International Journal of Assessment Tools in Education
2020, Vol. 7, No. 1, 130–144
https://dx.doi.org/10.21449/ijate.700244
Published at http://www.ijate.net http://dergipark.org.tr/en/pub/ijate *Research Article*

A Study on Developing Scale for Teacher Perceptions towards Spelling Rules

Ali Turkel [1,*]

[1] Dokuz Eylul University, Buca Education Faculty, Izmir, Turkey

ARTICLE HISTORY

Received: 01 October 2019
Revised: 14 February 2020
Accepted: 07 March 2020

KEYWORDS

Spelling mistakes,
Scale development,
Teacher perceptions,

Abstract: The purpose of this study is to develop a scale to determine Primary Education and Turkish Language teachers' perceptions regarding the frequency of spelling mistakes which their students make. Therefore, this experimental form made to serve this goal was presented to field specialists in terms of consultation; and each item in the form was regulated to ensure content validity rates in accordance with the feedback provided by the specialists. Items with the validity rate below 0.80 were omitted from the form. The trial form consisting of 34 items was administered to 232 Primary Education and Turkish Language teachers who teach at schools under the jurisdiction of Ministry of National Education (MoNE) via e-mail, and the gathered data were analyzed. With the help of Exploratory Factor Analyses (EFA), a four-dimensioned construct with 19 items including frequently made mistakes regarding acronyms, spelling of conjunctions and suffixes, spelling of capital letters, spelling of compound words, and the spelling of the words affected by word formation processes. In the analyses, the relations between the sub-scales of the original scale were taken into consideration, revealing that factors had positive and significant relationships and sub-dimensions were the constituents of the general structure named spelling mistakes, which made up the upper structure. Goodness of fit indices (GFI) of the model was detected to be quite high. Confirmatory Factor Analysis (CFA) administered to the second research group justified the EFA results. The internal consistency coefficient, calculated as .91, for the entire scale was found to be quite reliable.

1. INTRODUCTION

Writing, as acknowledged, is one of the four basic skills to be developed in the teaching of the first language and includes various dimensions. Both Turkish Language and Primary Education teachers aim at developing these various sub-dimensions in harmony with a careful consideration for balance in the process of teaching. One of the sub-dimensions to be developed is the skill for implementing spelling rules. In the teaching curricula (Ministry of National Education [MoNE], 2006, 2008), spelling rules planned to be taught were presented separately according to grade levels. Starting from the first grade in primary school, teaching of spelling rules gradually moves from easy to hard. However, the literature focusing on the application of these rules reflects certain complaints concerning the high number of problems related to learning these rules.

CONTACT: Ali Türkel ✉ ali.turkel@hotmail.com Dokuz Eylul University, Buca Education Faculty, Turkish Language Teaching Dept. İzmir, Turkey

ISSN-e: 2148-7456 /© IJATE 2020

According to Bayat (2013), language of writing (written language) is the language of education and enlightenment. In his study where he focused on pre-service teacher's writing process, Bayat (2013) advocates that the written language of pre-service teachers can provide clues about their teaching competencies. In a general perspective, writing skills of the students can generate outputs associated with their general thinking skills.

According to Aksoy (1985, 1990) language of writing is of three stages in terms of its narration features as follows: accurate writing, good writing and calligraphy. In *accurate writing*, which is the first prerequisite and step of writing, it is fundamental to express the aimed meaning in accordance with complete and precise language rules. Indeed, writing accurately is important for people to communicate correctly, to fulfill the functions necessary for their lives, and essentially to express themselves correctly in every field. Thusly, spelling rules and punctuations are the constructs which set a series of rules for the sake of accuracy to prevent misunderstandings and form a shared written language among people. In the up-to-date Turkish Dictionary of Turkish Language Institution (2019), *writing* is defined as *"Transcription of a language into written form by following certain rules, spelling"* while *spelling rules* is depicted as *"The rules determining the ways the words in a language are written"*. The ambiguity that emerges when spelling rules are not abided prevents the messages from being transferred as intended (Kıbrıs, 2010, p.128). Aksoy (1985,1990), who defined spelling mistake as "the spellings which do not follow the rules in spelling dictionary and the spelling in the word index of this dictionary", stated that there are various spelling mistakes committed.

In Turkey, rules concerning the spelling are set with efforts of committees elected by Turkish Language Institution (TLI), and these committees hold the right to make changes and updates to the rules from time to time. Even though these changes and updates can possibly cause confusions, it seems only natural for this institution to make changes in its views at some points in time. Whereas European countries like Germany, Portugal, Greece, and France possess such a dictionary as in Turkey, some countries like England and Italy do not have similar spelling dictionaries. It is acceptable that Turkey, who adopted the Latin alphabet not until 1928, experiences problems related to spelling more often than other Western countries do. Besides, it is known that Turkey faced debates regarding spelling even when she still used Arabic alphabet (Ünver, 2008; Kannas, 2019).

Considering the literature, the prominent reasons for students to make spelling mistakes are the lack of knowledge about the correct spelling of a word and forgetting the spelling of words due to learning them incorrectly or using them very rarely. These circumstances can be explained with lack of sufficient reading and writing. There can be uses defying the rules of grammar and Spelling Dictionary such as the misuse of suffixes (ağlıyan, yapmassa), failure to employ word formation processes (kitapı, ağaçı), writing of "ki, de" conjunctions and "mi" adverb, using slang and local dialects, incorrect hyphenation of words at the end of the lines, writing the words as pronounced in spoken language (peki, abi, etc.), the use of intensive adjectives (bembeyaz, sapsarı), confusion of resembling words (eğer-eyer, öyle-öğle, saç-sac), the use of circumflexes and apostrophes, writing of Arabic and Persian noun phrases, writing of proper nouns, titles, numbers, acronyms, and foreign words (Aktaş & Gündüz, 2003, p.115-116).

