



## Bedens Eğitimi ve Spor Dersi Sınıf Değerlendirme Atmosferi Ölçeği (Besdsdaö): Geçerlik ve Güvenirlik Çalışması

*Physical Education And Sport Course Class Assessment Atmosphere Scale (Pesccaas): Validity And Reliability Study*

Oğuz Kaan Esentürk\* - Aynur Yılmaz\*\*

**Abstract:** In this study, it is aimed to develop a measurement tool that can measure the classroom assessment atmosphere of physical education and sports lesson in a valid and reliable way. Exploratory Sequential Pattern, one of the mixed research approaches, was used in the research. The study was carried out on three different study groups attending in different high schools (Anatolian High School, Science High School and Vocational High School) in Kırıkkale province. A total of 466 students from different high schools in Kırıkkale participated in the study. The criterion sampling technique, one of the purposeful sampling methods, was used in the collection of data. Exploratory Factor Analysis (ECA) was performed with the data obtained from the first study group. The structure revealed by exploratory factor analysis was tested with Confirmatory Factor Analysis. Test-retest reliability analysis was performed with the data obtained from the third study group. As a result of exploratory factor analysis conducted within the framework of construct validity, a structure consisting of 19 items and two factors explaining 38.06% of the total variance was obtained. In order to test item discrimination, item test correlation and 27% sub-upper group comparison were performed. Within the scope of criterion-related validity, there were studies indicating a significant relationship between classroom assessment atmosphere and achievement orientation and burnout status. In this respect, the relationship between PELCAAS, Achievement Orientation Scale and Maslack Burnout Scale was examined. The reliability of the instrument was examined with Cronbach's alpha internal consistency coefficient and test-retest techniques. These test techniques provide evidence that the measuring instrument is highly reliable. Criterion validity analysis was used to provide evidence of validity. According to the current findings, it can be said that Physical Education and Sports Lesson Classroom Assessment Atmosphere Scale is a data collection tool that can make valid and reliable measurements.

\* Dr. Öğr. Üyesi, Erzincan Binali Yıldırım Üniversitesi Eğitim Fakültesi  
Asst. Prof. Dr., Erzincan Binali Yıldırım University/Faculty of Education  
ORCID 0000-0002-0566-838X  
esenturk954@gmail.com

\*\* Doç. Dr., Trabzon Üniversitesi Spor Bilimleri Fakültesi Spor Yöneticiliği Bölümü  
Assoc. Prof. Dr., Trabzon University School of Physical Education and Sports Department of Sports  
ORCID 0000-0001-7562-9469  
yilmazaynr@hotmail.com

**Cite as/ Atıf:** Esentürk O. K., Yılmaz, A. (2020). Beden eğitimi ve spor dersi sınıf değerlendirme atmosferi ölçeği (besdsdaö): geçerlik ve güvenilirlik çalışması, *Turkish Studies*, 15(1), 195-211.  
<https://dx.doi.org/10.29228/TurkishStudies.38205>

**Received/Geliş:** 01 October/Ekim 2019

**Accepted/Kabul:** 25 February/Şubat 2020

Copyright © MDE, Turkey

Checked by plagiarism software

**Published/Yayın:** 29 February/Şubat 2020

CC BY-NC 4.0

**Structured Abstract: Introduction** Lewin (1936) believed that human behavior is determined by the complex interaction of the individual and his environment. In other words, Lewin emphasized that interactions between human and environment are determinants of human behavior. Among these environmental conditions, one of the concepts that is considered as a critical environment for students' educational success is class (Anderson and Burns, 1989; Borich, 1988; Fraser and Walberg, 1991; Walberg, 1968). In addition, important findings have been identified that the classroom environment plays a critical role on the individual's behaviors (Bartha, Dunlap, Dane, Lochmanb and Wells, 2004). In a meta-analysis study in which classroom environment studies were used by Haertel, Walberg and Haertel (1981), it was found that students' perceptions of the classroom environment were a critical factor in determining certain aspects of students' learning outcomes such as motivation, achievement and student satisfaction.

Studies conducted to examine students' perceptions of the classroom environment revealed that a significant portion of the variance of learning outcomes was explained by the perceptions of the classroom environment (Dorman, 2001). Therefore, it is thought that observations specific to cognitive and affective behaviors of students should gain importance in classroom environment. One of the important components of the classroom environment is the classroom assessment atmosphere (Banks, 2012). Considering that a significant portion of the time in the classroom is devoted to assessment-related activities (Mertler, 2003), the assessment atmosphere in the classroom is considered to be one of the most important components of the classroom environment (Brookhart and DeVoge, 1999).

Ames (1992a) noted that the following classroom assessment practices are likely to reveal positive beliefs, impact and behavioral models in students: (a) designing assessment tasks that include difficulty, diversity, innovation and active participation; (b) giving students opportunities to make choices and decisions in the assessment process; (c) conducting assessment practices that evaluate progress, development and mastery and avoid social comparisons; and (d) allowing time in the assessment task to vary according to the nature of the task and the needs of the student.

In the light of these educational perspectives, it is thought that students' perceptions of the classroom assessment atmosphere affect their motivational beliefs and achievement (Ames, 1992b; Brookhart, 1997). As in all classes, it can be said that students' perceptions of classroom assessment atmosphere created in physical education and sports classes directly affect the gains to be obtained from the class. There is a scale in the national literature (İlhan and Çetin, 2014) that can be used to measure students' perceptions of classroom assessment atmosphere and a scale has also been found in the international literature (Alkharusi, 2009; 2011). However, the measurement tool that can directly determine the classroom assessment atmosphere for physical education and sports classes has not been found in national and international literature. The aim of this study is to determine the validity and reliability of the classroom assessment atmosphere scale for physical education and sports lesson.

### **Method**

In this study, mixed research approach was used. A total of 466 students from different high schools in Kırıkkale participated in the study. The criterion sampling technique, one of the purposeful sampling methods, was used in the collection of data. The construct validity of the measurement tool was provided by Criterion-related validity as well as EFA and CFA. Within the scope of criterion-related validity, there were studies indicating a significant relationship between classroom assessment atmosphere and achievement orientation and burnout status. In this respect, the relationship between PELCAAS, Achievement Orientation Scale and Maslack Burnout Scale was examined. The data were entered into SPSS 22 and Lisrel 8.8 package programs and necessary operations were performed. In this study, firstly, analysis of construct validity was performed. Within the scope of validity, EFA was first applied and CFA was used to evaluate the compatibility of the construct resulting from EFA with the data collected from a different study group. Item total test correlation and 27% sub-upper group comparisons were evaluated. The reliability of the construct determined by factor analysis was examined by Cronbach Alpha internal consistency coefficient and test-retest reliability methods.

### **Findings**

Construct Validity: EFA and CFA were applied within the scope of the construct validity of the measurement tool. "Principal Component Analysis" was used in the research. In this study, item factor load value was determined as 0.35. In the factor analysis, in the process of subtraction of items that do not measure

the same construct, it was paid attention that the difference between the load values is at least 0.10 when the factor load is less than 0.35 and when a material gives a high load value for two factors so that it does not cause overlap (Büyüköztürk, 2014). Considering the findings, it is seen that the measuring means consists of 19 items and two dimensions. "The Learning Focused Assessment" dimension consists of 13 items and explains 21.224% of the total variance. Factor loads of the items in this dimension ranged from 0.46 to 0.65, with an eigenvalue of 5.503. "The Performance Oriented Assessment" dimension consists of six items and explains 16.832% of the total variance. Material factor loads of the items in this dimension ranged between 0.56 and 0.70 and their eigenvalue was 1.727. Both dimensions account for 38.056% of the total specification.

In order to evaluate the construct validity of the model that emerged after EFA, CFA was performed (Kline, 2011). The indices of fit for the measurement tool were found as  $\chi^2 / sd = 2.15$ , RMSEA = 0.06, PGFI = 0.72, GFI = 0.90, AGFI = 0.87, PNFI = 0.81, IFI = 0.95, NFI = 0.91 and CFI = 0.95. When these fit indices are taken into consideration, it can be stated that the two-dimensional theoretical construct is confirmed by the collected data. Confirmatory factor analysis showed that the regressive link between observed variables and latent variables was between 0.48 and 0.73 in the "Learning Oriented Assessment" dimension and 0.46 and 0.64 in the "Performance Oriented Learning" dimension. Kline (2007) stated that the standardized weights around .30 reflect medium, and .50 and above reflect large effect sizes.

**Item Analysis:** In the context of item analysis, item total test correlation and 27% sub-upper groups were compared and examined. Item total test correlation explains the relationship between the scores obtained from test items and the total score of the test. Item total test correlation values are between 0.36 and 0.69. When the t-test (Kılıç Çakmak, Çebi & Kan, 2014) is examined for the significance of the difference between item scores of the groups, it is seen that t values vary between 2.05 and 10.44.

**Criteria Validity:** As a result of the analysis, it was found that there is a relationship between the scores obtained from the physical education lesson classroom assessment atmosphere scale and the achievement orientation. 43 ( $p < .001$ ) and Maslach burnout scale scores. 45 ( $p < .001$ ).

### Conclusion

As a result of the analyses conducted on the validity and reliability of the measurement tool, it can be said that the Physical Education and Sports Lesson Classroom Assessment Atmosphere (PESLCAA) is a data collection tool that can make valid and reliable measurements.

