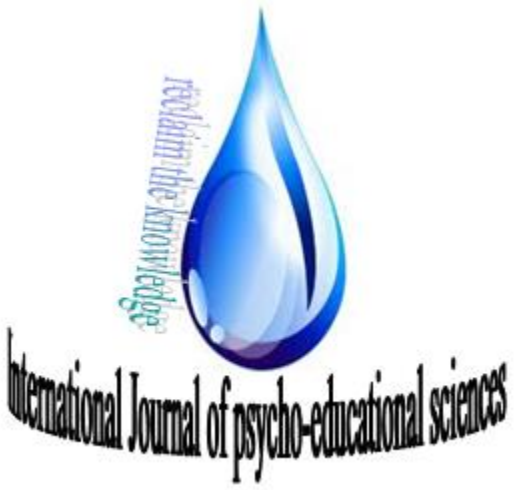


Development of Science Education Peer Assessment Scale: Validity and Reliability Study*

Adem Ozdemir[1], Yasemin Koc[2] & Kerim Gundogdu[3]

*The paper was orally presented in the XVIII. European Conference on Social and Behavioral Sciences Conference, Bucharest, Romania.

[1] Prof. Dr., Adnan Menderes University, Aydin-Turkey, ademozdemir@adu.edu.tr

[2] Assist Prof. Dr., Mustafa Kemal University, Hatay-Turkey, yaseminkoc@mku.edu.tr

[3] Prof. Dr., Adnan Menderes University, Aydin-Turkey, kerim.gundogdu@adu.edu.tr

Abstract

The aim of this research is to develop a scale that prospective science teachers in the Education Faculties compare themselves to their peers according to the "Science field Teacher and Professional Skills" courses. For this reason, 25 items related to Physics, Chemistry, Biology, Science Experiments and Professional Skills courses were prepared. Scale forms prepared in 5 scaled Likert types were administered to 298 students enrolled in 2nd, 3rd and 4th years in two universities. Data were analyzed in SPSS 22.00 for Exploratory Factor Analysis (AFA) and in the Lisrel 8.8 statistics software for Confirmatory Factor Analysis (DFA). As a result of the AFA, the factor load is not determined as its own factor by subtracting 1 item from the contingent item and 1 item. The remaining 23 items, obtained from the Varimax analysis, resulted in 5 sub-dimensions including four field education and one occupational skill dimension. The total variance explained by the structure is 72.62% and the factor loadings range from 0.66 to 0.89. BMD value 0.89 Barttlet test significance level $p<0.001$. The reliability coefficient of the scale is Cronbach $\alpha = 0.91$. The Gutman Split-Half and Spearman-Brown coefficients are also good for the scale and both coefficients are 0.92. According to the results of DFA, the compliance indices are quite good (RMSEA= 0.051, Chi-square= 387.75, df= 217 RMR= 0.04, CFI= 0.98, IFI= 0.98, RFI= 0.95, NFI= 0.96, NNFI). = 0.98, GFI = 0.90, AGFI = 0.87). According to these results, it may be said that this scale is a valid and reliable measurement tool that science teacher candidates can compare their knowledge and professional skills in this area with their peers.

Keywords: *Science education, peer assessment, scale development, validity, reliability.*

Introduction

The academic self-concept, which is an important aspect in the academic achievement and evaluation process, is defined by Arseven (1986) as the belief that a student develops his / her competence in relation to a certain academic occupation compared to other students. Research shows that students who have similar characteristics have success in learning process in favor of students having high academic self-concept (Saracaloğlu & Varol, 2007). Although the perceptions are important variable on academic success and attitudes towards science laboratory (Şenler, Karisan, Bilican, 2017) the concept of self does not only occur with one's own perceptions or expectations, but also the thoughts of the people around him or her. Peer opinions or opinions, especially in adolescence or young adulthood, are influential on the individual since adolescence is a period of intense emotional situations.

Social development is the period in which a person learns to act in harmony with the society in which he / she lives. According to Horroks, the peer group gives young people an opportunity to understand and interpret the world, and helps them to make sense of themselves and the universe (Demir et al., 2005). Self-concept is divided into two groups as academic and non-academic. The academic self-concept includes self-concept of language, social studies, mathematics and science courses developed by the individual about the courses of the school; The non-academic self is the self-concept of the individual's social relationships, emotional life and physical appearance (Arseven, 1986).

