

Evaluating the Suitability of Studies Published in the Field of Teaching of Turkish as a Foreign Language According to Qualitative Research Standards

Nurşat Biçer*

Amasya University, PO box 05000, Amasya, Turkey

Corresponding Author: Nurşat Biçer, E-mail: nursatbicer@gmail.com

ARTICLE INFO

Article history

Received: November 26, 2020

Accepted: January 19, 2021

Published: February 28, 2021

Volume: 12 Issue: 1

Advance access: February 2021

Conflicts of interest: None

Funding: None

Key words:

Teaching Turkish,
Language Training,
Qualitative Evaluation,
Literature Review,
Exploratory Design,
Rubric

ABSTRACT

The aim of this study is to investigate to what extent qualitative studies published in scientific journals between 2010 and 2017 on the teaching of Turkish as a foreign language, meet qualitative research criteria. For this purpose, exploratory design from mixed method was used in the study. For the study, 131 articles on the teaching of Turkish as a foreign language, published in Turkish were analyzed. During the course of the analysis of the articles, the Qualitative Research Evaluation Form (QREF), as prepared by the researcher, was used. With the internal consistency reliability analysis, it was found that the form had had a high level of reliability. In the analyses, the topic distributions of the articles and the arithmetic means for the suitability according to qualitative research criteria were designated. Using the variables of the number of authors, the publication year and research design, and the correlation between the article's qualitative research scores were calculated. Finally, a correlation analysis was conducted between the chapters of the articles. It was determined that most of the research was done on course materials and the problems encountered in the examined articles. As a result of the study, it is seen that the mean score of the findings section is high and the method section has a low mean. Articles are mostly 1 or 2 authors. There is a significant difference between article scores according to the publication year and research design. According to the correlation analysis between the sections of the articles, positive significant relationships were found. It is thought that although the articles published also have certain drawbacks about their suitability to the qualitative research criteria, the studies did in fact exhibit an improvement and that the future studies would therefore be of higher quality.

INTRODUCTION

Teaching Turkish to non-native speakers of the language (TFL) is a field that has gained popularity in recent years. The development of Turkish training activities has brought about the question of how this training could be conducted more efficiently. In accordance with the improving status of Turkish-language training, new methods, techniques, and materials, etc. have been introduced into the training process, and quality training is being pursued. Making benefit of the language training studies in the world as well as keeping the field open to recent developments would facilitate the maintenance of the quality of the training services. In this respect, the training activities have brought along with them an increase in scientific studies. These studies will determine what the drawbacks of Turkish language education are, as well as contribute to taking the necessary steps towards that. According to Büyükkiz (2014), universities and other institutions addressing subject of teaching Turkish to non-native speakers in accordance with recent international development, and includes the publishing of theses, articles, books, and presentations.

These studies are expected to exhibit a development in their quality and suitability for scientific methodology, in addition to exhibiting an increase in quantity. This element is indispensable for scientific research, which in turn is beneficial for the science world, researchers, practitioners, and other partners.

The theoretical background of a study, alongside the findings and results obtained in that study, should be reported with regard to certain criteria (Karasar, 2012). The suitability of the studies for scientific methodology and report writing principles has an impact on the contribution of these studies to theoretical and application knowledge in the field. A scientific study is brought into the world of science via reports prepared in accordance with the quality of the research. The academic communication is obtained through the sharing/publishing of new research and discussions in academic platforms with scientific qualities (Aktaş and Uzuner Yurt, 2015; Ozan and Köse, 2014). In order for scientific studies to have the expected impact, they must at least contain the introduction, method, findings and interpretation, discussion, results and recommendations

sections. The common attitude and understanding in the presentation of scientific reports is an important element that would by and large facilitate the scientific/academic communication (Karasar, 2012). It is natural that each field of science develop certain understandings peculiar to its own nature and structure, thus shaping the research in this respect. Although the ultimate goal for the scientific studies is to reach what is real, the differences in the topics addressed has generated new understandings about the courses and methods of the studies.

The paradigms/approaches/methods play a determining role in the preparation of scientific studies with different structures. Among these, the qualitative, quantitative, and mixed research methods guide each phase of the study from the beginning to both the realization and the presentation. Qualitative research methods have been used widely in the social sciences, particularly in research on education. Qualitative research methods provide the presentation of perceptions and events in their natural environment in a realistic and holistic manner using qualitative data collection methods such as observation, interviews, and document analysis (Yıldırım and Şimşek, 2008, p. 39). There may be some differences in the whole or in some parts of the qualitative studies with regard to the emphasis given. In addition, all reports discuss the nature of the problem at hand, the manner of conduct, and the findings obtained (Merriam, 2013, p. 238). Qualitative studies generally contain four main sections, including “introduction”, “method”, “findings” and “conclusion”. The introduction indicates the background and the significance of the study. In this section, the readers are provided with the context of the study, and the relevant studies are mentioned. The method section contains the design of the study, the rationale for the selected design, the roles of the researcher(s), the duration of implementation, the number and selection of the participants, data collection, data analysis methods, and the detailed explanation of this process. In the findings section, the data collected are transformed into a detailed narration and interpreted. Differently from quantitative research, themes and patterns are used in place of statistical results. The conclusion section is the reformulation of the primary focus of the study. The data is re-emphasized in a focused manner. A discussion is developed by importing the results of different studies. Suggestions for different studies and recommendations for practitioners are included in this section (McMillan and Schumer, 2010, p. 37-39; Maykut and Morehouse, 1994; Merriam, 2013; White, Woodfield and Ritchie, 2003).

