

PAPER DETAILS

TITLE: Baglama Performansi Ölçme Araci'nin (BPÖA) Gelistirilmesi: Geçerlik ve Güvenirlik Çalışması

AUTHORS: Alper BÖREKCI, Esra DALKIRAN

PAGES: 1726-1758

ORIGINAL PDF URL: <https://dergipark.org.tr/tr/download/article-file/1489087>


<http://kefad.ahievran.edu.tr>

Ahi Evran Üniversitesi Kırşehir Eğitim Fakültesi Dergisi

ISSN: 2147 - 1037

Development of the Bağlama Performance Measurement Tool(BPMT): Validity and Reliability Study

Alper Börekci
Esra Dalkıran

Article Information



DOI: 10.29299/kefad.854002

Received: 18.12.2019
Revised: 21.04.2020
Accepted: 12.06.2020

Keywords:

Bağlama Education,
Performance Measurement,
Performance Measurement Tool,
Validity,
Reliability.

Abstract

In this research, the aim was to develop a new measurement tool with high validity and reliability to positively contribute to the education-teaching process for end-semester exams in Individual Instrument Bağlama lessons in Fine Arts High Schools. A criterion skill list of 18 items in 4 subdimensions was prepared to measure Bağlama performance. To determine the reliability of the developed Bağlama Performance Measurement Tool (BPMT), Cronbach's alpha internal consistency test was applied and the reliability coefficient was found to be .91. To determine the consistency between judges, correlations were calculated between total points. The T test was used to determine whether the measurements made with the BPMT were sensitive or not. There was a significant difference found between measurements made with the measurement tool and with traditional methods. Scope validity of the tool was provided by expert opinions. Basic component (factor) analysis was used to identify the item validity of the developed bağlama performance measurement tool. The results of the analysis obtained a structure with one factor and 15 items explaining 78.838% of the total variance. Within the framework of these findings, it was identified that the BPMT is a measurement tool providing valid and reliable measurements.

Bağlama Performansı Ölçme Aracı'nın (BPÖA) Geliştirilmesi: Geçerlik ve Güvenirlik Çalışması

Makale Bilgileri



DOI: 10.29299/kefad.854002

Yükleme: 18.12.2019
Düzeltilme: 21.04.2020
Kabul: 12.06.2020

Anahtar Kelimeler:

Bağlama Eğitimi,
Performansın Ölçülmesi,
Performans Ölçme Aracı,
Geçerlik,
Güvenirlik.

Öz

Bu araştırmada, Güzel Sanatlar Liselerindeki Bireysel Çalgı Bağlama dersi yarıyıl sonu sınavlarında eğitim-öğretim sürecine olumlu yönde katkı sağlayacak geçerliği ve güvenilirliği yüksek olan yeni bir ölçme aracı örneği geliştirmek amaçlanmıştır. Bağlama performansını ölçmeye yönelik 4 alt boyuttan oluşan 18 maddelik ölçüt beceri listesi hazırlanmıştır. Geliştirilen Bağlama Performansı Ölçme Aracının (BPÖA) güvenilirliğinin belirlenmesi için Cronbach's Alpha iç tutarlılık testi uygulanmış ve güvenilirlik katsayısı .91 olarak bulunmuştur. Hakemler arası tutarlılığı belirlemek için de toplam puanlar arası korelasyon hesaplanmıştır. BPÖA'nın duyarlı ölçümler yapıp yapmadığının belirlenmesi için de t testinden yararlanılmış, ölçme aracı ile yapılan ölçme ile geleneksel yolla yapılan ölçme arasında anlamlı bir fark bulunmuştur. Ölçme aracının kapsam geçerliliği ise uzman görüşü ile sağlanmıştır. Geliştirilen bağlama performansı ölçme aracının madde geçerliğinin olup olmadığı temel bileşenler (Faktör) analizi ile tespit edilmiştir. Yapılan analiz sonucunda toplam varyansın %78.838'ini açıklayan 15 madde ve tek faktörden oluşan bir yapı elde edilmiştir. Bu bulgular çerçevesinde, BPÖA'nın geçerli ve güvenilir ölçümler sağlayan bir ölçme aracı olduğu tespit edilmiştir.

Sorumlu Yazar : Alper Börekci, Arş. Gör. Dr. Mehmet Akif Ersoy Üniversitesi, Türk Müziği Devlet Konservatuarı, Müzik Bölümü, Türkiye, aborekci@mehmetakif.edu.tr, ORCID ID: 0000-0002-5563-5383.

Esra Dalkıran, Prof. Dr., Mehmet Akif Ersoy Üniversitesi, Eğitim Fakültesi, Güzel Sanatlar Eğitimi Bölümü, Türkiye, edalkiran@mehmetakif.edu.tr, ORCID ID: 0000-0001-6829-7216.

Atf için: Börekci, A. & Dalkıran, E. (2020). Bağlama Performansı Ölçme Aracı'nın (BPÖA) geliştirilmesi: Geçerlik ve güvenilirlik çalışması. *Kırşehir Eğitim Fakültesi Dergisi*, 21(3), 1726- 1759.

Giriş

İnsanoğlu tarihsel süreç içerisinde doğasında var olan algı, bilinç, sorgulama, üretme, geliştirme, sürdürme, eğilimleri doğrultusunda, evrimsel bir değişim ve gelişimin en önemli parçası olmuştur. Bilgi aktarım eğilimi, en ilkel toplumlarda dahi, bir kurallar bütünü haline dönüşme yoluna girmiştir. Yerleşik düzenin gelişimiyle oluşan toplumlarda ise bu kurallı bilgi aktarımı, eğitim disiplini haline gelmiştir. Binlerce yıllık birikimin ardından toplumsal insanın eğitimi, bir davranış planı sunan kalıplaşmış bir bilgi aktarımı şeklini almıştır. Bu davranış planının karmaşık bir hale gelip bireye sunulmasına ise genel olarak “eğitim” adı verilir. Birey artık toplumsal bir varlıktır ve toplumun tercih ettiği davranış kalıplarını özümsemesi ve vakti geldiğinde edindiği bilgi birikimini sonraki kuşağa aktarması gerekmektedir. Eğitim kavramının çeşitli bilimsel tanımlarında da, “davranışsal hedeflerin planlı ve istendik olarak değiştirilmesi, geliştirilmesi, yetkinleştirilmesi ve bunların yasa ve ilkelerinin bulunması amacıyla teknikler geliştirilmesi sürecidir” (Erdem ve Akman, 1996; Tezcan, 1997; Uçan, 2005) şeklinde belirtilerek bu olguya işaret edilmiştir.

Eğitim sürecinin en önemli aşamalarından birisi ölçme boyutudur. Geniş anlamda ölçme, “herhangi bir objenin belirli bir niteliğini, belirli kurallara göre, sayarak, sınıflandırarak, derecelendirerek ya da birimlerle sayısal olarak ifade etme süreci olarak tanımlanabilir” (Özgüven, 2017, s.36). Günümüz toplumlarında başta bilimlerin gelişme süreci olmak üzere, hemen hemen her alanda ölçmenin yadsınamaz bir önemi vardır. Bireye, hedeflenen planlı davranışların kazandırılmasında, öngörülen değişikliklerin ne denli yerine getirildiğinin ortaya koyulmasında, ne düzeyde benimsenip özümsemişi, nasıl bir gelişme gösterdiğinin tespitinde ve bunun yanında bireyin, eğitimcinin ve eğitim sürecinin gelişimine yönelik tutarlı, doğru ve güvenilir kararların alınabilmesinde ölçme önemli bir rol oynamaktadır. Ölçme bu anlamda bir betimleme işlemi olarak görülebilir. “Ölçme ve değerlendirme yöntemleri öğretmene; (1) öğrencilerin başlangıç davranışları hakkında bilgi sağlamada, (2) her öğrenci için gerçeğe uygun amaçlar belirlemede, (3) başarı getiren amaçların başarı derecesini belirleyerek değerlendirmede ve (4) eğitimsel teknikleri belirlemede, değerlendirmede ve ayırt etmede yardımcı olur” (Mehrens ve Lehmann, 1991, s.8). Aynı zamanda öğrenci açısından bakıldığında ise, öğrenciye öğretmenin amaçlarını bildirme, öğrencinin eğitim öğretim sürecindeki istekliliğini ve etkinliğini artırma, doğru öğrenci alışkanlıkları ile öğrenci başarılarını olumlu yönde teşvik etme, öğrencinin etkili ve yetersiz yönlerine yönelik etkili geri dönütler verme hususunda mühim işlevlere sahiptir (Mehrens ve Lehmann, 1991).

Bağlama, Türk Halk Müziğinin (THM) zengin yapısının icrasında kullanılan ve Anadolu müzik geleneği içerisinde her türlü kültürel olgunun izdüşümü olarak görülen önemli halk sazlarımızdan birisidir. Böylesi bir görüş ve düşünce, Bağlama’yı geleneksel müzik yapısının taşıyıcı-aktarıcı unsuru haline getirmiştir. Geline süreçteki bu durum da THM Nazariyatı alanında bugüne değin yapılan araştırmaların ve çalışmaların genellikle, Bağlama temelli bir çerçevede yürütülmesine

sebebi olmuştur. Bu konudaki en büyük dayanak noktası bağlamanın perdeli bir çalgı olması ve yoğun bir şekilde icra edilmesinden kaynaklandığı düşünülmektedir (Börekci ve Nacakçı, 2019). Uzun yıllar boyunca usta-çırak ilişkisi içerisinde eğitim öğretim faaliyetlerini sürdüren bağlama çalgısı 'nın özellikle son yıllarda yapılan araştırmalar neticesinde öğretimine yönelik sistemli ve akademik anlamda çok çeşitli araştırmaların yapıldığı görülmektedir. Umut vadeditici düzeyde olan bu çalışmalar bağlama eğitimin hemen her kademesinde kendini göstermektedir. Ülkemizde bağlama eğitiminin verildiği resmî kurumlardan birisi Güzel Sanatlar Liseleridir. Lisans eğitimi öncesinde ele alınan bu formal eğitim dönemi, öğrencinin müzikal bilgi ve birikimini temel manada şekillendiren en önemli kurumların başında gelmektedir. Bu bağlamda, araştırmamızın ana konusunu oluşturan müzikal performans işlevi de her türlü eğitim kademesinde olduğu gibi GSL'de de önemli bir süreç olarak gösterilebilir.

