Modeling quality and prestige in applied linguistics journals: A bibliometric and synthetic analysis

Yiran Xu ✉
University of California, Merced, USA
https://orcid.org/0000-0002-9919-1032
yxu72@ucmerced.edu

Jingyuan Zhuang
Pennsylvania State University, University Park, USA
https://orcid.org/0000-0003-1148-9721
juz440@psu.edu

Ryan Blair
University of Arizona, Tuscon, USA
https://orcid.org/0000-0002-9919-1032
rblair1@arizona.edu

Amy I. Kim
Language Testing International, Tarrytown, NY, USA
https://orcid.org/0000-0001-6413-5439
amyk3860@gmail.com

Fei Li
China University of Geosciences, Wuhan, China
https://orcid.org/0000-0002-8438-2194
fei.li@cug.edu.cn

Rachel Thorson Hernández
Georgetown University, Washington, DC, USA
https://orcid.org/0000-0003-4116-5672
thordaughter@gmail.com

Luke Plonsky

Northern Arizona University, Flagstaff, USA

https://orcid.org/0000-0002-5791-1839

lukeplonsky@gmail.com

Abstract

The importance of academic journals in second language (L2) research is evident on at least two levels. Journals are, first of all, central to the process of disseminating scientific findings. Journals are also critical on a professional level as most L2 researchers must publish articles to advance their careers. However, not all journals are perceived as equal; some may be considered more prestigious or of higher quality and may, therefore, achieve a greater impact on the field. It is therefore necessary that we understand the identity and quality of L2 research journals, yet very little research (e.g., Egbert, 2007; VanPatten & Williams, 2002) has considered these issues to date. The current study sought to explore L2 journal identity and quality, and the relationship between these constructs. In order to do so, a database was compiled based on three different types of sources: (1) a questionnaire eliciting L2 researchers' perceptions of the quality and prestige of 27 journals that publish L2 research (N = 327); (2) manual coding of different types of articles (e.g., empirical studies, review papers), data (quantitative, qualitative, mixed), research settings, and authorship patterns (K = 2,024) using the same 27 journals; and (3) bibliometric and submission data such as impact factors, citation counts, and acceptance rates. Descriptive statistics were applied to explore overall quality and prestige ratings as well as publication trends found in each journal. The relationships between those patterns and subjective ratings were also examined. In addition, regression models were built to determine the extent to which perceptions of journal quality and prestige could be explained as a function of journal and article features. We discuss the findings of the study in terms of on-going debates concerning publication practices, study quality, impact factors, journal selection, and the "journal culture" in applied linguistics.

Keywords: Applied Linguistics journals; quality; prestige; perception; publication practices; journal culture

1. Introduction

Scholarly journals are an indispensable source for researchers to access and disseminate empirically based information. As pressure to publish has increased in the world of academia, the need to evaluate and carefully select scholarly journals has increased (Egbert, 2007; West & Rich, 2012). However, doing so can be very challenging especially for less experienced researchers, who may be less

familiar with journal scopes, guidelines, conventions, perceived quality, and publication history. In addition, the lack of sufficient guidance for authors, as well as other stakeholders such as promotion and tenure committees, regarding which scholarly journals to evaluate and select, can complicate matters further. As noted by Suber (2010), "if all journals were equal in prestige, or if all journals were equal in the eyes of their promotion and tenure (P&T) committee, most researchers would happily focus on their research and give very little thought to where it was published" (p. 115). Further complicating this matter is the fact that quantitative indices such as impact factors, which are generated automatically by large organizations (e.g., Thompson Reuters) and search engines such as Google Scholar only reflect partial information about the journals. Little do we know about how scholars perceive journal quality and prestige in applied linguistics.

The current study sought to better understand some of the submission conventions and practices within a set of journals in applied linguistics. We also aimed to gauge researchers' perceptions of the quality and prestige of these journals, both on their own and in relation to existing indices of journal impact as well as other journal characteristics such as citations and (co-)authorship patterns.

2. Journal quality and prestige among applied linguistics journals

The quality and prestige of scholarly journals are likely to play an important role in research and academic decision-making. Journal quality has been suggested to influence the visibility of data, researchers' decisions concerning where to submit their work, professional evaluation, advancement, and promotion, evaluation of scholarly and departmental performance, and the allocation of scholarly and departmental funding (Egbert, 2007; Johnstone, 2009; VanPatten & Williams, 2002; Weiner, 2001). Accordingly, efforts to examine what determines journal quality and prestige have been made by researchers across various fields (e.g., Bhuyan et al., 2020; Egbert, 2007; Haensly et al., 2008; Lee, et al., 2002; VanPatten & Williams, 2002; Yu et al., 2017; Zhao et al., 2017). The importance of selecting certain scholarly journals for review to guide academic research and development and the underdevelopment of methodology to determine journal quality have, in part, driven such inquiries. In addition, many researchers have acknowledged that the aim of studying journal quality should not simply be to rank journals but to be more informative and constructive for researchers in their field.

In the field of applied linguistics, there has seemed to be only a handful of studies examining perceived journal quality to date (e.g., Egbert, 2007; Smith & Lafford, 2009; VanPatten & Williams, 2002). VanPatten and Williams (2002), for example, surveyed 45 faculty members in research universities in the United

States and Canada to obtain perceived quality ratings of fifteen journals in the sub-field of second language acquisition (SLA). Ratings of quality among the 15 journals ranged on scale of 1-5 from 3.1 (*Applied Language Learning*) to 4.9 (*Studies in Second Language Acquisition*). The results also suggested that publications from the fifteen ranked journals had appeared to carry a considerable weight when evaluating tenure candidates. It was noted that such a ranking of journal quality may not be fixed due to the constantly shifting nature of the field and publishing outlets. Consequently, a periodic review of journal quality rankings was recommended.

Egbert (2007), in absence of a set methodology for determining journal quality, examined journal quality in the fields of teachers of English to speakers of other languages (TESOL) and applied linguistics with responses from 300 members of the TESOL Research Interest Section. Journal quality was examined based on citation analysis, rejection rate, time to publication, journal availability and accessibility, and other descriptive measures (i.e., journal target audiences, authors, and perceptions of readers). Egbert found that rather than converging on a unifying standard, respondents provided more than 25 criteria to judge the quality of journals, such as "relevance to context," "editorial board quality," and "my university's criteria." She thus speculated that bibliometrics and a set ranking of journals might not be as useful in TESOL and applied linguistics. However, it may be possible to roughly rank journals based on quality by noting patterns of consistency across multiple measures.

A third study related to journal quality was conducted by Smith and Lafford (2009) on journals in applied linguistics and the subfield of computer-assisted language learning (CALL). CALL scholars were asked to offer opinions on journal quality and their criteria for ranking for promotion and tenure decisions via two surveys. Smith and Lafford (2009) found that CALL scholars had shown a strong preference for journals that publish articles of high quality and relevance and make a significant contribution to CALL. Acceptance rate, impact factor, and circulation, on the other hand, were of less interest to CALL scholars.

