

Sınav Stresi Ölçeğinin Türkçeye Uyarlanması ve Ölçme Değişmezliğinin İncelenmesi

Adaptation of the Examination Stress Scale into Turkish and Examination of Measurement Invariance*

Büşra KARADUMAN**

Sevilay KİLMEN***

Öz

Bu çalışmada Sung ve Chao (2015) tarafından geliştirilmiş olan Sınav Stresi Ölçeğinin (SSÖ) Türkçeye uyarlanması yapılmış, ilgili ölçekten elde edilen ölçümlerin geçerlik ve güvenilirliğine ilişkin kanıtlar elde edilmeye çalışılmış ve ölçme değişmezliği test edilmiştir. Araştırmanın çalışma grubunu 2016-2017 eğitim-öğretim döneminde Balıkesir ilindeki liselerde öğrenim gören 1617 lise öğrencisi oluşturmaktadır. Araştırmanın ilk aşamasında yapılan doğrulayıcı faktör analizi sonucunda ölçeğin üç faktörlü yapısının doğrulandığı sonucuna ulaşılmıştır. Araştırmanın diğer aşamasında ise ölçeğin cinsiyet, okul türü ve sınıf düzeyi alt gruplarında ölçme değişmezliği incelenmiştir. Ölçeğin faktör yapısına ilişkin tanımlanan ölçme modelinin cinsiyet, okul türü ve sınıf düzeyi alt gruplarında ölçme değişmezliğinin sağlandığına ilişkin kanıtlar elde edilmiştir. Bu durum Sınav Stresi Ölçeğinin (SSÖ) cinsiyet, okul türü ve sınıf düzeyi alt gruplarında geçerli ve güvenilir ölçümler sağladığına işaret etmektedir.

Anahtar Kelimeler: Çoklu grup doğrulayıcı faktör analizi, ölçme değişmezliği, sınav stresi ölçeği

Abstract

In this study, Examination Stress Scale developed by Sung and Chao (2015) was adapted into Turkish, the proofs for validity and reliability were collected for the measures obtained from this scale, and measurement invariance was tested. The study group consists of 1617 students in the 9th, 10th, 11th and 12th grades of science high school, anatolian high school, technical high school and social sciences high school in Balıkesir in 2016-2017 education period. According to the confirmatory factor analysis conducted in the first phase of this research, it was found that the scale had three-factor structure. In the second phase of the study, the measurement invariance of the examination stress model was examined across the gender, school type and grade. Findings show that the Examination Stress Scale provides valid and reliable measures at different levels in sub-groups of gender, school type, and grade.

Keywords: Multi-group confirmatory factor analysis, measurement invariance, examination stress scale

GİRİŞ

Eğitim ve öğretim faaliyetleri içerisinde ölçme ve değerlendirme alanına hizmet eden sınavlar bireyin eğitim yaşantılarının her basamağında bulunmaktadır. Sözü edilen bu değerlendirilme durumunun bireyler üzerinde strese yol açtığı çeşitli araştırmalarla ortaya konmuştur (Arslan, 2016). Özellikle sınavlara gelecekteki kariyerlerini belirlemek için giren lise öğrencileri arasında sınav

*Bu çalışma, ilk yazarın, ikinci yazar danışmanlığında 2017 yılında tamamladığı "Sınav Stresi Ölçeğinin Uyarlanması ve Ölçme Değişmezliğinin İncelenmesi" isimli yüksek lisans tezinden üretilmiştir.

**Doktora öğrencisi, Gazi Üniversitesi, Eğitim Bilimleri Enstitüsü, Ankara-Türkiye, e-posta: busra_karaduman@yahoo.com ORCID ID: orcid.org/0000-0001-7565-1025

***Doç. Dr., Abant İzzet Baysal Üniversitesi, Eğitim Fakültesi, Bolu-Türkiye, e-posta: sevilaykilmen@ibu.edu.tr, ORCID ID: orcid.org/0000-0002-5432-7338

stresi önemli bir konudur (Sung ve Chao, 2015). Başarı odaklı toplumlarda bireylerin sosyal ortamlarda sıklıkla değerlendirilmeye maruz kalmaları, stresin ortaya çıkmasında temel faktör olarak rol oynayabilmektedir (Arslan, 2016). Toplumun öğrenciler üzerindeki başarı odaklı zorlayıcı tutumu öğrenciler üzerinde stres ve kaygıyla sonuçlanmakta, bu durum kişinin duygusal ve akademik durumunu etkilemektedir. Bireyin yaşadığı sınav stresinin bireyin sağlığı üzerindeki olumsuz etkileri çeşitli araştırmalarda (Costarelli ve Patsai, 2012; Steptoe, Wardle, Pollard, Canaan ve Davies, 1996) rapor edilmiştir.

Kavramsal olarak dünyada sınavlara yönelik yapılan çalışmalar incelendiğinde sınav stresinden çok sınav kaygısı üzerine odaklanıldığı göze çarpmaktadır. Oysa ki, stres, fiziksel ve duygusal yüklenmeler sonucu meydana gelen zorlanmalardır ve meydana gelen gerilimler ile birlikte stres bazen istenmedik tepkileri meydana getirir. Bu tepkilerden sadece bir tanesi kaygıdır (Altuntaş, 2003; Sarason ve Sarason, 1990). En tipik kaygı tepkileri terleme, korku, endişe gibi uyarılmalarıdır. Birçok araştırmacı tarafından kaygı, stres tepkisinin bir boyutu olarak tanımlanmıştır (Sung ve Chao, 2015).

Türkiye’de sınavlara yönelik çalışmalar incelendiğinde sınav kaygısına odaklanıldığı, sınav stresine yer verilmediği görülmektedir. Örneğin, Pazarlı (2009) araştırmasında, öğrencilerin öğrenme stilleri ile sınav kaygıları arasındaki ilişkiyi araştırmıştır. Çapulcuoğlu ve Gündüz (2012) ise araştırmalarında, sınav kaygısı, öğrenci tükenmişliği, akademik yetkinlik ve anne-baba tutumları değişkenleri arasındaki ilişkiyi incelemiştir. Bazı araştırmalarda da çeşitli değişkenlerin sınav kaygısı üzerindeki etkileri araştırılmıştır. Örneğin, Ulusoy, Yavuz, Esen, Umut ve Karatepe (2016) araştırmalarında, bilişsel müdahalelerin davranışçı müdahalelerden bağımsız olarak sınav kaygısı üzerindeki etkililiğini araştırmışlardır. İlgili alan yazın genel olarak incelendiğinde, Türkiye’de sınav kaygısına ilişkin hem tarama hem de deneme türü araştırmaların yer aldığı ancak sınav stresi kavramına hiç değinilmediği göze çarpmaktadır. Bu araştırmada alan yazındaki bu boşluğa odaklanılmış, Sung ve Chao (2015) tarafından geliştirilen SSÖ’nün Türkçeye uyarlaması yapılmış, ardından farklı gruplarda ölçme değişmezliği incelenmiştir.

Bu araştırmanın yapılması en az iki nedenden dolayı önemlidir. Birincisi, Türkiye’de sınav stresine ilişkin araştırmalara rastlanmamış olmasıdır. Bu durumun olası nedenlerinden biri sınav stresini ölçmeye yönelik Türkiye’de geliştirilen veya Türkçeye uyarlaması yapılan herhangi bir ölçeğin olmayışı olabilir. Sınav stresi ölçeğinin geliştirilmesi veya başka bir kültürde geliştirilen sınav stresi ölçeğinin Türk kültürüne uyarlanması Türkiye’de sınav stresine ilişkin araştırmalara katkı sağlayabilir. Nitekim öğrencilerin sınav stres düzeyinin ölçülmesi için bir araç geliştirilmesi, sınav stresi ile ilgili konuları araştırmaya yönelik ilk adımdır (Sung ve Chao, 2015). İkincisi ise uyarlanan SSÖ’ye dayalı olarak yapılacak grup karşılaştırmaları için ön koşul olan ölçme değişmezliğine kanıt aranmasıdır. Araştırmacılar psikolojik ölçümlerin geçerlik ve güvenilirliklerine dair kanıtlar elde ettiklerinde yapmış oldukları karşılaştırmalara güven duymaktadırlar. Oysa geçerlik ve güvenilirlik ölçme aracına dayanmaktan ziyade araçtan elde edilecek ölçümlere dayanmaktadır. Bu sebepten geçerlik ve güvenilirlik düzeyi ile ilgili bilgiler edinmek için hesaplanan test ve madde istatistikleri sadece grupta yer alan bireylerin özelliklerini yansıtmaktadır (Crocker ve Algina, 1986). Buradan farklı gruplarda yapılan ölçümlere ait geçerlik ve güvenilirlik kanıtlarının farklılaşabileceği anlaşılmaktadır. Elde edilen ölçümlerin psikometrik nitelikleri, bireylerin farklı özellikler taşımasından kaynaklanabilir ya da ölçme aracı bu farklılaşmaların sebebi olabilir. Bu sınırlılık nedeniyle, gözlenen değişkenlerin hangi durumlarda gruplar arasında geçerli ve güvenilir olduğu yönünde bir sorun gündeme gelmiş ve alan yazında ölçme değişmezliği çalışmaları yapılmaya başlanmıştır (Vandenberg ve Lance, 2000). Bu araştırmada SSÖ için farklı gruplarda ölçme değişmezliği çalışması yapılarak ölçme aracının gruplar arası karşılaştırmalar için uygun olup olmadığı test edilmiştir. Aşağıda konuyla ilgili kavramsal çerçeveye değinilmiştir.