In addition to the study conducted by Aktaş and Gündüz (2003) in which a classification was made regarding the topics where students committed mistakes, the literature includes studies that dealt with students' levels of accuracy in spelling. Among the studies, Bağcı (2011) revealed in his research concerning the 8th grades that, with a rate of 69%, topics where the students were least successful were the spellings of suffix "-ki", conjunction "de", suffix "-DA", and softening of the consonents (Bağcı, 2011, p.702). In another research, Karagül (2010) attempted to determine the 6th through 8th grade students' levels of implementing spelling and punctuation rules suggested by Turkish class syllabus. As for the results of the study, the 6th

grade students generated a success rate of 70% in implementing spelling rules whereas the 7th graders showed a success rate below 70% concerning "spelling the words that can be confused with one another", "spelling the numbers", and "spelling the capital letters". Additionally, considering the 8th graders, even though they were successful at spelling rules at 70% rate, they made frequent mistakes in the sense of spelling "ki conjunction" and "mi question tag".

Besides the students' justifications for committing spelling mistakes as determined by Aktaş (2003), a study carried out by Türkel, Yaman and Aksu (2017) accounts for the dimension of teachers.That is, the study by Türkel et al. (2017) focusing on Turkish language and primary education teachers' perceptions regarding spelling rules and spelling mistakes of students, revealed that teachers reported problems caused by the spelling rules of TLI. Accordingly, it was found that teachers regarded the spelling rules established by TLI as inconsistent at 66% rate. In connection, the teachers explained this inconsistency with frequently changing rules, excessive number of exceptions, the lack of consensus on the explanations, and inclusion of memorized items. Furthermore, to teachers, the increase in students' accurate spelling can be ensured by presenting reasonable and fixed rules, diminishing exceptions, encouraging reading books rather than emposing a rule-oriented nature, and avoiding frequent changes in rules in the spelling guide.

Bayat (2019), in the study that focused on text construction process, highlighted fundamental constituents of the process, and underlined the implementation of spelling rules as one of the three important factors that lead to exposing the quality of the text's internal structure.

While listing the stages employed in the writing process, Flower and Hayes (1981) mentioned revision process which included evaluation and editing sub-skills. Editing contains internal elements as well as external elements such as spelling and punctuation. Similar thought resurfaces in the last step of Kellog's writing model, and Kellog stated that this step comprises of reading and editing (Kellog, 1996). The editing concept appears as the last step of Zamel's (1983) Cognitive Writing Process. Cho (2003), Gebhard (1983), Sommers (1988), Alamargot and Chanquoy (2001), who dealt with writing process and attempted to determine how this process was handled in various writers, emphasized similar concepts by using certain terms.

To equip students with an effective written expression skill, it is vital to determine the mistakes committed in the context of writing. It is not other than teachers who would know about the frequency of mistakes committed by students in a sub-heading. Therefore, a scale to categorize teacher views related to students' spelling mistakes would be quite beneficial. To help reduce spelling mistakes, making sense out of the mistake process by categorizing the mistakes with the help of a teacher perceptions scale and identifying the steps to be taken for the sake of a solution might provide a considerable support. Motivated by this, a scale towards collecting Primary Education and Turkish Language teachers' perceptions was developed to determine the frequency of spelling mistakes committed by the students.

2. METHOD

In the methodology section of the research, the research model, participants, the development process of the data collection tool, production of the items, preparation of the trial form, ensuring the content validity, data collection process, and pre-test stage are included.

2.1. Research Model

The study takes on a survey research design which aims to determine the construct validity and internal consistency of the "Scale for Primary Education and Turkish Language Teacher Perceptions towards Spelling Rules" and to test the model. Survey research is used for the identification of a specific group's features (Büyüköztürk et al., 2016) and the depiction of the group's thoughts and perceptions about a subject area (Lodico, Spaulding, & Voegtle, 2006).

2.2. Research Group (Participants)

The research universe of the study includes Primary Education and Turkish Language teachers who teach at primary and secondary level schools under the jurisdiction of MoNE in Turkey during 2016-2017 academic years. Two independent participant groups took place in the research. For the both stages of the research, random sampling out of the sampling models was used to conduct sampling process. In the process of forming participant groups, the groups were confined simultaneously and some prerequisites were set as follows: they taught written expression course, were voluntary, and worked in different cities and schools for the sake of maximum diversity. Background information forms were exploited in the acquisition of personal data.

The first participant group included 226 teachers. Factorial structure of the scale was established and a reliability study was conducted on the data gathered from this group. 21, 32% (n=48) of these teachers was Primary Education while 78,8% (n=178) was Turkish Language teachers. In addition, 43, 4% (n=98) of the participants was male whereas 56,6% (n=128) was female teachers.

On the data gathered from the second participant group, whether the factorial structure plotted from the scale was confirmed was investigated. This participant group included 236 teachers, 20.3% (n=48) of which was Primary Education and 79.7% (n=188) was Turkish Language teachers. Moreover, 41.1% (n=97) of the group was male while 58.9% (n=139) was female teachers.

2.3. Ensuring the Appropriateness of the Data and Development of the Data Collection Tool

In scale development studies, it is recommended to take some common steps listed as follows: identifying the general need, establishing a theoretical structure, consulting with scholars, initializing the scale, piloting, administering EFA and CFA (Seçer, 2015). Based on this, in the development of the scale, initially, a literature review was conducted, and then a pool of items was constructed to be followed by the consultation with scholars. Following these steps, results of exploratory factor analysis and reliability computations were obtained to detect the reliability and validity of the research, and the model was tested via confirmatory factor analysis.

Prior to statistical analyses, assumptions related to normality, missing values, and outliers were taken into consideration for the model construction. Regarding the normal distribution of the data, measures of central tendency were detected to be close to each other. In addition, values obtained from the division of Skewness and Kurtosis amounts with the standard error were computed between -1.96 and +1.96 range. Lastly, Kolmogorov-Smirnov and Shapiro-Wilk p values did not generate any statistical significance (Can, 2013). Furthermore, Barlett Sphericty test suggested that the sample provided normal distribution conditions (Şencan, 2005; Brace, Kemp, & Snelgar, 2006; Tavşancıl, 2018). Considering the issue with regards to missing values, there was no missing values in the data due to online data collection procedures. To identify the outliers in the data set, Mahalanobis distances were calculated. It is required to use the value of $p<.001$ in Mahalanobis distances (Tabachnich & Fidell, 2007). Therefore, in this respect, six outliers were extracted from the sample of 226 participants. Normality values and outliers were screened to ensure the appropriateness of the data for the administration of Confirmatory Factor Analysis. For this purpose, data gathered from 236 participants were examined.

