those of Zhang (2020). Zhang reported that "while both quantitative and qualitative have remained relatively stable, the keyword mixed methods has made the biggest jump among all the keywords" (p. 216). All these records provide evidence for the popularity of MMR in AL. Yet, one of the best pieces of evidence showing the growing use of MMR in AL is the 304 studies published in top-tier AL journals over the last decade and reviewed in this article. The 304 MMR studies covered 15 different strands of AL, as presented in Table 5, showing its traction in AL strands. As stated in the results section, "language", "writing", "online", "foreign", "feedback", "perceptions", "motivation", and their collocations were the most frequent ones in the dataset. These key terms cover most of the research areas addressed by AAAL and BAAL.

If we accept that MMR is the trending methodology in AL for investigating various topic areas and language frontiers, the next step would be examining the quality and methodological transparency in MMR studies. Many sources now explain and discuss different aspects of MMR, from philosophical underpinnings to purposes for mixing methods and methods elements. Accordingly, MMR researchers are expected to allude to those sources when planning, conducting, and explaining their studies. A transparent explanation of the purpose, design structure, and method elements will add to the study's rigor. We discuss each aspect of the quality and transparency framework.

Our first quality criterion was documenting MMR studies using relevant literature. As presented in the results section, of 304 articles, more than half of them (n = 179, 58.9%) did not reference the MMR sources to feature their MMR study design. Surprisingly, these studies explained MMR without any reference to MMR literature. This lack of evidence-based explanation points to information literacy, which is the ability to "locate, analyze, evaluate, and use information resources to support research and learning" (International Society for Technology in Education, 2008). Only 125 articles (41.1%) documented their use of MMR by providing references to MMR literature. There were 226 MMR references cited in those 125 articles, which was promising. However, multidisciplinary MMR sources were the dominant category with 170 instances (75.22%), followed by discipline-specific (AL MMR sources) with 40 (17.7%) instances, and general research methodology sources with 16 instances (7%). As MMR is growing in AL, we expect more citations and discussions of discipline-specific MMR sources to crystalize the use of methodology and open new spaces for innovative use of the methodology. To reach that point, information literacy should be a key consideration: knowing

https://doi.org/10.1017/S0261444823000332 Published online by Cambridge University Press

Table 19. Purpose and methodological transparency observed in the MMR studies

	MMR purpose	MMR design	MMR priority or emphasis	MMR sampling
	N (%)	*N* (%)	*N* (%)	*N* (%)
Implicitly stated	201 (66)	214 (70)	252 (83)	304 (100)
Explicitly stated	103 (34)	90 (30)	52 (17)	0 (0)

about the MMR resources and using valid ones to document studies. Also, scientometric analysis of MMR sources (e.g., Amini Farsani et al., 2021) would greatly help future AL MMR researchers.

To discuss other quality and transparency aspects, we summarize our findings in Table 19.

As Table 19 presents, a low level of transparency was observed in the MMR studies. As shown in Table 19, only 103 articles (34%) explicitly stated a purpose for mixing methods, 90 (30%) explicitly explained design structures, 52 (17%) explicitly discussed the priority or emphasis they gave to either of the methods, and none of the studies explained their MMR sampling. AL MMR researchers may thus want to consult sources that discuss the purposes for mixing quantitative and qualitative methods (e.g., Bryman, 2006; Creswell & Plano Clark, 2011; Greene et al., 1989; Tashakkori & Teddlie, 2009). Stating the purpose for which the methods are mixed can help readers and future researchers to better understand one of the critical features of MMR studies. The initially stated purpose of MMR studies can also help the final integration of inferences and produce a meta-inference. An issue that echoes what Schoonenboom and Johnson (2017, p. 110) have referred to as "heightened knowledge and validity". There are situations where an inconsistency might arise between the initially stated purpose and the data analysis outcomes. As Mathison (1988, p. 17) said, "we end up with data that occasionally converge, but frequently are inconsistent and even contradictory". In such situations, researchers may argue that different constructs are involved in the quantitative and qualitative strands, and thus a shift to complementarity or initiation purposes would be plausible. They may even need to do more work trying to reconcile any contradictions in their findings.

We also found that among the five purposes for mixing the two methods suggested by Greene et al. (1989), studies with triangulation purposes were predominant. About half of the studies ($n = 148$, 48.7%) either explicitly or implicitly aimed to mix quantitative and qualitative methods to corroborate or triangulate results from one method with those from another. This finding is in line with Tazik et al. (2020), who found that convergent designs with triangulation purposes were used more frequently in 7,525 articles they reviewed in 10 AL journals from 1986–2015. Other purposes like complementarity, development, and initiation are yet to be used in AL research more systematically. Mixing methods to achieve other purposes can expand and enhance the scope of research questions. For example, complementarity will allow for conceptualizing research problems as multidimensional and addressing different dimensions using different data and analysis procedures. Researchers sometimes face conflicting results on the same topic due to methodological approaches and data collection and analysis methods. Such conflicting results can lead researchers to initiate an MMR study to collect and analyze quantitative and qualitative data in one study and provide a coherent perspective on the research topic. Very few ($n = 5$, 1.6%) studies followed an "initiation" purpose. Also, MMR studies might be conducted with a development purpose that can enable AL researchers in their attempts to develop scales, questionnaires, and rubrics in areas where such tools do not exist or are scarce.

Regarding design structure, it was interesting to see that AL researchers used both types of MMR designs, that is, core and complex designs. However, as presented in the results section, core designs ($n = 214$, 71.38%) were used significantly more than complex designs ($n = 77$, 25.32%). This finding implies that there is still room for AL researchers to use more complex designs to investigate language-related issues. There is a relationship between MMR purpose and design structure. As such, complex designs could be used to mix quantitative and qualitative methods for complementarity, development, or initiation purposes.

https://doi.org/10.1017/S0261444823000332 Published online by Cambridge University Press

On the other hand, among the core MMR designs, "explanatory sequential designs" ($n = 110$, 36%) were predominant in the 304 articles. This design is used when qualitative data and analysis are employed to explain the quantitative findings. Depending on the purpose of the study and research questions, other MMR designs can be considered in AL. For example, advanced designs like hybrid designs in which one or both strands of the MMR will include more than one data collection and analysis phase when researching more complex problems. As presented in Table 8, a few AL researchers used hybrid designs. One of the advantages of MMR is its design flexibility. While there are already design structures AL researchers might follow, there are many occasions AL researchers can be creative and design innovative studies. To reach that point, however, we need full familiarity with all MMR purposes and design structures, given AL research is yet to appreciate and apply MMR to investigate different research problems.