**Keywords:** Classroom assessment atmosphere scale, Physical education and sports lesson, Adaptation, Validity, Reliability

**Öz:** Bu çalışmada beden eğitimi ve spor dersi sınıf değerlendirme atmosferini geçerli ve güvenilir bir şekilde ölçebilecek ölçme aracı geliştirmek amaçlanmıştır. Araştırmada karma araştırma yaklaşımlarından "Keşfedici Sıralı Desen" kullanılmıştır. Araştırma Kırıkkale ilinde farklı liselerde (Anadolu Lisesi, Fen Lisesi ve Meslek Lisesi) öğrenim gören üç farklı çalışma grubu üzerinde gerçekleştirilmiştir. Toplam 466 öğrenci üzerinde gerçekleştirilen bu çalışmada amaçlı önekleme yöntemlerinde ölçüt örnekleme kullanılmıştır. Birinci çalışma grubundan elde edilen verilerle Açıklayıcı Faktör Analizi (AFA) yapılmıştır. İkinci gruptan elde edilen verilerle AFA sonucu ortaya konulan yapı Doğrulayıcı Faktör Analizi ile sınanmıştır. Üçüncü çalışma grubundan elde edilen verilerle test tekrar test güvenilirlik analizi yapılmıştır. Yapı geçerliği kapsamında yapılan AFA sonucunda toplam varyansın %38.06'sını açıklayan 19 madde ve iki faktörden oluşan bir yapı elde edilmiştir. Madde ayırt ediciliğini test etmek amacıyla madde test korelasyonu ve %27'lik alt-üst grup karşılaştırılması yapılmıştır. Ölçüte dayalı geçerlik analizi sonucunda beden eğitimi dersi sınıf değerlendirme atmosferi ölçeğinden alınan puanlarla başarı yönelimi .43 ( $p < .001$ ) ve Maslach tükenmişlik ölçeği puanları arasında .45 ( $p < .001$ ) ilişki bulunmuştur. Ölçme aracının güvenilirliği Cronbach Alfa iç tutarlık katsayısı ve test-tekrar test teknikleri ile incelenmiştir. Bu test teknikleri ölçme aracının yüksek düzeyde güvenilir olduğuna kanıt sağlamıştır. Geçerliği kanıt sağlamak amacıyla ölçüt geçerliği analizine yer verilmiştir. Mevcut bulgular doğrultusunda, "Beden Eğitimi ve Spor Dersi Sınıf Değerlendirme Atmosferi (BESDSDAÖ)"nin geçerli ve güvenilir ölçümler yapabilen bir veri toplama aracı olduğu söylenebilir.

**Anahtar Kelimeler:** Sınıf değerlendirme atmosferi, Beden eğitimi ve spor dersi, Uyarlama, Geçerlik, Güvenirlik.

## Giriş

Çevresel etkilerin bireyler üzerindeki önemini anlamaya yönelik teorik çerçeve, tarihsel olarak Lewin'in formülüne dayanmaktadır. Lewin (1936), insan davranışının, bireyin ve çevresinin karmaşık etkileşimi tarafından belirlendiğine inanmıştır. Diğer bir ifadeyle Lewin, insan ve çevre arasındaki karşılıklı etkileşimlerin insan davranışının belirleyicisi olduğunu vurgulamıştır. Bu çevresel şartlar arasında yer alan ve öğrencilerin eğitimsel başarısı için kritik bir ortam olarak kabul edilen kavramlardan birisi sınıftır (Anderson ve Burns, 1989; Borich, 1988; Fraser ve Walberg, 1991; Walberg, 1968). Ayrıca sınıf ortamının bireyin davranışları üzerinde kritik bir rol oynadığına dair önemli bulgular tespit edilmiştir (Bartha, Dunlap, Dane, Lochman ve Wells, 2004). Haertel, Walberg ve Haertel (1981) tarafından sınıf ortamı çalışmalarının kullanıldığı bir meta-analiz araştırmasında, öğrencilerin sınıf ortamı algılarının motivasyon, başarı ve öğrenci memnuniyeti gibi öğrencilerin öğrenme çıktılarının belirli yönlerini belirlemede kritik bir faktör olduğunu ortaya koymuştur.

Öğrencilerin sınıf ortamına yönelik algılarının incelenmesi amacıyla gerçekleştirilen araştırmalar, öğrenme çıktılarına ilişkin varyansın önemli bir kısmının sınıf ortamına yönelik algılar ile açıklandığını ortaya koymuştur (Dorman, 2001). Dolayısıyla sınıf ortamında öğrencilerin bilişsel ve duyuşsal davranışlarına özgü gözlemlerin önem kazanması gerektiği düşünülmektedir. Sınıf ortamının önemli bileşenlerinden biri de, sınıf değerlendirme atmosferidir (Banks, 2012). Sınıftaki zamanının önemli bir kısmının değerlendirme ile ilgili etkinliklere ayrıldığı dikkate alındığında (Mertler, 2003), sınıftaki değerlendirme atmosferinin sınıf ortamının en önemli bileşenlerinden biri olduğu düşünülmektedir (Brookhart ve DeVoge, 1999).

Brookhart (1997) sınıf değerlendirme atmosferini; öğretmen tarafından belirlenen değerlendirme amaçlarının, değerlendirme görevlerinin, değerlendirmede kullanılan performans ölçütlerinin ve sunulan geri bildirimlerin sınıf çevresine ilişkin öğrencilerde oluşturduğu algı şeklinde ifade etmektedir (Alkharusi, Aldhafri, Alnabhani ve Alkalbani, 2014). Bir başka deyişle, öğrencilerin sınıftaki değerlendirme uygulamaları hakkında sahip oldukları algılar sınıf değerlendirme atmosferini oluşturmaktadır (Brookhart ve DeVoge, 1999). Öğrenciler sınıfta çeşitli değerlendirme faaliyetlerine katılım sağlamaktadırlar. Eğitimciler, sınıfta sunulan bu etkinliklerin öğrencilere orada vurgulananlar hakkında önemli mesajlar ilettiklerini ve bunun sonucunda farklı eğitimsel sonuçların ortaya çıkabileceğini düşünmektedirler (Ames, 1992b; Ames ve Archer, 1988; Linnenbrink ve Pintrich, 2001).

Ames (1992a), aşağıdaki sınıf değerlendirme uygulamalarının, öğrencilerdeki olumlu inanç, etki ve davranış modellerini ortaya çıkarmasının muhtemel olduğuna dikkat çekmiştir: (a) zorluk, çeşitlilik, yenilikçilik ve aktif katılımı içeren değerlendirme görevlerini tasarlamak; (b) öğrencilere değerlendirme sürecinde seçimler ve kararlar alma fırsatları vermek; (c) ilerlemeyi, gelişmeyi ve ustalığı değerlendiren ve sosyal karşılaştırmalardan kaçınan değerlendirme uygulamaları yapmak; ve (d) değerlendirme görevinde zamanın görevin doğasına ve öğrencinin ihtiyacına göre değişmesine izin vermek.

Bu eğitim perspektifleri ışığında, öğrencilerin sınıf değerlendirme atmosferine ilişkin algılarının motivasyonel inançlarını ve başarılarını etkilediği düşünülmektedir (Ames, 1992b; Brookhart, 1997). Tüm derslerde olduğu gibi öğrencilerin beden eğitimi ve spor derslerinde oluşturdukları sınıf değerlendirme atmosferi algılarının dersten elde edilecek kazanımları doğrudan etkilediği söylenebilir. Ulusal literatürde öğrencilerin sınıf değerlendirme atmosferine yönelik algılarını ölçmek amacıyla kullanılacak bir ölçek (İlhan ve Çetin, 2014) bulunmakta olup uluslararası literatürde de bir ölçeğe (Alkharusi, 2009; 2011) rastlanmıştır. Ancak doğrudan beden eğitimi ve spor dersine yönelik sınıf değerlendirme atmosferini belirleyebilecek ölçme aracına ulusal ve uluslararası literatürde rastlanılmamıştır. Bu araştırmanın amacı beden eğitimi ve spor dersine yönelik sınıf değerlendirme atmosferi ölçeğinin geçerliği ve güvenilirliğini belirlemektir.

## Yöntem

Bu çalışmada; karma araştırma yaklaşımı kullanılmıştır. Bu yaklaşım; ne, nerede ve ne kadar gibi sayısal verilerin yer aldığı, araştırmacının daha somut ve objektif bilgilere ulaştığı (Yıldırım ve Şimşek, 2011) nicel araştırma ve araştırmaya dahil olan kişilerin bakış açılarıyla olay, olgu, norm ve değerleri incelemeye çalışan (Ekiz, 2003) nitel araştırma yaklaşımlarının birlikte kullanıldığı, bütünleştirildiği ya da harmanlandığı bir yaklaşımdır (Johson ve Onwuegbuzie, 2004). Karma araştırma farklı desenlerden oluşmaktadır. Bunlardan biri de çalışmada faydalanılan “Keşfedici Sıralı Desen”dir. Bu desende önce nitel verilerin toplanması ve çözümlenmesine fırsat verilip, ardından nicel araştırmayı uygulamaya koyup, birincil sonuçların test edilmesi ve genellemesine fırsat verir (Creswell ve Clark, 2015). Ölçek geliştirme amacıyla yapılan çalışmada; ilk olarak görüşme yapılmış, katılımcıların özellik hakkında sahip oldukları duygu ve düşünceler tespit edilmiştir. Ardından elde edilen veriler içerik analizine tabi tutularak, madde havuzunun oluşturulmasında keşfedici sıralı desenden faydalanılmıştır.