2.4. Writing the Items and Preparation of the Trial Form

To develop the scale, the related literature was primarily reviewed, and in the determination of the items, 37 perceptions that served the purpose of the research were written down by taking the mistakes made by students in their written expressions into consideration. A special

sensitivity was shown towards explicitness and legibility of the items by selecting expression with a single judgement.

Following this stage, specialist feedback was requested from the scholars working in Turkish Language Teaching and Turkish Language and Literature departments. A 36-item trial form was produced after making necessary revisions related to language and expression of the items based on the evaluation of the scholars.

The trial form including 36 items used 5-item Likert Scale model which expected participants' perceptions about the mistakes their students committed by responding to options as follows: "Never" (1), "Rarely" (2), "Sometimes" (3), "Often" (4), and "Always" (5). Instructions were adhered to the scale's information section which described the purpose of the scale and the correct way for responding to items.

2.5. Content Validity

To ensure content validity, the trial form of 36 items was presented to five scholars in the field for their perceptions. A three-level rating scale was used to receive scholars' feedback, requesting the scholars to choose among "appropriate", "partially appropriate", and "not appropriate" options for each item. The content validity of the items based on scholar perceptions was determined by referencing the content validity rate developed by Veneziano and Hopper (1997). According to the rate taken as reference, the main requirement was to take the minus one of the ratios of the number of scholars responding positively to the total number of the scholars. Thus, two items whose content validity rates were under .80 were omitted to establish the ultimate trial form.

2.6. Piloting Stage and the Collection of the Data

So as to test the reliability and validity of the draft perception scale with 34 items prepared, 232 Primary Education and Turkish Language teachers (the first participant group) that taught at different schools tied to MoNE during 2016-2017 academic year were administered a pilot application.

The draft scale planned for the piloting stage was delivered to 232 teachers in total who taught at MoNE schools via e-mail during 2016-2017 academic year, and their responses were saved online for individual analysis. Furthermore, personal information of teachers was collected through background data collection forms.

3. RESULT / FINDINGS

This section presents the evidence with regards to reliability and validity of the scale that was piloted.

3.1. Construct Validity

Construct validity is about the extent of measurability of the scores in relation with the concept (construct) intended to be measured (Cohen & Swerdlik, 2009; Çokluk, Şekercioğlu, & Büyüköztürk, 2016). In relation, Factor Analysis (FA) seeks answer to the question "Do the scores obtained by this test measure the construct that the test assumes to be measuring?". If the point in investigating the construct validity is to reveal the factorial status of the scale, "exploratory factor analysis" techniques are used. On the other hand, if the objective is to confirm the predetermined factorial structure of the scale, "confirmatory factor analysis" techniques are preferred (Büyüköztürk et al., 2016). In this direction, both Exploratory and Confirmatory Factor Analysis techniques were used to provide evidence for the construct validity of the scale in the research.

3.1. Exploratory Factor Analysis

In the identification of the factors in the study, Principal Components Analysis (PCA) with which a perspective from the components of the theoretical structure to the fundamental dimensions was provided was used. PCA is required to be used especially when the researcher's aim is to develop a scale (Şencan, 2005). A rotation is possible during the factor extraction (Büyüköztürk, 2016). Factorial structures are evaluated more easily through rotation procedures (Akbulut, 2010). For this reason, out of vertical rotation methods, maximized variation (varimax) was used.

Additionally, correlation values are observed to evaluate the appropriateness for EFA, which requires inter-item correlation values between .10 and .90 (Özdamar, 2016) and majority of these values above .30 (Field, 2009). Following the analyses, it was revealed that majority of the correlation values were .30 and all the correlations between the items were statistically significant (p<.05).

To determine the suitability of the data gathered, Kaiser-Meyer-Olkin (KMO) coefficient and Barlett Sphericty test were administered. The size of the sample can be deducted with the help of KMO value (Şencan, 2005; Büyüköztürk, 2016) while Barlett Sphericty test provides insight on the sample's conditions of normal distribution (Şencan, 2005; Brace, Kemp & Snelgar, 2006; Tavşancıl, 2018). As the result of the analysis, KMO coefficient was calculated as (.896). In addition to the KMO value that was considerably close to 1, and the result of Barlett's test of Sphericty calculated as (1976.141; p<.05) revealed that the samples of the study were sufficient and appropriate for the analysis (Akgül & Çevik, 2003). Moreover, a scree plot was generated to determine the number of factors that could illustrate the interaction among items. Regarding the eigenvalues' contribution to the variance, scree plot is regarded as quite useful for reducing the number of factors on the account of visuality (Çokluk et al., 2018). The related information concerning the scree plot used for determining the number of factors in the developed scale is presented in Figure 1.

