

Makalenin Türü / Article Type : Araştırma Makalesi / Research Article  
Geliş Tarihi / Date Received : 13.09.2018  
Kabul Tarihi / Date Accepted : 22.02.2019  
Yayın Tarihi / Date Published : 11.03.2019



<https://dx.doi.org/10.17240/aibuefd.2019.19.43815-459831>

## BİLİŞSEL OLMAYAN ÖĞRENME DURUM VE SÜREÇLERİNİ ÖLÇME VE DEĞERLENDİRMEDE BOYLAMSAL YAKLAŞIMLAR: DUYGU CETVELİ UYGULAMASI ÖRNEĞİ

Nilüfer KAHRAMAN<sup>1</sup>, Derya AKBAŞ<sup>2</sup>, Esra SÖZER<sup>3</sup>

### ÖZ

Ölçme desenlerinin boylamsal olarak tasarlanmalarının eğitim araştırmalarına getireceği katkılar, öğrenmenin sosyal-duygusal boyutlarının üniversite öğrencileri için çalışıldığı bir alan uygulaması olan Duygu Cetveli Uygulaması ile yapılan üç ayrı araştırmanın ölçme desenleri üzerinden örneklendirilmiştir. İlk çalışmada stres ve gerginlik duyguları arasındaki ilişkiler günlük tekrarlı gözlemler üzerinden karşılaştırılırken, ikinci ve üçüncü çalışmalarda duygu durumları ile uyku kalitesi ve psikolojik dayanıklılık arasındaki ilişkiler haftalık tekrarlı gözlemler kullanılarak modellenmiştir. Ölçme desenlerinin, ölçmeye konu özelliğin süreçsel yapısını daha iyi yansıtmasına izin verebilecek tekrarlı ölçümleri de içerecek şekilde genişletilerek tasarlanmalarının önemi ve bunun, özellikle de, dinamik faktör yapısına sahip özellikleri ölçme ve değerlendirme çalışmalarına getireceği katkılar tartışılmıştır.

**Anahtar Kelimeler:** Boylamsal, bilişsel olmayan, özellik, beceri, ölçme.

## LONGITUDINAL ASSESSMENT DESIGNS FOR NON-COGNITIVE CONSTRUCTS: APPLICATIONS USING TIME-SERIES DATA FROM AN EMOTION RULER FIELD STUDY

### ABSTRACT

This paper proposes a longitudinal assessment approach for the assessment of non-cognitive constructs that are prone to change over time. Focusing on how to evaluate “state” versus “trait” properties of propensities, a comparative narrative is provided attesting to the utility of employing a longitudinal perspective. Illustrations on how to collect, analyze and interpret time-series data are provided using three separate studies, conducted as a part of a larger longitudinal study, referred to as the Emotion Ruler Field Study, that has been conducted to investigate emotional aspects of learning for college students. The first study illustrates how inferences made about stress and tenseness levels of college students may differ when based on self-assessments taken over twelve days versus on those taken only once. The second and the third studies illustrate how inferences made about sleep quality and psychological resilience of college students may evolve when the variances observed over several weeks (within-persons) are taken into account in the analysis of individual differences (between-persons).

**Keywords:** Longitudinal, non-cognitive, trait, state, assessment

<sup>1</sup> Gazi Üniversitesi, Eğitim Fakültesi, kahramannilufer@gmail.com, <https://orcid.org/0000-0003-2523-0155>

<sup>2</sup> Adnan Menderes Üniversitesi, Eğitim Fakültesi, derya.akbas@adu.edu.tr, <https://orcid.org/0000-0001-9852-4782>

<sup>3</sup> Bartın Üniversitesi, Eğitim Fakültesi, esrasözer@gmail.com, <https://orcid.org/0000-0002-4672-5264>

## 1.GİRİŞ

Başarı kavramının "okul ortamında öğrenme" ve "bilişsel" alanla ilgili beceriler ile sınırlanmaması, uyum sağlama, duygularını kontrol etme ve motivasyon gibi "yaşam boyu öğrenme" ve "bilişsel olmayan" (duyuşsal) alanla ilgili becerileri de kapsayacak şekilde genişletilmesi gerektiği, içinde bulunduğumuz çağ için oldukça gecikmiş ve acil bir öngörüdür. Birleşmiş Milletler Sürdürülebilir Kalkınma Gündemi (United Nation's Agenda for Sustainable Development, 2030), örneğin, evrensel barışı sağlamak için önemsenmesi gereken önceliklerin ilk dördünü, yoksulluğun olmaması, açlığın olmaması, sağlık ve iyi-olma (well-being) ve kaliteli eğitim olarak listelemektedir (bkz. <https://sustainabledevelopment.un.org/sdgs>). Kaliteli eğitimin, şiddet, terör ve aşırı yoksulluğun ortadan kaldırılması konularına yardımcı olmadaki potansiyel rolü de düşünüldüğünde, neden birçok gelişmiş ülkenin eğitim programlarını, sosyal-duygusal beceriler olarak da adlandırılan, bilişsel olmayan becerileri de kapsayacak şekilde yeniden yapılandırmaya gittiği açıktır. Öğrencilerin 21. Yüzyılın iş dünyasında başarılı olabilecek rekabet gücüne sahip olmaları, yani motivasyon, girişimcilik, yaratıcılık, üretkenlik, dayanıklılık, yeniliklere ve değişen şartlara adapte olabilme, gibi temelde bilişsel olmayan beceriler ile donanmış olmaları, mevcut öğretim ve değerlendirilme süreçlerinin en önemli hedeflerinden biri olmalıdır (örn., PISA, <http://www.oecd.org/pisa/>; Budria and Carbonell, 2012; Gutman and Schoon, 2013; Jackson, 2012; Kautz ve ark., 2014). Bu, sadece bireylerin değil, tüm ekonomik sisteminin de geleceği için oldukça önemlidir. Ancak, bu konudaki girişimler, bilişsel olmayan becerilerin altında yatan sosyal-duygusal öğrenmelerin ölçülmesinde kullanılan yöntemlerin, böylesi zamanla gelişime/değişime açık olması çok muhtemel özelliklerin doğasını yakalamada ihtiyaç duyulacak şekilde, süreç değil de sonuç odaklı (Bloom, 1969, Scriven, 1967) olmalarından kaynaklanan nedenler ile oldukça sıkıntılı olabilmektedir. Bu durum, bilişsel olmayan özellikler ile ilgili becerilerin, öğretim programı değerlendirme çalışmalarında, genellikle ya çok fazla önemsenmemelerini, ya da önemsenseler bile başarı göstergeleri ile direk olarak ilişkilendirilemedikleri için arka planda kalmalarını beraberinde getirebilmektedir. Önemsendikleri ve bir şekilde de ön plana alındıkları araştırma ve uygulamalarda da (Örn., OECD, <http://www.oecd.org/pisa/publications/>), sağlıklı ve gerçekçi bir şekilde ölçülmelerinde uygulamada karşılaşılan "psikometrik" zorluklardan dolayı, anlamlı bir şekilde yorumlanmalarında ciddi zorluklar yaşanabilmektedir.

