

## Exploring Genre, Research Method, and Reliability Coefficients of the ESP Journal Articles Published Between 2010 and 2020

<sup>1</sup>Behzad Ghonsooly\*

<sup>2</sup>Tania Hossain

<sup>3</sup>Safoura Jahedizadeh

<sup>4</sup>Fadhil Shihan

Research Paper

IJEAP- 2203-1848

Received: 2022-03-02

Accepted: 2022-05-02

Published: 2022-05-12

### Abstract

The current study purports to carry out a systematic review of 279 articles published during the last ten-year period in the *ESP Journal*. It intends to (1) investigate how genre theory is viewed and used, (2) examine the most frequently used research methods, and (3) explore the general trends in the findings based on a bibliometric analysis of reliability coefficients. To this end, all the articles were initially codified based on the genres they established, type of reliability they reported, namely internal consistency (RA), interrater reliability (RB), and intra-rater reliability (RC), and the methods implemented, including Qualitative (MA), Quantitative (MB), and Mixed-Methods (MC). The findings obtained using SPSS software indicated that among the skills and subskills, writing received the highest attention (60.2%), and among various disciplines, business had the highest frequency (13.6%). Besides, Corpus Analysis was used by most of the researchers (43.4%). Furthermore, only 86 articles (30.8%) reported the reliability coefficients, and interrater reliability had the highest, and intra-rater reliability had the lowest frequencies. Among the three methods of research, Mixed Methods had the maximum frequency (52%) and Quantitative Methods had the lowest (11.5%). Besides, 36.6% of the articles used the Qualitative Method for their studies.

**Keywords:** Corpus Analysis, Genre, Reliability, Research Method, Systematic Review

### 1. Introduction

Studying the textual production of research papers has a long history in English for Specific Purposes (ESP) inquiry (Swales & Leeder, 2012). Besides, the number of articles in English has increased significantly in the last decades (Curry & Lillis, 2017; Hanauer & Englander, 2011; Solovova, Santos, & Veríssimo, 2018). This is because English language has become the lingua franca in the research world (Kuteeva & Mauranen, 2014). The widespread use of English is not restricted to writing mode; ESP lectures are also becoming more and more common in higher education (Fenton-Smith, Humphreys, & Walkinshaw, 2017; Hwang & Lin, 2010; Wächter & Maiworm, 2014). In other words, ESP has attracted researchers' attention in recent years (e.g., Javid & Mohseni, 2020).

As stated by Manzoor et al. (2020), scrutinizing academic writing is of great importance as researchers all around the world are exploring new approaches and innovations in their fields. Hence, genre analysis delves into the organization of texts to explore various methods and subject

---

<sup>1</sup>Full Professor of Applied Linguistics (Corresponding Author), ghonsooly@um.ac.ir, English Department, Ferdowsi University of Mashhad, Mashhad, Iran.

<sup>2</sup>Full Professor of Sociolinguistics, kstania2@waseda.jp, Faculty of letters, Arts and Sciences, Waseda University, Tokyo, Japan.

<sup>3</sup>Lecturer in Applied Linguistics, jahedi.s1310@gmail.com, English Department, Imam Reza International University, Mashhad, Iran.

<sup>4</sup>English Instructor, fadhilshihan2@gmail.com, Basra Oil Company, Basra, Iraq.

matters investigated in the academic world (Dudley-Evans, 1994; Swales, 1990). In other words, it is important for students and researchers to have this genre knowledge as it helps in generating discourse of the specific field as well as identifying the modes and purposes of texts (Manzoor et al., 2020). Consequently, investigating the genre of published articles seems crucial, since producing research texts has always been a challenging and demanding job, especially for non-native speakers of English (Saidi, & Cheraghi, 2020).

Genre analysis, since its introduction in the early 1980s, has attracted a lot of attention. Although much of the literature has focused on research articles (e.g., Marefat, Farahanynia, Bolouri, Chamani, & Soleimani, 2021; Swales, 1990, 2004), theses (e.g., Starfield, 2019), and legal discourse (e.g., Bhatia, 1993, 2000; Connor & Gladkov, 2004), there seems to be a gap in research concerning the genre of published articles in the ESP Journal in the last decade.

Another problem which may arise is when we are not certain about the acceptable reliability coefficient. It is unclear if limited and non-statistical effects in investigational revisions should be qualified to a type of dependent dimension with little consistency or to some other review features. Real or experimental reliability, of course, offers possibly a greater danger to the rationality of quantifiable L2 research results (Larson-Hall & Plonsky, 2016; Plonsky, 2016). Nevertheless, despite the reports given of low reliability in the context of EFL and ESP Journal (Chaudron, 2001), the teams of researchers seem not to be almost cognizant of the real consistency of the information gathering devices they implement in carrying out their research.

Research context of the ESP Journal has encountered numerous methodological inappropriateness and irregularities. In other words, several new meta-analyses and operational assessments (e.g., Peterson et al., 1985; Plonsky, 2016) have discovered reasonably unusual and infrequent reportage of reliability evaluations. Accordingly, the researchers and authors of primary studies do not have any empirically-based source of judgment for interpreting reliability coefficients in comparison with others in the context of second language as well as ESP Journal which is published quarterly.

To this end, at the outset, it is required to provide a clear and succinct definition of the reliability coefficient and reliability systematic review. For instance, according to Haertel (2006), when someone goes under test or is monitored several times, such as a student assessed at school in his/her class achievement test or a machinist witnessed while fixing train area paraphernalia, grooves reporting his or her presentation might or might be interpreted the same. Therefore, demanding the constancy of one single groove for managerial resolves, but the grades of persons might not agree altogether. Consequently, the goal of consistency studies is to assess the reliability of the obtained marks in a repeated observation process. Therefore, in the process of reliability coefficient estimate, the reliability between several capacities ranges from 0 to 1, and it does not differ in any sort of measurements whether it is research, test assessment, or any other types of measurements. Furthermore, Factor alpha (known as "Cronbach's alpha") as the greatest generally-used consistency coefficient provides evaluations of test score dependability from sole exam management by material from the relationship among test items internal to the test; henceforth, it is also named an internal consistency constant (Vieira & Gomes, 2010).

In addition to reliability considerations, it seems to be a lack of research on exploring the genre and research method of a large corpus which in turn can guide researchers towards filling more significant gaps in the field. For instance, which skills and subskills are ignored or which research methodology seems to be underestimated by ESP researchers. The present study thus hopes to provide a clear framework for researchers to know a better picture of research trends in a decade of research practice in the field of ESP and improve our research lens concerning possible weaknesses and strengths of research methods and statistical interpretation of research findings. Therefore, it is worth looking back at published articles in the ESP Journal in order to analyze the purposes mentioned above for a better improvement of the ESP field. Indeed, in the current study,

the researchers aim to shed more light on the systematic review of reliability coefficients and research methods in the ESP Journal, and to comparatively provide common consensuses on the acceptability of this subject. The following research questions were considered:

**Research Question One:** How is genre theory viewed and developed in a decade of research in *ESP Journal* from 2010 to 2020?

**Research Question Two:** What are the most frequently-used research methods in the research articles published in *ESP Journal* from 2010 to 2020?

**Research Question Three:** What are the general trends in the findings of *ESP Journal* based on a bibliometric analysis of reliability coefficients of the articles published from 2010 to 2020?

