Annual Review of Applied Linguistics (2024), **44**, 133–149
doi:10.1017/S0267190524000011

CAMBRIDGE UNIVERSITY PRESS

RESEARCH ARTICLE

The future of replication in applied linguistics: Toward a standard for replication studies

Kevin McManus

Pennsylvania State University, USA
Email: kmcmanus@psu.edu

Abstract

High-quality replication studies are widely understood to be critical to the growth and credibility of our discipline, as shown in commentaries and discussion since the 1970s, at least. Nevertheless, misunderstandings and limitations in the aims, designs, and reporting of replication research remain, thus reducing the usefulness and impact of replications. To address this issue and improve the rigor, quality, and conduct of replication studies in applied linguistics, a new standard for reporting replication studies that captures several critical features of replication research not discussed in current reporting guidelines is proposed. Importantly, this standard applies basic expectations in replication reporting so that outcomes can be better understood and evaluated. By applying this standard, replication studies will better meet their aims to confirm, consolidate, and advance knowledge and understanding within applied linguistics and second language research. In addition, readers will more easily understand how the replication study was carried out and be able to better evaluate the claims being made.

Keywords: replication; study quality; research methods; applied linguistics; reporting practices; second language acquisition

The quality, robustness, and credibility of our claims about language, its users and uses, and their underlying social and material conditions are established through *replication*, a research method that involves repeating a previous study's research design and methods with or without changes, collecting new data, and systematically comparing the previous study's findings with those from the new study (Nosek & Errington, 2020; Polio & Gass, 1997; Porte, 2012; Porte & McManus, 2019). This is why replication is widely understood to represent a powerful framework for confirming, consolidating, and advancing knowledge and understanding within empirical fields of study (Brandt et al., 2014; Isager et al., 2023; Long, 1993; Santos, 1989). Furthermore, replication allows us to better understand how a study's research data were collected, measured, analyzed, and interpreted while also providing a strong foundation for reconsidering, refining, extending, and sometimes limiting prior research findings.

© The Author(s), 2024. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted re-use, distribution, and reproduction in any medium, provided the original work is properly cited.

https://doi.org/10.1017/S0267190524000011 Published online by Cambridge University Press

Given the disciplinary benefits available from doing replication research, claims that replication studies are infrequent and poorly designed are troubling (Language Teaching Review Panel, 2008; Marsden et al., 2018; Polio, 2012; Porte & Richards, 2012), indicative of an uncritical approach to how a field accumulates evidence and builds theories. This is one reason why calls for replication are becoming more common (Nerenz, 2016; Plonsky, 2012; Plucker & Makel, 2021; Zwaan et al., 2018). Applied linguists have responded to this state of affairs by discussing and promoting replication studies in several ways, including books and reviews of the field (e.g., Porte, 2012; Yamashita & Neiriz, 2024), journal special issues (e.g., Atkinson, 2012; McManus, 2024; Tschichold, 2023), designated strands in journals for replication studies (e.g., *Language Teaching, Studies in Second Language Acquisition*), conference symposia and workshops (e.g., American Association for Applied Linguistics, 2009, 2020; International Symposium on Bilingualism, 2023; Symposium on Second Language Writing, 2010), and research funding earmarked for replication studies (e.g., Institute of Educational Sciences). Replication efforts like these are also in lockstep with significant and growing interest in Open Science initiatives that aim to make scholarship more open, inclusive, and transparent (e.g., Open Science badges, study preregistration, preprints, postprints; see Liu, 2023; Marsden, 2019; Plonsky, 2024). One field-specific initiative bringing together replication and Open Science is the *Instruments for Research into Languages* project (IRIS; Mackey, 2013; Mackey & Marsden, 2016; Marsden, 2013; Marsden & Mackey, 2014), established in 2011 with the explicit aim of supporting and facilitating replication. IRIS works toward this goal by providing a free, searchable, up- and downloadable collection of datasets and instruments, materials, and stimuli used to collect data for research into first, second, and foreign languages. Critically absent from this list of initiatives, however, is that very little guidance exists to support researchers in designing, conducting, and reporting replication studies, despite repeated calls to increase the frequency and improve the quality of replication studies (for exceptions, see Plonsky, 2015a; Porte, 2012; Porte & McManus, 2019).

While the field has developed resources to support the design, conduct, and reporting of applied linguistics research in general (e.g., Dörnyei, 2007; Mackey & Gass, 2022; McKinley & Rose, 2020), these general guidelines neglect critical aspects of replication studies that define their impact and usefulness. To be able to confront and revisit existing understanding with new evidence, replication reporting must include the following components, as a minimum: clear rationale for why the replication was carried out, full and transparent descriptions of differences and similarities in the design, methods, and results, and a framework for determining and evaluating the replicability of effect. Importantly, these components are neither encouraged nor discussed in general reporting guidelines for empirical research in applied linguistics. This is because the aims of most empirical studies (or extension studies) are different from those of replication studies. For one, extension studies aim to extend a current line of research in new ways (e.g., contexts, methods, populations), whereas a replication study aims to critically revisit one study to better understand its findings and impact on the discipline. This difference in aims impacts how the study is designed and reported, which explains why the information reported in replication studies is often piecemeal, incomplete, and difficult to understand and evaluate at best.

https://doi.org/10.1017/S0267190524000011 Published online by Cambridge University Press