Geleneksel ölçme desenlerinde bireyler ile ilgili kararlar genellikle tek veya test-tekrar test gibi birkaç seferlik ölçümler üzerinden alınmaktadır (Widaman, Ferrer & Cogner, 2010). Oysaki bir devamlılık içerisinde değil de, sadece belli zaman noktalarında, yani kesitsel alınan ölçümler ile bireylerin "gelişim" durumlarının sağlıklı bir şekilde izlenmesi mümkün olmayacaktır. Hele de, Eğitim ve Psikoloji alanlarında sıkça rastlandığı gibi, eğer ölçmeye konu özellikler doğaları gereği zaman içinde "değişim" göstermeye meyilli özellikler iseler (Borgonovi & Pál, 2016), elde edilecek ölçümlerin sağlıklı olabilmeleri için boylamsal ölçme desenleri ile alınmaları daha uygun olacaktır. Bilişsel olmayan becerilerin zaman içerisindeki gelişim ve değişimlerinin ve diğer değişkenlerle olan ilişkilerinin daha tutarlı ve anlamlı bir şekilde ölçülebilmeleri için kullanılan ölçme araç ve yöntemlerinin gelişim ve değişimi gözlemede kullanılabilecek şekilde, yani boylamsal olarak tasarlanmaları önemlidir. Örneğin, ister ideal olmayan uyku alışkanlıklarından, isterse psikolojik ya da fizyolojik sıkıntılardan kaynaklansın, uyku ile ilgili problemlerin dikkat toplama becerileri üzerindeki olumsuz etkileri ve uyku problemi yaşayan öğrencilerin okul başarılarının olumsuz etkilendiği bilinmektedir. (Örn., Curcio, Ferrara & Gennaro, 2006). Bu konuda bir okulda iyileştirmeye yönelik çalışmalar yapılacak ise eğer, bu çalışmanın öğrenciler için kullanışlı olabilecek destek mekanizmaları oluşturabilmesi için, boylamsal gözlemler yapması daha ideal olacaktır. Bir başka örnek, bu yüzyılın gerektirdiği önemli yaşam becerilerinden biri olarak görülmesi sebebi ile güncel eğitim araştırmalarında kendine gittikçe daha fazla yer bulan ve zamansal değişimlerin izlenmesinden faydalanacak bir diğer özellik ise psikolojik dayanıklılıktır (Masten, Cutuli, Herbers, Reed, 2009; Werner, 2013; Werner ve Smith, 1992). Oysa ki, bireylerin güçlülere ve stresli durumlara rağmen normların üzerinde davranabilmesine, başa çıkmasına ve eski haline dönmelerine yardımcı olan yetenek ve karakter özelliklerinin bir kombinasyonu (Rutter, 1993) olarak tanımlanan psikolojik dayanıklılık ölççekleri, uygulamalarda genellikle tek sefer veya ön-test/son test ölçümleri üzerine kuruludur (Collie ve ark., 2017; Bonanno, Galea, Bucciarelli ve Vlahov, 2007). Özetle, zaman faktörünün etkilerinin izlenebileceği boylamsal uygulamalar hem birey düzeyinde hem de grup düzeyinde dinamik analizlere olanak vermeler ile kesitsel uygulamalara göre oldukça daha kullanışlıdır (Conger ve Conger, 2002; Werner, 2013; Crane ve ark., 2012; Norris, Tracy, Galea, 2009). Duruma göre, yani bireyin yaşadığı bir süreç içerisinde nerede bulunduğu göre farklılık gösterebilecek özelliklerin tek sefer değil tekrarlı biçimde ölçülmeleri gerektiği gittikçe daha fazla kabul gören bir kanıdır (Funder, 2001).

### 1.1. Araştırmanın Amacı

Öğrencilerin kendi ve diğerlerinin duygularını anlama ve yönetme becerileri ile ilgili öğrenmelerinin bir bütünü olarak tanımlanan sosyal-duygusal öğrenmeleri (Elias, 2006) ve bu öğrenmeleri ile ilgili (bilişsel ve bilişsel-olmayan) yeterliliklerinin (Örn. WEF, 2016), eğitim programlarında ve bu programları yönlendirme amacı ile yapılan eğitim araştırmalarında daha çok yer almalarını engelleyen psikometrik zorlukları tanımlama ve aşmakta kullanılabilecek, süreç izlemeye yönelik boylamsal bir ölçme yaklaşımı önerilmiştir. Önerilen boylamsal ölçme

yaklaşımının, kullanılan yaklaşımının boylamsal olup olmamasına göre ölçümler üzerinden varılabilecek sonuçların nasıl farklılaşabileceğini, ölçümlerin tekrarlı gözlemler ile ve önceden belirlenmiş bir süreç içerisinde, yani boylamsal olarak alınmalarının, yapılabilecek çıkarımları nasıl zenginleştirebileceği üç ayrı uygulamada toplanan veriler üzerinden örneklendirilmiştir. Uygulamalar, geleceğin öğretmen adayları olan eğitim fakültesi öğrencilerinin sosyal-duygusal beceri profillerinin çalışıldığı daha büyük ölçekli boylamsal bir çalışma olan Duygu Cetveli Uygulamasının birer parçası olarak 2016-2017 yılları arasında farklı akademik dönemlerde yapılmış çalışmalardır.

Aşağıdaki ilk olarak, üniversite öğrencilerinin "stres" ve "gerginlik" duygu durumlarının zaman içerisindeki değişimlerinin günlük olarak izlendiği bir çalışmada toplanan veriler üzerinden boylamsal ölçme modellerinin Eğitim ve Psikolojide Ölçme ve Değerlendirme alanına getirebileceği potansiyel katkı örneklendirilerek tartışılmıştır. Takip eden bölümde, birincisi duygu durumları ve uyku kalitesi, ikinci ise duygu durumları ve psikolojik dayanıklılık arasındaki ilişkileri inceleyen iki ayrı boylamsal çalışmada elde edilen bazı özet bulgular paylaşılmıştır. Kullanılan ölçme deseninin boylamsal olmasının güvenilirlik ve geçerlilik kanıtları bağlamında elde edilen çıkarımlara katabileceği potansiyel zenginliği somutlaştıran bu iki uygulama örneğini takip eden bölümde, boylamsal ölçme desenlerinin daha yaygın olarak kullanılmalarının ölçme uygulamalarına getireceği katkılar tartışılmıştır. Son bölümde, ölçme ve değerlendirme alanında yapılacak yeni AR-GE çalışmalarının, tekrarlı gözlemleri içeren ve süreçsel güvenilirlik ve geçerlilik kanıtları ile desteklenmiş yeni ölçme desenlerinin geliştirilmesine odaklanması gerektiğine ve bunun Eğitim Bilimleri alanına getireceği katkıya vurgu yapılmıştır.