The findings regarding epistemological orientation revealed that AL MMR researchers might not be fully aware of the epistemological orientations in MMR. As presented in the findings section, only eight studies (i.e., 2.6%) of the 304 articles revealed the researchers' epistemological orientation and its role in designing their MMR studies. As Riazi (2017) rightly put it, "Mixed methods researchers' familiarity with different underlying worldviews, on the one hand, and their own stances as researchers, on the other, could enable them to better articulate their theoretical position when investigating more complex problems" (p. 32). It is, therefore, imperative that future AL MMR researchers familiarize themselves with different philosophical underpinnings (pragmatism, transformative, dialectical pluralism, and critical realism), and optimize their study within one of these epistemologies. Our third criterion was general issues or method features, including sampling, data sources, and data analysis. Our findings showed that MMR sampling transparency was utterly absent in the dataset. None of the reviewed studies reported the MMR sampling procedure, while they explained quantitative and qualitative samples. This is another crucial issue when designing and reporting an MMR study. Since two methods (quantitative and qualitative) are used in MMR, it is imperative to explain how the samples in the two strands are related to each other. Onwuegbuzie and Collins (2007) provide a helpful typology of sampling in MMR that AL researchers can use. MMR sampling has implications for purpose and the inferences made from mixing the two methods, so it goes beyond the sample per se. Our findings show that while the sample size in each strand was adequate, using the method's norms, the MMR sampling procedure was not transparent. This lack of transparency in MMR sampling would affect the researchers' meaningful meta-inferences. According to Collins (2010), Collins and Onwuegbuzie (2013), Collins et al. (2006), Collins, Onwuegbuzie, and Jiao (2007), and Onwuegbuzie and Collins (2014), the role of sampling in mixed methods research is still an underdeveloped area, so it needs careful attention in AL research, given its role in interpretive (in)consistency.

Another aspect of the method quality criterion is data sources and how clearly they are explained. Johnson and Turner (2003) discussed two data-source profiles. INTRAMETHOD data collection comprises gathering quantitative and qualitative data using one data collection procedure (e.g., questionnaires with both closed- and open-ended items). On the other hand, INTERMETHOD data collection is used when researchers collect quantitative and qualitative data using two or more data collection procedures (e.g., using both tests and interviews). The use of these two categories and the specific data collection methods depends on the purpose, research questions, and design structures of the MMR study. Intermethod data collection procedures were dominant in the articles we reviewed. More specifically, as presented in Table 13, the use of closed-ended questionnaires and an interview, representing the intermethod category, was more dominant in the MMR studies than in other combination patterns.

Still, another crucial point is that the data collection procedures and data sources must align with the purpose and design structure of the studies. As was noted before, triangulation purpose and explanatory sequential design were predominant. This purpose and design lend themselves to the questionnaire and interview data sources. The least frequent data sources were stimulated recall and think-aloud protocols. While they have functional differences, as Loewen and Plonsky (2016, p. 190) asserted, these two data sources provide researchers with a "window into the learners' attention and internal cognitive processing". As discussed earlier, different data sources could be used depending on the purpose, research questions, and design of the MMR study. However, this can be

https://doi.org/10.1017/S0261444823000332 Published online by Cambridge University Press

challenging for researchers, given the push for publishing quickly. Researchers may avoid such designs due to the level of commitment needed to data collection and analysis.

Regarding data analysis, we did not find serious problems with quantitative and qualitative data analysis. In the quantitative phase of the MMR studies, parametric and non-parametric tests were used, although parametric tests were dominant ($n = 295$, 77.63%). However, basic inferential statistics (t-test or ANOVA and their non-parametric equivalents) were prevalent in the MMR studies. Our findings showed that less than 10% of the studies used advanced statistics (like factor analysis, regression, or structural equation modeling). This trend is consistent with other findings (e.g., Khani & Tazik, 2019; Lazaraton, 2000, 2005). New data collection techniques like eye tracking may provide MMR researchers with more comprehensive data sets. As Gass (2009) put it, AL is becoming "more sophisticated in its use of statistics" (p. 19), and so future MMR studies may consider the use of such advanced procedures, too. As Brown (2015) argued, with advanced statistics, we may expand the types and number of variables, measure the variables more precisely, avoid the problem of multiple comparisons, boost statistical power, broaden AL research perspectives, and address multiple levels of analysis. These advanced analyses would help shape sound and solid meta-inferences in MMR studies.

Also, in the qualitative phase, data were analyzed using a variety of qualitative data analyses. We found that thematic analysis was the most frequently used qualitative data analysis ($n = 162$, 53.28%). However, we must remind ourselves that thematic analysis is a broad approach and encompasses different coding, categorization, and theme development approaches. Elaboration on the nuances of how the thematic analysis was conducted (e.g., inductive or deductive and the coding procedures with exemplifications) will add to the studies' rigor, transparency, and replicability. For example, there are two approaches to coding in thematic analysis. The two approaches are inductive, where the coding categories are developed from the data, and deductive, where available a priori categories are applied to the data. It will be beneficial to clarify the approach to coding and the thematic analysis and thus add to the rigor and transparency of the qualitative data analysis. One final crucial point is how the outcomes or inferences of individual data analysis procedures from quantitative and qualitative data analyses are integrated into broader conclusions or meta-inferences. This issue will be discussed below.

Our next quality criterion was data integration. A significant point to be discussed here is the integration procedure. Our analyses revealed that in 250 (82.2%) articles, the integration of quantitative and qualitative data and analysis was reported, which is excellent and another indication of progress in using MMR. However, AL MMR researchers preferred narrative integration ($n = 248$, 99.2%) over joint display ($n = 2$, 0.66%). Creswell et al. (2011) are among those who have discussed procedures for the systematic integration of the two datasets and analyses. One of the procedures they discuss is joint display using tables or figures that can display quantitative and qualitative data and analysis and how they could be related. Johnson et al. (2019) also developed and introduced a process for integrating and presenting qualitative and quantitative findings in a joint display. The process uses a four-stage technique called the Pillar Integration Process (PIP) by the authors. Although the integration process is developed for the health sciences, we believe it can also be used in AL. The four stages in the PIP include listing, matching, checking, and pillar building, which are completed once the quantitative and qualitative analyses are conducted separately. The key elements of this process revolve around data integration: presenting quantitative and qualitative data to get a unified message; visualizing and representing the integrated data through pillar charts or matrices; locating areas of convergence and divergence between quantitative and qualitative inferences; and synthesizing and elucidating potential areas that warrant further research.