Different opinions are proposed by researchers as to whether qualitative research, in which the perspectives and interpretations of the researchers stand out, would be shaped with regard to certain rules or not. While some scholars reject qualitative research having evaluation criteria (Bochner, 2000; Dixon-Woods, Shaw, Agarwal, & Smith, 2004; Guba and Lincoln, 2005), others argue that qualitative research should bear certain features (Cohen and Crabtree, 2008; Denzin, 2008; Beverland and Lindgreen, 2010). In many disciplines, experienced researchers encounter the

problem of how qualitative research ought to be evaluated with regards to being scientific. There is a confusion on how to evaluate the relevant studies with particular regard to objectivity, validity, and reliability (Spencer et al., 2003, p. 59). These problems, in turn, complicate the formulation of criteria to be used in the evaluation of these studies (Tracy, 2010). Scholars doing qualitative research particularly avoid standard criteria to guide the selection or the use of research methodology. Rather, they prefer to recognize the diversity and complexity of the research participants and contexts, as well as prefer to study about the limitations and the contexts of the study environments (Northcote, 2012, p. 103). Due to the hardship of formulating criteria for qualitative research, these studies occasionally try to make evaluations using the criteria for quantitative research (Cohen and Crabtree, 2008).

Why do scholars of qualitative research develop criteria despite certain criticism? The reason is that the most important feature of criteria is that they are useful. Rules and instructions facilitate learning and implementation, and contribute to the quality of the study (Tracy, 2010). The impact of certain rules is significant in writing more credible essays. The presence of common practices might help the students, the practitioners and researchers to access the content of the publications more efficiently (Hodge, 2016). The explicit and clear explanation of the procedures in the research process also increases the confidence in the study. The sufficiency of data, the extent of the analysis, transparency and reproducibility of the analysis could be appropriate criteria for the evaluation of qualitative research. A detailed explanation of all of steps taken and procedures implemented and the justification of conduct lie at the core of the qualitative research (Hannes, 2011; Stenius, Mäkelä, Mioviský and Gabrhelík, 2017). Methodological information such as research questions, theoretical knowledge, study design (e.g. the participants of the study, how the participants are selected, how data is collected and analysed) and its rationale are among the elements to be evaluated in the studies (Cohen & Crabtree, 2008). Shenton (2004, as cited in Yıldırım, 2010) asserts that providing flowcharts or diagrams that show how the study has been conducted, in order to ensure the auditability and comprehensibility of the study by the readers within a short period of time, would increase the quality of the study. In qualitative research, which is a scientific process, an evaluation using the criteria such as “meticulous” or “reliable” is essential (Spencer et al., 2003). Some researchers stated that qualitative research should constitute validity, reliability, and objectivity. Although there are different views on how qualitative research should be evaluated and what the criteria to be used in this evaluation should be, the evaluation instruments in qualitative research are acknowledged as being instruments that could be used as a part of the investigation and interpretation process, and that share the basic criteria (Hannes, 2011). The evaluation instruments, as prepared for the assessment of qualitative research to bear certain features scientific research, could be used.

It may be possible to determine the suitability of studies in certain fields for the nature of qualitative research via the criteria in order to be formulated for the evaluation of qualitative research. With the literature reviews in order to attain this aim, it would be easy to determine the impact force of the results of the present studies and to benefit from the fund of knowledge at hand. In this sense, literature reviews should be conducted in the discipline of education as in many other disciplines in order to benefit from these. These reviews would orientate new studies in addition to determining the status of the present scientific research (Erdem, 2011). In almost all disciplines, there are numerous studies that evaluate the subject areas under focus, the methodological features, and the findings presented. Those studies addressing tendencies in disciplines of the teaching of Turkish as mother tongue as well as a foreign language have become prominent over the past five years. These studies have generally been conducted as postgraduate theses. Şahin, Kana, and Varişoğlu (2013), Varişoğlu, Şahin, and Gökteş (2013), Büyükkiz (2014), Ercan (2014), Aktaş, and Uzuner Yurt (2015), Bozkurt and Uzun (2015), Biçer (2017), Özçakmak (2017), Boyacı and Demirkol (2018) and Türkben (2018), have all conducted their research on this subject with this purpose in mind. However, any study analyzing the articles on TFL or any other educational sciences disciplines for that matter could not be found. This study is acknowledged as being a new type of research with regard to article reviews.

When it is considered that the tradition of writing articles paying regard to research approaches in educational sciences, and in particular TFL in Turkey has made much progress in recent years, it is probable that some problems have been encountered in the use of these principles within the context of academic research. Therefore, the evaluation of studies on scientific research methods would thus contribute to the methodological improvement of the articles.

This study aims at determining the extent of compliance with the qualitative research and evaluate the different variables in qualitative research papers on the teaching of Turkish as a foreign language (TFL) published in Turkish scientific journals between 2010 and 2017. For this aim, the studies were scored with a rubric in terms of qualitative research criteria. The data obtained were used in statistical analysis. In this context, answers to the following research questions have been sought:

- What are the ratios of fulfilling the qualitative criteria?
- Is there any statistically significant difference in the suitability to qualitative research scores with regard to the number of authors?
- Is there any statistically significant difference in the suitability to qualitative research scores with regard to the publication year?
- Is there any statistically significant difference in the suitability to qualitative research scores with regard to the research design?
- Are there any correlations between the scores obtained for the sections?

METHOD

Research Design

The exploratory design from mixed method used in this study, among other qualitative research methods, investigates the articles on the teaching of TFL as written in accordance with the qualitative approach, and published between 2010 and 2017. Mixed method is a type of research in which quantitative and qualitative techniques are combined or mixed (Christensen, Johnson & Turner, 2015). In the exploratory design, a type of mixed method, qualitative information is first collected, analyzed, and this information is used to improve the quantitative follow-up phase of the data collection process (Creswell and Plano Clark, 2011). The researcher starts with exploring qualitative data, then uses these findings in the quantitative research dimension (Creswell, 2014). Quantitation was made based on the qualitative data collected in the study. Quantitative rubric was developed to evaluate qualitative data and qualitative data was evaluated with this tool. The suitability of these articles for the criteria identified in the Qualitative Research Evaluation Form (QREF), created by the researcher, was evaluated. Thus, quantitative evaluation of the articles prepared according to the qualitative research method was made.