## 1.2. Araştırmanın Önemi

Daha çok sosyal-duygusal öğrenmeler ile ilgili olduğu düşünülen ve alanyazında daha çok bilişsel olmayan beceriler kategorisi altında anılan becerilerin, bireylerin öğrenme yaşantılarını anlamlandırmasında ve hayata geçirmesinde önemli rolleri oldukları düşünülmektedir. Ancak öngörülen bu rolün düzeyi, süreçsel işleyişi ve diğer bilişsel ve duyuşsal becerileri ile olan ilişkileri hakkındaki önermeler uygulamada henüz çok da tutarlı bir şekilde tespit ve teyit edilememektedirler (Örn., Crossley, 2016; Muthén & Curran, 1997; Sayer & Cumsille, 2001). Bu durumun en önemli sebeplerinden biri de mevcut ölçme araç ve yöntemlerinin bilişsel becerilere göre, yapıları çok daha karmaşık ve değişken olabilen bilişsel-olmayan becerileri değerlendirmede yetersiz kalmalarıdır. Yapılan ölçümler ile ilgili güvenilirlik ve geçerlilik problemlerinin aşılması adına, mevcut "doğru" cevap temelli bilişsel alan ölçme yöntemlerinin, doğru yerine "yeğlenen" cevabın temele alındığı bilişsel olmayan alan ölçme yöntemlerine adapte edilmeleri konularında yıllardır birçok çalışma yapılmış olmasına rağmen, alanyazında gelinen nokta henüz istenenden oldukça uzaktır. Bu makaleye konu olan boylamsal bir ölçme yaklaşımı geliştirme çabası bu gereksinimden doğmuştur.

Önerilen ve örneklendirilen boylamsal ölçme yaklaşımı, önceden ve kuramsal bağlamda tanımlanmış bir süreç boyunca, birbirini eşit aralıklar ile takip eden zaman noktalarında, aynı bireylerden tekrarlı gözlemler yoluyla veri toplamaya odaklanan bir ölçme yaklaşımıdır. Bu yaklaşım, en temelde, tek sefer veya ön-test/son-test gibi birkaç sefer uygulaması olan geleneksel (kesitsel) ölçme uygulamalarına bir alternatif olarak düşünülebilir (Salkind, 2010). Boylamsal ölçme yaklaşımının, bireylerin ölçmeye konu özellikler bakımından zaman içinde değişimlerinin/gelişimlerinin ve var ise, yapılan değişiklik veya müdahalelerin sonrası gözlenen etki büyüklüklerinin ve zaman içerisindeki devamlılıklarının modellenmesine imkân vermesi bakımından, kesitsel ölçümleri içeren klasik yaklaşıma göre, pek çok avantajının olması beklenmektedir (Avey, Luthans ve Mhatre, 2008).

## 2. YÖNTEM

### 2.1. Araştırmanın Modeli

Duygu Cetveli Alan Uygulaması, ölçme uygulamalarında kesitsel olarak yapılan ölçümler kullanılarak keşfedilmeye çalışılan bireyler-arası farklılıkların, boylamsal araştırma desenleri ile desteklenerek bireyler-içi farklılıkları da hesaba katabilecek şekilde genişletilebildiği ölçme modellerinin uygulamalı olarak çalışılması için tasarlanmış çevrimiçi bir araştırma-geliştirme platformudur. Şu an, pilot uygulama aşamasında olan bu alan uygulamasında, görsel ve işitsel uyaranlarla zenginleştirilmiş yenilikçi madde türleri ve bir dizi boylamsal ölçme deseni tasarısı çalışılmaktadır.

Duygu Cetveli Alan Uygulamasında kullanılan psikometrik yaklaşım boylamsal ve bütüncül bir yaklaşımdır ve bu bağlamda birbirini tamamlayan iki bağdaşık önerme üzerine kurulmuştur. İlk önerme, motivasyon, adaptasyon gibi karmaşık ve yapıları eğilimlerin/becerilerin değişim/gelişim eğrilerinin izlenmesinin, ancak boylamsal bir yaklaşım ile mümkün olacaktır. Bu bağlamda, değişime açık yapıları becerilerin ölçülmesi ve değerlendirilmesinde boylamsal bir yaklaşımın kullanılması, yapılacak çıkarımları daha sağlıklı ve gerçekçi kılacaktır. İkinci önerme, çalışmanın üzerine inşa edildiği, değişime açık yani dinamik faktör yapıları eğilimlerin/becerilerin doğaları

gereği, zaman içerisinde açığa çıkabilecek bazı alt-boyut veya alt-grup farklılıklarına haiz olabileceği önermesidir. Bu bağlamda, becerileri oluşturduğu düşünülen farklı bilişsel ve bilişsel-olmayan özelliklerin bir arada ve bir süreç olarak gözlenmeleri ile, alan yazındaki kuramsal gelişmelerin ortaya koyduğu karmaşık yapıların test edilmesi, alt-boyut ve alt-grup karşılaştırmalarının yapılması mümkün olabilecektir. Böylece, ölçmede genellikle asıl konu olan bireyler-arası farklılıklar daha sağlıklı bir biçimde değerlendirilebilecektir. İdealde arzulanan, bireyler-arası farklılıklar çalışılırken, zaman içerisinde bireyde meydana gelebilecek değişimlerin, yani bireyler-içi farklılıkların da hesaba katılabileceği ölçme modelleridir. Nihai amaç, sonuç yerine süreç odaklı, tek veya birkaç zaman noktası yerine zaman serisi veri toplamaya izin verebilecek şekilde tasarlanabilecek olan yeni ölçme modellerinin geliştirilmesi ve yaygınlaşmasına katkıda bulunmaktır.

## 2.2. Veri Toplama Aşaması

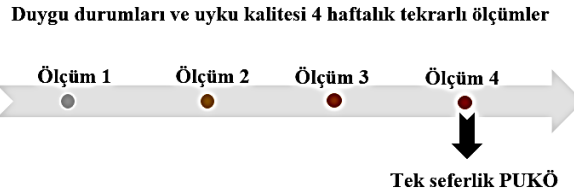
Bu çalışmada kullanılan örneklerin veri setleri, çağın getirdiği teknolojik imkânlardan maksimum düzeyde yararlanmak adına, çevrimiçi bir uygulama olarak tasarlanmış olan boylamsal bir alan uygulamasından gelmektedir. Bu alan uygulaması ile en temelde amaçlanan, boylamsal ölçme modeli geliştirme çalışmalarına uygulama ortamı yaratmak ve uygulama sonrası yapılan analizler ile güvenilirlik ve geçerlilik delilleri ile desteklenen modelleri alana kazandırmaktır.