Stres

Stres, bedenin kendi üstündeki baskıya gösterdiği “genel uyum sendromu” adı verilen tepkidir (Selye, 1986). Selye (1973)’ye göre stres üç aşamada gerçekleşmektedir. İlk aşama organizmanın tehlikeyi hissettiği andaki durumunu anlatan alarm aşamasıdır. İkinci aşama, tehlike organizma için

devam ederken, organizmanın fiziksel olarak tepkide bulunduğu direnç aşamasıdır. Tehlikenin geçmediği durumlarda organizmanın tükendiği aşama ise tükenme aşamasıdır.

Organizmanın bedensel ve ruhsal olarak algıladığı tehditler sonucunda stres meydana gelir. Bu durum organizmada bedensel ve psikolojik olarak değişikliklere sebep olmaktadır. Stres belirtileri kendi içinde dörde ayrılmaktadır:

Duygusal (psikolojik) belirtiler

Organizma stresin neden olduğu etkilere karşı bazı duygusal tepkiler verir. Rowshan (2000) duygusal stres belirtilerine; aşırı ağlama, depresyon, duygu değişikliği, hastalıklı gibi hissetme, kabus görme, kızgınlık, sinirsel gülmeler ve üzüntü örneklerini vermiştir.

Fiziksel belirtiler

Rowshan (2000) bir bireyde stresin başlayıp başlamadığının anlaşılması için fiziksel belirtilerin gözlenmesi gerektiğini ileri sürmektedir. Fiziksel stres belirtileri için solunum hızlanması, kalp atışı ve kan basıncının yükselmesi, kasların gerilmesi, göz bebeklerinin büyümesi, baş ağrısı, mide bulantısı, göğüste ağrı, yorgunluk, uykusuzluk, zararlı alışkanlıklara yönelme, iştah bozukluğu gibi belirtiler örnek olarak verilebilir (Baltaş ve Baltaş, 1999; Hançerlioğlu, 1988; Köknel, 1998).

Zihinsel belirtiler

Bireyin verdiği duygusal tepkilerin yanı sıra, birey ciddi stres nedenleriyle karşılaştığında bilişsel bozukluklar yaşayabilir (Atkinson, Atkinson, Smith ve Bem, 1999). Zihinsel stres belirtileri için iş kalitesindeki düşüş, hatalarda artış örnek olarak verilebilir (Braham, 1998).

Sosyal belirtiler

Stres bireyin çevresindeki insanlarla ilişkilerini etkileyebilir ve bu durumda stresin sosyal belirtileri ortaya çıkar. Sosyal stres belirtileri için topluma uyum sağlayamamak, ben merkezli olmak, yalnızlık hissi, insanlara karşı hoşgörülü olamamak, insanlarla kurulan iletişimde meydana gelen bozukluklar (Rowshan, 2000), işe devamsızlık gibi (Sabuncuoğlu ve Tüz, 1995) örnekler verilebilir.

Sınav stresi, sınavlarla ilgili olarak yaşanan stresi tanımlama çabasıdır. Bu nedenle özellikleri sınavlara özgü olarak tanımlanmıştır. Aşağıda sınav stresi ile ilgili genel bilgiler verilmektedir.

Sınav Stresi

Günümüzde yaşamın her döneminde, bireyler problemlerle veya durumlarla karşılaşmakta, sık sık stres kavramı dile getirilmekte ve bu kavram pek çok alanda kullanılmaktadır (Eryılmaz, 2009). Kullanılan alanlardan biri de eğitimidir. Eğitim bireyin doğduğu andan başlayarak yaşam boyu devam eden bir süreçtir. Genel olarak eğitimde stresin öğrenciler üzerinde çabuk sinirlenme, fazla uyuma, tahammül edememe, yalnızlık, okul başarısında düşme, arkadaşları ile iletişimde meydana gelen kopukluklar, öfke patlamaları gibi sonuçları olabilmektedir (Motavallı, 1997). Sınav stresi eğitim öğretim ortamlarında yaşanan stresin özel bir şeklidir. Sung ve Chao (2015), sınav stresini “fizyolojik kaygı tepkileri”, “bilişsel ve davranışsal tepkiler” ve “algılanan sosyal beklenti ve sosyal kıyas” olmak üzere üç boyutta ele almaktadırlar.

Fizyolojik kaygı tepkileri boyutu fiziksel hastalık veya rahatsızlık, uyku bozuklukları ve duygusal problemler olarak tanımlanmıştır. Bilişsel ve davranışsal tepkiler boyutu sınav stresiyle tetiklenen düşünceler ve davranışlar (örneğin; sınav puanlarına ilişkin endişeler ve kendilerini yoğun bir şekilde çalışmaya zorlamak) olarak ele alınmıştır. Algılanan sosyal beklenti ve sosyal kıyas boyutu birey

tarafından algılanan sosyal beklentiler ve kişilerarası karşılaştırmayı ele almaktadır (Sung ve Chao, 2015).

Ölçme Değişmezliği

Klasik test kuramında (KTK) hesaplanan test ve madde istatistikleri uygulandığı gruptan etkilenmektedir. Araştırmacılar, aynı ölçme aracını farklı gruplar üzerinde uygulayabilirler ve cinsiyet, okul türü gibi demografik özelliklerin etkilerinin ortadan kaldırılmadığı durumlarda elde edilen sonuçlar yorumlanırken hatalar yapılabilir. Bu durum KTK'nın bir sınırlılığıdır. Bu sınırlılık günümüzde de önemli bir konu olan ölçme değişmezliğini gündeme getirmiştir (Crocker ve Algina, 1986). Ölçme değişmezliği, bir ölçüğe ait özelliklerin farklı gruplarda değişmez olup olmadığı ile ilgilenir.

Karşılaştırmalarda gizil/örtük değişkenler olarak adlandırılan bilişsel yetenekler, kişilik gibi özellikler incelenmektedir (Somer, Korkmaz, Dural ve Can, 2009). Bireysel farklılıklardan hareketle, ölçme sonuçlarında gözlemlenen farklılığın sadece bireyin özelliğinden kaynaklandığını söylemek doğru değildir çünkü ölçümlerdeki farklılıklar ölçme aracının kendisinden de kaynaklanıyor olabilir (Cheung ve Rensvold, 2002).

Bu araştırmada ölçme değişmezliği, literatürde yaygın olarak kullanılan yapısal değişmezlik, metrik değişmezlik, ölçek değişmezliği ve katı değişmezlik olmak üzere dört ayrı hipotezin test edilmesi ile yürütülmüştür. Bahsedilen dört aşama (Milfont ve Fischer, 2010) aşağıda açıklanmaktadır:

Yapısal değişmezlik (configural invariance)

Ölçme değişmezliği çalışmalarının en temel düzeyidir ve daha sınırlayıcı modellerin karşılaştırılması için test edilmesi gerekmektedir (Wu, Li ve Zumbo, 2007). Ölçme değişmezliği hipotezler geliştirilerek aşamalı olarak test edilmektedir (Steenkamp ve Baumgartner, 1998). Test edilen ilk hipotez, psikolojik ölçme aracının faktör yapısının gruplar arasında değişmez olduğudur. Yapısal değişmezliğe ilişkin kanıt elde edilirse “gruplar arasında kavramsal yapı aynıdır ve maddeler gruplar arasında aynı psikolojik yapıyı ölçüyor” denilebilir. Bu durum psikolojik ölçümler için yapı geçerliğine ilişkin kanıtlar sağlamaktadır (Vandenberg ve Lance, 2000). Eğer yapısal değişmezliğe ilişkin kanıt sağlanamazsa ölçülen yapılar gruplar arasında farklılaştığı için grup farklılıkları testleri anlamlı olmayacaktır. Yapısal değişmezliğe ilişkin kanıtlar sağlandıktan sonra ölçme model parametrelerinin değişmezliğinin yorumlanabilmesi için bir sonraki aşama olan metrik değişmezliğe ilişkin hipotez test edilir (Vandenberg ve Lance, 2000).