To address this limitation and improve the rigor, quality, and conduct of replication research in applied linguistics, this paper proposes a new standard for reporting replication studies, consistent with increased interest in reporting standards, rigor, and study quality in the field more generally (see Awan et al., 2023; Norris et al., 2015; Plonsky, 2014; Riazi & Farsani, 2023). Indeed, a standard for replication studies is needed for multiple reasons. First, replication studies share many similarities with extension studies (e.g., motivation for the study, description, and justifications for research design and methods), but, importantly, replications require additional elements to be integrated into the design and reporting (e.g., justifications for study selection and variable modifications, comparative reporting in the design, methods, and results, determinations of replicability). Second, research indicates several known limitations in the design, conduct, and reporting of replication studies, including a lack of transparent labeling and inadequate reporting of between-study differences in the design, methods, and results (Marsden et al., 2018; McManus, 2022a; Polio, 2012; Polio & Gass, 1997). Third, most journals and professional associations in the field provide little to no guidance about how to report replication studies, contributing to their infrequency in the published literature. For these reasons, among others, a standard is needed that outlines basic expectations for the design, conduct, and reporting of replication research so that the field can better understand and evaluate the contribution of replication research to disciplinary growth. Without full and transparent information about why and how a replication study was designed and carried out, including how conclusions of replicability were reached, if at all, understanding the empirical results and impact of replication research is difficult, at best. Together, these reasons underscore why developing basic expectations for replication studies is critical to the development of the discipline.

In the sections that follow, guidelines for conducting and reporting replication studies from applied linguistics research methodology texts, journals, and specialized publications are critically reviewed to determine the types of guidance currently available. Using this review as a foundation, a new standard for replication studies that outlines basic expectations for reporting replication studies in applied linguistics is proposed for the first time in this article, which appears as part of the *Annual Review of Applied Linguistics* issue on "Research Methodology: Future Directions in the Field."

Guidelines for conducting and reporting replication studies in applied linguistics

Research methodology textbooks

General research methodology texts provide a key source of guidance for designing, conducting, and reporting empirical research in applied linguistics. However, other than some discussion about the importance, value, and ethics of replications, most of these texts provide very little guidance about how to conduct and report replication studies. Hatch and Farhady's (1982) text *Research design and statistics for applied linguistics*, for example, provides substantive discussion and guidance about approaches to research design, data analysis, and reporting empirical research but includes only passing remarks about replication research (for similar treatments, see Dörnyei, 2007; Plonsky, 2015b). Although some research methodology textbooks discuss the need

https://doi.org/10.1017/S0267190524000011 Published online by Cambridge University Press

and value of doing replication studies, prospective replication researchers are often provided with little-to-no guidance about conducting or reporting replications (e.g., Mackey & Gass, 2016; Rose et al., 2020). To illustrate, Mackey and Gass's (2022) influential text *Second language research: Methodology and design* reviews what replication research is, why replications are important, and, following Polio and Gass (1997), recommends including detailed appendices about the replication study's design and methods (e.g., data coding, instruments, biodata) due to space limitations in journals. However, in contrast to other research designs (e.g., quantitative, qualitative, mixed methods), no guidance is provided about design considerations for replications (although cross-referencing to other texts about replications is included).

Nevertheless, several edited volumes about general research methods in applied linguistics do provide more comprehensive treatments about replication studies (e.g., Gurzynski-Weiss & Kim, 2022; Mackey & Gass, 2011, 2023; Phakiti et al., 2018). In addition to reviewing ethical and conceptual issues in replication research, for example, Abbuhl's (2011) chapter in *Research methods in second language acquisition: A practical guide* includes practical advice about how to carry out replication. In a five-stage procedure, Abbuhl reviews how to select a study for replication, decide on the replication type, formulate research questions, and interpret and then write up the results (see also McManus, 2023).

In sum, even though general research methodology texts would be expected to represent a useful source of guidance for designing, conducting, and reporting replication studies, most of these texts either engage with replication very briefly or focus on the need and relevance of replications to the discipline. Some edited volumes about research methodology, however, do represent exceptions, likely because researchers with replication experience have contributed chapters to these volumes. Future research methodology textbooks must provide guidance about replication studies comparative to those provided for other research designs.

Journals

An additional source of guidance includes the author guidelines in academic journals, which specify the types of information to be included in manuscripts submitted to a particular journal. For example, the journal *Language Learning* publishes "rigorous, original empirical research as well as systematic critical literature reviews, innovative methodological contributions, and high-value replication research" (Language Learning, 2024). However, even though comprehensive author guidelines are provided, especially for the reporting of quantitative research findings, the only guidance directed to replication studies is that the term "replication" should be included in the study's abstract. In another prominent applied linguistics journal, *Studies in Second Language Acquisition*, although short descriptions of the different manuscript types are provided (e.g., Research Article, Replication Study, Methods Forum), no journal-specific guidelines are provided to authors for any manuscript type. Rather, all submissions to the journal "must conform to the requirements of the latest Publication Manual of the American Psychological Association" (Studies in Second Language Acquisition, 2024). Helpfully, though, these APA guidelines do include recommendations for replication studies under the category "reporting standards for special

https://doi.org/10.1017/S0267190524000011 Published online by Cambridge University Press

designs" (alongside reporting standards for quantitative meta-analysis as its own category). *TESOL Quarterly*'s author guidelines state explicitly that the journal publishes replication studies as *Brief Reports* (short empirical research reports no more than 3,400 words) and, similarly to *Studies in Second Language Acquisition*, recommends following the APA's journal reporting standards. *TESOL Quarterly* notes that "replications requiring full literature reviews or lengthy discussions should not be submitted as brief reports" (TESOL Quarterly, 2024). While not explicitly stated in the author guidelines, "full" replication studies should be submitted as *Articles* (personal communication, Charlene Polio, 02/09/2024). In addition, several other features are required depending on the study type (quantitative, qualitative, and mixed-methods studies but nothing for replication studies). For example, quantitative studies should include tabulated descriptive statistics, information regarding statistical significance (e.g., *p*-values, effect sizes), and data visualizations, while qualitative studies should include theoretical frameworks and descriptions of the data sources, including how the data were analyzed and coded. Another journal with a high impact factor in Linguistics is the *Modern Language Journal*, which, similarly to *Language Learning*, also provides comprehensive author guidance about, for example, using data-accountable graphics, reporting effect sizes, sharing key characteristics of instruments for questionnaire and interview-based studies, and correlational analyses. Although authors are directed toward the APA's journal reporting standards (as they are by *Studies in Second Language Acquisition*), no additional information on replication studies is included. Lastly, *Applied Linguistics* provides very little information about reporting expectations for any type of empirical study despite providing comprehensive information about formatting requirements and the presentation of references. The *Annual Review of Applied Linguistics* provides no guidance to authors in terms of manuscript preparation and reporting.

