Tosun, C., Taşkesenligil, Y. (2015). The instruments used in science education in Turkey: a descriptive content analysis. *Abant İzzet Baysal Üniversitesi Eğitim Fakültesi Dergisi, 15*(2), 364-383.

Geliş Tarihi: 27/11/2014 Kabul Tarihi: 28/11/2015

THE INSTRUMENTS USED IN SCIENCE EDUCATION IN TURKEY: A DESCRIPTIVE CONTENT ANALYSIS [*]

Cemal TOSUN[**]
Yavuz TAŞKESENLİGİL[***]

ABSTRACT

The aim of this descriptive content analysis study was to analyze recent research articles related to developing and adapting scale/achievement tests. For this purpose national journals were searched. Totally 62 articles published in 21 different journals between 2002 and 2013 were analyzed in terms of participants, subjects, validity and reliability studies, data collection tools, research methods/designs employed, sample sizes, data analysis methods, variance range, factor loading, and Cronbach-Alpha values. "Scale/Achievement Test Classification Form" developed by researchers was used to analyze the data. The results of the study revealed that science and technology discipline was the most common subject for the developed/adapted scales and achievement tests. The most common participant groups were undergraduate and secondary level students and the common sample sizes were around 101-200. In addition, the common research trend in scales was student attitudes. Exploratory factor analysis was the most preferred analysis (74 %) for the developed/adapted scales, and confirmatory factor analysis was run for a small number of studies (4%).

Key words: Achievement test, descriptive content analysis, scale adaptation, scale development, science education.

TÜRKİYE'DE FEN EĞİTİMİNDE KULLANILAN ÖLÇME ARAÇLARI: BETİMSEL İÇERİK ANALİZİ

ÖZ

Bu çalışmayla fen eğitimi alanında geliştirilen ve Türk kültürüne uyarlanan ölçeklerde ve başarı testlerinde yazarların yoğunlaştığı alanları, araştırmalarda tercih edilen konuları, araştırma metodunu, örneklemi ve verilerin analizinde tercih edilen veri analiz yöntemini, sonuç da ise geliştirilen/uyarlama yapılan aracı belirlemek amaçlanmıştır. Bu amaçla 2002-2013 yılları arasında 21 farklı ulusal dergide yayımlanmış fen eğitimi alanında geliştirilen/uyarlanan ölçeklerden/başarı testlerinden oluşan toplam 62 makalenin doküman incelemesi yapılmıştır. Geliştirilen/uyarlanan ölçeklerde ve başarı testlerinde daha çok fen ve teknoloji alanında yapılan çalışmaların ağırlık kazandığı, örneklem grubu olarak lisans ve ortaöğretim öğrencileriyle ve örneklem büyüklüğü olarak ise 101-200 aralığındaki örneklem büyüklükleriyle çalışmaların yapıldığı anlaşılmaktadır. Ayrıca daha çok öğrencilerin tutumlarını belirlemeye yönelik ölçeklerin geliştirildiği görülmektedir. Geliştirilen/adaptasyon çalışması yapılan ölçeklerin büyük bir çoğunluğunda açımlayıcı faktör analizinin (%74) tercih edildiği, çok az çalışmada ise doğrulayıcı faktör analizinin (%4) uygulandığı görülmektedir.

Anahtar Kelimeler: Başarı testi, betimsel içerik analizi, fen eğitimi, ölçek uyarlama, ölçek geliştirme.

---

[*] This paper was presented at XI. National Science and Mathematics Education Congress (XI. Ulusal Fen Bilimleri ve Matematik Eğitimi Kongresi - XI. UFBMEK), Adana, Turkey, 11-14 September 2014

[**] Assoc. Prof., Bartın University, Education Faculty, ctosun@bartin.edu.tr

[***] Prof. Dr., Atatürk University, Education Faculty, ytaskes@atauni.edu.tr

1. INTRODUCTION

The validity of the application of the theoretical hypotheses put forward in studies in the education should be supported with experimental researches. So as to attain scientificity and objectivity in the results, appropriate statistical tests and scales are needed. According to Turgut and Baykul (2012), the concept of scale is used for the mathematical properties of the number and symbols of showing the measurement results.

As well as the theoretical infrastructure of any study conducted in the education being appropriate, the statistical tests and scales to be used in the application of the study should be appropriate in order that expected results can be attained for the research. Scales are very commonly-used data collection tools in education researches. When literature studies are examined, it can be clearly seen that numerous scales have been developed recently in educational studies both at nationally and internationally so as to identify the attitudes, perceptions, motivations, anxiety and achievements of the students (e.g. Arıcak & Ilgaz, 2007; Bowen, 1999; Eddy, 2000; Koçak & Önen, 2012; Wells, 2003).

A valid measurement is the cornerstone of a scientific research (DeVellis, 2003) and has a key role in the evaluation of latent variables (Crook, Shook, Madden, & Morris, 2010; Reynolds, 2010). Besides, reliable and valid measurements enable the development of research. According to Reynolds (2010), the ability and skill of measuring the variables correctly has an important role in the advancement of science. Research articles that follow the methodological process carefully are referred to more often and an important factor in the advancement of science.

Developing a measurement tool that can enable the researchers to obtain valid and reliable results in education is time-consuming, tiring and requires thinking in detail. There has to be a consistent link between the scale to be developed and the feature to be measured. Therefore, it can be possible through items constituting the scale having high-level psychometric features. This process is done through experimental or theoretical processes. In the experimental process, a draft scale form is created in accordance with literature review and expert opinions and then a pilot implication is carried out on a sample group that bear similarities with the target group. In this manner, the psychometric features of scale items are defined and the final form is obtained from the ideal items. The most important features of this process are; that it has the characteristics of a quantitative study, a factor analysis is generally used in the process which requires having a big sample size. In the theoretical process, on the other hand, a qualitative study is done via getting expert opinions in case of not having a big sample size. The consistency between expert opinions is tested with a qualitative study. This current study focus on the experimental process and procedures followed on the experimental process are summarized as follows (Churchill, 1979; Slavec & Drnovsek, 2012): i) Content domain specification, ii) item pool generation, iii) content validity evaluation, iv) scale development and evaluation, v) translation and back-translation, vi) pilot study, vii) sampling and data collection, viii) dimensionality assessment, ix) reliability assessment and x) construct validity.

The scale development studies should start with the summarization of the sphere of influence of the scale developed. According to DeVellis (2003), the researcher must, initially, explain to what end the scaling is done. According to, Netemeyer, Bearden & Sharma (2003), defining the limits of the measurement area is important. Yet, while setting these limits, the area to be measured should not be narrowed and its important aspects

should not be ignored. For this, a detailed literature review is essential. In the second step, the researchers form a pool of items regarding the scale to be developed. According to Nunnally and Bernstein (1994), they endeavor to find out whether every item is fully consistent contents with the expert opinions (cited in Slavec & Drnovsek, 2012). These initial three steps put forward the theoretical importance of the scale to be developed.