People are actively involved in defining their identities. Identity formation processes that involve many different factors, such as parents and other powers, friends, institutions and groups, are quite complex. The views reflected by friends and friends, as a result of social comparisons, self-evaluation and identification with social groups are seen as important predictors of the concepts of identity of young people. Studies on identity development show that there is a positive and meaningful relationship between friendship, group identification and identity orientation (Aslan & Dönmez, 2013).

Age, profession, social situation, etc. in terms of each other is defined as *peer*. Peers who are defined as equal to each other, have an important role in many periods of life socially

and emotionally. Peer group is a forum where values and attitudes are discussed. In this way, it helps to gain personal independence (Çırpan and Çınar, 2013). Peer relations are defined as the whole of interaction with the same age, level of development or maturity, with continuity between the people sharing similar past, value, life, lifestyle and social context. Most often, peer relations and friendship concepts are mixed. However, friendship is an emotional bond with several peers. Of course, peer relations of all ages are different from each other (Gülay, 2010). Peer relations are a multi-faceted relationship that involves positive and negative behavioral examples and interacts with other social relations in society (Rodkin & Hodges, 2003).

Positive or negative perceptions are the research area of many subjects such as quantum thought. Whether it is explained by energy or a cognitive process, what one thinks about himself is so important that his peers think and compare his ideas about himself or vice versa. Most of our lives in the time of comparison to make conscious or unconscious. In fact, we do a comparison of everything we perceive physically (hot-cold, small-large, etc.). But of course, what is important in the self-perception of his self is about his comparisons. Social comparison theory was developed by Lean Festinger in 1954. Festinger argues that people are fully motivated when they are aware of their abilities, and that in order to understand this, the individual compares himself to other people. Social comparison theory can be summarized as follows (Karasakal & Aksu, 2014):

• People have the process of developing their own thoughts and abilities

• People develop themselves by comparing themselves with other people in the absence of physical standards

• In general, people prefer to encounter their own.

Feslinger's Social Comparison Theory has many explanations about the development of individual success, the causes of attitudes and judgments and the formation of the concept of self (Buunk and Mussweiler, 2001). The Social Comparison Theory was born as a theory which tried to understand the self-assessment of the individual in the 1950s. According to Feslinger (1954), there is a universal impulse to evaluate man's convictions and abilities. These evaluations should be as straightforward, objective and realistic as it is likely to have negative opinions or to misrepresent the ability to have a non-valid judgment. For this reason, people search for physical standards when evaluating (Teközel, 2007). Within the scope of this theory, when individuals compare themselves with others, they go towards “up” or “down ile comparisons; In the upward comparisons, it is stated that the individual compares himself with other individuals who are superior to himself, and in the case of downward comparisons, it is stated that the individual compares himself with the individuals with lower levels (1). As a result of these comparisons made in the light of continuous personal and other evaluations, the individual sets out his own social and personal value (charm, success, intelligence etc.) (Bilbek & Yılmaz, 2014).

There is consensus on that personal and motivational variables have an impact on learning (Karisan & Yilmaz-Tüzün, 2013). A student's peer comparison in a way he perceives himself and his friends has motivational consequences (Salmivalli, Ojanen, Haanpaa, & Peets, 2005). One way for a person to learn about himself is to make a comparison. The fact that the individual knows about himself is increasing during his adolescence (Aydın, 2005). Learning environments, in which students are active learners and construct their own knowledge through personal experiences, are seen as important places to achieve this goal (Karisan, Bilican Senler, 2017). In Turkey, the university environment is one of the most important individual development environments in which social diversity is maximized in

terms of social comparison and the development of the self-concept of the younger generation is ensured. Another comparison type in this learning environment where both social comparisons and other comparisons are made is to make peer comparisons according to academic interest. One aim of education is to ensure the self-confidence of people and to help them achieve a positive sense of self. Personality tests, autobiography and so on to help the person know his / her self. techniques are used. Recently, however, alternative evaluation approaches have emerged. Self-assessment and peer evaluation are among the alternative assessment approaches. In this study, peer comparison is discussed.