In sharp contrast, the journal *Language Teaching* provides extensive guidance for authors in the reporting of replication research. Compared to other journals in the field, this guidance is thorough, with different types of guidance for the study's introduction, methods and analysis, and results and discussion. These guidelines make clear, for example, that the introduction should explain the need for and the objectives of the replication study being reported. In the methods and analyses section, the reporting should clearly state what modifications to the initial study were made, if any. The results and discussion sections should comment on key comparisons with the initial study and any previous replications while also making suggestions for future replication research. *Language Teaching* also provides definitions for the different types of replication approaches (close, approximate, conceptual) so that authors can describe the replication approach taken. In this way, *Language Teaching* stands out in contrast to most other journals in the field by providing relatively comprehensive guidance for the reporting of replication studies.

In sum, given that the field has seen publications suggesting that replication studies are needed to improve the quality and rigor of its evidence base since the 1970s, it is surprising that so few journals provide any guidance at all. In some cases, authors are directed to external sources that might not always be relevant to applied linguistics research. The one journal that bucks this trend is *Language Teaching*, and it is telling that its editor (Graeme Porte) is a longtime advocate for replication research (see Porte, 2013, 2015; Porte & McManus, 2019).

https://doi.org/10.1017/S0267190524000011 Published online by Cambridge University Press

Specialized reporting guidelines

To complement this review of reporting guidelines provided by research methodology textbooks and academic journals, it is important to note that specialized resources about replication research also exist, even though they are small in number. Some of the earliest reflections on replication studies in applied linguistics can be found in commentaries by Tarone, Swain and Fathman (1976), Santos (1989), Long (1993), and Polio and Gass (1997). A unifying theme in these commentaries is the need to do replication research and the challenges that prevent replications from being carried out, especially inadequate or scant reporting in primary research. To address these challenges, Polio and Gass (1997) recommended (a) reserving space in journals for replication studies (see Valdman, 1993) and (b) allowing authors to submit detailed appendices about all aspects of the research design and methods, including instruments for data collection. Thus, early recommendations suggested that journals could play a pivotal role in supporting replication efforts by reserving space for replications and allowing authors to submit detailed appendices and supplementary materials.

In 2008, *Language Teaching* focused attention on the conduct and reporting of replication studies themselves through two critical contributions, the *Language Teaching Review Panel* (LTRP, 2008) and a Call for Papers for replication studies (Language Teaching, 2008). The LTRP outlined a series of questions and answers about the role and conduct of replications in applied linguistics, including why replications are rare and ways to design a replication study (e.g., exact, approximate, conceptual). In addition, the LTRP outlined recommendations for designing, conducting, and reporting replications in the field, thus representing an important turning point to support the conduct of high-quality replications that moved discussion beyond questions about the need, importance, and value of replication. A particularly important contribution of the LTRP was its guidance for authors on ways to identify a study for replication. In addition to the topic and findings needing to have continued relevance and importance in the field, for instance, other features for study selection included limitations in the design and methods of prior work, mixed findings, and the potential for the replication to contribute to theory-building. Echoing earlier commentary about reporting standards and the need to provide sufficient detail to allow for replication (see Polio & Gass, 1997 and Santos, 1989), the LTRP (2008) recommended that when space limitations apply, the authors should "focus on describing what changes they have made or how their studies differ […] the results of the original study should also be compared to those of the replication in the discussion section" (p. 8). The LTRP also recommended that authors make their data publicly available (e.g., CHILDES, see MacWhinney, 2000). *Language Teaching* (2008) then implemented the LTRP's recommendations by launching a new strand in the journal for replication studies and issued a Call for Papers for replication studies, noting the following requirements:

- The initial study should be important and significant, and its replication is needed.
- The replication should be conducted in a sound and thorough manner that sheds more light on the validity, reliability, and/or credibility of previous results.
- The replication should describe exactly what modifications were made and include enough detail to permit further replication.

https://doi.org/10.1017/S0267190524000011 Published online by Cambridge University Press

- The results and discussion/conclusion sections should comment comprehensively on key comparisons with the initial study and previous replications.

Through this work, *Language Teaching* has made and continues to make substantial contributions to the conduct of high-quality replications in applied linguistics.

An additional landmark contribution to replication efforts in applied linguistics and the development of reporting guidelines is Porte's (2012) edited volume, *Replication research in applied linguistics*, the first book dedicated to the topic of replication in the field with several chapters on the design, conduct, and reporting on replication research (e.g., Brown, 2012; Mackey, 2012; Polio, 2012). For example, Mackey (2012) reviews why, when, and how to do replication and discusses reporting limitations in primary research as an obstacle to replication, identifies several candidate topics ripe for replication, and uses a case-study approach to show how a replication project can be developed. Polio's (2012) chapter provides a historical and critical review of replication studies in the field. This important study is the first comprehensive account of replications in terms of topics, design features, and findings, followed by a series of recommendations to support and improve the conduct of replications (for an updated narrative review of replications in L2 research, see Marsden et al., 2018). Lastly, Brown's (2012) chapter presents a detailed account of the information that should be included in the write-up of a replication study. For example, Brown (2012) notes that the *Introduction* needs to set up the replication study by situating the initial study and any replications of it. In addition to establishing the importance and continued relevance of the topic, the introduction should indicate the need for and the aims of the replication. In the replication's *Method* section, comparisons with the initial study should be made throughout, including explanations for similarities and differences between the studies. In terms of results, Brown recommends that, alongside full reporting of the replication's analyses, "where necessary, the original study is described and its outcomes discussed in sufficient detail to understand similarities and differences" (p. 176). As can be seen, the contribution of this volume to replication efforts in applied linguistics is significant, providing much-needed support in many different areas of scholarship.