Data Collection

For the QRSF created for the review of the articles, first, the literature on the qualitative research methods was reviewed. In this respect, a draft form comprising of 34 items containing the criteria that should be present in qualitative research articles was prepared basing on studies such as those by Maykut and Morehouse (1994), Robson (2002), Des Jarlais, Lyles and Crepez (2004), Bogdan and Biklen (2007), Yıldırım and Şimşek (2008), McMillan and Schumacher (2010), Merriam (2013), Patton (2014), and Creswell (2015). These criteria were gathered under the following categories: Introduction, Method, Findings and Conclusion. The language and expression suitability of the items were checked. The internal validity of the form was determined. In order to determine the internal validity of the form, views of 8 experts were taken, 3 from the discipline of assessment and evaluation, 3 from the discipline of Turkish education, and 2 from the discipline of curriculum. The content validity ratios of the items in the forms were calculated in line with these expert views. The minimum values of the CVRs at the $\alpha=0.05$ significance level were tabulated in order to ease the calculation for testing the statistical significance of the CVRs (Vneziano and Hooper, 1997, as cited in Yurdugül, 2005). Accordingly, the minimum values about the number of experts and the content validity ratios obtained for the Qualitative Research Survey Form are presented in Table 1:

According to Table 2, the values that the items should bear for the content validity ratios are given according to the number of experts. Since 8 experts are used in this study, the items must have a minimum value of 0.78.

After calculating the CVR values, the Content Validity Index (CVI) was also calculated. The CVI is obtained using

the total CVR averages of the items that are significant at the $\alpha=0.05$ level, and that are to be included in the final survey form. CVI values are valid for the sub-dimensions, and are obtained for each sub-dimension considering the items in that sub-dimension (Yurdugül, 2005).

Nine items with low content validity scores were removed from the initial form with 34 items being in line with the expert views. Some corrections were made in some items in the final form with 25 items.

The items in the form were collected under four categories: introduction, method, findings and conclusion. The items were prepared as a five-point scale: “1=Very insufficient”, “2=Insufficient”, “3=Partially Sufficient”, “4=Sufficient” and “5=Very sufficient”. Ten articles were reviewed using the final form, whereupon the suitability of the items were evaluated. Some improvements were made in certain items in the form with the preliminary review process.

The Cronbach’s Alpha coefficient for the internal consistency interpretation of the form was found to be 0.908. When the confidence intervals of an assessment tool is considered,

it is seen that this ratio is very high (Büyüköztürk, 2011; Kılıç, 2016).

Research Inclusion Criteria

When determining the studies to be reviewed within the scope of this study, the articles on the teaching of TFL published in Turkish scholarly journals between 2010 and 2017 were based on. The Ulakbim, Google Academics, Sobiad and DergiPark indexing systems were used to assess the articles. Keywords such as “Turkish as a foreign language, teaching Turkish to foreigners, language training, Turkish training, teaching Turkish as a foreign language” were used in the searches. At the end of the search, 175 qualitative studies were found. Duplicate articles and conference presentations, as well as articles that were out of date range, and that did not have a method section were excluded, thereby leaving being 131 articles to be used in this study.

Validity and Reliability Study of the Encoding Process

Each article was evaluated according to the form upon encoding them as being M1, M2, and M3, etc. The relevant sections of the articles reviewed within the scope of this study were read and processed in the form. For the calculation of the reliability of the form, 25 articles were randomly selected from among the reviewed articles and were rated by a second rater. In order to determine the reliability between the two raters, the Pearson correlation analysis was conducted. Accordingly, the correlation coefficient between two raters was found to be 0.82. Since the correlation coefficient should be 0.70 at minimum for an assessment tool in order to exhibit stability (Karakoç and Dönmez, 2014), it could be argued that the reliability of the rating was high.

Data Analysis

While the data was being analyzed in this study using content analysis, the status of the articles for fulfilling the criteria was determined. The arithmetic means of the suitability of the articles to the criteria were calculated. In order to determine whether or not there was any statistically significant difference between the scores obtained and the independent variables of the number of the authors, publication date, and research

Table 1. Qualitative research survey form and content validity ratios

Item	CVR	Item	CVR
1	0,25	18	1,00
2	1,00	19	1,00
3	1,00	20	1,00
4	1,00	21	1,00
5	1,00	22	0,50
6	1,00	23	1,00
7	1,00	24	1,00
8	1,00	25	1,00
9	1,00	26	1,00
10	1,00	27	1,00
11	0,25	28	0,50
12	1,00	29	1,00
13	0,50	30	0,50
14	0,50	31	1,00
15	0,50	32	1,00
16	1,00	33	1,00
17	0,50	34	1,00

Table 2. Minimum values of the CVR’s at the $\alpha=0.05$ significance level

Number of Experts	Minimum Value	Number of Experts	Minimum Value
5	0.99	13	0.54
6	0.99	14	0.51
7	0.99	15	0.49
8	0.78	16	0.42
9	0.75	17	0.37
10	0.62	18	0.33
11	0.59	19	0.31
12	0.56	20	0.29

method, the unpaired t-test, and ANOVA were used. In order to determine the relationship between the criteria categories that the articles should have, the correlation analysis was conducted. IBM SPSS 17.0 software was used for these operations. Sample sections of the articles were also used in the investigation of the suitability to the article evaluation criteria.

FINDINGS

This section contains the findings about the data obtained. Table 3 presents the subject distribution of the articles reviewed:

Here, it is seen that course book (f: 20), teacher and student opinions (f: 18), problems encountered (f: 17), error analysis (f: 12) and material (f: 11) are frequently discussed in the articles reviewed. It is seen that the studies on course books are more frequent than studies on other subjects. It was found that history of teaching Turkish as a foreign lan-

guage (f: 1), curriculum (f: 1), teaching listening (f: 1) and academic Turkish (f: 1) were not popular. With reference to these data, it can be asserted that studies on these subjects are few in number, given that topics about the history of Turkish training are addressed mostly on theoretical grounds, that there is not any standard curriculum used in this discipline, that the listening skill is difficult to evaluate, and given that academic Turkish is fairly a new topic.

When the arithmetic means of the qualitative research methods used in the articles reviewed are considered (Figure 1), it is seen that the introduction section has 3.00, the method section has 2.58, the findings section has 3.22, and the conclusion section has an arithmetic mean of 2.70.