## 2. Methodology

### 2.1. Materials

A systematic review of the ESP Journal makes it possible to identify the trends and methods of scientific research applied frequently. *English for Specific Purposes*, selected as the material to be investigated in this study, is a global peer-reviewed journal publishing four issues yearly. However, issues such as second language acquisition in particular settings, need valuation, syllabus progress and assessment, resources training, discourse examination, metaphors of specialized varieties of English, instruction and challenging methods, the usefulness of several methods to language knowledge and language education, and the exercise or reskilling of educators for the teaching of ESP may be preserved from the viewpoint of English for specific purposes. Besides, the journal contains articles and negotiations that recognize features of *ESP* needing advance areas into which the practice of *ESP* may be extended, probable means of assistance among *ESP* curricula and learners' expert or occupational goods, and insinuations that the result from associated chastisements can have for the occupation of *ESP* (SCImago Journal & Country Rank).

A systematic review of *ESP Journal* was required to draw a representative sample of *ESP Journal*. The current study is not limited to the borders of a country or a geographical area in data collection phase. However, in order to gain a great amount and comprehensive illustration of consistency coefficients, a complete review of all the articles was assumed. Different criteria may be used in systematic reviews in order to analyze the corpus (e.g., Ghanizadeh & Jahedizadeh, 2015; Jahedizadeh & Al-Hoorie, 2021; Jahedizadeh, Ghonsooly, & Hosseini Fatemi, 2019), and for the present study the criteria were only the articles published between 2010 and 2020 in *ESP Journal* from *Elsevier Website*, written in English, and other academic works such as book chapters and reviews were extracted. This means that 279 articles published in this journal were systematically and individually examined to locate the type of reliability, genre, and method. Thus, a sample of 279 EFL articles published in time intervals 2010-2020 was selected for the systematic review investigation.

### 2.2. Procedure

To begin the survey and build bibliometric datasets about the research questions, all the *ESP Journal* articles published over a current 11-year extent of time were downloaded and the corpus was prepared. The researchers decided to analyze *ESP Journal*, since it publishes articles on various topics such as: second language acquisition in specialized contexts, needs assessment, curriculum development and evaluation, materials preparation, discourse analysis, descriptions of specialized varieties of English, teaching and testing techniques, the effectiveness of various approaches to language learning and language teaching, and the training or retraining of teachers for the teaching of *ESP*. Thus, it contains a thorough and comprehensive corpus to study.

For the next stage, all the titles of the EFL published articles were extracted and typed in an excel document file format, accompanied by the publication year. In order to facilitate the procedure, all articles were coded based on the publication year. That is, the number of each article

was combined with the last two digits of the publication year. For instance, the first article was entitled “Collaborative writing: Bridging the gap between the textbook and the workplace”, and it was published in 2010, so it was coded *110*. As another example, the last article entitled “A text analysis and gatekeepers’ perspectives of a promotional genre: Understanding the rhetoric of Fulbright grant statements” published in 2020 was coded 3020 (30 articles were published in 2020, and it was the last one). The remaining articles followed the same procedure.

### 2.3. Data Analysis

To obtain frequency of reliability estimate, SPSS software was utilized. Also, the frequency of reliability type (internal consistency, inter-rater reliability & intra-rater reliability) reported in the examined articles was estimated. Moreover, through SPSS the frequency of the employed method (qualitative method, quantitative method, & mixed methods) and genre (skill, discipline, and data collection design) in all *ESP* articles was investigated.

As stated by Lan et al. (2009), human processing might be affected by the drawback of subjectivity. Therefore, the process was carried out carefully and when needed, comments were sought from experts in this field to achieve a more objective classification and comparison. After a careful analysis of the articles in question, a precise meta-review of all articles was conducted by all the authors. Each article was analyzed to obtain the established genres. There are various types of genres in academic writing, including empirical, case studies, literature reviews, reports, reflective diaries, etc. In this regard, the genre of the studies was classified in terms of the skills, disciplines, and data collection methods reported in the published articles. This classification was based on the scarcity of systematic reviews considering these three criteria. Due to its variations, they were not coded by letters. Regarding the observed reliability which included *reliability estimates*, if the reliability type was reported in the articles, they were coded as (**Y**) and if not, it was coded as (**N**). The next step in the analysis was *reliability type* where it concentrates on the type of reliability utilized in the targeted articles which were coded as internal consistency (**RA**), inter-rater reliability (**RB**) and intra-rater reliability (**RC**). Finally, the last step in the analysis phase was the type of the methods implemented in the current study which was codified in three types and the qualitative method was codified as (**MA**), quantitative as (**MB**) and mixed methods as (**MC**). Reliability types and selected codes, method types along with genres and how they are coded accordingly are all displayed in Table 1 in order to facilitate comprehension of the coding procedure.

Table 1: Coding Scheme of the Articles

Method type	Reliability type
1. Qualitative Method (MA)	1. Internal Consistency (RA)
2. Quantitative Method (MB)	2. Inter-rater Reliability (RB)
3. Mixed Methods (MC)	3. Intra-rater Reliability (RC)

Once these data were gathered, we imported all data to SPSS software to obtain the frequency of each classification.

### 3. Results

Table 2 shows the frequency of all articles published between 2010 to 2020 in the *ESP* Journal. As can be seen, 2015 has the highest frequency followed by 2020.

Table 2: The Frequency of Published Articles During 2010-2020

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	2010	20	7.2	7.2	7.2
	2011	21	7.5	7.5	14.7
	2012	22	7.9	7.9	22.6
	2013	19	6.8	6.8	29.4
	2014	28	10.0	10.0	39.4
	2015	31	11.1	11.1	50.5
	2016	27	9.7	9.7	60.2
	2017	27	9.7	9.7	69.9
	2018	25	9.0	9.0	78.9
	2019	29	10.4	10.4	89.2
	2020	30	10.8	10.8	100.0
	Total	279	100.0	100.0	

Figure 1 also visualizes the frequencies in form of a bar graph and pie chart with the highest frequency in year 2015 (N=31), and the lowest publication frequency in the year 2013 (N=19).

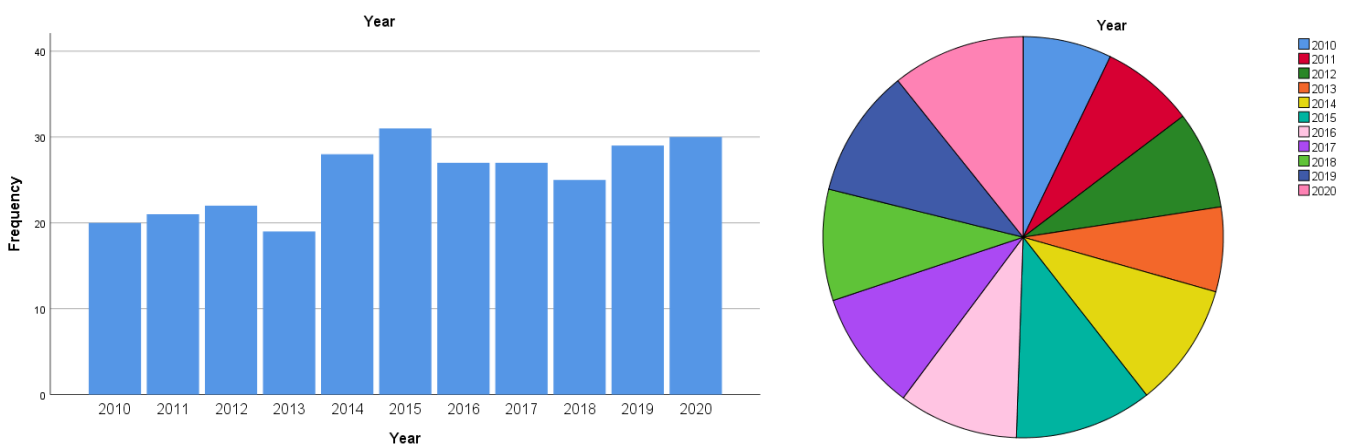


Figure 1: The Frequency of Published Articles During 2010-2020

In order to answer the first research question and explore the genre of published articles, three main categories are involved including the skill which was investigated, the discipline in which the study was conducted, and the method of data collection to reach the findings. Each of these sections was analyzed in terms of their frequency and percentage.