As previously noted, Marsden et al.'s (2018) systematic and narrative reviews of replication in L2 research can be seen as a revisiting and updating of Polio's (2012) initial review of replications in the field that reflects changes in quantitative methods and Open Science. As in Polio, Marsden et al.'s review also listed several recommendations for future replication studies. In line with the author guidelines in *Language Learning*, for instance, Marsden et al. (2018) recommend that replication studies "use more self-labeling with the term replication wherever appropriate" (p. 366). Also recommended is a principled, standard nomenclature for replications, which aims to account for the amount of intentional change introduced into the replication study design (e.g., direct replications when "no intentional change" is introduced). Marsden et al. (2018) also discuss the impact of "unclear reporting" in replication studies that can make it difficult for readers to ascertain the amount and magnitude of the similarities and differences between the initial and replication studies. To this end, authors of replications are recommended to "clarify the relationship with initial study (including descriptive statistics and effect sizes) and combine analyses where possible" (p. 367), which can be interpreted as using statistics to quantify between-study differences and

https://doi.org/10.1017/S0267190524000011 Published online by Cambridge University Press

similarities in results and integrating the initial study's data into the replication's analyses, when available. Lastly, in terms of recommendations for what studies should be replicated, Marsden et al. (2018) recommend "little or no top-down (e.g., journal or professional association) control [on studies that warrant replication], and researchers' agenda should drive what is replicated" (p. 367). Several additional recommendations are proposed by Marsden et al. (2018) to support replication efforts, including increasing the open availability of materials and data (see also Mackey & Marsden, 2016), supporting independent replications, multisite replication efforts, and changes to citation practices in which the replication study should be cited alongside the initial study.

Perhaps the most comprehensive set of guidelines for designing, conducting, and reporting replication studies in applied linguistics is Porte and McManus's (2019) volume, which provides a comprehensive walk-through of considerations and recommendations for designing, conducting, and reporting replication studies, including recommendations for how to select a study for replication, followed by detailed recommendations for how to carry out and write up a replication study using published articles as models. Porte and McManus recommend that decisions about what to replicate are informed by a critical and close reading of a study in the researcher's area of expertise, which may also be complemented by state-of-the-art reviews, calls for replication studies, as well as discussions in the future research and limitations sections of empirical research. In addition, Porte and McManus (2019) provide suggestions that can help pinpoint candidate studies in greater need for replication (e.g., surprising results, a study that has been frequently cited, questions about data handling, analysis, and/or other methodological shortcomings, as well as studies with high academic, theoretical, and societal impact).

In terms of designing and reporting a replication study, Porte and McManus (2019) use a case-study approach in which all practical aspects of the replication research process are described and considered in a step-by-step fashion. For example, the importance of justifying the need for the replication study and any variable modifications are outlined, followed by recommendations for how to communicate these aspects of the replication research process in the write-up. Indeed, the authors argue that these features constitute key features of replication studies. Also discussed are procedures for integrating the initial study's research question(s) into the replication, which ought to be the same given the shared aims of the studies. Procedures for doing this are outlined. Similarly, Porte and McManus (2019) show readers how to present the replication study's results in ways that highlight maximum comparability between the studies by presenting descriptive statistics from both studies in the write-up, thus pointing to a need for primary research to include descriptive statistics in the data analysis. A key point repeatedly made is that making systematic comparisons throughout the study is an essential and defining feature of the replication research process.

Taken together, the field of applied linguistics has developed several resources that can be used in designing, conducting, and reporting replication studies. While these resources have the potential to provide essential support for increasing the frequency and improving the quality of replication studies in the field, they are also relatively piecemeal and lack accessibility (or usability) and/or clarity. For example, few general research methodology texts and academic journals offer guidance about designing,

https://doi.org/10.1017/S0267190524000011 Published online by Cambridge University Press

conducting, and reporting replication studies. While several specialized resources exist, they are predominantly published as books that lack accessibility and usability due to high cost and length. In short, this situation is problematic for the discipline if it seeks to improve the rigor, quality, and conduct of replication studies.

That is not to say, of course, that commentaries about the state of replication in the field, calls for replication of specific studies and on particular topics, surveys of published replications, and book-length treatments about doing replication have not made essential and critical contributions to designing, conducting, and reporting replication studies in applied linguistics. However, replication has been an ongoing topic of scholarly discussion since the 1970s, at least, and there is no single resource that the field has been able to use to ensure standards in replication reporting. Furthermore, even though some journals direct researchers to recent APA recommendations, which do provide useful support in its journal article reporting standards, these are arguably less useable by the discipline at large, especially for researchers not doing experimental research (e.g., corpus linguists, discourse analysts).

In order to better support the field of applied linguistics in the design, conduct, and reporting of high-quality replication studies, there is, thus, a compelling need to develop and provide field-specific guidance that is comprehensive, accessible, and reflective of the types of research that applied linguists conduct. In doing so, the discipline will be better able to benefit from the potential of replication studies to meaningfully contribute to the growth and credibility of the discipline. In the next section, I begin this process by proposing a new standard for replication studies in applied linguistics, informed both by the replication resources reviewed in the previous sections as well as my own experiences conducting and evaluating replication studies as an applied linguistics scholar and research methodologist, including a recent guest-edited volume of replication studies in second language research (forthcoming in *Studies in Second Language Acquisition*).