The authors provide two examples of PIP that AL researchers may use to develop integration ideas for their studies. Guetterman et al. (2021) also provide a methodological review of visuals in joint displays that incorporate graphs, charts, maps, images, and other visuals in 33 MMR articles. They offer helpful recommendations to MMR researchers on how they might use joint visual displays. Since the process is creative and study-based, there are no fixed ways of jointly displaying quantitative and qualitative data and analyses. AL MMR researchers may consider different opportunities to communicate complex information in more streamlined and understandable ways.

https://doi.org/10.1017/S0261444823000332 Published online by Cambridge University Press

The final quality criterion was related to the integration of the two strands and constructing a meta-inference that alludes to how the initially stated purpose is actualized. In other words, the final meta-inference (integrating inferences from the quantitative and qualitative results) would allow researchers to show how the initial goal of mixing methods is achieved. We examined each study's purpose and traced the integration of the results in favor of a meta-inference. The integration of the results is usually presented in the discussion and conclusion sections of the articles. As Table 16 presented, we found a high consistency between the initially stated purpose and the final meta-inference made. For example, where the researchers had initially stated triangulation as their purpose, they integrated quantitative and qualitative results to convey a triangulation meta-inference in their report's discussion and/or conclusion section.

The high consistency between the initial goal and final integration patterns is promising and adds to the MMR studies' design, implementation, and reporting transparency. This is a breakthrough compared with the lack of meta-inference in AL MMR studies that Hashemi and Babaii (2013) reported. Only five studies did not meet this quality and transparency criterion. As presented in Table 17, for example, when a study stated an initiation purpose, it triangulated quantitative and qualitative data and analysis. There are two potential explanations here. The first is that the borderline between different purposes is not well recognized and is not thus attended to. This could be a possible reason, as discussed earlier, given the lack of an explicitly stated purpose. The second explanation is that a researcher may start with a particular goal (e.g., triangulation), but in the process and based on the data and analysis, they may find divergent rather than convergent results. In such cases, the researcher needs to discuss the situation explicitly. Schoonenboom et al. (2018) also discuss this issue and the possibility of coming up with different purposes in an MMR study.

As stated earlier, our six-pronged quality and transparency criteria were derived from six frameworks discussed in the literature review and aligned with Hirose and Creswell's (2023) framework. We cannot claim that the framework is comprehensive. However, it helped us to review and discuss the published MMR studies in AL journals and shed some light on rigor and transparency issues. According to Denzin (2008), "[w]e cannot afford to fight with one another … We need to find new strategic and tactical ways to work with one another. … We must expand the size of our tent; indeed, we need a bigger tent!" (p. 321). It seems that MMR is in a transitional quality movement. MMR quality and transparency criteria (e.g., those presented in Table 2) have been developed to help different stakeholders improve the design and implementation of MMR studies. As Creamer (2018) and Hirose and Creswell (2023) reported, their identified quality criteria may serve this purpose well, while others cannot be ignored. Such reorientation is highly in line with what Joseph Juran (1988) referred to as fit-for-use or conformance-to-requirements quality. As such, considering the specificity of research contexts and the fluid and dynamic nature of MMR, there is a need for more quality discussion in AL, which can inform and be informed by other disciplines (see Amini Farsani et al., 2021 for further discussion).

We want to end the discussion by reiterating that MMR methodological transparency can promote research quality and rigor. As Creamer (2018) observed, methodological transparency in MMR can promote "replication by providing explicit detail about the steps taken to complete data collection and data analysis as well as by delineating the link between results and the source of data" (p. 154). Through design transparency, AL MMR researchers can delineate the purpose for which they mixed the methods. They can also clarify the priority or emphasis put on either method, the type of MMR design (e.g., concurrent, sequential, or hybrid), the integration of quantitative and qualitative data and analysis, and the conclusions (meta-inference) reached in the study. Methodological transparency can also help systematic reviews of empirical studies.

6. Limitations of the review

We acknowledge some limitations regarding the systematic review we conducted. The first limitation is that we selected and included in our review only those articles that explicitly stated the use of MMR in their research report. We acknowledge that there could be other studies with an MMR approach without necessarily labelling themselves as such. However, the expectation is that AL researchers explicitly

https://doi.org/10.1017/S0261444823000332 Published online by Cambridge University Press

use MMR terminology if they are mixing quantitative and qualitative data and analysis. This can contribute to methodological discussion and transparency.

Our second limitation is that we used the classic Greene et al. (1989) framework to identify and discuss MMR purposes. There are some new publications that could be used to discuss purposes for mixing methods. One of these sources is that of Fetters (2020). Future AL researchers may use this and other recently published sources to discuss different aspects of MMR, including the purposes for mixing methods.

The third limitation of this review is that we did not examine research questions in the reviewed articles. Research questions in MMR studies can be critically analyzed to see how MMR researchers intend to integrate their quantitative and qualitative data and analysis. Future review studies may focus on research questions in MMR studies.

The fourth limitation is that we did not address ethics in the data set. Considering the recent emphasis on ethics in AL research (see Yaw et al., 2023), it is timely to address ethical issues in MMR and how these issues might be exercised.

Finally, future reviewers of MMR studies may plan for a sequential explanatory MMR review. That is, once they analyze the articles quantitatively, they recruit a sample of the articles' authors to interview about the major issues. Such an MMR review can add value to the findings and discussion.

7. Conclusions

The current review is a follow-up of the state-of-the-art review published in *Language Teaching* in 2014. Since the data included in the 2014 review covered published articles up to 2011, we decided to cover the whole decade of 2011–2020 (inclusive). As stated in the introduction and discussion sections of this article, MMR studies in AL are gaining momentum and growing at a high rate. As a case in point, we could locate 304 MMR studies in 20 top-tier AL journals for review and analysis. We may call the current situation the massification and methodological branding of MMR in AL. Like other methodologies' trajectories, it is time to focus on quality and transparency issues and help future MMR researchers consider rigor and transparency in their design, implementation, and reporting of studies. The current review was thus initiated to contribute to the quality discussion of MMR in AL.

In the concluding section of our review, we would like to list a number of recommendations future MMR researchers should consider.