3.1. Skills

Table 3 demonstrates the frequency and percentage of the skills which studied by the authors.

Table 3: The Frequency and Percentage of the Investigated Skills

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Writing	168	60.2	60.2	60.2
	Speaking	54	19.4	19.4	79.6
	Reading	2	.7	.7	80.3
	Listening	2	.7	.7	81.0
	Vocabulary	29	10.4	10.4	91.4
	Pronunciation	2	.7	.7	92.1
	Grammar	4	1.4	1.4	93.5
	General English	18	6.5	6.5	100.0
	Total	279	100.0	100.0	

As can be seen among the four skills (writing, speaking, reading, listening) and three sub-skills (vocabulary, pronunciation, and grammar), writing has the highest frequency which implies most of studies investigated this skill in order to do their analysis. General English also refers to those articles which considered all skills or the skill was not specified.

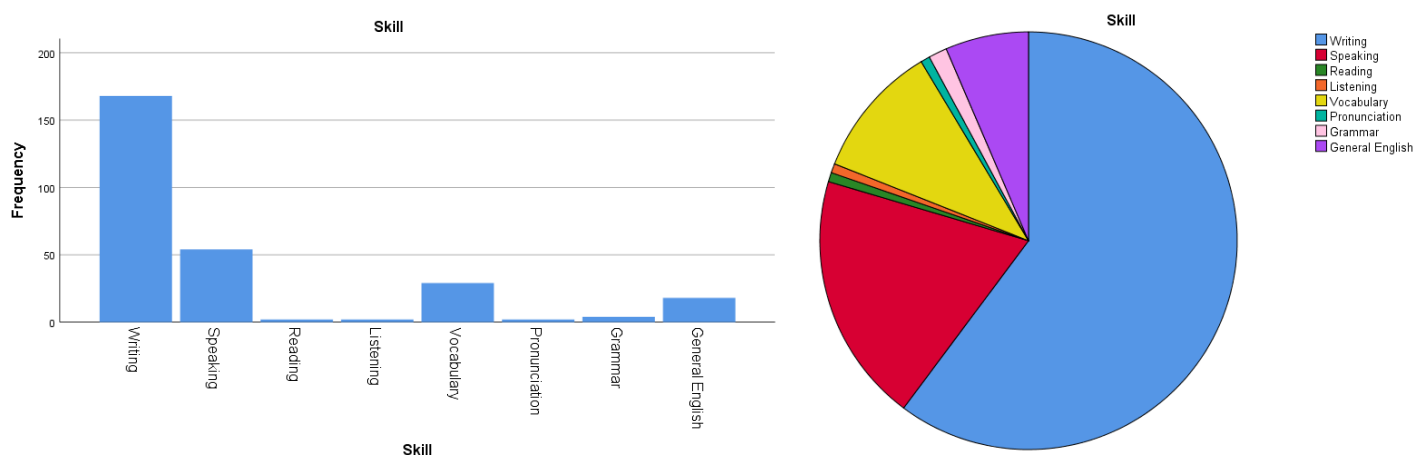


Figure 2: The Frequency of the Skills and Subskills Studied in the Articles

### 3.2. Disciplines

Table 4 depicts the frequency of disciplines on which the studies were conducted. The codification of its data was based on analyzing all fields mentioned in the articles. Other subject areas with only one occurrence and studies investigating numerous disciplines were included in the Miscellaneous category. As can be seen, Business has the highest frequency, followed by English and Health Sciences (Medicine, Nursing, Dentistry).

Table 4: The Frequency and Percentage of the Investigated Disciplines

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Business	38	13.6	13.6	13.6
	Industry	5	1.8	1.8	15.4
	Engineering	22	7.9	7.9	23.3
	Health Sciences	26	9.3	9.3	32.6
	Computer Sciences	5	1.8	1.8	34.4
	Law	13	4.7	4.7	39.1
	Education	3	1.1	1.1	40.1
	Biology	3	1.1	1.1	41.2
	Marketing	2	.7	.7	41.9
	English	26	9.3	9.3	51.3
	Social Sciences	12	4.3	4.3	55.6
	Linguistics and Applied Linguistics	15	5.4	5.4	60.9
	Economics and Finance	8	2.9	2.9	63.8
	Agriculture	2	.7	.7	64.5
	Physics	3	1.1	1.1	65.6
	Aviation	3	1.1	1.1	66.7
	Information Systems	3	1.1	1.1	67.7
	Mathematics and Statistics	6	2.2	2.2	69.9
	Technology	5	1.8	1.8	71.7
	Miscellaneous	74	26.5	26.5	98.2
	Chemistry	1	.4	.4	98.6
	Accounting	4	1.4	1.4	100.0
	Total	279	100.0	100.0	

Figures 3 and 4 visualize the same results.

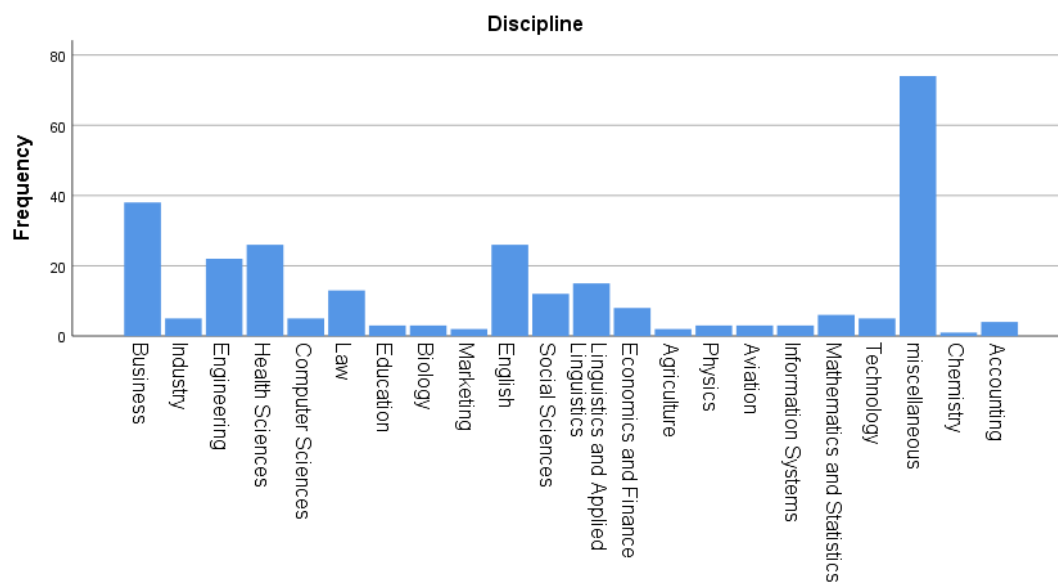


Figure 3: The Bar Chart of the Frequency of the Disciplines Studied in the Articles

