This article is an advanced online publication. It should be cited as:

Marefat, F., Farahanynia, M., Hamidi, F., Najjarpour, M., Banitalebi, Z., & Alamdar, P. (2025) Trends of replication studies in applied linguistics journals: a systematic review over half a century. *Canadian Journal of Applied Linguistics.* Advanced Online Publication.

Trends of Replication Studies in Applied Linguistics Journals: A Systematic Review Over Half a Century

Fahimeh Marefat
Allameh Tabataba'i University

Mahsa Farahanynia
Pishtazan Institute of Higher Education

Farzaneh Hamidi
Allameh Tabataba'i University

Mona Najjarpour
Allameh Tabataba'i University

Zahra Banitalebi
Allameh Tabataba'i University

Parvin Alamdar
Allameh Tabataba'i University

Abstract

Despite the importance of replication research in scientific fields, very few replications are conducted in applied linguistics (AL). To enhance language researchers' awareness of replications and provide a systematic evaluation of current replications, this study analyzed replication studies published in 92 AL leading journals from 1970 to 2021 based on five themes of replication labels, methodological orientations, research trends, authorship, and citation counts of replicators. The results reveal that replication labels have explicitly been mentioned since 2002, the replication of quantitative studies has predominantly been raised, studies on second language acquisition were frequently replicated, collaborative authorship has increased in replications, and influential AL scholars tend to conduct replication research. The study highlights the need for a well-established framework with thorough descriptions of various replication labels functioning as a reference for replication studies and calls for replication research in the areas and methodological orientations marginalized in AL. It is also recommended that prominent figures perform more replication research to consolidate its status in AL.

Résumé

Malgré l'importance de la recherche en réplication dans les domaines scientifiques, très peu de réplications sont menées en linguistique appliquée (LA). Afin de sensibiliser les chercheurs en linguistique à la nécessité des réplications et de fournir une évaluation systématique des réplications actuelles, cette étude a analysé les études de réplication publiées dans 92 revues de premier plan en LA de 1970 à 2021, en se basant sur cinq thèmes : les étiquettes de réplication, les orientations méthodologiques, les tendances de recherche, l'auteurship et les comptes de citations des réplicateurs. Les résultats révèlent que les étiquettes de réplication ont été explicitement mentionnées depuis 2002, que les réplications d'études quantitatives ont été principalement abordées, que les études sur l'acquisition d'une langue seconde ont été fréquemment répliquées, que la collaboration entre auteurs a augmenté dans les réplications, et que les chercheurs influents en LA ont tendance à mener des recherches en réplication. Cette étude met en évidence la nécessité d'un cadre bien établi avec des descriptions détaillées des différentes étiquettes de réplication, servant de référence pour les études de réplication, et appelle à la recherche en réplication dans les domaines et orientations méthodologiques marginalisés en LA. Il est également recommandé que des personnalités éminentes effectuent davantage de recherches en réplication pour consolider son statut en LA.

Trends of Replication Studies in Applied Linguistics Journals: A Systematic Review Over Half a Century

Background

Replication Definition and Function

Replication research is an umbrella term referring to repeating experiments in order to evaluate the reliability and generalizability of previous research findings (Kobrock & Roettger, 2023). Generally, replication studies (RSs) in which the same materials and procedures are used serve to test the reliability of original findings, and those in which specific methodological or participant characteristics are altered aim to validate the generalizability of original research findings (Marsden et al., 2018). Replication research can also offer better explanations of the phenomena under investigation, which enriches and validates existing theories and substantially contributes to pedagogy. These issues highlight the fact that replication research should be an integral part of any experimental science (Kobrock & Roettger, 2023), including applied linguistics (AL).

Replication Status in Applied Linguistics

Conducting replication studies has been largely sidelined in the history of AL for a variety of reasons (Abbuhl, 2018). For instance, the prioritization of innovation and originality, the lack of motivation and prestige for replication compared to original research by funding agencies, leading journals, editors, and reviewers in AL are among the many reasons that have led to the scarcity of this research genre in the field (Abbuhl, 2018; Porte, 2012). To offer a deeper insight into why replications are not welcomed in AL, McManus (2021) explored 354 researchers' perceptions of replication and conducting replication research. It was revealed that the majority of the participants evaluated RSs as relevant and precious to the field and rejected the widespread stereotypical assertion that RSs lack innovation, creativity, and originality. However, less chance of publication in well-established journals, the absence of explicit labeling, and the lack of training in replication methodology were among the most relevant issues to the scarcity of replication in this field of study. To estimate the replication position in experimental linguistics in a systematic study, Kobrock and Roettger (2023) assessed the replication rates for over 50,000 articles across 98 journals and found that under 1.6% of their corpus rates explicitly stated the search string "replicat*". Furthermore, only 1 in 1250 experimental linguistic studies was reported to be an independent direct replication (employing the same methodology but using new data and undertaking new data analyses). The authors concluded that the replication rate in experimental linguistics is not satisfactory and requires further consideration.

Recently, there has been a growing recognition of the significance of replications, leading to a shift in this trend. Abbuhl (2018) introduces three reasons for the recent surge of interest in conducting replications, one of them being the maturity of AL as a field. AL,

as a young area of research, has not only explored new dimensions of language learning but has also directed its attention to the importance of the confirmability of the existing results. The second one concerns the growing sophistication of AL in terms of statistics and their use and interpretation. As for replication research, primarily, it was preferred to replicate merely significant results to ensure their reliability or generalizability; but nowadays, scholars (Nassaji, 2012) believe that there should be replications to evaluate the validity of any given results, including both significant and non-significant ones. The third reason for the burgeoning replication movement is the current crisis of replicability caused by the prioritization of innovation. This crisis stems from the controversial claims that some of the published results are false, some are cases of fraud, and some have used questionable research practices (see John et al., 2012; King & Mackey, 2016; Pashler & Harris, 2012 for further information).

There have also been many repeated calls and recommendations for replication research in the field of second language (L2) research (Polio & Gass, 1997; Porte, 2012; Santos, 1989; Vandergrift & Cross, 2017). Recently, to facilitate many RSs in L2 research, an open-access repository of Instruments for Research into Second Language Learning called IRIS has been introduced where the authors can easily upload their data and instruments to this database (Marsden & Mackey, 2014; Mackey & Marsden, 2016: Marsden et al., 2016). Porte and McManus's (2019) book has been another big step toward uplifting the status of replication research in AL. This book exclusively focuses on replication research in AL, demystifies the most prevalent clichés about replication, underscores the necessity of carrying out RSs in AL, and highlights the substantial contributions of such research to the expansion of knowledge. It also provides the readers with exhaustive guidelines and concrete activities on how to conduct replication research and advances some practical recommendations for prospective authors on how to select target journals for publication and how to present a summary of their replicated works in conferences, seminars, and poster sessions.

Among other initiatives, the Journal of Studies in Second Language Acquisition (SSLA) has made it possible for researchers to submit RSs by including a type of manuscript known as a replication study in its submission guidelines. The RSs accepted and published in this journal have also been given a separate entry in the journal, called research studies. Similarly, the Journal of Memory and Language called for replication of the most influential findings in the field of language and memory, and dedicated a special issue in 2022, *Replicating Influential Findings in Memory and Language*, to publishing such studies. These strategic plans have offered a promising path for promoting replication research in the field of AL.

Role of Open Science Practices in Replication in Applied Linguistics

The practitioners of open science assume that as science generates observable evidence, the findings should be confirmed across numerous studies. For this purpose, researchers' practices, data, and behaviors should be entirely available, transparent, reproducible, and reusable so as to pave the way for other researchers to challenge the existing claims and

hence improve research quality (Popper, 1959). This issue underlines the importance of replication research in science, which can be facilitated by methodological transparency and open science practices, such as materials, code, and data sharing, study pre-registration, and large-scale collaborations (Bochynska et al., 2023; Del Ben et al., 2022; Marsden, 2020). Open science practices have found their way to different fields of study. For instance, in the field of psychology, the open science framework, established in 2011, provides free access to materials and resources to the general public to facilitate RSs in this field (Marsden & Mackey, 2014; Marsden et al., 2016). In the field of bilingualism, Del Ben et al. (2022) suggest that researchers acquaint themselves with such practices to expand their understanding of variations integral to bilingualism research and strengthen the science relevant to this domain. In fact, RSs and open science practices are two ways to contribute to the robustness of science (Nosek & Errington, 2020).

A burgeoning attention to open science practices for methodological transparency is witnessed in the field of AL. Marsden (2020) contends that methodological transparency has significant impacts on the quantity, scope, and quality of AL and more importantly on the quality of replicability, reproducibility, and synthesis of studies in this field. She refers to some developments and initiatives that have improved methodological transparency in AL. These practices include the introduction of a public website by the Center for Open Science for the preregistration of studies, the possibility of registering study reports prior to data collection to receive the reviewers' suggestions and modifications, the provision of peer reviewers' openness initiatives in which the reviewer refuses to evaluate the manuscript unless the detailed methodological information is available, and finally the promotion of some journal guidelines that require the submission of full methodological details through several stages. Marsden and Morgan-Short (2023) strongly argue that open research practices, such as open access publication, registered reports, open materials, data archives (IRIS repository), and multi-site collaboration research attempts, would improve the reliability, validity, replicability, sampling, quality, and scope of the existing language learning research in the distant future.