Purpose

In this study, it is aimed to develop the Science Education Peer Comparison Scale (SEPCS) in order to measure the comparison of Science Education students' their own academic study fields in science education. It is thought that the scale will help researchers in peer comparison studies.

Method

Participants

In the development of the SEPCS, 298 students from the 2nd, 3rd and 4th years of two teacher training institution in two public universities participated in the study. The research was conducted in the 2017-2018 academic year. According to Tabachnick and Fidel (2001), 300 people are considered sufficient for factor analysis. The numerical distributions of the participants are given in Table 1.

Table 1. Distribution of participants by university, class and gender.

	2nd year		3rd year		Last year		Total
	Female	Male	Female	Male	Female	Male	
University 1	37	10	36	11	39	10	143
University 2	45	8	44	9	39	10	155
Total	82	18	80	20	78	20	298

Scale

Science Education Peer Comparison Scale is prepared for the comparison of field education and vocational skills courses as required by the undergraduate program of Science Education. For this reason, a 25-item pool of 5 dimensions, namely physics, chemistry, biology, science experiments and vocational skills courses, was created. The level of participation in the items is 5-point Likert type and is rated as `Never (1), Rarely (2), Sometimes (3), Frequently (4) and Always (5)`. After these procedures, validity and reliability analyzes were computed.

Procedures

Construt, face and content validity of the scale were examined. Evaluation of face and content validity were realised with two academicians science field education and two educational sciences experts. As a result of the corrections made by the experts, the scale consisted of 25 items. After this process, 298 Science Education students being in 2nd, 3rd and 4th years were received the SEPCS. The application period of the scale varies between 5-10 minutes.

The construct validity of the scale was investigated by using Exploratory Factor Analysis (EFA). Kaiser-Meyer Olkin (KMO) coefficient and Barlett`s Sphericity test results were analyzed to determine the suitability of the data for factor analysis. Cronbach α, Gutman Split-Half and Spearman-Brown coefficients were calculated to provide evidence for reliability. In order to provide evidence for item validity, item correlations and item test correlations were examined. Confirmatory Factor Analysis (CFA) was performed to confirm the structure. Data for AFA were used for SPSS 22, CFA for Lisrel 8.8 statistical programs. In addition, the correlation coefficients between the factors forming the scale and the student transcript grades were also calculated.

Findings

The construct validity of the SEPCS was examined using principal component analysis which is a technique of the Factor Analysis. The suitability of the factor analysis (sample size and adequacy) and the appropriateness of the data to the normal distribution were tested before starting the analysis. Accordingly, the KMO value is 0.89 and this value is expressed by Pallant (2007), where the factorability condition is met when it is greater than 0.60. For the Barlett`s Sphericity test, the significance value of küçük2 = 5022.597 sd = 300 p <. In order to determine the suitability of each item for factor analysis, it was decided to include substances with an anti-image correlation value greater than 0.50 in factor analysis (Sipahi, Yurtkoru and Zinko, 2008). These results show that factor analysis can be performed on the scale.

Findings Related to Factor Analysis

The concept of construct validity of the scale is related to revealing the conceptual structure. Factor analysis technique is first applied to determine the construct validity (Field, 2005). Factor analysis data are shown in Table 2.

Factor analysis is used to reveal the basic structure of a large number of variables (Şencan, 2005). As a result of Varimax rotational factor analysis, it was seen that item 5 did not emerge in its own dimension and item 24 had high load value in more than one factor. Therefore, KMO value was calculated as 0.89 and Barlettitys Sphericity test P2 = 4549.737 sd = 253 P <0.001. Factor analysis of the scale structure initially designed by researchers physics (5 items), Chemistry (5 items), Biology (5 items), Science Experiments (4 items) and Professional Skills (4 items) were observed to overlap with the 5-dimensional structure. A factor with multiple factors states that at least 3 items should be present (Little, Lindenberger & Nesselroade 1999; Velicer & Fava, 1998). It is recommended to remove substances with a load value of less than 0.30 (Field, 2005). Factor load values vary between 0.66-0.89. Factor load value explains the relationship between the factor and factor is expected to be high (Kline, 1994). It was noted that the common variance was greater than 0.50 (Köklü, 2002; Çokluk et al, 2010). The smallest common variance is 0.60. The variance explained for the whole scale was 72.62%. Cronbachmaktads α internal consistency coefficient is 0.91 and in Table 2, common variances, anti-image correlation coefficients, factor loads, Cronbach's coefficients, eigenvalues and substances of sub-dimensions are presented. The Gutman Split-Half and Spearman-Brown coefficients were also good and the coefficient was 0.92.