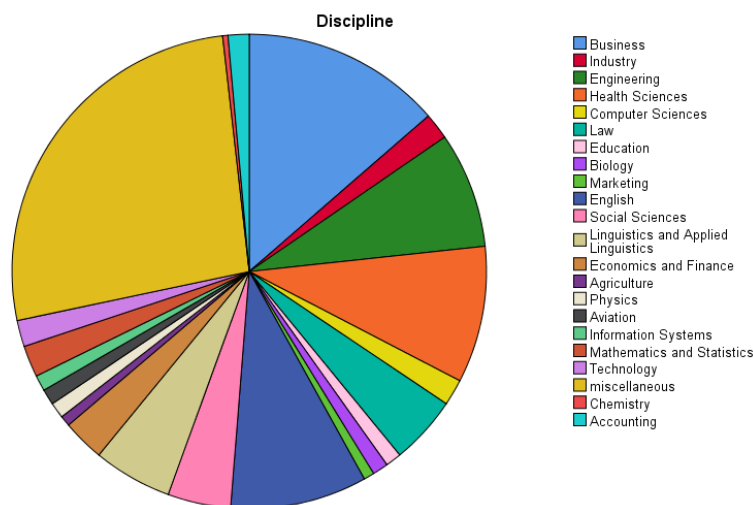


Figure 4: The Pie Chart of the Frequency of the Disciplines Studied in the Articles

### 3.3. Data Collection Design

The 279 studies utilized a variety of designs to conduct the analysis. Due to considerable variation, the most effective methods were extracted and coded. Other designs were included in the Other Methods category. This group also consists of articles that reported numerous designs. The results can be seen in Table 5.

Table 5: The Frequency and Percentage of the Reported Data Collection Designs

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid Corpus Analysis	121	43.4	43.4	43.4
Genre Analysis	15	5.4	5.4	48.7
Textbook Analysis	4	1.4	1.4	50.2
Both Questionnaire and Interview	21	7.5	7.5	57.7
Questionnaire or Test	10	3.6	3.6	61.3
Interview	24	8.6	8.6	69.9
Narratives	3	1.1	1.1	71.0
Case Study	18	6.5	6.5	77.4
Comparative Analysis	5	1.8	1.8	79.2
Field Notes	3	1.1	1.1	80.3
Other Methods	39	14.0	14.0	94.3
Experimental	7	2.5	2.5	96.8
Reflections	2	.7	.7	97.5
Move Analysis	7	2.5	2.5	100.0
Total	279	100.0	100.0	

Figures 5 and 6 demonstrate these results in the forms of a bar and a pie chart.



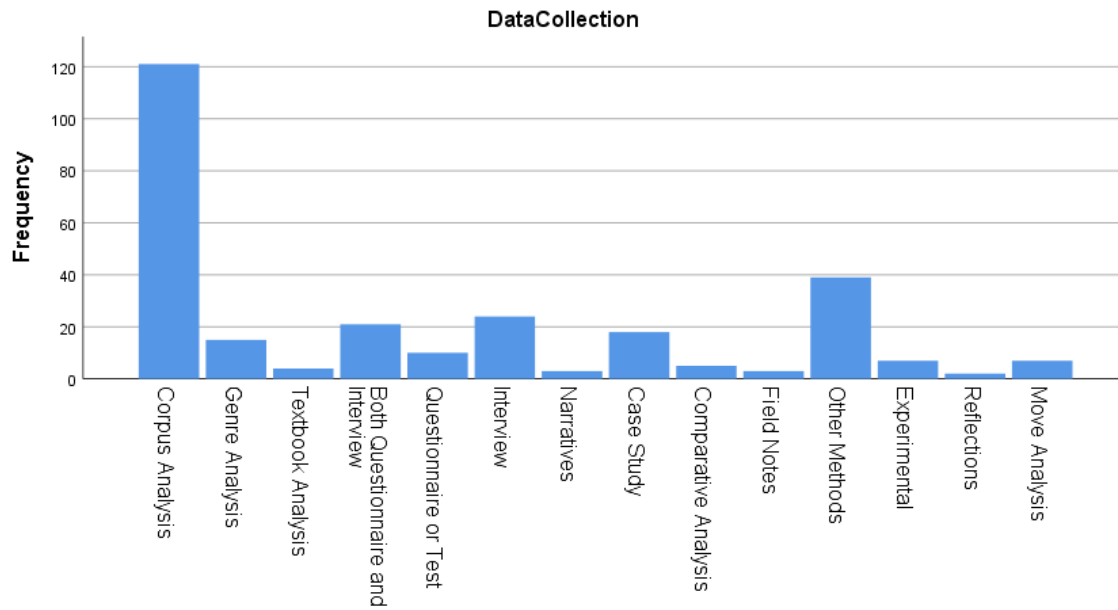


Figure 5: The Bar Chart of the Frequency of the Data Collection Designs Reported in the Articles

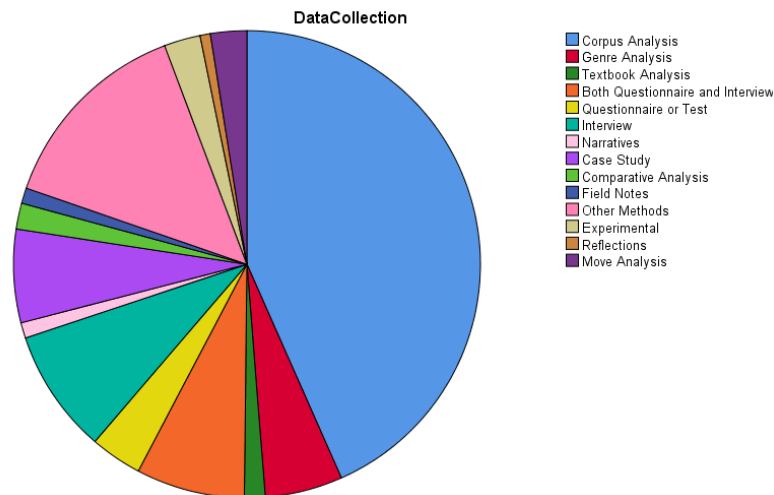


Figure 6: The Bar Chart of the Frequency of the Data Collection Designs Reported in the Articles

In order to answer the second research question, descriptive analysis was used to analyze the frequency and percentage of methods of research in the *ESP* articles. Table 6 shows the frequency and percentage of the methods of research employed by the authors.

Table 6: The Frequency and Percentage of the Methods of Research

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Qualitative Method	102	36.6	36.6	36.6
	Quantitative Method	32	11.5	11.5	48.0
	Mixed Methods	145	52.0	52.0	100.0
	Total	279	100.0	100.0	

Three types of methods of research were explored among **279** articles which are Qualitative Method, Quantitative Method, and Mixed Methods. As it can be seen in Table 6, among three methods of research, Mixed Methods has the highest frequency (f=145, p=52%) and Quantitative Method has the lowest frequency (f=32, p=11.5%). In addition, results of

descriptive analysis indicated that 36.6% of the articles (f=102) used Qualitative Method for their study. Figure 7 shows the bar and pie charts of the frequency of three types of methods of research.