It can be argued that the item (Table 4) "The purpose of the study is expressed explicitly" has the highest mean with 3.88, and that the articles generally have a fair mean in fulfilling this criterion. It can also be stated that, with the item "the insufficiencies in the literature are presented" with the 1.83 mean, the articles have significant insufficiencies in terms of fulfilling this criterion. When the overall status of the introduction section is considered, it can be concluded that the introduction sections are partially sufficient. When it is considered that the introduction section is written with great rigor in numerous fields of science, it can be argued that there is not a significant problem in the articles reviewed. An example for expressing the purpose of the study explicitly in the articles reviewed is presented below:

This study aims at reviewing the mobile applications developed for teaching Turkish to foreigners, and to put on the mobile application markets with regard to developers, extent, method, and language of instruction. (M16).

It is seen that the item (Table 5) "the data collection tools and their properties are suitable for the purpose of the study"

Table 3. Subject distribution of the articles reviewed

Subject	f
Course book	20
Teacher and student opinions	18
Problems encountered	17
Error analysis	12
Material	11
Grammar teaching	9
Acculturation	8
Psychological features	6
Teaching writing	6
Model suggestions	5
Teaching reading	5
Teaching speaking	4
Teaching vocabulary	2
Literary texts	2
Needs analysis	2
History of Turkish-language education	1
Curriculum	1
Teaching listening	1
Academic Turkish	1
Total	131

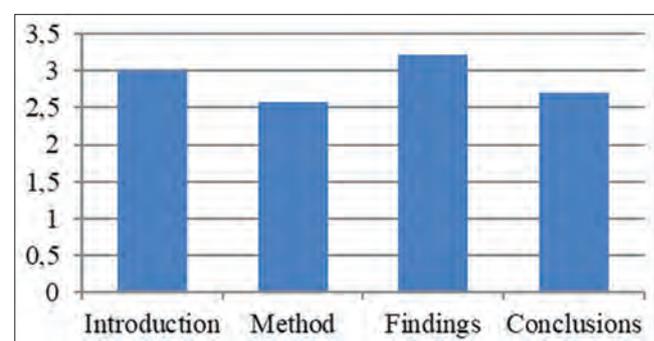


Figure 1. Arithmetic Means of the Sections of the Studies

Table 4. Arithmetic means of the items in the introduction section

Sub-dimension	Item No.	Items	Mean
Introduction	1.	The introduction contains the general extent of the study.	3.70
	2.	The literature is relevant to the purpose of the study.	3.63
	3.	The insufficiencies in the literature are presented.	1.83
	4.	The rationale of the study is expressed explicitly.	2.50
	5.	The purpose of the study is expressed explicitly.	3.88
	6.	The research questions are related to the purpose of the study.	2.45
General Arithmetic Mean			3.00

Table 5. Arithmetic means of the items in the method section

Sub-dimension	Item No.	Items	Mean
Method	7.	The research design is expressed explicitly.	3.21
	8.	The rationale for the selection of the research design is explained.	2.19
	9.	The participants/sample and their properties are elaborated.	3.17
	10.	The selection method for participants/sample is expressed.	1.77
	11.	The data collection process is elaborated in accordance with the design (i.e. duration, use of resources, etc.)	2.72
	12.	The data collection tools and their properties are suitable for the purpose of the study.	3.40
	13.	Information about the validity and reliability studies for the assessment tools are provided.	1.94
	14.	The data analyses procedures are elaborated.	2.74
	15.	Information on how validity and reliability are ensured in the analyses is provided.	2.11
General Arithmetic Mean			2.58

in the method section has the highest mean with the score of 3.40, and it is thought that there is no serious problem in the suitability of the data collection tools for the studies. However, it is seen that the item “The selection method for participants/sample is expressed” item has mean of 1.77, whereupon insufficient results have been obtained. Accordingly, it is understood that there are serious problems in providing the methods used in the selection of the sample/study group. When the overall status of the method section is considered, it is concluded that the mean is at the insufficient level. Certain insufficiencies can be regarded as being reasonable considering that the method section has only been recently focused on field training studies. A sample extract relevant to the suitability of the data collection tools and their properties to the purpose of the study is presented below:

In accordance with the purpose of the study, the data were collected using a semi-structured interview form. By means of the interaction, flexibility, and probes from the interview provided to the researcher (...) an attempt was made to reveal the perceptions and beliefs of Bosnian instructors about teaching Turkish as a foreign language. The 10 items in the form were prepared by the researchers and submitted to the evaluation of four experts, teaching Turkish to foreign students. At the end of the evaluation, two questions, which were thought to be out of scope, and one question, which was considered to contain the same sense were excluded from the semi-constructed interview form. A preliminary study was conducted by administering the interview form comprising of seven questions to three instructors, who had not participated in the study, and the interview form was made final. (M11).

Table 6. Arithmetic means of the items in the findings section

Sub-dimension	Item No.	Items	Mean
Findings	16.	Different participant views/documents are included.	3.23
	17.	Quotations are given from the participant views/documents are included.	2.95
	18.	Personal opinions of the researcher are included.	3.16
	19.	The findings are presented in accordance with the purpose of the study.	3.41
	20.	The findings are supported with visual elements (tables, charts, photographs, etc.).	3.33
General Arithmetic Mean			3.22

When the criteria in the findings section are evaluated (Table 6), it is seen that they have close and average ratios in general. It is found that the item “the findings are presented in accordance with the purpose of the study” has the highest mean with a score of 3.41, and that the most important item of the findings, which is the suitability to purpose, is higher than the other items. On the other hand, it is understood that

the item “quotations are given from the participant views/ documents” has the lowest mean with a score of 2.95, and this element, which is important for the reliability of the data, is paid less attention. When all of the items of the findings section are considered, it is concluded it has a partially sufficient mean with 3.22. When compared to other sections, it is understood that the highest mean is in the findings section, and that the findings are considered to be relatively more important than other sections. However, the flexibility of the qualitative studies in presenting the findings causes the researchers to construct their narrations in a rather different manner. An example about including the personal opinions of the researcher is presented below:

What instructors should do is to orientate the learners towards the correct answer by allowing them to think instead of correcting them immediately after they give a wrong answer. They should try to make the students find the correct answer by asking different questions. Thus, in doing so, learning can be more lasting and effective. (M70).