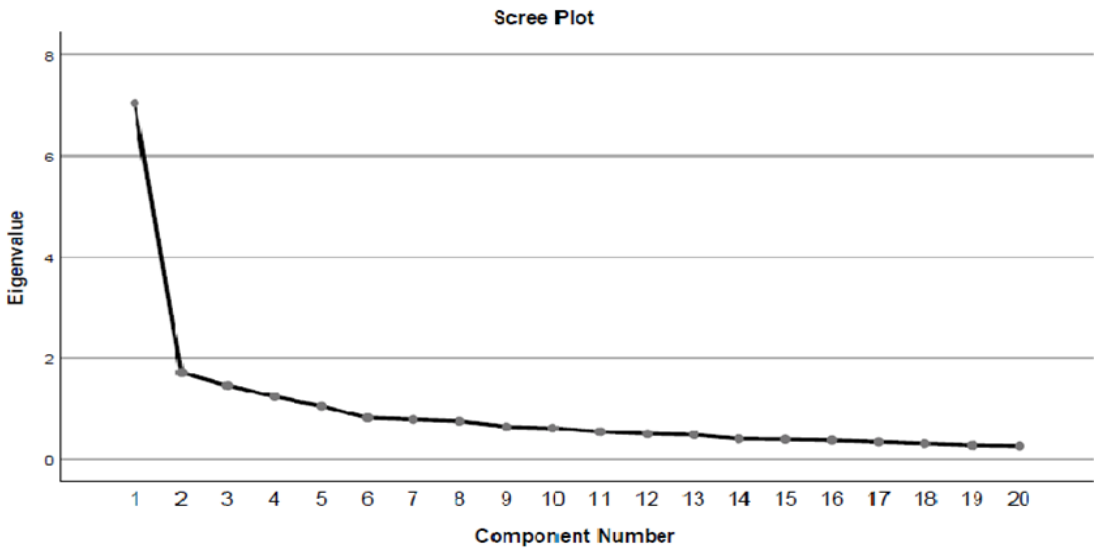

Figure 1. Screen Plot Graphic

Considering the graphic in Figure 1, it was detected that the break point (the point where the highly-accelerated rapid decrease stops) actualized after the fourth factor. It is explained that the breaking point indicates the factorizing owing to the examination of scree plot graphic. Moreover, it is recommended that, according to Kaiser Criterion, factors with eigenvalues above 1 be kept during factor extraction (Büyüköztürk, 2016). Thus, a structure with four factors was established ensuring that it was related to the theoretical construct. 15 items were

removed from the analysis due to either factor loading values being under 0.40 (Ferguson & Takane, 1989) or loadings for more than one factor with a loading value below .10 (Büyüköztürk, 2016). Table 1 presents the results of factor analysis and reliability analysis for the clusters downgraded to 19 items and loaded around 4 sub-dimensions.

Table 1. Factor Analysis Results after Varimax Rotation

Factors and Items	Explained Variance (%)	Eigenvalue	$\bar{X}$	SD	Item total r	Factor load
Factor 1 (α=.83)						
Item 1	14.509	2.757	2.98	.802	.45	.73
Item 2			2.40	.890	.46	.74
Item 3			2.55	.884	.56	.79
Item 4			2.17	1.042	.60	.77
Factor 2 (α=.83)						
Item 9	16.434	3.122	3.50	.818	.55	.70
Item 12			3.36	.790	.52	.74
Item 14			3.03	.821	.54	.69
Item 15			2.65	.877	.56	.72
Item 16			2.96	.847	.54	.72
Factor 3 (α=.82)						
Item 7	15.919	3.025	2.34	1.008	.40	.64
Item 10			2.49	.818	.44	.56
Item 20			2.26	.997	.52	.74
Item 21			1.87	.849	.50	.73
Item 22			2.07	.879	.52	.73
Factor 4 (α=.82)						
Item 28	15.784	2.999	2.31	.835	.45	.687
Item 29			2.71	.958	.41	.619
Item 30			2.44	.879	.53	.788
Item 31			2.87	.833	.48	.739
Item 32			2.33	.914	.51	.668
Overall (α=.91)	62.646	11.894	2.59	10.449		

According to Table 1, reliability values for Factor 1 through Factor 4 were calculated as .83, .83, .82, and .84 in respective orders whereas reliability coefficient of .91 was generated for the whole scale. As Bayram (2004) pinpointed, a Cronbach's Alpha value above .70 can be regarded as appropriate in terms of reliability. This outcome, therefore, indicates that the scale has a high reliability level.

Additionally, it was revealed that the finalized scale was comprised of 4 factors in total and these factors explained 62.6% of the variance. Taking item contents into consideration, the four dimensions were categorized as the most commonly made mistakes on 1- *spelling of acronyms, conjunctions, and suffixes*, 2- *spelling of capital letters*, 3- *spelling of compound letters*, and 4-

spelling of words that went through word formation processes. In social sciences, the ideal range for the explained variance is acknowledged as between 40-60% (Scherer, 1988).

Furthermore, it is suggested that correlations between sub-scale and the total scale scores should be reported (Pallant, 2011). It is observed that the correlation between sub-scales ranges from .45 to .59. Whereas, the correlation score between sub-scales and the total scale in the range of .75 and .83 indicates a statistically significant relationship. Findings regarding the correlations between sub-scales and the total score are described in Table 2.

Table 2. Correlations between Sub-scales and Total Score

Sub-dimensions*	Factor 1	Factor 2	Factor 3	Factor 4	Total Score
Factor 1	1				
Factor 2	.45**	1			
Factor 3	.53**	.53**	1		
Factor 4	.59**	.54**	56**	1	
Total Score	.75**	.79**	.83**	.81**	1

**n=226; **p<.001*

As can be referenced in Table 2, significant correlation in a positive direction is observed among sub-dimensions of the scale and between each factor and the whole scale. The obtained results can serve as evidence for the construct validity. Following these, factor-based discrimination procedures commenced. Item discrimination procedure involves the scores gained by the comparison of the scores of those from upper and lower quarters (27%) via independent sample t-test. The main point of this procedure is to display whether a response given to a specific item has changed between upper and lower groups, thusly, indicating the power of discrimination (Büyüköztürk, 2012). Thus, in this context, an independent samples t-test was used to determine whether there was a statistically significant difference between arithmetic means of upper and lower 27% groups, and the results of item-total scores were screened as shown in Table 3.

Table 3. Total Score Lower-Upper 27% Findings

Group	n	$\bar{X}$	SD	df	t	*p*<	η^2
Lower 27%	61	39.2623	4.99967	120	-13.819	.05	.61*
Upper 27%	61	49.0328	2.34497				

**Large effect size (Büyüköztürk, 2016)*

Upon implementing the independent samples t-test to determine statistical significance between lower and upper 27% groups separately designated for the discrimination of scale total scores, differences among all groups indicated statistical significance (p<.05). It is suggested that impact scale be taken into consideration to underline the power of statistical significance (Akbulut, 2010). Regarding the scale of impact in terms of total scores, a wide impact scale is observed (Büyüköztürk, 2016).