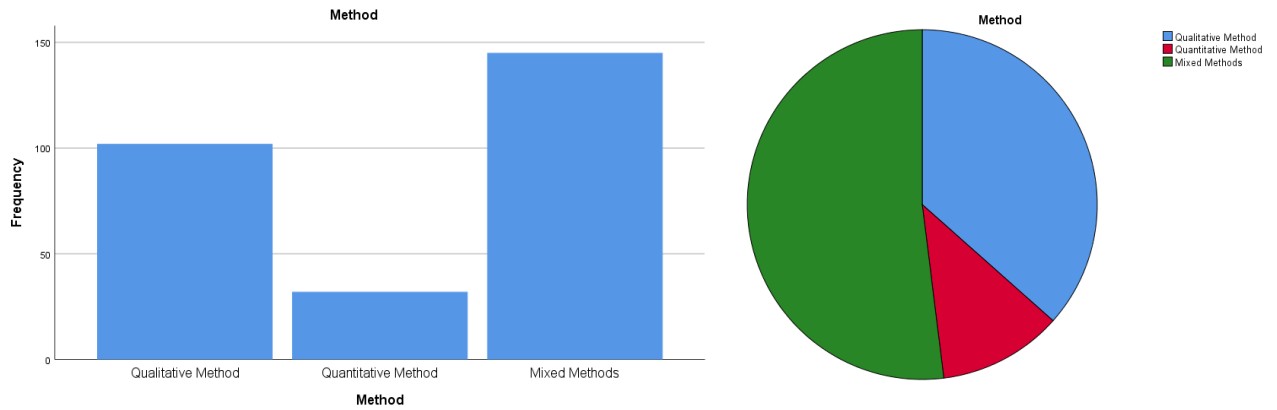


Figure 7: The Bar and Pie Charts of the Frequency of Three Types of Methods of Research

In demand to answer the third research question, descriptive analysis was used to analyze the frequency and percentage of the reliability coefficients of the 279 articles. Table 7 shows the frequency and percentage of the articles which reported and did not report the reliability coefficients.

Table 7: The Regularity and Proportion of the Articles Reporting and not Reporting Dependability Coefficients

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Not reported	193	69.2	69.2	69.2
	Reported	86	30.8	30.8	100.0
	Total	279	100.0	100.0	

As it is shown in Table 7, among 279 articles, only 86 articles (30.8%) reported the reliability coefficients for their study. It also showed that 193 articles (69.2%) did not report the reliability coefficients. Figure 8 shows the pie chart of the percentage of articles which reported and did not report the reliability coefficients.

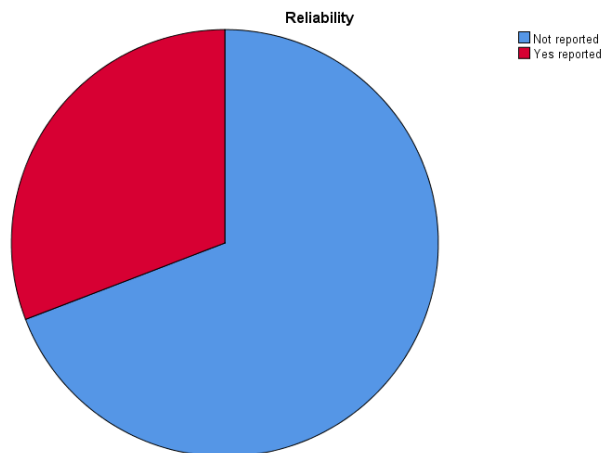


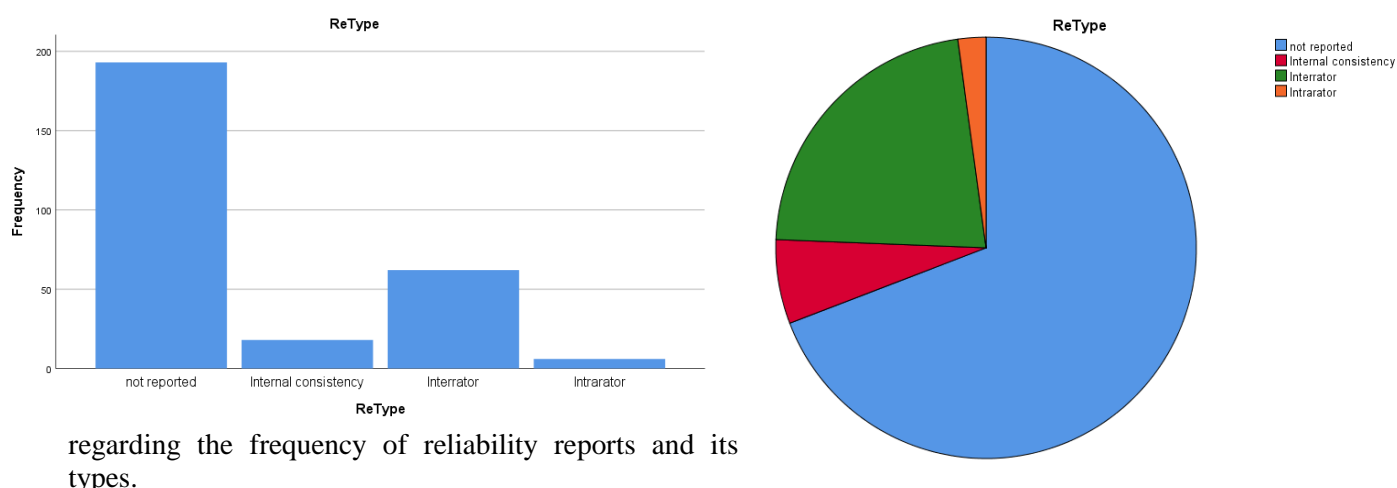
Figure 8: Pie Chart for Reporting the Reliability Coefficients

In order to find out which type of reliability coefficient was reported in these 86 articles, descriptive analysis was used. Table 8 shows the frequency and percentage of reported type of reliability.

Table 8: The Frequency and Percentage of the Types of Reliability

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Not reported	193	69.2	69.2	69.2
	Internal consistency	18	6.5	6.5	75.6
	Inter-rater reliability	62	22.2	22.2	97.8
	Intra-rater reliability	6	2.2	2.2	100.0
	Total	279	100.0	100.0	

Three types of reliability were explored among these articles which are Internal Consistency, Inter-rater reliability and Intra-rater reliability. As it can be seen in Table 8, among three types of reliability in these 86 articles, Inter-rater reliability has the highest frequency ( $f=62$ ) and intra-rater reliability has the lowest frequency ( $f=6$ ). In addition, results of descriptive analysis indicated that 18 articles reported Internal Consistency for reliability. These results can also be seen in Figure 9



regarding the frequency of reliability reports and its types.

Figure 9: The Bar and Pie Charts of the Frequency of Three Types of Reliability

#### 4. Discussion

The present research aimed at examining the tendencies of researchers whose articles have been printed in the arena of ESP during the last ten years. In this regard, 279 papers published in the *ESP Journal* between 2010 and 2020 were studied. Then, each paper was coded and analyzed based on its corresponding category. As the results of the first research question indicated, the genre of most of the articles was corpus analysis. According to McEnery and Hardie (2012), English has been studied via the application of corpora. This is due to the fact that the field of English has developed in the US and UK as the two most critical English-speaking countries. Besides, most of the studies in the *ESP Journal* investigated the writing skill. The fact that this skill is the most significant ability in higher education is undeniable (Walsh, 2010). Hence, researchers attempt to scrutinize it from different angles to provide solutions for students' writing challenges and problems (e.g., McDowell & Liardét, 2020; Wette, 2019). The results also indicated that Business is the most frequent discipline studied in the articles. This can be attributed to the wide use of English as a lingua franca in business (Nickerson, 2005, 2013). In other words, the communication that takes place in business

contexts is of great importance, and thus the issue of written and spoken business discourse has been considered extensively (e.g., Chan, 2019; Handford, 2010; Koester, 2006, 2010; McCarthy, 2020; McCarthy & Handford, 2004).