### 2.2.1. Örnek çalışma 1 için ölçme deseni, örneklem, veri toplama araçları ve verilerin analizi

2016 yılında yapılan bu çalışmada gönüllü katılımcılarının oluşturduğu çalışma grubunda yaklaşık 64 Eğitim Fakültesi öğrencisi yer almaktadır. Çalışma 1'in ölçme deseninde, katılımcılardan 13 gün boyunca her akşam çevrimiçi bir ölçeği doldurmaları istenmiştir. 13 gün boyunca, öğrencilerin her gün için işaretledikleri stres ve gerginlik düzeyleri (1: hiç ya da çok az - 5: Çok fazla) görselleştirilerek, günlük değişimlerin bireyler-arası karşılaştırmalar için ne anlama gelebileceği örneklendirilmiştir.

### 2.2.2. Örnek çalışma 2 için ölçme deseni, örneklem, veri toplama araçları ve verilerin analizi

Örnek çalışma 2'nin ölçme deseni ve uygulaması 4 haftalık bir zaman dilimi için ve haftalık olarak planlanmıştır. 69 gönüllü üniversite öğrencisinden oluşan çalışma grubuna iki ayrı çevrimiçi ölçek uygulanmıştır. Bu ölçeklerden ilki sağlık kuruluşlarında yaygın kullanımı olan ve cevaplayıcıların bir önceki ayı düşünerek uyku kaliteleri ile ilgili soruları cevaplamalarını gerektiren Pittsburgh Uyku Kalitesi Ölçeğidir (PUKÖ, Buysse vd., 1988). İkinci ölçek ise, bu "aylık ve tek seferlik" olarak düşünülebilecek olan Pittsburgh Uyku Kalitesi Ölçeğine bir alternatif olarak ve bu çalışma kapsamında tasarlanmış olan, "4 haftalık-tekrarlı uygulamalı Uyku Cetveli" Ölçeğidir. Çalışma grubu 4 hafta boyunca her hafta sonu Uyku Cetveli Ölçeğini cevaplamışken, Pittsburgh Uyku Kalitesi Ölçeğini sadece bir kez ve 4. haftada cevaplamıştır. Çalışmanın ölçme deseni Şekil 1' de gösterilmiştir. Bireylerin, bu 4 hafta boyunca "hiç" için 1'den "çok" için 5'e kadar rakamlar vererek işaretledikleri yorgunluk ve stres düzeyleri ile ilgili maddeler de dikkate alınarak, uyku kalitesi ile ilgili 4 haftalık ve aylık ölçümler karşılaştırmalı olarak değerlendirilmiştir.



Şekil 1. Çalışma 2'nin Ölçme Deseni

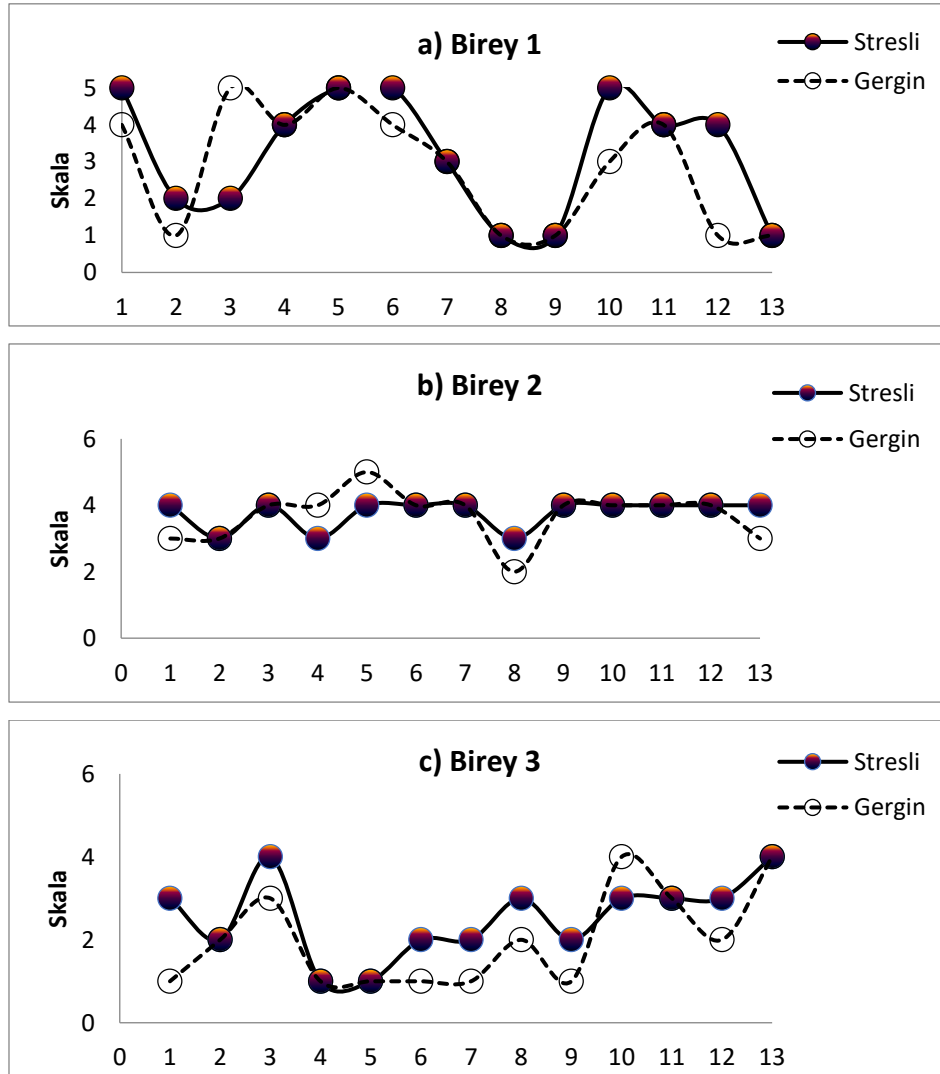
### 2.2.3. Örnek çalışma 3 için ölçme deseni, örneklem, veri toplama araçları ve verilerin analizi

Örnek çalışma 3 için ölçme deseni sekiz haftalık bir zaman dilimini kapsamaktadır. Bu sekiz haftada, 80 gönüllü üniversite öğrencisinden oluşan katılımcı grubuna birisi boylamsal olan, üç ayrı ölçek çevrimiçi olarak uygulanmıştır. Bu ölçeklerden tek uygulamalı olan iki ölçek Eğitim Bilimleri alanyazınında dayanıklılığın ölçülmesinde oldukça yaygın kullanılan ve bu çalışmada uygulamaya da bu nedenle seçilen Connor-Davidson Psikolojik Dayanıklılık (Campell-Sills ve Stein, 2007) ve Kısa Psikolojik Dayanıklılık (Smith, Dalen, Wiggins, Tooley, Christopher ve Bernard, 2008) ölçekleridir. Boylamsal olan üçüncü ölçek ise, 8 hafta boyunca ve haftalık uygulanan Duygu Cetveli Ölçeğidir. Bu ölçekte, katılımcılardan her hafta başı, "hiç" için 1'den "çok" için 5'e kadar rakamlar vererek o günkü psikolojik dayanıklılık düzeylerini ve bir dizi diğer duyu durumlarını belirtmeleri istenmiştir. Katılımcılara, ayrıca her hafta ciddi bir yaşam zorluğu yaşayıp yaşamadıkları da sorulmuştur. Bunun sebebi ise, psikolojik dayanıklılığın zorluk, travma, trajedi, tehdit ve stres yaşama ve bu durumlara uyum sağlama süreci ile ilgili olmasıdır (APA, 2014). Psikolojik dayanıklılığı ölçmek üzere kullanılacak ölçme modelinin bir zorluk/stres yaşayıp yaşamama durumunu ve bu durumun sonrasındaki süreci kapsayacak nitelikte olmasının, ilgili özellik hakkında toplanacak bilgiyi ve elde edilen ölçümlerin gerçeği yansıtmaya oranını arttıracığı düşünülmüştür.