The second phase consists of four steps. In this phase, whether the data collected in relation with the scale to be developed is appropriate and its representation force is identified. Since the fourth and the fifth steps of this phase depend on the type of the research, they are optional. The fifth step that involves translation and back translation is one that has to be carried out in places where different languages are spoken. According to Dillman, Smyth and Christian, (2009), with the pilot study that constitutes the sixth step, potential problems regarding the scale are determined by the researchers beforehand (cited in Slavec & Drnovsek, 2012). It is important to know that with the scale developed within the seventh step, data is collected from the sample group.

The last three steps are those in which the dimensionality, reliability and validity of the scale is evaluated. Netemeyer, Bearden & Sharma, (2003) states that the dimensionality of the scale is in relation with the homogeneity of the scale items. The reliability of the scale, on the other hand, is the replicability of the measurement tool (Carmines & Zeller, 1979). Structural validity is the degree at which the scale can measure what is to be measured (Haynes, Richard & Kubany, 1995; Hinkin, 1998).

In this study, we are aiming at researching the important steps in scale developing, pointing at certain problems confronted in scale development studies and giving suggestions about what should be paid attention to while conducting scale development and adaptation studies. It is expected to guidance to future researches in the area with this study. It aims to identify priority areas providing a rich resource for researchers, decision-makers and practitioner. The researchers who concerned studies developing and adapting scale/achievement tests in science education provided the opportunity to use their time effectively is desirable. Teachers who want to follow developments in science education research aim to inform.

1.1.The Purpose of the Study

The purpose of this study is to identify what subject matters science educators focus on, what path they follow while developing and adapting scale/achievement tests, what sample groups they perform their applications with and which analysis methods they prefer while carrying out data analyses in scales and achievement tests developed and adapted to Turkish culture by Turkish science educators. Thus, the trends of the scales and achievement tests developed/adapted in the science education through validity and reliability studies were examined. Illustrating the current situation, these studies were compiled under one roof and were evaluated.

1.2.Research Questions

- How is the distribution of the scales/achievement tests developed/adapted in science education between 2002-2013?
- How is the distribution of the scales developed/adapted in science education between 2002-2013 according to main discipline?

- How is the distribution of the scales developed/adapted in science education between 2002-2013 according to subject matters?
- What methods were followed in the development/adaptation of scales/achievement tests whose descriptive content analyses were done?
- What are the sample and sample sizes of scales/achievement tests developed/adapted between 2002-2013?
- What are the data analyses methods used for the scales/achievement tests developed/adapted between 2002-2013?
- What are the variance range, factor loading and Cronbach-Alpha value of the scales and achievement tests whose descriptive content analyses were performed?

2. METHODOLOGY

62 articles examined by using the descriptive content analyses method. Descriptive content analyses are systematic studies that examined of the studies done on a certain subject and trends and the results of study will be evaluated in a descriptive manner (Çalık vd., 2008; Sözbilir, Kutu & Yaşar, 2012). Also the general trends in the field are determined which qualitative and quantitative studies that they made independently are regulated by examining (Selçuk, Palancı, Kandemir & Dündar, 2014).

2.1.Data Resource

For this study, a total of 62 articles published in 21 different journals consisting of scales/achievement tests developed/adapted in science education were put through descriptive content analysis. All articles are published in the 32 journals were examined one by one to determine the studies to be included in the descriptive content analysis. It was defined as criteria title or purpose of the article which scales/achievement tests developed/adapted in science education. So it was determined that published of the studies to be included in the research between the years 2002-2013.

The title of each journal and the number of articles are given in Appendix 1. *"Scale/Achievement Test Classification Form"* developed by the researchers was used. This form was created by benefiting *"Paper Classification Form-CPF"* developed by Sözbilir, Kutu & Yaşar, (2012). This form consists of nine parts. The part A includes the descriptive information of the paper. The part B comprises classification of the paper according to the main discipline that paper belonged such as physics, biology, chemistry, science and technology etc. The part C deals with the subject matters studied. The part D comprises simply information about research methods. In the part E, what kind of a method was followed in the development and adaptation of the scale/achievement tests? In the part F, each paper was categorized according to their data collection instrument. The part G includes the sample and sample size. The part H comprises the data analysis methods and techniques benefited in the studies. Lastly, the part I include variance range, factor loading, Cronbach-Alpha values.

2.2.Data Analysis

So as the maintain the reliability of the investigation, numerous pilot practices were performed during the development of *"Scale/Achievement Test Classification Form"* developed by the researchers. The form was finalized according to the results obtained from the analyses of the first 15 articles that were put through descriptive content analysis. In order to achieve a reliable of the papers, the authors initially worked together. The disagreements were discussed and resolved, and then the rest of the papers were classified by the first author. The data were collected via filling the *"Online Scale/Achievement Test Classification Form"* prepared using "Google Drive" for every paper. This program enabled the data to be presented to Microsoft Excel regularly. Consequently, the results were transformed into graphic, frequency and percentage tables and presented in a descriptive form.

3. FINDINGS

When the articles whose descriptive content analyses were performed are examined in terms of authors numbers, the average author number of the 39 articles with scale development $\overline{X}$ = 2.38, the average author numbers of 11 articles with adaptation studies $\overline{X}$ = 2.09 and the average author number of 12 articles with achievement test development studies $\overline{X}$ = 2.0 and the average author number of all articles (a total of 62) $\overline{X}$ = 2.25.

The studies subject to descriptive content analysis were classified. According to this, 63% of the studies whose descriptive content analysis was made were scale/inventory development studies while 18% were adaptation of the scale to Turkish culture and 19% were achievement test development.

So as to show the distribution of the scales and achievement tests developed or adapted in science education in the journals they were published by the years, Table 1 was created. When Table 1 is examined, it can be seen that science educators have performed the scales and achievements tests developed or adapted studies in science education since 2002.

Table 1.
The distribution of articles by years (2002-2013; N=62).

	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	Total
Scale/Inventory Development	1	--	1	3	2	3	4	9	5	3	6	2	39
Scale Turkish Culture Adaptation	1	1	--	1	2	1	--	2	--	1	2	--	11
Achievement Test Development	--	1	--	--	--	1	--	1	1	7	1	--	12
Total	2	2	1	4	4	5	4	12	6	11	9	2	62

When examined, Table 1 reveal that scale development studies were more numerous in 2009 and more achievement test development studies were carried out in 2011 when compared with other years.

Table 2 was created as a result of the classification of main discipline that articles examined in accordance with the *"Scale/Achievement Test Classification Form".*

Table 2.
Classification of articles according to their main discipline

	Biology	Chemistry	Science and Technology	Physics	Environmental Science	Mixed	Total
Scale/Inventory Development	6	6	13	6	7	1	39
Scale Turkish Culture Adaptation	2	5	3	1	---	---	11
Achievement Test Development	2	4	---	6	---	---	12
Total	10	15	16	13	7	1	62

Table 2 shows that a significant proportion of papers (%26) are published in science and technology education. Paper focused on chemistry education make up %24 of the total, followed by physics education (%21), biology education (%16) and environmental education (%11). Papers belonging to more than one discipline are classified as mixed.