A methodological foundation in research and language acquisition is required in order to assess the rationality of privileges in the empirical literature, and also to make knowledgeable choices. Since language acquisition is a broad domain and an interdisciplinary field with theoretical, historical, and methodological ties to education, psychology and linguistics, it is valuable manner in attention that confident areas and research question types call for certain methodological approaches. In quantitative research question and design, researchers focus on description, contextualization, and more importantly frequency of occurrences of data; they compute and study their expressive statistics such as revenue, average deviation, assurance interludes, and occurrence amounts.

Considering the second research question, the findings imply that researchers in the ESP field seem to be more interested in mixed methods research, especially in the last three years. In other words, the adoption of mixed methods has been growing since it assists the researcher to understand the phenomena under study deeply and accurately (Brown, 2014; McKim, 2017). Moreover, as Plastow (2016) contends, in mixed-methods research, the strengths of one research method compensate for the weaknesses of another. Through mixed methods research design, the effectiveness of a study is maximized by using the strength of both quantitative (what question) and qualitative (how question) research (Phakite, De Costa, Plonsky & Starfield, 2018).

In spite of all the advantages of qualitative and quantitative research methods, there are drawbacks in both designs. According to Christensen and Johnson (2012), qualitative investigators opinion the social biosphere as being active and border their results to a certain cluster of people being calculated in its place of generalizing them. Moreover, the particular technique used by qualitative researchers might be inaccurate, mistaken, or confusing (Cohen & Morrison, 2011). However, with regard to some barriers in mixed-methods design, the researchers at the beginning of the last decade were not that much interested in employing this type of research. This might be due to the limitations researchers might face such as the high time required and cost involved in the process of data collection, analysis and interpretation.

According to the obtained results on the employment of qualitative research method, about 36.6 percent of researchers appoint qualitative research design which is the second preferred method. There might be several factors that encourage researchers to opt for qualitative research methodology; commonly talking, qualitative educations are more exposed and inductive than their quantitative complements.

A second justification relies on non-numerical data such as pictures and words primarily and provides factual and descriptive information through qualitative analysis of data. Hence, the scheme through which information are recovered in qualitative research method is measured as being exclusive (Daniel, 2016).

Thirdly, an advantage of qualitative research approach is that theory emerges from data. In other words, when theory emerges out of data, it allows the investigator to concept and renovate systems where it is vital; emergence of theory would be based on the data researchers generate, rather than testing data generated by other researchers elsewhere (Maxwell, 2013).

Eventually, according to Lichman (2013), a benefit of using qualitative research approaches is that they scrutinize human beings and conduct in a public setting, that is, interactions, thought, reasoning and norms are examined holistically. Moreover, the connection between the investigator and members facilitate the contribution participants have in order to shape the research.

Taking quantitative research method into consideration, an incommmodity might be the linear and nonflexible nature of this research approach in which the researcher is required to follow a certain order, he/she is in fact at the "driver's seat", and participants have no room to contribute to the study (Bryman, 2001 , p. 286).

Furthermore, the quantitative research design does not require or encourage creative and critical thinking, since it is organized with determined variables, hypotheses and design (Bryman, 2001; Creswell, 2009). As Bryman (2001) argues, the positive point of quantitative study method is the use of numerical information as an instrument for valid time and possessions; that is, it decreases the stretch and energy the researchers would have capitalized in telling the results of their studies. Moreover, statistics, parts and quantifiable statistics can be considered and directed by a processor, which would save a lot of energy and resources (Connolly, 2007).

Since this research approach spaces stress on numbers and statistics in the gathering and investigation of data, it can be viewed as being scientific in nature (Gorard, 2001). Furthermore, researchers might opt for quantitative method of data collection and analysis since generalizations are possible; communication made with one collection can be comprehensive to and be reflective of a wider group. That is, the study results need not be seen as a simple accident when researchers are interpreting their results (Cohen et al., 2011).

According to Lichtman (2013), another benefit with the use of quantitative research approach is its replicability. Simply put, this method essentially depends on theories challenging, thus the investigator wants to follow strong ideas and guidelines rather than doing intelligent guesswork. Shank and Brown (2007) argue that this type of study is showed in a universal or community style due to its clear objective, therefore, can be replicated in other times or places leading to the same results.

Quantitative research approaches might be employed more often by researchers since they give room for the use of regulator and education clusters (Johnson, & Christensen, 2012). To clarify, using governor clusters, for instance, the investigator might agree to split the members into clusters giving them the similar instruction, but using different instruction approaches, behavior in attention the issues that he is learning. In the teaching context, the clusters can be collected and the investigator can then examine the problem-solving capability of the pupils and would be able to admit the teaching technique that has the greatest impacts on students' problem-solving abilities.

Finally, quantitative research is sometimes described as "researcher detachment" (Denscombe, 1998); that is, the researcher might not be in direct contact with the participants, collecting data through telephone, internet, or questionnaires. There is a filled controller for substitutes such as interpretation, explanation, and inferences (Creswell, 2009). Therefore, researchers might select quantitative research design in order to benefit from this issue.

Since reliability concerns the faith a researcher can have in the obtained data and deals with the stability of findings, it needs to be presented precisely (Altheide & Johnson, 1994). With regard to the third research question, the corpus was analyzed and the findings suggested that merely 30.8 percent of the articles have reported reliability coefficient, and about 69.2 percent have left the reliability untouched.

Moreover, when the reliability coefficient is not clearly reported, the readers of the research would not be able to interpret the study results satisfactorily. As Plonsky (2015) insists, confirming inner rationality in research needs, amongst other circumstances, dependable arrangement, however, second language (L2) investigators frequently flop to statement and even more often fail to understand dependability. The problems with reliability coefficient report might vary in different

fields and contexts. Gliem and Gliem (2003), for instance, analyzed the reliability measures of articles published in an agricultural journal, and the suggested results stand in sharp contrast with the results of the current study; that is, the typical 90 percent of the agricultural articles have successfully reported reliability coefficient, though only 7 percent could analyze the reliability correctly. Since reliability is an important feature of a test and indicates the quality and usefulness of that test, researchers need to report the reliability estimates that are relevant for a particular test. When some test raters evaluate responses to questions, differences in judgments might lead to variations in test scores. In order to make sure how consistent test scores are, inter-rater reliability is reported. Based on the findings of this study, inter-rater reliability is reported more often by researchers in comparison with other types of reliability. These results seem not to be in agreement with what Wilhelm, Rouse, and Jones, (2018) mentioned in a study they carried out on differences in measurement and reporting reliability; they suggested that reliability estimates other than inter-rater reliability is focused more by the researchers.

On the other hand, through internal consistency reliability, the extent to which items on a test measure the same thing is indicated, that is, a high internal consistency reliability coefficient means that the items on the test are homogenous. The inner reliability technique estimates how fine the set of objects on an exam associate with one another. According to the results, internal consistency reliability had the second frequency. However, not reporting internal consistency does not necessarily indicate low measures of reliability. As Plonsky (2015) contends, observed reliability is of more considerable reputation than whether or not such data are informed. Therefore, it cannot be claimed that the articles failing to report inner reliability have certainly used unreliable data, though it is of paramount importance to precisely calculate and report this type of reliability.