To summarize, although journal quality is important in a practical sense, only a small number of studies have sought to gauge researchers' perceptions of the quality of different journals in applied linguistics. There is also little evidence that the samples in this body of work are representative of the population of interest. Further complicating this matter is the fact that researchers are hesitant if not doubtful about the usefulness and applicability of purely perception data. In order to get a clearer view of journal quality and prestige, which is probably best understood as a latent (non-directly observable) variable (see Larsson et al., 2022), multiple measures are likely needed.

3. Empirical and objective measures of quality and prestige

In the previous section, we described several studies that examined researchers' perceptions of different journals. A number of empirical and/or objective metrics have also been employed to assess the quality and prestige of academic journals. These include automatically generated indices such as citation counts, the Thompson Reuters impact factor (IF), Google Scholar metrics, the h-index, and others that are collected manually such as acceptance rates (Haensly et al., 2008).

Scholars have warned against the dangers of evaluating research led by metrics rather than by judgment (Hicks et al., 2015) or clarified the intended function of these metrics (Varki, 2017). For some researchers, such indices might seem unimportant since each journal tends to have certain unique aims and audiences. Nevertheless, a small-but-growing body of work within applied linguistics has begun to explore citation-based indices and other similar metrics as indicators of quality, prestige, and influence. In some instances, data from automatically produced bibliometric indices such as cite score (i.e., total citations of a given journal) are considered in relation to data collected through methodological synthesis, a type of secondary research that systematically codes and analyzes research and reporting practices in a given domain (see Chong & Plonsky, 2023; Plonsky & Gönülal, 2015). Amini Farsani et al. (2021), for example, collected 3,992 articles from 18 journals, each of which was coded for a number of methodological features including study design and the type(s) of data that were analyzed (e.g., quantitative, qualitative, mixed). Among other results, citations were greater among quantitative than qualitative studies. Secondary research such as narrative reviews and meta-analyses were also among the most highly cited reports in their sample. The authors also examined authorship patterns and found a higher rate of co-authorship among quantitative research than in qualitative or mixed-methods studies.

Al-Hoorie and Vitta (2019) provide another example of a combination of traditional, bibliometric data such as impact factors and hand-coded, synthetic data. This study focused on the reporting practices related to statistical analyses in 150 articles found in 30 different journals that publish L2 research. One of the study's research questions specifically addressed the relationship between journal statistical quality (i.e., thorough and transparent reporting of statistical results, such as reporting reliability, non-significant results, effect sizes etc.) and Scopus citation analysis metrics. The study revealed a positive and moderate correlation between these variables (r ranged from .339 to .414), indicating that journals of higher statistical quality (or at least one facet thereof) are also cited more often. This finding falls in light with what we might hope or expect to find. Nevertheless, such indicators cannot full account for the notions of quality and prestige.

4. The current study

The overarching picture is that journal quality and prestige in applied linguistics might best be illustrated with a holistic approach that involves triangulation across measures. An approach which includes perceptual, objective, as well as synthetic data will likely, we believe, provide the most appropriate, comprehensive, and accurate view. Taken from a different angle, coverage of these different data sources, as intended by the present study, allows, we believe, for greater content validity for the constructs of interest. The current study investigates the following three research questions:

1. How do L2 scholars perceive different applied linguistics journals in terms of prestige and quality? (RQ1)
2. What is the relationship between objective (e.g., IF) and subjective (i.e., ratings) measures? Which objective measure is most closely related to what researchers in the field think? (RQ2)
3. To what extent do different journal features predict perceived journal prestige and quality? (RQ3)

5. Method

This section presents a detailed description of the different data sources that were employed to address our research questions. Because our data were drawn from four different types of sources (researcher survey; editor survey; synthetic coding of primary studies; automatically generated citation-related indices), we provide in Table 1 an overview of each source, the type of information it yielded, and the items collected in each.

Table 1 Overview of data types and sources

Types	Sources	Items
Participant background	Survey (2017)	Current position Continent of PhD Home department
Perceptions of journal prestige and quality*	Survey (2017)	Ratings of journal prestige Ratings of journal quality Criteria for ratings Open-ended comments
Submission information**	Journal editors (2016)	Number of submitted manuscripts Number of accepted manuscripts Acceptance rate
Bibliometric Information**	Web of Science; Google Scholar (2015)	Impact factor Eigenfactor H5 H5-median Total citations
Synthetic Information**	Journal websites (2017)	Journal starting year Journal guidelines Mission statement Editorial board affiliation Transparency (number of entries on IRIS)
	Manual coding of studies (2014-2016)	Article title Number of authors per article Author affiliation Type of paper Type of data Participant's first language Target language

Note. Measures considered to be subjective or objective are denoted with * and **, respectively.

All data were collected in 2017. Submission and bibliometric information were only available for 2016 and 2015, respectively, at the time of data collection. We chose to code studies published between 2014 and 2016, the most recent, intact three-year longitudinal data from the data collection point. We recognize that our data are somewhat obsolete and may not reflect the current state of the field. However, many of the patterns present in our results are, we believe, fairly stable in that we do not expect that they have changed much since the data were collected. Furthermore, rather than providing any definitive findings on the quality and prestige in applied linguistics and the predictors thereof, we envision this study as a kind of exploration into these relationships. Future research taking a more confirmatory approach can and should be concerned with more current data.

5.1. Subjective measures: Perceived journal quality and prestige

5.1.1. Survey questionnaire

In order to gauge how L2 scholars rate applied linguistics journals in terms of prestige and quality, we employed an online survey questionnaire with two main sections. The first part of the survey was concerned with participant background. Respondents were asked about their academic position, affiliated academic department, and place where they earned or were pursuing their PhD. The second part of the survey was developed to understand participants' perceptions of L2 journal prestige and quality. Respondents were asked to rate L2 journals in terms of prestige and quality on a 6-point Likert scale, with 1 indicating "very low prestige" and 6 indicating "very high prestige." The ratings for prestige and quality were organized separately. For each journal, respondents could also check the box "I am not familiar enough with this journal to evaluate it" instead of providing a rating. Besides, respondents were allowed to name two unlisted journals and provide corresponding ratings if they chose to do so. Next, respondents were asked to indicate the criteria used in their ratings of the L2 journals: They were given a list of criteria (e.g., acceptance rate, impact factor, intended audience, methodological rigor, relevance to context) and were guided to select all that apply. Again, the selection of the ranking criteria was done separately for prestige and quality. In the end, participants optionally provided comments about the survey to complement the quantitative findings. The survey, which was piloted, revised, and finalized before the formal data collection, will be made available on IRIS upon publication (Marsden et al., 2016).