Bireyin kelimelerle anlatamadığı duygu ve düşüncelerini, notalar ve sesler yardımıyla belli bir ahenk içerisinde ifade ettiği bir bilim ve sanat dalı olan müziğin eğitim sürecinde, bireyin enstrümantal icra ve seslendirmeye ilişkin sergilediği davranışlar müziksel performans olarak adlandırılmaktadır. "Bu bağlamda, müzikal performansın içerdiği çalma ve söyleme aşamaları bireyin psikomotor becerilerini sergilediği bölümdür" (Atak Yayla, 2004). Müziğin sanatsal bir performans olarak görülmesi ve müzik eğitimi süreci açısından da bu performansın ölçülmesi birçok sebepten dolayı önemli görülen bir husustur. Çünkü "müziksel performansa dayalı öğretimin ve bu öğretim sürecinin sınanmasının nesnel bir temele dayandırılması bir süreç olan müzik eğitimi açısından gereklidir" (Schleuter, 1996, s.176). Bu yüzden performans ölçme araçları, öğrencinin eğitim süreci içerisindeki müzikal gelişiminin doğru bir şekilde tespiti, öğrenciye özenli sistematik geri bildirimlerin verilmesi, gelişimin aşamalarının düzenli olarak takip edilmesi için gereklidir. Bireyin değerlendirilme sürecine ilişkin geliştirilen ölçme araçları, bilişsel, duyuşsal, ya da devinişsel olmak üzere ölçülecek davranışın niteliğine göre çeşitlilik göstermektedir. Sönmez'e (2001) göre devinişsel (psikomotor) alan, "öğrenci becerilerinin kodlandığı alandır. Bu alan bilişsel ve duyuşsal alanla iç içedir. Devinişsel alanla ilgili öğrenilmiş davranışlar ya ürün ya süreç ya da hem ürün hem süreç açısından değerlendirilmelidir." Bu çerçevede öğrencinin gelişim sürecinin ölçülmesi ve değerlendirilmesi için, ne kadar hedef davranış varsa belirlenmeli ve bunlara uygun ölçme koşulları geliştirilmelidir. Bu bağlamda, eğitimde yapılan ölçümlerden elde edilen verilerin değerlendirilmesi, ileriye dönük olarak verilecek kararların, yapılacak düzenlemelerin, hazırlanacak olan eğitim programlarının, eğitim sürecinde ihtiyaç duyulan/duyulabilecek olan materyallerin belirlenmesine ve süreç devam ederken tekrar gözden geçirilerek geliştirilmesine önemli katkı sağlayacaktır (Akçay, 2011).

Problem Durumu

“Eğitimde, eğitim programının sağlam olup olmadığını anlama, öğretimde başvuru metotların etkililik derecesini saptama, öğrencileri başarılı olabilecekleri alanlara yönlendirme, öğrenme güçlüklerini teşhis etme, öğrenci başarısını saptama, ölçme ve değerlendirme öğesinin iyi işleyip işlemediğini görme gibi amaçlarla yapılan değerlendirmelerin hepsi ölçme sonuçlarına dayanır” (Baykul, 2000, s.89). “Müzik eğitiminde olduğu gibi, özellikle bireysel olarak yapılan ve devinışsel davranışların ölçüldüğü sınavlarda ölçülmek istenen davranışların tespit edilmemesi, bu davranışların ağırlıklı puanlarının belirlenmemesi kısaca yazılı bir ölçme aracı oluşturulmaması ölçme sonuçlarının nesnellikten uzak olmasına sebep olmaktadır” (Dalkıran, 2006, s.14).

Sınav performans ölçümlerinde, derecelendirilmiş ve etkililiği sınanmış yazılı bir ölçme aracı kullanılmaması, öğrencinin sınav performansına ilişkin herhangi bir geri bildirim almasını olumsuz yönde etkilemekte ve bunun yanında gelişimi ile ilgili sistematik ve sağlıklı bilgi edinilememesine sebep olmaktadır. Fiske’ ye (1983, aktaran Ciorba ve Smith, 2009) göre, pek çok değerlendirici de performans derecelerini nasıl belirlediklerini bilmemektedir. Ayrıca eğitimsel bir süreç içerisinde düşünüldüğünde, öğretim boyutunun gelişen- gelişmeyen, işleyen- işlemeyen yanları tespit edilmesi oldukça güçleşecektir. Yapılan analizler ve uzman görüşleri doğrultusunda, günümüzde müzik eğitimi alanında, ölçme araçlarının geliştirilmesi ile ilgili çeşitli çalışmalar yapılmasına rağmen, halk çalgıları ve milli kültürümüzün en önemli simgesi olan Bağlamanın ölçme boyutuna ilişkin geçerli ve güvenilir bir ölçme aracının geliştirilmediği saptanmıştır. Bahsedilen bu hususlar çerçevesinde, bağlama dersi veren öğretim elemanları ve öğretmenlerden elde edilen görüşlerden hareketle bireysel çalgı bağlama dersi yarıyıl sonu sınavlarında kullanılmak üzere yeni bir performans ölçme aracı oluşturulması ve bu aracın kullanılabilirliğinin değerlendirilmesi bu çalışmanın ana konusunu oluşturmaktadır.

Bu bağlamda, araştırmanın problem cümlesi “Millî Eğitim Bakanlığı’na (MEB) bağlı Güzel Sanatlar Liseleri Bireysel Çalgı Bağlama dersinin performans boyutuna ilişkin geliştirilmiş ölçme aracı geçerli ve güvenilir midir?” şeklinde belirlenmiştir.

Alt problemler: Araştırmanın amacı, önemi ve belirtilen problem durumuna ilişkin incelenen alt problemler şu şekilde kurgulanmıştır.

1. Araştırma kapsamında geliştirilen ölçme aracı güvenilir bir ölçme aracı mıdır?
2. Araştırma kapsamında geliştirilen ölçme aracı geçerli bir ölçme aracı mıdır?

Araştırmanın Amacı ve Önemi

Bu çalışmada, Güzel Sanatlar Liselerindeki “Bireysel Çalgı Bağlama” dersi yarıyıl sonu sınavlarında yapılan ölçmeye olumlu yönde katkı sağlamak ile geçerliği ve güvenilirliği yüksek olan yeni bir ölçme aracı örneği geliştirmek amaçlanmıştır.

Bu araştırma, yeni bir ölçme aracı geliştirerek bireysel çalgı bağlama dersi alan her öğrencinin yarıyıl sonu sınav performansına ilişkin ölçümün daha tutarlı, nesnel ve objektif bir biçimde yapılmasını sağlaması bakımından önemli görülmektedir. Geliştirilen ölçme aracı sayesinde öğrencilerin öğrenme sürecindeki güçlükleri, eksiklikleri ve yanlışlarının ütesinden gelmeye yönelik verimli ve etkili geribildirimler vererek daha pozitif bir rehberlik yapılabileceğine ve eğitim öğretim sürecindeki gelişimin daha etkin değerlendirileceğine inanılmaktadır. Ayrıca performans ölçümünde, bağlama dersi veren öğretim elemanları ve öğretmenlere ölçüm sistematığı üzerinden tutarlılık ile kolaylık sunacağı ve bu çerçevede de halk çalgılarına yönelik gelecekte oluşturulacak ölçme araçlarına örnek olacağı düşünülmektedir.

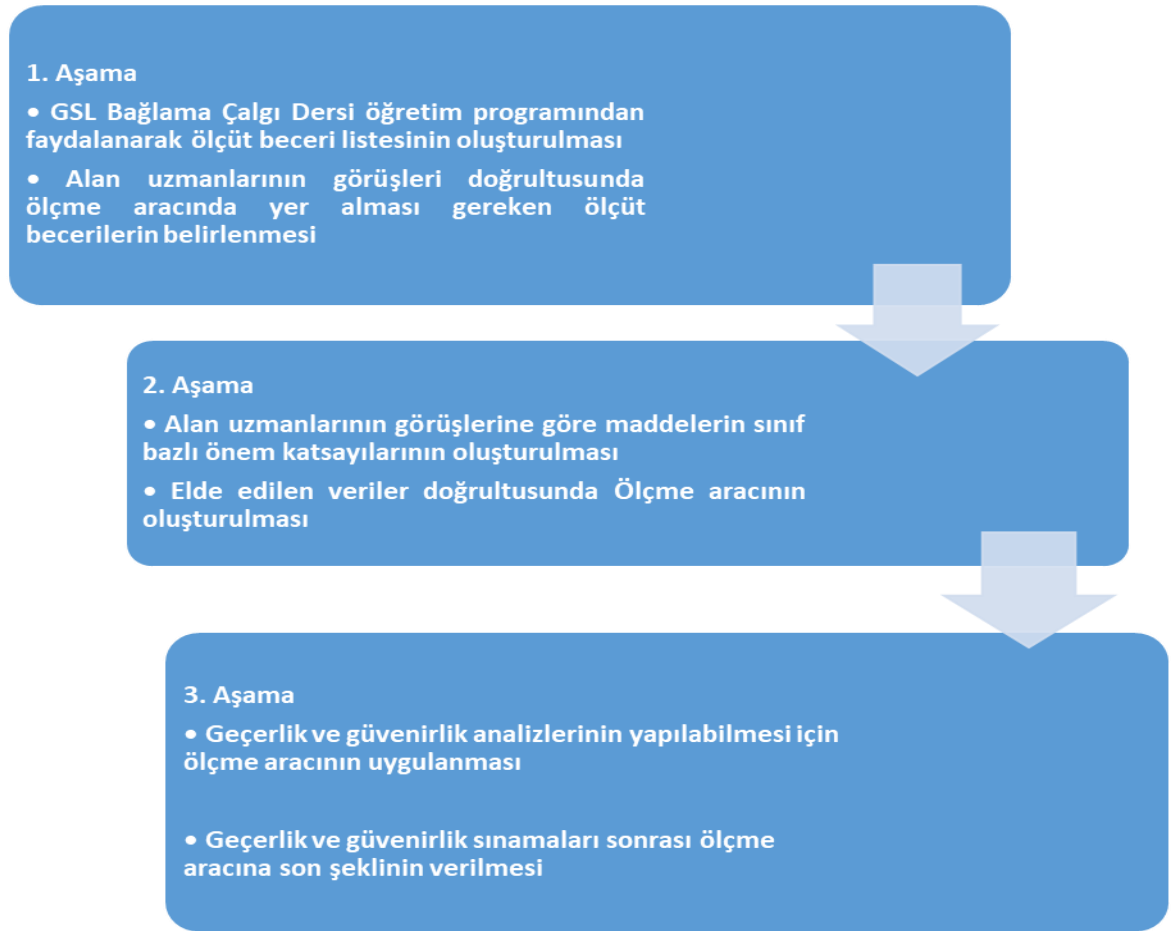
Yöntem

Bu bölümde, araştırmanın modeli ve deseni, çalışma grubu, araştırmada kullanılan verileri toplamada kullanılan veri toplama araçları ve ölçme aracının geliştirilme aşamaları, verilerin toplanması, toplanan verilerin çözümlenmesinde kullanılan istatistiksel yöntem ve teknikler ayrıntılı olarak açıklanmıştır.

Araştırmanın Modeli ve Deseni

Bu araştırma, Güzel Sanatlar Liselerinde “Bireysel Çalgı Bağlama” dersi yarıyıl sonu sınavlarında öğrenci performanslarını değerlendirmek amacıyla geçerliği ve güvenilirliği yüksek olan bir ölçme aracının oluşturularak kullanımının sağlanması amacıyla betimsel niteliktedir. Araştırmanın yürütülmesinde genel tarama modeli kullanılmıştır. “Tarama modelleri, geçmişte ya da halen var olan bir durumu var olduğu şekliyle betimlemeyi amaçlayan araştırma yaklaşımlarıdır” (Karasar, 2009). Araştırmanın temellendirilmesi ve madde havuzunun oluşturulması aşamasında belgesel tarama yapılarak konuyla ilgili olan kaynaklar belirlenip toplanmış ve akabinde ilgili ders programının analiz edilip çözümlenmesiyle veriler elde edilmiştir.

Araştırmanın deseni ve deseni belirleyen aşamalar şöyledir;



1. Aşama: GSL Bireysel Çalgı Bağlama dersi öğretim programında yer alan hedef kazanımlara göre madde havuzu oluşturulmuş, oluşturulan bu madde listesi alan uzmanlarının görüşleri alınarak ölçme aracında yer alması gereken ölçüt beceriler belirlenmiştir.