To accentuate the essence of open science practices and replication for the generalizability of findings, Plonsky (2023) investigated the quality of several sample-related variables observed in 308 reports of synthetic AL research and demonstrated that the insufficiency of information about sample demographics like age, target language, language proficiency, and instructional settings would extremely limit the generalizability of findings to other samples. He concluded that open science practices would pave the way for performing replication research with various samples, which raises the likelihood of result generalization. Isbell et al. (2022) explored the utilization of open science practices and replication as a preventive policy and a safeguard mechanism to hinder fraudulent and false findings in AL. To this end, they surveyed a group of 351 applied linguists for questionable research practices (QRP) in quantitative research. Seventeen percent of the participants declared one or more scientific malpractices and 94% accepted one or more QRPs pertinent to quantitative research. These results bear witness to the merit of boosting open science practices like study preregistration, and codes and materials sharing to enhance the methodological transparency, open a path for further inspection by unbiased

parties, avoid the expansion of disingenuous findings, and encourage researchers to conduct replication research in the field of AL.

Replication Controversies

The replication of different methodological (qualitative or quantitative) orientations of original studies arouses controversy. There is a debate about whether replications are possible or even desirable for qualitative studies (Casanave, 2012; Matsuda, 2012; Sasaki, 2012). The argument against such kind of replication lies in the fact that the verification of results in this type of inquiry is reductionist and fruitless. According to Schofield (2002), the particularities of such research types, including the participants, researchers, and research environments, would make it difficult to replicate a qualitative study. Moreover, it is almost impossible to obtain the same results so as to confirm the generalizability of the findings. In this regard, Abbuhl (2018) claimed that the concept of success and failure of replication in the replication of some types of qualitative studies needs to be reconceptualized, which means that the current controversies about the concept of transferability in the replication of qualitative studies should be dispelled. Additionally, some years earlier, Porte and Richards (2012) had declared the necessity of reconceptualizing the concept of transferability in qualitative replications. It means that we should draw our attention away from identifying whether the findings are transferable toward modifying the notion of transferability itself. As an example, one may ask in what respect or for what reasons the cases in the original qualitative studies might be transferable. This approach is advocated by many qualitative researchers (Golden, 1995; Markee, 2017; Schofield, 2002). Markee (2017), for instance, presents a strategy for replicating qualitative studies in L2 classroom research, including comparative reproduction research and a common practice in conversation analysis.

The aforementioned barriers do not restrict the replication of quantitative studies since it is claimed that experimental studies are methodologically more adaptable to get replicated (Porte & McManus, 2019). However, there are controversies surrounding the methodological transparency of original quantitative studies. To tackle this issue, Larson-Hall and Plonsky (2015) highlight the importance of reporting the following quantitative data types: descriptive statistics, effect sizes and confidence intervals, instrument reliability, visual displays of data, and raw data. They claim that such transparency in data reporting can facilitate the process of data interpretations, meta-analysis, and replication research.

Current Replication Taxonomies

To date, various taxonomies of replication have been propounded in different fields of study. For instance, in psychology, Lykken (1968) introduces three types of literal, operational, and constructive replications. In literal replication, additional participants are added to the same study. In operational replication, the same method and procedures, often with novel participants, are used. While the conceptual meaning remains unchanged in this type, the dependent variable is operationalized anew. In constructive replication, the

relation between the constructs is re-examined through other methods than the ones employed in the original study. The taxonomy, proposed by Makel et al. (2012) and Schmidt (2009), has only two subtypes: direct and conceptual. In direct replication, there is no intentional alteration to the original study while, in conceptual replication, there are some intentional adaptations to the initial study.

Concerning L2 research, a tripartite distinction is usually made between exact, approximate, and conceptual replication (Abbuhl, 2012a, 2012b; Language Teaching Review Panel, 2008; Marsden et al., 2018; Polio, 2012; Porte, 2012; Porte & Richards, 2012). Exact or close replication, which is the least common, albeit well-known, type, is defined as the identical repetition of the original study. In approximate or partial replication, as a more frequently undertaken type, the methodology of the original study is kept unchanged; some modifications are made to non-major variables, such as participants or research settings, in order to evaluate the generalizability of the original findings. The last type is conceptual replication in which a new research design is used to investigate the concepts or hypotheses put forth by the original study. The purpose of conducting conceptual replication is to verify whether the results of the original study are the artifacts of its own methodology or generalizable to a new context.

Research on Replication in Applied Linguistics

Replication research has not received the attention it deserves in social (Makel & Pluker, 2014) or experimental sciences (Kobrock & Roettger, 2023). For example, in an assessment of the quantity and nature of replication efforts in the field of education, Makel and Pluker (2014) reported only 0.13% replication endeavors. Likewise, there is a paucity of replication research in AL, primarily due to the low academic payoff (Kobrock & Roettger, 2023). For example, Marsden et al. (2018) provided a systematic review of RSs in 26 journals from 1973 to 2016 by examining the existing nomenclature used to identify self-labeled RSs and the extent to which the labels reflect the number of changes between the original and replication studies. Their discoveries identified only 67 (52%) self-labeled replications out of 129 study reports, of which 67% of the studies were labeled as "replicat*", 25% as a replication, and 21% as a partial replication in their titles. Their results did not find any cases of direct or exact replication types of study. In fact, they found only replicated studies that could be described as close (1%), replication (3%), and conceptual replications (4%). Accordingly, the authors claimed that a majority of the replicated studies did not explicitly carry the label type in the title, and merely by examining the entire article could one identify its type. To Marsden and her team, this lack of explicit nomenclature is a weakness of RSs. They recommended using more self-labeling studies with the term 'replication' and applying principled, standard nomenclatures like direct, partial, and conceptual replications in these studies to ensure that any potential diversity and circumstantial details are fully documented by replicators. Marsden et al. (2018) also investigated another factor associated with the reproducibility of original findings: authorship, that is, how closely the results of replicated studies resemble those of the original studies when the original studies are replicated by the same or other authors.

Their results demonstrated that the degree of reproducibility of original findings was 48% when the same authors were involved in the replication, while it was considerably lower (28%) when there was no overlap between the authors of the original and replicated studies. Their results were further supported by the findings of Kobrock and Roettger (2023) as they found that independent replications were less frequent than those conducted by the authors of the original study. Therefore, overlap in authorship can be an intervening factor affecting the results of replicated studies.

Significance of the Study and Research Questions

Given the plethora of recommendations offered for conducting replications in L2 research, one would expect to witness such studies quite frequently. However, lack of prestige, excitement, and originality serve as barriers to undertaking such research in AL (Marsden et al., 2018), resulting in a scarcity of RSs. It is therefore necessary to carry out a more extensive systematic review of the existing RSs to make the stakeholders aware of this situation and request the promotion of RSs in AL. To the best of our knowledge, just one systematic review (Marsden et al., 2018) has been conducted on replication research in AL. Nonetheless, Marsden et al. (2018) reviewed merely the studies whose titles carried the term "replication". Therefore, more research is needed to delve into the whole contents of the studies for "replication" to provide a clearer picture of replication research in AL. Moreover, through such systematic studies, authorities, stakeholders, and journal editors will be informed of the status quo of replication research, which may lead to some modifications in their submission guidelines and consequently promote researchers' academic motivation to undertake replications. To bridge this gap, contribute to this strand of inquiry, and raise awareness among L2 researchers and editors about the implementation of replication research, the present systematic review was conducted. More specifically, we aimed at perusing RSs (with the term 'replication' mentioned in the text or title) in terms of replication labels, methodological orientations, research trends, authorship orientations, and research replicators' citations in 92 AL leading journals within the last 50 years. Investigating these issues, left untouched by Marsden et al. (2018), provides a more comprehensive account of RSs in the literature.

Porte (2012) introduces three types of exact, approximate, and conceptual replication; Marsden et al. (2018) add some other replication labels like true, partial, and quasi-replication. Coding the data, we had these nomenclatures in mind. Methodological orientations or inclinations toward particular research approaches or paradigms, i.e., qualitative, quantitative, or mixed methods (Creswell, 2014; Denscombe, 2008), were examined to determine which types were more frequently replicated.

As for the research trends, the research focus of RSs was explored to unravel which research areas have been targets for replication. To this end, Stapleton and Shao's (2017) classification of research focus in L2 learning and teaching, including instructional effects, teacher cognition, exploratory practice, teacher practices, and SLA mechanisms, was examined (See Appendix A).

Furthermore, in their study, Riazi et al. (2020) explored authorship orientations in the context of English for Academic Purposes (EAP) and found an increasing trend of collaborative authorship (both co-authorship and multiple authorship) in the EAP studies from 2012 to 2019. Their findings were in line with the global trend of increasing collaboration among authors across institutions and the growing trend among researchers to participate in interdisciplinary and multidisciplinary research (Henriksen, 2018; Hyland, 2015). Following Riazi et al. (2020), this study also investigated whether this pattern of authorship is present among replicators when doing RSs. In this study, collaborative authorship behavior was defined as the number of researchers involved in replication research. Ultimately, we approached replicators' citations to explore whether or not they are among the influential figures in AL. By showcasing the impact of replicators, more researchers are encouraged to follow their footsteps and replicate more research in AL. Therefore, the following research questions led the course of this systematic review:

> How have replication studies labeled themselves in the 92 leading applied linguistics journals from 1970 to 2021?
> What are the methodological orientations (quantitative, qualitative, mixed method) of replication studies in the 92 leading applied linguistics journals from 1970 to 2021?
> What are the research trends of replication studies in the 92 leading applied linguistics journals from 1970 to 2021?
> What are the authorship orientations (single authorship, co-authorship, multiple authorship) in conducting replication research in the 92 leading applied linguistics journals from 1970 to 2021?
> Who are the most frequently cited replication researchers in applied linguistics and what does this disclose about replication?