- Given the dynamic nature of AL, AL researchers have initiated addressing multilayered and multidimensional problems using MMR. Most of these studies are, however, oriented to SLA. Future studies may address research issues related to ideology, culture, corpus linguistics, technology, and AI, using big data, in addition to SLA-related issues. These areas are highly advantageous if prospective researchers address relevant issues through the lens of principled and innovative MMR.
- The studies we reviewed have significantly contributed to the AL content areas. With the increasing use of MMR to investigate research questions, it is critical that AL researchers also contribute to emerging methodologies like MMR. There are increasing publications on MMR in AL, as listed in the reference section. There is also to be a distinct section on MMR in the second edition of the *Encyclopedia of Applied Linguistics*. Therefore, future AL researchers can be more strategic in contextualizing their MMR with AL methodological sources. Based on our findings, the bulk of references to MMR was multidisciplinary. Nevertheless, in the pursuit of theorizing research methods in AL (see McKinley, 2020), journal gatekeepers may provide more space for MMR methodological-based articles. Such articles can discuss different aspects of MMR as well as issues and challenges involved in this methodology. The publication of the *Journal of Research Methods in Applied Linguistics* and the research section of *Studies in Second Language Acquisition* has helped bridge this void. It is also possible to establish a new MMR SIG in AL in the American Association for Applied Linguistics (AAAL) or the British Association of Applied Linguistics (BAAL).
- The transparency of the purpose for which methods were mixed was not adequately exercised in the studies we reviewed. Accordingly, future AL researchers who intend to use MMR may bridge

https://doi.org/10.1017/S0261444823000332 Published online by Cambridge University Press

this gap by more carefully examining MMR purposes and practices. In this review, we followed Greene et al.'s (1989) widely-used purpose typology, and our findings highlighted the frequent use of triangulation and complementarity. Future AL researchers may use more recent sources (e.g., Fetters, 2020) to develop and advance context-specific MMR purposes.

- Our results revealed the dominance of core MMR designs in the reviewed studies. However, the rise of complex MMR designs was also evident. We thus recommend that future researchers apply more complex MMR designs to address multilayered and multidimensional language-related problems, as stated in the first recommendation. This helps AL researchers to stay away from default designs to employ more innovative MMR designs, leading to novel insights and fresh perspectives in the field. To achieve this goal, discipline-specific journal editors may also explicitly set MMR criteria and request researchers to abide by those criteria.
- The representation of philosophical underpinnings of MMR was not evident in the studies reviewed. There were only eight studies that explicitly stated their epistemological orientation (four pragmatic, two transformative, one dialectical pluralism, and one critical realism). Prospective AL MMR researchers are urged to consider different MMR philosophical underpinnings and frame their studies accordingly. MMR, with different philosophical underpinnings, can advantage AL researchers by enabling them to address less frequently addressed topics. For example, recently applied linguists have charted the issues on political and economic concepts and related ideologies such as neoliberalism. One possibility would be to investigate topics dealing with neoliberalism from the lens of transformative MMR. Using MMR with other philosophical underpinnings, such as transformative, critical realism, or dialectical pluralism, will add value to the currently predominant pragmatic MMR approaches.
- MMR sampling procedure explanations were entirely absent in the studies reviewed. While there was information about quantitative and qualitative sampling procedures, the relationship between the two (MMR sampling) was lacking. In MMR studies, it is crucial to explain how the samples of the two strands are related. While the explanation of MMR sampling was problematic, data collection procedures and analysis were relatively adequately discussed. Future AL researchers need to consider the relationship between the two sampling procedures if they intend to embark on MMR studies.
- Finally, while the strength of MMR is highly contingent on the strength of each phase, the integration of quantitative and qualitative data adds rigor and quality to the MMR study. Of the three modes of data integration (narrative, data transformation, and joint display), the narrative was the dominant mode, while only a few studies used joint display. So, while narrative data integration has its own merits, future AL MMR researchers may consider joint display and data transformation as potential procedures for contributing to the rigor of MMR studies.

Supplementary material. To view supplementary material for this article, please visit the journal website.

Notes

[1] We used IF (based on JCR report) and Cite Score (based on Scopus report) of the journals reported in 2018.

[2] The unit of analysis is individuals. We left out those with other units of analysis such as documents, images, films, etc.

References

Alise, M., & Teddlie, C. (2010). A continuation of the paradigm wars? Prevalence rates of methodological approaches across the social/behavioral sciences. *Journal of Mixed Methods Research*, *4*(2), 103–126. doi:10.1177/1558689809360805

Alwaleedi, M. A., Gillies, R. M., & Obaidul Hamid, M. (2019). Collaborative writing in Arabic as a second language (ASL) classrooms: A mixed-method study. *Language, Culture and Curriculum*, *32*(2), 157–172. doi:10.1080/07908318.2018.1521422

American Education Research Association. (2006). Standards for reporting on empirical social science research in AERA publications. *Educational Researcher*, *35*(6), 33–40. doi:10.3102/0013189X035006033

American Psychological Association. (2019). *Publication manual of the American Psychological Association, 2019*. American Psychological Association.

https://doi.org/10.1017/S0261444823000332 Published online by Cambridge University Press

Amini Farsani, M., Babaii, E., Beikmohammadi, M., & Babaii Farsani, M. (2022). Mixed-methods research proficiency for applied linguists: A PLS-path modelling approach. *Quality & Quantity*, *56*(5), 3337–3362. doi:10.1007/s11135-021-01268-7

Amini Farsani, M., Jamali, H. R., Beikmohammadi, M., Daneshvar Ghorbani, B., & Soleimani, L. (2021). Methodological orientations, academic citations, and scientific collaboration in applied linguistics: What do research synthesis and bibliometrics indicate? *System*, *100*, 102547. doi:10.1016/j.system.2021.102547

Amini Farsani, M., & Jamali, H. R. (in press). Collaboration network of applied linguistics research articles with different methodological orientations. *Studies in Second Language Learning and Teaching*.

Amini Farsani, M., & Mohammadi, V. (2022). Mixed-methods research in an EFL Context: A quality assessment perspective. *Journal of Modern Research in English Language Studies*, *9*(2), 99–122. doi:10.21203/rs.3.rs-44890/v1

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77–101. doi:10.1191/1478088706qp063oa

Brown, J. D. (2014). *Mixed methods research for TESOL*. Edinburgh University Press.

Brown, J. D. (2015). Why bother with learner advanced quantitative methods in L2 research? In L. Plonsky (Ed.), *Advancing quantitative methods in second language research* (pp. 9–20). Routledge.