2. Aşama: Bağlama dersi veren öğretim elemanları ve öğretmenlerin görüşleri doğrultusunda, ölçme aracında yer alan her bir maddenin sınıf düzeyine göre önem katsayıları belirlenmiş ve ölçme aracına deney süreci öncesinde son hali verilmiştir.

3. Aşama: Ölçme aracının denence sürecine geçilmiş, geçerlik ve güvenirlik durumlarının sınanması yapılmıştır. Bunun sonucunda ölçme aracına son şekli verilmiştir.

Çalışma Grubu

Araştırmanın çalışma grubu öğrenciler ve öğretim elemanlarından oluşmaktadır. Bu bağlamda ölçme aracının denel kısmında kullanılan öğrenciler ile hazırlanan ölçüt becerilerin yapı ve kapsam geçerliliği kısmında görüşlerinden faydalanılan öğretim elemanları ile ilgili bilgiler Tablo 1, 2 ve 3'te ayrıntılı olarak sunulmuştur. Buna göre, öğrencilerin bulundukları liseler ve sınıf düzeylerine göre sayıları şu şekildedir;

Tablo 1. Araştırmaya katılan öğrencilerin bulundukları liselere göre sınıf düzeyleri ve sayıları

Okul Adı	Katılımcı Öğrencilerin Toplam Sayıları	Sınıf Düzeyleri			
		9	10	11	12
Çankırı GSL	11	3	4	2	2
Burdur GSL	9	2	2	3	2
Antalya ATSO GSL	8	2	2	3	1
Bolu GSL	9	1	1	3	4
Isparta GSL	16	4	4	5	3
Toplam	53				

Tablo 1'e göre, araştırmaya 53 öğrenci katılmıştır. Özel yetenek gerektiren GSL' ne günümüzde başvuru sayısının oldukça az olması hatta %71'inin belirlenen kontenjanlarını dolduramadığı (Aksu, 2014) göz önüne alındığında öğrenci sayısının az olması dikkat çekmektedir.

Ölçme aracında yer alması gereken maddelerin ve sınıf düzeylerine göre önem katsayılarının belirlenmesi için çeşitli üniversiteler ve MEB'e bağlı güzel sanatlar liselerinde görev yapmakta olan bağlama dersini yürüten öğretim elemanları ve öğretmenlerin görüşlerine başvurulmuştur. Ankete katılan öğretim elemanları ve öğretmenlere ait demografik bilgiler ve görev yaptığı kurumlar Tablo 2 ve Tablo 3'te sunulmuştur.

Tablo 2. Ankete katılan öğretim elemanları ve öğretmenlerin görev yaptığı kurumlar

Kurumlar	f	Yüzde (%)
Antalya Atso GSL	1	10
Çankırı GSL	1	10
Burdur GSL	1	10
Ege Üniversitesi	2	20
Muğla Sıtkı Kocaman Üniversitesi	1	10
Mehmet Akif Ersoy Üniversitesi	1	10
Akdeniz Üniversitesi	1	10
Erzurum Atatürk Üniversitesi	1	10
Necmettin Erbakan Üniversitesi	1	10
Toplam	10	100

Tablo 2' ye göre, ankete görüş bildiren katılımcılar; %10 Antalya GSL, %10 Çankırı GSL, %10 Burdur GSL, %20 Ege Üniversitesi, %10 Muğla Sıtkı Kocaman Üniversitesi, %10 Mehmet Akif Ersoy Üniversitesi, %10 Akdeniz Üniversitesi, %10 Erzurum Atatürk Üniversitesi ve %10 Necmettin Erbakan Üniversitesi olmak üzere çeşitli üniversite ve güzel sanatlar liselerindeki öğretim elemanları ve öğretmenlerden oluşmaktadır.

Tablo 3. Ankete katılan öğretim elemanları ve öğretmenlerin demografik bilgileri

		f	Yüzde (%)
Unvan	Profesör	1	10
	Doçent	3	30
	Öğretim Elemanı	3	30
	Öğretmen	3	30
	Toplam	10	100
Deneyim	1-5 yıl	2	20
	6-10 yıl	1	10
	11-15 yıl	3	30
	16-20 yıl	1	10
	21 yıl veya daha fazla	3	30
	Toplam	10	100
Eğitim Seviyesi	Doktora	5	50
	Yüksek Lisans	3	30
	Lisans	2	20
	Toplam	10	100

Tablo 3'e göre, ankete katılan bağlama öğretim elemanları ve öğretmenlerin unvan bilgilerine göre, %10 Profesör, %30'u Doçent, %30 Öğretim görevlisi ve %30 Öğretmenlerden oluşmaktadır. Kıdem düzeylerine bakıldığında ise, %20 1-5 yıl, %10 6-10 yıl, %30 11-15 yıl, %10 16-20 yıl ve %30 21 yıl ve üzeri olduğu görülmektedir. Öğrenim düzeyleri dikkate alındığında ise, %50'sinin Doktora, %30'unun yüksek lisans ve %20'sinin lisans derecesine sahip olduğu tespit edilmiştir. Buna göre, araştırmada görüş bildiren katılımcıların kıdem ve öğrenim düzeyleri dikkate alındığında yeterince deneyimli oldukları söylenebilir.

Veri Toplama Araçları

Araştırmada, MEB tarafından hazırlanan Bireysel Çalgı Bağlama dersi öğretim programı çözümlenerek ve ilgili literatür taranarak;

1. Geliştirilen ölçme aracında yer alacak ölçüt becerilerin ve önem derecelerinin belirlenebilmesi için Bireysel Çalgı Bağlama Dersi öğretim elemanları ve öğretmenlerine uygulanan "Anket Formu",
2. Araştırmacılar tarafından geliştirilen "Bağlama Performansı Ölçme Aracı" (BPÖA) ile birlikte iki adet veri toplama aracı kullanılmıştır.

Anket formu: Performans ölçme aracının hazırlanması aşamasında, MEB tarafından hazırlanan Bireysel çalgı bağlama dersi öğretim programı incelenerek hedefler ve program bünyesinde kazandırılması planlanan ölçüt beceriler (davranışlar) 18 madde olarak belirlenmiştir. Hazırlanan Anket Formu aracılığıyla alan uzmanlarının belirlenen 18 maddenin geçerliliğine ve sınıf düzeylerine göre önem katsayılarına ilişkin görüşleri alınmış, bu görüşler doğrultusunda 18 madde içinden yarıyıl sonu sınavlarındaki performans ölçümlerinde kullanılmak üzere 15 ölçüt beceri (davranış) belirlenmiştir.

Bağlama performans ölçme aracı: Ölçme aracı ölçülecek olan Programın Boyutu, Sınav Performans Boyutu ve Yarıyıl İçi Durum Boyutu olmak üzere 3 ana boyuttan; dönem içi performans, programın gereklerine uygunluk ile teknik ve müzikal beceriler olmak üzere 4 alt boyuttan oluşturulmuştur.

Alan uzmanlarına uygulanan anket sonucunda ölçme aracında yer alan boyut, alt boyut ve katsayılar belirlenmiş olup, ölçme aracının derecelendirme bölümünde “1-çok düşük, 2-ortanın altında, 3-Orta, 4-ortanın üstünde, 5-çok iyi” seçeneklerinden oluşan 5’ li Likert tipi ölçek kullanılmıştır. Buna ek olarak “9. Sınıf, 10. Sınıf, 11. Sınıf ve 12. Sınıf” olmak üzere her bir madde için sınıf düzeylerine göre önem katsayıları eklenmiştir. Puan hesaplamaları öğrencinin sınıf düzeyine göre bu önem katsayıları üzerinden yapılmaktadır.

Ölçme aracının yapı geçerliliğinin belirlenebilmesi için temel bileşenler analizi uygulanmıştır. Yapılan analiz sonucunda geliştirilen ölçme aracının tek bir temel yapıyı ölçtüğü sonucuna varılmıştır. Temel bileşenler analizi sonuçlarına dayanarak ölçme aracı ile amaçlanan bağlama performansının ölçülmesi temel boyutunun ölçülebildiği saptanmıştır. Ölçme aracının Cronbach Alfa güvenirlik katsayısı .91 olarak belirlenmiştir. Ölçme aracının, bağlama performansının ölçülmesine ilişkin varyansın %78,83’ünü açıkladığı tespit edilmiştir. Diğer bir ifade ile, ölçme aracının her bir maddesi, içeriğin genel amacı olan BPÖA’na yüksek oranda hizmet etmektedir.

Verilerin Analizi

Araştırmanın, belirlenen amaç ve problem durumu doğrultusunda kurgulanan alt problemler çerçevesinde elde edilen veriler IBM SPSS Statistics 22 programı ile analize tabi tutulmuştur. İstatistiksel anlamlılığın ölçümünde .05 anlamlılık düzeyi seçilmiştir. Oluşturulan BPÖA’nın geçerlik ve güvenirliği 53 kişilik bir öğrenci grubu üzerinde uygulanarak sınanmıştır. Bu aşamada veriler, eş zamanlı olarak araştırmacı ve bir alan uzmanı tarafından ayrı ayrı puan verilmek suretiyle elde edilmiştir.

Güvenirlik: “Güvenirlik, bir ölçme aracının ve dolayısıyla da ölçme sonuçlarının hatalardan arınlık derecesi olarak tanımlanabilir” (Bahar, Nartgün, Durmuş ve Bıçak, 2006). Diğer bir ifadeyle güvenirlik, ölçümün farklı zaman dilimleri içerisinde, farklı mekânlarda ve farklı kişilerle tekrarlandığında aynı sonuçların elde edilmesidir (Ruiz-Primo ve Shavelson, 1996; Johnson, Penny ve Gordon, 2009). Ölçme aracının güvenirliğinin ortaya koyulması için öncelikle iç tutarlılığı incelenmiştir. Edinilen değerlendirme puanlarına ait iç tutarlılığın incelenmesinde, Cronbach’s Alpha formülü, puanlayıcıların verdiği puanların ortalamaları üzerinden kullanılmıştır (Büyüköztürk, 2010). İstatistiksel olarak, puanların ortalamalarının alınıp alınamayacağının belirlenmesi için de iki puanlayıcı arasındaki uyum düzeyinin hesaplanmasını sağlayan Cohen’s Kappa formülünden yararlanılmıştır (Kutlu, Doğan ve Karakaya, 2010).

Ölçme aracının diğer bir türden güvenirliğinin incelenmesi için toplam puanlar üzerinden puanlayıcılar arası tutarlılığa bakılmıştır. Geliştirilen ölçme aracının duyarlılığının sınanmasında,

puanlayıcıların BPÖA kullanarak ve kullanmadan (geleneksel yolla) verdiği puanlar arasındaki fark, İlişkili Örneklem t-testi kullanılarak değerlendirilmiştir.

Geçerlik: “Geçerlik, testin bireyin ölçülmek istenen özelliğini diğer özelliklerle karıştırmadan ne derece doğru ölçtüğüyle ilgilidir” (Büyüköztürk, Çakmak, Akgün, Karadeniz ve Demirel, 2009, s.115). BPÖA’nın geçerliğinin belirlenmesinde ilk olarak kapsam geçerliği irdelenmiştir. Bu amaç çerçevesinde, ölçme aracının içerik, ifade, puanlama vb. özellikleri hakkında hazırlanan anket vasıtasıyla alan uzmanlarının görüşleri alınmıştır. Alan uzmanları ölçme aracının kapsam bakımından, Bağlama dersi yarıyıl sonu sınav performansının ölçülmesine yönelik geçerli bir ölçme yaptığını belirtmişlerdir. “Ölçme aracının ölçülmek istenen davranış bağlamında soyut bir kavramı (faktörü) doğru bir şekilde ölçebilme becerisine yapı geçerliği denir” (Büyüköztürk, 2010). Ölçme aracının yapı geçerliğinin olup olmadığının incelenmesi için Temel Bileşenler Analizi yapılarak faktör yüklerine sonrasında da madde toplam korelasyonlarına bakılmıştır. Madde analizi yapılarak, ölçme aracında yer alan maddelerin geçerlilikleri incelenmiştir. Bu amaçla ölçme aracından elde edilen toplam puanlara göre, oluşturulan alt %27 ve üst %27’lik grupların madde ortalama puanları arasındaki farkların incelenmesi için İlişkisiz Örneklem t-testi kullanılmıştır.