Method

Inclusion Criteria and Study Selection

In this study, 214 replication studies published in 92 AL leading journals from 1970 to 2021 were included. The starting point was chosen to be 1970 since the first replication study was reported to be published then (Marsden et al., 2018). To select the top-tier journals, we consulted Marsden et al. (2018), and following Hashemi and Babaii (2013) and Khany and Tazik (2019), we referred to Weber and Campbell (2004), Jung (2004), Egbert (2007), who listed well-established AL journals. We initially excluded the repetitive journals in the aggregated inventory. Then, journals in sister fields, like education and psychology, were eliminated. To do so, first, we sorted out the journals with no language-related titles, and then searched through these journals to find language-related studies. The journals with at least one language-related study were included. In the end, the 92 leading journals, all indexed in SSCI, were identified (see Appendix B). Justifiably, SSCI journals are reliable concerning their strict blind review procedures and the methodological aspects

in the field (Hiver et al., 2021; Zou et al., 2020); moreover, concentration on methodological procedures gains more significance in research studies compared to other empirical research (Porte & McManus, 2019).

To locate the relevant studies in these journals, keywords like “replicat*”, “exten*”, “revisi*”, “replication”, “replicated”, “revision”, “revisited”, “extension”, and “extended” were all used to search in the title, abstract, keywords, or body of articles in each journal, as such truncations encompass the word families of replication, revision, and extension. The start and finish cutoff points were instituted to be 1970 and 2021, respectively. Four researchers, Ph.D. candidates majoring in AL, independently searched the selected journals’ webpages for the studies meeting the following criteria:

Written in English.
Published as an article in a journal or a book chapter (conference proceedings, newsletter, editorials, and incomplete articles were not included).
Empirical in nature and a genuine replication of previous studies.

Then, the duplicates were eliminated after rechecking. This initial screening resulted in 238 RSs conforming to the criteria. To assure the objectivity of this manual search, a random selection of total journals (N = 6) was double-checked by two of these researchers, and no new study was added in the second round. To ensure consistency in the inclusion of the studies, the screened articles were re-evaluated by an experienced applied linguist who had not been involved in the initial searching procedure. After a meeting with her, 24 studies were found to violate the last criterion and were excluded. This further screening yielded 214 empirical articles for analysis.

Coding Procedure

The papers chosen for inclusion were coded in terms of five main factors: replication labels, methodological orientations, research trends, authorship orientation, and replicators’ citations. In the pilot coding, 10 randomly selected articles were independently coded in terms of the above-mentioned variables by the four coders. To explore how the authors of RSs labeled their studies, the nomenclatures proposed by Marsden et al. (2018) were used. In their study, Marsden et al. (2018) devised a coding scheme based on a narrative review and literature on replication in L2 research and finally manually coded their corpus. This categorization has the potential of detailing the studies’ replication labels and providing richer information on what terms they have used to reveal the replication nature of their research. The coders read the whole studies to detect replication-related terms and classified them based on Marsden et al.’s (2018) replication labels manually. Studies using qualifiers with the term “replication” were coded under the categories that specified replication types, such as exact replication, partial replication, etc. Studies using the term “replication” without any qualifiers were coded under replicat* category. Moreover, studies containing the terms other than replicat*, such as revisit* and exten*, to avoid labeling their study as a replication (Marsden et al., 2018) were classified under their

corresponding categories. Eventually, studies with other labels than the aforementioned ones were put under the category of "other". In this way, any possible overlap between categories was eliminated. It is worth mentioning that to identify genuine replication and distinguish the replication from non-replication studies (despite the presence of the keywords), the coders analyzed the co-text around the search words included in each study to see whether the original study was explicitly or implicitly mentioned. As for the methodological orientations, the studies were manually categorized based on the major research paradigms of quantitative, qualitative, and mixed methods (Denscombe, 2008). To classify the selected empirical studies with respect to research trends, the framework proposed by Stapleton and Shao (2017) was used (see Appendix A). They used Chaudron's (2001, as cited in Stapleton & Shao, 2017) framework for preliminary independent coding to establish codes under "Main Focus". They coded the articles manually, identified the categories, and collectively wrote descriptors until an agreed-upon final rubric with 12 types of research trends in L2 learning and teaching was established. Using this rubric, the coders in the present study manually coded the articles. With regard to authorship orientation and collaborative behavior of replicators, the corpus was coded in terms of the number of authors. To be more precise, studies with one, two, and more than two authors were categorized as single authorship, co-authorship, and multiple authorship, respectively. In the pilot study, an inter-coder Kappa agreement of 0.81 was observed.

After that, each coder manually analyzed one-fourth of the studies (n = 53 each) in terms of the aforementioned variables. During the coding procedure, some problems were raised and rectified. To further appraise the reliability of coding, another SLA researcher coded almost half of the articles (N = 100, 46% of the sample). The inter-coder reliability (κappa = 0.76, $p < .001$) was acceptable (McHugh, 2012).

Moreover, citation counts for influential researchers conducting replication in the field of AL were meticulously extracted using Google Scholar's Altmetrics on 5th February 2023. For each identified replicator, we retrieved the number of citations they had accrued across the field. This involved conducting individual searches for each replicator's name, aggregating their citation counts from multiple sources, and cross-referencing the data to ensure accuracy. In order to compare the number of citations of these replicators, the 100 most-cited applied linguists were sought from AL most cited authors (Exaly, 2023). This open access database provides rankings of scholars, journals, and papers from the emergence of a field to the current day (Amjad et al., 2018). All the coded data were merged in an Excel file (see Supplementary Information for the coding sheet[1]).

Data Analysis

Data analysis of the present study exclusively drew on descriptive statistics, including the number and percentages of articles and the citation counts of the replicators. This is because, as in any typical review study, this systematic review aimed to identify and report the potential trends of replications in the field and provide reasonable accounts for them.

Results

The first research question addressed the number of self-labeled RSs published from 1970 to 2021 and the replication labels used in these studies (Table 1).

Table 1
Terms Used to Label Replication Studies over Time (1970-2021)

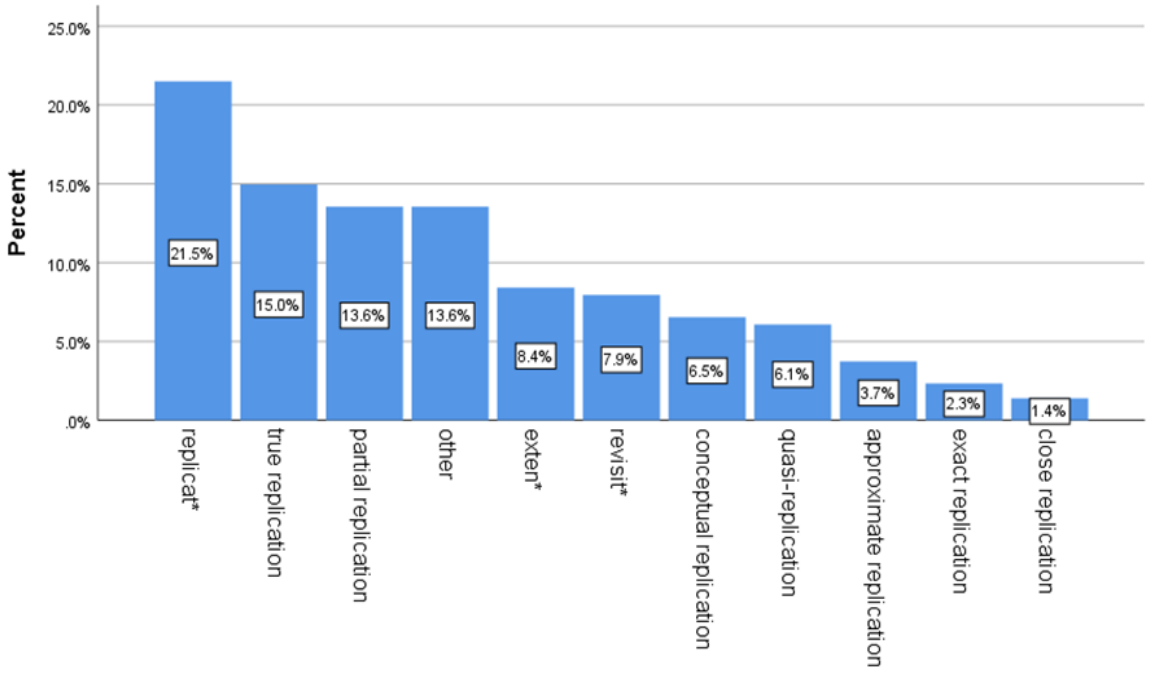
	Publication Year											
	1970-1980**		1981-1990		1991-2001		2002-2011		2012-2021		Total	
Labels	*N**	%	*N*	%	*N*	%	*N*	%	*N*	%	*N*	%
true replication	1	0.5	3	1.4	7	3.3	7	3.3	14	6.5	32	15.0
exact replication	0	0.0	0	0.0	0	0.0	1	0.5	4	1.9	5	2.3
close replication	0	0.0	0	0.0	1	0.5	0	0.0	2	0.9	3	1.4
approximate replication	0	0.0	0	0.0	0	0.0	1	0.5	7	3.3	8	3.7
quasi-replication	0	0.0	1	0.5	0	0.0	3	1.4	9	4.2	13	6.1
partial replication	1	0.5	2	0.9	4	1.9	7	3.3	15	7.0	29	13.6
conceptual replication	0	0.0	0	0.0	0	0.0	0	0.0	14	6.5	14	6.5
exten*	1	0.5	5	2.3	2	0.9	2	0.9	8	3.7	18	8.4
replicat*	2	0.9	4	1.9	7	3.3	13	6.1	20	9.3	46	21.5
revisit*	0	0.0	1	0.5	4	1.9	8	3.7	4	1.9	17	7.9
Other	0	0.0	3	1.4	2	0.9	9	4.2	15	7.0	29	13.6
Total	5	2.4	19	8.9	27	12.6	51	23.8	112	52.3	214	100

* *N*=number of articles; %=percentage of articles
**It is worth noting that some timespans are 11 years and some are 10 years.