Metrik değişmezlik (metric invariance)

Metrik değişmezlik farklı grupların maddeleri aynı şekilde yanıtlayıp yanıtlamadıklarını test eder. Diğer bir deyişle, belli ölçek maddeleri ve bu maddelerin temelini oluşturan yapılar arasındaki ilişkinin gücünün aynı olup olmadığını sınar. Metrik değişmezlik faktör yüklerinin (λ) gruplar arasında aynı olacak şekilde sınırlandırılmasıyla test edilir (Milfont ve Fischer, 2010). Bollen (1989) faktör yüklerinin gözlenen değişkenleri gizil/örtük değişkenlere bağladığını ve bu nedenle gizil/örtük değişkendeki en küçük değişimin, gözlenen değişkeni etkilediğini belirtmiştir. Eğer metrik değişmezliğe ilişkin kanıt elde edilemezse, maddelerin tüm gruplar için anlamlarının aynı olmadığına yani maddelerin bir ya da birden fazla gruba karşı yanlı davrandığına ilişkin yorum yapılabilir.

Ölçek değişmezliği (scalar invariance)

Ölçek değişmezliği gizil/örtük ortalamaların karşılaştırılmasını gerektirir. Ölçek değişmezliğinin sağlanması gözlenen puanların ve gizil/örtük puanların ilişkili olduğunu göstermektedir. Diğer bir deyişle, gizil/örtük yapıda aynı puanı alan bireylerin –hangi gruba mensup olup olmadığına

bakılmaksızın- gözlenen değişkenler için de aynı puanı alacağı anlamına gelmektedir. Ölçek değişmezliği sabitin gruplar arasında aynı olacak şekilde sınırlandırılmasıyla test edilir (Milfont ve Fischer, 2010).

Katı değişmezlik (strict invariance)

Hata varyansları, doğrudan gözlemlenemeyen gizil/örtük değişkenlerin altında yatan değişkenliğin açıklanamayan kısmıdır ve gözlenen değişkenler arasındaki korelasyon büyüklüğünü etkilemektedirler. Katı değişmezlik aşamasında gruplar arasında ölçüğe ait hata varyanslarının değişip değişmediği test edilir. Burada önceki değişmezlik aşamalarındaki faktör yapısı, faktör yükleri ve madde sabitlerine ek olarak, hata varyanslarının da gruplar arası aynı olduğu sınırlaması getirilmektedir (Hirschfeld ve von Brachel, 2014; Milfont ve Fischer, 2010).

Araştırmanın Amacı

Bu araştırmanın amacı Sung ve Chao (2015) tarafından geliştirilen SSÖ'nün Türkçeye uyarlamasını gerçekleştirdikten sonra bu ölçme aracının faktör yapısı için oluşturulan ölçme modelinin veri ile uyum düzeyini inceleyip, aracın farklı gruplarda (cinsiyet, okul türü ve sınıf düzeyi) aynı yapıyı ölçüp ölçmediğini diğer bir deyişle gruplar arası ölçme değişmezliğini belirlemektir. Bu genel amaca ulaşmak için dört araştırma sorusu oluşturulmuştur:

1. SSÖ'nün üç faktörlü yapısı doğrulanmakta mıdır?
2. SSÖ'nün cinsiyet açısından ölçme değişmezliği sağlanmakta mıdır?
3. SSÖ'nün okul türü açısından ölçme değişmezliği sağlanmakta mıdır?
4. SSÖ'nün sınıf düzeyi açısından ölçme değişmezliği sağlanmakta mıdır?

YÖNTEM

Çalışma Grubu

Araştırmanın çalışma grubunu 2016-2017 eğitim-öğretim yılının güz döneminde Balıkesir ilindeki fen lisesi, anadolu lisesi, mesleki ve teknik anadolu lisesi ve sosyal bilimler liselerinde 9., 10., 11. ve 12. sınıfa devam eden 1617 öğrenci oluşturmuştur. Tablo 1'de öğrencilerin cinsiyet, sınıf düzeyi ve okul türlerine göre dağılımları verilmiştir.

Tablo 1. Çalışma Grubundaki Öğrencilerin Cinsiyet, Okul Türü ve Sınıf Düzeylerine göre Dağılımı

		9. sınıf		10. sınıf		11. sınıf		12. sınıf		Toplam
		E	K	E	K	E	K	E	K	
Okul türü	Fen lisesi	45	61	38	57	36	63	48	33	381
	Anadolu lisesi	58	52	53	81	80	87	37	46	494
	Mesleki ve teknik anadolu lisesi	96	40	61	73	48	80	36	52	486
	Sosyal bilimler lisesi	35	74	27	64	-	-	25	31	256
	Genel toplam	461		454		394		308		1617

E: Erkek K: Kız

Çalışma grubuna 894 (%55.3) kız öğrenci ve 723 (%44.7) erkek öğrenci dahil edilmiştir. Tablo 1'de görüldüğü gibi, öğrencilerin 381'i (%23.6) fen lisesinde, 494'ü (%30.5) Anadolu lisesinde, 486'sı (%30) mesleki ve teknik anadolu lisesinde ve 256'sı (%15.8) da sosyal bilimler lisesinde öğrenim görmektedir. Öğrencilerin sınıf düzeylerine ilişkin dağılımları incelendiğinde 9. sınıflardan 461

(%28.5), 10. sınıflardan 454 (%28.1), 11. sınıflardan 394 (%24.4) ve 12. sınıflardan 308 (%19) öğrencinin çalışmaya dahil edildiği görülmektedir.

Veri Toplama Aracı

Araştırmada Sung ve Chao (2015) tarafından geliştirilmiş olan SSÖ veri toplama aracı olarak kullanılmıştır. SSÖ beşli likert tipinde (hiç katılmıyorum:1, tamamen katılıyorum:5), 27 maddeden oluşmaktadır. Ölçekte sınav stresi; fizyolojik kaygı tepkileri, bilişsel ve davranışsal tepkiler ve algılanan sosyal beklenti ve sosyal kıyas şeklinde toplam üç boyutta ele alınmaktadır.

Fizyolojik kaygı tepkileri: Ölçekte fizyolojik kaygı tepkileri boyutu fiziksel hastalık veya rahatsızlık, uyku bozuklukları ve duygusal sıkıntı gibi tepkiler olarak tanımlanmıştır (Örnek: Sınavlara hazırlanırken, kendimi çoğunlukla fiziksel olarak rahatsız hissedirim). Bahsedilen boyutta 10 madde yer almaktadır. Bu boyuta ait Cronbach alfa değeri 0.89 olarak bulunmuştur.

Bilişsel ve davranışsal tepkiler: SSÖ'de bilişsel ve davranışsal tepkiler boyutu sınav stresiyle tetiklenen düşünceler ve davranışlar olarak tanımlanmıştır (Örnek: Tüm sınavlardan aldığım puanları önemserim). Bahsedilen boyutta 8 madde yer almaktadır. Bu boyuta ait Cronbach alfa değeri 0.85 olarak bulunmuştur.

Algılanan sosyal beklenti ve sosyal kıyas: Ölçekte algılanan sosyal beklenti ve sosyal kıyas boyutu öğrenciler tarafından algılanan beklentiler ve başkalarıyla karşılaştırılmaya ilişkin olarak tanımlanmıştır (Örnek: Sınav puanlarım ile ilgili olarak ailemin beklentileri beni rahatsız eder). Bahsedilen boyutta ise 9 madde yer almaktadır. Bu boyuta ait Cronbach alfa değeri 0.88 olarak bulunmuştur (Sung ve Chao, 2015).