Bryman, A. (2006). Integrating quantitative and qualitative research: How is it done? *Qualitative Research*, *6*(1), 97–113. doi:10.1177/1468794106058877

Byrnes, H. (2013). Notes from the editor. *Modern Language Journal*, *97*(4), 825–827. doi:10.1111/j.1540-4781.2013.12051.x

Cai, S., & Zhu, W. (2012). The impact of an online learning community project on university Chinese as a foreign language students' motivation. *Foreign Language Annals*, *45*(3), 307–329. doi:10.1111/j.1944-9720.2012.01204.x

Caracelli, V. J., & Greene, J. C. (1993). Data analysis strategies for mixed-method evaluation designs. *Educational Evaluation and Policy Analysis*, *15*(2), 195–207. doi:10.3102/01623737015002195

Cerezo, L. (2016). Type and amount of input-based practice in CALI: The revelations of a triangulated research design. *Language Learning & Technology*, *20*(1), 100–123.

Chapelle, C. A. (Ed.) (forthcoming). *Encyclopedia of applied linguistics*. (2nd ed.). Wiley Blackwell.

Charmaz, K. (2006). *Constructing grounded theory: A practical guide through qualitative analysis*. Sage.

Cohen, L., Manion, L., & Morrison, K. (2002). *Research methods in education*. Routledge.

Collins, K. M. (2010). Advanced sampling designs in mixed research. In A. Tashakkori & C. Teddlie (Eds.), *The handbook of mixed methods in social and behavioral research* (pp. 353–377). Sage.

Collins, K. M., Onwuegbuzie, A. J., & Sutton, I. L. (2006). A model incorporating the rationale and purpose for conducting mixed-methods research in special education and beyond. *Learning Disabilities: A Contemporary Journal*, *4*(1), 67–100.

Collins, K. M. T., & Onwuegbuzie, A. J. (2013). Establishing interpretive consistency when mixing approaches: Role of sampling designs in evaluations. *New Directions for Evaluation*, *2013*(138), 85–95. doi:10.1002/ev.20060

Collins, K. M. T., Onwuegbuzie, A. J., & Jiao, Q. G. (2007). A mixed methods investigation of mixed methods sampling designs in social and health science research. *Journal of Mixed Methods Research*, *1*(3), 267–294. doi:10.1177/1558689807299526

Cooper, H. (2016). *Research synthesis and meta-analysis: A step-by-step approach*. Sage.

Copland, F., Garton, S., & Burns, A. (2014). Challenges in teaching English to young learners: Global perspectives and local realities. *TESOL Quarterly*, *48*(4), 738–762. doi:10.1002/tesq.148

Creamer, E. G. (2018). *An introduction to fully integrated mixed methods research*. Sage.

Creswell, J., & Plano Clark, V. (2007). *Designing and conducting mixed methods research*. Sage.

Creswell, J. W. (1994). *Research design: Qualitative and quantitative approaches*. Sage.

Creswell, J. W. (2009). Mapping the field of mixed methods research. *Journal of Mixed Methods Research*, *3*(2), 95–108. doi:10.1177/1558689808330883

Creswell, J. W. (2015). *A concise introduction to mixed methods research*. Sage.

Creswell, J. W., & Creswell, J. D. (2017). *Research design: Qualitative, quantitative, and mixed methods approaches*. Sage.

Creswell, J. W., Klassen, A. C., Plano Clark, V. L., & Smith, K. C. (2011). *Best practices for mixed methods research in the health sciences*. Retrieved from https://www2.jabsom.hawaii.edu/native/docs/tsudocs/Best_Practices_for_Mixed_Methods_Research_Aug2011.pdf

Creswell, J. W., & Plano Clark, V. (2011). *Designing and conducting mixed methods research*. Sage.

Creswell, J. W., & Plano Clark, V. L. (2018). *Designing and conducting mixed methods research*. Sage.

Creswell, J. W., & Tashakkori, A. (2007). Editorial: Developing publishable mixed methods manuscripts. *Journal of Mixed Methods Research*, *1*(2), 107–111. doi:10.1177/1558689806298644

Crosby, P. B. (1986). *Running things—The art of making things happen*. American Society for Quality Control.

Denzin, N. K. (2008). The new paradigm dialogs and qualitative inquiry. *International Journal of Qualitative Studies in Education*, *21*(4), 315–325. doi:10.1177/1077800410364608

Dörnyei, Z. (2007). *Research methods in applied linguistics*. Oxford University Press.

Fàbregues, S., & Molina-Azorin, J. F. (2017). Addressing quality in mixed methods research: A review and recommendations for a future agenda. *Quality and Quantity*, *51*(6), 2847–2863. doi:10.1007/s11135-016-0449-4

Fàbregues, S., Molina-Azorin, J. F., & Fetters, M. D. (2021). Virtual special issue on "quality in mixed methods research". *Journal of Mixed Methods Research*, *15*(2), 146–151. doi:10.1177/15586898211001974

Fàbregues, S., Paré, M. H., & Meneses, J. (2019). Operationalizing and conceptualizing quality in mixed methods research: A multiple case study of the disciplines of education, nursing, psychology, and sociology. *Journal of Mixed Methods Research*, *13*(4), 424–445. doi:10.1177/1558689817751774

Fetters, M. D. (2020). *The mixed methods research workbook: Activities for designing, implementing, and publishing projects*. Sage.

Fetters, M. D., Curry, L. A., & Creswell, J. W. (2013). Achieving integration in mixed methods designs-principles and practices. *Health Services Research*, *48*(6), 2134–2156. doi:10.1111/1475-6773.12117

Fetters, M. D., & Freshwater, D. (2015). The 1+1=3 integration challenge. [Editorial]. *Journal of Mixed Methods Research*, *9* (2), 115–117. doi:10.1177/1558689815581222

Fetters, M. D., & Molina-Azorin, J. F. (2019a). A checklist of mixed methods elements in a submission for advancing the methodology of mixed methods research. *Journal of Mixed Methods Research*, *13*(4), 414–423. doi:10.1177/1558689819875832

Fetters, M. D., & Molina-Azorin, J. F. (2019b). New requirements to include the methodological contribution in articles published in the *Journal of Mixed Methods Research*. *Journal of Mixed Methods Research*, *13*(2), 138–142. doi:10.1177/1558689819834753

Fraschini, N., & Caruso, M. (2019). "I can see myself…." A Q methodology study on self-vision of Korean language learners. *System*, *87*, 102147. doi:10.1016/j.system.2019.102147

Gánem-Gutiérrez, G. A., & Gilmore, A. (2018). Tracking the real-time evolution of a writing event: Second language writers at different proficiency levels. *Language Learning*, *68*(2), 469–506. doi:10.1111/lang.12280

Gass, S. (2009). A survey of SLA research. In W. Ritchie & T. Bhatia (Eds.), *Handbook of second language acquisition* (pp. 3–28). Emerald.