## Çalışma Grubu

Araştırma Kırıkkale ilinde farklı liselerde (Anadolu Lisesi, Fen Lisesi ve Meslek Lisesi) öğrenim gören üç farklı çalışma grubu üzerinde yapılmıştır. Birinci çalışma grubunu 139 (%69,2)’u kadın ve 62(%30,8)’si erkek olmak üzere toplam 201 lise öğrencisi oluşturmuştur. Bu öğrencilerin 70’i (%34,8), 9. sınıf, 56’sı (%27,9) 10. sınıf, 43’ü (%21,4) 11. sınıf ve 32’si (%15,9) 12. sınıfta öğrenim görmektedir. Öğrencilerin 57’si (%28,4) Anadolu Lisesi, 46’sı (%22,9) Fen Lisesi ve 98’i (%48,8) Meslek Lisesi’nde okumaktadır. Bu çalışma grubundan elde edilen veriler üzerinde Açıklayıcı Faktör Analizi yapılmıştır. İkinci çalışma grubunu yine Kırıkkale’deki farklı liselerde öğrenim gören 123’ü (%58,57) kadın ve 87’si (%41,43) erkek olmak üzere toplam 210 lise öğrencisi oluşturmuştur. Araştırmadan elde edilen verilerin güvenilirliği kanıt sağlamak amacıyla elde edilen ölçme aracı üç hafta ara ile 65 (35’i kadın ve 30’u erkek) lise öğrencisine iki kez uygulanmış, böylece verilerin tutarlığı incelenmiştir. Ölçek geliştirme çalışmalarında AFA ve DFA analizlerinin farklı örneklem gruplarından elde edilen veriler üzerinden yapılması (Cabrera-Nguyen, 2010) ve farklı çalışma grupları üzerinde yürütülen analizlerin farklı sayıda katılımcı ile gerçekleştirilmesinin geçerliğe yönelik daha fazla kanıt sağlayabilmektedir.

## Madde Havuzunun Oluşturulma Süreci

Ölçek geliştirme çalışmalarında araştırmacı ölçmek istediği özelliği tanımlaması gerekmektedir (Cohen ve Swerdlik, 2010). Bu amaçla literatürde sınıf değerlendirme atmosferi özelliği ile ilgili yapılan çalışmalar (Alkharusi, 2007; 2011; Brookhart & DeVoge, 1999; Church, Elliot & Gable, 2001; İlhan & Çetin, 2014; Stiggins & Conklin, 1992) incelenmiştir. Alkharusi (2011) sınıf değerlendirme atmosferini ölçmek amacıyla geliştirdiği ölçme aracının dayanağını başarı hedefi teorisinden almıştır. Onun çalışmasında bu özellik işlevsel olarak öğrenme ve performans odaklı değerlendirme ortamı olarak kullanılmıştır. Ölçme aracının açtığı özelliğin ilk boyutu olan öğrenmeye dayalı değerlendirme ortamında; öğrencilerin değerlendirilmesine yönelik görev ve sorumluluklar orta düzeydedir. Bu ortamda değerlendirme kriterleri açık ve anlaşılır bir şekilde öğrencilere sunulmaktadır. Öğretmenler tarafından öğrencilere eksikliklerini görebilecekleri fırsatlar sunulmakta ve bu eksikleri nasıl giderebileceklerine dair dönütler sağlanmaktadır. Ayrıca öğrenme esnasında öğrencinin hata yapması normal kabul edilmekte ve öğrencilere hatalarını düzeltme şansı verilmektedir. Ölçme aracını oluşturan sınıf değerlendirme atmosferinin ikinci boyutu olan performans dayalı öğrenme ortamında ise yukarıda bahsedilen değerlendirme kriterlerinin aksine öğrencilerin değerlendirme ile ilgili görev ve sorumluluklarının zor olduğu değerlendirme kriterleri mevcuttur. Değerlendirme sürecinde öğrencinin kazanımlar elde etmesinden ziyade öğrencinin aldığı notun önemli görüldüğü, öğrencilerin bireysel başarı ve performansının değerlendirilmesi yerine birbirleriyle performanslarının karşılaştırıldığı değerlendirilme ortamı sunulmaktadır. Beden eğitimi dersi sınıf değerlendirme atmosferi ölçeği madde havuzu

oluşturulurken, Alkharusi (2011)'in çalışmasını temellendirdiği yukarıda bahsedilen iki boyuttan faydalanılmıştır (Tablo 1).

**Tablo 1:** Sınıf Değerlendirme Atmosferi Ölçeğinin Alt Boyutları ve Göstergeleri

Kategoriler	Göstergeler	İşlevsel Tanımlar*	Maddele r
Öğrenmeye Dayalı Ortam	Öğrenme sürecinde öğrencinin gösterdiği çaba ve sorumluluğun ön planda tutulduğu ortam (Süreç Odaklı)	✓ Üründen çok öğrenmeye ilişkin çabaların yer alması	M <sub>1</sub> ,M <sub>2</sub> ,M <sub>3</sub> ,M <sub>4</sub> ,M <sub>5</sub> ,M <sub>6</sub> ,M <sub>7</sub> ,M <sub>8</sub> ,M <sub>9</sub> ,M <sub>10</sub> ,M <sub>11</sub> ,M <sub>12</sub> ,M <sub>13</sub> ,M <sub>14</sub> ,M <sub>15</sub> ,M <sub>16</sub> ,M <sub>17</sub> ,M <sub>18</sub>
		✓ Öğrencilere sunulan görev ve sorumlulukların orta seviyede olması	
Performansa Dayalı Ortam	Değerlendirmede sonucun ön planda tutulduğu ortam (Ürün Odaklı)	✓ Bireysel gelişim ve öğrenmelerin ön planda olması	M <sub>19</sub> ,M <sub>20</sub> ,M <sub>21</sub> ,M <sub>22</sub> ,M <sub>23</sub> ,M <sub>24</sub> ,M <sub>25</sub>
		✓ Öğrencinin gelişimine fırsat sağlayan olanakların sunulması	
		✓ Öğrenmede geri bildirim sağlanması	
		✓ Hatanın öğrenme sürecinin doğal bir parçası olması	
		✓ Öğrenciye değerlendirme süreci ile ilgili tercih yapma ve karar alma fırsatının sunulması	
		✓ Değerlendirmede kullanılacak ölçütlerin açık ve anlaşılır olarak öğrencilere sunulması	
		✓ Değerlendirmede süreçten ziyade sonucun (not) dikkate alınması	
		✓ Gösterilen çabadan ziyade ortaya konulan performansın anlamlı olması	
		✓ Bireysel değerlendirme (mutlak) yerine sosyal karşılaştırılmanın (bağıl) yapılması	

\*Alkharusi, 2009, 2011; Amess, 1992; McMillan & Workman, 1998; Stiggins & Conklin, 1992

Tablo Sınıf değerlendirme atmosferini konu edinen çalışmaların incelenmesi neticesinde, bu özelliği açıklayan boyut ve göstergeler tespit edilmiştir (Tablo 1). Ardından sınıf değerlendirme atmosferinin derslerde nasıl olduğunu tespit etmek amacıyla sorular hazırlanmış, deneyimli beden eğitimi öğretmeni (n=2), sınıf yönetimi alanında uzman öğretim üyesi (n=1) ve türkçe eğitimi alanından uzmana (n=1) hazırlanan sorular sunulmuştur. Onlardan gelen dönütler doğrultusunda sorulara son şekli verilmiştir. Çalışma grubu ile benzer özelliklere sahip olan 10 öğrenciye bu sorular yöneltilerek, beden eğitimi öğretmenin derslerde ve sınavlarda değerlendirmeye ilişkin nasıl uygulamalar yaptıkları tespit edilmeye çalışılmıştır. Elde edilen görüşme kayıtları yazı aktarılmış ve sınıf değerlendirme atmosferi ile ilgili olan ifadeler madde havuzuna eklenmiştir. Literatür taraması ve görüşmeler neticesinde 28 maddeden oluşan deneme formu hazırlanmıştır. Deneme formu pilot çalışmaya konulmadan önce uzman görüşüne sunulmuştur. Bu formda yer alan her bir maddenin uzmanlar tarafından “uygun”, “uygun değil” ve “geliştirilmesi gerekir” şeklinde değerlendirilmesi istenmiştir. Uzmanlar tarafından kapsama uygun olmadığı düşünülen üç madde ölçme aracından çıkartılmıştır. Gerekli düzeltmelerin yapılmasının ardından, maddelerin dil ve anlatım bakımından uygunluğunu sağlamak amacıyla, iki Türk Dili uzmanından görüş alınmıştır.

Maddelerin açıklığı ve anlaşılabilirliği, herkes tarafından aynı anlamı ifade etmesi ve uygulama süresinin belirlenmesi amacıyla farklı lise türlerinden tesadüfen seçilen 20 öğrenciye 25 maddelik deneme formu uygulanmıştır. Onlardan gelen dönütler neticesinde anlaşılır olmayan iki madde düzeltilmiş ve deneme formu uygulamaya hazır hale getirilmiştir. Bu deneme formu 25 maddeden oluşmuştur. Maddelerin 18'i olumlu ve 7'si olumsuzdur.

### Verilerin Toplanması

Verilerin toplanmasında amaçlı örnekleme yöntemlerinden ölçüt örnekleme tekniğinden faydalanılmıştır. Buradaki amaç, farklı akademik başarı seviyesine sahip olan öğrencilerin beden eğitimi dersinde yapılan değerlendirmelere yönelik görüşlerinin ölçme aracının oluşturulmasında geçerliliği arttıracak olmasıdır. Gruplar arasındaki heterojenliğin sağlanması elde edilen yapıya farklı özellikleri temsil gücünü arttıracaktır. Gerekli izinler alınarak, okullara gidilmiş, araştırmanın amacı öğrencilere belirtilmiş, gönüllü katılım sağlamak isteyenlere deneme formu uygulanmıştır.