In the scales and achievement tests developed/adapted within science education, which subject matters are preferred/put more emphasis on is among the research questions in this study.

Table 3.
Frequently ınvestigated subject matters

	Self efficacy	Anxiety	Perception	Attitude	Motivation	Scientific Process Skills	Constructivist Approach	Environmental Attitude	Learning environment	Achievement Test Development	Other	Total
Scale/Inventory Development	3	1	1	17	1	2	2	3	1	--	8	39
Scale Turkish Culture Adaptation	3	1	--	3	2	--	--	--	1	--	1	11
Achievement Test Development	--	--	--	--	--	--	--	--	--	12	--	12
Total	6	2	1	20	3	2	2	3	2	12	9	62

When Table 3 is examined, of the subject the descriptive content analyses of which were performed, 32% were attitude, 19% were achievement tests, 10% were self-efficacy, 5% were motivation and environmental attitude, 3% were anxiety, learning environment, scientific process skills and constructive approach and 2% were perception studies. Furthermore, 15% of the total papers classified as other were in confidence in learning by oneself, measurement and evaluation adequacies, association with daily life, teacher qualities, learning difficulties in project based learning, inquiry learning skills, contentment, environmental awareness and sensitivity and classroom environment studies.

One of the research questions within this study is what kind of a method is followed while developing/adapting the scales/achievement tests. To answer this question, Figures 1 and 2 were formed. In Figure 1, at what extend the noteworthy processes were complied with during the development process of 51 studies developed and put through descriptive content analysis in science education.

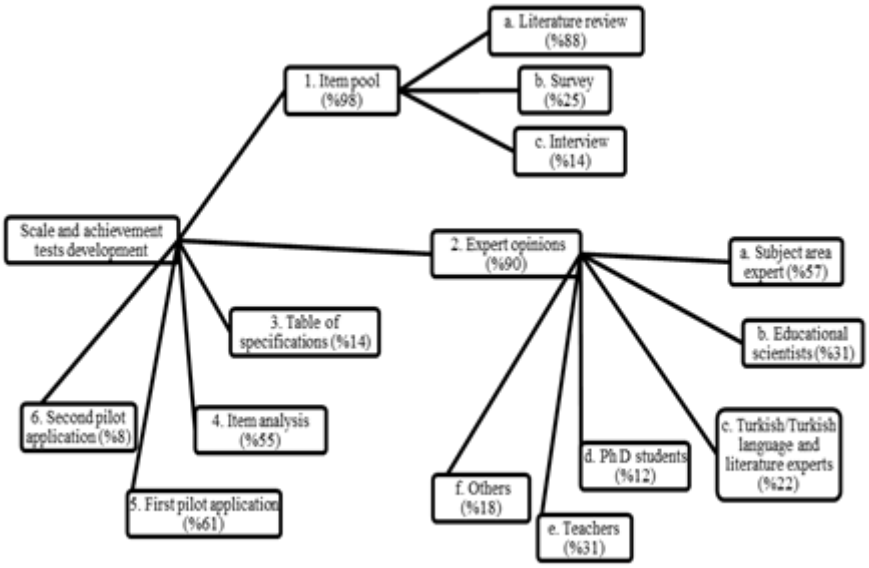

Figure 1. The process for the scales and achievement tests development studies

When Figure 1 is examined, it can be seen that while forming the scale items of almost all (98%) of the 51 scales and achievement tests, an item pool was created. While forming the scale items of these 51 studies, expert opinions were consulted to for 90%, a table of specifications was made for 14%, an item analysis was made for 55%, first pilot application was done for 61% and second pilot application was done for 8%. On the other hand, it is revealed that a literature review was performed for 88% of the studies, a survey was carried out on an appropriate sample group for 25% of the studies and for 14%, and an interview was made. Moreover, it is understood that in 10% of the studies with surveys, open-end questions were asked and in 18%, students were asked to write a composition. Of the studies with interview, 6% included informal interview, 4% included semi-structured interview and 2% included unstructured interview. In 51 studies developed and put through descriptive content analysis, in 57% subject area experts, in

31% educational scientists, in 22% Turkish/Turkish Language and Literature experts, in 12% PhD students, in 31% teachers and in 18% other experts were consulted.

In figure 2, to what extend the processes to be noticed in the adaptation steps of 11 studies that were adapted to Turkish culture were complied with.

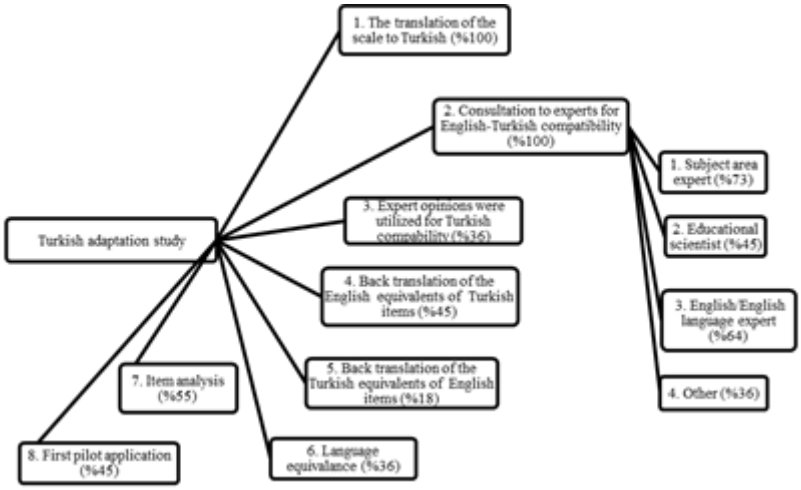

Figure 2. The process for the adapted scales

When examined, Figure 2 makes visible that for all of the 11 adapted scales, the translation of the scale to Turkish and consultation to experts for English-Turkish compatibility phases were followed. In 36% of the adapted scales, expert opinions were utilized for Turkish compatibility and in 45%, back translation of the English equivalents of Turkish items and in 18%, back translation of the Turkish equivalents of English items was performed. In 36%, of the scales adapted to the stage at which language equivalence was tested via applying to the certain sample group at intervals and in 55%, item analysis was performed while in 45%, a first pilot application was carried out. On the other hand, at the stage of consulting to experts for English-Turkish compatibility, for 73% of the adapted scales, subject area experts, for 45% educational scientists, for 64%, English/English Language experts and for 36% experts classified as "others" are consulted.

The following Table 4 shows the frequently studied samples. Of the total sample groups in studies subject to descriptive content analyses, 38% were undergraduate students, 33% were secondary students (9-12), 20% were primary students (6-8), 6% were teachers, 1.5% were primary school students (1-5) and post graduate students. Pre-school students, families and administrators were not preferred for sample groups. Moreover, since two different sample groups were used in two studies, the total number of articles with descriptive content analyses was given as 64 in Table 4.