Drawing upon the replication labels, as shown in Table 1, 11 categories were identified. Overall, there was an increase in the use of all categories; “close replication”, “quasi-replication”, “revisit*”, and “exten*” enjoyed more fluctuation within the period, though. The results also demonstrated that exact, approximate, and conceptual replications were non-existent before 2002. From 2002 to the end of 2011, true replication, conceptual replication, and “exten*” remained constant while partial replication, “replicat*” (studies that did not use any qualifiers for the type of replication), and revisit* almost doubled. Since 2012, the authors have shown more tendency to mention the types in their studies. The term “replicat*” without any qualifiers was the most frequently used label over these decades (N = 20, 9.3%). Taken altogether, the lion’s share of studies (N = 46, 21.5%) was labeled “replicat*”; only 1.4% (N = 3) and 2.3% (N = 5) of studies, however, were found with the labels of “close replication” and “exact replication”, respectively. Figure 2 illustrates these findings.

Figure 2

Use of Different Replication Terms in Studies



The second research question dealt with the methodological orientation of replications. As observed in Table 2, during the first two decades of publication (1970-1990), quantitative research was the dominant category (N = 5, 2.3% and N = 18, 8.4%, respectively) with only one case of qualitative research and no mixed-methods study. However, the publication of qualitative research soared in the following decades (from 2002 onwards, N = 21, 9.8% and N = 29, 13.6%, in order). Unlike qualitative research, mixed-methods replications increased only slightly by the end of the period (from 0% to a total of 10, 4.7%). The highest occurrences of quantitative studies (N = 76, 35.5%) were

observed during the last decade (2012-2021). All in all, there is a growing trend in the use of all three methodological orientations over the decades, with the predominance of quantitative research. One can also notice the overall domineering weight given to quantitative research (*N* = 150, 70.1%), compared to qualitative research (*N* = 54, 25,2%) and mixed-methods orientations (*N* = 10, 4.7%) in RSs.

Table 2
Methodological Orientations Over 1970-2021

Methodological Orientation	Publication Year											
	1970-1980		1981-1990		1991-2001		2002-2011		2012-2021		Total	
	*N***	%	*N*	%	*N*	%	*N*	%	*N*	%	*N*	%
Quan*	5	2.3	18	8.4	22	10.3	29	13.6	76	35.5	150	70.1
Qual	0	0.0	1	0.5	3	1.4	21	9.8	29	13.6	54	25.2
Mixed	0	0.0	0	0.0	2	0.9	1	0.5	7	3.3	10	4.7
Total	5	2.3	19	8.9	27	12.6	51	23.8	112	52.3	214	100

*QUAN = quantitative studies; QUAL = qualitative studies; MIXED = mixed-methods studies
** *N* = number of articles; % = percentage of articles.

The results of the third research question concerning the research trends of the studies, as reported in Table 3, indicated that the number of studies focusing on "instructional effects", "learner behaviors", "learner cognition", "theoretical essays, narratives, and descriptions", "second language acquisition (SLA)", "instructional materials", and "assessment and validation" had a rising trend in recent years (from 2011 to 2021). However, "teacher practices", "curriculum and program development", and "reviews, surveys, meta-studies" were the least replicated categories (*N* = 1, 0.5%). As also illustrated in Figure 3, SLA was by far the most commonly replicated area of research (*N* = 91, 42.5%). Except for "learner behavior" (*N* = 29, 13.6%), other research trends were within a limited range of 9 (4.2%) to 21 (9.8%) in frequency.

Table 3

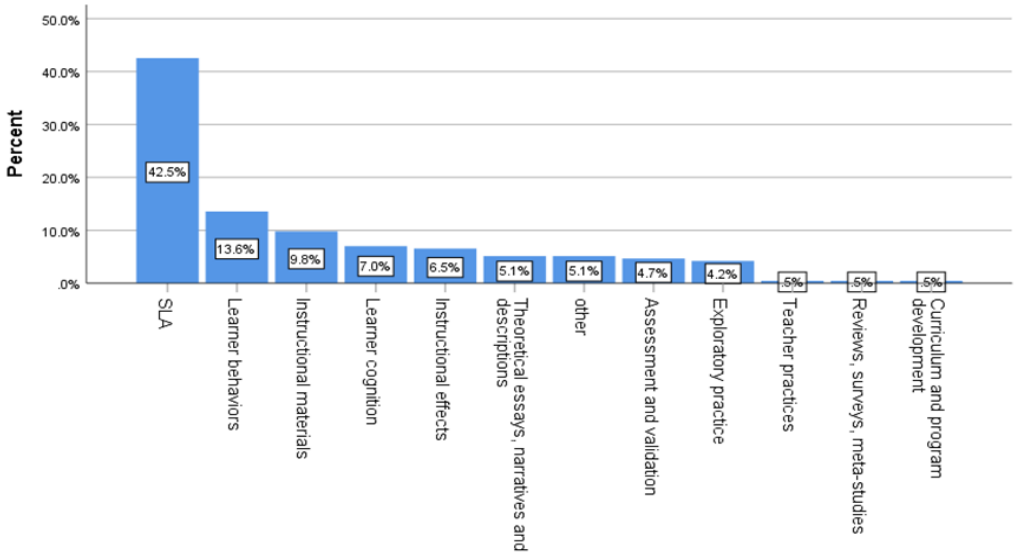
Trends of Replication Studies over Time (1970-2021)

	Publication Year											
	1970-1980		1981-1990		1991-2001		2002-2011		2012-2021		Total	
Research Trends	*N**	%	*N*	%	*N*	%	*N*	%	*N*	%	*N*	%
Instructional effects	0	0.0	1	0.5	2	0.9	4	1.9	7	3.3	14	6.5
Exploratory practice	1	0.5	1	0.5	0	0.0	0	0.0	7	3.3	9	4.2
Learner behaviors	0	0.0	4	1.9	3	1.4	6	2.8	16	7.5	29	13.6
Learner cognition	1	0.5	0	0.0	3	1.4	5	2.3	6	2.8	15	7.0
Theoretical essays, narratives, and descriptions	0	0.0	0	0.0	1	0.5	7	3.3	3	1.4	11	5.1
Teacher practices	0	0.0	0	0.0	0	0.0	1	0.5	0	0.0	1	0.5
Second language acquisition	2	0.9	9	4.2	12	5.6	18	8.4	50	23.4	91	42.5
Instructional materials	1	0.5	2	0.9	5	2.3	3	1.4	10	4.7	21	9.8
Assessment and validation	0	0.0	2	0.9	0	0.0	1	0.5	7	3.3	10	4.7
Curriculum and program development	0	0.0	0	0.0	0	0.0	1	0.5	0	0.0	1	0.5
Reviews, surveys, meta-studies	0	0.0	0	0.0	0	0.0	0	0.0	1	0.5	1	0.5
Other	0	0.0	0	0.0	1	0.5	5	2.3	5	2.3	11	5.1
Total	5	2.3	19	8.9	27	12.6	51	23.8	112	52.3	214	100

* *N* = number of articles; % = percentage of articles

Figure 3

Research Focuses of Replication Studies



The fourth research question investigated the authorship orientation of the studies. Table 4 shows a striking hike in both co-authorship (from *N* = 15, 7% to *N* = 45, 21%) and multiple authorship (from *N* = 15, 7% to *N* = 39, 18.2%) across the last two decades (2002-2021); while sole authorship increased more gently (from *N* = 21, 9.8% to *N* = 28, 13.1%) during these two decades. The findings indicated that overall single-authored, co-authored, and multiple-authored articles were almost similarly frequent (*N* = 72, 33.6%, *N* = 75, 35.05%, and *N* = 67, 31.3%, respectively).

Table 4

Authorship of Replication Studies over Time (1970-2021)

	Publication Year											
	1970-1980		1981-1990		1991-2001		2002-2011		2012-2021		Total	
Authorship	*N**	%	*N*	%	*N*	%	*N*	%	*N*	%	*N*	%
Single authorship	2	0.9	10	4.7	11	5.1	21	9.8	28	13.1	72	33.6
Co-authorship	0	0.0	2	0.9	13	6.1	15	7.0	45	21.0	75	35.0
Multiple authorship	3	1.4	7	3.3	3	1.4	15	7.0	39	18.2	67	31.3
Total	5	2.3	19	8.9	27	12.6	51	23.8	112	52.3	214	100

**N* = number of articles; % = percentage of articles.

The last research question addressed the most influential researchers conducting replication in AL. To this end, those with more than 1000 citations were considered major contributors to the field.