Araştırmanın Etik İzinleri

Yapılan bu çalışmada “Yükseköğretim Kurumları Bilimsel Araştırma ve Yayın Etiği Yönergesi” kapsamında uyulması belirtilen tüm kurallara uyulmuştur. Yönergenin ikinci bölümü olan “Bilimsel Araştırma ve Yayın Etiğine Aykırı Eylemler” başlığı altında belirtilen eylemlerden hiçbirisi gerçekleştirilmemiştir.

Etik kurul izin bilgileri:

Araştırmanın verileri 2020 yılı öncesinde toplanmıştır. Bu çerçevede araştırma ile ilgili alınan herhangi bir etik izin belgesi bulunmamaktadır. Buna karşın, araştırmada elde edilen veriler etik ihlaller yapılmadan kurallara uygun bir şekilde toplanmıştır.

Bulgular

Bu bölümde, Güzel Sanatlar Liselerinde Bireysel çalgı bağlama dersi yarıyıl sonu sınavlarında kullanılmak üzere, geçerli ve güvenilir bir ölçme aracı oluşturulması ve bu aracın kullanılabilirliğinin test edilmesi ile ilgili belirlenen alt problemler çerçevesinde toplanan verilerin istatistiksel çözümlenmesiyle elde edilen bulgular ve bu bulgulara ilişkin yorumlar yer almaktadır.

Birinci Alt Probleme İlişkin Bulgular ve Yorum

İç tutarlılık: Öğretim elemanı ile araştırmacının eş zamanlı olarak yapmış olduğu puanlama ile elde edilen puanların iç tutarlılığının sınanması amacıyla hakemlerin her bir madde için vermiş olduğu puanların ortalaması alınarak Cronbach’s Alpha İç Tutarlılık katsayısı hesaplanmıştır. Ancak iki hakemin verdiği puanların ortalamasının alınabilmesi için hakemlerin her bir madde için verdiği

puanlar arası tutarlılığa bakmak gerekmektedir. Bu tutarlılığın istatistiksel olarak hesaplanması için Cohen's Kappa Formülü kullanılmıştır. Elde edilen sonuçlar Tablo 4'te sunulmuştur.

Tablo 4. Her bir madde için puanlayıcılar arası kappa katsayıları

Maddeler	n	Kappa
Madde 1	53	.974
Madde 2		.949
Madde 3		.924
Madde 4		.951
Madde 5		.926
Madde 6		.921
Madde 7		.923
Madde 8		.924
Madde 9		.923
Madde 10		.951
Madde 11		.899
Madde 12		.876
Madde 13		.925
Madde 14		.875
Madde 15		.926

Yapılan hesaplamalar doğrultusunda, elde edilen Kappa katsayılarının .60'tan büyük olması puanlayıcılar arası tutarlılığın olduğunun bir kanıtı olarak kabul edilebilir (Kutlu ve diğerleri, 2010). Bu bulgudan hareketle, hakemlerin her bir madde için vermiş olduğu puanların ortalaması alınmış ve Cronbach's Alpha iç tutarlılık katsayısı hesaplanmıştır. Her bir maddenin ortalaması üzerinden uygulanan Cronbach Alpha Testi sonucunda güvenilirlik katsayısı .91 ($p < .05$, $n=53$) olarak bulunmuştur. Elde edilen bu değer, performans değerlendirme araçlarında kabul edilen minimum değer ($.70$) üzerinde olduğundan (Büyüköztürk, 2010), BPÖA' nın iç tutarlılığının olduğundan söz edilebilir.

Toplam puanlar arası tutarlılık: BPÖA' nın güvenilirliğinin bir başka yönden sınanması için toplam puanlar arası tutarlılık korelasyonu incelenmiştir. Uygulanacak olan korelasyon formülünün belirlenmesi için ise toplam puanlar üzerinden puanlayıcılar arası çarpıklık durumu incelenmiştir. Yapılan çarpıklık analizinin sonuçları Tablo 5'te sunulmuştur.

Tablo 5. Toplam puanlar üzerinden puanlayıcılar arası çarpıklık

Puanlayıcılar	n	Çarpıklık
Araştırmacı	53	-.410
Öğretim Elemanı		-.479

Tablo 5 incelendiğinde toplam puanlar arası çarpıklık değerlerinin dağılımının normal olduğu tespit edilmiştir. Bu sonuçlar doğrultusunda uygulanacak olan korelasyon formülü Pearson Momentler Çarpımı Korelasyon Katsayısı olarak belirlenmiştir. Yapılan analiz Tablo 6'da sunulmuştur.

Tablo 6. Pearson momentler çarpımı korelasyon katsayıları

Puanlayıcılar	n	$\bar{x}$	ss	p	Korelasyon
Araştırmacı	53	65,2830	20,51770	.000*	.995
Öğretim Elemanı	53	66,8113	19,97600		

Tablo 6 incelendiğinde, 53 öğrencinin toplam puanları üzerinden yapılan Pearson Momentler Çarpımı Korelasyon Katsayıları sonuçlarına göre hakemlerin puanlamaları arasında “.995” oranında anlamlı bir ilişki olduğu tespit edilmiştir. Bu durum, hakemlerin yaptıkları ölçümler arasındaki tutarlılığın ve BPÖA’ nın güvenilir olduğunun bir başka kanıtı olarak gösterilebilir.

Duyarlılık: Geliştirilen ölçme aracının, duyarlı ölçümler yapan bir ölçme aracı olup olmadığının sınanması için, ölçme aracı ile verilen puanlarla, geleneksel yöntemle (ölçme aracı olmadan) verilen puanların ortalamaları arasında anlamlı bir fark olup olmadığı incelenmiştir. Ölçme aracı ile verilen puanlarla geleneksel yöntemlerle verilen puanların ortalamaları üzerinden yapılan ilişkili örneklem t-testi sonuçları Tablo 7’de belirtilmiştir.

Tablo 7. Ölçme aracı kullanılarak verilen puanlar ve geleneksel yöntemle verilen puanlar arası ilişkili örneklem t-testi sonuçları

Ölçme Yöntemi	n	$\bar{x}$	ss	p	t
Ölçme Aracı ile Yapılan Ölçme	53	66.0472	20.22262	.000*	-10.640
Geleneksel Ölçme	53	69.6226	20.17589		

$p < .05$

Yapılan analizin sonuçlarına göre, 53 öğrenci üzerinde Ölçme Aracı ile verilen puanlar ve Ölçme Aracı olmadan (Geleneksel Yolla) verilen puanlar arasında anlamlı fark tespit edilmiştir ($p < .05$). Bu, ölçme aracı ile verilen puan ortalamalarının, geleneksel yolla verilen puan ortalamalarından yaklaşık 3 puan daha düşük olması sebebiyle ortaya çıkmaktadır. P değeri, standart sapma ve puan ortalamaları değerleri dikkate alındığında, puan ortalamaları arasındaki farkın, ölçme aracı kullanılarak yapılan ölçümlerin daha duyarlı olmasından kaynaklandığı düşünülmektedir. Bu tespitler doğrultusunda ölçme aracının duyarlı bir ölçüm yaptığı söylenebilir.

İkinci Alt Probleme İlişkin Bulgular ve Yorum

Kapsam geçerliliği: Araştırma kapsamında geliştirilen BPÖA’ nın kapsam geçerliğini belirlemek amacıyla, ders programının çözümlemesi akabinde 18 madde olarak oluşturulan ölçüt beceriler listesi hakkında, alan uzmanlarının görüşleri alınmıştır. Bağlama öğretim programının çözümlenmesi sonucunda belirlenen ve BPÖA’nda bulunması gerektiği düşünülen ölçüt beceriler, alan uzmanlarından elde edilen ölçme aracının ifade, içerik ve puanlamasına yönelik görüşleri doğrultusunda revize edilerek 15 madde olarak tekrar düzenlenmiştir. Görüşleri alınan alan uzmanları, geliştirilen ölçme aracının kapsam geçerliği çerçevesinde bağlama performansının ölçülmesine yönelik ölçme yaptığını belirtmişlerdir.

Yapı geçerliliği: Araştırma kapsamında geliştirilen ölçme aracının yapı geçerliğinin sınanması için temel bileşenler (faktör) analizi yapılmıştır. Yapılan temel bileşenler analizi sonuçları Tablo 8’de verilmiştir. Temel bileşenler Analizinin yapılmasının gerekliliği için KMO-Bartlett testi yapılmıştır. KMO testi sonucunda .915 değeri elde edilmiş ve bu veri faktör analizinin mükemmel bir biçimde

kullanılacağını göstermiştir. “p” değeri incelendiğinde ise $p=.000$ bulunmuş ve bu değer $<0,05$ ’ten küçük olduğundan Temel Bileşenler Analizinin yapılabileceği sonucuna varılmıştır.

Tablo 8. Temel bileşenler analizi sonuçları

No	Ölçüt Beceriler	Faktör yükü
5	Perde ve pozisyonlar arasındaki geçişlerde hâkimiyeti sağlayabilme	.885
6	Eserde yer alan notaları usulüne uygun tezene kalıplarıyla seslendirebilme	.882
4	Eseri parmak numaralarına uygun bir biçimde seslendirebilme	.872
3	Sağ ve sol el birlikteliğini uyum içerisinde yapabilme	.839
10	Eseri gerçek metronomuna uygun bir hızda seslendirebilme	.838
9	Eserde yer alan notaları sürelerine uygun bir biçimde doğru seslendirebilme	.825
1	Bağlamayı uygun oturuş pozisyonunda, sağ el ve sol el tutuş tekniklerini doğru bir biçimde uygulayarak konumlandırabilme	.824
15	Dönem boyunca bağlama dersine gerekli ilgiyi göstererek titiz davranabilme, verilen ödevleri çalışarak zamanında ve doğru seslendirebilme	.813
13	Eseri bir bütünlük içerisinde seslendirebilme	.810
2	Aşağı ve yukarı tezene vuruşlarını temiz ve dengeli yapabilme	.797
11	Eseri temposunda kalarak seslendirebilme	.783
14	Bulunduğu bağlama dersi programının gereklerine uygun düzeyde eserler seslendirebilme	.762
6	Eserde yer alan notaları usulüne uygun tezene kalıplarıyla seslendirebilme	.725
7	Eserdeki yöresel icra(tavır) özelliklerini, doğru ve akıcı bir biçimde seslendirebilme	.679
12	Eserde yer alan nüansları(varsa) etkili bir biçimde uygulayabilme	.493
Açıklanan toplam varyans		78.838%
Eigen Value		11.826

Tablo 8’de görüldüğü üzere araştırma kapsamında geliştirilen BPÖA, amaçlanan tek bir temel yapıyı ölçmektedir. Her bir maddenin ölçme aracının genel amacına hizmet etme derecesini bildiren bileşen yükleri .493 ile .885 arasında değişmektedir. Az maddeli testlerde faktör yük değerlerindeki alt sınırın .45 (Büyüköztürk, 2010) olarak kabul edildiği göz önüne alındığında tüm faktör yüklerinin bu değerin üzerinde olduğu görülmeye karşın, 12. Madde’nin faktör yükündeki düşüklüğün ise araştırmaya katılan GSL seviyesindeki öğrencilerde eserlerde geçen nüansları uygulayabilme becerisinin henüz gelişmemiş olmasından kaynaklı olduğu düşünülmektedir. Ölçme aracının bağlama performansının ölçülmesine ilişkin toplam varyansın %78.838’sini açıkladığı görülmüş ve Eigen değeri 11.826 olarak tespit edilmiştir. Bu sonuçlar ışığında, BPÖA’nda yer alan tüm maddelerin bağlama performansının ölçülmesine yönelik ölçüm yaptığı ifade edilebilir.