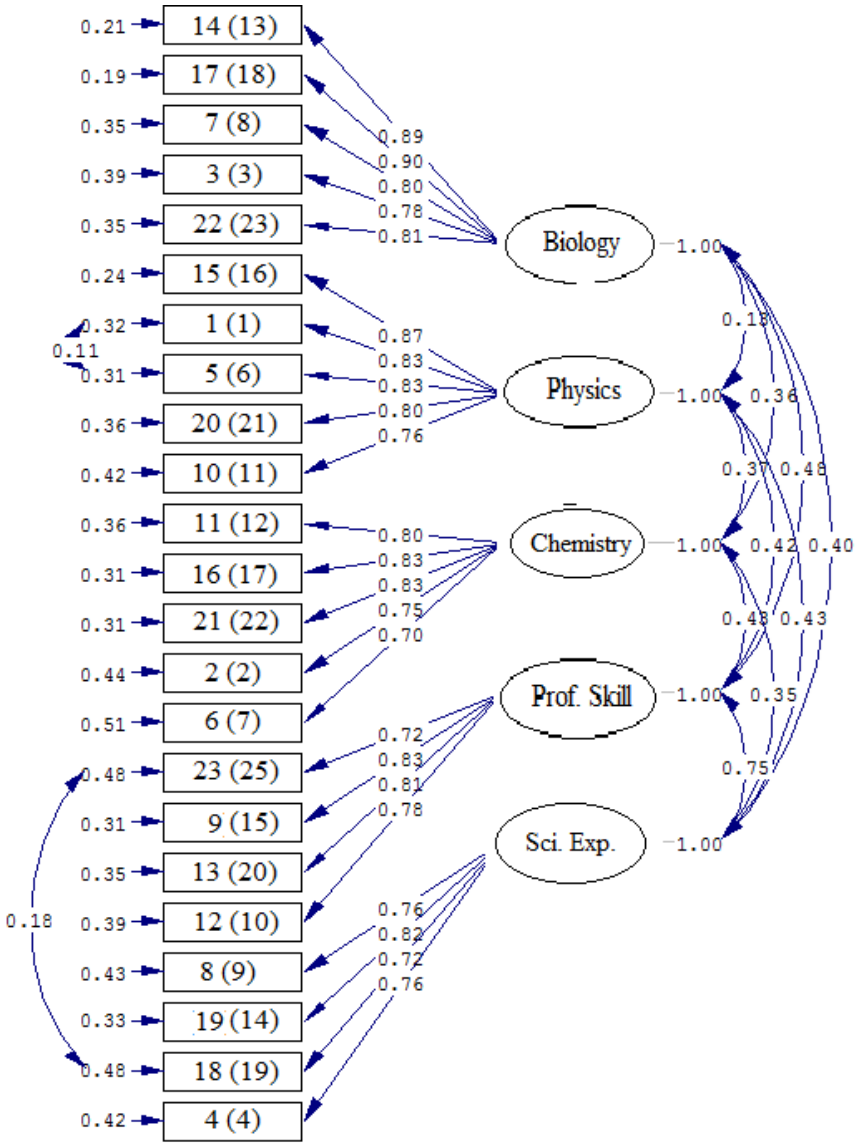
Table 2. Results of Factor Analysis of Science Education Peer Comparison Scale.