Table 5
Ten Most Influential Researchers Conducting Replication in the Field

Rank	Replicator	Citation Count
1	Rod Ellis	123991
2	Ellen Bialystok	72919
3	Merrill Swain	35770
4	Bill VanPatten	21155
5	Robert M. DeKeyser	18118
6	Kim McDonough	6633
7	Emma Marsden	6622
8	Kara Morgan-Short	3507
9	Paul Stapleton	2555
10	Martha C. Pennington	2384

Among the researchers conducting replication, as is shown in Table 5, authors like Rod Ellis and Ellen Bialystok are of the most significance to the field of AL. Citation counts of these 10 replicators range from 1000-140000. According to 100-most-cited applied linguists, Rod Ellis is the 19th most influential author in AL (Exaly, 2023).

Discussion

The current study aimed to provide a comprehensive review of the RSs over half a century by addressing how replication research is labeled, what methodologies are used, what research areas are of most interest, what authorship orientation of replicators is, and who the most influential researchers conducting replication in the field of AL are. Applied linguistics, as a subfield of linguistics, likely faces a replication crisis (Sönning & Werner, 2021). Our aim was to bring more insight into the matter.

Replication Labels

The results on how the authors labeled their RSs indicated a gradual increase in the use of various word families of “replication”, “extension”, and “revision” over the decades, suggesting a greater inclination of AL replicators to introduce their research as replication and explicitly mention its type. It seems that AL researchers feel less intimidated and do not consider labeling their work as “replication” a threat to their academic stance. However, as Irvine (2021) observed, scholars still refrain from using the word “replication” to shield their research from the connotations attached to it. Consequently, many RSs are covert since “replication” is used neither in the titles nor throughout the text (Abbuhl, 2018; Irvine, 2021; Marsden et al., 2018).

Nevertheless, using more replication labels by AL researchers over time, as observed in the current study, reveals that AL researchers have started to alter their attitude toward this type of research. Although there is still a long way to go for the general acceptance of replication research in AL, the field has witnessed a shift in the culture of AL research from overemphasizing research novelty to raising awareness about replication practices. One of the reasons for this shift might stem from the growing popularity of open science across different fields, including AL, in recent years. As a matter of fact, open science facilitates accessibility to research data and enhances methodological transparency, which paves the way for undertaking replication research (Plonsky, 2023). Moreover, a number of studies (e.g., Del Ben et al., 2022; Kobrock & Roettger, 2023; Marsden, 2020; Marsden et al., 2018) have directed AL researchers' attention to the utility of RSs, boosting researchers' tendency to perform such research.

Among different replication labels, the terms "true replication", "partial replication", and "replicat*" have been used more frequently, especially since 2002. The least frequent labels were close and exact replications. This can be attributed to the difficulty associated with exactly replicating a given study, keeping all the conditions of the initial study intact. Low methodological transparency of original studies, which hinders their exact or close replication, may be another demotivating factor. It can also be stated that due to its low academic payoff, AL researchers are either reluctant to carry out exact and close replication or tend to hide the exact or close replication nature of their research (Abbuhl, 2018; Irvine, 2021; Marsden et al., 2018).

Clearly, a wide range of replication nomenclatures are used by the authors, and there are some conceptual overlaps between them (e.g., close and approximate replication). This indicates the need for a well-established replication framework with precise replication labels, their definitions, and operationalizations in AL. This would enhance the researchers' tendency to include not only the word "replication" but its type in their works.

Methodological Orientation

The analysis of results on the methodological orientation of RSs in the past fifty years revealed that most of the replicated studies tended toward the quantitative paradigm. The reason may lie in the fact that many scholars in AL attempt to familiarize researchers with replication (e.g., Abbuhl, 2018; Porte, 2012; Porte & McManus, 2019) accentuate the replication of quantitative studies, especially the experimental ones. In these discussions, there are always mentions of concepts related to experiments and other quantitatively-oriented research, such as testing hypotheses and effect sizes. In a similar vein, in other disciplines like psychology, replication is confined to quantitative studies, and many frameworks proposed to evaluate replication success (e.g., Heirene, 2021; Held et al., 2022; Ioannidis & Trikalinos, 2007; Kelly, 2019) have centered on comparing the initial and replicated effect sizes.

Interestingly, one-quarter of the corpus in the current study were inclined toward the qualitative paradigm, which further supports the replicability of this research approach. One such example is Mullock (2006) which has qualitatively (partially) replicated a mixed-

methods paper by Gatbonton (2000). The objective of the replication was to examine whether the patterns of pedagogical knowledge found in the initial study would also be found in other teaching situations. Methodologically, the initial research was a mixed-methods one, yet its replication was qualitative research. The author of this specific example has only used the initial paper's stimulated recall methodology and coding features. Furthermore, it has extended the initial study by incorporating a number of key changes such as wider databases and different characteristics of the participants. One of the main differences between the replication and the initial study was the nature of the classes under investigation. The classes in the initial study were specifically designed for the research; nonetheless, its replication counterpart examined intact classes with their teachers. Moreover, the proficiency level of the participants in Gatbonton (2000) was lower-intermediate whereas Mullock (2006) considered classes with different proficiency levels. All in all, Mullock (2006) has studied the same issue on a broader and deeper level. Finally, the findings of the replication supported the main findings of the initial study and attributed the minor differences to the sample and context. This particular example corroborates the discussion put forward by Makel et al. (2022) stating that replication aligns with the transparency and intentionality of a qualitative study. Also, minor discrepancies may exist due to various reasons such as the settings and breadth of the investigation (Makel et al., 2022).

Generally, qualitative research in AL involves intricate decisions about how to measure, operationalize, and interpret linguistic phenomena (Coretta et al., 2023); therefore, there is a deeply rooted debate about whether qualitative research can -and should- be replicated. To settle the hot debate over the replicability of this paradigm, there is a need for some modifications as to how to assess replication in this approach. For instance, the notions of transferability and comparability in replication should be reconceptualized in a way to be more compatible with this approach (Abbuhl, 2018). To be more precise, the focus should be shifted to what aspects of the original study can be transferred to the replicated study and the ways and reasons the cases might be transferable (Porte & Richards, 2012). As obvious, qualitative studies are context-based and heavily rely on the researcher for data interpretation (Coretta et al., 2023; Silberzahn et al., 2018); hence, by replication, we can explore whether the subjective results specific to one context are observed in other similar contexts or not. Furthermore, numerous alternative interpretations of the same phenomenon are possible, and the replicators can delve into the neglected areas of that phenomenon. Therefore, we recommend that qualitative researchers do not refrain from conducting replication. It seems that conceptual replication presents a welcoming ground for replicating qualitative studies because it involves using different methods or data sources to test the same phenomenon, thereby enhancing the validity of the initial findings and providing a more comprehensive and nuanced understanding of the phenomenon. Nassaji (2021) also contends for measuring effect sizes for qualitative research, which encourages replication in the qualitative paradigm and paves the way for evaluating replication success. It is worth noting that the results of this study do not guarantee that the initial studies were qualitative in nature. Future studies can explore

whether the initial studies, which were replicated using a qualitative approach, employed a qualitative methodology themselves.

The results also demonstrate a growing interest in mixed-methods replication. While in the first three decades, this methodology in replication was missing, in the past two decades some mixed methods replications, though very few, have been conducted. One reason for the emergence of mixed-methods replication can be attributed to an increase in the number of original mixed-methods studies over these decades (Magnan, 2006). Moreover, as the mixed-methods approach involves the qualitative paradigm, these results support our previous claim that qualitative research is replicable.

It is noteworthy that the results suggest the greater tendency to replicate studies in quantitative, qualitative, and mixed-method approaches over the past decades. One possible reason for such an increase is a rise in the number of original articles in these paradigms over this period of time, which requires further research.

Overall, our study's findings suggest that the landscape of replication extends beyond distinct research paradigms. Replication exhibits robustness across all three paradigms, notably thriving within the quantitative approach. Qualitative and mixed-methods paradigms also show signs of growth, calling for heightened attention in future lines of AL replication research.

Research Trends, Authorship Orientation, and Influential Replication Researchers

Our results on the research trends in RSs signified that various areas of inquiry were of interest to the researchers for replication. SLA (42%), described as "studies whose primary function is to better understand how second languages are acquired" (Stapleton & Shao, 2017, p. 357), was by far the most frequently explored research trend, followed by learner behavior (13.6%). Replications in other trends (like “teacher practices”, “curriculum and program development”, and “reviews, surveys, meta-studies”) were less than 10%. Although there are not many original studies on SLA, as observed by Stapleton and Shao (2017), apparently this area could attract replicators’ attention to a great extent. Arguably, the SLA area is claimed to be more theory-oriented (Jordan, 2005), which has great potential for experimental research; hence, this subject matter yields itself to replication more readily (Abbuhl, 2018). Moreover, the replicators seemed to be more inclined to assess the previous findings on how second languages are learned so as to contribute to the existing SLA theories.

The second most investigated area was “learner behavior”, studies focusing on characteristics like self-reflection, dominance behavior, beliefs about learning, and self-monitoring. These topics are usually approached through the qualitative paradigm. Although it is generally believed that qualitative research is not replicable, the results of this study suggest that the researchers tend to replicate qualitative studies on these topics. To gain a deeper understanding of replication research in AL, future studies can precisely explore whether the source of differences across these areas stems from their yielding themselves to specific methodological orientations.

With respect to the pattern of authorship among AL replicators, it was found that the single-authored, co-authored, and multiple-authored research articles were similarly frequent. Starting from the 1970s, single-authored replication research outnumbered co-authorship and multiple authorship, but it underwent a decreasing trend toward 2021, and it was the least frequent one in 2012-2021. On the contrary, both co-authorship and multiple authorship demonstrated exponential growth and tripled proportionally.