### 3. BULGULAR

#### 3.1. Örnek çalışma 1 - Üniversite öğrencilerinin günlük stres ve gerginlik durumlarının 13 günlük seyri için bulgular

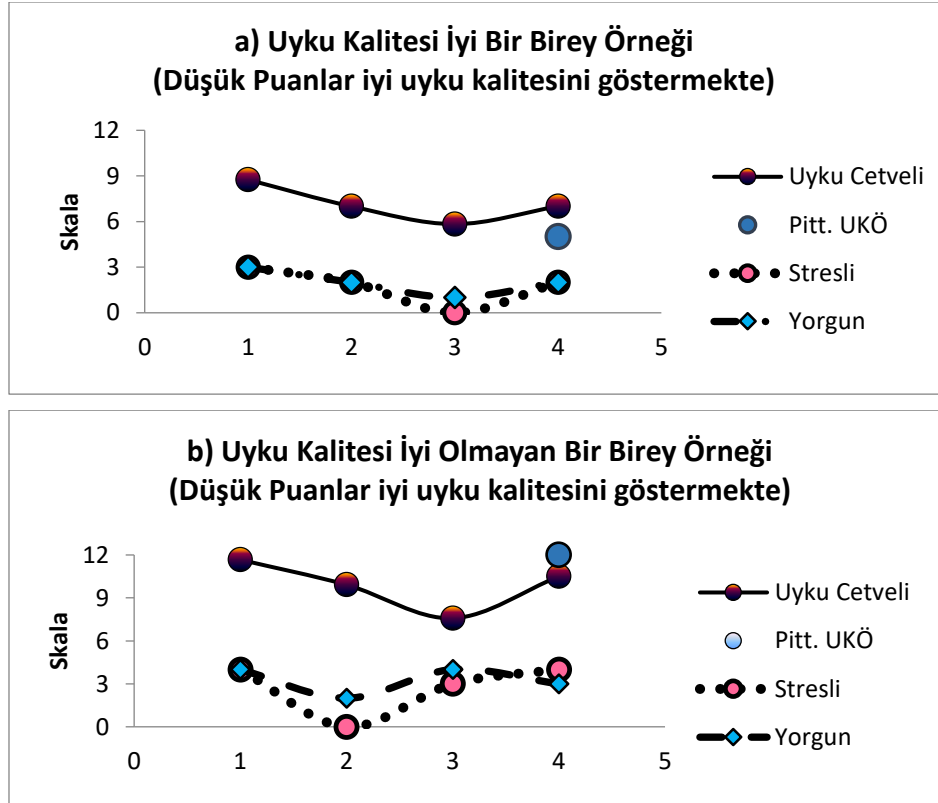
Tekrarlı ölçümlerin alınması durumunda, ilgilenilen özellikler ile ilgili ölçme sonuçlarının çizebileceği resmin, tek sefer ya da sınırlı sayıda alınan ölçümlere göre, ne kadar daha detaylı olabileceğini (Örn., Sliwinski, et al. 2009) örneklendiren bu örnek çalışmanın konusu, stres ve gerginlik duygularının 13 günlük seyrinin izlenmesi ile ilgilidir. Şekil 2, çalışma grubundaki gönüllü 64 Eğitim Fakültesi öğrencisinden tesadüfi olarak belirlenen üç bireyin 13 gün boyunca her akşam o gün için işaretledikleri stres ve gerginlik düzeylerini göstermektedir. Görülmektedir ki, bu üç birey, stres ve gerginlik düzeyleri ile ilgili olarak birbirlerinden oldukça farklıdır, ancak, bu farklılıklar bireyler-arası farklılıklarla sınırlı kalmamıştır. Birey 1'in stres ve gerginlik düzeyleri 1-5 skalasında birkaç günlük döngüler halinde inip çıkmışken, Birey 2 için kısmen yüksek ve sabit kalmış, Birey 3 için ise ilk 3 gün için küçük ve hızlı bir yükselme göstermiş, sonra minimuma düşmüş ancak takip eden 8-9 günde yavaşça yükselerek, maksimuma ulaşmasa da, tırmanmıştır. Bir başka deyişle, bu üç birey hissettikleri stres ve gerginlik düzeyleri açısından birbirlerinden farklı oldukları kadar, hissettikleri stres ve gerginlik düzeylerinde (kendi içlerinde) gösterdikleri zamansal değişimler açısından da birbirlerinden farklıdır. Bireylerin duygularının günlük yaşantılarında karşılaştıkları güçlüklerle ilgili olarak değişkenlik gösterebileceği gerçeği göz önüne alındığında, 13 gün boyunca alınan tekrarlı gözlemler ile çizilen bu resmin, bu bireylerin (verilen 13 gün aralığından seçilecek herhangi bir gün için alınmış olabilecek) tek seferlik gözlemleri ile çizilebilecek bir resimden çok daha gerçekçi ve anlamlı olduğu açıktır. Dahası, eğer bu üç bireye de stres ve gerginlikle baş etme konularında yardım ve destek gerekse idi, tek seferlik bir ölçüm üzerinden hazırlanabilecek bir destek ve yardım programı yerine, tekrarlı alınmış ölçümler üzerinden hazırlanacak "bireyselleştirilmiş" bir destek ve yardım programı çok daha etkili sonuçlar verebilecektir.