For the purpose of this study, L2 journals were defined as journals that publish empirical inquiry into the learning, usage, knowledge, teaching, and assessment

of non-native languages, or in short, journals that primarily publish L2 research. Since including an exhaustive list of L2 journals in the survey was impractical, 27 journals were selected for the current investigation: *Annual Review of Applied Linguistics, Applied Psycholinguistics, Applied Linguistics, Bilingualism: Language and Cognition, CALICO Journal, Canadian Modern Language Review, Computer Assisted Language Learning, English Language Teaching, English Language Teaching Journal, Foreign Language Annals, International Journal of Bilingualism, Journal of English for Academic Purposes, Journal of Multilingual and Multicultural Development, Journal of Second Language Writing, Language Assessment Quarterly, Language Awareness, Language Learning, Language Learning & Technology, Language Teaching, Language Teaching Research, Language Testing, Modern Language Journal, RELC Journal, Second Language Research, Studies in Second Language Acquisition, System, TESOL Quarterly*. These journals were selected based on a combination of sources and considerations, comprising journals included in closely related studies (e.g., VanPatten & Williams, 2002), the top 150 journals in the linguistics category by Thomson Reuters, the top 20 journals by h-5 index in several relevant categories ("language and linguistics," "foreign language learning," and their general category "humanities, literature & arts") within Google Scholar metrics, and potential familiarity of the journal among participants. Some journals publish research in domains that are somewhat broader than L2 research (e.g., *Applied Linguistics*), and others publish in subdomains within L2 research (e.g., *Journal of Second Language Writing, Language Testing, Computer Assisted Language Learning*). The focus for all included journals, however, remains L2 research.

5.1.2. Procedure

Data collection took place between April 2017 and August 2017. First, we sent the survey through Qualtrics to approximately 2,000 potential L2 scholars worldwide. These scholars were identified because they are affiliated with an applied linguistics or SLA department or they had one or more articles published in a major L2 journal from which their email addresses were extracted. To ensure an adequate level of familiarity and knowledge about the L2 journals mentioned in the survey, only scholars who have completed or are currently completing doctoral degree were asked to fill out the survey. Scholars were asked to forward the survey to qualified L2 researchers they felt would potentially be interested in taking the survey. Participation was voluntary, and respondents gave their consent before taking the survey. In the end, we collected a total of 543 responses. However, some had no responses or stopped prior to completing both sections of the survey. Additionally, three respondents chose not to give their consent to participate in the initial consent

screen. This resulted in removing 171 incomplete cases. Thus, the final sample consisted of responses from 372 respondents.

Table 2 shows the background information of the L2 scholars who participated in the survey. They represented a range of academic positions (e.g., professor, associate professor, assistant professor, lecturer) as well as PhD/EdD students. Note that some endorsed the "Other" category, comprising 11.29% of the sample (N = 42), which included postdoctoral researchers, professors emeriti, administrators, independent researchers, and non-tenure track faculty (e.g., adjunct professor, research professor). The academic units that the survey respondents were affiliated with at the time of data collection were diverse: English (N= 67, 18.01%), education (N= 55, 14.78%), and applied linguistics (N= 50, 13.44%), followed by other departments shown in the table. Approximately 16.40% (N = 61) of the respondents endorsed the 'Other' category to indicate their affiliated academic departments, including neuroscience, cognitive science, international communication, TESOL, rhetoric, and also non-profit organizations/research centers. The survey also gauged where respondents earned or were pursing their PhD. The percentage of respondents who earned their degree in North America comprised the majority of the sample (N = 209, 56.18%) compared to other areas, such as Europe (25.54%), Asia (6.99%), and Oceania (6.72%). The results of the study are, therefore, limited in their generalizability outside of North America.

Table 2 Participant background information

		Number of respondents (N)	%
Academic position (N = 370)	Professor	97	26.08
	Associate professor	78	20.97
	Assistant professor	53	14.25
	PhD/EdD student	52	13.98
	Lecturer	29	7.80
	Senior lecturer	18	4.84
	Reader	1	0.27
	Other (e.g., postdoctoral fellow)	42	11.29
Departments (N = 371)	English	67	18.01
	Education	55	14.78
	Applied linguistics	50	13.44
	Linguistics	45	12.10
	Foreign languages (e.g., Arabic)	42	11.29
	SLA/SLS	38	10.22
	Psychology	13	3.49
	Other (e.g., neuroscience)	61	16.40
Place where PhD was earned or being pursued (N = 370)	North America	209	56.18
	Europe	95	25.54
	Asia	26	6.99
	Oceania	25	6.72
	Africa	1	0.27
	South America	0	0.00
	Other (e.g., Turkey, Middle East)	14	3.76

5.2. Objective measures: Submission, bibliometric and study information

A synthetic coding scheme was developed to extract information from the 27 included journals and the studies published in these journals from 2014 to 2016. As recommended in the synthetic literature (e.g., Plonsky & Oswald, 2015), this instrument was developed through an iterative process that incorporated information from bibliometric studies such as those reviewed above and the results of pilot coding. The synthetic scheme was piloted with four randomly selected journals by four co-authors and was then revised and finalized.

The final synthetic scheme consisted of submission information as well as bibliometric and synthetic journal and article information (see Table 1). Submission information was obtained directly from journal editors through email. The data for this variable reflects the acceptance rate of 2016, calculated based on the total number of submitted manuscripts divided by the total number of accepted manuscripts. Bibliometric information includes impact iactor, eigenfactor, h5, h5-median, and total number of citations by 2015. Synthetic information includes two parts: (1) journal-related information was directly obtained from the journal websites; (2) article information, such as author affiliation, type of paper, and type of data (quantitative, qualitative, or mixed), was obtained through manual coding of 2,124 randomly selected articles published in the 27 journals by eight applied linguists involved in the project.

5.3. Analysis

Before analyzing the data to answer our research questions, several rounds of data aggregation or combination were carried out. Initially, the survey yielded a participant-level data file, while the synthetic scheme yielded a journal-level data file and a study-level data file. All data were processed and analyzed using a combination of Microsoft Excel and SPSS (version 27, IBM Corp.).

To process the survey data, we aggregated prestige-rating and quality-rating data in the participant-level data file and added them to the journal-level data file. Specifically, based on all participants' prestige ratings of a certain journal, descriptive statistics of the prestige rating (including mean, standard deviation, maximum value, minimum value) as well as the number of participants who indicated insufficient familiarity with this journal to give a rating were calculated. The same process was then repeated for all participants' quality ratings.

To process data coded in the synthetic scheme, we first aggregated numerical study-level data and then added them to the journal-level data file. Specifically, among all coded studies for a given journal, the percentages that each

category of a variable accounted for were calculated (e.g., percentage of articles with a single author; percentage of articles that are empirical, classroom-based studies). These percentages were then added to the data for each journal, along with existing items such as impact factor, in the journal-level data file.

The two data files, one participant-level data file and the journal-level data file, were then used for different analyses to answer our research questions. The first research question, regarding L2 scholars' perception of journal prestige and quality, was addressed through simple descriptive statistics. The second research question, regarding the relationship between objective and subjective measures, was answered using Pearson's correlation analyses. Finally, we employed multiple regression to address the last research question regarding the extent to which the journal features predict perceived journal prestige and quality.