Ölçme aracının yapı geçerliliği AFA ve DFA yanısıra ölçüt bağımlı geçerlik ile de sağlanmıştır. Ölçüt bağımlı geçerlik kapsamında sınıf değerlendirme atmosferi ile başarı yönelimi ve tükenmişlik durumu arasında anlamlı ilişki olduğunu ortaya koyan çalışmalara rastlanmıştır. Bu doğrultuda BEDSDAÖ, Başarı Yönelimi Ölçeği ve Maslach Tükenmişlik Ölçeği arasındaki ilişki incelenmiştir. Ölçüt bağımlı geçerliliği test etmek amacıyla çalışmada kullanılan ölçme araçlarının özellikleri aşağıda sunulmuştur.

2x2 Başarı Yönelimleri Ölçeği: Elliot ve Murayama (2008) tarafından geliştirilen ölçeğin Türkçe uyarlaması Aslan ve Akın (2015) tarafından yapılmıştır. Ölçme aracı 4 faktör (öğrenme-yaklaşma yönelimi, öğrenme-kaçınma yönelimi, performans-yaklaşma yönelimi ve performans-kaçınma yönelimi) ve 12 maddeden oluşmaktadır. Ölçeğin güvenilirlik katsayısı öğrenme-yaklaşma yönelimi alt boyutu için .72, öğrenme kaçınma yönelimi boyutu için .68, performans-yaklaşma yönelimi boyutu için .62 ve performans-kaçınma yönelimi alt boyutu için .69 olarak bulunmuştur.

Mevcut çalışma kapsamında ölçeğin geçerlik ve güvenilirlik analizleri sınanmıştır. Elde edilen güvenilirlik değerleri ölçeğin öğrenme-yaklaşma yönelimi alt boyutu için .76, öğrenme kaçınma yönelimi boyutu için .68, performans-yaklaşma yönelimi boyutu için .73 ve performans-kaçınma yönelimi alt boyutu için .70 olarak bulunmuştur. Yapı geçerliliğini test etmek için yapılan Doğrulayıcı Faktör Analiz sonucunda elde edilen uyum indeksleri  $\chi^2=1.23$ , RMSEA= .07, NNFI=.95, NFI=.96, CFI= .95, IFI= .95, RFI= .93, GFI=.88, olduğu tespit edilmiştir. Elde edilen uyum indeksleri ölçme aracının mevcut çalışma grubu üzerinde de dört boyutlu yapısını koruduğunu göstermektedir.

Maslach Tükenmişlik Envanteri Öğrenci Formu: Schaufeli, Martinez, Marques-Pinto, Salanova ve Bakker (2002a) tarafından geliştirilen Maslach Tükenmişlik Envanteri-Öğrenci Formu'nun (MTE-ÖF) Türkçe uyarlaması Çapri, Gündüz ve Gökbakan (2011) tarafından yapılmıştır. Ölçme aracı; "Tükenme", "Duyarsızlaşma" ve "Yetkinlik" olmak üzere üç alt boyut ve 13 maddeden oluşmaktadır. MTE-ÖF'nin Cronbach Alfa iç tutarlık katsayısı birinci boyut için 0.76, ikinci boyut için 0.82, üçüncü boyut için ise 0.61 olarak hesaplanmıştır. Ayrıca, ölçeğin test-tekrar test güvenilirlik sonuçları sırasıyla 0.76, 0.74 ve 0.70 olarak bulunmuştur.

Mevcut çalışma kapsamında ölçeğin geçerlik ve güvenilirlik analizleri sınanmıştır. Elde edilen güvenilirlik değerleri ölçeğin "Tükenmişlik" boyutu için 0.79, "Duyarsızlaşma" boyutu için 0.75, ve "Yetkinlik" boyutu için 0.72 olarak bulunmuştur. Yapı geçerliliğini test etmek için yapılan Doğrulayıcı Faktör Analiz sonucunda elde edilen uyum indeksleri  $\chi^2=1.35$ , RMSEA= .06, NNFI=.97, NFI=.95, CFI= .98, IFI= .97, RFI= .95, GFI= .89, olduğu tespit edilmiştir. Elde edilen uyum indeksleri ölçme aracının mevcut çalışma grubu üzerinde de üç boyutlu yapısını koruduğunu göstermektedir.

### Verilerin Analizi

Veriler SPSS 22 ve Lisrel 8.8 paket programlarına girilmiş ve gerekli işlemler yapılmıştır. Araştırmada öncelikle yapı geçerliğine ilişkin analizler gerçekleştirilmiştir. Yapı geçerliği kapsamında ilk olarak, AFA uygulanmış; AFA sonucunda ortaya çıkan yapının farklı bir çalışma grubu üzerinden toplanan verilerle uyumunu değerlendirmek amacıyla DFA' dan yararlanılmıştır. Madde toplam test korelasyonu ve % 27'lik alt-üst grup karşılaştırmaları değerlendirilmiştir. Faktör analizi ile tespit edilen yapının güvenilirliği, Cronbach Alfa iç tutarlık katsayısı ve test-tekrar test güvenilirlik yöntemleri ile incelenmiştir.

### Bulgular

#### Yapı Geçerliği

Ölçme aracının ölçmeye çalıştığı teorik yapıyı ölçebilme derecesi olarak tanımlanan yapı geçerliği, ölçekte yer alan her bir maddenin birbiri ile olan ilişkisini ortaya koyar (Seçer, 2015). Hazırlanan bu ölçekteki maddelerin ölçülmek istenen yapıyı ne derece ölçtüğü sorunu, yapı geçerliği ile ilgilidir. BEDSDAÖ' den elde edilen ölçümlerin yapı geçerliğini test etmek amacıyla AFA ve DFA uygulanmıştır.

#### Açımlayıcı Faktör Analizi (AFA)

Çalışma grubundan elde edilen verilerin faktör analizi için uygun olup olmadığı Kaiser-Mayer-Olkin (KMO) katsayısı ve Bartlett testi ile açıklanabilir (Büyüköztürk, 2014). Bu durumda, Bartlett testi sonucunun anlamlı çıkması ve KMO değerinin 0.50'den büyük çıkması beklenmektedir. KMO'nun 0.60'dan yüksek, Bartlett testinin de anlamlı çıkması verilerin faktör analizi için uygun olduğunu göstermekte (Tavşancıl, 2014; Çokluk ve ark., 2012) olup, verilerin çok değişkenli normal dağılımdan geldiğini ve değişkenler arasında faktörleştirme tekniklerinin kullanılabileceğini ortaya koymaktadır (De Vellis, 2014; Fraenkel ve Wallen, 2000). Bu araştırmada, KMO örneklem uygunluk değeri 0.85 ve Bartlett-Sphericity testi ki kare değeri 1000,743 ( $p < 0.01$ ) olarak bulunmuştur. Bu değerler çalışma grubundan alınan verilerin faktör analizine uygun olduğunu göstermektedir.

Araştırmada "Temel Bileşenler Analizi" kullanılmıştır. Uygulanan Temel Bileşenler Analizi'nde faktör sayısına karar verilirken her bir faktörün toplam varyansa yaptığı katkı ile Kaiser-Guttman ilkesi uyarınca özdeğeri 1'den büyük olan örtük değişkenler dikkate alınmıştır (Murphy ve Davidshofer, 1991). Tabachnick ve Fidell (2007), ölçek faktörlerinin birbiriyle ilişkili olduğu durumlarda gerçekleştirilecek eksen döndürme işlemlerinde direct oblimin döndürme yaklaşımının tercih edilebileceğini belirtmektedir. Mevcut araştırmada, ölçülmesi hedeflenen özelliğin alt boyutlarının ilişkili olduğu ön görüldüğünden dolayı, varimax dik döndürme tekniği kullanılmıştır.

Bir çok araştırmacıya (De Vellis, 2014; Büyüköztürk ve ark., 2012; Tavşancıl, 2014) göre, maddelerin faktör yük değerinin 0.30 değerinden yüksek olması gerekmektedir. Bu çalışmada madde faktör yük değeri 0.35 olarak belirlenmiştir. Faktör analizinde aynı yapıyı ölçmeyen maddelerin çıkarılması sürecinde, faktör yükü 0.35 değerinin altında olan ve bir maddenin iki faktör için yüksek yük değeri verdiğinde, yük değerleri arasındaki farkın binişiklik yaratmaması için en az 0.10 olmasına dikkat edilmiştir (Büyüköztürk, 2014). Bu ölçütleri karşılayamayan dokuz madde (2-4-9-10-17-19) ölçme aracından çıkarılmıştır. AFA sonucunda elde edilen faktörler, faktörleri temsil eden maddeler, madde faktör yük değerleri ve ortak varyans değerleri Tablo 2'de verilmiştir.