Gilmore, A., & Ganem-Gutierrez, G. A. (2020). Investigating complexity in L2 writing with mixed methods approaches. In G. G. Fogal & M. Verspoor (Eds.), *Complex dynamic systems theory and L2 writing development* (pp. 183–206). John Benjamins.

Greene, J. C., Caracelli, V. J., & Graham, W. F. (1989). Toward a conceptual framework for mixed-method evaluation designs. *Educational Evaluation and Policy Analysis*, *11*(3), 255–274. doi:10.3102/01623737011003255

Guetterman, T. C. (2015). *The development, design, and test of a self-assessment instrument of mixed methods research proficiency* [Unpublished Ph.D. Dissertation]. University of Nebraska.

Guetterman, T. C. (2017). What distinguishes a novice from an expert mixed methods researcher? *Quality and Quantity*, *51* (1), 377–398. doi:10.1007/s11135-016-0310-9

Guetterman, T. C., Fàbregues, S., & Sakakibara, R. (2021). Visuals in joint displays to represent integration in mixed methods research: A methodological review. *Methods in Psychology*, *5*, 100080. doi:10.1016/j.metip.2021.100080

Han, J., & Hiver, P. (2018). Genre-based L2 writing instruction and writing-specific psychological factors: The dynamics of change. *Journal of Second Language Writing*, *40*, 44–59. doi:10.1016/j.jslw.2018.03.001

Harrison, R. L., Reilly, T. M., & Creswell, J. W. (2020). Methodological rigor in mixed methods: An application in management studies. *Journal of Mixed Methods Research*, *14*(4), 473–495. doi:10.1177/1558689819900585

Harvey, L., & Green, D. (1993). Defining quality. *Assessment & Evaluation in Higher Education*, *18*(1), 9–34. doi:10.1080/0260293930180102

Hashemi, M. R. (2020). Expanding the scope of mixed methods research in applied linguistics. In J. McKinley & H. Rose (Eds.), *The Routledge handbook of research methods in applied linguistics* (pp. 39–51). Routledge.

Hashemi, M. R., & Babaii, E. (2013). Mixed methods research: toward new research designs in applied linguistics. *Modern Language Journal*, *97*(4), 828–852. doi:10.1111/j.1540-4781.2013.12049.x

Hashemi, M. R., & Gohari Moghaddam, I. (2019). A mixed methods genre analysis of the discussion section of MMR articles in applied linguistics. *Journal of Mixed Methods Research*, *13*(2), 242–260. doi:10.1177/1558689816674626

Heyvaert, M., Maes, B., & Onghena, P. (2013). Mixed methods research synthesis: Definition framework, and potential. *Quality and Quantity*, *47*(2), 659–676. doi:10.1007/s11135-011-9538-6

Hirose, M. (2018). *Evaluating a behavioral community support program for families with youth with developmental disorders: A mixed methods approach* [Unpublished Ph.D. dissertation]. Kwansei Guikin University.

Hirose, M., & Creswell, J. W. (2023). Applying core quality criteria of mixed methods research to an empirical study. *Journal of Mixed Methods Research*, *17*(1), 12–28. doi:10.1177/15586898221086346

International Society for Technology in Education (ISTE). (2008). *Standards for teachers*. Retrieved from www.iste.org/standards/ISTE-standards/ standards-for-teachers.

Ivankova, N. V., & Greer, J. L. (2015). Mixed methods research and analysis. In B. Paltridge & A. Phakiti (Eds.), *Research methods in applied linguistics: A practical guide* (pp. 73–91). Bloomsbury.

Jacobsen, N. (2015). A cognitive linguistic analysis of English conditionals in English for Academic Purposes (EAP) instruction: Implications from Sociocultural Theory. In K. Masuda, C. Arnett, & A. Labarca (Eds.) , *Cognitive linguistics and sociocultural theory: Applications for second and foreign language teaching* (pp. 103–126). De Gruyter Mouton.

Jang, E. E., Wagner, M., & Park, G. (2014). Mixed methods research in language testing and assessment. *Annual Review of Applied Linguistics*, *34*, 123–153. doi:10.1017/S0267190514000063

Johnson, B. (2006). Editorial: Note from the guest editor. *Research in the Schools*, *13*(1), v–vi.

Johnson, B., & Turner, L. A. (2003). Data collection strategies in mixed methods research. In A. Tashakkori & C. Teddlie (Eds.), *Handbook of mixed methods in social and behavioral research* (pp. 297–319). Sage.

Johnson, R. B., & Christensen, L. (2019). *Educational research: Quantitative, qualitative, and mixed methods research*. (7th ed.). Sage.

Johnson, R. B., Onwuegbuzie, A. J., & Turner, L. A. (2007). Toward a definition of mixed methods research. *Journal of Mixed Methods Research*, *1*(2), 112–133. doi:10.1177/1558689806298224

Johnson, R. E., Grove, A. L., & Clarke, A. (2019). Pillar integration process: A joint display technique to integrate data in mixed methods research. *Journal of Mixed Methods Research*, *13*(3), 301–320. doi:10.1177/1558689817743108

Juran, J. M. (1988). *Juran on planning for quality*. Collier Macmillan.

Khany, R., & Tazik, K. (2019). Levels of statistical use in applied linguistics research articles: From 1986 to 2015. *Journal of Quantitative Linguistics*, *26*(1), 48–65. doi:10.1080/09296174.2017.1421498

Kim, A. A., Lee, S., Chapman, M., & Wilmes, C. (2019). The effects of administration and response modes on grade 1–2 students' writing performance. *TESOL Quarterly*, *53*(2), 482–513. doi:10.1002/tesq.495

Kwan, N., & Dunworth, K. (2016). English as a lingua franca communication between domestic helpers and employers in Hong Kong: A study of pragmatic strategies. *English for Specific Purposes*, *43*, 13–24. doi:10.1016/j.esp.2016.02.001

Lazaraton, A. (2000). Current trends in research methodology and statistics in applied linguistics. *TESOL Quarterly*, *34*(1), 175–181. doi:10.2307/3588103

Lazaraton, A. (2005). Quantitative research methods. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning* (pp. 109–224). Erlbaum.