6. Results

Before presenting the results for our three research questions, we provide the descriptive statistics of the objective and collected measures of the journals under investigation, including submission, bibliometric and synthetic information.

On average, the journals in our sample received and accepted close to 300 and 50 submissions per year, respectively. These values vary substantially across our sample, however, as indicated in the large SDs. The journals also varied substantially on a number of other measures such their presence on IRIS (used as an indicator of journal transparency, one facet of study quality; see Gass et al., 2021) and on the various citation-based indices.

While the journals included board members across different continents, over 75% had North American (46.4%) or European (31.6%) affiliations. A similar pattern was observed in author affiliations, with 42.1% of the authors whose work appeared as journal articles in our sample affiliated with institutions in North America, followed by Europe (27.3%), Asia (22.3%), Oceania (6.4%), South America (1%), and Africa (0.8%). These numbers make a strong case for the fact that scholars from institutions in geographical regions other than North America and Europe are underrepresented in prominent applied linguistics journals.

The data in Table 3 also show that almost half the articles published in the journals sampled here were non-classroom-based empirical studies (47.1%). A tendency toward publishing articles with both quantitative and qualitative data (54%) over articles with purely quantitative (19%) or qualitative data (26.3%) was observed, as was the tendency toward English as the target language of the study (62.2%). No marked preference for either sole- or co-authorship was found in this data set.

Table 3 Submission, bibliometric and synthetic information across journals

Journal Characteristics	M	SD	N (Journals*)
Submission data			
Number of submitted manuscripts (2016)	279.609	135.202	23
Number of accepted manuscripts (2016)	50.696	48.564	23
Acceptance rate (2016)	18.803	12.815	23
Citation data			
Impact factor (2015)	1.375	0.687	24
Total citations (2015)	1064.261	699.088	23
Eigenfactor (2015)	0.002	0.001	23
Google h5 (2015)	24.115	5.955	26
Google h5-median (2015)	37.577	11.748	26
Synthetic data			
Number of entries on IRIS [4]	52.667	69.380	27
Journal start year	1979	21.156	27
Board member affiliation – Africa [5]	0.009	0.026	27
Board member affiliation – Asia [5]	0.130	0.130	27
Board member affiliation – Europe [5]	0.316	0.215	27
Board member affiliation – North America [5]	0.464	0.279	27
Board member affiliation – Oceania [5]	0.071	0.079	27
Board member affiliation – South America [5]	0.008	0.019	27
Average number of authors per article	1.810	0.376	27
Sole-authored articles [1]	0.482	0.162	27
Author affiliation – Africa [3]	0.008	0.015	27
Author affiliation – Asia [3]	0.223	0.202	27
Author affiliation – Europe [3]	0.273	0.151	27
Author affiliation – North America [3]	0.421	0.213	27
Author affiliation – Oceania [3]	0.064	0.051	27
Author affiliation – South America [3]	0.010	0.026	27
Study type – empirical, non-classroom-based [1]	0.471	0.224	27
Study type – empirical, classroom-based [1]	0.174	0.186	27
Study type – literature review, meta-analysis [1]	0.012	0.020	27
Study type – literature review, non-meta-analysis [1]	0.084	0.178	27
Data type – mixed [2]	0.540	0.248	27
Data type – qualitative [2]	0.263	0.194	27
Data type – quantitative [2]	0.190	0.170	27
Target language – English [2]	0.622	0.282	27

Note. *Data were not available for all items for all journals in our sample. (1) For these characteristics, percentages (i.e., number of articles that have this feature divided by the total number of coded articles for this journal) are used for calculating the mean and SD. (2) For these characteristics, percentages (i.e., number of articles that have this feature divided by the total number of coded articles that are applicable for coding this feature) are used. For instance, literature review articles are not applicable for coding data type. (3) For these characteristics, percentages (i.e., number of authors who are affiliated with an educational institution in this region divided by the total number of authors from all coded articles for this journal) are used. (4) The number of entries each journal has on IRIS database (Marsden et al., 2016) as of August 8, 2017, which is adopted as a measure of transparency. (5) For these characteristics, percentages (i.e., number of editorial board members who are affiliated with an educational institution in this region divided by the total number of board members for this journal) are used.

6.1. RQ1: How do L2 scholars perceive different applied journals in terms of prestige and quality?

Table 4 presents the prestige and quality ratings and ranking of the 27 applied linguistics journals in our sample, based on the 317 participant responses received. One of the salient patterns that emerged from the result is the significant overlap between the prestige and quality ratings. Despite some minor variations in the order of the ranking, the top journals in both rankings were *Language Learning, Studies in Second Language Acquisition, Applied Linguistics, Modern Language Journal, Bilingualism: Language and Cognition*, and *Language Testing*. It is also worth noting that one of the recurring themes in the comment section had to do with the similarities in participants' perception of prestige and quality. For example, many comments included statements like, "Quality and prestige go hand in hand," or "I guess responses to these two will be highly correlated," indicating the perceived overlap between the two constructs. Another participant wrote, "I'm not confident that my responses to 'prestige' are substantively distinguishable for my answers to 'quality.' To me, if a journal publishes evidently high-quality work, my estimation of its 'prestige' rises." By and large, this trend carried over to the remainder of the ranking. It should be noted that many of these higher-ranked journals were fairly broad in scope (e.g., *Language Learning, Applied Linguistics*). Domain-specific journals (e.g., *Computer Assisted Language Learning)*, on the other hand, tended to fall a bit lower in the rankings and to vary a bit more in terms of prestige and quality.

Additionally, higher-ranked journals tended to have a greater number of responses, indicating that these journals were more familiar to participants. In particular, many of the higher-ranked journals, such as *Language Learning, Modern Language Journal, TESOL Quarterly*, and *Studies in Second Language Acquisition*, received responses from a majority of the survey respondents. Some of the journals that received a lower number of responses across prestige and quality include *Journal of Multilingual and Multicultural Development, Language Awareness*, and *English Language Teaching*. The number of responses these journals received was generally less than half of the sample. To further explore the relationship between familiarity on the one hand and prestige and quality on the other, we calculated the correlation between each of these pairs of variables. Both were fairly strong. The number of respondents for each journal and their corresponding ratings of prestige and quality were found to correlate at $r = .64$ and $.61$, respectively.

Figure 1 and Table 5 summarize the frequency of the criteria used by respondents to evaluate the prestige and quality of the 27 L2 journals included in the survey. One of the notable findings is that in terms of both prestige and quality, methodological quality/rigor outranked all other criteria. This suggests that, overall,

methodological quality/rigor was considered to be one of the most important criteria in determining a journal's prestige and quality. Attesting to these quantitative data, the comment section included statements such as, "Quality and prestige should be related to the rigor of the peer-review process," or "Methodological quality/rigor equals quality in my opinion, and is the only meaningful evaluation criteria for me." A number of other criteria were also ranked similarly to judge journal prestige and quality, such as time to publication, accessibility, quality and timeliness of the review process, and topic of articles.