Sub Factors	Item No*	Items	Common Variance	Anti İmage r	Factor Load
BIOLOGY	14 (13)	I'm as good as my friends in biology.	0.84	0.88	0.89
	17 (18)	I'm as successful as my friends in biology.	0.83	0.86	0.87
	7 (8)	I'm as good at biology as my friends.	0.76	0.88	0.85
	3 (3)	I'm not as scared of biology as my friends.	0.68	0.90	0.80
	22 (23)	I like biology topics as much as my friends	0.70	0.91	0.79
	Cronbach α= 0.92 Explained Variance: 16.904 Eigenvalue: 3.388				
PHYSICS	15 (16)	I'm as skillful as my friends in physics.	0.79	0.91	0.86
	1 (1)	I'm as good as my friends in physics.	0.79	0.87	0.86
	5 (6)	I'm as good as my friends in physics.	0.80	0.86	0.84
	20 (21)	I like physics subjects as much as my friends.	0.73	0.89	0.83
	10 (11)	I'm not as afraid of physics as my friends.	0.68	0.88	0.79
	Cronbach α= 0.91 Explained Variance: 16.380 Eigenvalue: 3.767				
CHEMISTRY	11 (12)	I like chemistry as much as my friends.	0.75	0.83	0.84
	16 (17)	I'm as good as chemistry with my friends.	0.76	0.88	0.84
	21 (22)	I'm as good at chemistry as my friends.	0.74	0.92	0.81
	2 (2)	I'm as successful as chemistry in my friends.	0.71	0.88	0.76
	6 (7)	I'm not as scared of chemistry as my friends.	0.61	0.89	0.75
	Cronbach α= 0.89 Explained Variance: 15.172 Eigenvalue: 3.490				
PROFESSIONAL SKILLS	23 (25)	I'm not as scared of my profession skills as my friends.	0.69	0.89	0.76
	9 (15)	I'm as skillful as my friends in professional skills.	0.72	0.91	0.74
	13 (20)	I'm as successful as my friends in professional skills.	0.71	0.88	0.73
	12 (10)	I'm as good as my friends in vocational skills.	0.66	0.91	0.72
	Cronbach α= 0.86 Explained Variance: 12.365 Eigenvalue: 3.844				
SCIENCE EXPERIMENTS	8 (9)	I like to do science experiments as much as my friends.	0.77	0.85	0.85
	19 (14)	I'm as successful as my friends in doing science experiments.	0.75	0.92	0.77
	18 (19)	I'm not as scared of doing science experiments as my friends.	0.60	0.89	0.67
	4 (4)	I'm as skillful as my friends in doing science experiments.	0.63	0.94	0.66
	Cronbach α= 0.84 Explained Variance: 11.800 Eigenvalue:2.714				
Total Cronbach α= 0.91 Explained Variance : %72.62					

* The old item numbers are given in brackets.

Findings Related to Confirmatory Factor Analysis

A five-item, 23-item construct with AFA was tested with DFA. Compliance indices calculated as a result of the analysis were re-calculated by adding two modification corrections to the model and the fit indices of the model were recalculated. Accordingly, RMSEA= 0.051, Chi-Square= 387.75, df= 217 RMR= 0.04, CFI= 0.98, IFI= 0.98, RFI= 0.95, NFI= 0.96, NNFI= 0.98, GFI= 0.90, AGFI= 0.87 was calculated. Figure 1 shows the standardized values and error variances of the model.


14 (13)
17 (18)
7 (8)
3 (3)
22 (23)
15 (16)
1 (1)
5 (6)
20 (21)
10 (11)
11 (12)
16 (17)
21 (22)
2 (2)
6 (7)
23 (25)
9 (15)
13 (20)
12 (10)
8 (9)
19 (14)
18 (19)
4 (4)
0.21
0.19
0.35
0.39
0.35
0.24
0.32
0.11
0.31
0.36
0.42
0.36
0.31
0.31
0.44
0.51
0.48
0.31
0.35
0.39
0.43
0.33
0.48
0.42
0.18
0.89
0.90
0.80
0.78
0.81
0.87
0.83
0.83
0.80
0.76
0.80
0.83
0.83
0.75
0.70
0.72
0.83
0.81
0.78
0.76
0.82
0.72
0.76
Biology
Physics
Chemistry
Prof. Skill
Sci. Exp.
1.00
1.00
1.00
1.00
1.00
0.18
0.36
0.37
0.48
0.42
0.40
0.48
0.43
0.35
0.75