Lei, L., & Liu, D. (2019). Research trends in applied linguistics from 2005 to 2016: A bibliometric analysis and its implications. *Applied Linguistics*, *40*(3), 540–561. doi:10.1093/applin/amy003

Loewen, S., & Plonsky, L. (2016). *An A-Z of applied linguistics research methods*. Palgrave.

Long, M. H. (2005). *Second language needs analysis*. Cambridge University Press.

Macaro, E., & Lee, J. H. (2013). Teacher language background, codeswitching, and English-only instruction: Does age make a difference to learners' attitudes?. *TESOL Quarterly*, *47*(4), 717–742. doi:10.1002/tesq.74

MacIntyre, P. D., & Legatto, J. J. (2011). A dynamic system approach to willingness to communicate: Developing an idiodynamic method to capture rapidly changing affect. *Applied Linguistics*, *32*(2), 149–171. doi:10.1093/applin/amq037

Mackey, A., & Bryfonski, L. (2018). Mixed methodology. In A. Phakiti, P. De Costa, L. Plonsky, & S. Starfield (Eds.), *The Palgrave handbook of applied linguistics research methodology* (pp. 103–121). Palgrave.

Marsden, E., Mackey, A., & Plonsky, L. (2016). Breadth and depth: The IRIS repository. In A. Mackey & E. Marsden (Eds.), *Advancing methodology and practice: The IRIS repository of instruments for research into second languages* (pp. 1–21). Routledge.

Mathison, S. (1988). Why triangulate? *Educational Researcher*, *17*(2), 13–17. doi:10.3102/0013189X017002013

McKinley, J. (2020). Theorizing research methods in the golden age of applied linguistics research. In J. McKinley & H. Rose (Eds.), *The Routledge handbook of research methods in applied linguistics* (pp. 1–12). Routledge.

Mertens, D. M. (2011). Publishing mixed methods research. *Journal of Mixed Methods Research*, *5*(1), 3–6. doi:10.1177/1558689810390217

Miles, M. B., Huberman, A. M., & Saldaña, J. (2014). *Qualitative data analysis: A methods sourcebook*. Sage.

Mirhosseini, S. A. (2018). Mixed methods research in TESOL: Procedures combined or epistemology confused? *TESOL Quarterly*, *52*(2), 468–478. doi:10.1002/tesq.427

Moeller, A. J., & Theiler, J. (2014). Spoken Spanish language development at the high school level: A mixed-methods study. *Foreign Language Annals*, *47*(2), 210–240. doi:10.1111/flan.12085

Murray, L., Giralt, M., & Benini, S. (2020). Extending digital literacies: Proposing an agentive literacy to tackle the problems of distractive technologies in language learning. *ReCALL*, *32*(3), 250–271. doi:10.1017/S0958344020000130

Newman, I., & Ramlo, S. (2010). Using *Q*-methodology and *Q* factor analysis in mixed methods research. In A. Tashakkori & C. Teddlie (Eds.), *SAGE handbook of mixed methods in social & behavioral research* (pp. 505–530). Sage.

NIH, Office of Behavioral and Social Sciences. (2011). *Best practices for mixed methods research in the health sciences*. National Institutes of Health.

NIH, Office of Behavioral and Social Sciences. (2018). *Best practices for mixed methods research in the health sciences* (2nd ed.). National Institutes of Health.

Norris, J. M., & Ortega, L. (2006). The value and practice of research synthesis for language learning and teaching. In J. M. Norris & L. Ortega (Eds.), *Synthesizing research on language learning and teaching* (pp. 3–50). John Benjamins.

Omidian, T., Shahriari, H., & Siyanova-Chanturia, A. (2018). A cross-disciplinary investigation of multi-word expressions in the moves of research article abstracts. *Journal of English for Academic Purposes*, *36*, 1–14. doi:10.1016/j.jeap.2018.08.002

Onwuegbuzie, A. J., & Collins, K. M. T. (2014). Using Bronfenbrenner's ecological systems theory to enhance interpretive consistency in mixed research. *International Journal of Research in Education Methodology*, *5*(2), 651–661. doi:10.24297/ijrem.v5i2.3910

Onwuegbuzie, A. J., & Collins, K. M. T. (2007). A typology of mixed methods sampling designs in social science research. *The Qualitative Report*, *12*(2), 281–316. doi:10.46743/2160-3715/2007.1638

Onwuegbuzie, A. J., & Corrigan, J. A. (2014). Improving the quality of mixed research reports in the field of human resource development and beyond: A call for rigor as an ethical practice. *Human Resource Development Quarterly*, *25*(3), 273–299. doi:10.1002/hrdq.21197