Subsequently, from a global perspective, it is evident that these patterns of authorship are successfully comparable to the global trend of increasing collaboration among authors (Henriksen, 2018; Hyland, 2015). It can also be argued that since most of the RSs identified in our corpus are quantitative, and empirical quantitative studies are more likely to be conducted by two or more authors due to the simplicity of the division of labor in these studies (Borry et al., 2006; Henriksen, 2018), we expect to witness more co-authorship and multiple authorship than single authorship.

In greater depth, these patterns of authorship are comparable to those of Riazi et al. (2020) regarding the propensity of researchers to engage in more collaborative research in the EAP context, indicating an increasing trend of collaborative authorship in EAP studies in RSs. Elaborating on the benefits of collaboration in research, Henriksen (2018, p. 1401) believes that "research production by sharing tasks and profiting from each other's expertise" is optimized, and such teamwork lowers "the risk of having a no-publishing period for scholars." Research collaboration would also lead to conducting more meticulous research studies as researchers pool their resources together. Being aware of these advantages, researchers are more willing to conduct research in collaboration with others. This collaborative behavior in research is indicative of changes in the research culture (Henriksen, 2018), which has also influenced replication researchers.

This study also identified the most influential researchers in the area of AL who have conducted replications. It is assumed that the more a researcher is cited, the more impact they have on a scientific discipline (Robinson & Adler, 2004). We found that some of the most influential people such as Rod Ellis and Merrill Swain have undertaken RSs, which implies that replication is an accepted craft for prominent scholars in the field with a critical and strong stance. Having such leading figures carrying out replication leaves no doubt that it needs to be practiced by professional AL researchers more often. Even novice researchers can benefit from replicating the works of pioneering scholars with rigorous methodology because over the course of replication research, not only do they contribute to the field by assessing previous findings, but they also learn a lot about how to design decent original research (Moreau & Wiebels, 2023).

Conclusion and Recommendations

Replication has gained momentum in scientific fields as it reminds us of 'human fallibility' and sampling and measurement errors that may lead to false positives, false negatives, and other potential errors in the results reported (Porte & McManus, 2019). Indeed, replication allows the readers of the research articles to question and calibrate their certainty in the robustness of the claimed findings and to "repeat it in a particular way to

establish its stability in nature and eliminate the possible influence of artifacts or chance findings" (Porte, 2012, p. 4). Due to the significance of replication attached to any scientific field including AL, this study undertook a systematic review of the RSs conducted over 50 years in order to reveal the current status of replication research in AL and make some recommendations on how to promote its status in the field.

The findings of the current study revealed that before 2002, there was no mention of some types of replications. Moreover, since 2012, scholars have been more encouraged to self-label their papers without being afraid of damaging the prestige of their projects. Quantitative, qualitative, and mixed-methods paradigms were respectively the most frequently observed methodological orientations in RSs. During the last two decades, the number of qualitative studies soared; however, RSs using a mixed-methods approach are still rare. SLA was by far the most researched theme while other research trends were marginalized. Furthermore, research replication has witnessed a surge of co-authorship and multiple authorship since 2012 to the detriment of single authorship. It was found that the most prominent figures in the realm of AL, like Rod Ellis, officially recognize replication research and its value and have made some attempts to perform it. It can be concluded that over the decades, replication research has been embraced more by the field, though there is a long way to go for its general acceptance.

There are some caveats to the obtained results as this study is not free from limitations. Our search strategy could not identify studies with covert replication, i.e., studies in which the word families of "replication", "extension", or "revision" were not included in the title, abstract, or body. To tackle this issue, all articles in the field should go through an in-depth analysis. As it is a herculean task to perform such assessments on large-scale data, future studies can confine themselves to a specific journal or research trend. This study did not provide replication rates in AL in terms of the variables under investigation. For instance, it is not clear whether the predominance of quantitative RSs in the last period is because of replicators' greater tendency to assess the findings of original quantitative studies or due to a substantial rise in original quantitative studies in AL in this period. Interested researchers can investigate this issue in their future works. The present study did not compare the initial and replicated studies to explore whether the replication types claimed by the replicators in their studies truly matched and whether the methodological orientations of the initial and replicated studies were similar. These comparisons can provide a clearer picture of replication research in AL. In this regard, another fruitful line of research could be an investigation of the extent to which the results of the initial studies and those of their replications are similar or different. However, having very few replications in some areas of study is concerning. The negligence of some areas of study could be due to the scholars' reluctance to conduct replication in these areas or the theme of research being unreplicable. Either way, the reason for this negligence must be investigated. The other point meriting further research is examining whether studies uploading their materials in open science repositories enjoy a greater chance of being replicated or not. Moreover, there are other kinds of replication, such as research that checks computational reproducibility (i.e., doing the same analysis on the same data) or

analytical robustness (i.e., doing a different analysis on the same data); therefore, further research is needed to examine whether or not such types of replications exist in AL studies.

We bring the article to an end with some recommendations to all scholars in AL. Our first recommendation for RSs is to include the word "replication" in their works, especially in the titles. This would grab the readers' and researchers' attention to "replication", which enhances researchers' incentive to design such research studies as they see research articles of this kind being published by journals. Interested researchers' access to RSs would increase and replication rates would be boosted, accordingly (Marsden et al., 2018). It is also recommended to raise the researchers' awareness of different replication types in AL. Given the different nomenclatures of replication and their imprecise conceptualization in the literature (Abbuhl, 2018; Porte & McManus, 2019) and the diversity of replication labels (correctly or erroneously) used by authors in AL studies, there is a need to put forth a comprehensive well-designed framework in AL for replication research. By providing a set of replication types and their detailed definitions, this framework can function as a replication reference. It can also include the goals of replication, the feasibility of replicating the initial study, the availability of resources, and the relevance and utility of the replication to the field and should allow for flexibility and adaptation to specific research contexts. This framework offers a number of benefits: 1) informing AL researchers about the complex process of replication, 2) assisting researchers to make appropriate choices regarding which aspects of the initial study to replicate and which to modify, 3) evaluating the validity and reliability of previous findings in a systematic way, and 4) giving researchers the courage to mention the type of replication in their papers without worrying about their academic stance. Furthermore, by providing a well-established framework for replication in AL, replication research can find its official place in the field, academic incentive systems (e.g., journals) may be more welcoming to RSs, and AL researchers may be more inclined to conduct replication.

It seems that not only is there a small number of replications in AL (Marsden et al., 2018), but the existing body of replication is also limited in scope to certain subjects (e.g., SLA) and in methodology to the quantitative approach in the field. Hence, it is proposed that AL researchers interested in replication should not limit themselves to specific areas of inquiry or a particular methodological orientation and should replicate the original studies even in the marginalized research trends in replication (like "teacher practices" and "curriculum and program development") and those with the low-attended methodological orientations in replication (qualitative and mixed-methods approaches). Observing the replicators' growing tendency for collaborative authorship behavior, we also recommend that replication research not be performed in collaboration with one or some of the authors of the original studies although they may provide better insights into the phenomenon. The reason is that this collaboration may lead to bias, and to avoid such bias, independent direct replications are to be conducted. Or else the collaborators should take the right action to minimize this bias. It is also recommended that more prominent scholars in the field undertake replication research so as to consolidate the position of replication in AL.

Academic incentive systems are also responsible for enhancing the status quo of replication. Therefore, higher education institutions should develop a positive image of

replication for both early-career and senior researchers rather than the stigmatized common belief since such research can be an invaluable starting point in these researchers' professionalism (Moreau & Wiebels, 2023). Funding agencies and journals across various research trends in AL can modify their submission guidelines to enable researchers to confidently design RSs using any methodological orientation. For instance, journals can dedicate special issues to such research and encourage replicators to include the word "replication" and the types in the title and body of their works.

Another point meriting a mention is to promote open science practices in AL. By sharing all information about original studies, such as study design, methodology, data collection, coding sheets, and materials in open-access repositories (Del Ben et al., 2022), the original authors can facilitate the accessibility of their colleagues to useful sources for replication. Evidently, such an undertaking can increase the rate of RSs, and positively affect the stance of replication in the field.

Correspondence should be addressed to Mahsa Farahanynia
Email: m.farahanynia85@yahoo.com

Notes:

[1] iris-database.org/details/v8Z33-qw75S

References

Abbuhl, R. (2012a). Practical methods for teaching replication to applied linguistics studies. In G. Porte (Ed.), *Replication research in applied linguistics* (pp. 135-150). Cambridge University Press.

Abbuhl, R. (2012b). Why, when, and how to replicate research. In A. Mackey & S. Gass (Eds.), *Research methods in second language acquisition* (pp. 296-312). Wiley-Blackwell. https://doi.org/10.1002/9781444347340.ch15

Abbuhl, R. (2018). Replication research. In A. Phakiti, A., P. De Costa, L. Plonsky, & S. Starfield (Eds.), *The Palgrave handbook of applied linguistics research methodology* (pp. 145-162). Palgrave MacMillan.