Subsequently, to detect the discrimination impact of scale items, an item-based upper-lower 27% analysis was implemented. Results of the item analysis were shown in Table 4.

Table 4. Item-Based Lower-Upper 27% Findings

Factor	Item	t	$p<$	η^2
Factor 1	Item 1	-2.665	.01	.056*
	Item 2	-2.393	.05	.046*
	Item 3	-2.916	.01	.066**
	Item 4	-2.693	.01	.057*
Factor 2	Item 9	-6.556	.001	.264***
	Item 12	-5.050	.001	.175***
	Item 14	-3.873	.001	.111**
	Item 15	-4.609	.001	.150***
	Item 16	-4.189	.001	.128**
Factor 3	Item 7	-2.702	.01	.057*
	Item 10	-5.329	.001	.191***
	Item 20	-3.260	.01	.081**
	Item 21	-2.835	.01	.063**
	Item 22	-3.793	.001	.107**
Factor 4	Item 28	-3.332	.01	.085**
	Item 29	-4.088	.001	.122**
	Item 30	-4.132	.001	.125**
	Item 31	-4.106	.001	.123**
	Item 32	-3.827	.001	.109**

** Small effect size; ** Medium effect size; ***Large effect size (Büyüköztürk, 2016)*

Following the independent samples t-test run to unearth statistical significance between lower and upper 27% groups that were individually assigned for the discrimination of all scale items, all group differences were found to be statistically significant ($p<.05$). Considering the item-based analyses, apart from the significant difference between all items' upper and lower scores, a small impact for 4 items, a moderate impact for 11 items, and a wide impact for the remaining 4 items were discovered (Büyüköztürk, 2016).

3.2. Confirmatory Factor Analysis

The aim of the confirmatory factor analysis is to test the model generated as a result of the exploratory factor analysis (Seçer, 2015). Originally, in scale development studies, a confirmatory factor analysis of the model obtained after the exploratory factor analysis is necessary as an additional technique. In the analysis of the data obtained in the research, LISREL 8.7 software was used to test the model. Standardized results concerning the model are presented in Figure 2.

In circumstances where the objective is to increase the fit indices, some modifications can be made to serve this purpose (Seçer, 2015). These modifications need to be in accord with the theoretical framework (Sümer, 2000). In this sense, some modifications were made in Item 12, Item 19, Item 20, and Item 22 in accordance with the theoretical basis. Modifications for the items that were theoretically significant and were accumulated under the same factor were employed. During the implementation of these modifications, it is recommended that the decline in chi-square values be taken into consideration (Şencan, 2000). Therefore, the all modifications were performed with these delicate considerations.

To prove model appropriateness, fit values need to be examined (Seçer, 2015; Büyüköztürk, 2016). In the research, Chi-Square Goodness index, Goodness of Fit Index (GFI), Adjusted

Goodness of Fit Index (AGFI), Comparative Fit Index (CFI), Normed Fit Index (NFI), Non-normed Fit Index (NNFI), Parsimony Goodness of Fit Index (PGFI), Root Mean Square Error of Approximation (RMSEA), Root Mean Square Residual (RMR), Incremental Fit Index (IFI), Relative Fit Indices (RFI), and Standardized Root Mean Square Residuals (SRMR) values were transferred. Fit values are presented in Table 5.