Table 4.
Frequently studied samples

	Primary (1-5)	Primary (6-8)	Secondary (9-12)	Undergraduate	Postgraduate	Teacher	Total
Scale/Inventory Development	1	9	12	13	---	4	39
Scale Turkish Culture Adaptation	---	2	3	6	---	---	11
Achievement Test Development	---	2	6	5	1	---	14
Total	1	13	21	24	1	4	64

Table 5 was formed in relation with sample sizes. When Table 5 is examined, it can be seen that student groups of 101-200 are most frequently preferred as sample size for scale and achievement test development studies. Besides, it is visible that the minimum and the maximum sample groups are in achievement test development studies. It can be understood that the average sample group in adaptation to Turkish studies is averagely larger in comparison with scale and achievement test development studies.

Table 5.
Frequently studied sample sizes

Minimum	Maximum	Mean		1-100	101-200	201-300	301-400	401-500	501-1000	Over 1000	Total
82	969	363	Scale/Inventory Development	1	10	7	6	8	7	---	39
255	659	398	Scale Turkish Culture Adaptation	---	---	3	3	2	3	---	11
78	1898	312	Achievement Test Development	---	7	2	1	1	---	1	12
		360	Total	1	17	12	10	11	10	1	62

When Table 5 is examined, the proportions of sample size in scales and achievement tests developed and adapted are; 101-200 in 27%, 201-300 in 19%, 401-500 in 18%, 301-400 and 501-1000 in 16% and 1-100 - and over 1000 in 2%.

To identify whether a factor analysis study for the structural validity of developed and adapted scales has been performed and, if so, what type of factor analysis was performed, Table 6 were created. When Table 6 is examined, it can be seen that of the total of 50 scales developed and adapted (12 studies subject to achievement test analysis not included), 74 % were put through only exploratory factor analysis; 4% were subject to only confirmatory factor analysis is and 14% were both exploratory and confirmatory factor analysis. No factors analyses were performed for 8% of the studies.

Table 6.
Frequently factor analysis

	Scale/Inventory Development	Scale Turkish Culture Adaptation	Total
Exploratory Factor Analysis	29	8	37
Confirmatory Factor Analysis	1	1	2
Exploratory + Confirmatory Factor Analysis	6	1	7
No Factor Analysis	3	1	4
Total	39	11	50

On the other hand, in the scales developed and adapted, how many options of scale items were included was also specified. An examination of Table 7 reveals that in 80% of the developed and adapted scales there are five points; in 8%, there are four points; in 4%, there are three points and in 2%, there are 7 Likert type scales. Besides, in 6% of the developed and adapted scales, what kinds of scales were used is not specified.

Table 7.
Scale types

	3 points	4 points	5 points	7 points	Unspecified	Total
Scale/Inventory Development	2	3	31	---	3	39
Scale Turkish Culture Adaptation	---	1	9	1	---	11
Total	2	4	40	1	3	50

In the scales and achievement tests developed/adapted in science education, total item counts were specified (See Table 8).

Table 8.
Total items

	Total Mean
Scale/Inventory Development	29.4
Scale Turkish Culture Adaptation	29.7
Achievement Test Development	20.1

When we examine Table 8, the average item number in the scales developed $\overline{X}$ = 29.4, the ones adapted to Turkish culture $\overline{X}$ = 29.7 and in achievement test studies $\overline{X}$ = 20.1.

In Table 9, how many dimensions the scales developed in science education and adapted to Turkish culture had the percentage of the total variance each sub dimension defined,

the total variance, factor loadings for each sub dimension and scale general and Cronbach-Alpha values were included. An examination of Table 9 reveals that 4% of 50 studies developed and adapted are single or multi-dimensional is not specified. 68% of the studies subject to descriptive content analysis were multi-dimensional while 28% of them were single dimensional.

It is seen that in single dimensional scales, minimum 30% and maximum 51% of the total variance are given. As for multi-dimensional scales, minimum 35% and maximum 78% of the total variance were defined (see Table 9). When the factor loading of the scales developed and adapted to Turkish culture in science education are examined, it can be seen that it is between .28-.95. In 2% of the total of 50 studies excluded from achievement test development studies subject to content analysis, factor loadings are between .20-.29 minimum; in 28% it is between .30-.39, in 36%, it is between .40-.49 and in 4%, it is between .60-.69. In 14% of the studies, the factor loading was not specified.

It is visible that the Cronbach-Alpha values of the sub dimensions of the scales (except 12 achievement test studies) are between .55-.97 and in the scale's general, it is .74-.98 (See Table 9). The Cronbach-Alpha levels of 6% of the studies subject to descriptive content analysis are not given. The Cronbach-Alpha values in general of the scales are; .70-.79 for 8%. .80-.89 for 46% and .90-.99 for 40%.

Table 9.
Variance Range, Factorial Loading and Cronbach-Alpha Values (T=Unidimensional; B=Unspecified)

			Variance ranges								Factorial loadings							Cronbach-Alpha							
	Item number	Number of dimensions	1. Sub dimensions	2. Sub dimensions	3. Sub dimensions	4. Sub dimensions	5.Sub dimensions	6. Sub dimensions	7. Sub dimensions	Total	1. Sub dimensions	2. Sub dimensions	3. Sub dimensions	4. Sub dimensions	5.Sub dimensions	6. Sub dimensions	7. Sub dimensions	1. Sub dimensions	2. Sub dimensions	3. Sub dimensions	4. Sub dimensions	5.Sub dimensions	6. Sub dimensions	7. Sub dimensions	Total
SCALE/INVENTORY DEVELOPMENT	1	4	54	11	7	5				78	.71-.95	.45-.73	.72-.89	.73-.77				.90	.80	.72					.93
	2	3	23	18	15					58	.57-.74	.59-.77	.54-.79					.88							.80
		3	17	14	12					44	.38-.69	.50-.72	.53-.81					.80							
	3	5	12	10	10	8	8			50	.57-.72	.51-.72	.44-.66	.51-.67	.46-.65			.79	.75	.75	.63	.60			.88
	4	3	37	5	5					48	UNSPECIFIED							.94	.89	.83					.96
	5	3	UNSPECIFIED								.46-.66	.43-.71	.40-.66												.94
	6	2	19	15						35	.45-.58	.46-.72						.91	.78						.86
	7	4								71	.45-.82	.37-.76	.53-.77	.53-.78											.98
	8	T	UNIDİMENSIONAL								.37-.66														.89
	9	7	34	8	5	4	3	3	3	62	.49-.79	.52-.82	.44-.73	.64-.85	.65-.79	.65-.79	.59-.82								.94
	10	4	18	14	10	10				54	UNSPECIFIED							.88	.85	.77	.70				.94
	11	4	31	16	15	12				75	.68-.93	.82-.90	.72-.87	.62-.92				.97	.94	.92	.88				.92
	12	3	46	10	5					62	.69-.80	.55-.70	.76-.78					.93	.91	.89					.95
	13	5	21	7	7	5	4			47	.53-.68	.47-.68	.51-.64	.49-.73	.52-.70			.75	.68	.56	.55	.59			.80
	14	T	UNIDİMENSIONAL							30	.47-.68														.84
	15	B	UNSPECIFIED								UNSPECIFIED														.81
	16	4	19	15	14	12				61	.53-.79	.81-.91	.73-.78	.72-.78				.79	.86	.74	.68				.76
	17	3	UNSPECIFIED								.36-.58	.41-.62	.36-.61					.67	.73	.71					.84
	18	3	37	8	5					52	.56-.71	.49-.73	.47-.79					.87	.87	.78					.92
	19	5	39	9	8	7	5			70	.71-.85	.71-.79	.59-.80	.69-.88	.47-.76			.91	.84	.81	.86	.88			.93
	20	4								48	.42-.82	.50-.71	.35-.44	.39-.66				.83	.79	.74	.71				.87
	21	2	28	23						52	.55-.74	.60-.81						.90	.87						.93
	22	T	UNIDİMENSIONAL							51	.58-.80														.97
	23	2	24	10						35	.32-.71	.45-.76													.74
	24	3	38	8	5					52	.49-.77	.67-.72	.60-.77					.92	.82	.57					.87
	25	5	27	10	7	6	6			56	.43-.57	.42-.74	.47-.69	.43-.71	.42-.90										.87