Table 4 Ratings of applied linguistics journals

	Prestige				Quality			
	No of articles	M	SD	ranking	No of articles	M	SD	ranking
Language Learning	320	5.35	.94	1	280	5.30	.88	1
Studies in Second Language Acquisition	318	5.23	.95	2	278	5.12	.96	2
Modern Language Journal	327	5.18	.98	3	290	4.98	1.02	6
Applied Linguistics	332	5.16	1.06	4	302	5.06	1.04	3
Bilingualism: Language and Cognition	222	5.03	.99	5	190	5.04	.93	4
Language Testing	191	5.02	.93	6	170	5.02	.99	5
TESOL Quarterly	329	4.92	1.08	7	288	4.70	1.15	8
Journal of Second Language Writing	234	4.71	.99	8	206	4.68	1.07	10
Applied Psycholinguistics	215	4.70	1.06	9	191	4.88	1.00	7
Second Language Research	250	4.63	1.03	10	223	4.55	1.09	12
Language Assessment Quarterly	158	4.57	1.15	11	134	4.69	1.18	9
Language Teaching Research	214	4.54	.94	12	187	4.55	1.09	12
Annual Review of Applied Linguistics	289	4.46	1.23	13	266	4.58	1.24	11
Language Teaching	199	4.36	1.17	14	175	4.46	1.18	14
Language Learning & Technology	181	4.32	1.13	15	160	4.36	1.15	15
System	306	4.25	1.17	16	269	4.23	1.26	17
Journal of English for Academic Purposes	185	4.12	1.10	17	156	4.34	1.08	16
International Journal of Bilingualism	167	4.06	1.04	18	150	4.15	1.11	18
Canadian Modern Language Review	272	3.87	1.07	19	244	4.12	1.09	19
Journal of Multilingual and Multicultural Development	116	3.86	1.09	20	98	3.97	1.17	21
Foreign Language Annals	228	3.82	1.21	21	200	3.90	1.23	22
Computer Assisted Language Learning	205	3.82	1.19	21	171	3.99	1.16	20
English Language Teaching Journal	180	3.79	1.28	23	160	3.87	1.29	24
Language Awareness	148	3.72	1.14	24	129	3.88	1.21	23
RELC Journal	184	3.62	1.19	25	152	3.61	1.29	27
English Language Teaching	165	3.60	1.29	26	123	3.78	1.26	25
CALICO Journal	210	3.51	1.23	27	172	3.78	1.30	25

Despite the overlap, other evaluation criteria varied considerably across prestige and quality. For example, scholars mentioned (in descending order) impact factor, frequency of citation, acceptance rate, and editorial board members as critical factors that influence their evaluation of journal prestige. On the other hand, they frequently chose (in descending order) article topic, impact factor, relevance to context, and frequency of citation as evaluation criteria for journal quality. While impact factor and frequency of citation were both important in their perceived journal ratings, the criteria were also different. For example, prestige ratings tended to be

influenced by more objective and/or holistic measures, such as acceptance rate and editorial board (e.g., authority of the editor, board members' records of publication), while quality tended to be associated with context-specific factors at a narrower level, such as topic of articles or relevance to context.
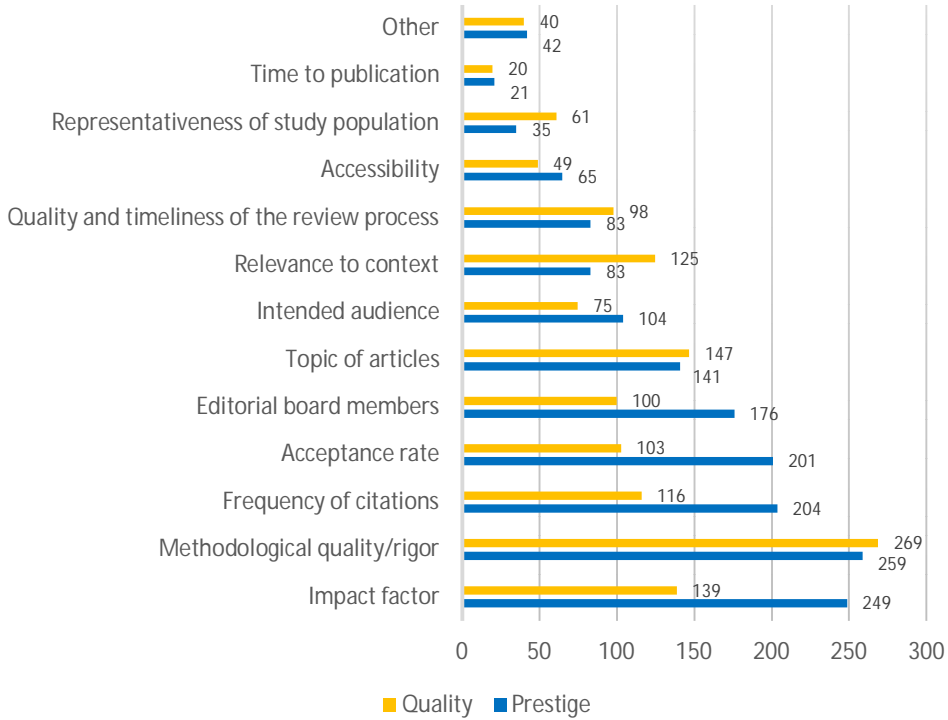

Figure 1 Criteria used by L2 scholars in ranking journal prestige and quality

Table 5 Frequency of criteria used by L2 scholars in ranking journal prestige and quality

	Prestige		Quality	
	Frequency	Ranking	Frequency	Ranking
Methodological quality/rigor	259	1	269	1
Impact factor	249	2	139	3
Frequency of citations	204	3	116	5
Acceptance rate	201	4	103	6
Editorial board members	176	5	100	7
Topic of articles	141	6	147	2
Intended audiences	104	7	75	9
Relevance to context	83	8	125	4
Quality and timeliness of the review process	83	8	98	8
Accessibility	65	10	49	11
Representativeness of study population	35	11	61	10
Time to publication	21	12	20	13
Other	42	10	40	12

Some participants provided specific comments about the discrepancy: "Some journals have acquired high levels of academic prestige, but have rather lost touch with reality," and "I find that some of the journals that are always discussed as being high quality tend to publish research on more restricted topics." Additional and related comments explaining participants' ratings are as follows:

> *This is very difficult to judge because some journals have good reputations, and deservedly so based on study quality, but they publish within a very narrow band of research, so they may end up publishing research which is less relevant generally (though still of a very high quality).*
> *Some journals have high status but the quality can be very uneven . . .*
> *Impact factors are difficult to judge . . . based on how easily accessible an article is and the topic, not always the quality of the research.*
> *. . . sometimes journals link prestige to the author(s) names and not so much to the quality of the study.*

Of course, some respondents mentioned in the survey that quality and prestige measure the same construct and thus cannot be differentiated: "I find the difference, at an operational level, not significant since prestige is based on quality." Some commented that the peer review process has substantially changed their perception of the journal quality, despite being a prestigious one: "I've had bad experiences with [X] – I don't think their review process is fair and unbiased. It's all a bit of a lottery to some extent – acceptance/rejection is too dependent on reviewer personality/attitude. Reviewer responses to the same paper can vary widely – some people are just more generous & encouraging while others see the reviewing process as a chance to stifle potential competition, I feel."