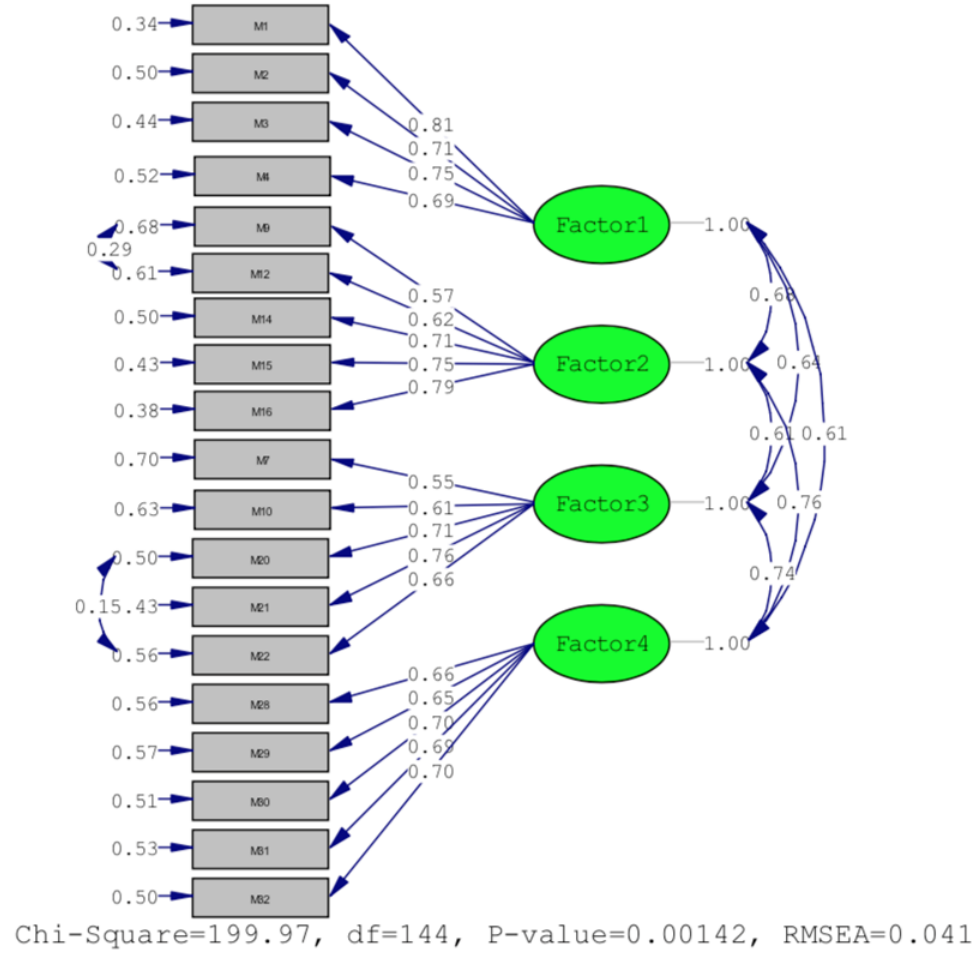

Figure 2. CFA with Standardized Results

Table 5. Fit values of the scale

Fit Index	Obtained Value	Reference Vlues	Source
Chi-Square	199.97	N/A	N/A
df	144	N/A	N/A
Chi-Square/df	1.39	≤ 2	(Tabachnick and Fidell, 2007)
GFI	0.92	≥0.90	(Sümer, 2000)
AGFI	0.89	≥0.90	(Sümer, 2000)
NFI	0.96	≥0.95	(Sümer, 2000)
NNFI	0.99	≥0.95	(Sümer, 2000)
CFI	0.99	≥0.95	(Sümer, 2000)
PGFI	0.70	≥0.50	(Mulaik et al., 1989)
RFI	0.96	≥0.90	(Marsh and Hau, 1996)
IFI	0.99	≥0.90	(Marsh and Hau, 1996)
RMSEA	0.041	≤ 0.05	(Sümer, 2000)
RMR	0.037	≤ 0.05	(Brown, 2006)
SRMR	0.046	≤ 0.05	(Brown, 2006)

Upon consideration of the table, it is explicit that fit indices for the CFA model overlaps with cutting points/reference values provided by the literature. Although AGFI value is barely above the limit, considering that fit indices should be dealt with as a whole (Jöreskog & Sörbom, 1993), it can be stated that the model showed an appropriate fit.

4. DISCUSSION, CONCLUSION and RECOMMENDATIONS

As the product of the research, a perception scale with 4 sub-dimensions (spelling of acronyms, conjunctions, and suffixes; spelling of capital letters; spelling of compound words, spelling of words that went through word formation processes) and 20 items to be used to determine students' spelling mistakes was developed. Initially, the pool of 36 items was reduced to 34 items based on the feedback (content validity rate) from the scholars in the field to produce the pilot form. Subsequently, the research carried on with two participant groups, and the first group including 232 primary education and Turkish language teachers were administered the 34-item scale. The factor analysis revealing 4 sub-dimensions led to a re-run of the factor analysis which generated 19 items whose factor loading values were above 0.40. The final version of the scale which included 19 items and 4 sub-categories was observed to be explaining the 62,6% of the total variance, therefore, qualifying as ideal (Scherer, 1988).

On another note, exploratory and confirmatory factor analyses of the scale were run for the construct validity of the scale. In the sense of EFA, two outcomes were reached. First, KMO test result (0.896) revealed that the scale was sufficient for factor analysis in terms of sample size (Pallant, 2011). Second, the data was ensured to be appropriate based on both carrying the prerequisite of 10 observations per variable (Şencan, 2005) and Barlett Sphericty test result (1976.141; $p<.05$) in the sense of normal distribution (Tavşancıl, 2018; Brace, Kemp & Snelgar, 2006). In addition, even though inter-item correlation values above .30 are recommended for the majority, the research included some correlations below .10. Upon the examination of the related literature, it can be deduced that sufficiency for the correlation matrix was obtained due to Barlett Sphericty test being statistically significant ($p<.05$) (Ho, 2006). On the other hand, item factor loading values generated appropriate results (Ferguson & Takane, 1989), and the explained variance was acceptable for the field of social sciences (Scherer, 1988). Moreover, according to the scree plot graphic, factorization ended right after the breaking point (Büyüköztürk, 2016) and the eigen values for each factor were computed above 1 (Büyüköztürk, 2016) in accordance with Kaiser criterion. Thusly, it is possible to state that the criteria suggested by the literature concerning the EFA were met. Additionally, it is explicit that correlations between the scale's sub-factors were linear to one another and they indicated statistically significant relations. Correlation values can be deemed as appropriate.

It can be confirmed that the findings reached after the confirmation of the construct obtained as a consequence of EFA with CFA were aligned with the reference values and that majority of them were above acceptable values (Mulaik et al., 1989; Marsh & Hau, 1996; Sümer, 2000; Brown, 2006; Tabachnick & Fidell, 2007). Although the AGFI value was below the acceptable level, it can be evidenced that the value was around the threshold. It is suggested that fit indices should rather be assessed as a whole, not individually (Jöreskog & Sörbom, 1993). In this perspective, it can be deduced that fit indices in the study, when examined as a whole, meet the criteria dictated by the literature.

Cornbach Alpha reliability coefficient of the scale was calculated as .91. Four sub-dimensions of the scale, on the other hand, were computed as .83, .83, .82, and .82 respectively. As a result of this analysis, consistency of the items in the whole scale and the items in the corresponding factors was tested. The obtained coefficient can be between 0 and 1, and the closer it gets to 1, the more reliable it becomes as a value (Ural & Kılıç, 2006). According to these findings, it can be validated that the scale is reliable in terms of both the whole scale and its sub-dimension

levels (Özdamar, 2016). Regarding the lower-upper 27% analyses, it was observed that all items had statistically significant differences. In the sense of both item and total scores, mostly medium and large effect sizes were traced (Büyüköztürk, 2016). In addition, considering the items' total correlations; it was observed that all of them contributed to the scale with correlation values below .40. In this context, it is suggested that correlations above .30 can be accepted as functioning values (Büyüköztürk, 2016). Therefore, it is plausible to state that the scale meets the criteria observed by the literature in terms of reliability and validity.

It is observed in the literature that scale development studies on written expression are mainly conducted on attitudes, anxiety, belief, and self-efficacy subjects. No scale development research towards spelling mistakes is present. Perceptions of teachers concerning various teaching circumstances can make great contributions in terms of functionality and quality of the education. Besides, identifying teacher perceptions regarding wrongdoings of knowledge and practical skill towards spelling rules which make up a great deal of the writing process can elevate the quality of development of writing skill. In this sense, this scale, which could be used to solve problems related to categorization of the frequency of spelling mistakes and reduction of students' spelling mistakes during the writing skill acquisition process, can be administered to ensure that teachers categorize spelling mistakes and related findings in accordance with the four sub-dimensions to channel their activities towards these mistake types.