When the criteria in the conclusion section are considered (Table 7), it is seen that the distribution of the criteria differ from each other. It is found that the items “making suggestions for application” (3.60), “the suggestions being relevant to the findings” (3.18), and “presentation of the results in accordance with the findings” (3.13) have a higher mean than the other items. However, it is found that the items “comparison with the findings of the previous studies” (2.05) and “making suggestions for researchers” (1.56) have very low ratios. When the overall status of the conclusion section is considered, it is concluded that it is insufficient with a score of 2.70. This suggests that the articles reviewed did not have any opinions about leading other researchers. An example extract about making suggestions for researchers is presented below:

It is considered that the researchers might regard the steps in Taba-Tyler’s curriculum development model as being study topics based upon this study for future research to be conducted on curriculum development steps in the teaching of TFL. (M60).

When the number of the authors of the articles that were reviewed are considered, it is seen that the number varies between 1 and 4. It is understood that the articles mostly have 1 or 2 authors, and there is a small number of other articles as well. When Table 8 is examined, it is seen that articles with 2 authors have the highest mean. The lowest mean, on the other hand, is seen in articles with 4 authors. Articles with either 1 or 3 authors fall between these two categories. When this situation is considered, it can be concluded that

generally articles with more than one authors have higher means.

The One Way Analysis of Variance (ANOVA) was conducted to understand whether there was a statistically significant difference between the suitability to qualitative research scores of the articles with regard to the number of authors (Table 9).

The variance value of the suitability of the qualitative research scores of the articles with regard to the number of authors variable ($F=2.219$; $p>0.05$) was not found to be statistically significant. This finding indicates that there is no statistically significant difference between the suitability to qualitative research scores of the articles with regard to the number of the authors. The descriptive statistics of the article scores with regards to the publication year are presented in Table 10.

It can be argued that the highest number of articles published between 2010 and 2017 were published in 2015, and the lowest number of articles were published in 2011. The number of the studies are higher in recent years, and there were few studies published between 2010 and 2012. When Table 10 is examined, it is seen that the articles published in 2017 have the highest means, and that the articles published in 2010 have the lowest means. It is intriguing that the mean of the articles published in 2016 is also higher. When the overall condition is considered, it can be argued that the mean scores of the articles have increased over recent years. In this respect, it is understood that the articles have become more and qualified with each passing year, and likewise the research methodology has also improved as well.

The One Way Analysis of Variance (ANOVA) was conducted in order to understand whether there was a statistically significant difference between the suitability to qualitative research scores of the articles with regards to the publication year variable (Table 11).

The variance value of the suitability to qualitative research scores of the articles with regard to the publication year variable ($F=5.426$; $p<0.05$) was found to be statistically significant. Accordingly, it is seen that there is a statistically significant difference between the suitability for qualitative research scores of the articles with regards to the publication year. In order to understand amongst in which groups the difference is significant, the Scheffe test was conducted. At the end of the test, it is seen that the difference is seen between mean scores of 2010 and 2016, and 2016 and 2017. Accordingly, it is concluded that the articles published in 2010 are at a low level, and that the articles published in 2016 and 2017 are at a higher level. In this sense, it is understood that articles, in which meth-

Table 7. Arithmetic means of the items in the conclusion section

Sub-dimension	Item No.	Items	Mean
Conclusion	21.	The results are presented in accordance with the study findings.	3.13
	22.	The results are compared with the findings of the previous studies.	2.05
	23.	Suggestions for application are made.	3.60
	24.	Suggestions for researchers are made.	1.56
	25.	Suggestions are relevant to the findings of the study.	3.18
General Arithmetic Mean			2.70

odological criteria tends to be used more, have been more recently authored.

The descriptive statistics of the article scores with regard to the research design are presented in Table 12.

When the research designs of the articles are considered, it is seen that phenomenology, case study, action study, descriptive research and document analysis are used. It is seen that document analysis and descriptive study designs are used more frequently than other designs. When Table 12 is examined, it is seen that articles using phenomenology as research design have the highest means, and given that articles using document analysis have the lowest means. When the means of the research designs used in the articles are examined, it is understood that those that are less frequently used have higher means, where as that are frequently used have lower means. The One Way Analysis of Variance (ANOVA) was conducted to understand whether there was a statistically significant difference between the suitability of qualitative research scores of the articles with regards to the research method variable (Table 13).

The variance value of the suitability to qualitative research scores of the articles with regard to the research design variable (Table 13) ($F=12.699$; $p<0.05$) was found to be statistically significant. Accordingly, it is seen that there is a statistically significant difference between the suitability to qualitative research scores of the articles with regard to the research design. In order to understand among which groups the difference is significant, the Scheffe test was conducted. At the end of the test, it is seen that the difference is seen between mean scores of articles using phenomenology/case study and descriptive study/document analysis designs. Moreover, correlation analysis was conducted in order to understand whether or not there is a correlation between the categories of the suitability to qualitative research scores of the articles (Table 14).

According to the correlation analysis conducted between the scores for the introduction, method, findings and con-

clusion sections of the articles, alongside the total score, positive correlations with a significance level of $p>0.05$ were found between the sections (Table 14). Accordingly, as the mean score of the sections of the articles increases, so the mean scores of other sections as well. The correlation level between the mean score of the method section and the total mean score is higher than other sections. This indicates that the method section determines the overall status of the article.

CONCLUSION AND SUGGESTIONS

When the results of this study are considered, it is seen that problems encountered are prominent with regard to the subject of the articles (e.g. course books, teachers, student opinions). On the other hand, it can be argued that the topics such as the history of Turkish-language training, teaching listening, curriculum, and academic Turkish are less popular. When the subjects of the articles are considered in general, it becomes apparent that the researchers wish to study the most popular subject. Bozkurt and Uzun (2015) found that articles on materials such as course books and literacy skills were conducted frequently. Büyükkiz (2014) stated that grammar reviews were in the majority; Biçer (2017) argued that material suggestions were studied more. However, it is found that listening and speaking skills were studied less in the articles (Büyükkiz, 2014; Özçakmak, 2017; Boyacı and Demirkol, 2018).