**Tablo 2: BESDSDAÖ Faktör Yükleri ve Ortak Varyans Değerleri**

Boyut	Maddeler	Ortak Varyans	ÖOD	POD	
Öğrenme Odaklı Değerlendirme	1-Beden eğitimi ve spor dersinde uygulanan değerlendirme etkinlikleri, öğrencilerin güçlü ve zayıf yönlerini görebilmelerini sağlar.	,486	,550	,313	
	3-Beden eğitimi ve spor dersinde öğrencilere performansları ile ilgili eksiklerini tamamlamalarını sağlayacak geri bildirimler (dönüt) verilir.	,429	,508	,268	
	5-Beden eğitimi ve spor dersinde uygulanan değerlendirme etkinlikleri, öğrencilere sağlıklı yaşam biçimi davranışları kazandırmaya yöneliktir.	,461	,466	,315	
	6-Beden eğitimi ve spor dersinde uygulanan değerlendirme sürecinde öğretmen diğer öğrencilerin değerlendirmelerini de dikkate alır.	,392	,490	,227	
	7-Beden eğitimi ve spor dersinin değerlendirme sürecinde öğretmen hareketin esaslarını bozmadan öğrencilerin beceriyi kendilerine özgü (kendinden bir şey katmak) şekilde sunmalarına fırsat verir.	,490	,512	,389	
	8-Beden eğitimi ve spor dersinde öğrencilerin değerlendirilmesinde öğrencilerin kişisel etkinliği dikkate alınır.	,610	,652	,292	
	11-Beden eğitimi ve spor dersinde değerlendirme aşamasında öğrencilerin okul içi ve dışı etkinliklere katılımı dikkate alınır.	,466	,516	,201	
	12-Beden eğitimi ve spor dersinde değerlendirme sürecinde başarısız öğrencilere yeni deneme fırsatları (farklı zamanda performansı değerlendirmek) verilir.	,448	,622	,212	
	13-Beden eğitimi ve spor dersindeki sınav sonuçlarında öğrencilerin gelişimine yönelik bilgi verilir.	,481	,614	,235	
	14-Beden eğitimi ve spor dersinde uygulanan değerlendirme etkinlikleri öğrencileri yeni beceriler kazandırmaya teşvik eder.	,492	,540	,318	
	15-Beden eğitimi ve spor dersinde uygulanan değerlendirme etkinliklerinde değerlendirme öncesinde öğrencilerin sınava hazır olup olmadığı (ruhsal, fizyolojik, psikolojik vb.) dikkate alınır.	,395	,542	,107	
	16- Beden eğitimi ve spor dersinde önceden belirlenen hedeflere ilişkin değerlendirmeler yapılır.	,396	,544	,213	
	18- Beden eğitimi ve spor dersinde değerlendirme aşamasında öğrenci seviyesine uygun materyal (fiziksel özellikler, kız ve erkeklerde farklı büyüklük, ağırlık vb) kullanılır.	,461	,596	,132	
	<b>Açıklanan Varyans</b>		<b>%21,224</b>		
	Performans Odaklı Değerlendirme	*20-Beden eğitimi ve spor dersinde öğrencilerin beceriye ilişkin çabasından ziyade, aldığı not önemli görülür.	,610	,107	,706
		*21- Beden eğitimi ve spor dersinde ders konularıyla değerlendirilecek beceri tutarlı/ilişkili değildir.	,427	,103	,567
		*22-Beden eğitimi ve spor dersinde uygulanan puanlama sistemi açık değildir.	,512	,170	,619
		*23-Beden eğitimi ve spor dersinde değerlendirme sürecinde bireysel performanstan ziyade, diğer öğrencilerin performansları dikkate alınmaktadır.	,521	,105	,640
*24-Beden eğitimi ve spor dersinde öğrencinin sınıftaki diğer öğrencilerden düşük performans göstermesi eleştirilmesine neden olur.		,594	,131	,691	
*25- Beden eğitimi ve spor dersinde uygulanan değerlendirme etkinlikleri sonuç (aldığı not) odaklıdır.		,471	,146	,592	
<b>Açıklanan Varyans</b>		<b>%16,832</b>			
<b>Açıklanan Toplam Varyans</b>		<b>%38,056</b>			
Faktörler	Ort	s.s	a	Korelasyon ÖOD	POD
Öğrenme Odaklı Değerlendirme (ÖOD)	3,59	0,64	0,81	1	
Performans Odaklı Değerlendirme (POD)	3,17	0,81	0,71	-,040	1
Beden Eğitimi Dersi Sınıf Değerlendirme Atmosferi	3,38	0,50	0,74	,599**	,776**

\* İşaretili maddeler ters puanlanacaktır.

Tablo 2’deki bulgular dikkate alındığında; ölçme aracının 19 madde ve iki boyuttan meydana geldiği görülmektedir. “Öğrenme Odaklı Değerlendirme” boyutu 13 maddeden (1-3-5-6-7-8-11-12-13-14-15-16-18) oluşmakta ve toplam varyansın % 21,224’ünü açıklamaktadır. Bu boyutta yer alan maddelerin faktör yükleri 0.46 ile 0.65 arasında değişmekte olup, özdeğeri 5,503’tür. “Performans Odaklı Değerlendirme” boyutu dört maddeden (20-21-22-23-24-25) oluşmakta ve toplam varyansın % 16,832’ sini açıklamaktadır. Bu boyutta yer alan maddelerin madde faktör yükleri 0.56 ile 0.70 arasında değişmekte ve özdeğeri 1,727’ dir. Her iki boyut toplam özelliğinin %38,056’sını açıklamaktadır. Ölçme aracı 5’li likert tarzında (Kesinlikle Katılmıyorum-Kesinlikle Katılıyorum) uygulanmaktadır. Ölçme aracında toplam puan elde edilmektedir. Performans Odaklı Değerlendirme boyutundaki maddeler ters puanlanmaktadır.

### Doğrulayıcı Faktör Analizi (DFA)

AFA sonrasında ortaya çıkan modelin, yapı geçerliliğini değerlendirmek için DFA yapılmıştır (Kline, 2011). DFA’ da sınanan modelin yeterliliğini ortaya koymak üzere Tablo 3’de sunulan değerler hesaplanmıştır. İlgili literatür doğrultusunda uyum indekslerine ilişkin kabul edilebilir ve iyi varsayılan değer aralıkları ve araştırmada elde edilen değerler Tablo 3’ de verilmiştir.

**Tablo 3:** Doğrulayıcı Faktör Analizine İlişkin Elde Edilen Sonuçlar

Model Uyum İndeks	Mükemmel Aralık	Kabul Edilebilir Aralık	*BEDSDAÖ
$X^2/sd$	$0 < X^2/sd < 2$	$2 < X^2/sd < 3$	2.13
RMSEA	$0.00 < RMSEA < 0.05$	$0.05 < RMSEA < 0.10$	0.06
PGFI	$0.95 < PGFI < 1.00$	$0.50 < PGFI < 0.95$	0.72
PNFI	$0.95 < PNFI < 1.00$	$0.50 < PNFI < 0.95$	0.81
GFI	$0.90 < GFI < 1.00$	$0.85 < GFI < 0.90$	0.90
AGFI	$0.90 < AGFI < 1.00$	$0.85 < AGFI < 0.90$	0.87
IFI	$0.95 < IFI < 1.00$	$0.90 < IFI < 0.95$	0.95
NFI	$0.95 < NFI < 1.00$	$0.90 < NFI < 0.95$	0.91
CFI	$0.95 < CFI < 1.00$	$0.90 < CFI < 0.95$	0.95

Tabachnick ve Fidell, 2007; Jöreskog ve Sörbom, 1993; Schumacher ve Lomax, 1996; Hooper, Coughlan ve Mullen, 2008. \*Beden Eğitimi Dersi Sınıf Değerlendirme Atmosferi Ölçeği

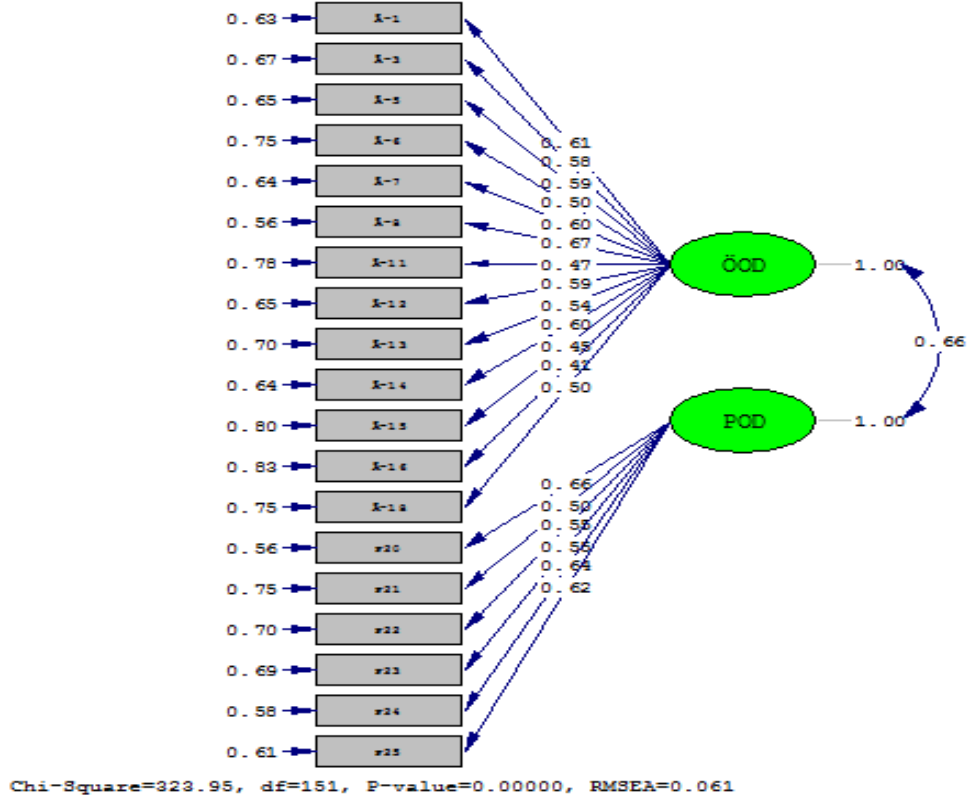
Tablo 3’ de ölçme aracına ilişkin uyum indeksleri;  $x^2/sd=2.15$ ,  $RMSEA=0.06$ ,  $PGFI=0.72$ ,  $GFI=0.90$ ,  $AGFI=0.87$ ,  $PNFI=0.81$ ,  $IFI=0.95$ ,  $NFI=0.91$  ve  $CFI=0.95$  olarak bulunmuştur. Bu uyum indeksleri dikkate alındığında, iki boyutlu teorik yapının toplanan verilerle doğrulandığı ifade edilebilir.