Madde geçerliliği: Her bir maddenin, bağlama performansı başarı düzeyini alt grupta ve üst grupta yer alan öğrenciler için istatistiksel olarak ayırt edebilmesi, o maddenin amacına uygun ölçme yaptığı için bir göstergesi olarak kabul edilebilir. Bu durumun saptanması için, ölçme aracı kullanılarak bağlama dersi performansları değerlendirilen 53 öğrenci, başarı düzeyleri bakımından %27’lik alt ve üst gruplara ayrılarak bu gruplar arasında, her bir maddeye ilişkin ölçümler üzerinden İlişkisiz Örneklem t-testi yapılmıştır. Yapılan İlişkisiz Örneklem t-testi sonuçları Tablo 9’da belirtilmiştir.

Tablo 9. Bağlama dersi performansına göre belirlenen %27'lik alt ve üst gruplar arası her bir ölçüt beceri için ilişkisiz örneklem t-testi sonuçları

Ölçüt Beceriler	Grup	n	X	Ss	t	sd	p
1	Alt Grup	14	2.5714	.64	12.612	26	.000*
	Üst Grup	14	4.9286	.26			
2	Alt Grup	14	2.4286	.85	9.815	26	.000*
	Üst Grup	14	4.8571	.36			
3	Alt Grup	14	2.3571	.84	10.202	26	.000*
	Üst Grup	14	4.8571	.36			
4	Alt Grup	14	2.0714	.73	10.146	26	.000*
	Üst Grup	14	4.5000	.51			
5	Alt Grup	14	2.0714	.91	8.392	26	.000*
	Üst Grup	14	4.4286	.51			
6	Alt Grup	14	2.4286	.93	7.000	26	.000*
	Üst Grup	14	4.4286	.51			
7	Alt Grup	14	1.7857	.69	7.450	26	.000*
	Üst Grup	14	3.8571	.77			
8	Alt Grup	14	2.2143	.97	5.339	26	.000*
	Üst Grup	14	4.0000	.78			
9	Alt Grup	14	2.4286	.93	6.378	26	.000*
	Üst Grup	14	4.3571	.63			
10	Alt Grup	14	2.3571	.92	7.071	26	.000*
	Üst Grup	14	4.5000	.65			
11	Alt Grup	14	2.3571	1.00	6.683	26	.000*
	Üst Grup	14	4.5000	.65			
12	Alt Grup	14	1.6429	.74	6.871	26	.000*
	Üst Grup	14	3.9286	.99			
13	Alt Grup	14	2.2143	.89	7.744	26	.000*
	Üst Grup	14	4.5000	.65			
14	Alt Grup	14	2.2143	.97	7.816	26	.000*
	Üst Grup	14	4.6429	.63			
15	Alt Grup	14	2.3571	1.00	7.816	26	.000*
	Üst Grup	14	4.7857	.57			

* $p < .05$

Tablo 9'da görüldüğü üzere ölçme aracında yer alan her bir madde için %27'lik alt ve üst başarı gruplarının başarı puanları arasında anlamlı bir farklılık tespit edilmiştir. Elde edilen p değerleri göz önüne alındığında, geliştirilen ölçme aracının madde geçerliğinin olduğu söylenebilir.

Sonuç ve Öneriler

Sonuçlar

Ölçme aracının güvenilirliğine ilişkin sonuçlar: BPÖA' nın güvenilirliğinin belirlenmesi için, ilk olarak iç tutarlılığı incelenmiştir. İç tutarlılığın incelenmesinde her bir madde üzerinden araştırmacı ve bir alan uzmanı tarafından eş zamanlı yapılan değerlendirmede verilen puanların ortalaması kullanılmıştır. Puanlayıcılar tarafından yapılan puanlamaların ortalamasının alınıp alınamayacağının belirlenmesine yönelik Cohen's Kappa testi uygulanmış ve testin sonucunda puanlayıcılar arası tutarlılığın var olduğu görülmüştür.

Puanlayıcıların verdiği puanların ortalaması üzerinden yapılan Cronbach Alpha testi güvenilirlik testi sonucunda elde edilen tutarlılık kat sayısının kabul edilen minimum değerden oldukça yüksek olduğu saptanmıştır. Elde edilen bu iç tutarlılık kat sayısı, oluşturulan ölçme aracının iç tutarlılığının olduğunun bir kanıtı olarak kabul edilebilir.

Geliştirilen ölçme aracının iç güvenilirliğinin diğer bir sınaması, elde edilen toplam puanlar üzerinden puanlayıcılar arası tutarlılık ölçülerek yapılmıştır. Bu bağlamda, toplam puanlar arası tutarlılığın incelenmesinde kullanılacak korelasyon formülünün belirlenmesi için öncelikle puanlayıcılar arası çarpıklık değerleri incelenmiş ve dağılım normal bulunmuştur. Buna göre, Pearson Momentler Çarpımı Korelasyon katsayısı formülü kullanılmasıyla elde edilen korelasyon katsayısı, ölçme aracının toplam puanlar arası tutarlılığının olduğunun bir göstergesi olarak kabul edilebilir.

Ölçme aracı ile yapılan ölçmelerdeki duyarlılığın incelenmesi için, performans değerlendirme sürecinde ölçme aracı ile yapılan puanlama ile ölçme aracı olmadan geleneksel yolla yapılan puanlama arasında anlamlı bir fark olup olmadığı araştırılmıştır. Yapılan analizin sonuçlarına göre, Ölçme aracı ile elde edilen puanlar ve Ölçme aracı olmadan elde edilen puanlar arasında anlamlı fark tespit edilmiştir. Bu durumda ölçme aracı ile kullanılarak elde edilen puanların ortalamasının yaklaşık 3 puan daha düşük olduğu tespit edilmiştir. P değeri, standart sapma ve puan ortalamaları değerleri dikkate alındığında, puan ortalamaları arasındaki farkın, ölçme aracı kullanılarak yapılan ölçümlerin daha duyarlı olmasından kaynaklandığının göstergesi olarak kabul edilebilir. Tüm bu sonuçlar göz önüne alındığında, araştırma kapsamında geliştirilen ölçme aracının güvenilir ölçümler yaptığı söylenebilir.

Ölçme aracının geçerliğine ilişkin sonuçlar: Araştırma kapsamında geliştirilen ölçme aracının öncelikle kapsam geçerliği alan uzmanlarının da görüşleriyle sağlanmıştır. Bu amaç çerçevesinde, görüşleri alınan alan uzmanları, hazırlanan ölçüt becerilerin ifade, içerik ve puantajı ile sınıf düzeyine göre katsayıları bakımından bağlama icra becerisini ölçmek için yeterli olduğunu belirtmişlerdir.

Ölçme aracının yapı geçerliğinin sınanması için yapılan temel bileşenler analizi sonucunda ölçme aracının bağlama performansına yönelik tek bir temel yapıyı ölçtüğü sonucuna ulaşılmıştır. Elde edilen faktör yükleri incelendiğinde maddelere ait faktör yüklerinin kabul edilebilir sınırdan daha yüksek olduğu saptanmıştır. Bu durum, geliştirilen ölçme aracının bağlama performansının ölçülmesine yönelik ölçüm yaptığının ve aracın yapı geçerliğinin olduğunun bir göstergesi olarak kabul edilebilir.

Madde geçerliğine yönelik yapılan analiz çerçevesinde, BPÖA ile yapılan değerlendirme sonucunda öğrenciler, başarı düzeyleri bakımından %27'lik alt ve üst guruplara ayrılmış ve bu iki gurup arasında her bir ölçüt beceri üzerinden t-testi yapılmıştır. Yapılan t-testi sonucunda ölçme aracında yer alan her bir ölçüt beceri için öğrenci gurupları arasında anlamlı bir farklılık saptanmıştır.

Elde edilen sonuçlar çerçevesinde geliştirilen ölçme aracının madde geçerliğinin olduğu söylenebilir. Tüm bu sonuçlar dikkate alındığında, geliştirilen ölçme aracı geçerli bir ölçme aracıdır denebilir.

Öneriler

Araştırmada elde edilen sonuçlara ilişkin oluşturulan öneriler aşağıda sunulmuştur.

1. Araştırma kapsamında geliştirilen Bağlama performans ölçme aracı, güzel sanatlar liseleri, bireysel çalgı bağlama eğitimi yarıyıl sonu sınavlarında kullanıldığı takdirde yapılacak ölçme işleminin, daha nesnel, tutarlı, geçerli ve güvenilir olacağı göz önünde bulundurulması gereken bir noktadır.
2. Geliştirilen ölçme aracı, öğrenci performansına yönelik geribildirimler sağlaması ve öğrencilerdeki tespit edilen eksikliklerin giderilmesine ile ilgili öğretim elemanları ve öğretmenlere katkı ve kolaylıklar sağlaması açısından önemsenmelidir.
3. BPÖA ile yapılan değerlendirilmelere göre, bağlama öğretim programında gerekli görüldüğü takdirde değişiklikler yapılmasının, verimli bir eğitim süreci için sağlıklı olacağı düşünülmektedir.
4. Geliştirilen ölçme aracı, Güzel Sanatlar Liselerinde yapılan yetenek sınavlarında bağlama performansını ölçmek için kullanılabilir.
5. Geliştirilen ölçme aracı, yükseköğretim kurumlarında verilen bağlama eğitimi için yeniden yapılandırılıp, yarıyıl sonu bağlama sınavlarında da kullanılabilir.
6. Geliştirilen bu performans ölçme aracından yararlanılarak Türk müziği çalgıları için de yeni ölçme araçları oluşturulabilir.
7. Gerekli görüldüğü takdirde geliştirilen ölçme aracı genişletilerek daha fazla öğrenci grubu üzerinde denenebilir.
8. Türk ve batı müziği çalgılarında performansının ölçülmesi için bu ve benzeri ölçme araçlarının geliştirilmesinin, daha nesnel, tutarlı, geçerli ve güvenilir bir sınav değerlendirmesi için verimli olacağı düşünülmektedir.



<http://kefad.ahievran.edu.tr>

Ahi Evran Üniversitesi Kırşehir Eğitim Fakültesi Dergisi

ISSN: 2147 - 1037

ENGLISH VERSION

Introduction

During the historical process of humanity, evolutionary change and development has become the most important thing in line with perceptions, awareness, questioning, production, development, sustainability and inclinations present in nature. The inclination to share information has transformed into integrated rules, even in the most primitive societies. In societies formed by the development of settlement patterns, this information sharing of rules became an educational discipline. After thousands of years of accumulation, education of social humans has become a form of information transfer with fixed behavioral plan. This behavioral plan is complicated, and is transferred to individuals under the general name of “education”. The individual is now a social asset and needs to adopt the behavior modes chosen by society, and when the time comes, pass on the accumulation of knowledge to the next generation. There are a variety of scientific definitions of the education concept with the case indicated by stating it is “the process of planned and desired change, development, and authorization of behavioral targets, and development of techniques with the aim of knowing their laws and principles” (Erdem and Akman, 1996; Tezcan, 1997; Uçan, 2005).