Based on the discussion provided on the matters, the following recommendations can be presented:

1. This scale can be administered to teachers working in different levels of educational contexts, and situation reports for spelling rules in accordance with various steps can be generated.
2. This study can be upgraded to be administered in higher education levels after an adaptation process.
3. A similar version of this study can be developed in a different field to detect the situation in punctuation marks.

Declaration of Conflicting Interests and Ethics

The authors declare no conflict of interest. This research study complies with research publishing ethics. The scientific and legal responsibility for manuscripts published in IJATE belongs to the author(s).

ORCID

Ali Türkel https://orcid.org/0000-0003-4743-8766

5. REFERENCES

Akbulut, Y. (2010). *SPSS applications in social sciences: frequently used statistical analyses and SPSS solutions with explanations*. İstanbul: İdeal Kültür Pub.

Akgül, A., & Çevik, Ç. (2003). *İstatistiksel analiz teknikleri [Statistical analysis techniques]*. İstanbul: Emek Pub.

Aksoy, Ö. A. (1985). *Yine dil yanlışları 270 sözün eleştirisi [Again language mistakes 270 criticism of the word]*. Ankara: Öğretmen Pub.

Aksoy, Ö. A. (1990). *Dil yanlışları [Language mistakes]*. Ankara: Yalçın Emel Pub.

Aktaş, Ş., & Gündüz, O. (2003). *Yazılı ve sözlü anlatım kompozisyon sanatı [Written and oral expression composition art]*. Ankara: Akçağ Pub.

Alamargot, D., & Chanquoy, L. (2001). *Through the models of writing*. London: Kluwer Academic Pub.

Bağcı, H. (2011). Elementary 8th grade student's level of ability to apply spelling rules and punctuation marks. *Turkish Studies, 6*(1), 693-706.

Bayat, N. (2013). Choice of vocabulary and syntax mistakes of novice teachers in writing. *Electronic Journal of Social Sciences, 12*(43), 116-144.
Bayram, N. (2004). SPSS data analysis in social sciences, Bursa: Ezgi Pub.
Brace, N., Kemp, R., & Snelgar, R. (2006). *SPSS for psychologists.* New York: Palgrave Macmillan Press.
Brown, T. A. (2006). *Confirmatory factor analysis for applied research.* New York: The Guilford Press.
Büyüköztürk, Ş. (2016). *Handbook of data analysis for social sciences.* Ankara: Pegem Academy Pub.
Büyüköztürk, Ş., Kılıç Çakmak, E., Akgün, Ö.E, Karadeniz Ş., Demirel, F. (2016). *Bilimsel araştırma yöntemleri [Scientific research methods].* Ankara: Pegem Academy Pub.
Cohen, R. J., & Swerdlik, M. E. (2009). *Psychological testing and assessment: An introduction to tests and measurement.* Boston: McGraw-Hill.
Cho, Y. (2003). Assessing writing: are we bound by only one method. *Assessing Writing*, *8*(3), 165-191. Doi:10.1016/S1075-2935(03)00018-7
Çokluk, Ö., Şekercioğlu, G., & Büyüköztürk, Ş. (2018). *Multiple variable statistics, SPSS and LISREL applications for social sciences.* Ankara: Pegem Academy Pub.
Flower, L., & Hayes, J. R. (1981). A cognitive process theory of writing. *College Composition and Communication, 32*(4), 365-387.
Ferguson, F., & Takane, Y. (1989). *Statistical analysis in psychology and education*. McGraw Hill Book Company.
Field, A. (2009). *Discovering statistics using SPSS*. London: Sage Publications.
Gebrhard, R.C. (1983). Writing processes, revision and rhetorical problems: a note on three recent articles. *College Composition and Communication*, *34*(3), 294-296.
Ho, R. (2006). *Handbook of univariate and multivariate data analysis and interpretation with SPSS.* London & New York: Chapman & Hall/CRC.
Jöreskog, K. G., & Sörbom, D. (1993). *LISREL 8: Structural equation modeling with the simplis command language.* Lincolnwood: Scientific Software International, Inc.
Kansas, S. (2019). *Fransızca yazım kılavuzu [French spelling guide].* Paris: Hatier Pub.
Karagül, S. (2010). *Application level of the noted writing and its rules in Turkish lesson teaching programme of the elemantary 6-8. Class students.* Unpublished MA Dissertation, Dokuz Eylul University, İzmir.
Kellogg, R.T. (1996). A model working memory in writing. In C.M. Levy & S. Ransdell (Eds.). *The science of writing: therories, methods, individual differences and applications* (p.57-72). Mahwaha, NJ: L. Erlbaum Associates.
Kıbrıs, İ. (2010). *Türkçe 1: yazılı anlatım. [Turkish 1: written expression].* Ankara: Kök Pub.
Lodico, M. G., Spaulding, D. T., & Voegtle, K. H. (2006*). Methods in educational research: from theory to practice.* San Francisco: Jossey-Bass.
Marsh, H. W., & Hau, K. T. (1996). Assessing goodness of fit: is parsimony always desirable. *The Journal of Experimental Education, 64*(4), 364-390.
MEB (2006). *İlköğretim Türkçe Dersi Öğretim Programı (6-8. Sınıflar) [Primary Education Turkish Course Curriculum (Grades 6-8)].* Ankara: MEB Pub.
MEB (2009). *İlköğretim Türkçe Dersi Öğretim Programı (1-5. Sınıflar) [Primary Education Turkish Course Curriculum (Grades 1-5)]. Ankara*: MEB Pub.
MEB (2018). *İlköğretim Türkçe Dersi Öğretim Programı (6-8. Sınıflar) [Primary Education Turkish Course Curriculum (Grades 1-5)].* Ankara: MEB Pub.
Mulaik, S. A., James, L. R., Van Alstine, J., Bennett, N., Lind, S., & Stillwell, C. D. (1989). An evaluation of goodness of tit indices for structural equation models. *Psychological Bulletin*, *105*(3), 430-445.