Allwright, D. (2003). Exploratory practice: Rethinking practitioner research in language teaching. *Language Teaching Research*, *7*(2), 113-141. https://doi.org/10.1191/1362168803lr118oa

Amjad, T., Daud, A., & Aljohani, N. R. (2018). Ranking authors in academic social networks: A survey. *Library Hi Tech, 36*(1), 97-128. https://doi.org/10.1108/LHT-05-2017-0090

Bochynska, A., Keeble, L., Halfacre, C. A. E., Casillas, J. V., Champagne, I., Chen, K., Röthlisberger, M., Buchanan, E. M., Roettger, T. B. (2023). Reproducible research practices and transparency across linguistics. MetaArXiv. https://doi.org/10.31222/osf.io/rcews

Borry, P., Schotsmans, P., & Dierickx, K. (2006). Author, contributor or just a signer? A quantitative analysis of authorship trends in the field of bioethics. *Bioethics, 20*(4), 213-220. https://doi.org/10.1111/j.1467-8519.2006.00496.x

Casanave, C. (2012). Heading in the wrong direction? A response to Porte and Richards. *Journal of Second Language Writing, 21*(3), 296-297. https://doi.org/1060374312000422

Coretta, S., Casillas, J. V., Roessig, S., Franke, M., Ahn, B., Al-Hoorie, A. H., Al-Tamimi, J., Alotaibi, N. E., AlShakhori, M. K., Altmiller, R. M., Arantes, P., Athanasopoulou, A., Baese-Berk, M. M., Bailey, G., Sangma, Ch. B. A., Beier, E. J., Benavides, G. M., Benker, N., BensonMeyer, E. P., … & Roettger, T. B. (2023). Multidimensional signals and analytic flexibility: Estimating degrees of freedom in human speech analyses. *Advances in Methods and Practices in Psychological Science*, *6*(2). https://doi.org/10.1177/25152459231162567

Creswell, J. W. (2014). *Research design. qualitative, quantitative and mixed method approaches* (4th ed.). SAGE Publications.

Del Ben, R., Brouillard, M., Gonzalez-Barrero, A.M., Killam, H., Kremin, L.V., Quirk, E., Sander-Montan, A., Schott., E., Tsui, R., & Byers-Heinlein, K. (2022). How open science can benefit bilingualism research: A lesson in six tales. *Bilingualism: Language and Cognition,25*(5), 913-920. https://doi.org/10.1017/S1366728922000256

Denscombe, M. (2008). Communities of practice: A research paradigm for the mixed methods approach. *Journal of Mixed Methods Research, 2*(3), 270-283. https://doi.org/10.1177/1558689808316807

Egbert, J. (2007). Quality analysis of journals in TESOL and applied linguistics. *TESOL Quarterly, 41*(1), 157-171. https://www.jstor.org/stable/40264335

Exaly (2023). Applied Linguistics: Top Authors. Retrieved on May 9, 2023, from https://exaly.com/journal/19954/applied-linguistics/top-authors

Gatbonton, E. (2000). Investigating experienced ESL teachers' pedagogical knowledge. *Canadian Modern Language Review, 56*(4), 585-616. https://doi.org/10.3138/cmlr.56.4.585

Golden, M. (1995). Replication and non-quantitative research. *Political Science and Politics*, 28(3), 481-483. https://doi.org/10.2307/420313

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

Heirene, R. M. (2021). A call for replications of addiction research: Which studies should we replicate and what constitutes a 'successful' replication? *Addiction Research & Theory, 29*(2), 89-97. https://doi.org/10.1080/16066359.2020.1751130

Held, L., Micheloud, C., & Pawel, S. (2022). The assessment of replication success based on relative effect size. *The Annals of Applied Statistics, 16*(2), 706-720. https://doi.org/10.1214/21-AOAS1502

Henriksen, D. (2018). What factors are associated with increasing co-authorship in the social sciences? A case study of Danish economics and political science. *Scientometrics, 114*(3), 1395-1421. https://doi.org/10.1007/s11192-017-2635-0

Hiver, P., Al-Hoorie, A. H., Vitta, J. P., & Wu, J. (2021). Engagement in language learning: A systematic review of 20 years of research methods and definitions. *Language Teaching Research, 28*(1), 201-230. https://doi.org/10.1177/13621688211001289

Hyland, K. (2015). *Academic publishing: Issues and challenges in the construction of knowledge.* Oxford University Press.

Ioannidis, J. P., & Trikalinos, T. A. (2007). An exploratory test for an excess of significant findings. *Clinical Trials, 4*(3), 245-253. https://doi.org/10.1177%2F1740774507079441

Irvine, E. (2021). The role of replication studies in theory building. *Perspectives on Psychological Science, 16*(4), 844-853. https://doi.org/10.1177/1745691620970558

Isbell, D. R., Brown, D., Chen, M., Derrick, D. J., Ghanem, R., Gutierrez Arvizu, M. N., Schnur, E., Zhang, M., & Plonsky, L. (2022). Misconduct and questionable research practices: The ethics of quantitative data handling and reporting in applied linguistics. *The Modern Language Journal. 106*(1), 172-195. https://doi.org/10.1111/modl.12760

John, L., Loewenstein, G., & Prelec, D. (2012). Measuring the prevalence of questionable research practices with incentives for truth telling. *Psychological Science, 23*(5), 524-532. https://doi.org/10.1177%2F0956797611430953

Jordan, G. (2005). *Theory construction in second language acquisition*. John Benjamins.

Jung, U. O. H. (2004). Paris in London revisited or the foreign language teacher's topmost journals. *System, 32*(3), 357-361.

Kelly, C. D. (2019). Rate and success of study replication in ecology and evolution. *PeerJ, 7*. https://doi.org/10.7717/peerj.7654

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

King, K., & Mackey, A. (2016). Research methodology in second language studies: Trends, concerns, and new directions. *Modern Language Journal, 100*(s1), 209-227. https://doi.org/10.1111/modl.12309

Kobrock, K., & Roettger, T. B. (2023). Assessing the replication landscape in experimental linguistics. *Glossa Psycholinguistics, 2*(1). http://dx.doi.org/10.5070/G6011135

Language Teaching Review Panel. (2008). Replication studies in language learning and teaching: Questions and answers. *Language Teaching, 41*(1), 1-14. https://doi.org/10.1017/S0261444807004727

Larson-Hall, J., & Plonsky, L. (2015). Reporting and interpreting quantitative research findings: What gets reported and recommendations for the field. *Language Learning, 65*(s1), 127-159. https://doi.org/10.1111/lang.12115

Lykken, D. T. (1968). Statistical significance in psychological research. *Psychological Bulletin, 70*(3), 151-159. https://psycnet.apa.org/doi/10.1037/h0026141

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

Magnan, S.S. (2006). From the editor: The MLJ turns 90 in a digital age. *The Modern Language Journal, 90*(1), 1-5. https://www.jstor.org/stable/3588810

Makel, M. C., Meyer, M. S., Simonsen, M. A., Roberts, A. M., & Plucker, J. A. (2022). Replication is relevant to qualitative research. *Educational Research and Evaluation, 27*(1-2), 215-219. https://doi.org/10.1080/13803611.2021.2022310

Makel, M., & Plucker, J. (2014). Facts are more important than novelty: Replication in the education sciences. *Educational Researcher, 43*(6), 304-316. https://doi.org/10.3102/0013189X14545513

Makel, M., Plucker, J., & Hegarty, B. (2012). Replications in psychology research: How often do they really occur? *Perspectives on Psychological Science, 7*(6), 532-542. https://doi.org/10.1177%2F1745691612460688

Markee, N. (2017). Are replication studies possible in qualitative second/foreign language classroom research? A call for comparative re-production research. *Language Teaching, 50*(3), 367-383. https://doi.org/10.1017/S0261444815000099

Marsden, E. J. (2020). Methodological transparency in applied linguistics and its consequences for the quality and scope of research. In J. McKinley, & H. Rose (Eds.), *Routledge handbook of research methods in applied linguistics* (pp. 15-28). Routledge.

Marsden, E., & Mackey, A. (2014). IRIS: A new resource for second language research. *Linguistic Approaches to Bilingualism, 4*(1), 125-130. https://doi.org/10.1075/lab.4.1.05mar

Marsden, E., Mackey, A., & Plonsky, L. (2016). The IRIS repository: Advancing research practice and methodology. In A. Mackey, & E. Marsden (Eds.), *Advancing methodology and practice: The IRIS repository of instruments for research into second languages* (pp. 1-21). Routledge. https://doi.org/10.4324/9780203489666

Marsden, E., & Morgan-Short, K. (2023) (Why) are open research practices the future for the study of language learning? *Language Learning, 73*(s2), 344-387. https://doi.org/10.1111/lang.12568

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. https://doi.org/10.1111/lang.12286

Matsuda, P. (2012). On the nature of second language writing: Replication in a postmodern field. *Journal of Second Language Writing, 21*(3), 300-302. http://dx.doi.org/10.1016%2Fj.jslw.2012.05.006

McHugh, M. L. (2012). Interrater reliability: The kappa statistic. *Biochemia Medica, 22*(3), 276-282. https://pmc.ncbi.nlm.nih.gov/articles/PMC3900052/

McManus, K (2021). 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 https://doi.org/10.1017/S0272263121000838

Moreau, D., & Wiebels, K. (2023). Ten simple rules for designing and conducting undergraduate replication projects. *PLoS Computer Biology, 19*(3). https://doi.org/10.1371/journal.pcbi.1010957

Mullock, B. (2006). The pedagogical knowledge base of four TESOL teachers. *The Modern Language Journal, 90*(1), 48-66. https://doi.org/10.1111/j.1540-4781.2006.00384.x

Nassaji, H. (2012). Significance tests and generalizability of research results: A case for replication. In G. Porte (Ed.), *Replication research in applied linguistics* (pp. 92-115). Cambridge University Press.