In addition to the listed criteria, some participants selected the "Other" category to explain their responses. Other factors that reportedly influenced scholars' ranking of journal prestige included the publisher, theoretical rigor, prestige of author(s), practical relevance, professionalism of editors and reviewers, and reputation among peers. In terms of journal quality, they mentioned reporting of statistics (e.g., inclusion of effect sizes), theoretical relevance, quality/style of writing, novelty, and reputation among peers.

6.2. RQ2: What is the relationship between objective (e.g., IF) and subjective (i.e., ratings) measures? Which objective measure is most closely related to what researchers in the field think?

To examine the relationship between objective (i.e., submission information, bibliometric data, synthetic data) and subjective (journal quality and prestige ratings)

measures in RQ2, we first checked the four assumptions to run Pearson's correlation analyses following the procedures in Larson-Hall (2015): (1) independence of observations; (2) normally distributed data; (3) a linear relationship between each pair of variables; and (4) homoscedasticity (constant variance). First, data collected for each journal were independent and did not influence each other. Then, to check the normality of data, we examined data skewness and kurtosis and subjected all data to the Shapiro-Wilk test. We kept data with a skewness value between ±1 and a kurtosis value between ±2 (George & Mallery, 2020, pp. 114-115) as well as data with a non-significance ($p > .05$) result in the Shapiro-Wilk test. As a result, we excluded 18 variables[1] that did not meet the normality benchmarks. Next, we used scatterplots to rule out non-linear relationships between objective and subjective measures and further excluded one variable (i.e., board member affiliation – Europe). Finally, we checked data homoscedasticity and confirmed that all remaining variables' residuals were equally distributed. In the end, two subjective measures and 11 objective measures were retained for this phase of the analysis: journal quality rating (QR), journal prestige rating (PR), impact factor (IF), Google h5 index (GH5), Google h5 median (GH5M), board member affiliation – North America (BNA), sole author percentage (SA), the average number of authors per article (A/A), author affiliation percentage – North America (ANA), author affiliation percentage – Europe (AE), data type percentage – quantitative (DQ), data type percentage – qualitative (DL), target language percentage – English (TE).

Table 6 shows the results of bivariate Pearson's correlation analysis. Three bibliometric variables (impact factor, Google h5 index, and Google h5 median) showed significant positive correlations with the journals' mean prestige and quality ratings ($.536 \leq r \leq .579$, $p < .01$). According to Plonsky and Oswald's (2014) empirically based and domain-specific benchmarks for r (.25 small, .4 medium, .6 large), these correlations are considered moderate. Not surprisingly, journals with higher impact factor values and higher h5-index or median tend to be considered of a higher quality and prestige by applied linguistics scholars. In addition, the results indicate that journals of higher prestige and quality tended to publish more articles that were collaborated. This is evident by the fact that the percentage of sole-authored articles negatively correlated with prestige ($r = -.439$, $p < .05$) and quality ($r = -.496$, $p < .01$) ratings and that the average number of authors per article positively correlated with prestige ($r = .421$, $p < .05$) and quality

---

[1] The 18 excluded variables are: acceptance rate; total citations; Eigen factor score; the number of entries on IRIS; journal starting year; board member affiliation percentage – Africa, Asia, Oceania, South America; author affiliation percentage – Africa, Asia, Oceania, South America; study type percentage – empirical non-classroom-based, empirical classroom-based, literature review meta-analysis, literature review non-meta-analysis; data type percentage – mixed.

(r = .473, p < .05) ratings. We also find that journals showing strong preference for quantitative articles generally received higher ratings in journal prestige (r = .500, p < .01) and quality (r = -.528, p < .01).

Table 6 Bivariate Pearson's correlation coefficient between objective and subjective measures

	Average journal prestige rating	Average journal quality rating
Impact factor	.54**	.54**
Google h5 index	.56**	.54**
Google h5 median	.58**	.58**
Board member affiliation – North America	.20	.25
Sole author percentage	-.44*	-.50**
Average number of authors per article	.42*	.47*
Author affiliation percentage – North America	.28	.32
Author affiliation percentage – Europe	.31	.34
Data type percentage – quantitative	.50**	.53**
Data type percentage – qualitative	-.34	.39*
Target language percentage – English	-.80	-.11

Note. ** p < .01; * p <.05 (2-tailed)

Not surprisingly, objective and subjective measures that gauge similar constructs show moderate to strong positive correlations. For example, journal prestige and quality were strongly correlated (r = .983, p < .01), the three bibliometric variables were significantly moderately correlated (.528 ≤ r ≤ .832, p < .01), and the two variables related to author collaboration were strongly correlated (r = .938, p < .01). What appears surprising is that some objective measures exhibit strong correlations. For example, journals with more board members from North America published more articles written by North American authors (r = .832, p < .01) as well as studies with English as a target language (r = .54, p < .01).

As mentioned in the assumption checks, some variables (e.g., acceptance rate, total citations, number of entries on IRIS, study type) were excluded from the correlation analyses, which could have offered a more comprehensive picture into this new field. Given the exploratory nature of our study, we hope to offer an introduction to these variables that applied linguistics researchers can fruitfully investigate in future studies.

6.3. RQ3: To what extent do different journal features predict perceived journal prestige and quality?

Our last research question sought to understand perceived journal prestige and quality as a function of various journal features simultaneously. Toward this end,

we applied multiple regression analysis. Specifically, we chose to employ stepwise regression, given the exploratory nature of our analysis and the lack of theoretical guidance with respect to the relative importance of different predictors (Jeon, 2015). As the ratings of prestige and quality were highly correlated (r = .983, p < .01), we averaged the two and obtained a single composite rating, which was then used as the dependent variable in the regression. After checking the assumptions of normality, multicollinearity, linearity, and homoscedasticity, four of the 11 journal features retained in RQ2 were excluded from the remaining analyses. As a result, seven journal features were entered into the regression as independent variables, namely Google h5 index, sole author percentage, impact factor, author affiliation percentage – North America, author affiliation percentage – Europe, data type percentage – qualitative, and target language percentage – English.

The results of the regression analysis showed that four of the seven journal features that were entered into the model (Google h5 index, sole author percentage, impact factor, author affiliation percentage – North America) collectively explained 58 percent of the variance in journal prestige and quality ratings (R^2 = 0.58, F (4, 22) = 7.51, p = .001). The full results of this regression model are reported in Table 7. Among the four predictor variables in this regression model, one (Google h5 index) emerged as a statistically significant predictor of journal prestige and quality composite ratings.