(continued)

Table 9. (continued)

Variance Range, Factorial Loading and Cronbach-Alpha Values (T=Unidimensional; B=Unspecified)

SCALE/INVENTORY DEVELOPMENT	26	5	12	10	9	9	8			50	.45-.77	.40-.72	.58-.76	.45-.75	.42-.77			.85	.84	.85	.79	.79			.91
	27	5								44	.35-.68	.37-.71	.50-.67	.36-.68	.38-.73			.95	.92	.93	.86	.83			.97
	28	3	21	12	7					42	.49-.68	.48-.69	.39-.58					.89	.80	.68					.83
	29	5	15	13	12	11	7			60	.52-.77	.54-.76	.59-.75	.55-.76	.64-.78										.86
	30	T	UNPERFORMED								UNSPECIFIED							UNSPECIFIED							
	31	6	27	9	6	5	5	4		59	.47-.71	FOR ALL SIZES													.90
	32	4	19	12	11	9				52	.44-.86	.43-.66	.55-.73	.55-.72				.91	.82	.85	.69				.93
	33	3	31	15	10					57	.45-.83	.51-.75	.69-.83					.92	.80	.70					.88
	34	2	20	19						40	.33-.85	.39-.75						.91	.90						.92
	35	6	21	5	5	4	4	3		44	.34-.77	FOR ALL SIZES													.87
	36	T	UNIDİMENSIONAL							47	.39-.84														.94
	37	3	39	7	4					51	.54-.77	.45-.68	.46-.73					.93	.75	.80					.93
	38	2	UNPERFORMED								UNSPECIFIED							.80	.82						
	39	B	UNPERFORMED								UNPERFORMED														.83
	40	3	8	9	54					71	.32-.74	.52-.73	.52-.80					.82	.76	.80					.81
SCALE TURKİSH ADAPTATİON	41	4	6	9	11	39				66	.61-.85	.76-.87	.67-.85	.76-.88				.88	.87	.86	.87				
	42	3	12	15	21					58	.50-.70	.50-.76	.50-.70					.93	.90	.88					.94
	43	T	UNSPECIFIED								UNSPECIFIED														.78
	44	T	UNIDİMENSIONAL							39	.52-.85	FOR ALL SIZES													.76
	45	2	14	23						37	.43-.73	.28-.71						.86	.69						.85
	46	2	12	28						40	.34-.75	.51-.77						.71	.77						.80
	47	4	CONFİRMATORY FACTOR ANALYSİS								.75-.87	.69-.83	.66-.78	.51-.79				.84	.81	.78	.68				.88
	48	6								56	.64-.83	.50-.64	.43-.73	.63-.83	.45-.74	.41-.74		.71	.85	.74	.54	.77	.77		.87
	49	5	4	5	6	8	21			46	.55-.78	.43-.78	.45-.75	.42-.64	.40-.61			.87	.81	.76	.75	.61			.86
	50	4	5	7	11	26				50	.40-.69	.54-.77	.39-.79	.67-.72				.80	.66	.74	.61				.82

4. CONCLUSION AND DISCUSSION

This descriptive content analysis study examines the tendencies of the scales and achievement tests developed/adapted in science education through a validity and reliability study. To illustrate the current situation, compile and gather these studies under a single roof, a total of 62 articles that consist of scales/achievement tests developed/adapted in science education, published in 21 different journals between 2002-2013, and were examined.

It is understood that in our country, a tendency towards scale development/adaptation studies as of the year 2002; scale development studies increased in number in 2009 and achievement test development endeavors increased in 2011.

It can be seen that in scales and achievement tests developed/adapted, studies in science and technology were weighted and that undergraduate and secondary education students were mostly preferred as sample groups. This situation is in compliance with the finding by Göktaş et al. (2012), who investigates the tendencies in educational studies in Turkey that in educational investigations in Turkey, undergraduate students and teachers are preferred for sample groups.

As the sample size in scales and achievement tests developed/adapted, 101-200 were preferred most. In today's world, in which not the upper limits but the lower limits of observation numbers are questions of debate, a lot of studies have been performed on this topic. According to Pedhazur (1997), the ratio of item and observation numbers should be 1/15 or 1/30 (cited in Osborne & Costello, 2004). That is, Pedhazur suggests a ratio of 15 or 30 observations per one item. According to Gorsuch (1983), on the other hand, defines the minimum ratio as 1/5 adding, though, that this is a bottom limit and anything below this ratio must not be used and the higher one goes beyond this level, the better it is (cited in Osborne & Costello, 2004). Nunnally and Bernstein (1994) assert this ratio would best be 1/10. Costello and Osborne (2003) examined 1076 published articles in a study they conducted and observed that in 40.5% of the articles, the item-observation ratio in the factor analyses were below 1/5 (cited in Osborne & Costello, 2004). According to Guadagnoli and Velicer (1988), who take the absolute observation number as the standard, the adequate absolute observation number must be between 100-200. Aleamoni (1976) asserts that for factor analysis, the sample size is 400 (cited in Guadagnoli & Velicer, 1988). According to Comfrey and Lee (1992), the absolute observation size: 50-(too weak), 100-(weak), 200-(neutral), 300-(good), 500-(very good) and 1000 and above is ideal (cited in Osborne & Costello, 2004). In the study carried out with the Monte Carlo Simulation by Osborne and Costello (2004) who accept the item number / observation number ratio, absolute observation number and expected factor / observation number ratio. According to Osborne and Costello (2004), for the item number / observation number, 1/10, for absolute observation number, 1000 and for item per factor, 11 are the ideal values.