A new standard for replication studies in applied linguistics

Before proposing the standard for replication studies in applied linguistics, I will describe what a standard is, what a standard for replication studies might look like, and what it might need to include. First and foremost, a standard represents an agreed way of doing something, such as making a product, managing a process, or delivering a service (National Institute of Standards and Technology, 2022). In this way, a standard is a text that outlines the requirements, specifications, guidelines, or characteristics that can be used to ensure that materials, products, and processes are fit for purpose. Thus, by proposing a standard for replication studies, my aim is to lay the groundwork to guide researchers in meeting and communicating basic expectations for doing replication research in applied linguistics.

In terms of specifying what a standard for replication studies might look like, my aim is to outline what the basic expectations for replication studies are. For example, in addition to describing and justifying why an initial study was selected for replication along with justifications for the replication approach and variable changes, it is important that replication studies provide information about the precise process by which data for replication study were collected, extracted, coded, analyzed, and

https://doi.org/10.1017/S0267190524000011 Published online by Cambridge University Press

reported. At the same time, each replication study needs to create a replication data set that includes all information necessary to understand, evaluate, and build upon the study, made publicly available using but not limited to public repositories (e.g., IRIS, Open Science Framework). For quantitative research, this might include data, specialized computer programs, analytic code/rubrics, coding books, and explanatory notes (e.g., a Read.Me file), all of which are made publicly available by the research team (see In'nami et al., 2022)

Taken together, then, my aim in proposing a standard for replication studies in applied linguistics is to outline and specify the types of information required so that readers can understand, evaluate, and build upon the replication study conducted without needing to request additional information from the authors. In this way, my proposal aims to support and improve the conduct and reporting of high-quality replication studies in the field of applied linguistics that can meaningfully contribute to the growth and credibility of the discipline.

In the next sections, several reporting recommendations for replication studies are proposed for the following key aspects of empirical studies in applied linguistics: title, abstract, introduction and background, research questions and/or hypotheses, method, results, discussion and conclusions, and data and materials. Table 1 summarizes these recommendations. As a reminder, this standard is not designed to replace general reporting guidelines currently used in the field (e.g., APA, MLA). The focus is on components specific to the conduct and reporting of replication studies that (a) researchers can use in planning for and reporting replication studies, (b) educators can use in teaching about reporting expectations for replications, and (c) journals can add to their author guidelines.

These proposed standards for replication studies in applied linguistics were devised following a critical review and synthesis of the following: Abbuhl (2011), Appelbaum et al. (2018), Atkinson (2012), Brandt et al. (2014), Brown (2012), Errington et al. (2021), In'nami et al. (2022), Isager et al. (2023), Language Teaching Review Panel (2008), Marsden et al. (2018), Mackey (2012), McManus (2022a, 2022b, 2023), McManus and Liu (2022), Morrison (2022), Nosek and Errington (2020), Plucker and Makel (2021), Polio (2012), Polio and Gass (1997), Porte (2012), Porte and McManus (2019), Tschichold (2023), Yamashita and Neiriz (2024), Zwaan et al. (2018).

Title

In addition to identifying the topic, main variables, and/or theoretical issues investigated in the study, the title should also state the replication approach adopted in the study (e.g., *exact replication, close replication, approximate replication*; for definitions, see Porte & McManus, 2019). Doing so helps increase the discoverability and use of replication studies in the field and informs readers about the magnitude of change between the initial study and the replication.

Abstract

The abstract should follow general guidelines for abstracts by providing information about the study's aims/objectives, the data sample, the method, findings, and

https://doi.org/10.1017/S0267190524000011 Published online by Cambridge University Press

Table 1. Standards for replication studies in applied linguistics

Paper section	Features to be included
Title	• Identify the main topic, variables, and/or theoretical issues under study. • State the replication approach (e.g., close replication, approximate replication).
Abstract	• State the motivation for conducting the replication, the study that was replicated, and the replication approach (e.g., close, approximate). • Describe modifications implemented in the replication. • Present key findings from the replication with comparisons to the initial study.
Introduction and background	• Justify why the initial study was selected for replication. • Review the initial study (design, method, findings) and other replications of it. • Define the replication approach. • Describe and justify changes made in the replication study, including features that are considered major and minor.
Research questions and/or hypotheses	• Use the initial study's research questions and/or hypotheses, if available, with appropriate and clearly identified edits that reflect any modifications made. • Note in the text if no research questions were included in the initial study and that new ones were created in the replication.
Method	• Describe all similarities and differences between the initial study and the replication in terms of data sample, research design, methods, data collection procedures, data preparation, and data analysis. • Specify and justify the sample size used in the replication. • State whether the initial study's data collection materials were available and used in the replication and how these were accessed. Modifications to the initial study's materials should be described, justified, and provided. New data collection materials created for the replication study should be described, and procedures for determining the comparability of the new data collection materials in light of those described in the initial study should be outlined. • Describe to what extent data coding and analysis procedures are the same as those used in the initial study. Additional and modified procedures implemented in the replication should be described and justified. • Describe procedures for determining the replicability of an effect.
Results	• Report results to allow for maximum comparison with the initial study, including narratively, in tables of results, as well as in figures. • Use supplementary materials if text length restrictions prevent full and transparent reporting of between-study similarities and differences.
Discussion and conclusions	• Revisit research questions and/or hypotheses in light of the results. • Discuss similarities and differences in findings between the initial and replication studies, as well as those reported in prior replications. • Interpret results from the replication study, accounting for modifications made to the initial study. • Discuss limitations in the replication study's design and methods • Review and propose future studies and replications that can build on and advance the line of research investigated.
Data and materials	• Make all information necessary to understand, evaluate, and build upon the study publicly available using but not limited to public repositories, including data collection materials, data, data preparation procedures, and analytical procedures.

conclusions. At the same time, the abstract should also state the motivation for conducting the replication study (as part of the aims and objectives). The abstract should clearly identify the study selected for replication and the replication approach

https://doi.org/10.1017/S0267190524000011 Published online by Cambridge University Press

(e.g., close, approximate). Additional information should be provided about the key modifications implemented in the replication (e.g., modifications to the population, context, methods). The abstract should state key findings from the replication in light of those reported in the initial study. The conclusion should state what has been learned from doing the replication.