**Tablo 4:** Doğrulayıcı Faktör Analizi Sonucu Elde Edilen t Değerleri

Madde	t	Madde	t
1	11.04**	20	11.67**
3	10.33**	21	8.36**
5	10.58**	22	9.28**
6	8.82**	23	9.39**
7	10.86**	24	11.28**
8	12.38**	25	10.83**
11	8.21**		
12	10.58**		
13	9.63**		
14	10.76**		
15	7.66**		
16	7.00**		
18	8.66**		

Tablo 4 incelendiğinde, DFA sonucu elde edilen iki boyutlu modele ilişkin t-testi değerlerinin “Öğrenme Odaklı Değerlendirme” boyutunda 7.00 ile 12.38 arasında ve “Performans

Odaklı Öğrenme” boyutunda 8.36 ile 11.67 arasında olduğu görülmektedir. Literatürde, t değerlerinin 1.96’dan büyük olması .05 düzeyinde; 2.58’den büyük olması ise .001 düzeyinde anlamlı olduğunu göstermektedir (Kline, 2011). Buna göre, DFA’da elde edilen tüm t değerlerinin .01 düzeyinde anlamlı olduğu tespit edilmiştir. Anlamlı olmayan t değerleri, söz konusu t değerlerine ilişkin maddelerin modelden çıkarılması gerektiğine işaret etmekte veya araştırmadaki katılımcı sayısının faktör analizi için yetersiz olduğunun göstergesi olarak değerlendirilmektedir (Byrne, 2013). Dolayısıyla, DFA sonucunda elde edilen t değerleri, ortaya konan yapının kuramsal tanımlamalara uygun olduğunu ve herhangi bir maddenin analizden çıkarılmasına gerek olmadığını göstermiştir.



Şekil 1: Madde ve faktör ilişkisi

Şekil 1, gözlenen değişkenler ile gizil değişkenler arasındaki regresif bağı “Öğrenme Odaklı Değerlendirme” boyutunda 0.48 ile 0.73 ve “Performans Odaklı Öğrenme” boyutunda ise 0.46 ile 0.64 arasında olduğunu göstermektedir. Kline (2007) .30 civarındaki standartlaştırılmış ağırlıkların orta; .50 ve üzerindeki ağırlıkların ise büyük etki büyüklüklerini yansıttığını belirtmektedir. Bu doğrultuda, DFA sonucunda elde edilen madde faktör yüklerinin büyük etki büyüklüklerini yansıttığı söylenebilir.

### Madde Analizleri

Madde analizi kapsamında madde toplam test korelasyonu ve % 27 alt-üst grupları karşılaştırılarak incelenmiştir. Ölçekte yer alan her bir maddenin, ölçmek istediği özelliği ölçüp ölçmediği ve ölçtükleri özellik açısından kişileri ayırt etmede ne kadar yeterli olduklarının belirlenmesi amacıyla, ilk olarak madde-toplam test korelasyonu hesaplanmıştır (Büyüköztürk ve ark., 2012). Eğer bir maddenin toplam puanla olan korelasyonu düşük ise, bu durum o maddenin

testteki diğer maddelerden farklı bir niteliği ölçtüğünü gösterir (Karasar, 2014). Uzmanlar madde-toplam test korelasyonu 0.30 ve daha yüksek olan maddelerin iyi derecede ayırt edicilik özelliğe sahip olduğunu belirtmektedir (Büyüköztürk, 2014; Erkuş, 2014). Madde toplam test korelasyonu, test maddelerinden alınan puanlar ile testin toplam puanı arasındaki ilişkiyi açıklamaktadır. Madde toplam test korelasyon değerleri 0.36 ile 0.69 arasında yer almaktadır.

**Tablo 5:** Beden Eğitimi Dersi Sınıf Değerlendirme Atmosferi Ölçeğinin Madde Analizleri

Boyutlar	No	% 27 Üst Grup n=54		% 27 Alt Grup n=54		t	p	Madde Toplam Test Korelasyonu	
		X	Ss	x	Ss				
Öğrenme Odaklı Değerlendirme	1	4,46	0,69	3,29	1,17	6,28	0,00	,533	
	3	4,16	1,04	2,98	1,23	5,38	0,00	,499	
	5	4,35	0,85	3,33	1,19	5,09	0,00	,461	
	6	4,05	0,97	2,70	1,22	6,34	0,00	,463	
	7	4,22	1,04	2,68	1,09	7,47	0,00	,593	
	8	4,48	0,74	2,68	1,21	9,28	0,00	,699	
	11	3,88	1,31	2,81	1,27	4,31	0,00	,410	
	12	4,27	0,94	2,37	0,95	10,44	0,00	,605	
	13	3,85	1,26	2,85	1,21	4,18	0,00	,424	
	14	4,46	0,71	3,22	1,17	6,61	0,00	,537	
	15	3,96	1,24	2,53	1,20	6,04	0,00	,441	
	16	3,83	1,27	2,75	1,04	4,79	0,00	,471	
	18	4,31	1,00	2,85	1,49	5,96	0,00	,544	
	Performans Odaklı Değerlendirme	20	3,94	1,18	2,61	1,30	5,54	0,00	,366
		21	3,64	1,31	3,09	1,85	2,30	0,00	,403
		22	3,72	1,07	2,87	1,22	3,83	0,00	,387
		23	3,92	1,09	2,85	1,25	4,74	0,00	,411
		24	4,18	1,04	2,68	1,38	6,34	0,00	,490
25		3,24	1,45	2,72	1,12	2,05	0,00	,485	

Tablo 5’de toplam puana göre üst % 27 ve alt % 27’lik grupların madde puanları arasındaki farkın anlamlılığı için t-testi (Kılıç Çakmak, Çebi ve Kan, 2014) incelendiğinde, t değerlerinin 2,05 ile 10,44 arasında değiştiği görülmektedir. Alt ve üst grup arasındaki farklara ilişkin t değerlerinin anlamlı olması maddenin ayırt ediciliği için bir kanıt olarak değerlendirilmektedir (Erkuş, 2014). Üst ve alt grup ayırt ediciliğini tespit etmek amacıyla yapılan analizde tüm maddelerin ayırt edici olduğu belirlenmiştir.

### Ölçüt Geçerliliği

Beden eğitimi dersi sınıf değerlendirme atmosferi ölçeğinin ölçüt geçerliliğini test etmek için Aslan ve Akın (2015) tarafından Türkçe’ye uyarlanan “2x2 Başarı Yönelimleri Ölçeği” ve Çapri, Gündüz ve Gökbakan (2011) tarafından Türkçe’ye uyarlaması yapılan “Maslach Tükenmişlik Envanteri Öğrenci Formu” kullanılmıştır. Analiz sonucunda beden eğitimi dersi sınıf değerlendirme atmosferi ölçeğinden alınan puanlarla başarı yönelimi .43 ( $p < .001$ ) ve Maslach tükenmişlik ölçeği puanları arasında .45 ( $p < .001$ ) ilişki bulunmuştur. Elde edilen sonuca göre beden eğitimi dersi sınıf değerlendirme atmosferi ölçeği, başarı yönelimi ve Maslach tükenmişlik öğrenci formu ölçeği ile orta düzeyde anlamlı bir ilişkisinin olduğu, ölçüt geçerliliğinin gerekliliğini yerine getirdiği söylenebilir.

**Tablo 6:** BEDSDAÖ, Başarı Yönelimi ve Maslak Tükenmişlik Ölçeği Arasındaki İlişki

	Ort/ s.s	Başarı Yönelimi	Tükenmişlik
Öğrenme Odaklı Değerlendirme (ÖOD)	3,59±0,64	,34**	-,39**
Performans Odaklı Değerlendirme (POD)	3,17± 0,81	-,04*	,24*
Beden Eğitimi Dersi Sınıf Değerlendirme Atmosferi	3,38±0,50	,43**	-,45**

Ölçüt bağıntılı geçerlik için “Başarı yönelimi” ve “Maslak tükenmişlik öğrenci boyutu ölçeği” puanları ile “Beden Eğitimi Dersi Sınıf Değerlendirme Atmosferi” alt ölçekleri arasındaki ilişkiyi tespit etmek amacı ile Pearson korelasyon analizi kullanılmıştır. ÖOD boyutu ile başarı yönelimi ( $r=.34$ ,  $p<0.01$ ) arasında, BEDSDAÖ ile başarı yönelimi ( $r=.43$ ,  $p<0.01$ ), arasında pozitif yönlü anlamlı ilişki tespit ediliken, ÖOD ile tükenmişlik ( $r=.44$ ,  $p<0.01$ ) ve BEDSDAÖ ile tükenmişlik ( $r=-.45$ ,  $p<0.01$ ) arasında negative yönlü ilişki saptanmıştır. Ölçeğin bir diğer boyutu olan PDO ile tükenmişlik ( $r=.24$ ,  $p<0.05$ ) arasında pozitif yönlü ilişki olduğu, ancak başarı yönelimi ( $r=-.04$ ,  $p<0.05$ ) ile arasında negative yönlü bir ilişki olduğu tespit edilmiştir.