Özdamar, K. (2016). *Structural equation modelling for scale and test development*. Eskişehir: Nisan Pub.

Pallant, J. (2011). *SPSS survival manual: a step by step guide to data analysis using SPSS.* Crow's Nest: Allen & Unwin Pub.

Seçer, İ. (2015). *Process of psychological test development and adaptation: SPSS and LISREL applications.* Ankara: Anı Publishing.

Şencan, H. (2005). *Validity and reliability in social and behavioral measurements.* Ankara: Seçkin Publishing.

Sümer, N. (2000). Structural equation models: basic concepts and sample applications. *Turkish Psychology Articles, 3*(6), 49-74.

Ural, A., & Kılıç, İ. (2006). *Bilimsel araştırma süreci ve SPSS ile veri analizi [Scientific research process and data analysis with SPSS].* Ankara: Detay Pub.

Ünver, İ. (2008). Suggestion on standart writing in transcription. *Turkish Studies, 3*(6),1-46.

Tavşancıl, E. (2018). *Measuring the attitudes and data analysis via SPSS.* Ankara: Nobel Academy Pub.

TDK Türkçe Sözlük (2019). *Turkish Dictionary of Turkish Language Institution.* http://www.tdk.gov.tr/index.php?option=com_gts&view=gts (Access: 05.03.2019).

Türkel A., Yaman B., Aksu C. (2017). *Perceptions of classroom and Turkish teachers' towards spelling rules and students' writing mistakes.* 10.UTEOK, Okan Üniversitesi, İstanbul.

Veneziano L., & Hooper J. (1997). A method for quantifying content validity of health-related questionnaires. *American Journal of Health Behavior, 21*(1), 67-70.

6. APPENDIX

Table A1. The Scale of Perceptions of Teachers' on Spelling Mistakes (Original [Turkish] version).

Please indicate how often your students perform the actions and situations given below by marking the relevant figure as (X).

Maddeler	Her zaman	Bazen	Sıklıkla	Nadiren	Hiçbir zaman
1.Öğrencilerim, özel adların ilk harfini küçük yazarak yanlışlar yapar (atatürk, ayşe, Paris caddesi, türkçe vb.).	5	4	3	2	1
2.Öğrencilerim, cins adların ilk harfini büyük yazarak yanlışlar yapar (Kuş, Araba, Oyuncak vb.).	5	4	3	2	1
3.Öğrencilerim, tümce başlarındaki ilk sözcüğe küçük harfle başlayarak yanlışlar yapar (kedim arkamdan geldi vb.).	5	4	3	2	1
4.Öğrencilerim sözcük ortasında büyük harf kullanarak yanlışlar yapar (seFer, aDa vb.).	5	4	3	2	1
5.Öğrencilerim, yazılarında ağız özelliklerini kullanarak yanlışlar yapar (gelirem, meğersem, halbüsem, gidek vb.).	5	4	3	2	1
6.Öğrencilerim, bağlaç olan "dE" ve ek olan "-DE"nin yazımında yanlışlar yapar (Ben de para yok; bende gelmek istiyorum, Ahmet de bizimle gelecek vb.).	5	4	3	2	1
7.Öğrencilerim, ünlülerle ilgili ses olaylarının gerçekleştiği (ünlü daralması, ünlü düşmesi, ünlü türemesi)ne uğrayan sözcüklerde yanlışlar yapar (bekleyor, ağızının, azcık, sapsağlam vb.).	5	4	3	2	1
8.Öğrencilerim, bağlaç olan "ki" yazımında yanlışlar yapar (Çalışki başarasın; eminimki, Senki en yakın arkadaşımsın...).	5	4	3	2	1
9.Öğrencilerim, bitişik yazılması gereken ad ve sıfatları ayrı yazarak yanlışlar yapar (vatan sever, uyur gezer, dedi kodu, bir kaç, çok bilmiş vb.).	5	4	3	2	1
10.Öğrencilerim, bitişik yazılması gereken birleşik eylemleri ayrı yazarak yanlışlar yapar (gele bilirim, red ettim kaçı verdi, şaşa kaldı vb.).	5	4	3	2	1
11.Öğrencilerim, ayrı yazılması gereken birleşik eylemleri bitişik yazarak yanlışlar yapar (farketti, terketti vb.).	5	4	3	2	1
12.Öğrencilerim, sözcükteki veya sözcüğün ekindeki "ğ" sesini yazmayarak yanlışlar yapar (fotoraf, öretmen, baktında, oldunu, adamcaz vb.)	5	4	3	2	1
13.Öğrencilerim, bazı sözcüklere fazladan ünlü veya ünsüz harfi ekleyerek yanlışlar yapar (bağazı, hayyal, messela vb.).	5	4	3	2	1
14.Öğrencilerim, bazı sözcüklerdeki harfleri düşürerek yanlışlar yapar (galba, heralde, kavaltı vb.).	5	4	3	2	1
15.Öğrencilerim, ölçü birimlerinin yazımında yanlışlar yapar (sm, lt, vb.).	5	4	3	2	1
16.Öğrencilerim, yer-yön adlarının yazımında yanlışlar yapar (içerde, dışarda, burda, ilerde vb.).	5	4	3	2	1
17.Öğrencilerim, temel kısaltmaları küçük yazarak ve/veya harflerin aralarına nokta koyarak yazımında yanlışlar yapar (tbmm, T.D.K vb.).	5	4	3	2	1
18.Öğrencilerim, coğrafi adların yazımında ilk addan sonra gelen adlara küçük harfle başlayarak yanlışlar yapar (Asya yakası, İstanbul boğazı vb.).	5	4	3	2	1
19.Öğrencilerim, özel adlara ek getirilmesi ile ilgili yazım yanlışları yapar (Konağ'a, Burağ'a, Serab'a vb.)	5	4	3	2	1