Şekil 2. 13 Günlük Stres ve Gerginlik Düzey ve Değişimleri

### 3.2. Örnek çalışma 2 - Bireylerin uyku alışkanlıkları ve duygu durumları arasındaki ilişkiye boylamsal bir bakış için bulgular

Şekil 3, çalışma grubundan (a) uyku kalitesi iyi olan ve (b) olmayan olarak ayrılan iki grubu örneklendirmek üzere seçilen birer kişi için elde edilen haftalık bulguları göstermektedir (Sözer, Çavdar & Kahraman, 2017). Şekil 3'ün iki kişi üzerinden özetlediği genel bulgular göstermektedir ki, (1) bireylerin uyku kaliteleri haftadan haftaya değişim gösterebilmektedir, (2) uyku kalitesi düşük olan grupta yer alan bireyin uyku kalitesi, zamanla değişime daha açıktır ve (3) bireylerin 4. haftada aldıkları Pittsburgh Uyku Kalitesi Ölçek puanları, beklendiği gibi önceki 4 haftalık zamandan ziyade, 4. haftadaki Uyku Cetveli Ölçeği puanlarına yakındır. Bu ölçümlere eşlik eden 4 haftalık stres ve yorgunluk düzeyi ölçümleri ve bu ölçümlerde 4 haftalık zamanda gözlenen değişimler de haftalık uyku kalitesi ölçümlerinin, tek seferlik (aylık) uyku kalitesi ölçümüne göre bireylerin uyku kalitesi durumlarını tasvir etmede ve duygu durumlarında meydana gelen değişimleri yakalamada daha kullanışlı olabileceğini destekler niteliktedir.



Şekil 3. Uyku Kalitesi İyi Olan ve İyi Olmayan İki Ayrı Birey İçin Alınan Haftalık Gözlemler

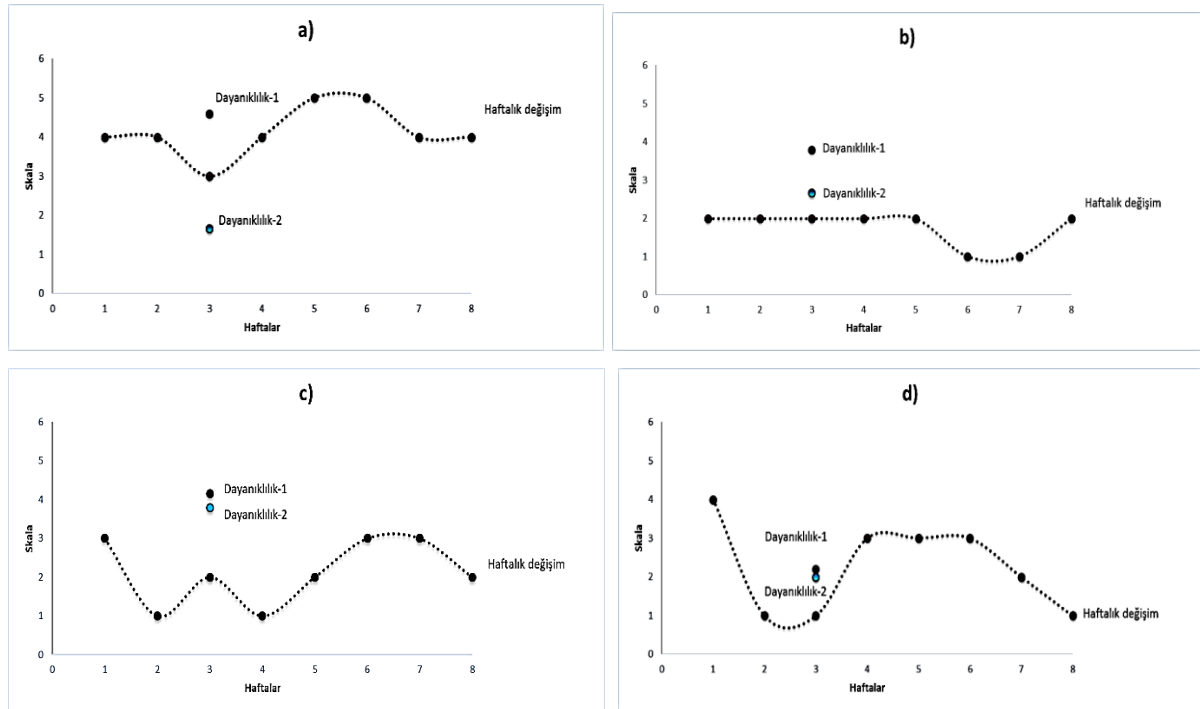
### 3.3. Örnek çalışma 3 - Bireylerin psikolojik dayanıklılık düzeylerinin boylamsal analizi için bulgular

Bireylerin karşılaştıkları zorluklarla baş edebilmesine yardımcı olan yetenek ve karakter özelliklerinin bir kombinasyonu olan psikolojik dayanıklılığın bireyin içinde bulunduğu durumlara ve zamana göre değişebildiği teorik ve ampirik olarak desteklenmektedir (Masten, Cutuli, Herbers ve Reed, 2009; Werner, 2013; Werner ve Smith, 1992). Dinamik bir yapısı olan bu özelliğin, tek seferlik ölçme uygulamaları ile ölçülmesi, bir birey için psikolojik dayanıklılığın birey-durum etkileşiminin yalnızca tek bir kesitinin alınmasına sebep olmaktadır. Tek sefer uygulamalar yerine tekrarlı ölçümleri içeren boylamsal bir ölçme modeli kullanıldığında ise birey-durum etkileşiminin farklı noktalarındaki kesitler alınarak birey hakkında yapılacak çıkarımlara dayanak oluşturan bilgi zenginliği artırılabilir. Bu araştırma örneğinde bireylerin psikolojik dayanıklılık düzeyleri tekrarlı ve tek sefer ölçülerek, bu özelliğin zaman içinde değişen bir özellik olup olmadığı, eğer değişiyorsa bu değişim örüntüsünün bireyler arası farklılık gösterip göstermediği ve tek sefer ve tekrarlı ölçümler arasında fark olup olmadığı incelenmiştir. Yanıt aranan bu sorulara cevap bulabilmek için bireylerin zaman içinde değişimini açıkça yansıtan grafikler kullanılmıştır.

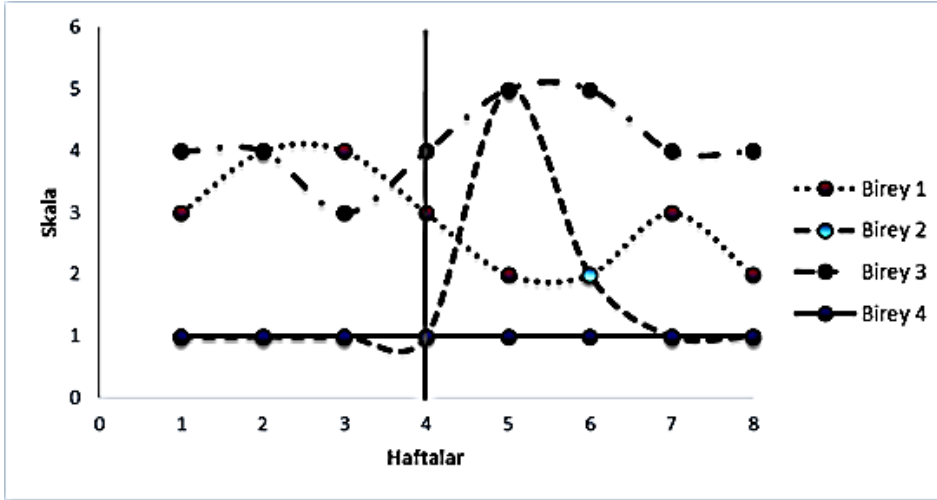
Psikolojik dayanıklılığın zaman içinde değişen bir özellik olup olmadığını, (varsa) bu değişimin bireyler-arası farklılaşıp farklılaşmadığını ve tekrarlı ölçümler ile tek sefer ölçümler arasındaki uyumu görebilmek adına grafiksel analizler yapılmıştır. Şekil 4 bu analizlerde sıkça gözlenen bir durumun görselini, çalışma grubundan amaçlı seçilen dört ayrı birey üzerinden göstermektedir. Şekil 4'te ilk göze çarpan, bu bireylerin dayanıklılık düzeylerinin haftadan haftaya azımsanmayacak düzeyde bir değişim gösterdiği.