When the fact that the average item number in scales developed $\overline{X}$ = 29.4, in those adapted to Turkish culture $\overline{X}$ = 29.7 and in achievement test studies $\overline{X}$ = 20.1 is taken into consideration, the fact that in studies subject to descriptive content analyses, the 27% study in which as sample size of 101-200 was preferred and the 19% study in which a sample size of 201-300 was preferred creates doubts should be remembered.

On the other hand, it is observed that in studies subject to descriptive content analysis, scales aiming at determining the attitudes of the students are mainly developed (32%). In most of the scales developed/adapted (68%), the explanatory factor analysis was preferred and in a few studies (4%), the confirmatory factor analysis was used. It is visible that in single-dimensioned scales minimum 30% and maximum 51% of the total variance was identified while in multi-dimensional scales, minimum 35% and maximum 78% of the total variance was identified. When the factorial loadings of the scales developed and adapted to Turkish culture in science education are examined, it can be seen that these values are between .28-.95. On the other hand, the Cronbach-Alpha values of the sub dimensions of the scales subject to descriptive content analysis are observed to be .55-.97 while it is .74-.98 in scale general. This situation parallels with Price and Mueller's (1986) remarks that the minimum limit of the Cronbach-Alpha value should be at least 70.

It is desirable to recover from again and scatter educational researches. Teachers and researchers who have the problem and access to researches should be informed with such studies to enable them to follow educational studies. The reliability and validity a measurement tool is the cornerstone of scientific researches. Such studies are needed that it is summarized what steps are to be followed in the validity and reliability studies.

This study is limited to 62 articles published in 21 journals. It is recommended to be given in the results to be obtained with the diversity of resources (dissertation, proceedings etc.).

REFERENCES

Arıcak, O.T. & Ilgaz, G. (2007). The investigation of construct validity of biology course attitude scale with exploratory and confirmatory factor analysis. *Eurasian Journal of Educational Research, 28*, 1-8.

Bowen, C. W. (1999). Development and score validation of a chemistry laboratory anxiety instrument (CLAI) for college chemistry students. *Educational and Psychological Measurement, 59* (1), 171-185.

Çalık, M., Ünal, S., Çoştu, B. & Karataş, F.Ö. (2008). Trends in Turkish science education. *Essays in Education,* Special Edition, 23-45.

Carmines, E. G. & Zeller, R. A. (1979). *Reliability and Validity Assessment* (pp.71). Sage Publications, Beverly Hills, London.

Crook, T. R., Shook, C. L., Madden, T. M. & Morris, M. L. (2010). A review of current construct measurement in entrepreneurship. *International Entrepreneurship and Management Journal, 6* (4), 387-398.

Churchill, G. A. (1979). A paradigm for developing better measures of marketing constructs. *Journal of Marketing Research, 16* (1), 64-73.

DeVellis, R. F. (2003). *Scale Development: Theory and Applications* (2nd ed. Vol. 26). Sage Publications, Thousand Oaks, London.

Eddy, R. M. (2000). Chemophobia in the college classroom: Extent, sources, and student characteristics. *Journal of Chemical Education. 77* (4), 514-517.

Göktaş, Y., Akçay, A., Hasançebi, F., Bayrak, N., Varisoğlu, B., Baran, M. & Sözbilir, M. (2012). Trends in educational research in Turkey: A content analysis. *Educational Sciences: Theory & Practice, 12* (1), 443-460.

Guadagnoli, E. & Velicer, W.F. (1988). Relation of sample size to the stability of component patterns. *Psychological Bulletin, 103* (2), 265-275.

Haynes, S. N., Richard, D. C. S. & Kubany, E. S. (1995). Content validity in psychological assessment: A functional approach to concepts and methods. *Psychological Assessment, 7* (3), 238-247.

Hinkin, T. R. (1998). A brief tutorial on the development of measures for use in survey questionnaires. *Organizational Research Methods, 1* (1), 104-121.

Koçak, C. & Önen, A.S. (2012). A study for the development of attitude scale towards daily life chemistry. *Hacettepe University Journal of Education, 43,* 318-329.

Netemeyer, R. G., Bearden, W. O. & Sharma, S. (2003). *Scaling Procedures: Issues and Applications.* Thousand Oaks, CA: Sage Publications.

Osborne, J. W. & Costello, A. B. (2004). Sample size and subject to item ratio in principal components analysis. *Practical Assessment, Research & Evaluation, 9* (11), Retrieved March 9, 2014 fromhttp://PAREonline.net/getvn.asp?v=9&n=11.

Price, J.P., & Mueller, C.W. (1986). *Handbook of organizational measurement.* Marshfield, MA: Pittman.

Reynolds, C. R. (2010). Measurement and assessment: An editorial view. *Psychological Assessment, 22* (1), 1-4.

Selçuk, Z., Palancı, M., Kandemir, M. & Dündar, H. (2014). Eğitim ve bilim dergisinde yayınlanan araştırmaların eğilimleri: İçerik analizi. *Eğitim ve Bilim, 39* (173), 430-453.

Slavec, A. & Drnovsek, M. (2012). A perspective on scale development in entrepreneurship research. *Economic and Business Review, 14* (1), 39-62.

Sözbilir, M., Kutu, H. & Yaşar, M.D. (2012). *14. Science education research in Turkey: A content analysis of selected features of published papers.* In D. Jorde & J. Dillon (Eds). *Science Education Research and Practice in Europe: Retrospective and Prospective* (pp.341-374). Rotterdam: Sense Publishers.

Turgut, M. F. ve Baykul, Y. (2012). *Eğitimde Ölçme ve Değerlendirme. 4. Baskı*, Pegem Akademi, 424 s, Ankara.

Wells, R.R. (2003). *The development of an instrument to assess chemistry perceptions.* Doctoral Dissertation. Texas Tech University.

Yıldırım, A. & Şimşek, H. (2006). *Sosyal Bilimlerde Nitel Araştırma Yöntemleri. (6. Baskı),* Seçkin Yayınevi, 366 s, Ankara.

GENİŞ ÖZET

1.Çalışmanın Amacı

Bu çalışmayla fen eğitimi alanında geliştirilen ve Türk kültürüne uyarlanan ölçeklerde ve başarı testlerinde yazarların hangi alanlara yoğunlaştıklarını, araştırmalarında hangi konuları tercih ettiklerini, ölçek/başarı testi geliştirme ve uyarlama çalışmalarını yaparken nasıl bir yol izlediklerini, uygulamalarını hangi örneklem grubuyla gerçekleştirdiklerini ve verilerin analizinde hangi veri analiz yöntemlerini tercih ettiklerini, sonuç da ise ne tür bir araç geliştirdiklerini/uyarlama çalışması yaptıklarını belirlemek amaçlanmıştır. Böylece ülkemizde fen eğitimi alanında geçerlik ve güvenirlik çalışması yapılarak geliştirilen/uyarlanan ölçeklerin ve başarı testlerinin eğilimleri incelenmiştir. Mevcut durumun resmi çizilerek, bu çalışmalar derlenmiş, bir çatı altında toplanıp değerlendirilmiş ve aşağıdaki araştırma sorularına cevap aranmıştır.