Together, these results indicated that a journal was likely to be rated with higher prestige and quality when it had a higher Google h5 index (β = 0.38), a smaller percentage of single-authored articles (β = -0.30), a higher impact factor (β = 0.29), as well as a larger percentage of authors who were affiliated with North American institutions (β = 0.24). It should be noted that given the small sample size (i.e., the number of journals that we examined; N = 27), the current findings are suggestive rather than conclusive.

Table 7 Regression results for perceived journal prestige and quality

Variable	B	SE	95% CIs	β	t	p	Cumulative R^2
Google h5 index	0.03	0.02	[0.00, 0.06]	0.38	2.29	.03	0.30
Sole author	-0.99	0.48	[-1.99, 0.01]	-0.30	-2.05	.05	0.47
Impact factor	0.24	0.13	[-0.04, 0.52]	0.29	1.78	.09	0.53
Author affiliation – North America	0.59	0.36	[-0.16, 1.34]	0.24	1.63	.12	0.58

Note. CI = confidence interval

7. Discussion

The overall results of the study can be summarized by four main points. First, we set out to describe the landscape of L2 journals and journal articles published

from 2014 and 2017. One of the notable findings is that L2 scholars affiliated with North American institutions were well represented both in terms of authorship and editorial boards. Second, we found that there is a considerable consensus among L2 scholars about the perceived prestige and quality of journal in the field. The ratings of journal prestige and quality were similar, especially when it comes to domain-general journals. The overlap between the constructs of prestige was evident in the ratings and rankings as well as in participants' open-ended comments. Despite the similarities in ratings, prestige ratings tended to be influenced by more objective measures (e.g., acceptance rate, editorial board), while quality tended to be associated with context-specific measures (e.g., topic of articles, relevance to context). Third, we also found that L2 scholars' perceived quality/prestige ratings were strongly correlated with objective measures, such as impact factor, Google h5 index, and percentage of authors affiliated with North American institutions. Lastly, we found that high Google h5 index (β = 0.38), a percentage of single-authored articles (β = -0.30), impact factor (β = 0.29), and percentage of authors who were affiliated with North American institutions (β = 0.24) are all predictive of higher perceived ratings of journal prestige/rating.

Expanding on this final point, the four features (high Google h5 index, percentage of single-authored articles, impact factor, and percentage of authors affiliated North American institutions), together, explained 58% of the variance in journal prestige and quality ratings. This finding is, in part, consistent with the results of correlation analyses which indicated that journals with a high impact factor or high Google h5 index were considered of higher prestige/quality by L2 scholars surveyed.

Beyond simply describing researchers' perceptions of different journals and the variables that predict those perceptions, our findings may have implications for those handling tenure/promotion cases, for whom the perceived prestige/quality of publication venues may be an important factor in evaluating their colleagues. Here, though, we would urge caution. Similar to AAAL's guidelines (American Association for Applied Linguistics, 2019) on promotion and tenure, however, we would not recommend relying too heavily on journal rankings, whether provided by automatically generated tools such as Google Scholar or a survey such as ours. To be sure, both perceived/subjective and objective journal prestige/quality indicators are influenced by other context-specific factors, depending on the subdomains, research orientation, and research context. Furthermore, a scholar's contribution to the field cannot and should not be viewed in solely quantitative terms. All that said, our findings could also serve as a resource for those less familiar with L2 research, those who might be new to the field, or those who may need an entry point into where different journals stand in terms of prestige and quality.

The study also provides evidence for a relationship between authorial/editorial affiliation and journal prestige/quality. To begin with, the descriptive statistics showed a dominant presence of scholars affiliated with North American institutions in terms of authorship and editorial boards. Authors and members of editorial boards from other areas of the world, especially South America, Asia, and Africa, were underrepresented. Furthermore, the correlational analysis indicated that the journals with more board members from North American institutions tended to publish more articles written by authors from North American institutions (as well as studies with English as a target language). It is not entirely clear whether there is a bias toward North American authors/editorial board members, or whether authors from North America tend to submit more to journals with North American editorial boards. While it is premature to make any claims about the North American bias, the findings show limited diversity in publications and editorial boards, that is, overrepresentation of North America-based researchers in the landscape of L2 journals. Building on the emerging body of research that focuses on the issues of accommodating inclusion/diversity in the field of applied linguistics (e.g., Bhattacharya et al., 2020), the current study will, we hope, prompt action on the part of these journals to diversify their boards and the voices of those scholars whose work they publish.

Another noteworthy finding is that journals that publish more collaborative work and less single-authored work were perceived as more prestigious and of higher quality. As collaboration and cross-disciplinary work has been encouraged in the field (Amini Farsani et al., 2021), we view this finding optimistically. As explained in VanPatten and Williams (2002), joint projects can be advantageous in that they can result in greater productivity for all those involved. One way to understand this finding is that collaboration may increase the quality and productivity of work of scholars with different types of expertise, which may also increase the likelihood of getting the article accepted. Further studies into the nature and process of collaboration in applied linguistics are also needed.

The results of our study also showed that journals with a high Google h5 index, as well as those that are geared toward publishing research associated with quantitative paradigms, tend to be perceived as more prestigious and of higher quality. It should be noted that Google h5 index numbers may to some extent indicate influence over quality and do not include a system for discounting self-citations. Additionally, the fact that, overall, the correlations between journal features and prestige and quality ratings were not particularly strong suggests that the participants in this study may have had their own criteria and judgments in mind when determining the prestige or quality of the different journals sampled here.

The results here may also indicate the existence of a sort of echo chamber when it comes to the make-up publication boards and the institutional affiliation

of the authors they tend to publish. The majority of the journals sampled here had boards made up of members from North America and Europe, and the majority of the authors whose work appeared in the sample were from these same regions. Additionally, 62% of the articles sampled here focused on the study of English. This begs the question as to whether quality and prestige, which this study has shown can be predicted by h5-index ratings, are unduly connected with North American and European research paradigms and culture in the minds of many academics. Additionally, the fact that English was the target language of focus for the majority of the articles sampled here is an important reminder for North American boards and scholars, as well as those from other regions, of the necessity to be allies in the effort to "maximize the knowledge and resources available to maintain and strengthen revitalization efforts of the Indigenous languages around the world" (McIvor, 2020, p. 93). If journal boards continue to overwhelmingly publish North American and European authors working on English, the field must ask itself what space is left for scholars from other regions and/or those working on languages other than English. This same issue of circularity and lack of representation extends to the tenure process and may influence who ultimately ends up getting the type of job that leads to inclusion in a prestigious journal board.