Bununla birlikte, psikolojik dayanıklılık düzeylerindeki değişim bireyden bireye farklılık göstermektedir ki bu durum boylamsal ölçme modelinin hem birey-içi hem de bireyler arası farklılıkları yansıtmada sağladığı katkıları göstermektedir. Diğer dikkate değer durum ise, bu bireylerin Connor-Davidson Psikolojik Dayanıklılık ve Kısa Psikolojik Dayanıklılık ölçeğinden elde edilen psikolojik dayanıklılık puanlarının hem kendi içlerinde birbirleri ile hem de Duygu Cetveli Ölçeğindeki dayanıklılık düzeyi geneli ile çok da tutarlı olmadığıdır (Çavdar, Sözer & Kahraman, 2017).

Şekil 5, çalışma grubundan seçilen dört birey üzerinden, yapılan analizlerde gözlenen dört ayrı dayanıklılık örgüsünü temsil eden bireyleri örneklendirmektedir. Bu bireyler, 8 haftalık gözlem sırasında 4. hafta ciddi bir yaşam zorluğu yaşadıklarını ifade etmiş bireylerden seçilmiştir. Şekil 5, bu bireylerin 4. hafta yaşadıkları zorluk öncesi ve sonraki haftalar için hissettikleri dayanıklılık düzeyleri göstermektedir. Birey 1, hissettiği dayanıklılık duygu durumu bakımından haftalar arası hafif dalgalanmalar yaşasa da, kendini dayanıklı hisseden bir bireyken, yaşadığı zorluk sonrası, kendisini daha dayanıksız hissetmektedir. Birey 2, zorluk öncesinde kendini dayanıksız hissetmekte, ancak zorluklarla karşılaştığında kendini daha dayanıklı hissetmektedir. Birey 3, zorluk yaşamamasına rağmen kendini genel olarak dayanıklı hissetmektedir. Birey 4 ise, zorluk öncesi ve sonrası fark etmeksizin kendini aynı ve dayanıksız hissetmektedir. Bulgular göstermektedir ki bireylerin dayanıklılık düzeyleri, ilgili ölçümlerin alındığı zaman zarfında gerçek hayatlarında deneyimledikleri ciddi bir zorluk olup olmadığına göre, zaman içerisinde değişim gösterebilmektedir. Buna ek olarak, zorluk sonrasındaki süreç izlendiğinde bireyler farklı değişim örüntülerine sahiptirler. Buna göre, psikolojik dayanıklılığın ölçülmesinde hem birey-içi hem de bireyler-arası farklılıklar ile bu farklılıkların sebeplerini belirlemeye olanak sağlayan boylamsal bir yaklaşımın izlenmesi (Baltes ve Nesselroade, 1979) geçerliğin artırılması bakımından oldukça önemlidir.



Şekil 4. Dört bireyin sekiz hafta boyunca dayanıklılık durumu ve bu bireyin tek seferlik uygulamalı Kısa (Dayanıklılık-1) ve Connor-Davidson (Dayanıklılık-2) Psikolojik Dayanıklılık ölçeklerinden aldığı psikolojik dayanıklılık puanları



Şekil 5. Sekiz haftalık bir zaman diliminde, 4. haftada ciddi bir yaşam zorluğu yaşadığını ifade eden dört ayrı bireyin her hafta 5 üzerinden ifade ettiği dayanıklılık durumları

#### 4. TARTIŞMA ve SONUÇ

Eğitimcilerin uygulamada karşılaştıkları zorluklardan belki de en önemlisi, sosyal-duygusal öğrenmeler ve çıktıkları olarak paydaşlar tarafından önemsenen bilişsel-olmayan becerileri ve bu becerileri etkileyen faktörleri sağlıklı ve anlamlı bir şekilde ölçme ve değerlendirme konusundadır. Geleneksel ölçme desenlerinin (özellikle de zaman içerisinde değişim/gelişim göstermesi beklenen bilişsel ve duyuşsal beceriler söz konusu olduğunda) bireysel farklılıkları değerlendirilmede artık çok da ideal olmayan uygulamalar oldukları düşünülmektedir. Bu makalede önerilen çözüm ölçme uygulamalarının bir süreç olarak tasarlandığı ve uygulandığı boylamsal ölçme yaklaşımıdır. Boylamsal ölçme modellerinin kullanılması ile, ölçümlerin tek-seferlik alınmasının uygulamalara getirdiği sınırlılıklar tespit edilebilecek, değerlendirmeye konu olan özellikler zaman içerisindeki değişim/gelişimleri açısından izlenebilecek ve güçlü geçerlilik kanıtları elde edilebilecektir. Kısacası, bireylerin daha değişken olan duygu durumları ile daha durağan olan duyuşsal becerileri, her ikisi arasındaki geçişleri ve zaman içerisindeki değişimleri, diğer bilişsel beceriler veya demografik faktörler ile ilişkileri daha sağlıklı bir biçimde tespit edilebilecek ve değerlendirilebilecektir.

Ölçme uygulamaları ile hedeflenen, elde edilen ölçümlerin ölçülmek istenen özelliğin gerçek doğasını yansıtabilmesidir. Ölçme sonuçlarından elde edilen puanların sağlıklı bir şekilde yorumlanmaları ve pratikte uygulamalar için anlamlılıkları hem araştırmacılar hem de yöneticiler açısından büyük önem taşımaktadır (Messick, 1989, 1995). Bu nedenle, eğitim ve psikolojide ölçmeye konu olan özelliklerin gelişimini bir süreç olarak izlemeye daha uygun olan bir ölçme deseni (boylamsal) kullanılarak, elde edilen sonuçların (bir süreç içinde anlamlı olabilecek) güvenilirlik ve geçerlik kanıtları ile desteklenmesi önemlidir. Bu makalede önerilen ve örnekler üzerinden tartışılan boylamsal ölçme yaklaşımı, söz konusu ihtiyacı karşılamaya yönelik AR-GE aşamasında olan yeni bir girişimdir. Bireylerin okul, iş ve özel yaşamlarındaki "başarı" kavramının "iyi oluş" kavramına evrilmesi gerektiği konusunda hem fikir olduğumuz çağımız paradigmasında (New Visions for Education, World Economic Forum®, <https://www.weforum.org/reports>), sosyal-duygusal beceriler ve önemli öğrenme çıktıları ile ilişkileri araştıran araştırmacıların, gruplar-arası, bireyler-arası ve bireyler-içi farklılıkları keşfetmede kullanışlı, değişime duyarlı, belli güvenilirlik ve geçerlilik kriterlerini sağlayan alternatif ölçme araçlarına ve değerlendirme modellerine her zamankinden çok daha fazla ve acilen ihtiyaçları vardır.