When the reviewed sections of the articles are considered, it is seen that the findings section has the highest mean, and that the method section has the lowest mean. It is thought that while the finding section in qualitative research has a higher means as it enables flexibility and lacks a standard format, the method section has a low mean given that it requires a certain level of technical knowledge and skill. However, it is noteworthy that there are problems in the method section, which is significant in order for an article to be deemed scientific. It is seen that in most of the articles reviewed by Aktaş and Uzuner Yurt (2015), the abstract lacked information about methodology. According to Dönmez and Gündoğdu (2016), it was found that in some of the studies reviewed, the method of the research is not explicitly expressed, and that the studies lacked the method section. It can be argued that the problems encountered about the methodology in the studies originate from a lack of knowledge (Aktaş and Uzuner Yurt, 2015; Dönmez and Gündoğdu, 2016). When the introduction sections of the articles are examined, it is implicit that although the articles are at a high level in expressing the purpose of the study explicitly, they are insufficient at expressing the deficiencies in the litera-

Table 8. Arithmetic mean and standard deviation values of articles' suitability to qualitative research scores with regard to the number of the authors

Number of Authors	N	\bar{X}	SD
1	60	67.2333	15.04836
2	59	74.8475	20.09359
3	9	73.1111	12.92715
4	3	62.0000	9.53939
Total	131	70.9466	17.60783

Table 9. Variance analysis of the suitability to qualitative research scores of the articles with regard to the number of authors

	Sum of squares	s.d.	Mean of squares	F	p.
Intergroup	2007.377	3	669.126	2.219	0.089
Intragroup	38297.249	127	301.553		
Total	40304.626	130			

ture. However, it is worth mentioning that the articles are at a higher level with regard to the purpose of the study, thus indicating that the studies are conducted in accordance with a determined purpose. Tracy (2010) states that the literature reflecting the subject should be given richly.

When the method sections of the articles are evaluated, it is seen that the item about data tools being suitable for the purpose of the study has a higher mean than other items. On the other hand, it can be argued that there are some insufficiencies in terms of expressing the selection method of the study group. As data collection tools are important in achieving the aims of the study, the researchers have little in the way of problems in terms of selecting data collection tools suitable to the purpose of the study. However, it can be asserted that attention is not paid to the selection of the participants, and that the studies are conducted with the participation of the persons in the immediate circle. By elaborating upon the selection of the participants, the data collection and data analysis phases, the reliability of the study is increased (Anfara, Brown and Mangione, 2002). According to Devers (1999), the inclusion of information about the selection of participants is one of the qualitative research criteria. Tracy (2010) states that detailed information should be given about data collection and analysis procedures. Cohen and Crabtree (2008) argue that appropriate and rigorous methods should be used. In this respect, it can be argued that the insufficiencies in the method section directly affects the scientific qualification of the qualitative studies.

When the findings section, which provides information about the solutions to the problem (Karasar, 2012) it is seen that the items in this category have both a higher mean than other categories, and that they are more balanced. Tracy (2010) states that the data should be suitable for the purpose and have anesthetic. According to Cohen and Crabtree

(2008), the research report should be clear and consistent. The item about presenting the findings in accordance with the purpose of the study has the highest mean, whereas the item about providing quotations from participant opinions has a lower mean. According to Creswell (2015, p. 219), the researchers allow for participant opinions, in addition to encoding the text with the language of qualitative research. When this insufficiency is considered, it is considered that avoiding quotations poses a reliability problem for the studies. Spencer et al. (2003) state that good qualitative research should use an interpretive research framework.

When the conclusion sections of the articles are reviewed, it is seen that while providing suggestions about practice has a high mean score, whereas the item about providing suggestions for the researchers has a very low mean score. This suggests that the articles are mostly written to guide the practitioners. However, through suggestions for researchers, one of the suggestion types for enabling or facilitating the solution of the problem based on the judgements (Karasar, 2012), the articles pioneer the researchers following the article.

It is understood that the articles are mostly with 1 or 2 authors, and a small number of other articles. Parallel to this finding, Varışoğlu, Şahin and Göktaş (2013) and Biçer (2017) in Turkish language education articles, and Dönmez and Gündoğdu (2016) and Ozan and Köse (2014) in curriculum discipline had all found that the articles generally have either 1 or 2 authors. At the end of the t-test, which conducted to determine whether or not there was a difference between the scores of the articles with regard to the number of authors, no statistically significant difference could be found. Although it was thought that different perspectives would be reflected in the qualitative study with the increase in the number of authors, it is understood that it does not have any impact on the qualifications of the articles. At the end of the ANOVA test conducted to understand whether there was a difference between the scores of the articles with regards to the publication year of the articles, a statistically significant difference between those articles published in 2010 versus those published between 2016 and 2017. According to this finding, it can be claimed that the articles have further improved with regard to the qualitative research criteria. It is seen that document analysis and descriptive study designs are used more frequently than other designs. Ercan (2014) reviewing the postgraduate theses on the teaching of TFL, argued that they are predominantly conducted using document analysis. Şahin, Kana and Varışoğlu (2013) found that descriptive theses are predominant in Turkish training postgraduate theses. At the end of the analysis conducted to see whether or not there was a difference between the scores of the articles with regard to the research design of the articles,

Table 10. Arithmetic mean and standard deviation values of articles' suitability to qualitative research scores with regard to the publication year

Publication Year	N	\bar{X}	SD
1. 2010	9	52.7778	8.64259
2. 2011	5	62.2000	10.28105
3. 2012	9	63.7778	15.25433
4. 2013	17	71.5882	14.91668
5. 2014	21	68.8571	14.92744
6. 2015	36	67.7222	15.31127
7. 2016	16	82.6875	18.07288
8. 2017	18	83.8889	19.98790
Total	131	70.9466	17.60783

Table 11. Variance analysis of the suitability to qualitative research scores of the articles with regard to the publication year

	Sum of squares	s.d.	Mean of squares	F	p.	Difference
Intergroup	9509.588	7	1358.513	5.426	0.000	1-7,
Intragroup	30795.038	123	250.366			1-8
Total	40304.626	130				

Table 12. Arithmetic mean and standard deviation values of articles' suitability to qualitative research scores with regard to the research design