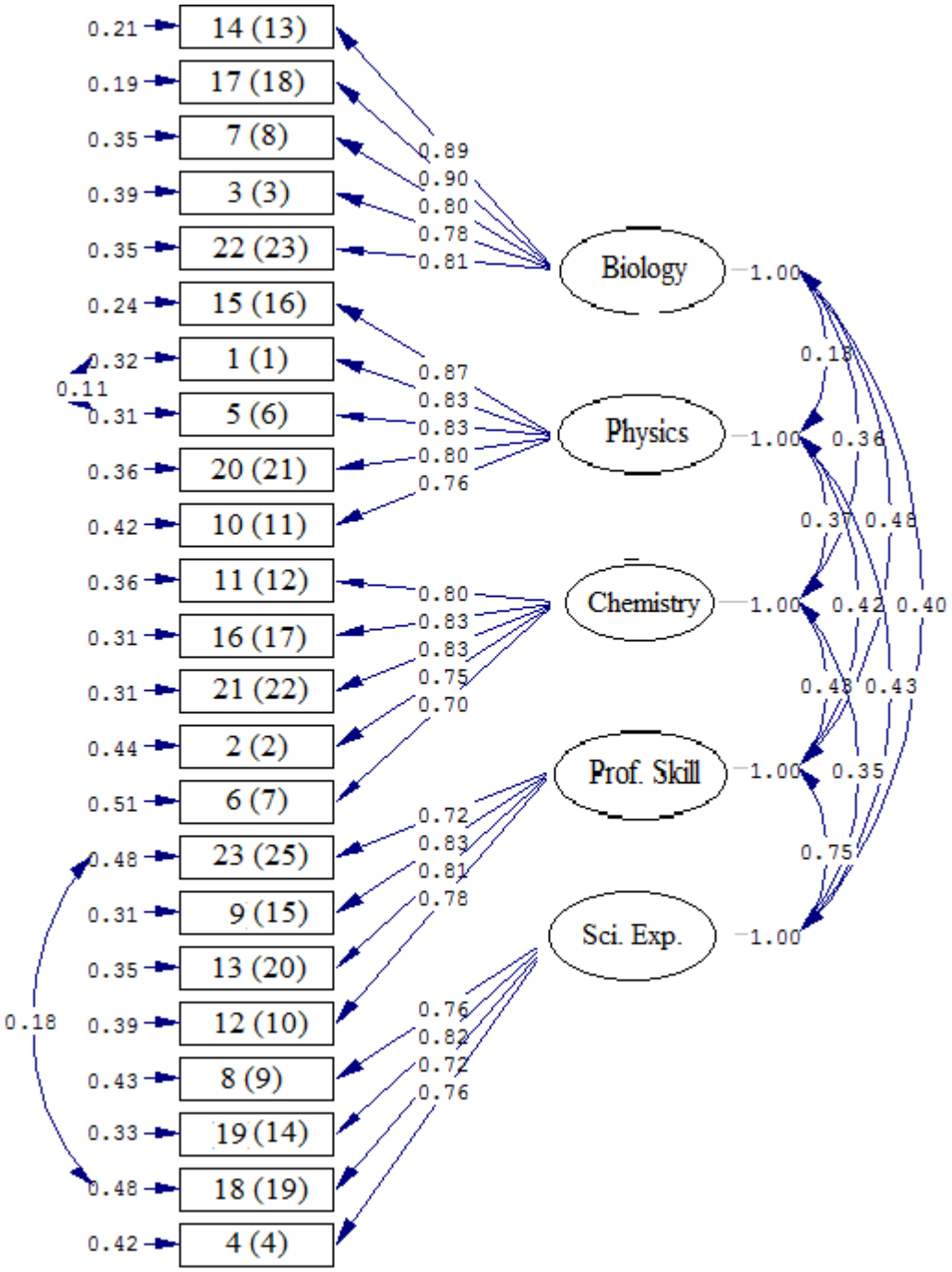


Figure 1. DFA Analysis Standard Values and Error Variances.

Results

In this study, Science Education Peer Comparison Scale was developed for the undergraduate students of the science teacher tranings programs. In order to determine the psychometric properties of the scale, the validity and reliability studies of the scale were done with 23 items. In line with the expert opinions the scale was finalized. This final version of the scale was implementated with a total of 298 prospective teachers from two teacher training faculty in two different universities on a voluntary basis.