Verilerin Analizi

Ölçeğin dilsel eşdeğerliğini incelenmek amacıyla ölçeğin orijinal formu ile Türkçe formundan elde edilen puanlar arasındaki ilişki hesaplanmıştır. Toplam 50 kişiden elde edilen verilerle ölçeğin toplam puan ve alt boyutları için Pearson Momentler Çarpım Korelasyonu hesaplanmış ve bağımlı örneklem *t*-testi yapılmıştır. Ölçümlerin güvenilirliğine ilişkin kanıtlar elde etmek için Cronbach alfa değeri hesaplanmıştır. Ölçekteki maddelerin faktör yükleri ve hata varyanslarının değiştiği göz önüne alınarak çok boyutlu ölçekler için hesaplanan ve Cronbach alfa katsayısına göre daha güçlü bir güvenilirlik değeri olarak belirtildiği için bileşik güvenilirlik katsayısı (Raykov, 1997) da ek olarak hesaplanmıştır. Geçerliğe ilişkin kanıtlar elde etmek için ise doğrulayıcı faktör analizi (DFA) kullanılmıştır. Bu analiz için LISREL 8.7 paket programından yararlanılmıştır. Model veri uyumunu kestirebilmek için Ki kare, Ki kare/sd, RMSEA, CFI, NFI, IFI ve NNFI model uyum indeksleri kullanılmıştır (Hu ve Bentler, 1999). Ölçme değişmezliği ise çoklu grup doğrulayıcı faktör analizi (ÇGDFA) ile incelenmiştir (Jöreskog ve Sörbom, 1993). Ölçme değişmezliği ile ilgili analiz sonuçlarının doğru bir şekilde yorumlanabilmesi için verilerle analize geçilmeden önce varsayımların karşılanıp karşılanmadığına yönelik bazı ön analizlerin yapılması gerekmektedir. Bu nedenle araştırmada kayıp değerler, uç değerler ve normalliğe bakılmış ayrıca çoklu bağlantı sorunu da incelenmiştir. Veriler varsayımları karşıladığı için tahmin yöntemi olarak en çok olabilirlik yöntemi kullanılmıştır. ΔCFI ve $\Delta RMSEA$ (Chen, 2007; Cheung ve Rensvold, 2002) ise farklı gruplarda ölçme değişmezliğine ilişkin aşamalı testlerin analizinde karar ölçütü olarak kullanılmıştır. Chen (2007) tarafından yapılan simülasyon araştırması sonucunda 300'den büyük örneklem söz konusu olduğunda $-0.010 \leq \Delta CFI$ ve $\Delta RMSEA \leq 0.015$ değerleri değişmezlik kararı için kesim noktası olarak önerilmektedir. Bu nedenle bu araştırmada bu değerler ölçme değişmezliğinin sağlanıp sağlanmadığına ilişkin değerlendirmelerde kesim noktası olarak kullanılmıştır.

BULGULAR

SSÖ'nün Dilsel Eşdeğerliğine İlişkin Bulgular

Uyarlama çalışması kapsamında 27 maddelik SSÖ İngilizce ile Türkçe dillerine hakim beş uzman tarafından, birbirinden bağımsız olacak şekilde Türkçeye çevrilmiştir. Çevirisi yapılan formlar birbiri ile karşılaştırılmış, benzer çeviriye sahip olduğu düşünülen maddeler ise görüş birliği ile belirlenmiştir. İngilizceden Türkçeye çevrilen form dört uzman tarafından tekrardan geri çeviri yöntemi kullanılarak İngilizceye çevrilmiştir. Düzenlenen ölçek ön uygulama için 50 kişilik bir öğrenci grubuna uygulanmış, ön uygulamadan alınan geri bildirimler sonucu düzeltmeler yapılarak ölçeğin asıl formu oluşturulmuştur. SSÖ'nün orijinal formu (İngilizce) ile Türkçe formundan elde edilen puanlar arasındaki ilişki Pearson Momentler Çarpım Korelasyonu ile bu puanlar arasında fark olup olmadığı ise bağımlı örneklem t testi ile incelenmiştir.

Tablo 2. SSÖ'nün Dilsel Eşdeğerliğine İlişkin Bulgular

	Fizyolojik kaygı tepkileri	Bilişsel ve davranışsal tepkiler	Algılanan sosyal beklenti ve sosyal kıyas	Toplam
Fizyolojik kaygı tepkileri	0.93	-	-	-
Bilişsel ve davranışsal tepkiler	-	0.87	-	-
Algılanan sosyal beklenti ve sosyal kıyas	-	-	0.96	-
Toplam	-	-	-	0.97

Tablo 2 incelendiğinde ölçeğin orijinal formu ile çeviri formunun fizyolojik kaygı tepkileri, bilişsel ve davranışsal tepkiler ve algılanan sosyal beklenti ve sosyal kıyas boyutlarının korelasyon katsayıları sırasıyla 0.93, 0.87 ve 0.96 olarak bulunmuştur. Tüm ölçekten elde edilen korelasyon katsayısı ise 0.97 olarak bulunmuştur. Korelasyon katsayılarına bakıldığında katsayıların pozitif yönde yüksek olduğu görülmektedir. Ölçeğin orijinal formu ile Türkçe formunun uygulamaları arasında ilişkinin yüksek olması, ölçeğin dilsel eşdeğerliğinin sağlandığının göstergesidir. Bu bulguya ek olarak, orijinal ve alt ölçek formlarından elde edilen uygulamaların ortalamaları arasındaki farkın anlamlılığını test etmek için bağımlı örneklem t testi yapılmıştır. Tablo 3'de elde edilen bulgulara yer verilmiştir.

Tablo 3. SSÖ'nün Orijinal ve Türkçe Formundan Elde Edilen t Testi Sonuçları

Faktör	Formlar	n	sd	\bar{X}	t
1.Boyut	Orijinal	50	49	31.50	1.06
	Türkçe	50	49	32.00	
2.Boyut	Orijinal	50	49	16.58	-1.65
	Türkçe	50	49	16.08	
3.Boyut	Orijinal	50	49	26.92	-1.80
	Türkçe	50	49	26.34	
Toplam	Orijinal	50	49	75.20	-1.33
	Türkçe	50	49	74.42	

$p > 0.05$

Tablo 3'te t değerleri incelendiğinde SSÖ'nün orijinal formu ile çeviri formundan elde edilen puanların farklarının manidar olmadığı görülmektedir. Bu durum SSÖ'nün orijinal dilinden Türkçeye uygun şekilde çevrildiğinin ve dilsel eşdeğerliğinin sağlandığının bir göstergesidir.

Betimsel İstatistiklere ve Ayırt Edicilik Değerlerine İlişkin Bulgular

SSÖ'nün dilsel eşdeğerliği sağlandıktan sonra ölçeğin geçerliğini ve güvenilirliğini sınamak için ölçeğin asıl uygulaması gerçekleştirilmiştir. Ölçek, 1617 kişiden oluşan çalışma grubuna uygulanmıştır. Tablo 4'te ölçek uygulamasından elde edilen betimsel istatistiklere yer verilmiştir. Tablo 4'e göre, SSÖ'nün her bir maddesi için hesaplanan ortalama puan 2.01 ile 3.62 arasında değişmektedir. SSÖ'den elde edilen veriler üzerinde her bir maddeye ve alt boyutlara ilişkin test istatistikleri hesaplandıktan sonra, ölçekteki her bir maddenin ayırt ediciliğini belirlemek için madde toplam korelasyonları hesaplanmıştır. Maddelerin ayırt edicilik değerlerinin 0.28 ile 0.68 arasında değiştiği görülmektedir. Maddelerin genelde yüksek ayırt edicilik düzeylerine sahip olduğu gözlenmiştir.