Araştırma Soruları

- 2002-2013 yılları arasında fen eğitimi alanında geliştirilen/uyarlanan ölçeklerin/başarı testlerinin yıllara göre dağılımı nasıldır?
- 2002-2013 yılları arasında geliştirilen/uyarlanan ölçeklerin/başarı testlerinin konularına göre dağılımı nasıldır?
- 2002-2013 yılları arasında fen eğitimi alanında geliştirilen/uyarlanan ölçeklerin/başarı testlerinin alanlarına göre dağılımı nasıldır?
- Betimsel içerik analizi yapılan ölçeklerin/başarı testlerinin geliştirilmesinde ve uyarlanmasında nasıl bir yol takip edilmiştir?
- 2002-2013 yılları arasında geliştirilen/uyarlanan ölçeklerin/başarı testlerinin örneklem ve örneklem büyüklüğü nedir?
- 2002-2013 yılları arasında geliştirilen/uyarlanan ölçeklerde/başarı testlerinde kullanılan veri analiz yöntemleri nelerdir?
- Betimsel içerik analizi yapılan ölçeklerin ve başarı testlerinin analiz sonrası varyans aralıkları, faktör yük değerleri ve Cronbach-Alpha değerleri nasıldır?

2.Yöntem

Betimsel içerik analiz yöntemiyle incelemesi yapılan 62 makalenin düzenli ve detaylı incelemesi yapılmıştır. Betimsel içerik analizi belirli bir konu üzerinde yapılan çalışmaların ele alınıp eğilimlerinin ve çalışma sonuçlarının betimsel bir şekilde değerlendirildiği sistematik çalışmalardır (Çalık vd., 2008; Sözbilir, Kutu & Yaşar, 2012). Ayrıca bu tür çalışmalar ile birbirinden bağımsız olarak yapılan nitel ve nicel çalışmalar incelenip düzenlenmekte ve alandaki genel eğilimler belirlenmeye çalışılmaktadır (Selçuk, Palancı, Kandemir, & Dündar, 2014).

2.1.veri kaynağı

Bu çalışma için 21 farklı dergide yayınlanmış fen eğitimi alanında geliştirilen/uyarlanan ölçeklerden/başarı testlerinden oluşan toplam 62 makalenin doküman incelemesi yapılmıştır. Her bir derginin adı ve dergideki geliştirilen/uyarlanan ölçeklerle/başarı testleriyle ilgili incelemesi yapılan makale sayıları Ek-1'de verilmiştir.

2.2.veri toplama aracı

Veri toplama aracı olarak araştırmacılar tarafından geliştirilen "Ölçek/Başarı Testi Sınıflandırma Formu" kullanılmıştır. Bu form ölçeğin/başarı testinin kimliği hakkında tanımlayıcı bilgi, ölçeğin alanı, konusu, ölçeğin/başarı testinin geliştirilmesinde ve uyarlanmasında kullanılan yöntem ve nasıl bir yol takip edildiği, veri toplama araçları, örneklem ve örneklem büyüklüğü, veri analiz yöntemi ve sonuç da ise nasıl bir ölçek/başarı testi elde edildiğini gösterir bilgi olmak üzere toplam dokuz alt boyuttan oluşmaktadır.

2.3.veri analizi

Araştırmanın güvenirliğini sağlamak amacıyla araştırmacılar tarafından geliştirilen "Ölçek/Başarı Testi Sınıflandırma Formunun" geliştirilme sürecinde birçok defa pilot uygulaması yapılmıştır. Formun son haline içerik analizi yapılan ilk 15 makalenin analizinden elde edilen sonuçlara göre karar verilmiştir. Veriler "Google Drive" yardımıyla hazırlanan çevrimiçi ölçek/başarı testi sınıflandırma formunun, her bir çalışma için doldurulmasıyla gerçekleştirilmiştir. Bu program verilerin Microsoft Excel ortamına düzenli olarak sunulmasını sağlamıştır. Böylece sonuçlar grafik, frekans ve yüzde tablolarına dönüştürülerek, betimsel biçimde sunulmuştur.

3.Sonuç ve Tartışma

Bu içerik analizi çalışmasıyla ülkemizde fen eğitimi alanında geçerlik ve güvenirlik çalışması yapılarak geliştirilen/uyarlanan ölçeklerin ve başarı testlerinin eğilimleri incelenmiştir. Mevcut durumun resmini çizebilmek, bu çalışmaları derleyip, bir çatı altında toplayıp değerlendirebilmek için ise 2002-2013 yılları arasında 21 farklı dergide yayınlanmış (bkz. Ek-1) fen eğitimi alanında geliştirilen/uyarlanan ölçeklerden/başarı testlerinden oluşan toplam 62 makalenin doküman incelemesi yapılmıştır.

Ülkemizde fen eğitimi alanında ölçek geliştirme/adaptasyon çalışmalarına 2002 yılından itibaren bir ilginin olduğu, ölçek geliştirme çalışmalarının 2009 yılında, başarı testi geliştirme çalışmalarının ise 2011 yılında diğer yıllara göre daha fazla sayıda olduğu anlaşılmaktadır.

Geliştirilen/adaptasyon çalışması yapılan ölçeklerde ve başarı testlerinde daha çok fen ve teknoloji alanında yapılan çalışmaların ağırlık kazandığı, örneklem grubu olarak lisans ve ortaöğretim öğrencileriyle yapılan çalışmaların daha çok tercih edildiği anlaşılmaktadır. Bu durum Türkiye'deki eğitim araştırmalarındaki eğilimleri inceleyen Göktaş *vd.,* (2012)'ın çalışmasında Türkiye'deki eğitim araştırmalarında örneklem grubu olarak lisans öğrencileri ve öğretmenlerin tercih edildiği bulgusuyla uyum içerisindedir.

Geliştirilen/adaptasyon çalışması yapılan ölçeklerde ve başarı testlerinde örneklem büyüklüğü olarak ise 101-200 aralığındaki örneklem büyüklükleriyle yapılan çalışmaların daha çok tercih edildiği anlaşılmaktadır. Gözlem sayılarının üst sınırından daha çok alt sınırlarının tartışma konusu olduğu günümüzde bu konuda birçok çalışma yapılmıştır. Pedhazur (1997)'e göre madde ve gözlem sayılarının oranı 1:15 veya 1:30 şeklinde olmalıdır (akt. Osborne & Costello, 2004). Yani madde başına 15 ya da 30 gözlem şeklinde bir oran önermiştir. Gorsuch ise minimum oranı 1:5 olarak ifade etmekte ancak bunun bir alt sınır olduğunu, bu oranın aşağısının kesinlikle kullanılmaması gerektiğini ve bu oranın ne kadar üzerine çıkılırsa o kadar iyi olacağını ifade etmektedir