One of the most important stages of education is the measurement dimension. In a broad sense, measurement “can be defined as the process of counting, classifying, grading or numerically expressing in units a certain quality of any object according to certain rules” (Özgüven, 2017, p.36). Currently, measurement has incontrovertible importance in nearly every area in societies in the development process of the sciences mainly. For the individual to gain the targeted planned behavior, measurement plays an important role in revealing how much they have completed the predicted changes, to what level they have adopted them, in identifying what development occurs and making consistent, correct and reliable decisions about both the development of the trainer and the education process. Measurement in this sense appears to be a descriptive process. “Measurement and assessment methods aid the teacher in (1) providing information about the student’s initial behavior, (2) determining appropriate aims in accordance with reality for each student, (3) evaluating the degree of success in achieving these aims and (4) determining, assessing and differentiating educational techniques” (Mehrens and Lehmann, 1991, p. 8). At the same time, when examined from the student’s

perspective, the aims of the teacher have important functions of informing the student, increasing the efficacy and motivation of the student about the education-teaching process, encouraging students correctly about student successes and habits, and providing effective feedback about the student's strong and weak aspects (Mehrens and Lehmann, 1991).

Baglama is used within the execution of the rich structure of Turkish Folk Music (TFM) and is one of the important folk instruments seen as a projection of all types of cultural phenomena within the Anatolian music tradition. Thus, opinion and thought has made baglama a carrying-transmitting element within the traditional music structure. This situation has led to research and studies performed to date in the theoretical field of TFM mainly being completed on a baglama-based framework. The main basis for this topic is thought to be that the baglama is a fretted instrument and is intensely played (Börekci and Nacakçı, 2019). With education-teaching activities within a master-apprentice relationship for many years, it appears research about the baglama has included a wide variety of research about teaching in systematic and academic sense, especially in recent years. These studies at promising levels are present for all stages of baglama education. One of the official organizations where baglama education is given is Fine Arts High Schools (FAHS). Being the formal education period before undergraduate education makes these organizations one of the most important in shaping the student's musical knowledge and accumulation in a basic sense. In this context, the musical performance forming the main topic of the research can be shown to be an important process in FAHS, just as with all types of education stages.

The scientific and arts branch of music involves feelings and thoughts that the individual cannot express in words being expressed within a certain harmony with the aid of notes and sounds. During the music education process, all behavior displayed in relation to instrumental execution and vocalization of an individual are called musical performance. "In this context, the stages of playing and singing involved in a musical performance are the section where the individual displays their psycho-motor skills" (Atak Yayla, 2004). Music is seen as an artistic performance and in terms of the music education process, measurement of this performance is an important element for many reasons. "Musical performance-based teaching and testing of this teaching process on an objective basis is necessary for the process of music teaching" (Schleuter, 1996, p.176). As a result, performance measurement tools are needed to accurately identify the musical development of the student within the education process, to provide attentive systematic feedback to the student and to regularly monitor the development stages. Measurement tools developed in relation to the process of evaluating an individual display some differences according to the quality of the behavior to be measured as cognitive, affective or psycho-motor. According to Sönmez (2001), the psycho-motor area "is the area coding student skills. This area is integrated with the cognitive and affective areas. Behavior learned in relation to the psycho-motor area should be assessed in terms of product or process or both product and process." Within this framework, the target behavior should be measured to the extent to which it

exists and appropriate test situations should be developed to measure the student's development process. In this context, measurement in education and evaluation of data obtained from these measurements provides important contributions to decisions made in relation to the future, organizations needed and curriculums to be prepared, in addition to determining the material needed or that will be needed during the educational process and developing reviews while the process continues (Akçay, 2011).

Problem Statement

In education, assessment with aims like understanding whether a curriculum is reliable or not, identifying the degree of efficacy of methods used in teaching, directing the student towards areas where they will be successful, identifying learning difficulties, identifying student success, and seeing whether measurement and assessment elements function well or not are all based on measurement results" (Baykul, 2000, p.89). "As in music education, especially in tests which are individually performed and where psycho-motor behavior is measured, not identifying the desired behavior, not giving weighted points for this behavior, in short not creating a written measurement tool, leads to measurement results being distant from objectivity" (Dalkıran, 2006, p.14).

Test performance measurements where a written scale tool, without grading or tested efficacy, is not used negatively affects the student receiving any feedback related to the test performance and additionally leads to inability to obtain systematic and healthy information about development. According to Fiske (1983, cited Ciorba and Smith, 2009), many evaluators do not know how they determine a performance grade. Additionally, considering the educational process, it makes it very difficult to identify those who are/are not developing and aspects that are/are not functioning in the teaching dimension. In line with analyses and expert opinions, currently in spite of a variety of studies to develop measurement tools in the music education area, it was identified that no valid and reliable measurement tool was developed for measurement of the most important folk instrument and symbol of Turkish national culture of the baglama. Within the framework of these elements, based on interviews with lecture staff and teachers giving baglama lessons, creation of a new performance measurement tool for use in end-semester exams for individual instrument baglama lessons and assessment of the utility of this tool formed the main topic of this study.

In this context, the problem statement for the research was determined as "is the measurement tool developed for the performance dimension of individual instrument baglama lessons in Fine Arts High Schools linked to the Ministry of National Education (MEB) valid and reliable?".

Sub problems: The subproblems investigated in relation to the aim, importance and problem statement in the research were conceptualized as follows;

3. Is the measurement tool developed in the scope of the research reliable?

4. Is the measurement tool developed in the scope of the research valid?

Aim and Importance of the Research

In this research, the aim was to develop a new measurement tool with high validity and reliability to positively contribute to measurement in education-teaching during end-semester exams for individual instrument baglama lessons in Fine Arts High Schools. This research is important in terms of developing a new measurement tool to contribute to more practical and objective measurement related to end-semester test performance of students in the individual instrument baglama class. It is believed the developed measurement tool will allow more effective assessment of the student by the teacher, allow the teacher to provide more effective guidance due to more efficient feedback to resolve deficiencies, mistakes and difficulties in the learning process, and provide more effective development within the education-teaching process. Additionally, the performance measurement will be convenient due to ensuring consistency in measurement systematics by lecture staff and teachers giving baglama lessons and is considered to be an example for measurement tools to be created in the future in relation to folk instruments.

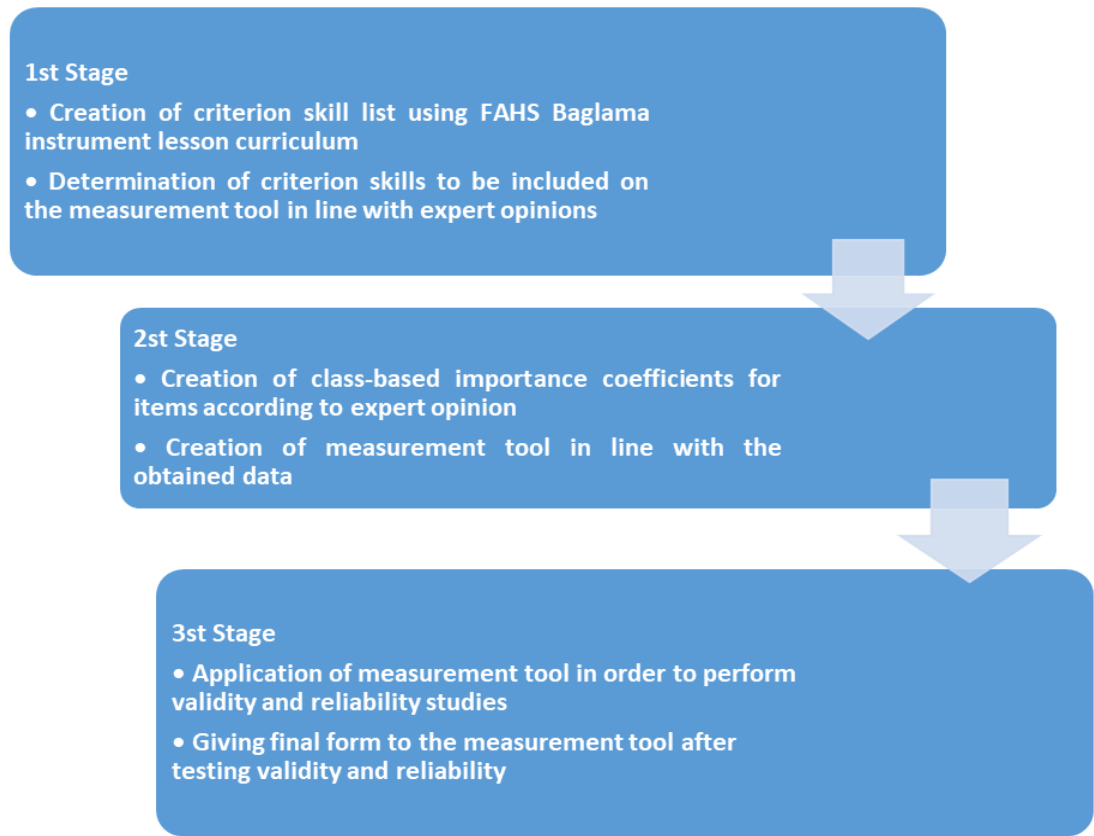
Method

In this section, the research model and pattern, study group, tools used for data collection in the research and development stages of the measurement tool, collection of data, and statistical methods and techniques used for analysis of data are explained in detail.

Model and Pattern of Research

This research has descriptive quality with the aim of developing and using a measurement tool with high validity and reliability in exams performed to assess student performance at the end of the semester in the "Individual Instrument Baglama" lesson in Fine Arts High Schools. A general screening model was used when completing the research. "Screening models are a research approach aiming to describe a situation as it is, whether in the past or still existing" (Karasar, 2009). To create a foundation for the research and an item pool, documentary screening was performed to determine and collect literature related to the problem and the relevant curriculum was investigated and analyzed to obtain data.

The research pattern and stages determining the pattern are as follows;



1st stage: Creation of criterion skill list using FAHS baglama instrument lesson curriculum, determination of criterion skills to be included on the measurement tool in line with expert opinions.

2nd stage: Determination of class-based importance coefficients for items according to opinions of lecture staff and teachers giving baglama lessons and giving final form to the measurement tool before the experimental process.

3rd stage: Testing of validity and reliability status of the measurement tool after trials. At the end of this, the measurement tool was given its final form.

Study Group

The study group of the research consists of students and lecturers. In this context, the information about the students used in the experimental part of the measurement tool and the lecturers whose views are benefited from the structure and content validity part of the criterion skills are presented in Table 1, 2 and 3 in detail. Accordingly, the number of students according to high school and class levels are given below;

Table 1. *Class level and number according to high school of students participating in the research*

School name	Total number of Participating Students	Class levels			
		9	10	11	12
Çankırı FAHS	11	3	4	2	2
Burdur FAHS	9	2	2	3	2
Antalya ATSO FAHS	8	2	2	3	1
Bolu FAHS	9	1	1	3	4
Isparta FAHS	16	4	4	5	3
Total	53				

According to Table 1, 53 students participated in the research. The low number of students is noteworthy considering that the schools cannot even fill 71% of their quota (Aksu, 2014) due to very low current numbers of applications to FAHS requiring special talent.