Tablo 4. SSÖ'nün Madde ve Alt Ölçeklerine İlişkin Hesaplanan Betimsel İstatistikler

	\bar{X}	Mod	Medyan	SS	Min	Max	Madde toplam korelasyonu
M1	3.26	4.00	3.00	1.26	1	5	0.55
M2	2.07	2.00	2.00	1.08	1	5	0.55
M3	2.92	3.00	3.00	1.21	1	5	0.59
M4	3.38	4.00	4.00	1.26	1	5	0.61
M5	3.62	4.00	4.00	1.18	1	5	0.65
M6	3.77	4.00	4.00	1.16	1	5	0.68
M7	3.33	4.00	4.00	1.27	1	5	0.63
M8	3.36	4.00	4.00	1.33	1	5	0.52
M9	3.10	4.00	3.00	1.32	1	5	0.64
M10	2.32	2.00	2.00	1.25	1	5	0.45
M11	2.56	1.00	1.00	0.82	1	5	0.43
M12	2.54	2.00	2.00	1.14	1	5	0.46
M13	2.01	1.00	2.00	1.09	1	5	0.45
M14	2.64	1.00	1.00	0.93	1	5	0.45
M15	2.72	2.00	3.00	1.25	1	5	0.48
M16	2.25	1.00	2.00	1.22	1	5	0.58
M17	2.04	1.00	2.00	1.17	1	5	0.49
M18	2.24	1.00	2.00	1.21	1	5	0.28
M19	2.77	2.00	3.00	1.29	1	5	0.51
M20	2.96	1.00	3.00	1.54	1	5	0.38
M21	3.09	4.00	3.00	1.27	1	5	0.48
M22	2.61	2.00	2.00	1.32	1	5	0.55
M23	2.59	2.00	2.00	1.32	1	5	0.37
M24	2.97	4.00	3.00	1.38	1	5	0.55
M25	2.68	2.00	3.00	1.33	1	5	0.56
M26	3.39	5.00	4.00	1.35	1	5	0.36
M27	3.26	5.00	3.00	1.42	1	5	0.49

Ölçümlerin Güvenirliğine İlişkin Bulgular

27 maddelik SSÖ'den elde edilen ölçümlerin güvenilirliğine ilişkin kanıtlar sağlamak amacıyla Cronbach alfa katsayısı ile bileşik güvenilirlik katsayısı (Raykov, 1997) hesaplanmıştır. Ölçeğin alt boyutlarından elde edilen Cronbach alfa değerlerinin 0.75 ile 0.87 arasında değiştiği gözlenmiştir. Ölçümlerin güvenilirliğine ilişkin bir diğer güvenilirlik ölçüsü olarak bileşik güvenilirlik katsayısı hesaplanmıştır. Bu değerler Tablo 5'te gösterilmektedir.

Tablo 5. Ölçümlerin Farklı Gruplardan Elde Edilen Bileşik Güvenirlik Katsayıları

Gruplar	Fizyolojik kaygı tepkileri	Bilişsel ve davranışsal tepkiler	Algılanan sosyal beklenti ve sosyal kıyas
Kız	0.86	0.76	0.80
Erkek	0.87	0.75	0.77
Fen lisesi	0.91	0.76	0.80
Anadolu lisesi	0.88	0.78	0.80
Mesleki ve teknik anadolu lisesi	0.83	0.75	0.72
Sosyal bilimler lisesi	0.89	0.76	0.84
9. sınıf	0.88	0.78	0.78
10.sınıf	0.86	0.76	0.79
11.sınıf	0.88	0.77	0.79
12.sınıf	0.86	0.75	0.80
Tüm grup	0.88	0.77	0.79

Tablo 5 incelendiğinde hesaplanan bileşik güvenirlik katsayılarının tamamının güvenirlik düzeyinin alt sınırı olan 0.70'in üzerinde değerler aldığı görülmektedir. Bu durum ölçümlerin güvenirliliğinin bir kanıtı olarak gösterilebilir.

Yapı Geçerliğine İlişkin Bulgular

Yapı geçerliğine ilişkin kanıtlar sağlamak için öncelikle ölçeğin faktör yapısına ilişkin tanımlanan ölçme modelinin veri ile uyumu DFA kullanılarak test edilmiştir. Analiz sonuçlarını doğru yorumlamak ve anlamlı çıkarımlar yapabilmek için ilk olarak veri içerisinde, kayıp verilerin olup olmadığı incelenmiştir. Araştırmada, başlangıçta lise öğrencilerinden 1640 adet veri toplanmıştır. Ancak veri incelendiğinde 23 öğrencinin, birden fazla maddeyi yanıtızsız bıraktığı görülmüştür. Bu durumun analiz sonuçlarını etkileyeceği göz önünde tutulmuş, az sayıda katılımcının kayıp değere sahip olduğu düşünülerek araştırmada kayıp veri yöntemlerinden liste bazında silme yöntemi kullanılarak araştırma 1617 öğrenci üzerinden yürütülmüştür.

Kayıp veriler analiz dışında bırakıldıktan sonra araştırmada tek değişkenli uç değerler belirlenmiştir. $N > 100$ gibi geniş örneklemelerde normal dağılım için z puan aralığının ± 4 arasında olması istenmektedir (Çokluk, Şekercioğlu ve Büyüköztürk, 2010). Araştırmada kullanılan veri sayısı 1617 olduğu için örneklem büyüklüğü dikkate alınarak hesaplanan z puanlarının ± 4 puan aralığında olup olmadığı kontrol edilmiştir. Elde edilen z puanlarının -2.28 ile 3.34 arasında değiştiği belirlenmiş ve ± 4 sınırları dışında olan herhangi bir uç değere rastlanmamıştır. Çok değişkenli normallik varsayımı için her gözlenen değişkende tek değişkenli normallik varsayımının gözlenmesi gerekmektedir (Çokluk, Şekercioğlu ve Büyüköztürk, 2010). Bu doğrultuda araştırmada yer alan her bir bağımsız değişkendeki gözlenen değişken için basıklık, çarpıklık değerleri dikkate alınmıştır. Çarpıklık ve basıklık değerlerinin $p < 0.05$ için ± 1.96 'dan büyük; $p < 0.01$ için ± 2.58 'den büyük değerler alması, verideki manidar basıklık ve çarpıklığa işaret etmektedir (Harrington, 2009). Araştırmada yer alan bağımsız değişkenlerdeki gözlenen değişkenlere ilişkin basıklık ve çarpıklık değerleri hesaplanmış ve ± 1.96 değerini aşan herhangi bir değer bulunmadığı için normallik varsayımı karşılanmıştır.

Çoklu bağlantı problemi değişkenler arası korelasyonların yüksek olması durumunda ortaya çıkmaktadır. Bu nedenle değişkenlere ilişkin VIF (Variance Inflation Factor; varyans şişkinlik faktörü) değeri incelenmiş ve 10'dan oldukça küçük olduğu görülmüştür. VIF değerinin 10'a eşit olması ya da daha büyük olması durumunda çoklu bağlantı probleminden bahsedilebilir. Elde edilen VIF sonuçlarına göre araştırmada çoklu bağlantı sorununun veriler için söz konusu olmadığı görülmektedir.

Ölçeğin 27 maddelik Türkçe formunun üç faktörlü yapısına ilişkin tanımlanan ölçme modelinin veriye uyumunu incelemek için yapılan DFA sonucunda elde edilen t değerleri, hata varyansları ve faktör yüklerine ait değerler Tablo 6'da belirtilmiştir. Tablo 6 incelendiğinde, maddelere ait faktör yüklerinin 0.29 ile 0.75 arasında değiştiği görülmektedir. Maddelere ait t değerleri 2.56'yı aştığından

27 madde için elde edilen tüm t değerleri 0.01 düzeyinde manidardır. Buradan hareketle yapılacak analizde, ölçekte bulunan herhangi bir maddenin çıkarılmasına gerek yoktur. Ancak kesin karara varmadan önce hata varyansları da incelenmelidir. Gözlenen değişkenlere ilişkin hata varyansları incelendiğinde 0.90 üzerinde sadece M18'in (0.92) yer aldığı görülmektedir. Hata varyansı yüksek olan bir madde düşük faktör yükü vermektedir. M18'in faktör yükünün 0.29 olduğu görülmektedir. Diğer maddeler ile kıyaslandığında M18 en düşük faktör yüküne sahip maddedir. M18'in hata varyansı yüksek olmasına rağmen söz konusu maddeye ilişkin manidar t değerleri elde edildiği için ve faktör yük değerinin 0.25 üzerinde olmasından dolayı maddenin model içinde yer almasına ilişkin karar verilebilmektedir (Çokluk, Şekercioğlu ve Büyüköztürk, 2010; Önen, 2009). Bu nedenle M18 ölçekten çıkarılmamıştır. Model uyum indeksleri incelendiğinde elde edilen uyum iyiliği indekslerinin kabul edilebilir aralıklarda yer aldığı söylenebilir (RMSEA <0.08, CFI>0.90, NFI >0.90, IFI >0.90 ve NNFI >0.90). Modeldeki gözlenen değişkenlerin (maddelerin) ilgili yapının iyi birer temsilcisi olduğu ve veriye uyum sergilemesi, ölçekten elde edilen ölçümlerin yapı geçerliğine ilişkin kanıtlar sağlamaktadır.