Introduction and background

The introduction and background play an important role in motivating the replication study (i.e., why was a replication study carried out) and justifying several aspects of the replication approach, including what aspects of the initial study's design and/or method were modified and why. The background section should also provide a justification for why the initial study was selected for replication by drawing on multiple factors (e.g., theoretical, academic, societal impact of the study, methodological uncertainties). The use of clear subheadings in this section is recommended to highlight these important components of the replication study (e.g., "The initial study," "Motivation for replication").

Furthermore, the replication study should provide a close review of the initial study's research design, methods, and findings, with a discussion of how the initial study fits in with broader work and scholarly discussion in the field. If other studies have replicated the initial study, these should also be discussed. It should also be made clear how a replication of the initial study has the potential to contribute to disciplinary knowledge (e.g., perhaps replication is needed because the initial study's claims are theoretically important and valuable, but the empirical evidence underlying the claims is ambiguous or unclear).

The background section should also provide a definition for the replication approach (e.g., close, approximate), with references, given wide variations in nomenclature in current usage. For example, if the replication study is defined as a *close replication*, one definition could be as follows: "a close replication allows only one major variable to be modified. All other aspects of the previous study are unchanged" (McManus, 2023, p. 337). In the example of a close replication, the text should also describe what the major modification was and justify why it is considered a major rather than a minor modification. All other modifications that are determined to be minor in scope should be described and justified. For example, perhaps changes to the location (urban vs. rural United States) or the forms of data collection (online vs. face-to-face) are considered by the research team to constitute minor modifications.

Research questions and/or hypotheses

The replication study should include research questions and/or hypotheses/predictions consistent with conventions in the field in order to give structure and focus to the replication study. Given that a replication study systematically revisits a prior study by repeating its research design and methods with or without changes, it is expected that the same general research questions will be used, given the shared aims of the study. However, appropriate edits that reflect modifications to the initial study should be implemented. In short, all modifications to the initial study's research questions

https://doi.org/10.1017/S0267190524000011 Published online by Cambridge University Press

and/or hypotheses should be clear. If the initial study included no research questions, this information should be noted in the text.

Method

The replication study's design, methods, and analyses should be clearly reported in line with current conventions in the field. In addition, full and transparent reporting of all differences between the initial study and the replication is required, especially in terms of data sample, research design, methods, procedures, data handling, and analyses. This information can also be summarized in a table.

The replication should indicate whether the initial study's data collection materials were available and used in the replication and how these were accessed (e.g., from the authors, an open repository). If the initial study's materials were modified, including translations, the text should describe how they were modified. If new data collection materials were created, the procedures for how these were created should be described. Any new materials should be made publicly available. Procedures for determining the comparability of the new data collection materials in light of those described in the initial study should be outlined.

The replication study should indicate to what extent the data coding and analysis procedures are the same as those used in the initial study. Additional and modified procedures implemented in the replication should be described and clearly justified. In addition, the replication should describe the procedures for determining the replicability of an effect (i.e., how will we know if the initial study's results are confirmed in the replication?). For quantitative research, for instance, such indices of replicability could include: does the point estimate of the replication study occur within the 95% confidence intervals of the initial study and vice versa?

Results

In addition to following field-specific reporting standards, the replication study should report its results in a manner that allows for maximum comparison with the initial study as possible, including narratively, in tables of results, as well as in figures. For example, describing the main findings from the initial study before presenting those from the replication is one way to compare results. Results may also be reported to match the order of presentation in the initial study. In addition, descriptive results from the initial study, where available, should be presented alongside those from the replication in tables, and figures can be used to visualize results. If text length restrictions prevent full and transparent reporting of between-study similarities and differences, these should be reported in supplementary materials and made publicly available. However, journals should ensure that article lengths for replication studies are not shorter than those for regular articles.

Discussion and conclusions

In addition to following field-specific conventions, the replication study should revisit its research questions and/or hypotheses in light of its results and provide statements of

https://doi.org/10.1017/S0267190524000011 Published online by Cambridge University Press

support, nonsupport, etc. Similarities and differences between the initial and replication studies should be discussed, as well as those reported in prior replications, if appropriate. Results from the replication study should be interpreted, taking into account modifications made to the initial study's research design and methods. Limitations in the replication study's design and methods should be discussed. Lastly, the replication should review and propose future studies and replications of the current study that can build on and advance the line of research investigated.

Data and materials

The replication study should make all information necessary to understand, evaluate, and build upon the study publicly available by using public repositories (e.g., IRIS, OSF), including data collection materials, data (raw and coded), data preparation procedures, and analytical procedures and code (e.g., R script).

Conclusion and the future of replication in applied linguistics

This article has proposed a new standard for replication studies in applied linguistics with the goal of supporting researchers in conducting and reporting high-quality replication studies. The proposed new standard aims to improve the rigor, quality, and conduct of replication studies in applied linguistics. As previously noted, numerous resources exist to support researchers in how research studies are reported (e.g., APA, MLA), but very few resources are available for reporting replication studies. This proposal includes recommendations for all aspects of research reporting, including the replication study's title, abstract, introduction and background, research questions and/or hypotheses, method, results, discussion and conclusions, and data and materials. For each section, recommendations are made that aim to communicate unique aspects of the replication approach that are not captured in general reporting guidelines, such as clearly identifying the type of replication study, why the replication study was conducted, and what modifications were made to the initial study's design and methods with justifications. Another recommendation is that researchers clearly draw comparisons between the studies in the methods, results, and discussions, as well as making all information connected to the study publicly available (e.g., data, materials, analysis protocols). Through this standard, replication studies in applied linguistics can better meet their aims to confirm, consolidate, and advance knowledge and understanding. Furthermore, this proposed new standard can help readers more easily understand how the replication study was carried out to evaluate its claims.