To determine the items to be included on the measurement tool and importance coefficients according to class level, opinions were sought from lecture staff and teachers giving baglama lessons in a variety of universities and fine arts high schools linked to the MEB. The organizations and demographic information of lecture staff and teachers participating in the survey are presented in Table 2 and Table 3.

Table 2. *Organizations of lecture staff and teachers participating in the survey*

Organization	f	Percentage (%)
Antalya ATSO FAHS	1	10
Çankırı FAHS	1	10
Burdur FAHS	1	10
Ege University	2	20
Muğla Sıtkı Kocaman University	1	10
Mehmet Akif Ersoy University	1	10
Akdeniz University	1	10
Erzurum Atatürk University	1	10
Necmettin Erbakan University	1	10
Total	10	100

According to Table 2, participants completing the survey comprised lecture staff and teachers from a variety of universities and fine arts high schools with 10% from Antalya ATSO FAHS, 10% Çankırı FAHS, 10% Burdur FAHS, 20% Ege University, 10% Muğla Sıtkı Kocaman University, 10% Mehmet Akif Ersoy University, 10% Akdeniz University, 10% Erzurum Atatürk University and 10% Necmettin Erbakan University.

Table 3. Demographic information of lecture staff and teachers participating in the survey

		f	Percentage (%)
Title	Professor	1	10
	Associate Professor	3	30
	Teaching Assistant	3	30
	Teacher	3	30
	Total	10	100
Experience	1-5 years	2	20
	6-10 years	1	10
	11-15 years	3	30
	16-20 years	1	10
	21 years or more	3	30
	Total	10	100
Educational Level	Doctorate	5	50
	Masters	3	30
	Degree	2	20
	Total	10	100

According to Table 3, based on title of baglama lecture staff and teachers participating in the survey comprised 10% professor, 30% associate professor, 30% teaching assistant and 30% teachers. When years of experience are examined, there were 20% with 1-5 years, 10% with 6-10 years, 30% with 11-15 years, 10% with 16-20 years and 30% with 21 years or more. When education level is noted, 50% had doctorates, 30% had masters and 20% had undergraduate degrees. Accordingly, when the title and educational level of participants in the research are noted, it can be said they had sufficient experience.

Data Collection Tools

In the research the curriculum for the Individual Instrument Baglama lesson prepared by the MEB was analyzed and the relevant literature was screened.

1. To determine criterion skills and significance grading included on the developed measurement tool, the “survey form” was applied to the individual instrument baglama lecture staff and teachers

2. Two data collection forms were used including the “Baglama Performance Measurement Tool (BPMT)” developed by the researchers.

Survey form: In the stage of preparing the performance measurement tool, the curriculum for the individual instrument baglama lessons prepared by MEB was investigated with 18 items determined for the criterion skills (behavior) targeted and to be gained during the program. Opinions of experts about the validity of the 18 items and the importance coefficients according to class level were obtained with the prepared Survey Form. In line with these opinions, 15 criterion skills (behavior) were chosen for use on the performance measurement during end-semester exams from the 18 items.

Baglama performance measurement tool: The measurement tool comprises 3 main dimensions of curriculum dimension, exam performance dimension and in-semester status dimension with 4 subdimensions of appropriateness to curriculum requirements, in-semester performance and technical and musical skills.

The dimensions, subdimensions and coefficients in the measurement tool were determined after surveys of relevant experts. The grading of the tool used a 5-point Likert type scale of "1 - very low, 2 – below average, 3 – average, 4 – above average, 5 - very good". Additionally, importance coefficients were added according to class level for each item as 9th class, 10th, class, 11th class and 12th class. Point calculations for students were made based on these importance coefficients according to class level.

With the aim of determining the structural validity of the measurement tool, basic components analysis was applied and at the end of the analysis, the measurement tool was identified to have a single factor structure. Based on basic component analysis results, it was identified that the measurement tool can measure the basic dimension of the targeted baglama performance. The Cronbach alpha reliability coefficient for the measurement tool was determined as .91. The measurement tool appeared to explain 78.83% variance in relation to measurement of baglama performance. In other words, each item and content on the measurement tool serves the general aim of the BPMT at a high rate.

Analysis of Data

Data collected within the framework of the subproblems conceptualized in line with the aim and problem statement of the research were analyzed with IBM SPSS Statistics 22 program. For measurement of statistical significance, .05 was chosen as significance level. The BPMT was applied to a 53-person student group to test the validity and reliability. In this application, points were given separately by the researcher and an expert in the field simultaneously.

Reliability: "Reliability may be defined as the degree of sterility of a scale tool, and hence scale results, from errors" (Bahar et al., 2006). In other words, reliability implies that the scale will provide the same results when repeated with different people at different times and places (Ruiz-Primo and Shavelson, 1996; Johnson, Penny and Gordon, 2009). Firstly, internal consistency was investigated for reliability of the scale tool. For investigation of the internal consistency for points obtained from the assessment, Cronbach's alpha formula was used on the means of points given by the scorers (Büyüköztürk, 2010). Statistically for determination of whether mean points were obtained or not, Cohen's kappa formula was used to calculate the compatibility between two scorers (Kutlu et al., 2010).

For investigation of another type of reliability of the measurement tool, consistency between scorers was examined for total points. To test the sensitivity of the developed measurement tool, the

difference in success level between points when scorers used the BPMT and when they did not (traditional route) was evaluated with the related samples t test.

Validity: “Validity is related to how accurately a test measures the desired characteristic in an individual without confusing it with other characteristics” (Büyüköztürk et al., 2009, p.115). To determine the validity of the BPMT, firstly scope validity was investigated. Within the framework of this aim, opinions were sought from experts with a survey prepared about content, statements, points etc. features of the measurement tool. In terms of the scope of the measurement tool, experts stated it was a valid tool for measurement of exam performance in end-semester exams for baglama lessons. “The ability of a measurement tool to measure the desired behavior in the context of an abstract concept (factor) is called the structural validity” (Büyüköztürk, 2010). To test whether the developed measurement tool had structural validity or not, basic component analysis was performed with the item total correlations examined after factor loading. Item analysis was performed to investigate item validity. With this aim, according to total points obtained from the tool, the unrelated samples t test was used to investigate differences between item mean points in the lower 27% and upper 27% groups.

Ethical Permission for the Research

This study abided by all rules stated in the “Higher Education Institutions Scientific Research and Publication Ethics Directives”. None of the actions stated under the heading “Actions Contravening Scientific Research and Publication Ethics” in the second section of the directive occurred.

Ethics committee permission: data in the research were collected before 2020. Within this framework, there was no ethics committee document obtained related to the research. Contrary to this, data obtained in the research were collected in accordance with regulations without any ethical violations.

Results

In this section, findings obtained from creation of a valid and reliable measurement tool for use in end-semester exams for individual instrument baglama lessons in Fine Arts High Schools, statistical analysis of data collected within the framework of the sub problems determined to test the utility of the measurement and interpretations related to these findings are presented.

Findings and Interpretations related to the first sub problem

Internal consistency: With the aim of testing the internal consistency of points obtained after simultaneous scoring by lecture staff and the researcher, means were obtained for points given to each item by the judges and Cronbach’s alpha internal consistency coefficient was calculated. However, in order to obtain the mean of points given by two judges, it is necessary to examine the consistency between points given for each item. Cohen’s kappa formula was used to statistically calculate this consistency. The results obtained are presented in Table 4.

Table 4. *Kappa coefficients between scorers for each item*

Items	n	Kappa
Item 1	53	.974
Item 2		.949
Item 3		.924
Item 4		.951
Item 5		.926
Item 6		.921
Item 7		.923
Item 8		.924
Item 9		.923
Item 10		.951
Item 11		.899
Item 12		.876
Item 13		.925
Item 14		.875
Item 15		.926

In line with the calculations, the obtained Kappa coefficients are larger than .60 which is accepted as evidence of consistency between scorers (Kutlu et al., 2010). Based on this finding, the mean of points given for each item by judges was taken and Cronbach's alpha internal consistency coefficient was calculated. The Cronbach alpha test results applied to the mean for each item found the reliability coefficient was .91 ($p < .05$, $n=53$). This value shows the BPMT has internal consistency as the value for the performance assessment tool is above the minimum accepted value (.70) (Büyüköztürk, 2010).

Consistency between total points: The consistency correlation between total points was investigated as another way to test the reliability of the BPMT. To determine the correlation formula to apply, the skewness between scorers for total points was investigated. The skewness analysis results are presented in Table 5.

Table 5. *Skewness between scorers for total points*

Judges	n	Skewness
Researcher	53	-.410
Lecture Staff		-.479

When Table 5 is investigated, it was identified the distribution of skewness values between total points was normal. In line with these results, the correlation formula to be applied was determined as the Pearson Moments Multiplication correlation coefficient. The analysis is presented in Table 6.

Table 6. *Pearson moments multiplication correlation coefficients*

Judges	n	$\bar{x}$	ss	p	Correlation
Researcher	53	65,2830	20,51770	.000*	.995
Lecture Staff	53	66,8113	19,97600		

When Table 6 is investigated, according to the Pearson Moments Multiplication correlation coefficient results based on the total scores of 53 students, there was a significant correlation identified

at a rate of “.995” between judges’ scores. This situation can be shown as further proof of consistency between measurements made by judges and that the BPMT is reliable.

Sensitivity: To test whether the developed measurement tool makes sensitive measurements or not, the difference between mean points given by the measurement tool and given by traditional methods (without the measurement tool) was investigated for significance. The related samples t test results for mean points given with the measurement tool and with traditional methods are given in Table 7.

Table 7. *Related samples t test results for points given using the measurement tool and points given with traditional methods*

Measurement method	n	$\bar{x}$	ss	p	t
Measurements with measurement tool	53	66.0472	20.22262	.000*	-10.640
Traditional measurement	53	69.6226	20.17589		

$p < .05$

According to the analysis results, the difference between points given with the measurement tool and without the measurement tool (traditional methods) made on 53 students was identified to be significant ($p < .05$). It appeared that this difference was due to mean points given with the measurement tool being nearly 3 points lower than points given using traditional methods. Considering the p value, standard deviation and mean point values, the difference in mean points is considered to be due to the measurement made with the developed tool being more sensitive. In line with these identifications, the measurement tool can be said to be significant.

Findings and Interpretations related to the Second Sub Problem

Scope validity: With the aim of determining the scope validity of the BMPT, expert opinions were sought about the list of criterion skills comprising the 18 items determined by analysis of the curriculum. The criterion skills determined after analysis of the baglama curriculum and considered necessary for the BPMT were revised in line with opinions about statements, content and points on the measurement tool obtained from experts and reorganized in 15 items. Experts whose opinions were sought stated that the developed measurement tool measured baglama performance within the framework of scope validity.

Structural validity: Basic component (factor) analysis was performed to test the structural validity of the measurement tool developed in the scope of the research. Basic component analysis results are given in Table 8. The KMO-Bartlett test was performed as required for basic component analysis. The KMO test result obtained .915 value and this data shows factor analysis can be used perfectly. When the “p” value is investigated, $p = .000$ and as this value is < 0.05 , it was concluded that basic component analysis should be performed.