Tablo 6. Sınav Stresi Ölçme Modeline İlişkin Tüm Veriden Elde Edilen t Değerleri, Hata Varyansları ve Faktör Yükleri

Maddeler	t	Hata varyansları	Standartlaştırılmış beta katsayıları
M1		0.63	0.61
M2	22.87	0.63	0.61
M3	22.90	0.58	0.65
M4	24.33	0.54	0.68
M5	25.30	0.48	0.72
M6	26.49	0.43	0.75
M7	24.15	0.52	0.69
M8	19.38	0.68	0.56
M9	25.03	0.52	0.69
M10	17.45	0.75	0.50
M11		0.72	0.53
M12	14.96	0.71	0.54
M13	14.06	0.74	0.51
M14	15.44	0.73	0.52
M15	14.63	0.69	0.56
M16	16.98	0.51	0.70
M17	17.07	0.60	0.64
M18	8.89	0.92	0.29
M19		0.70	0.55
M20	14.49	0.83	0.41
M21	16.02	0.71	0.54
M22	17.73	0.62	0.61
M23	13.73	0.81	0.44
M24	20.75	0.62	0.62
M25	18.86	0.54	0.68
M26	13.69	0.79	0.46
M27	17.38	0.64	0.60

$\chi^2=2808.25$, $sd=321$, $\chi^2/sd=8.74$, RMSEA=0.068 (0.066-0.071), CFI=0.94, NFI=0.93, NNFI=0.93, IFI=0.94

Modelin tüm grup ile veri uyumu incelendikten sonra, ayrıca modelin her bir grupta ayrı ayrı veri ile uyumunun incelenmesi gerekmektedir. Araştırmanın amacı doğrultusunda cinsiyete, okul türüne ve sınıf düzeyine göre model veri uyumuna ilişkin değerler Tablo 7'de verilmiştir.

Tablo 7. Sınav Stresi Ölçme Modelinin Cinsiyet, Okul Türü ve Sınıf Düzeyine İlişkin Veri Uyum Değerleri

	χ^2	s.d.	$\chi^2/s.d.$	RMSEA (CI)	CFI	NFI	IFI	NNFI
Kız	1682.78	321	5.24	0.068 (0.065-0.071)	0.93	0.92	0.93	0.93
Erkek	1471.56	321	4.58	0.069 (0.065-0.072)	0.93	0.91	0.93	0.92
Fen	956.70	321	2.98	0.068(0.063-0.074)	0.95	0.93	0.95	0.95
Anadolu	113.60	321	3.53	0.069 (0.065-0.074)	0.94	0.92	0.94	0.93
Meslek	1044.24	321	3.25	0.064 (0.059-0.068)	0.92	0.88	0.92	0.91
Sosyal	772.09	321	2.40	0.068 (0.061-0.075)	0.95	0.92	0.95	0.95
9.sınıf	1157.15	321	3.60	0.072 (0.067-0.077)	0.94	0.91	0.94	0.93
10.sınıf	1041.93	321	3.24	0.064 (0.059-0.069)	0.94	0.91	0.94	0.93
11.sınıf	973.37	321	3.03	0.070 (0.065-0.075)	0.94	0.91	0.94	0.93
12.sınıf	789.82	321	2.46	0.065 (0.058-0.071)	0.94	0.90	0.94	0.94

RMSEA güven aralığı parantez içerisinde verilmiştir.

Tablo 7 incelendiğinde, sınav stresi ölçme modelinin cinsiyet, okul türü ve sınıf düzeyi değişkenlerine ilişkin uyum indekslerinin, model veri uyumunu değerlendirmede kullanılan ölçütlerin, mesleki ve teknik anadolu lisesindeki NFI değeri hariç, kabul edilebilir aralıklarda olduğu görülmektedir (RMSEA<0.08, CFI>0.90, NFI>0.90, IFI >0.90, NNFI >0.90). Modelde kızlar ve erkekler; fen lisesi, anadolu lisesi, mesleki ve teknik anadolu lisesi ve sosyal bilimler lisesi; 9., 10., 11. ve 12. sınıflar için hesaplanan uyum indekslerinde ilgili gruplar birbiri ile karşılaştırıldığında model uyum indekslerinin benzer olduğu görülmektedir. Tüm gruplardan ayrı ayrı elde edilen model veri uyumu indeksleri sonuçlarına bakılarak cinsiyet, okul türü ve sınıf düzeyi gruplarına göre üç boyutlu bir yapı olarak sınav stresi ölçme modeli, 27 madde ve üç boyuttan oluşan bir model olarak doğrulanmıştır. Bu bulgular kızlar ve erkekler; fen lisesi, anadolu lisesi, mesleki ve teknik anadolu lisesi ve sosyal bilimler lisesi; 9., 10., 11. ve 12. sınıf öğrenci grupları için yapı geçerliğine ilişkin kanıtlar sağlamaktadır.

Ölçme Değişmezliğine İlişkin Bulgular

Modelin tüm alt gruplarda veri ile iyi uyum sergilemesinin ardından farklı gruplarda ölçme değişmezliği çalışmalarına geçilmiştir. Alt gruplara göre ölçme değişmezliği bulguları Tablo 8’de gösterilmektedir.

Tablo 8. Alt Gruplardan Elde Edilen Ölçme Değişmezliği Bulguları

	χ^2	s.d.	RMSEA (CI)	Δ RMSEA	CFI	Δ CFI
Cinsiyet						
Yapısal	3073.5	642	0.075 (0.073-0.078)	-	0.93	-
Metrik	3126.4	666	0.074 (0.072-0.077)	-0.001	0.93	0
Ölçek	3144.1	672	0.074 (0.072-0.077)	0	0.93	0
Katı	3401.4	699	0.077 (0.074-0.079)	0.003	0.92	-0.01
Okul türü						
Yapısal	3852.3	1284	0.076 (0.073-0.078)	-	0.93	-
Metrik	3959.9	1356	0.074 (0.072-0.077)	-0.002	0.93	0
Ölçek	4012.5	1374	0.074 (0.072-0.077)	0	0.93	0
Katı	4319.8	1455	0.075 (0.073-0.078)	0.001	0.93	0
Sınıf düzeyi						
Yapısal	3856.5	1284	0.076 (0.074-0.079)	-	0.93	-
Metrik	4168.2	1356	0.072 (0.070-0.074)	-0.004	0.93	0
Ölçek	3998.0	1374	0.075 (0.072-0.077)	0.003	0.93	0
Katı	4319.8	1455	0.074 (0.072-0.077)	-0.001	0.93	0

RMSEA güven aralığı parantez içerisinde verilmiştir.

Tablo 8 incelendiğinde yapısal değişmezlik aşaması için, model uyumunun değerlendirilmesinde kullanılan uyum ölçütlerinin tüm gruplar için kabul edilebilir sınırlar içerisinde yer aldığı söylenebilir (RMSEA<0.08, CFI>0.90, NFI>0.90, NNFI>0.90, IFI>0.90). Yapısal değişmezlikte modele ilişkin faktör yükleri, faktörler arası korelasyon ve hata varyansları parametreleri alt gruplarda serbest bırakıldığı için, sınav stresi ölçme modelinin yapısının cinsiyet, okul türü ve sınıf düzeyi alt gruplarında aynı olduğu söylenebilir ve söz konusu alt gruplarda yer alan öğrencilerin, ölçek maddelerine cevap verirken kullanmış oldukları kavramsal bakış açısı aynıdır. Yapısal değişmezlik aşaması sağlandığı için bir sonraki metrik değişmezlik aşamasına geçilmiştir.

Metrik değişmezlik aşamasında, alt gruplarda faktör yüklerinin aynı olma sınırlaması getirilmiştir. Elde edilen uyum ölçütleri incelenmiş ve modelin veri ile iyi uyum sergilediği sonucuna ulaşılmıştır. Metrik değişmezliği test etmek için yapısal değişmezlik ve metrik değişmezlik aşamalarında elde edilen CFI ve RMSEA değerleri arasındaki fark incelenmiş ve metrik değişmezlik için ΔCFI ve $\Delta RMSEA$ 'nın kabul edilebilir sınırlar içerisinde yer aldığı görülmüştür ($\Delta CFI \leq 0.01$; $\Delta RMSEA \leq 0.015$). Bu bulgu modele alınan değişkenlerin faktör yüklerinin cinsiyet, okul türü ve sınıf düzeyi alt gruplarında değişmediğine işaret etmektedir. Metrik değişmezlik aşaması sağlandığı için bir sonraki ölçek değişmezliği aşamasına geçilmiştir.