Nassaji, H. (2021). Effect sizes in quantitative and qualitative research. *Language Teaching Research, 25*(5), 681-684. https://doi.org/10.1177%2F13621688211040882

Nosek, B.A., & Errington, T. M. (2020). The best time to argue about what a replication means? Before you do it. *Nature, 583*, 518-520. https://doi.org/10.1038/d41586-020-02142-6

Pashler, H., & Harris, C. (2012). Is the replicability crisis overblown? Three arguments examined. *Perspectives on Psychological Science, 7*(6), 531-536. https://doi.org/10.1177%2F1745691612463401

Plonsky, L. (2023). Sampling and generalizability in Lx research: A second-order synthesis. *Languages, 8*(1), 75. https://doi.org/10.3390/languages8010075

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

Polio, C., & Gass, S. (1997). Replication and reporting: A commentary. *Studies in Second Language Acquisition, 19*(4), 499-508. https://www.jstor.org/stable/44487991

Popper, K. R. (1959). *The logic of scientific discovery.* Hutchinson.

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

Porte, G., & McManus, K. (2019). *Doing replication research in applied linguistics.* Routledge. https://doi.org/10.4324/9781315621395

Porte, G., & Richards, K. (2012). Focus article: Replication in second language writing research. *Journal of Second Language Writing, 21*(3), 284-293. https://doi.org/10.1016/j.jslw.2012.05.002

Riazi, A. M., Ghanbar, H., & Fazel, I. (2020). The contexts, 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*. https://doi.org/10.1016/j.jeap.2020.100925

Robinson, L. M., & Adler, R. D. (2004). Business research in eight business disciplines. *Journal of College Teaching & Learning, 1*(2), 69-80. https://doi.org/10.19030/tlc.v1i2.1916

Santos, T. (1989). Replication in applied linguistics research. *TESOL Quarterly, 23*(4), 699-702. https://doi.org/10.2307/3587548

Sasaki, M. (2012). An alternative approach to replication studies in second language writing: An ecological perspective. *Journal of Second Language Writing, 21*(3), 303-305. https://doi.org/10.1016/j.jslw.2012.05.004

Schmidt, S. (2009). Shall we really do it again? The powerful concept of replication is neglected in the social sciences. *Review of General Psychology, 13*(2), 90-100. https://doi.org/10.1037/a0015108

Schofield, J. (2002). Increasing the generalizability of qualitative research. In A. Huberman, & M. Miles (Eds.), *The qualitative researcher's companion* (pp. 171-203). SAGE. https://doi.org/10.4135/9781412986274

Silberzahn, R., Uhlmann, E. L., Martin, D. P., Anselmi, P., Aust, F., Awtrey, E., Bahník, Š., Bai, F., Bannard, C., Bonnier, E., Carlsson, R., Cheung, F., Christensen, G., Clay, R., Craig, M. A., Dalla Rosa, A., Dam, L., Evans, M. H., Flores Cervantes, I., Fong, N., … & Nosek, B. A. (2018). Many analysts, one data set: Making transparent how variations in analytic choices affect results. *Advances in Methods and Practices in Psychological Science*, *1*(3), 337-356. https://doi.org/10.1177/2515245917747646

Sönning, L., & Werner, V. (2021). The replication crisis, scientific revolutions, and linguistics. *Linguistics*, *59*(5), 1179-1206. https://doi.org/10.1515/ling-2019-0045

Stapleton, P., & Shao, Q. (2017). Research in language teaching over two decades: A retrospective of the first 20 volumes of Language Teaching Research. *Language Teaching Research, 22*(3), 350-369. https://doi.org/10.1177/1362168816688929

Vandergrift, L., & Cross, J. (2017). Replication research in L2 listening comprehension: A conceptual replication of Graham & Macaro (2008) and an approximate replication of Vandergrift & Tafaghodtari (2010) and Brett (1997). *Language Teaching, 50*(1), 80-89. https://doi.org/10.1017/S026144481500004X

Weber, M., & Campbell, C. M. (2004). In other professional journals. *Modern Language Journal*, *88*(1), 130-139. https://doi.org/10.1111/0026-7902.00010

Zou, D., Luo, S., Xie, H., & Hwang, G. (2020). A systematic review of research on flipped language classrooms: Theoretical foundations, learning activities, tools, research topics and findings. *Computer Assisted Language Learning, 35*(8) 1811-1837. https://doi.org/10.1080/09588221.2020.1839502

Appendix A
Research Focuses in the Field of Language Second Language Teaching and Learning
(Adopted from Stapleton & Shao, 2017, p. 357)

Main Focus	Descriptor
Instructional effects	Refers to studies whose primary function is to describe (often innovative) methods, techniques or approaches that lead to improved L2 performance of students.
Teacher cognition	Refers to studies whose primary function is to investigate a wide range of elements including the teacher's attitude, beliefs, confidence, self-image, and emotional disposition, among other similar traits, and how their cognition impacts on their teaching practices. The focus is on what teachers think.
Exploratory practice	Refers to articles that report 'practitioners working to understand what they want to understand …by using normal pedagogical practices as investigative tools.' (Allwright, 2003, p. 127)
Teacher education	Refers to studies that investigate language teacher education programs and life-long career developmental programs. Usually, novice or student teachers' improvement in their teaching practices through those programs is a focus; i.e. the focus is often on novice or pre-service teachers' practices and what they learn and how they grow.
Learner behaviors	Refers to learners' behaviors inside and outside the classroom including such traits as their level of participation, degree of dominance behavior, learning style, learning strategies, degree of self-reflection, self-monitoring and attitudes towards and beliefs about learning. The focus is on what learners do.
Learner cognition	Refers to studies that primarily focus on the learner's attitudes, beliefs, confidence, self-image, and emotional disposition, among other similar traits, and how this impacts their learning behaviors and patterns. The focus is on what learners think.
Theoretical essays, narratives, and descriptions	Includes articles that propose new theories, frameworks, models, methods, descriptions of practices, advocacy, narratives of experience teaching/researching, or critiques.
Teacher practices	Refers to studies that primarily investigate teachers' practices often via observations and interviews in order to better understand how they conduct classes. The focus is on what teachers do.

Second language acquisition (SLA) mechanism	Refers to studies whose primary function is to better understand how second languages are acquired (in contrast to what pedagogical techniques are best used).
Instructional materials	Refers to studies that describe or investigate materials such as textbooks or wordlists for pedagogical purposes.
Assessment	Refers to studies that primarily focus on the role of exams or tests in language teaching.
Curriculum and program development	Refers to studies that describe or investigate language teaching programs or curriculums and their reform.
Reviews, surveys, meta-studies	Includes articles that review a collection of studies or survey the profession or aspect of the profession from a broad perspective.

Appendix B
The Complete List of AL Journals Consulted for Replication Studies

1. Asian Journal of English Language Teaching
2. Annals of Dyslexia
3. Annual Review of Applied Linguistics
4. Applied Language Learning
5. Applied Linguistics
6. Canadian Journal of Experimental Psychology
7. DFL Bulletin
8. Assessing Writing
9. Babel
10. Bilingualism: Language and Cognition
11. CALICO Journal
12. CALL
13. Applied Psycholinguistics
14. Canadian Modern Language Review/La Revue canadienne des langues vivantes
15. Cognition
16. Computers and Composition
17. Educational Action Research
18. Educational and Psychological Measurement
19. Educational Research
20. Educational Researcher
21. Language Teaching Research
22. Linguistics and Education
23. Recall
24. Journal of Educational Psychology
25. Journal of Experimental Education
26. Journal of Experimental Psychology: General
27. The Modern Language Journal
28. Journal of Experimental Psychology: Learning, Memory, and Cognition
29. Journal of Child Language
30. Journal of Memory and Language
31. International Journal of Qualitative Studies in Education
32. Journal of Pidgin and Creole Languages
33. Journal of Pragmatics
34. Journal of Psycholinguistic Research
35. Journal of Second Language Writing
36. Language
37. Language and Cognitive Processes
38. Language Awareness
39. Language Change and Variation
40. Language in Society
41. Language Learning
42. Language Learning & Technology
43. Language Teaching
44. South African Journal of Psychology
47. Educational Technology, Research, and Development
48. Educational Technology Systems
49. ELT Journal
50. English for Specific Purposes
51. English Today
52. English World-Wide: A Journal of Varieties of English
53. Foreign Language Annals
54. French Review
55. Harvard Educational Review
56. Hispania
57. International Journal of Applied Linguistics
58. International Journal of Intercultural Relations
59. International Journal of Lexicography
60. Journal of Experimental Psychology: Human Perception and Performance
61. IRAL
62. JALT Journal
63. Journal of Applied Psychology
64. Journal of Language and Social Psychology
65. Journal of Cross-Cultural Psychology
66. Journal of Educational Measurement
67. Research in the Teaching of English
68. RELC Journal
69. Second Language Research
70. Studies in Second Language Acquisition
71. TESL Reporter
72. Japan Association for Language Teaching (JALT) Journal
73. Language Testing
74. International Review Applied Linguistics in Language Teaching
75. Etudes de Linguistique Appliquee
76. Language and Speech
77. Journal of Multi-lingual and Multicultural Development
78. Anthropological Linguistics
79. Bilingual Research Journal
80. Discourse Processes
81. English Language Teaching Journal
82. The Language Learning Journal
83. Issues in Applied Linguistics
84. TESOL Quarterly
85. Journal of English for Academic Purposes
86. Journal of Language, Identity, and Education
87. Korean Journal of Applied Linguistics
88. Korea TESOL Journal (KOTESOL)
89. Language Assessment Quarterly
90. TESL-EJ