### Güvenirlilik Analizleri

Güvenirlilik analizleri sonucunda, ölçme aracının Cronbach Alfa güvenirlilik katsayısı, ölçeğin geneli için 0.74 olarak hesaplanmıştır (Tablo 2). Ölçeğin alt boyutları için bu değer sırasıyla 0.81 ve 0.71 olarak hesaplanmıştır. Ayrıca, ölçme aracının güvenirliliğine kanıt sağlamak amacıyla test-tekrar test tekniği uygulanmıştır. Ölçümlerin test-tekrar test güvenirliliği için 77 (42 kadın, 35 erkek) öğrenciden toplanan veriler üzerinden üç hafta arayla iki uygulama yapılmıştır. Ölçme aracındaki maddelere verilen yanıtlar arasındaki tutarlılığın sınıandığı bu aşamada, puanlar arasındaki korelasyon katsayıları hesaplanmıştır. Ölçeğin genelinde hesaplanan test-tekrar test puanları arasındaki korelasyon katsayısı 0.74 olarak bulunmuşken, bu değer alt boyutlar için sırasıyla 0.77 ve 0.72 olarak tespit edilmiştir. Test-tekrar test tekniğine yönelik elde edilen bulgular incelendiğinde, ölçme aracının tutarlı ölçümler sağladığı şeklinde ifade edilebilir.

### Tartışma ve Sonuç

Öğrencilerin beden eğitimi ve spor dersi sınıf değerlendirme atmosferine yönelik görüşlerini ölçebilecek geçerli ve güvenilir bir ölçme aracı geliştirmek mevcut çalışmada amaçlanmıştır. Bu amaç doğrultusunda ilk olarak sınıf değerlendirme atmosferinin literatürde nasıl tanımlandığı incelenmiş ve özelliğın kuramsal yapısı ortaya konulmuştur. Alkharusi (2011) sınıf değerlendirme atmosferini başarı hedefi teorisine temelinde, öğrenme odaklı değerlendirme ve performans odaklı değerlendirme olmak üzere iki boyutta incelenmiştir. Bu teori ve sınıf değerlendirme atmosferi ile ilgili yapılan çalışmalar (Alkharusi, 2009, 2011; Amess, 1992; McMillan & Workman, 1998; Stiggins & Conklin, 1992) doğrultusunda ölçme aracının boyutları tespit edilmiştir. Bu boyutlar öğrenme odaklı performans ve performans odaklı değerlendirme olarak sınıflandırılmıştır. Bu yapı ve uzman görüşleri doğrultusunda 28 maddeden oluşan madde havuzu oluşturulmuş ve deneme formu hazırlanmıştır.

Ölçme aracının yapı geçerliğini sınamak amacıyla ilk olarak AFA yapılmış, ortaya koyan iki boyutlu yapı DFA ile doğrulanmıştır. Araştırmanın birinci grubundan elde edilen verilerin faktör analizi için uygunluğu Kaiser-mayer-Olkin (KMO) ve Barlett testi ile incelenebilir (Büyüköztürk, 2014). Bu çalışmada, KMO örneklem uygunluk değeri 0.85 ve Barlett-Sphericity testi ki kare değeri 1000,743 ( $p<0.01$ ) olarak bulunmuştur. Bu değerler çalışma grubundan alınan verilerin faktör analizine uygun olduğunu göstermektedir. Ölçme aracının yapısını ortaya koymak amacıyla yapılan AFA neticesinde özelliğe ilişkin açıklanan varyansın %38,056’sını açıkladığı, ölçme aracının 19 madde ve iki boyuttan oluştuğu saptanmıştır. Ölçeğin birinci boyutu olan “Öğrenme Odaklı Değerlendirme” 13 maddeden (1-3-5-6-7-8-11-12-13-14-15-16-18) oluşmakta ve toplam varyansın % 21,224’ünü açıklamaktadır. Bu boyutta yer alan maddelerin faktör yükleri 0.46 ile 0.65 arasında

değişmekte olup, özdeğeri 5,503'tür. İkinci boyut olan "Performans Odaklı Değerlendirme" dört maddeden (20-21-22-23-24-25) oluşmakta olup, toplam varyansın % 16,832' sini açıklamaktadır. Bu boyuttaki maddelerin madde faktör yükleri 0.56 ile 0.70 arasında değişmekte ve özdeğeri 1,727' dir. Faktörlerin isimleri belirlenirken Alkharusi (2009 ve 2011)'nin sınıf değerlendirme atmosferi yapısından faydalanılmıştır.

Özelliğe (Sınıf Değerlendirme Atmosferi) ilişkin kuramsal yapı ve ilgili çalışmalar doğrultusunda oluşturulan maddelerin yapı geçerliğinin ilk aşaması olan AFA sonucunda, mevcut kuramsal yapı ile benzerlik gösterdiği saptanmıştır. Elde edilen modelin farklı çalışma grubundan elde edilen verilerle doğrulanıp doğrulanmadığı DFA ile ortaya konulmuştur. Elde edilen uyum indeks değerleri  $\chi^2/sd=2.15$ , RMSEA=0,06, PGFI=0,72, GFI=0,90, AGFI=0,87, PNFI=0,81, IFI=0,95, NFI=0,91 ve CFI=0,95 olarak bulunmuştur. Bu uyum indeksleri dikkate alındığında, iki boyutlu teorik yapının toplanan verilerle doğrulandığı ifade edilebilir. DFA sonucunda elde edilen t değerleri dikkate alındığında; bu değerler "Öğrenme Odaklı Değerlendirme" boyutu için 7.00 ile 12.38 arasında ve "Performans Odaklı Öğrenme" boyutunda 8.36 ile 11.67 arasında olduğu görülmektedir. Bu durum yapı geçerliğine kanıt sağlamakta olup, ölçme aracında yer alan maddelerin her birinin yer aldığı boyuta yeterince hizmet ettiğini göstermektedir. DFA'da "Öğrenme Odaklı Değerlendirme" boyutu için hesaplanan madde faktör yük değerlerinin 0.41 ile 0.67 arasında değerler almakta iken; "Performans Odaklı Öğrenme" boyutunda bu değerler 0.50 ile 0.66 arasında olduğu görülmektedir. Ölçme aracının yapı geçerliğine kanıt sağlamak amacıyla yapılan AFA ile DFA sonucunda elde edilen bulgular ölçme aracının geçerliği olduğunu ortaya koymaktadır.

Madde analizi kapsamında madde toplam test korelasyonu ve % 27 alt-üst grupları karşılaştırılarak incelenmiştir. Ölçekte yer alan her bir maddenin, ölçmek istediği özelliği ölçüp ölçmediği ve ölçtüklere özellik açısından kişileri ayırt etmede ne kadar yeterli olduklarının belirlenmesi amacıyla, ilk olarak madde-toplam test korelasyonu hesaplanmıştır (Büyüköztürk ve ark., 2012). Eğer bir maddenin toplam puanla olan korelasyonu düşük ise, bu durum o maddenin testteki diğer maddelerden farklı bir niteliği ölçtüğünü gösterir (Karasar, 2014). Uzmanlar madde-toplam test korelasyonu 0.30 ve daha yüksek olan maddelerin iyi derecede ayırt edicilik özelliğe sahip olduğunu belirtmektedir (Büyüköztürk, 2014; Erkuş, 2014). Madde toplam test korelasyonu, test maddelerinden alınan puanlar ile testin toplam puanı arasındaki ilişkiyi açıklamaktadır. Madde toplam test korelasyon değerleri 0.36 ile 0.69 arasında yer almaktadır.

Beden eğitimi ders sınıf değerlendirme atmosferi ölçeğinin ölçüt geçerliğini test etmek için Aslan ve Akın (2015) tarafından Türkçe'ye uyarlanan "2x2 Başarı Yönelimleri Ölçeği" ve Çapri, Gündüz ve Gökbakan (2011) tarafından Türkçe'ye uyarlanması yapılan "Maslach Tükenmişlik Envanteri Öğrenci Formu" kullanılmıştır. Analiz sonucunda beden eğitimi dersi sınıf değerlendirme atmosferi ölçeğinden alınan puanlarla başarı yönelimi .43 ( $p < .001$ ) ve Maslach tükenmişlik ölçeği puanları arasında .45 ( $p < .001$ ) ilişki bulunmuştur. Elde edilen sonuca göre Beden eğitimi dersi sınıf değerlendirme atmosferi ölçeği, başarı yönelimi ve Maslach tükenmişlik öğrenci formu ölçeği ile orta düzeyde anlamlı bir ilişkinin olduğu, ölçüt geçerliğinin gerekliliğini yerine getirdiği söylenebilir.