References

Abbuhl, R. (2011). Why, when, and how to replicate research. In A. Mackey & S. M. Gass (Eds.), *Research methods in second language acquisition: A practical guide* (pp. 296–312). Wiley Blackwell.

American Association for Applied Linguistics. (2009). Encouraging replication research in the field of AL and SLA. [Colloquium]. In *American Association for Applied Linguistics Conference*. Denver, Colorado, USA.

American Association for Applied Linguistics. (2020). Replication research in AL: What needs to be replicated, why, and how to do it. [Colloquium]. In *American Association for Applied Linguistics Conference*. Denver, Colorado, USA. [Cancelled due to COVID-19].

https://doi.org/10.1017/S0267190524000011 Published online by Cambridge University Press

Appelbaum, M., Cooper, H., Kline, R. B., Mayo-Wilson, E., Nezu, A. M., & Rao, S. M. (2018). Journal article reporting standards for quantitative research in psychology: The APA Publications and Communications Board task force report. *American Psychologist*, *73*(1), 3–25.

Atkinson, D. (2012). Disciplinary dialogues: Introduction. *Journal of Second Language Writing*, *21*(3), 283.

Awan, S., Yahya, U., & Arif, M. (2023). Quality standards of qualitative research in applied linguistics: A conceptual review. *VFAST Transactions on Education and Social Sciences*, *11*(2), 68–75.

Brandt, M. J., Ijzerman, H., Dijksterhuis, A., Farach, F. J., Geller, J., Giner-Sorolla, R., … van't Veer, A. (2014). The Replication Recipe: What makes for a convincing replication? *Journal of Experimental Social Psychology*, *50*, 217–224.

Brown, J. D. (2012). Writing up a replication report. In G. K. Porte (Ed.), *Replication research in applied linguistics* (pp. 173–197). Cambridge University Press.

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

Errington, T. M., Mathur, M., Soderberg, C. K., Denis, A., Perfito, N., Iorns, E., & Nosek, B. A. (2021). Investigating the replicability of preclinical cancer biology. *eLife*, *10*, e71601.

Gurzynski-Weiss, L., & Kim, Y. (Eds.) (2022). *Instructed second language acquisition research methods*. Benjamins.

Hatch, E. M., & Farhady, H. (1982). *Research design and statistics for applied linguistics*. Newbury House.

In'nami, Y., Mizumoto, A., Plonsky, L., & Koizumi, R. (2022). Promoting computationally reproducible research in applied linguistics: Recommended practices and considerations. *Research Methods in Applied Linguistics*, *1*(3), 100030.

International Symposium on Bilingualism. (2023). Symposium 27. In *International Symposium on Bilingualism. Macquarie University*. Sydney, Australia.

Isager, P. M., Van Aert, R. C. M., Bahník, Š., Brandt, M. J., DeSoto, K. A., Giner-Sorolla, R., … Lakens, D. (2023). Deciding what to replicate: A decision model for replication study selection under resource and knowledge constraints. *Psychological Methods*, *28*(2), 438–451.

Language Learning. (2024). *Author guidelines: Instructions for contributors*. Retrieved March 21, 2024, from https://onlinelibrary.wiley.com/page/journal/14679922/homepage/forauthors.html

Language Teaching. (2008). Call for papers: Replication research studies. *Language Teaching*, *41*(1), iii–iv.

Language Teaching Review Panel. (2008). Replication studies in language learning and teaching: Questions and answers. *Language Teaching*, *41*(1), 1–14.

Liu, M. (2023). Whose open science are we talking about? From open science in psychology to open science in applied linguistics. *Language Teaching*, *56*(4), 443–450.

Long, M. H. (1993). Assessment strategies for second language acquisition theories. *Applied Linguistics*, *14*(3), 225–249.

Mackey, A. (2012). Why (or why not), when, and how to replicate research. In G. K. Porte (Ed.), *Replication research in applied linguistics* (pp. 21–46). Cambridge University Press.

Mackey, A. (2013). Instruments for research into second languages. *TIRF Today*, *4*, 2.

Mackey, A., & Gass, S. M. (2011). *Research methods in second language acquisition: A practical guide*. Wiley Blackwell.

Mackey, A., & Gass, S. M. (2016). *Second language research. Methodology and design* (2nd ed.). Routledge.

Mackey, A., & Gass, S. M. (2022). *Second language research. Methodology and design* (3rd ed.). Routledge.

Mackey, A., & Gass, S. M. (2023). *Current approaches in second language acquisition research: A practical guide*. Wiley Blackwell.

Mackey, A., & Marsden, E. (Eds.). (2016). *Advancing methodology and practice: The IRIS repository of instruments for research into second languages*. Routledge.

MacWhinney, B. (2000). *The CHILDES Project: Tools for analyzing talk*. Lawrence Erlbaum.

Marsden, E. (2013). Instruments for research into second language learning (IRIS). *ELT Research*, *28*, 34.

Marsden, E. (2019). Open science and transparency in applied linguistics research. In C. A. Chapelle (Ed.), *The encyclopedia of applied linguistics* (pp. 1–10). Wiley.

Marsden, E., & Mackey, A. (2014). IRIS: A new resource for second language research. *Linguistic Approaches to Bilingualism*, *4*(1), 125–130.

Marsden, E., Morgan-Short, K., Thompson, S., & Abugaber, D. (2018). Replication in second language research: Narrative and systematic reviews and recommendations for the field. *Language Learning*, *68*(2), 321–391.

McKinley, J., & Rose, H. (2020). *The Routledge handbook of research methods in applied linguistics*. Routledge.

https://doi.org/10.1017/S0267190524000011 Published online by Cambridge University Press

McManus, K. (2022a). Are replication studies infrequent because of negative attitudes? Insights from a survey of attitudes and practices in second language research. *Studies in Second Language Acquisition*, *44*(5), 1410–1423.