Tablo 8'e göre, ölçek değişmezliği aşamasında, uyum indekslerinin kabul edilebilir sınırlar içerisinde yer aldığı söylenebilir. Ölçek değişmezliğini test etmek için yapı değişmezliğinden elde edilen CFI ve RMSEA değerleri ile ölçek değişmezliğinden elde edilen CFI ve RMSEA değerleri arasındaki farklar incelenmiştir. Bulgular incelendiğinde sınav stresi ölçme modelinin ölçek değişmezliğini sağladığı görülmüştür ($\Delta CFI \leq 0.01$; $\Delta RMSEA \leq 0.015$). Maddeler için oluşturulan regresyon denklemlerindeki sabitlerin cinsiyet, okul türü ve sınıf düzeyi alt gruplarında değişmez olduğu doğrulanmıştır. Bu bulgudan hareketle maddeler bazında herhangi bir yanlılığın bulunmadığı ve gözlenen değişkenlerdeki ortalama farklılıkların gizil/örtük yapılarıdaki farklılıktan kaynaklandığı sonucuna ulaşılabılır. Ölçek değişmezliği aşaması sağlandıktan sonra bir sonraki katı değişmezlik aşamasına geçilmiştir.

Tablo 8'e göre, katı değişmezlik için uyum indekslerinin kabul edilebilir sınırlar içerisinde yer aldığı söylenebilir. Yapı değişmezliği ve katı değişmezlik aşamalarından elde edilen CFI ve RMSEA değeri arasındaki fark incelendiğinde cinsiyet, okul türü ve sınıf düzeyi alt gruplarında sınav stresi ölçme modelinin katı değişmezlik aşamasını sağladığı görülmüştür ($\Delta CFI \leq 0.01$; $\Delta RMSEA \leq 0.015$).

Sınav stresine ilişkin oluşturulan ölçme modeli, tüm değişmezlik aşamalarını sağlamıştır. Bu durum ölçme modelinin cinsiyet, okul türü ve sınıf düzeyi alt gruplarında ölçme değişmezliğinin sağlandığına işaret etmektedir.

SONUÇ

Bu araştırmanın amacı doğrultusunda, ilk olarak ölçeğin İngilizceden Türkçeye çevirisinin uygun olup olmadığını belirlemek amacıyla dilsel eşdeğerlik çalışması yapılmıştır. Ölçeğin İngilizce formundan elde edilen puanlar ile Türkçe formundan elde edilen puanlar karşılaştırılmıştır. Elde edilen bulgular çerçevesinde SSÖ'nün İngilizce formunun Türkçeye uygun bir biçimde çevrildiği sonucuna ulaşılmıştır.

27 maddelik SSÖ'den ölçümlerin güvenilirliğine ilişkin kanıtlar sağlamak amacıyla Cronbach alfa katsayısı ile bileşik güvenilirlik katsayısı hesaplanmıştır. Ölçeğin alt boyutlarından elde edilen ve tüm ölçekten elde edilen Cronbach alpha ve birleşik güvenilirlik katsayıları incelendiğinde katsayı değerlerinin yüksek olduğu görülmüştür. Elde edilen güvenilirlik katsayıları, ölçeğin orijinal formundan elde edilen katsayılar ile yakın değerlere sahiptir. Elde edilen katsayıların 0.70'den düşük değerler almadığı görülmektedir. Psikolojik bir test için hesaplanan güvenilirlik katsayısının 0.70 ve üstü olması durumunda elde edilen güvenilirlik düzeyi yeterli görülmektedir (Büyüköztürk, 2003). Bu bulgulardan hareketle ölçeğin güvenilir bir araç olduğu söylenebilir.

27 maddelik SSÖ'nün geçerliğine ilişkin kanıtları sağlamak için tüm gruplardan elde edilen veri üzerinde DFA yapılmıştır. Oluşturulan sınav stresi ölçme modelinin tüm grup ile veri uyumu incelendikten sonra ayrıca her bir grupta ayrı ayrı veri ile uyumu incelenmiştir. Tüm gruplardan ayrı ayrı model veri uyumu indeksleri sonuçlarına bakılarak cinsiyet, okul türü ve sınıf düzeyi gruplarına göre ölçeğin sınav stresini üç boyutlu bir yapı olarak ölçebildiğine işaret eden yapı eşitliği sağlanmıştır ve sınav stresi ölçme modeli, 27 madde ve üç boyuttan oluşan bir ölçme modeli olarak doğrulanmıştır. Aynı zamanda kızlar ve erkekler; fen lisesi, anadolu lisesi, mesleki teknik ve anadolu lisesi ve sosyal bilimler lisesi; 9., 10., 11. ve 12. sınıf öğrenci grupları için yapı geçerliğine ilişkin ayrı ayrı kanıtlar sağlanmıştır.

Model veri uyumu sağlandıktan sonra cinsiyet, okul türü ve sınıf düzeyine dayalı olarak yapılan ölçme değişmezliği sonuçları ΔCFI ve $\Delta RMSEA$ değerleri incelenerek belirlenmiştir ($\Delta CFI < 0.01$, $\Delta RMSEA \leq 0.015$). Sınav stresi ölçme modelinin cinsiyet, okul türü ve sınıf düzeyine göre alt gruplarda yapısal, metrik, ölçek ve katı değişmezlik aşamalarının tümünü sağlayarak tam değişmezlik koşulunu yerine getirdiği sonucuna ulaşılmıştır.

Genel olarak bulgular değerlendirildiğinde bu araştırmadan çıkarılacak sonuçlar şöyle sıralanabilir: SSÖ Türkçe formu kullanılmaya uygun geçerli ve güvenilir bir ölçme aracıdır. Ölçeğin orijinal ve Türkçe formundan elde edilen sonuçlar benzerdir. Ayrıca yapılan ölçme değişmezliği analizleri sonucunda ölçeğin cinsiyet, okul türü ve sınıf düzeyi değişkenlerine dayalı olarak karşılaştırılabilir olduğu ifade edilebilir.

Her araştırmada olduğu gibi bu araştırmanın da bazı sınırlıkları bulunmaktadır. Bu sınırlıklardan yola çıkılarak gelecekte yapılacak olan araştırmalara bazı önerilerde bulunulabilir. Bu çalışma Balıkesir ilinde yer alan ortaöğretim kurumlarının oluşturduğu bir çalışma grubu üzerinden yürütülmüştür. Sonuçların genellenebilirliğini arttırmak amacıyla araştırmalar, farklı okul türleri üzerinde yürütülebilir. Araştırma, çalışma grubu üzerinden gerçekleştirilmiştir. Farklı örneklem büyüklüklerinde, örneklem büyüklüğünün ölçme modeline etkisini test etmek amacıyla araştırmalar yürütülebilir. Araştırmada SSÖ kullanılarak cinsiyet, okul türü ve sınıf düzeyi değişkenleri ile farklı gruplarda ölçme değişmezliği test edilmiştir. Cinsiyet, okul türü ve sınıf düzeyi değişkenleri dışında farklı değişkenler kullanılarak ölçme değişmezliğine ilişkin araştırmalar yapılabilir.

KAYNAKÇA

- Altuntaş, E. (2003). *Stres yönetimi*. İstanbul: Alfa Yayınları.
- Arslan, S. (2016). *Üniversitelere hazırlanan öğrencilerde stres düzeylerinin duyguları yönetme becerisine etkisi*. Yayınlanmamış Yüksek Lisans Tezi, Nişantaşı Üniversitesi, Sosyal Bilimler Enstitüsü, İstanbul.
- Atkinson, R., Atkinson, R., Smith, E., & Bem, D. (1999) *Psikolojiye giriş*. (Çev. Y. Alogan). Ankara: Arkadaş Yayınevi.
- Baltaş, A., & Baltas, Z. (1999). *Stres ve başa çıkma yolları*. İstanbul: Remzi Kitabevi.
- Bollen, K. A. (1989). *Structural equations with latent variables*. New York: Wiley Interscience.
- Braham, B. (1998). *Stres yönetimi* (Çev. V. Diker). İstanbul: Hayat Yayıncılık.
- Büyüköztürk, Ş. (2003). *Sosyal bilimler için veri analizi el kitabı*. Ankara: Pegem.
- Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling*, 14(3), 464–504.
- Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling*, 9 (2), 233-255.
- Costarelli, V., & Patsai, A. (2012). Academic examination stress increases disordered eating symptomatology in female university students. *Eat Weight Disord*, 17 (3), 164-169.
- Crocker, L., & Algina, J. (1986). *Introduction to classical and modern test theory*. Philadelphia: Harcourt Brace Jovanovich College Publishers.
- Çapulcuoğlu, U., & Gündüz, B. (2012). Öğrenci tükenmişliğini yordamada stresle başa çıkma, sınav kaygısı, akademik yetkinlik ve anne-baba tutumları. *Eğitim Bilimleri Araştırmaları Dergisi Uluslararası E-Dergi*, 3 (1), 201–218.
- Çokluk, Ö., Şekercioğlu, G. & Büyüköztürk, Ş. (2010). *Sosyal bilimler için çok değişkenli istatistik (SPSS ve LISREL uygulamaları)*. Ankara: Pegem Akademi.