In order to determine the factor structure, exploratory factor analysis and confirmatory factor analysis were performed to determine the accuracy of this structure. As a result of exploratory factor analysis, explaining 72.62% of total variance; A total of 23 items were collected in a five-factor structure including biology, physics, chemistry, profession skills and science experiments. Items 3, 7, 14, 17 and 22 Biology 3; items 1, 5, 10, 15 and 20 Physics 5; items 2, 6, 11, 16 and 21 Chemistry 2; items 9, 12, 13 and 23 Professional Skills 9; items 4, 8, 18 and 19 were included in the 'Science Experiments 4 dimensions. The score which can be taken from the five-point Likert scale varies between 23 and 115. The low scores of the trainees in the scale were weak when compared with their peers in terms of the factors in the scale dimensions; The higher the scores they get, the more they can be interpreted as being perceived as more adequate. The fact that there is a significant relationship between the students' academic achievement and FEAKÖ contributes to the validity of the scale. It can be said that Science Education Peer Comparison Scale which is obtained as a result of this study is a valid and reliable measurement tool which can be used in peer comparison studies of Science Teacher Education programs.

References

Arseven, D. A. (1986). Çocukta benlik gelişimine ailenin etkisi ve çocuğun okuldaki başarısı [The effect of the family on the development of self in the child and the success of the child in school]. *Education and Science. 10* (60), 11-17.

Arseven, A. D. (1986). Benlik tasarımı okul gelişimi ve başarısı ile ilişkisi [The relationship between academic self-concept and school development and success]. *Hacettepe Üniversity Education Faculty Journal,* 1, 15-26.

Aslan S. ve Dönmez A. (2013). Gençlerde kimlik yönelimi, toplumsal karşılaştırma, arkadaşlara bağlılık ve grupla bütünleşme [Identity orientation, social comparison, commitment to friends and integration with groups]. *Mülkiye Dergisi, 37*(3), 141-173.

Aydın, B. (2005). *Çocuk ve ergen psikolojisi [Child and adolescent psychology, 2nd ed.]*. İstanbul: Atlas Yayınları.

Bilbek, M. ve Yılmaz, C. Y. (2014). Beden eğitimi ve spor yüksekokulu öğrencilerinin sosyal karşılaştırma düzeylerinin çeşitli değişkenler açısından incelenmesi [Examining social comparison levels of physical education and sports college students in terms of various variables.], *Beden Eğitimi ve Spor Bilimler Dergisi*, *12* (2), 105-112.

Buunk, B. P., & Mussweiler, T. (2001). New directions in social comparison research. *European Journal of Social Psychology, 31*(5), 467-475.

Çırpan, F.ve Çınar, S. (2013). Sağlık hizmetleri meslek yüksekokulu öğrencilerinde akran desteği ile akademik başarı arasındaki ilişkinin değerlendirilmesi [Evaluation of the relationship between peer support and academic achievement in vocational high school students of health services]. *Marmara Üniversitesi Sağlık Bilimleri Enstitüsü Journal.* 3, 191–199.

Çokluk, Ö., Şekercioğlu, G. & Büyüköztürk, Ş. (2010). *Sosyal bilimler için çok değişkenli istatistik [Multivariate statistics for social sciences]*. Ankara: Pegem Akademi.

Demir, N. Ö., Baran, A. G., & Ulusoy, D. (2005). Türkiye'de ergenlerin arkadaş-akran grupları ile ilişkileri ve sapmış davranışlar: Ankara örneklemi [relationships and deviant behavior with adolescents friend-peer groups in Turkey: Ankara sample]. *Bilig/Türk Dünyası Sosyal Bilimler Dergisi*, 32, 83-108.

Field, A. (2005). *Discovering statistics using SPPS*. London: Sage Publications.

Gülay, H. (2010). *Okul öncesi dönemde akran ilişkileri [Peer relations in pre-school period]*. Ankara: Pegem Akademi.

Karasakal, N. K. & Aksu, B. (2014). Meslek yüksekokulu teknik ve sosyal program öğrencilerinin sosyal karşılaştırılması sonucu girişimci kişilik eğilimlerini ortaya koymaya yönelik bir alan araştırması [A field study aimed at revealing entrepreneurial personality trends as a result of the social comparison of vocational and social program students.]. *Organizasyon ve Yönetim Bilimleri Dergisi, 6* (2), 102-116.