(Gorsuch, 1983 p.332, akt. Osborne & Costello, 2004). Nunnally ve Bernstein (1994) ise bu oranın 1:10 olacağını söylemektedir. Costello ve Osborne ise yaptıkları çalışmada 1076 yayınlanmış makaleyi incelemişler ve makalelerin %40.5'inde faktör analizlerindeki madde gözlem oranının 1:5'in altında olduğunu gözlemlemişlerdir (akt. Osborne & Costello, 2004). Mutlak gözlem sayısını ölçüt olarak kabul eden Guadagnoli & Wayne (1988)'ye göre ise yeterli mutlak gözlem sayısı 100-200 arasında olmalıdır. Aleamoni (1976)'a göre ise faktör analizi için örneklem genişliği 400 olarak belirtilmektedir (akt. Guadagnoli & Velicer, 1988). Comfrey ve Lee (1992)'e göre ise mutlak gözlem genişliği: 50-(çok zayıf), 100-(zayıf), 200-(kararsız), 300-(iyi), 500-(çok iyi) ve 1000 ve yukarısı ise ideal olarak kabul edilmektedir (akt. Osborne & Costello, 2004). Madde sayısı : gözlem sayısı oranı, mutlak gözlem sayısı ve beklenen faktör : gözlem sayısı oranını da ölçüt olarak kabul eden Osborne & Costello, (2004)'nun Monte Carlo simülasyonu ile yaptıkları çalışmalarında madde sayısı: gözlem sayısı için 1:10, mutlak gözlem sayısı için 1000 ve faktör başına düşen madde sayısı için de 11'in ideal değerler olduğunu belirtmişlerdir.

Geliştirilen ölçeklerdeki ortalama madde sayısının $\overline{X}$ = 29.4, Türkçe'ye uyarlananlar da $\overline{X}$ = 29.7 ve başarı testi çalışmalarında ise $\overline{X}$ = 20.1 olduğu düşünüldüğünde içerik analizi yapılan çalışmalarda 101-200 aralığındaki örneklem büyüklüğünün tercih edildiği %27 oranındaki çalışmanın ve 201-300 aralığındaki örneklem büyüklüğünün tercih edildiği %19 oranındaki çalışmanın ilgili literatürle çok da uyuşmadığı anlaşılmaktadır.

Diğer taraftan içerik analizine tabi tutulan çalışmalarda daha çok öğrencilerin tutumlarını belirlemeye yönelik ölçeklerin geliştirildiği görülmektedir (%32). Geliştirilen/adaptasyon çalışması yapılan ölçeklerin büyük bir çoğunluğunda açımlayıcı faktör analizinin tercih edildiği (%68), çok az bir çalışmada ise doğrulayıcı faktör analizinin (%4) tercih edildiği görülmektedir. Tek boyutlu ölçeklerde toplam varyansın minimum %30'unun maksimum ise %51'inin açıklandığı çok boyutlu olan ölçeklerde ise toplam varyansın minimum %35'inin ve maksimum ise %78'inin açıklandığı görülmektedir. Fen eğitimi alanında geliştirilen ve Türk kültürüne uyarlanan ölçeklerin faktör yük değerlerine bakıldığında .28-.95 aralığında olduğu görülmektedir. Diğer taraftan içerik analizi yapılan ölçeklerin alt boyutlarının Cronbach-Alpha değerlerinin .55-.97 aralığında, ölçeğin genelinde ise .74-.98 aralığında olduğu görülmektedir. Bu durum Price & Mueller (1986)'a göre Cronbach-Alpha değerinin minimum sınırı .70 olmalıdır şeklindeki görüşleriyle paralellik göstermektedir.

Eğitim araştırmalarını tekrardan ve dağınıklıktan kurtarabilmek arzu edilmektedir. İş yükü ve araştırmalara ulaşma sıkıntısı yaşayan öğretmenler ve araştırmacıların eğitim alanındaki gelişmeleri takip etmelerini sağlamak için bu tür çalışmalarla bilgilendirilmeleri gerekir. Bilimsel araştırmaların köşe taşlarından birisi geçerliği ve güvenirliği sağlanmış ölçme araçlarıdır. Durum böyle olunca geçerliği ve güvenirliğinin sağlanmasında hangi tür aşamalardan geçildiğinin özetlendiği bu tür çalışmalara ihtiyaç duyulmaktadır.

Fen eğitimi alanında geliştirilen ve uyarlanan ölçeklerin ve başarı testlerinin betimsel içerik analizinin yapıldığı bu çalışma 21 dergide yayımlanmış toplam 62 makale ile sınırlı olup, farklı kaynak çeşitliliğine (bildiri, tez vb.) gidilerek yapılacak çalışmalardan elde edilecek sonuçlara yer verilmesi önerilmektedir.

Appendix-1 : *List of Journals*

	Journal Title	Scanned years	f
1	Bogazici University Faculty of Education	1974-2010	---
2	Ankara University Journal of Faculty of Educational Sciences	1985-2012	---
3	Hacettepe University Journal of Education	1986-2012	6
4	Marmara University Journal of Ataturk Educational Faculty of Educational Sciences	1989-2012	---
5	Pamukkale University Journal of Education	1996-2013	3
6	Abant Izzet Baysal University Journal of Education Faculty	2000-2012	---
7	Cukurova University Faculty of Education Journal	2000-2013	1
8	Eurasian Journal of Educational Research (EJER)	2000-2013	5
9	Kastamonu University Kastamonu Education Journal	2000-2013	5
10	Ahi Evran University Journal of Kırşehir Education Faculty	2000-2013	5
11	Journal of Uludag University Faculty of Education	2001-2012	2
12	Gazi University Journal of Gazi Educational Faculty (GUJGEF)	2001-2013	1
13	Educational Sciences: Theory & Practice (ESTP)	2001-2013	---
14	Educational Sciences & Practice	2002-2012	1
15	Erzincan University Journal of Education Faculty	2002-2012	2
16	Electronic Journal of Social Sciences	2002-2013	---
17	Elementary Education Online	2002-2013	5
18	Turkish Online Journal of Educational Technology	2002-2013	1
19	Ataturk University Journal of Kazım Karabekir Education Faculty	2003-2012	---
20	Gazi University Journal of Turkish Educational Sciences	2003-2012	---
21	Yüzüncü Yıl University Journal of Education Faculty	2004-2012	2
22	Inonu University Journal of the Faculty of Education	2004-2013	2
23	Journal of Turkish Science Education	2004-2013	9
24	Dokuz Eylul University Journal of the Buca Education Faculty	2005-2012	4
25	Journal of Theory and Practice in Education	2005-2013	1
26	Mersin University Journal of the Faculty of Education	2005-2013	1
27	Faculty of Necatibey Education Electronic Journal of Science and Mathematics Education	2007-2012	3
28	Education and Science	2007-2013	---
29	Mehmet Akif Ersoy University Journal of Education Faculty	2009-2013	1
30	Contemporary Educational Technology	2010-2013	---
31	Ondokuz Mayıs University Journal of Education Faculty	2011-2012	---
32	Mustafa Kemal Universitesi Sosyal Bilimler Enstitü Dergisi	2004-2013	2
Total			62