Güvenirlilik analizleri sonucunda, ölçme aracının Cronbach Alfa güvenirlilik katsayısı, ölçeğin geneli için 0.74 olarak hesaplanmıştır (Tablo 2). Ölçeğin alt boyutları için bu değer sırasıyla 0.81 ve 0.71 olarak hesaplanmıştır. Ayrıca, ölçme aracının güvenirliliğine kanıt sağlamak amacıyla test-tekrar test tekniği uygulanmıştır. Ölçümlerin test-tekrar test güvenirliliği için 77 (42 kadın, 35 erkek) öğrenciden toplanan veriler üzerinden dört hafta ara ile iki uygulama yapılmıştır. Ölçme aracındaki maddelere verilen yanıtlar arasındaki tutarlılığın sınıandığı bu aşamada, puanlar arasındaki korelasyon katsayıları hesaplanmıştır. Ölçeğin genelinde hesaplanan test-tekrar test puanları arasındaki korelasyon katsayısı 0.74 olarak bulunmuşken, bu değer alt boyutlar için sırasıyla 0.77 ve 0.72 olarak tespit edilmiştir. Test-tekrar test tekniğine yönelik elde edilen bulgular incelendiğinde, ölçme aracının tutarlı ölçümler sağladığı şeklinde ifade edilebilir.

---

**Kaynakça**

- Alkharusi, H. (2009). Classroom assessment environment, self-efficacy, and mastery goal orientation: A causal model. Paper Presented at the 2nd International Conference of Teaching and Learning (ICTL 2009), INTI University College, Malaysia.
- Alkharusi, H. (2011). Development and datametric properties of a scale measuring students' perceptions of the classroom assessment environment. *International Journal of Instruction*, 4(1), 1308-1470.
- Alkharusi, H., Aldhafri, S., Alnabhani, H. ve Alkalbani, M. (2014). Modeling the relationship between perceptions of assessment tasks and classroom assessment environment as a function of gender. *Asia Pacific Education Researcher*, 23(1), 93-104.
- Ames, C. (1992a). *Achievement goals and the classroom motivational climate*. In D. H. Schunk & J. Meece (Eds.), *Student perceptions in the classroom* (pp. 327-348). Hillsdale, NJ: Erlbaum.
- Ames, C. (1992b). Classrooms: Goals, structures, and student motivation. *Journal of Educational Psychology*, 84, 261 – 271.
- Ames, C., & Archer, J. (1988). Achievement goals in the classroom: Students' learning strategies and motivation processes. *Journal of Educational Psychology*, 80, 260-267.
- Anderson, L. W., & Burns, R. B. (1989). *Research in classroom*. Oxford: Pergamon Press.
- Arslan, S., & Akın, A. (2015). 2x2 Başarı yönelimleri ölçeği (revize formu): Geçerlik ve güvenilirlik çalışması. *Sakarya Üniversitesi eğitim Dergisi*, 5(1), 7-15.
- Banks, S. R. (2012). *Classroom assessment: Issues and practices*. IL: Waveland Press Inc.
- Bartha, J. M., Dunlap, S. T., Dane, H., Lochmanb, J. E., & Wells, K. C. (2004). Classroom environment influences on aggression, peer relations, and academic focus. *Journal of School Psychology*, 42(2), 115-133.
- Borich, G. D. (1988). *Effective teaching methods*. Columbus, Ohio: Merrill Pub. Company.
- Brookhart, S. M., & DeVoge, J. G. (1999). Testing a theory about the role of classroom assessment in student motivation and achievement. *Applied Measurement in Education*, 12(3), 409-425.
- Brookhart, S. M. (1997). A theoretical framework for the role of classroom assessment in motivating student effort and achievement. *Applied Measurement in Education*, 10(2), 161-180.
- Büyüköztürk, Ş. (2014). *Sosyal bilimler için veri analizi el kitabı: İstatistik, araştırma deseni SPSS uygulamaları ve yorum* (Genişletilmiş 20. Baskı). Ankara: Pegem A.
- Büyüköztürk, Ş., Kılıç-Çakmak, E., Akgün, Ö. E., Karadeniz, Ş., & Demirel, F. (2012). *Bilimsel araştırma yöntemleri* (Geliştirilmiş 11. Baskı). Ankara: Pegem.
- Byrne, B. M. (2013). *Structural equation modeling with Mplus: Basic concepts, applications, and programming*. New York: Routledge
- Cabrera-Nguyen P. (2010). Author guidelines for reporting scale development and validation results in the journal of the Society for Social Work and Research. *J Soc Soc Work Res*. 1(2), 99-103.
- Church, M. A., Elliot, A. J., & Gable, S. L. (2001). Perceptions of classroom environment, achievement goals, and achievement outcomes. *Journal of Educational Psychology*, 93(1), 43-54. <http://dx.doi.org/10.1037/0022-0663.93.1.43>

- Cohen R. J., & Swerdlik M. E. (2010). *Psychological testing and assessment*. Boston: McGraw Hill Companies
- Creswell, J. W., & Clark, V. L. P. (2007). *Designing conducting mixed methods research*. Thousand Oaks, CA: Sage.
- Çapri, B., Gündüz, B. ve Gökçakan, Z. (2011). Maslach Tükenmişlik Envanteri-Öğrenci Formu (MTE-ÖF)' nin Türkçe'ye uyarlaması: geçerlik ve güvenilirlik çalışması. *Çukurova Üniversitesi Eğitim Fakültesi Dergisi*, 1(40), 134-147.
- Çokluk, Ö., Şekercioğlu, G., & Büyüköztürk, Ş. (2012). *Sosyal bilimler için çok değişkenli istatistik SPSS ve LISREL uygulamaları* (2. Baskı). Ankara: Pegem Yayınları.
- De Vellis, R. F. (2014). *Ölçek geliştirme: Kuram ve uygulamalar* (Ed. T. Totan). Ankara: Nobel Yayıncılık.
- Dorman, J. P. (2001). Associations between classroom environment and academic efficacy. *Learning Environments Research*, 4(3), 243-257.
- Ekiz, D. (2009). *Bilimsel araştırma yöntemleri*. Ankara: Anı Yayıncılık.
- Erkuş, A. (2014). *Psikolojide ölçme ve ölçek geliştirme-I: Temel kavramlar ve işlemler* (2. Baskı). Ankara: Pegem.
- Fraenkel, J. R., & Wallen, N. E. (2000). *How to design and evaluate research in education*. New York: McGraw.
- Fraser, B. J. & Walberg, H. J. (1991). *Educational environments: Evaluation, antecedents and consequences*. Oxford: Pergamon Press.
- Haertel, G. D., Walberg, H. J., & Haertel, E. H. (1981). Socio-psychological environments and learning: A quantitative synthesis. *British Educational Research Journal*, 7, 27-36.
- Hooper, D., Coughlan, J., & Mullen, M. (2008). Structural equation modelling: Guidelines for determining model fit. *Electronic Journal of Business Research Methods*, 6(1), 53-60
- İlhan, M. ve Çetin, B. (2014). Sınıf değerlendirme atmosferi ölçeğinin (SDAÖ) geliştirilmesi: Geçerlik ve güvenilirlik çalışması. *Eğitim ve Bilim*, 39(176), 31-50. Doi: 10.15390/EB.2014.3334
- Johnson, R., & Onwuegbuzie, A. (2004). Mixed methods research: A research paradigm whose time has come. *Educational Researcher*, 33(7), 14-26.
- Jöreskog, K. G., & Sörbom, D. (1993). *LISREL 8: Structural equation modeling with the simplis command language*. Lincolnwood: Scientific Software International, Inc.
- Karasar, N. (2014). *Bilimsel araştırma yöntemleri: Kavramlar, teknikler ve ilkeler* (27. Baskı). Ankara: Nobel.
- Kline, R. B. (2011). *Convergence of structural equation modeling and multilevel modeling*. M. Williams & W. P. Vogt (Eds.), *Handbook of methodological innovation in social research methods* (pp. 562-589). London: Sage
- Kılıç Cakmak E., Çebi, A., Kan A. (2014). Developing a social presence scale for e learning environments. *Kuram ve Uygulamada Eğitim Bilimleri*, 14,764-768.
- Linnenbrink, E. A., & Pintrich, P. R. (2001). Multiple goals, multiple contexts: The dynamic interplay between personal goals and contextual goal stresses. In S. Volet & S. Jarvela (Eds.), *Motivation in learning contexts* (pp. 251-270). Amsterdam: Pergamon.



- McMillan, J. H., & Workman, D. J. (1998). *Classroom assessment and grading practices: A review of the literature*. ERIC Document Reproduction Service.
- Mertler, C. A. (2003, October). *Preservice versus inservice teachers' assessment literacy: Does classroom experience make a difference?* Paper presented at the meeting of the Mid-Western Educational Research Association, Columbus, OH.
- Murphy K. R., & Davidshofer, R. (1991). *Psychological testing: Principles and applications*. New Jersey: Prentice-Hall.
- Seçer, İ. (2015). *SPSS ve LISREL ile pratik veri analizi: Analiz ve raporlaştırma*. Ankara: Anı Yayıncılık.
- Stiggins, R. J., & Conklin, N. F. (1992). *In teachers' hands: Investigating the practices of classroom assessment*. Albany: State University of New York Press.
- Schumacher, R. E., & Lomax, R. G. (1996). *A beginner's guide to SEM*. New Jersey: Mahwah.
- Tabachnick, B. G., & Fidell, L. S. (2007). *Multivariate analysis of variance and covariance. Using multivariate statistics*. Boston: Allyn and Bacon
- Tavşancıl, E. (2014). *Tutumların ölçülmesi ve SPSS ie veri analizi*. Ankara: Nobel.
- Walberg, H. J. (1968). Structural and affective aspects of classroom climate. *Psychology in the Schools*, 5, 247-253.
- Yıldırım, A., & Şimşek, H. (2011). *Sosyal bilimlerde nitel çalışma yöntemleri* (8. Baskı). Ankara: Seçkin Yayınları.