Research design	N	\bar{X}	SD
1. Phenomenology	5	100.4000	9.58123
2. Case study	28	81.4286	20.09870
3. Action research	2	97.5000	21.92031
4. Descriptive Research	40	67.5750	13.53985
5. Document Analysis	56	64.5357	13.36204
Total	131	70.9466	17.60783

Table 13. Variance analysis of the suitability to qualitative research scores of the articles with regard to the research design

	Sum of squares	s.d.	Mean of squares	F	p.	Difference
Intergroup	11580.365	4	2895.091	12.699	0.000	1-4,5
Intragroup	28724.261	126	227.970			2-4,5
Total	40304.626	130				

Table 14. The correlation between the category scores of the articles

		Introduction	Method	Findings	Conclusion	Total
Introduction	r	-				
Method	r	0.480**	-			
Findings	r	0.452**	0.562**	-		
Conclusion	r	0.597**	0.565**	0.502**	-	
Total	r	0.767**	0.874**	0.753**	0.795**	-

** . $P < 0.01$

it is seen that there is a significant difference between the descriptive research – document analysis designs and both the phenomenology and case study designs. Accordingly, it is understood that technical features are paid more attention in the phenomenology and case study, which has become popular very recently, the document analysis and descriptive studies are conducted using the format of ordinary analysis.

When the correlation analysis between the scores obtained for the sections of the articles, positive significant correlations are found between the sections. Accordingly, the scores of each section of the articles increase as the scores of other sections increase. The fact that the method section has the highest correlation with the total score indicates that the method section is at the core of the article. The method section determines the validity and reliability of the study by providing details on how the study is conducted (Büyüköz-türk et al., 2011).

The evaluation of qualitative research according to specific criteria depends on the research field as well as the research paradigm and epistemological beliefs (Northcote, 2012). According to the results of the study, articles in the field of teaching Turkish as a foreign language were evaluated in terms of their compliance with qualitative research criteria. It was determined that most of the research was done on course materials and the problems encountered in the examined articles. As a result of the study, it is seen that the mean score of the findings section is high and the

method section has a low mean. Articles are mostly 1 or 2 authors. There is a significant difference between article scores according to the publication year and research design. According to the correlation analysis between the sections of the articles, positive significant relationships were found.

When the findings of the study are evaluated, the positive and negative elements of articles on teaching TFL are determined with regard to their suitability to qualitative research criteria. It is important to see that the articles exhibit an improvement for suitability for qualitative research criteria with regards to years. In addition, the effective use of these criteria in the recently popularized designs is encouraging for the future studies. Providing courses on this topic to students wanting to become TFL teachers/experts during their graduate and post-graduate studies would contribute in the publication of better-structured articles. In addition, enabling researchers with a pre-assessment of their articles with regard to suitability to scientific qualifications would also remove many insufficiencies. This kind of study review in the discipline of TFL and/or any other discipline is considered valuable for research methodology.

REFERENCES

- Aktaş, E. & Uzuner Yurt, S. (2015). An content analysis for article abstracts in Turkish education area. *Turkish Studies - International Periodical for the Languages,*