8. Limitations

The study was exploratory in nature, and while offering a novel glimpse into how our perceptions of journal prestige/quality relate to journal characteristics, it is not without limitations. First, since the bibliometric measures (e.g., impact factor) are collected at the journal level (i.e., in aggregated form), they do not necessarily represent the quality of individual papers. For example, the constituents of "quality" or "importance" most of these measures refer to do not consider the methodological rigor or novelty of each paper, which some users of the information may value more. Rather, they are concerned with the attention the articles in the journal have received or, more specifically, the formal recognition or contribution (e.g., citation) the articles have received. Therefore, caution must be taken when interpreting the bibliometric measures and using them with a belief that the numbers represent the true "prestige" or "quality" of journals in all aspects. Second, given the small sample size, the current findings are suggestive rather than conclusive. The number of journals examined in the current study is 27. These 27 journals were selected based on previous research (e.g., VanPatten & Williams, 2002), yet they cannot be seen as representative of all L2 journals. Future studies should compile and re-run some of these analyses

based on a more exhaustive list of L2 journals. It should be also noted that not every study from each journal published between 2014 and 2016 was coded. Studies coded for each journal were randomly selected for the analysis. For example, as regards the *International Journal of Bilingualism*, we selected 25 articles from 2014, and 25 articles from 2015-2016 for coding. This was not consistent across the journal. It is hoped that the findings reported in the study offer some preliminary insights into L2 studies.

At least two additional limitations remain. In parallel to our limited sampling of journals and of the articles in those journals, our sampling of scholars in L2 research may also be limited. That is, although the sample size was fairly large for survey-based L2 research ($N = 371$), our sampling techniques almost certainly led to a bias toward North American and European participants. Future studies should seek to address this weakness by targeting participants outside of these regions. Finally, due to logistical constraints, we were not able to estimate inter-rater reliability for our synthetic coding. We believe, however, that our piloting and the low inference nature of most items in our coding scheme will protect against error that may have entered our dataset at this stage of the study.

9. Conclusion

Quality and prestige are often discussed in applied linguistics but are very rarely examined empirically. The study sought to explore these two constructs using a novel approach that involves several different types of data and data sources including subjective ratings, editor-provided and journal website data, synthetically coded data, and automatically generated citation data. Our results found a number of considerable relationships among these variables, and they provide, we hope, a foundation for others interested in these topics to better understand quality and prestige in the field. The perspective we offer here also serves to reveal certain biases which we hope will also encourage scholars, journal board members, and students to consider the issues of circularity and overrepresentation of certain cultures and research paradigms in applied linguistics.

References

Al-Hoorie, A. H., & Vitta, J. P. (2019). The seven sins of L2 research: A review of 30 journals' statistical quality and their CiteScore, SJR, SNIP, JCR Impact Factors. *Language Teaching Research*, *23*, 727-744.

American Association for Applied Linguistics. (2019). *Promotion & tenure guidelines*. https://www.aaal.org/promotion-and-tenure-guidelines

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

Bhattacharya, U., Jiang, L., & Canagarajah, S. (2020). Race, representation, and diversity in the American Association for Applied Linguistics. *Applied Linguistics*, *41*(6), 999-1004.

Bhuyan, R., Park, Y., & Lee, E. (2020). Should journal rankings matter? Assigning "prestige and quality" in the neoliberal academy. *Journal of Women and Social Work*, *35*(2), 157-159.

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

Egbert, J. (2007). Quality analysis of journals in TESOL and applied linguistics. *TESOL Quarterly*, *41*(1), 157-171.

Gass, S., Loewen, S., & Plonsky, L. (2021). Coming of age: The past, present, and future of quantitative SLA research. *Language Teaching*, *54*, 245-258.

George, D., & Mallery, P. (2020). *IBM SPSS statistics 26 step by step: A simple guide and reference* (16th ed.). Routledge.

Haensly, P. J., Hodges, P. E., & Davenport, S. A. (2008). Acceptance rates and journal quality: An analysis of journals in economics and finance. *Journal of Business & Finance Librarianship*, *14*(1), 2-31.

Hicks, D., Wouters, P., Waltman, L., De Rijcke, S., & Rafols, I. (2015). The Leiden Manifesto for research metrics. *Nature, 520*, 429-431.

Jeon, E. H. (2015). Multiple regression. In L. Plonsky (Ed.), *Advancing quantitative methods in second language research* (pp. 131-159). Routledge.

Johnstone, B. (2009). In the profession: Choosing the right journal for your manuscript. *Journal of English Linguistics, 37*, 375-379.

Larson-Hall, J. (2015). *A guide to doing statistics in second language research using SPSS and R*. Routledge.

Larsson, T., Plonsky, L., & Hancock, G. R. (2022). Latent variables in learner corpus research: Conceptual and methodological considerations. *International Journal of Learner Corpus Research*, *8*, 237-260. https://doi.org/10.1075/ijlcr.21007.lar

Lee, K. P., Schotland, M., Bacchetti, P., & Bero, L. A. (2002). Association of journal quality indicators with methodological quality of clinical research articles. *JAMA, 287*(21), 2805-2808.

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.

McIvor, O. (2020). Indigenous language revitalization and applied linguistics: Parallel histories, shared futures? *Annual Review of Applied Linguistics, 40*, 78-96.

Plonsky, L., & Gönülal, T. (2015). Methodological synthesis in quantitative L2 research: A review of reviews and a case study of exploratory factor analysis. *Language Learning, 65*, Supp. *1*, 9-36.

Plonsky, L., & Oswald, F. L. (2014). How big is "big"? Interpreting effect sizes in L2 research. *Language Learning, 64*, 878-912.

Plonsky, L., & Oswald, F. L. (2015). Meta-analyzing second language research. In L. Plonsky (Ed.), *Advancing quantitative methods in second language research* (pp. 106-128). Routledge.

Smith, B., & Lafford, B. A. (2009). The evaluation of scholarly activity in computer-assisted language learning. *Modern Language Journal*, *93*, 868-883.

Suber, P. (2010). Thoughts on prestige, quality, and open access. *Logos*, *21*(1-2), 115-128.

VanPatten, B., & Williams, J. (2002). *Research criteria for tenure in second language acquisition: Results from a survey of the field*. (Unpublished manuscript). The University of Illinois of Chicago, IL.

Varki, A. (2017). Rename the impact factor. *Nature*, *548*(7668), 393-393.

Weiner, G. (2001). The academic journal: Has it a future? *Education Policy Analysis Archives*, *9*(9), 1-19.

West, R. E., & Rich, P. J. (2012). Rigor, impact and prestige: A proposed framework for evaluating scholarly publications. *Innovative Higher Education*, *37*(5), 359-371.

Yu, D., Wang, W., Zhang, S., Zhang, W., & Liu, R. (2017). A multiple-link, mutually reinforced journal-ranking model to measure the prestige of journals. *Scientometrics, 111*, 521-542.

Zhao, J., Beckett, G. H., & Wang, L. L. (2017). Evaluating the research quality in education journals in China: Implications for increasing global impact in peripheral countries. *Review of Educational Research*, *87*(3), 583-618.