Patton, M. Q. (2014). *Qualitative research & evaluation methods: Integrating theory and practice*. Sage.
Plano Clark, V. L., & Ivankova, N. V. (2015). *Mixed methods research: A guide to the field*. Sage.
Plonsky, L. (2013). Study quality in SLA: An assessment of designs, analyses, and reporting practices in quantitative L2 research. *Studies in Second Language Acquisition*, *35*(4), 655–687. doi:10.1017/S0272263113000399
Plonsky, L. (2014). Study quality in quantitative L2 research (1990-2010): A methodological synthesis and call for reform. *Modern Language Journal*, *98*(1), 450–470. doi:10.1111/j.1540-4781.2014.12058.x
Plonsky, L. (2017). Quantitative research methods. In S. Loewen & M. Sato (Eds.), *The Routledge handbook of instructed second language acquisition* (pp. 505–521). Taylor and Francis.
Plonsky, L., & Oswald, F. L. (2015). Meta-analyzing second language research. In L. Plonsky (Ed.), *Advancing quantitative methods in second language research* (pp. 106–128). Routledge.
Razfar, A., & Simon, J. (2011). Course-taking patterns of Latino ESL students: Mobility and mainstreaming in urban community colleges in the United States. *TESOL Quarterly*, *45*(4), 595–627. doi:10.5054/tq.2011.268060
Révész, A., Michel, M., & Lee, M. (2019). Exploring second language writers' pausing and revision behaviors: A mixed-methods study. *Studies in Second Language Acquisition*, *41*(3), 605–631. doi:10.1017/s027226311900024x
Riazi, A. M. (2016). Innovative mixed-methods research: Moving beyond design technicalities to epistemological and methodological realizations. *Applied Linguistics*, *37*(1), 33–49. doi:10.1093/applin/amv064
Riazi, A. M. (2017). *Mixed-methods research in language teaching and learning*. Equinox.
Riazi, A. M. (2018). Mixed methods approaches to studying second language writing. *The TESOL Encyclopedia of English Language Teaching*, 1–6. doi:10.1002/9781118784235.eelt0562
Riazi, A. M., & Candlin, C. N. (2014). Mixed-methods research in language teaching and learning: Opportunities, issues, and challenges. *Language Teaching*, *47*(2), 135–173. doi:10.1017/s0261444813000505
Riazi, A. M., Ghanbar, H., & Fazel, I. (2020). The contexts, the theoretical and methodological orientation of EAP research: Evidence from empirical articles published in the *Journal of English for Academic Purposes*. *Journal of English for Academic Purposes*, *48*, 100925. doi:10.1016/j.jeap.2020.100925
Riazi, A. M., Ghanbar, H., Marefat, H., & Fazel, I. (in press). Review and analysis of empirical articles published in *Tesol Quarterly* over its lifespan. *Studies in Second Language Learning and Teaching (SSLLT)*.
Riazi, A. M., Shi, L., & Haggerty, J. (2018). Analysis of the empirical research in the journal of second language writing at its 25th year (1992–2016). *Journal of Second Language Writing*, *41*, 41–54. doi:10.1016/j.jslw.2018.07.002
Sampson, R. (2012). The language-learning self, self-enhancement activities, and self-perceptual change. *Language Teaching Research*, *16*(3), 317–335. doi:10.1177/1362168812436898
Sato, M. (2013). Beliefs about peer interaction and peer corrective feedback: Efficacy of classroom intervention. *Modern Language Journal*, *97*(3), 611–633. doi:10.1111/j.1540-4781.2013.12035.x
Schoonenboom, J., & Johnson, R. B. (2017). How to construct a mixed methods research design. *Kolner Zeitschrift fur Soziologie und Sozialpsychologie*, *69*(Suppl 2), 107. doi:10.1007/s11577-017-0454-1
Schoonenboom, R., Johnson, R. B., & Froehlich, D. E. (2018). Combining multiple purposes of mixing within a mixed methods research design. *International Journal of Multiple Research Approaches*, *10*(1), 271–282. doi:10.29034/ijmra.v10n1a17
Schwartz, M., & Gorbatt, N. (2017). "There is no need for translation: She understands": Teachers' mediation strategies in a bilingual preschool classroom. *Modern Language Journal*, *101*(1), 143–162. doi:10.1111/modl.12384
Singleton, D., & Pfenninger, S. E. (2015). Insights from a mixed methods approach with respect to age and long-term instructed language learning. In J. Navracsics & S. Bátyi (Eds.), *First and second language: Interdisciplinary approaches* (pp. 11–21). Tinta Könyvkiadó.
Tahamtan, I., Afshar, A. S., & Ahmadzadeh, K. (2016). Factors affecting number of citations: A comprehensive review of the literature. *Scientometrics*, *107*(3), 1195–1225. doi:10.1007/s11192-016-1889-2
Tashakkori, A., & Teddlie, C. (2009). Integrating qualitative and quantitative approaches to research. In A. Tashakkori & C. Teddlie (Eds.), *Handbook of mixed methods in social and behavioral research* (pp. 283–318). Sage. doi:10.4135/9781483348858.n9
Tashakkori, A., & Teddlie, C. (2010). Putting the human back in "human research methodology": The researcher in mixed methods research. *Journal of Mixed Methods Research*, *4*(4), 271–277. doi:10.1177/1558689810382
Tazik, K., Khani, R., & Aliakbari, M. (2020). Trends of research methods in applied linguistics: The case of research articles published between 1986 and 2015. *Scientific Journal of Language Research*, *12*(36), 87–108.
Teddlie, C., & Tashakkori, A. (2006). A general typology of research designs featuring mixed methods. *Research in the Schools*, *13*(1), 12–28.
Tian, J., & Low, G. D. (2012). To what extent are postgraduate students from China prepared for academic writing needed on UK master's courses? *Language, Culture and Curriculum*, *25*(3), 299–319. doi:10.1080/07908318.2012.734313
Tsushima, R. (2015). Methodological diversity in language assessment research: The role of mixed methods in classroom-based language assessment studies. *International Journal of Qualitative Methods*, *14*(2), 104–121. doi:10.1177/160940691501400202
Van Kemenade, E., Pupius, M., & Hardjono, T. W. (2008). More value to defining quality. *Quality in Higher Education*, *14*(2), 175–185. doi:10.1080/13538320802278461

Watts, S., & Stenner, P. (2005). Doing Q-methodological research: Theory, method and interpretation. *Qualitative Research in Psychology*, *2*(1), 67–91. doi:10.4135/9781446251911

Yaw, K., Plonsky, L., Larsson, T., Sterling, S., & Kytö, M. (2023). Research ethics in applied linguistics. *Language Teaching*, 1–17. doi:10.1017/S0261444823000010

Zhang, X. (2020). A bibliometric analysis of second language acquisition between 1997 and 2018. *Studies in Second Language Acquisition*, *42*(1), 199–222. doi:10.1017/s0272263119000573

Zheng, Y., Lu, X., & Ren, W. (2019). Profiling Chinese university students' motivation to learn multiple languages. *Journal of Multilingual and Multicultural Development*, *40*(7), 590–604. doi:10.1080/01434632.2019.1571074

Zheng, Y., Lu, X., & Ren, W. (2020). Tracking the evolution of Chinese learners' multilingual motivation through a longitudinal Q methodology. *Modern Language Journal*, *104*(4), 781–803. doi:10.1111/modl.12672

A. Mehdi Riazi is a Professor of Applied Linguistics at the College of Humanities & Social Sciences of Hamad Bin Khalifa University, Qatar. He is also an honorary professor at Macquarie University, Australia. Riazi's areas of interest include research methodology (especially mixed methods research), second language writing, and testing and evaluation. He is the author of *The Routledge encyclopedia of research methods in applied linguistics* and *Mixed methods research in language teaching and learning* (Equinox). He is also the mixed methods research section editor of the second edition of the *Encyclopedia of applied linguistics* (Wiley).

Mohammad Amini Farsani is an Assistant Professor of applied linguistics at the Iran University of Science and Technology. Mohammad's areas of interest include the quality of research in applied linguistics, especially mixed methods research. He has published bibliometric reviews on different aspects of research in applied linguistics, aiming to advance the theoretical understanding of research methods in the field. He has published in different journals, including *System*, *Argumentation*, *Studies in Second Language Learning and Teaching*, *Language, Culture, and Curriculum*, *Quality and Quantity*, *Mixed Methods Research*, and *Writing & Pedagogy*.

Cite this article: Riazi, A. M., & Farsani, M. A. (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