- Eryılmaz, A. (2009). Ergenlik döneminde stres ve başa çıkma. *Yüzüncü Yıl Üniversitesi Eğitim Fakültesi Dergisi*, 6 (2), 20-37.
- Hançerlioğlu, O. (1988). *Ruhbilim sözlüğü*. İstanbul: Remzi Kitabevi.
- Harrington, D. (2009). *Confirmatory factor analysis*. New York: Oxford University Press, Inc.
- Hirschfeld, G. & von Brachel, R. (2014). Multiple-Group confirmatory factor analysis in R –A tutorial in measurement invariance with continuous and ordinal indicators. *Practical Assessment, Research & Evaluation*, 19(7), 1-11.
- Hu, L., & Bentler, P. (1999). Cut off criteria for fit indices in covariance structure analysis: conventional criteria versus new alternatives. *Structural Equation Modeling*, 6, 1-55.
- Jöreskog, K. G. & Sörbom, D. (1993). *Lisrel 8: Structural equation modeling with the simplis command language*. Lincolnwood: Scientific Software International, Inc.
- Köknel, Ö. (1998). *Zorlanan insan (Kaygı çağında stres)*. İstanbul: Altın.
- Milfont, T. L. & Fischer, R. (2010). Testing measurement invariance across groups: Applications in cross-cultural research. *International Journal of Psychological Research*, 3(1), 111-121.
- Motavalli, N. (1997). Çocukluk çağında görülen “Travma sonrası stres bozukluğunun” klinik özellikleri ve seyri. *Yeni Sempozyum*, 35, 92-95.
- Önen, E. (2009). *Ölçme değişmezliğinin yapısal eşitlik modellemesi teknikleri ile incelenmesi*. Yayınlanmamış Doktora Tezi, Ankara Üniversitesi, Eğitim Bilimleri Enstitüsü, Ankara.
- Pazarlı, S. (2009). *Öğrenme stilleri ile sınav kaygısı arasındaki ilişki (İstanbul ili örneği)*. Yayınlanmamış Yüksek Lisans Tezi, Yeditepe Üniversitesi, Eğitim Bilimleri Enstitüsü, İstanbul.
- Raykov, T. (1997). Estimation of composite reliability for congeneric measures. *Applied Psychological Measurement*, 21(2), 173-184.
- Rowshan, A. (2000). *Stres yönetimi*. (Çev. Ş. Cüceloğlu). İstanbul: Sistem Yayıncılık.
- Sabuncuoğlu, Z., & Tüz, M. (1998). *Örgütsel psikoloji*. Bursa: Alfa.
- Sarason I. G., & Sarason B. R. (1990). Test anxiety. In Leitenberg H. (Ed.), *Handbook of social and evaluation anxiety* (pp. 475–496). New York, NY: Plenum Press.
- Selye, H. (1973). The evolution of the stress concept. *American Scientist*, 61(6), 692-699.
- Selye, H. (1986). History and present status of the stress concept. Goldberger, L. S. & Breznitz, S. (Ed.), *Handbook of stress*. New York: The Free Press.
- Somer, O., Korkmaz, M., Dural, S., & Can, S. (2009). Ölçme eşdeğerliğinin yapısal eşitlik modellemesi ve madde cevap kuramı kapsamında incelenmesi. *Türk Psikoloji Dergisi*, 24(64), 61-75.
- Steenkamp, E. M., & Baumgartner, H. (1998). Assessing measurement invariance in cross-national consumer research. *Journal of Consumer Research*, 25(1), 78-90.
- Steptoe, A., Wardle, J., Pollard, T. M., Canaan, L., & Davies, G. J. (1996). Stress, social support and health-related behavior: a study of smoking, alcohol consumption and physical exercise. *Journal of Psychosomatic Research*, 41(2), 171–180.
- Sung, T. S., & Chao, T. Y. (2015). Construction of the examination stress scale for adolescent students. *Measurement and Evaluation in Counseling and Development*, 48(1), 44-58.
- Ulusoy, S., Yavuz, K. F., Esen, F. B., Umut, G. & Karatepe, H. T. (2016). Sınav kaygısına yönelik bilişsel grup terapisi. *Journal of Cognitive-Behavioral Psychotherapy and Research*, 5(1), 28-37.
- Vandenberg, R. J., & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: suggestions, practices and recommendations for organizational research. *Organizational Research Methods*, 3(1), 4-70.
- Wu, D. A., Li, Z., & Zumbo, B. D. (2007). Decoding the meaning of factorial invariances and updating the practice of multi-group confirmatory factor analysis: A demonstration with TIMSS data. *Practical Assessment Research and Evaluation*, 12(3), 1-26.

EXTENDED ABSTRACT

Introduction

Within the scope of this research, validity and reliability studies are carried out by adapting the Examination Stress Scale developed by Sung and Chao (2015) into Turkish and a model for examination stress is tested. In the second phase of the study, the measurement invariance was examined across gender, school type and grade.

Method

Data collection tool, the Examination Stress Scale, is a five-point likert scale consisting of 27 items. Firstly, the original form of the scale was translated into Turkish by five experts. The English and Turkish form of the scale were applied to a group knowing both languages well. Pearson Moment Product Correlation Coefficient was estimated to examine the linguistic equivalence between the scores obtained from the original form and Turkish form.

The study group consists of 1617 students in the 9th, 10th, 11th and 12th grades of science high school, Anatolian high school, technical high school and social sciences high school in Balıkesir in 2016-2017 education period. Test and item analyses were performed on the data. Item –total correlations were calculated for item discrimination and reliability was calculated by the Cronbach alpha internal consistency coefficient for all subscales of the scale separately. In addition, considering the differentiation of factor loadings and error variances, the composite reliability coefficient is calculated because it is considered that the composite reliability coefficient calculated for multidimensional scales has an appropriate reliability value than the Cronbach alpha coefficient. Confirmatory factor analysis, for the validity of the scale, was conducted to test the three-factor structure of the scale. Measurement invariance of the examination stress measurement model was examined according to the gender, school type and grade. Multi-group confirmatory factor analysis technique was used to examine the measurement invariance of the examination stress model across the gender, school type and gender. The measurement invariance tests were interpreted based on ΔCFI and $\Delta RMSEA$ values.

Results

Analyses that were conducted to examine the linguistic equivalence between the scores obtained from the original form of the scale and the scores obtained from the Turkish form showed that correlation coefficients were quite high. Correlation coefficients obtained from subscales were found as .93, .87, .96, respectively. Difference between the scores obtained from Turkish and English form was examined via the paired sample t-test. According to the result of the analysis, significant difference was not found. After the linguistic equivalency was achieved, the examination stress measurement model was tested by confirmatory factor analysis. Findings revealed that the Examination Stress Scale was composed of 27 items and three dimensions confirmed in Turkish sample. After it was determined that all the data fit well with the model, the model data fit was tested separately for each group in the sub-groups. It has been seen that modeling in each of sub-groups is well adapted. These findings were taken as evidence of validity of the scale. After this phase, multi-group confirmatory factor analysis technique was performed to examine the measurement invariance of the examination stress model across gender, school type and grade. Measurement invariance results showed that acceptable measurement invariance (configural, metric, scalar and strict invariance) exists across gender, school type and grade. The results were evaluated generally, it has been suggested that the Examination Stress Scale is a valid and reliable measurement tool and it can be used to measure examination stress at different levels in sub-groups of gender, school type, and grade.