McManus, K. (2022b). Replication research in instructed SLA. In L. Gurzynski-Weiss & Y. Kim (Eds.), *Instructed second language acquisition research methods* (pp. 103–122). John Benjamins.

McManus, K. (2023). How and why to conduct a replication study. In A. Mackey & S. M. Gass (Eds.), *Current approaches in second language acquisition research* (pp. 334–351). Wiley.

McManus, K. (2024). Replication in second language research. *Studies in Second Language Acquisition*.

McManus, K., & Liu, Y. (2022). Using elicited imitation to measure global oral proficiency in SLA research: A close replication study. *Language Teaching*, *55*(1), 116–135.

Morrison, K. (2022). Conceptual replications, research, and the "what works" agenda in education. *Educational Research and Evaluation*, *27*(1–2), 35–60.

National Institute of Standards and Technology. (2022). Standards & Measurements. Retrieved March 21, 2024, from https://www.nist.gov/services-resources/standards-and-measurements

Nerenz, A. (2016). Editor's message. *Foreign Language Annals*, *49*(1), 7–9.

Norris, J. M., Plonsky, L., Ross, S. J., & Schoonen, R. (2015). Guidelines for reporting quantitative methods and results in primary research: Guidelines for reporting quantitative methods. *Language Learning*, *65*(2), 470–476.

Nosek, B. A., & Errington, T. M. (2020). What is replication? *PLoS Biology*, *18*(3), e3000691.

Phakiti, A., De Costa, P., Plonsky, L., & Starfield, S. (Eds.) (2018). *The Palgrave handbook of applied linguistics research methodology*. Palgrave.

Plonsky, L. (2012). Replication, meta-analysis, and generalizability. In G. K. Porte (Ed.), *Replication research in applied linguistics* (pp. 116–132). Cambridge University Press.

Plonsky, L. (2014). Study quality in quantitative L2 research (1990–2010): A methodological synthesis and call for reform. *The Modern Language Journal*, *98*(1), 450–470.

Plonsky, L. (2015a). Quantitative considerations for improving replicability in CALL and applied linguistics. *Calico Journal*, *32*(2), 232–244.

Plonsky, L. (Ed.) (2015b). *Advancing quantitative methods in second language research*. Routledge.

Plonsky, L. (Ed.) (2024). *Open science in applied linguistics*. Applied Linguistics Press.

Plucker, J. A., & Makel, M. C. (2021). Replication is important for educational psychology: Recent developments and key issues. *Educational Psychologist*, *56*(2), 90–100.

Polio, C. (2012). Replication in published applied linguistics research: A historical perspective. In G. K. Porte (Ed.), *Replication research in applied linguistics* (pp. 47–91). Cambridge University Press.

Polio, C., & Gass, S. M. (1997). Replication research and reporting. A commentary. *Studies in Second Language Acquisition*, *19*(4), 499–508.

Porte, G. K. (Ed.) (2012). *Replication research in applied linguistics*. Cambridge University Press.

Porte, G. K. (2013). Who needs replication? *Calico Journal*, *30*(1), 10–15.

Porte, G. (2015). Replication research in quantitative research. In J. D. Brown & C. Coombe (Eds.), *The Cambridge guide to research in language teaching and learning* (pp. 140–145). Cambridge University Press.

Porte, G. K., & McManus K. (2019). *Doing replication research in applied linguistics*. Routledge.

Porte, G. K., & Richards, K. (2012). Replication in second language writing research. *Journal of Second Language Writing*, *21*(3), 284–293.

Riazi, A. M., & Farsani, M. A. (2023). Mixed-methods research in applied linguistics: Charting the progress through the second decade of the twenty-first century. *Language Teaching*, 1–40.

Rose, H., McKinley, J., & Briggs Baffoe-Djan, J. (2020). *Data collection research methods in applied linguistics*. Bloomsbury.

Santos, T. (1989). Replication in applied linguistics research. *TESOL Quarterly*, *23*(4), 699–702.

Studies in Second Language Acquisition. (2024). *Author instructions: Preparing your materials*. Retrieved March 21, 2024, from https://www.cambridge.org/core/journals/studies-in-second-language-acquisition/information/author-instructions/preparing-your-materials

Symposium on Second Language Writing. (2010). *Crossing Disciplinary Boundaries*. [Symposium]. University of Murcia, Spain.

Tarone, E., Swain, M., & Fathman, A. (1976). Some limitations to the classroom applications of current second language acquisition research. *TESOL Quarterly*, *10*(1), 19–32.

https://doi.org/10.1017/S0267190524000011 Published online by Cambridge University Press

TESOL Quarterly. (2024). *TESOL Quarterly Submission Guidelines.* Retrieved March 21, 2024, from https://onlinelibrary.wiley.com/page/journal/15457249/homepage/forauthors.html

Tschichold, C. (2023). Replication in CALL. *ReCALL*, *35*(2), 139–142.

Valdman, A. (1993). Replication study. *Studies in Second Language Acquisition*, *15*, 505.

Yamashita, T., & Neiriz, R. (2024). Why replicate? Systematic review of calls for replication in *Language Teaching*. *Research Methods in Applied Linguistics*, *3*(1), 100091.

Zwaan, R. A., Etz, A., Lucas, R. E., & Donnellan, M. B. (2018). Making replication mainstream. *Behavioral and Brain Sciences*, *41*, e120.

Cite this article: McManus, K. (2024). The future of replication in applied linguistics: Toward a standard for replication studies. *Annual Review of Applied Linguistics*, *44*, 133–149. https://doi.org/10.1017/S0267190524000011

https://doi.org/10.1017/S0267190524000011 Published online by Cambridge University Press