Through intra-rater reliability, a metric for rater's self-consistency in the scoring of subjects, the same assessment is completed by the same rater on two or more occasions. As the results of this study indicate, reporting intra-rater reliability had the lowest frequency in comparison with other types of reliability. One line of explanation for this finding is that journals might not take this issue so seriously; that is, portion of the accountability deceits with journal assessors and publishing supervisor who have not forced very strict necessities on writers, or have not complete so satisfactorily, with reverence to journalism and construing this type of reliability in their research. Since intra-rater dependability states to the grade of arrangement between repeated supervisions of an exam, if this dimension cannot be achieved consistently, it is hard to assign variations in the reliant on variable to the properties of the self-determining variable. Therefore, articles not reporting intra-rater reliability might not be as reliable as expected and suffer from low levels of consistency. In other words, measurement would be worthless if there is no arrangement among repetitive capacities. As Koo and Li (2016) contend, a low intra-rater reliability could not only replicate the little grade of rater or dimensions arrangement, but might also lead to different interpretations. Koo and Li (2016) go further to mention that it is authoritative for investigators to report full evidence about intra-rater reliability estimates.

Considering the few articles that included reports of intra-rater reliability, it is worthy to mention that authors might need to know more about the significance of this issue. According to Portney and Watkins (2000), reporting intra-rater reliability leads to better communication among researchers; therefore, it is suggested that reporting intra-rater reliability includes software information, model, type, and definition selection.

## 5. Conclusion

The current study was conducted due to a paucity of systematic review on articles published in the field of ESP. The results of this study have several implications and applications in the realm of ESL and EFL, for both researchers and journal editors. The findings of the present study are hoped

to contribute to the current literature on genre, method, and reliability estimates of second/foreign language learning research. Researchers might be well aware of the significant role reliability measures play in yielding consistent, reproducible estimates of what is assumed to be an underlying true score (Plonsky, 2015); however, they sometimes fail to report the reliability estimates and the interpretations of the study results will go under question. Therefore, the results of this study give a clear picture of how much ESP researchers feel obliged to report the reliability coefficients in their studies, and it will help future researchers to be cautious when interpreting the results and employing them in their own research.

Moreover, journal editors and reviewers might benefit from the findings of the current research, since they will be provided with a broad view of ESP articles published in the recent ten years, and might make revisions in rules and regulations of the journals concerning the three issues investigated in the present study.

## **6. Implications and Suggestions for Future Research**

In the current study, some limitations were observed that seem to be unavoidable and would better be considered by future researchers. The current study has only considered a dataset of those articles published between 2010-2020 and the articles published before and after this date were not taken into consideration. This study considered only the ESP Journal, other journals are not examined, so it can be considered as a limitation for this study. The scope of this research, size of sample, and the length of the era were limited by the number of ESP Journals available related to the field of EFL, and which is regarded a limitation in this study. Due to time and cost problems, it will not be possible to cover other research works including book reviews or theses. In addition, regarding the research designs, authors might benefit from the results of this study that pinpoint the low use of quantitative research designs by researchers in the field of ESP. The systematic review of an 11-year time span signified how much interested ESP researchers are in mixed methods research; hence, they might give it a second thought when selecting their research design in future, in order to let the strengths of one research method compensate the weaknesses of another through mixing them. The findings of this study offer insights to the analysis of SLA published articles with regard to the reporting of reliability. Therefore, it is suggested that studies be conducted in other areas of second language acquisition, and the results be compared and contrasted with those of the current research.

Second, this study purported to find out the research designs employed by ESP researchers. It is also interesting to find out whether the same is true about other fields of study or other domains of second/foreign language learning. Third, the current study examined articles published in the recent ten years and excluded others due to limitations of time and space. It would be of great interest if future researchers investigate the same trends in other time spans and compare the results to figure out the similarities and differences. Moreover, the corpus collected for the purpose of systematic review was extracted from ESP Journal in the current research. It is highly recommended that other ESP Journals will be examined within the same period of time and compare and contrast the findings. Last but not least, the coding system is highly dependent on the researcher's choice and creativity. It would be noteworthy to design a novel and different coding system and take other variables except reliability estimates and method types into consideration.

## **7. Acknowledgement**

I would like to thank all those who helped us conduct the research.

## **Declaration of Conflicting Interests**

We do not have any conflicts of interest to declare.

## **Funding Details**

This research did not receive any funding from any organization.

## References

- Altheide, D. L., Johnson, J. M. & Singh, D (1994). Criteria for assessing interpretive validity in qualitative research. In N. K. Denzin & Y. S. Lincoln (Eds.). *Handbook of qualitative research*, (pp. 485-499). Thousand Oaks, CA: SAGE.
- Bhatia, V. K. (1993). *Analysing genre: Language use in professional settings*. London: Longman.
- Bhatia, V. K. (2000). Discourse of philanthropic fundraising. *New Directions for Philanthropic Fundraising*, 22, 95-110.
- Brown, J. D. (2014). *Mixed-methods research for TESOL*. Edinburgh: Edinburgh University Press.
- Bryman, A. (2001). *Social research methods*. New York: Oxford University Press.
- Chan, C. (2019). Long-term workplace communication needs of business professionals: Stories from Hong Kong senior executives and their implications for ESP and higher education. *English for Specific Purposes*, 56, 68-83.
- Chaudron, C. (2001). Progress in language classroom research: Evidence from. *The Modern Language Journal, Modern Language Journal*, 85, 57-76.
- Cohen, L., Manion, L. & Morrison, K. (2011). *Research methods in education*. (7th ed). London: Routledge.
- Connolly, P. (2007). *Qualitative data analysis in education: A critical introduction using SPSS*. London: Rutledge.
- Connor, U., & Gladkov, K. (2004). Rhetorical appeals in fundraising direct mail letters. In U. Connor, & T. A. Upton (Eds.), *Discourse in the professions: Perspectives from corpus linguistics* (pp. 257-286). Amsterdam/Philadelphia: John Benjamins.
- Creswell, J. W. (2009). *Research design qualitative, quantitative and mixed methods approach*. (3rd ed). London: SAGE Publication.
- Curry, M. J., & Lillis, T. (2017). Problematizing English as the privileged language of global academic publishing. In M. J. Curry, & T. Lillis (Eds.), *Global academic publishing: Policies, perspectives, and pedagogies* (pp. 1-22). Bristol, UK: Multilingual Matters.
- Daniel, E. (2016). The usefulness of qualitative and quantitative approaches and methods
- Denscombe, M. (1998). *The good research for small –scale social research project*. Philadelphia: Open University Press.
- Dudley-Evans, T. (1994). Genre analysis: An approach to text analysis for ESP. In C. Malcolm (Ed.), *Advances in written text analyses* (pp. 219-228). London: Routledge.
- Fenton-Smith, B., Humphreys, P., & Walkinshaw, I. (Eds.). (2017). *English medium instruction in higher education Asia-Pacific: From policy to pedagogy. Multilingual Education* (vol. 21) Cham, Switzerland: Springer.
- Ghanizadeh, A., & Jahedizadeh, S. (2015). Teacher burnout: A review of sources and ramifications. *British Journal of Education, Society & Behavioural Science*, 6(1), 24-39.
- Gliem, A., & Gliem, R., (2003). Calculating, Interpreting, and Reporting Cronbach's Alpha Reliability Coefficient for Likert-Type Scales; Continuing, and Community Education, 2003 Midwest Research to Practice Conference in Adult, Continuing, and Community Education, 83-85.
- Gorard, S. (2001). *Quantitative methods in educational research: The role of numbers made easy*. London: The Tower Building.