Karisan, D., Yilmaz-Tuzun, O. (2013). An Exploration Of Undergraduate Engineering, Education, Art's and Sciences Students' Chemistry Laboratory Anxiety Levels. *International Journal on New Trends in Education and Their Implications,* 4 (4), 75-86.

Karisan, D., Bilican, K., Senler, B. (2016). Yansıtıcı Sorgulayıcı Laboratuvar Etkinliklerinin

Sınıf Öğretmen Adaylarının Bilimsel Süreç Becerilerine Etkisinin İncelenmesi. *YYÜ Eğitim Fakültesi Dergisi (YYU Journal Of Education Faculty),*13(1), 123-145.

Kline, P. (1994). *An easy guide to factor analysis*. UK: Routledge.

Little, T. D., Lindenberger, U. & Nesselroade, J. R. (1999). On selecting indicators for multivariate measurement and modeling with latent variables: when good ındicators are bad and bad ındicators are good. *Psychological methods, 4*(2), 192-211.

Pallant, J. (2007). *SPSS survival manual: a step-by-step guide to data analyses using SPSS for Windows.* Philedelphia, PA: Open University Press.

Köklü, N. (2002). *Açıklamalı istatistik terimleri sözlüğü [Glossary of explanatory statistical terms]*. Ankara: Nobel.

Rodkin, P. C., & Hodges, E. V. (2003). Bullies and victims in the peer ecology: Four questions for psychologists and school professionals. *School Psychology Review*, *32*(3), 384-400.

Salmivalli, C., Ojanen., Haanpaa., J., & Peets, K. (2005). "I'm ok but you're not" and other peer-relational schemas: explaining individual differences in children's social goals", *Developmental Psychology, 41*(2), 363-375.

Saracaloğlu, A. S & Varol, S. R.(2007). Beden Eğitimi Öğretmeni Adaylarının Yabancı Dile Yönelik Tutumları Ve Akademik Benlik Tasarımları İle Yabancı Dil Başarıları Arasındaki İlişki [The Relationship Between Prospective Teachers' Attitudes Towards Foreign Language And Academic Self-Concept and Foreign Language Achievements]. *Eğitimde Kuram ve Uygulama.* 3 (1):39-59.

Sipahi, B., Yurtkoru, E.S. & Çinko, M. (2008). *Sosyal bilimlerde SPSS'le veri analizi [Data analysis in social sciences with SPSS]*. İstanbul: Beta.

Şencan, H. (2005). *Sosyal ve davranışsal ölçümlerde güvenirlik ve geçerlik* [Reliability and validity in social and behavioral measurements]. Ankara: Seçkin.

Şenler, B**.,** Karışan, D., & Bilican, K. (2017). Sınıf Öğretmeni Adaylarının Fen ve Teknoloji Laboratuvarına Yönelik Algı ve Tutumlarının İncelenmesi. Mehmet Akif Ersoy Üniversitesi Eğitim Fakültesi Dergisi, (42), 105-122. Doi: http://dx.doi.org/10.21764/efd.14921

Tabachnick, B. G. & Fidel, L. S. (2001). *Using multivariate statistics.* Needham Heights, MA: Allyn & Bacon.

Teközel, İ. M. (2007). *Gerçekliği inşa etkinliği olarak sosyal karşılaştırma sosyal karşılaştırma süreçlerinde gerçekçi ve kurgusal karşılaştırma enformasyonu kullanma eğilimlerinin incelenmesi [To examine the tendencies of using realistic and fictitious comparison information in social comparison processes as social realization activity.]*. Unpublished doctoral thesis, Ege University, İzmir-Turkey.

Velicer, W. F. & Fava, J. L. (1998). Effects of Variable and Subject Sampling on Factor Pattern Recovery. *Psychological Methods, 3*(2), 231-251.

TDK (2018). *Akran*. 25.08.2018 http://www.tdk.gov.tr/index.php?option=com_bts&view=bts&kategori1=veritbn&kelimesec=8948