- Literature and History of Turkish or Turkic*, 10(7) 73-96, <http://dx.doi.org/10.7827/TurkishStudies.8121>
- Anfara, V. A., Brown, K. M., & Mangione, T. L. (2002). Qualitative research on stage: Making the research process more public. *Educational Researcher*, 31(7), 28-38.
- Beverland, M., & Lindgreen, A. (2010). What makes a good case study? A positivist review of qualitative case research published in *Industrial Marketing Management*, 1971–2006. *Industrial Marketing Management*, 39(1), 56–63.
- Biçer, N. (2017). An analysis study on the published articles relating to the field of teaching Turkish to foreigners. *Pamukkale University Journal of Social Science Institute*, 27, 236-247.
- Bochner, A. (2000). Criteria against ourselves. *Qualitative Inquiry*, 6, 266-272.
- Bogdan, R. C. & Biklen, S. K. (2007). *Qualitative research for education* (Fifth edition). London: Pearson.
- Boyacı, S. & Demirkol, S. (2018). An examination of doctoral dissertations in the field of teaching Turkish. *Journal of Mother Tongue Education*, 6(2), 512-531.
- Bozkurt, B. Ü. & Uzun, N. E. (2015). A review of literature on Turkish language education: Trends in international conferences, *Journal of Language Education and Research*, 1(2), 1-15
- Büyükkikiz, K. K. (2014). An investigation on graduate dissertations written on Turkish teaching as a foreign language, *Mustafa Kemal University Journal of Social Sciences Institute*, 11(25), 203-213.
- Büyükköztürk, Ş. (2011). *Manual data analysis for the Social Sciences*. Ankara: Pegem A Publications.
- Büyükköztürk, Ş. et. al. (2011). *Scientific research methods*. Ankara: Pegem Academy.
- Christensen, L. B., Johnson, R. B., & Turner, L. A. (2014). *Research methods, design, and analysis*. Pearson.
- Cohen, D. J. & Crabtree, B. F. (2008). Evaluative criteria for qualitative research in health care; controversies and recommendations. *Annals of Family Medicine*, 6(4), 331-39.
- Creswell, J. (2014). *Research design: qualitative, quantitative, and mixed methods approaches*. (4th ed.). Thousand Oaks, CA: Sage.
- Creswell, J. W. (2015). *Qualitative inquiry & research design, choosing among five approaches*. Ankara: Siyasal Publishing.
- Creswell, J. W., & Plano Clark, V. L. (2011). *Designing and conducting mixed methods research*. (2nd ed.). Thousand Oaks, CA: Sage.
- Denzin, N. K. (2008). The new paradigm dialogues and qualitative inquiry. *International Journal of Qualitative Studies in Education*, 21, 315-325.
- Des Jarlais, D. C, Lyles, C., & Crepaz, N. (2004). Improving the reporting quality of nonrandomized evaluations of behavioral and public health interventions: *The TREND Statement*. *American Journal of Public Health*, 3(94), 361-366.
- Devers, K. J. (1999). How will we know “Good” qualitative research when we see it? beginning the dialogue in health services research. *HSR: Health Services Research*, 34(5), Part II, 1153-1188.
- Dixon Woods, M., Shaw, R. L., Agarwal, S. & Smith, J. A. (2004). The problem of appraising qualitative research. *QSHC*, 13, 223-225.
- Dönmez, B. & Gündoğdu, K. (2016). Content analysis of articles and theses on Turkish lesson curriculum published between the years of 2000-2016. *International Journal of Turkish Literature Culture and Education*, 5(4), 2109-2125.
- Ercan, A. N. (2014). Yabancılar Türkçe öğretimi üzerine yapılmış lisansüstü tezlerin çeşitli değişkenler açısından analizi. M. V. Coşkun (Ed.), *Türkçenin Eğitimi-Öğretimine Yönelik Çalışmalar*, Ankara: Pegem Akademi.
- Erdem, D. (2011). Türkiye’de 2005–2006 yılları arasında yayımlanan eğitim bilimleri dergilerindeki makalelerin bazı özellikler açısından incelenmesi: Betimsel bir analiz. *Journal of Measurement and Evaluation in Education and Psychology*, 2(1), 14-147.
- Guba, E. G., & Lincoln, Y. S. (2005). Paradigmatic controversies, contradictions, and emerging confluences. In N. K. Denzin & Y. S. Lincoln (Eds.), *The Sage handbook of qualitative research* (3rd ed., pp. 191-216). Thousand Oaks, CA: Sage.
- Hannes, K. (2011). Critical appraisal of qualitative research. In: Noyes J, Booth A, Hannes K, Harden A, Harris J, Lewin S, Lockwood C (Eds.), *Supplementary guidance for inclusion of qualitative research in cochrane systematic reviews of interventions*. Available from URL <http://cqrng.cochrane.org/supplemental-handbook-guidance>
- Hodge, D. R. (2016). The art and craft of writing qualitative manuscripts a primer. *The Indian Journal of Social Work*, 77(3), 253-270.
- Karakoç, F. Y. and Dönmez, L. (2014). Basic principles of scale development. *Tip Eğitimi Dünyası*, 40, 39-49.
- Karasar, N. (2012). *Scientific research method*. Ankara: Nobel Publisher.
- Stenius, K, Mäkelä, K, Miovský, M and Gabrhelik, R. (2017). How to write publishable qualitative research. In: Babor, T. F., Stenius, K., Pates, R., Miovský, M., O’Reilly, J. and Candon, P. (eds.) *Publishing Addiction Science: A Guide for the Perplexed*, pp. 155–172. London: Ubiquity Press.
- Kılıç, S. (2016). Cronbach’s alpha reliability coefficient. *Journal of Mood Disorders*, 6(1), 47-48.
- Maykut, P. and Morehouse, R. (1994). *Beginning qualitative research a philosophic and practical guide*. London: The Falmer Press.
- McMillan, J. H. and Schumacher, S. (2010). *Research in education: Evidence-based inquiry*. (7th Edition). London: Pearson.
- Merriam, S. B. (2013). *Qualitative research, A guide to design and implementation*. Ankara: Nobel Publisher.
- Northcote, M. (2012). Selecting criteria to evaluate qualitative research. In M. Kiley (Ed.), *Narratives of Transition: Perspectives of Research Leaders, Educators & Postgraduates*. Paper presented at the 10th Quality in Postgraduate Research Conference, Stamford Grand, Adelaide, pp. 99-110. Canberra, Australia: The Centre for Higher Education, Learning and Teaching. The Australian National University. Retrieved from <http://www>

- qpr.edu.au/wp-content/uploads/2015/09/QPR_2012_proceedings-1.pdf
- Ozan, C. and Köse, E. (2014). Research trends in curriculum and instruction. *Sakarya University Journal of Education*, 4(1), s. 116-136.
- Özçakmak, H. (2017). New trends of the postgraduate researches on Turkish language education (2011-2015). *International Journal of Turkish Literature Culture and Education*, 6(3), 1607-1618.
- Patton, M. Q. (2014). *Qualitative research & evaluation methods*. Ankara: Pegem Academy Publishing.
- Robson, C. (2002). *Real world research*. Oxford: Blackwell Publishing.
- Spencer, L., Ritchie, J., Lewis, J., & Dillon, L. (2003). *Quality in qualitative evaluation: A framework for assessing research evidence*. London: National Centre for Social Research, Government Chief Social Researcher's Office, UK.
- Şahin, E. Y., Kana, F. & Varışoğlu, B. (2013). The research trends of postgraduate dissertations in Turkish education departments. *International Journal of Human Sciences*, 10(2), 356-378.
- Tracy, S. J. (2010). Qualitative quality: Eight "Big-Tent" criteria for excellent qualitative research. *Qualitative Inquiry*, 16(10) 837–851
- Türkben, T. (2018). Evaluation of postgraduate studies conducted on teaching Turkish as a foreign language. *International Journal of Turkish Literature Culture and Education*, 7(4), 2464-2479.
- Varışoğlu, B., Şahin, A. & Göktaş, Y. (2013). Trends in turkish education studies. *Educational Sciences: Theory & Practice*, 13(3), 1767-1781.
- White, C., Woodfield, K. and Ritchie, J. (2003). Reporting and presenting qualitative data. Ritchie, J. and Lewis. J. (Eds.) *Qualitative Research Practice: A Guide for Social Science Students and Researchers*. Sage Publications, London.
- Yıldırım, A. & Şimşek, H. (2008). *Qualitative research methods in social sciences*. Ankara: Seçkin Publishing.
- Yıldırım, K. (2010). Raising the quality in qualitative research. *Elementary Education Online*, 9(1), 79-92.
- Yurdugül, H. (2005). *Ölçek geliştirme çalışmalarında kapsam geçerliği için kapsam geçerlik indekslerinin kullanılması*, XIV. Ulusal Eğitim Bilimleri Kongresi, Pamukkale Üniversitesi Eğitim Fakültesi, 28–30 Eylül Denizli, Retrieved from <http://yunus.hacettepe.edu.tr/~yurdugul/3/indir/PamukkaleBildiri.pdf> adresinden alınmıştır.