- Haertel, E. H. (2006). Reliability. In R. L. Brennan (Ed.), *Educational measurement* (4th ed., pp. 65–110). Westport, CT: American Council on Education and Praeger.
- Hanauer, D. I., & Englander, K. (2011). Quantifying the burden of writing research articles in a second language: Data from Mexican scientists. *Written Communication*, 28(4), 1-14.
- Handford, M. (2010). *The language of business meetings*. Cambridge: Cambridge University Press.
- Hwang, Y., & Lin, S. (2010). A study of medical students' linguistic needs in Taiwan. *Asian ESP Journal*, 6(1), 35-58.
- Jahedizadeh, S., & Al-Hoorie, A. (2021). Directed motivational currents: A systematic review. *Studies in Second Language Learning and Teaching*, 11(4), 517-541.
- Jahedizadeh, S., Ghonsooly, B., & Hosseini Fatemi, A. (2019). Student evaluation apprehension: An interdisciplinary review of determinants and ramifications. *Polish Psychological Bulletin*, 50(3), 226-236.
- Javid, M., & Mohseni, A. (2020). English for law enforcement purposes: ESP needs analysis of border guarding officers. *Iranian Journal of English for Academic Purposes*, 9(4), 89-111.
- Johnson, B. & Christensen, L. (2012). *Educational Research, Qualitative, Quantitative and Mixed Approach*. (4th ed). California: SAGE Publication.
- Koester, A. (2006). *Investigating workplace discourse*. London: Routledge.
- Koester, A. (2010). *Workplace discourse*. London: Continuum.
- Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine*, 15, 155-163.
- Kuteeva, M., & Mauranen, A. (2014). Writing for publication in multilingual contexts: An introduction to the special issue. *Journal of English for Academic Purposes*, 13(1), 1-4.
- Lan, L., Lian, Z.W., Pan, L., Ye, Q. (2009). Neurobehavioral approach for evaluation of office workers' productivity: the effects of room temperature. *Build. Environ.* 44(8), 1578-1588.
- Larson-Hall, J., & Plonsky, L. (2016). Reporting and interpreting quantitative research findings: What get reported and recommendations for the field. *Language Learning*, 65(1), 127–159.
- Leedy, P. D., & Ormrod, J. E. (2015). *Practical research. Planning and design* (11th ed.). Boston, MA: Pearson. Jürgen Rudolph.
- Lichtman, M. (2013). *Qualitative research in education: A user's guide*. (3rd ed). USA: SAGE Publication.
- Manzoor, H., Majeed, A., & Munaf, M. (2020). Genre analysis of Civil Engineering's research article introductions. *International Journal of English Linguistics*, 10(2), 322-330.
- Marefat, F., Farahanynia, M., Bolouri, M., Chamani, F., & Soleimani, T. (2021). Generic structure of literature reviews in research articles: Iranian and international journals. *Iranian Journal of English for Academic Purposes*, 10(3), 33-50.
- Maxwell, A. (2013). *Qualitative research design: An interactive approach*. SAGE: Los Angeles.
- McCarthy, M. (2020). Vague language in business and academic contexts. *Language Teaching*, 53(2), 203-214.
- McCarthy, M., & Handford, M. (2004). "Invisible to us": A preliminary corpus-based study of spoken business English. In U. Connor & T. A. Upton (Eds.), *Discourse in the professions: Perspectives from corpus linguistics* (pp. 167-201). Amsterdam: John Benjamins

- McDowell, L., & Liardét, C. (2020). Towards specialized language support: An elaborated framework for Error Analysis. *English for Specific Purposes*, 57, 16-28.
- McEnery, T., & Hardie, A. (2012). *Corpus linguistics: Methods, theory and practice*. Cambridge: Cambridge University Press.
- McKim, A. (2017). The Value of Mixed Methods Research: A Mixed Methods Study. *Journal of Mixed Methods Research*, 11(2), 202-222.
- Nickerson, C. (2005). English as a lingua franca in international business contexts. *English for Specific Purposes*, 24(4), 367-380.
- Nickerson, C. (2013). English for specific purposes and English as a lingua franca. In B. Paltridge & S. Starfield (Eds.), *Handbook of English for specific purposes* (pp. 445-460). Malden, MA: John Wiley & Sons.
- Peterson, R. A., Albaum, G., & Beltramini, R. F. (1985). A meta-analysis of effect sizes in consumer behavior experiments. *Journal of Consumer Research*, 12(1), 97-103.
- Phakiti, A., De Costa, P., Plonsky, L., Swafield, S. (2018). Applied linguistics research: current issues, methods, and trends. In A. Phakiti, P. De Costa, L. Plonsky, S. Swafield (eds). *The Palgrave handbook of Applied Linguistics research methodology* (pp. 5-30). London: Palgrave MacMillan.
- Plastow, N.A. (2016). Mixing-up research methods: A recipe for success or disaster? *South African Journal of Occupational Therapy*, 46(1), 89-90.
- Plonsky, L. (2016). A meta-analysis of reliability coefficient in second language research. *Language Learning*, 61, 993-1038.
- Portney, L.G. and Watkins, M.P. (2000). *Foundations of clinical research: Applications to practice*. 2nd Edition, Prentice Hall Health, Upper Saddle River.
- Saidi, M. & Cheraghi, F. (2020). Genre analysis of artificial intelligence research article abstracts: Local versus international journal. *Global Journal of Foreign Language Teaching*, 10(2), 111-119.
- SCImago Journal & Country Rank (2022). University of Granada, accessed 22 April 2022, <https://www.scimagojr.com/aboutus.php>.
- Shank, G. & Brown, L. (2007). *Exploring educational research literacy*. New York: Routledge.
- Solovova, O., Santos, J., & Veríssimo, J. (2018). Publish in English or perish in Portuguese: Struggles and constraints on the semiperiphery. *Publications*, 6(2), 25.
- Starfield S, 2019, 'Thesis and dissertation writing in a second language: Context, identity, genre', *Journal of Second Language Writing*, <http://dx.doi.org/10.1016/j.jslw.2018.10.002>
- Swales, J. (2004). Research genres. *Explorations and applications*. Cambridge: Cambridge University Press.
- Swales, J. M. (1990). *Genre analysis: English in academic and research settings*. Cambridge: Cambridge University Press.
- Swales, J. M., & Leeder, C. (2012). A reception study of articles published in *English for Specific Purposes* from 1990 to 1999. *English for Specific Purposes*, 31, 137-146.
- Vieira, E & Gomes, J. (2010). Citations to scientific articles: Its distribution and dependence on the article features. *Journal of Informetrics*, 4, 1-13.

Wächter, B., & Maiworm, F. (Eds.). (2014). *English-taught programmes in European higher education: The state of play in 2014*. Bonn, Germany: Lemmens.

Walsh, K. (2010). *The importance of writing skills: Online tools to encourage success*. Retrieved from <http://www.emergingedtech.com/2010/11/the-importance-of-writing-skills-online-tools-to-encourage-success/>.

Wette, R. (2019). Embedded provision to develop source-based writing skills in a Year 1 health sciences course: How can the academic literacy developer contribute? *English for Specific Purposes*, 56, 35-49.

Wilhelm, A., Rouse, G., & Jones, F. (2018). Exploring differences in measurement and reporting of classroom observation inter-rater reliability. *Practical Assessment, Research & Evaluation*, 23(4), 1-16.