Table 8. *Analysis results of basic components*

No	Criterion skills	Factor load value
5	Ability to confidently perform transitions between frets and positions	.885
6	Ability to play using plectrum forms in accordance with notes in the piece	.882
4	Ability to play appropriately with the finger numbering in the piece	.872
3	Ability to use right and left hand associations compatibly	.839
10	Ability to play the piece at appropriate speed for true metronome	.838
9	Ability to play notes in the piece accurately according to duration	.825
1	Ability to hold the baglama in appropriate sitting position, and accurately apply right hand and left hand holding techniques correctly	.824
15	During the semester, ability to behave carefully showing necessary care during baglama lessons, complete homework on time and play correctly	.813
13	Ability to play the piece as a whole	.810
2	Ability to perform upward and downward plectrum strikes cleanly and in balance	.797
11	Ability to play the piece in tempo	.783
14	Ability to play pieces at appropriate level required by the curriculum of the baglama lesson attended	.762
6	Ability to reveal the motifs and phrasing in the piece	.725
7	Ability to play the local execution (manner) features in the piece accurately and fluently	.679
12	Ability to effectively apply the nuances (if present) in the piece	.493
<i>Explained Total Variance</i>		<i>78.838%</i>
<i>Eigen Value</i>		<i>11.826</i>

As seen in Table 8, the BMPT developed within the scope of the research measured the targeted single basic structure. The power of each item in relation to the basic aims of the measurement tool, or in other words, the degree to which they serve the general aims of the measurement tool, is represented by component loads varying from .493 to .885. As the lower limit for factor load values in tests with few items is accepted as .45 (Büyüköztürk, 2010), all factor loads appear to be above this value. It appears the measurement tool explains 78.838% of the total variance related to measurement of baglama performance and the Eigen value was identified as 11.826. In light of these results, all items on the BPMT represent measurements about baglama performance.

Item validity: The ability of each item to statistically differentiate students in the lower group and upper group for baglama performance success is accepted as a marker of whether that item measures appropriately. To identify this situation, 53 students with baglama lesson performance assessed using the scale tool were divided into lower and upper 27% groups, and the unrelated samples t test was performed for measures related to each item between these groups. The results of the unrelated samples t test are given in Table 9.

Table 9. Unrelated samples *t* test results for each criterion skill between lower and upper 27% groups determined with the baglama performance tool

Criterion Skill	Group	n	X	Ss	t	sd	p
1	Lower	14	2.5714	.64	12.612	26	.000*
	Upper	14	4.9286	.26			
2	Lower	14	2.4286	.85	9.815	26	.000*
	Upper	14	4.8571	.36			
3	Lower	14	2.3571	.84	10.202	26	.000*
	Upper	14	4.8571	.36			
4	Lower	14	2.0714	.73	10.146	26	.000*
	Upper	14	4.5000	.51			
5	Lower	14	2.0714	.91	8.392	26	.000*
	Upper	14	4.4286	.51			
6	Lower	14	2.4286	.93	7.000	26	.000*
	Upper	14	4.4286	.51			
7	Lower	14	1.7857	.69	7.450	26	.000*
	Upper	14	3.8571	.77			
8	Lower	14	2.2143	.97	5.339	26	.000*
	Upper	14	4.0000	.78			
9	Lower	14	2.4286	.93	6.378	26	.000*
	Upper	14	4.3571	.63			
10	Lower	14	2.3571	.92	7.071	26	.000*
	Upper	14	4.5000	.65			
11	Lower	14	2.3571	1.00	6.683	26	.000*
	Upper	14	4.5000	.65			
12	Lower	14	1.6429	.74	6.871	26	.000*
	Upper	14	3.9286	.99			
13	Lower	14	2.2143	.89	7.744	26	.000*
	Upper	14	4.5000	.65			
14	Lower	14	2.2143	.97	7.816	26	.000*
	Upper	14	4.6429	.63			
15	Lower	14	2.3571	1.00	7.816	26	.000*
	Upper	14	4.7857	.57			

* $p < .05$

As seen in Table 9, there were significant differences identified for each item on the measurement tool between the success points of the lower and upper 27% success groups. Considering the *p* values obtained, it can be said the developed measurement tool has item validity.

Conclusion and Recommendations

Conclusions

Results related to reliability of the measurement tool: For determination of the reliability of the BPMT, firstly internal consistency was investigated. Investigation of internal consistency used mean points given with an assessment performed simultaneously by the researcher and an expert for each item. Cohen's kappa test was applied to determine whether there were differences in the mean points given by scorers and it was observed there was consistency between scorers at the end of the test.

Cronbach alpha test reliability results for the mean points given by scorers obtained a consistency coefficient that was identified to be very high compared to the minimum accepted value. The internal consistency coefficient obtained is accepted as proof that the created measurement tool has internal consistency.

Further proof of the internal reliability of the measurement tool was obtained by measuring the consistency between scorers in terms of total points. For determination of the correlation formula to be used for investigation of consistency between total points, firstly the skewness values between scorers were investigated and the distribution was found to be normal. Accordingly, the correlation coefficient obtained by using the Pearson moments multiplication correlation coefficient formula was accepted as an indication of consistency between total points on the measurement tool.

For investigation of the sensitivity of measurements performed with the measurement tool, the difference between mean points given using the measurement tool during the performance evaluation process and points given without the tool using traditional methods was investigated for significance. According to the analysis results, a significant difference was identified between points given with the measurement tool and points given without it. This difference was due to mean points given by the scale tool being 3 points lower. Considering the p value, standard deviation and mean point values, the difference in mean points was accepted as being due to higher sensitivity of measures performed with the measurement tool. Considering all these results, it can be said that the measurement tool developed in the scope of the research provides reliable measurements.

Results related to the validity of the measurement tool: Within the scope of the research, opinions were sought from experts for scope validity initially. Within the framework of this aim, experts providing opinions stated that the expression, content and pointing of criterion skills included on the measurement scale and coefficients according to class level were sufficient in terms of measuring baglama execution skills.

The results of basic components analysis performed to test the structural validity of the scale tool concluded the measurement tool measured baglama performance with a single basic structure. When the obtained factor loads are investigated, it was identified that the factor loads were higher than the acceptable limits. This situation is accepted as showing the measurement tool measured baglama performance and the structure of the tool was valid.

To test the item validity, at the end of the assessment using the BPMT students were separated into lower and upper 27% groups in terms of success levels and the t test was performed for each criterion skill between these two groups. The t test results identified significant difference between each criterion skill on the measurement tool for the student groups. The obtained results are evidence that the developed measurement tool has acceptable item validity. When all these results are considered, it can be said the measurement tool developed in the scope of the research is valid.

Recommendations

According to the results obtained in the research, the following recommendations are made;

1. It is necessary to note that if the baglama performance measurement tool developed in the scope of the research is used for end-semester exams in fine arts high schools for individual instrument baglama education, the measurement process will be more objective, consistent, valid and reliable.
2. The developed measurement tool is important in terms of contribution and convenience for lecture staff and teachers to ensure feedback about student performance and resolve deficiencies of the student.
3. According to assessment with the BPMT, if changes are made as necessary to the baglama curriculum it is considered to contribute to the health of an efficient education process.
4. The developed measurement tool can be used to measure baglama performance for talent exams performed in Fine Arts High Schools.
5. The developed measurement tool can be restructured for baglama education given in higher education organizations and used for end-semester baglama exams.
6. New measurement tools can be created based on this performance scale tool for other Turkish music instruments.
7. If considered necessary, the developed scale tool can be expanded and used for more student groups.
8. Development of this and similar scale tools to measure instrument performance in Turkish and western music is thought to be productive for more objective, consistent, valid and reliable test assessment.

References

- Akçay, Ş. Ö. (2011). *Gitar eğitiminde performans ölçeği geliştirme çalışması*, Yüksek Lisans Tezi, Atatürk Üniversitesi, Eğitim Bilimleri Enstitüsü, Erzurum.
- Aksu, C. (2014). Güzel Sanatlar Liseleri Müzik Bölümlerine Başvuran Öğrenci Sayılarının İncelenmesi: Karamsar Bir Tablo, *İdil Sanat ve Dil Dergisi*, 3(13), 87-100.
- Atak Yayla, A. (2004). *Müziksel performansın ölçülmesi*. 1924-2004 Musiki Muallim Mektebinden Günümüze Müzik Öğretmeni Yetiştirme Sempozyumu, Isparta.
- Bahar, M., Nartgün, Z., Durmuş, S. and Bıçak, B. (2006). *Geleneksel-alternatif ölçme ve değerlendirme*, Pegem Yayıncılık : Ankara.
- Baykul, Y. (2000). *Eğitimde ve psikolojide ölçme: klasik test teorisi ve uygulaması*, Cem Web Ofset: Ankara.
- Börekci, A. and Nacakçı, Z. (2019). Bağlama’ da transpoze icra sırasında 24 ton eşit temperaman (24-tet) sisteminin uygulanabilirliği. in. Sağır, Arapgirlioğlu, Akgül ve Sürmeli. (Ed.), *Uluslararası sanad kongresi bildiri kitabı* (pp. 15-25). Ankara: Gece Akademi.
- Büyüköztürk, Ş., Çakmak, E. K., Akgün, Ö. E., Karadeniz, Ş. and Demirel, F. (2009). *Bilimsel araştırma yöntemleri*, Pegem Akademi: Ankara.
- Büyüköztürk, Ş. (2010). *Sosyal bilimler için veri analizi el kitabı*. Pegem Akademi: Ankara.
- Ciorba, C. R., and Smith, N. Y. (2009). Measurement of instrumental and vocal undergraduate performance juries using a multidimensional assessment rubric, *Journal of Research in Music Education*, 57(1), 5-15.
- Dalkıran, E. (2006). *Keman eğitiminde performansın ölçülmesi*. Doktora Tezi, Gazi Üniversitesi, Eğitim Bilimleri Enstitüsü, Ankara.
- Erden, M. and Akman, Y. (1996). *Eğitim psikolojisi*, Arkadaş Yayınevi: Ankara.
- Johnson, R. L., Penny, J. A. and Gordon, B. (2009). *Assessing performance*. Guilford Press: New York.
- Karasar, N. (2009). *Bilimsel araştırma yöntemi*. Nobel Yayın Dağıtım: Ankara.
- Kutlu, Ö., Doğan, D.C. and Karakaya, İ. (2010). Öğrenci başarısının belirlenmesi: Performansa ve portfolyoya dayalı durum belirleme. Pegem Akademi: Ankara.
- Mehrens, W., Lehmann, A. and Irvin J. (1991). *Measurement and evaluation in education and psychology*. Fourth Edition, Harcourt Brace Collage Publishers, USA.
- Özgüven, İ. E. (2017). *Psikolojik testler*. (14. Ed.) Nobel Yayıncılık: Ankara.
- Ruiz-Primo, M. and Shavelson, R. (1996). Rhetoric and reality in science performance assessment. *Journal of Research in Science Teaching*, 33(10), 1045-1063.

- Schleuter, S. L. (1996). *A sound approach to teaching instrumentalist*. Second Edition, Schirmer Books, An Imprint of Simon and Schuster Macmillan, Prentice Hall International: New York.
- Smith, N. Y. (2006). *The effect of learning and playing songs by ear on the performance of middle school band students* (Doctoral dissertation, University of Hartford). Dissertation Abstracts International, 47, 1230A.
- Sönmez, V. (2001). *Program geliştirmede öğretmen el kitabı*. Anı Yayıncılık: Ankara.
- Tezcan, M. (1997). *Eğitim sosyolojisi*. Ferhat Matbaacılık: Ankara.
- Uçan, A. (2005). *Müzik eğitimi*. Önder Matbaacılık: Ankara.