Tartışmasızdır ki, boylamsal ölçme modellerinin bireysel farklılıkları ölçme ve değerlendirmede kullanılan yöntem ve elde edilen sonuçlar için yapılacak güvenilirlik ve geçerlilik argümanlarına getireceği zenginlik, ölçmeye konu özelliklerin zaman içerisinde değişim/gelişime ne kadar açık oldukları ile ilgilidir (Boylamsal modeller ile ilgili daha fazla bilgi için bkz. Muthén and Curran, 1997). Eğer ölçmeye konu özelliklerin zamanla değişebilir/gelişebilir olduklarını destekleyen teori veya deneyimler varsa, boylamsal ölçme desenleri toplanan veriler üzerinden yapılacak bireyler-arası karşılaştırmaları, gözlenen birey-içi değişim/gelişim eğrilerini ekleyerek zenginleştirecektir (Avey, Luthans ve Mhatre, 2008). Ancak, ilgilenilen özelliklerin değişebilir/gelişebilir olduklarını destekleyen teori veya deneyimler yoksa veya teoriye bağlı olarak beklenen değişim eğrileri ampirik olarak desteklenmiyorlar ise, elbette ki, boylamsal modellerin tek seferlik ölçme modelleri üzerine katacağı bir değer olamayacaktır. Planlanan araştırma ve geliştirme çalışmalarının odaklandığı ölçme problemlerinin ve mevcut diğer şartlara göre uygulanabilirliği olan araştırma yöntemlerinin her uygulama için farklı olması kaçınılmazdır. Bu anlamda, boylamsal ölçümleri de içerecek olan araştırma desenlerinin tasarlanması ve uygulanması aşamalarında, araştırmacıların mevcut bilgi ve birikimleri sorgulamaya açık olmaları, çok-disiplinli

ve yeniliğe/gelişime açık bir bakış açısı kullanmaları oldukça önemlidir. Elde edilen bulguların mevcut literatürü tam olarak desteklemediği durumlarda ihtiyaç duyulacak gerçekçi ve güncel çıkarımlara ulaşma ve sentez becerisi tüm eğitim camiası için elzem bir beceridir.

**KAYNAKÇA**

- American Psychological Association. (2014). *The road to resilience*. Washington, DC: American Psychological Association. 24.9.2018 tarihinde <http://www.apa.org/helpcenter/road-resilience.aspx> adresinden alınmıştır.
- Avey, J. B., Luthans, F., & Mhatre, K. H. (2008). A call for longitudinal research in positive organizational behavior. *Journal of Organizational Behavior: The International Journal of Industrial, Occupational and Organizational Psychology and Behavior*, 29, 705-711.
- Baltes, P. B., & Nesselroade, J. R. (Eds.) (1979). *History and rationale of longitudinal research*. In J. R. Nesselroade & P. B. Baltes, *Longitudinal research in the study of behavior and development* (ss. 21–27). New York: Academic Press.
- Bloom, B.S. (Ed) (1969). Some theoretical issues relating to educational evaluation. In R. W. Taylor, *Educational evaluation: New roles, new means: The 68th yearbook of the National Society for the Study of Evaluation, Part II* (ss. 26-50). Chicago: University of Chicago Press.
- Bonanno, G. A., Galea, S., Bucciarelli, A. ve Vlahov, D. (2007). What predicts psychological resilience after disaster? The role of demographics, resources, and life stress. *Journal of Consulting and Clinical Psychology*, 75, 671-682.
- Borgonovi, F. & Pál, J. (2016). A framework for the analysis of student well-being in the PISA 2015 study. Paris: OECD Publishing. 22.2.2019 tarihinde <http://disde.minedu.gob.pe/bitstream/handle/123456789/4863/A%20Framework%20for%20the%20Analysis%20of%20Student%20Well-Being%20in%20the%20Pisa%202015%20Study%20Being%2015%20In%202015.pdf?sequence=1&isAllowed=y> adresinden alınmıştır.
- Budria, S., & Ferrer-i-Carbonell, A. (2012). *Income comparisons and non-cognitive skills*. Discussion Paper Series, 1-29. 24.9.2018 tarihinde <https://www.econstor.eu/bitstream/10419/59024/1/715684280.pdf> 24 adresinden alınmıştır.
- Buysse, D. J., Reynolds, C. F., Monk, T. H., Berman, S.R. & Kupfer, D.J. (1988). The Pittsburgh Sleep Quality Index: A New Instrument for Psychiatric Practice and Research. *Psychiatry Research*, 28, 193-213.
- Campbell-Sills, L. & Stein, MB. (2007). Psychometric analysis and refinement of the Connor Davidson Resilience Scale (CD-RISC): Validation of a 10 item measure of resilience. *Journal of Trauma Stress*, 20, 1019-1028.
- Collie, R. J., Martin, A. J., Bottrell, D., Armstrong, D., Ungar, M. ve Liebenberg, L. (2017). Social support, academic adversity and academic buoyancy: a person-centered analysis and implications for academic outcomes. *Educational Psychology*, 37, 550-564.
- Conger, R. D. ve Conger, K. J. (2002). Resilience in Midwestern families: Selected findings from the first decade of a prospective, longitudinal study. *Journal of Marriage and Family*, 64, 361-373.
- Crane, M., Lewis, V., Cohn, A. D., Hodson, S. E., Parslow, R., Bryant, R. A. ve Forbes, D. (2012). A protocol for the longitudinal study of psychological resilience in the Australian Defence Force. *Journal of Military and Veterans Health*, 20, 36.
- Cranford, J. A., Shrout, P. E., Iida, M., Rafaeli, E., Yip, T., & Bolger, N. (2006). A procedure for evaluating sensitivity to within-person change: Can mood measures in diary studies detect change reliably? *Personality and Social Psychology Bulletin*, 32, 917-929.
- Crossley, J. G. (2017). Assessing the Non-Cognitive Domains: Measuring What Matters Well. In *Innovative Practices for Higher Education Assessment and Measurement* (ss. 348-372). IGI Global.
- Curcio, G., Ferrara, M. & Gennaro, L.D. (2006). Sleep loss, learning capacity and academic performance. *Sleep Medicine Reviews*, 10, 323–337.
- Çavdar, D., Sözer, E. & Kahraman, N. (2017, Nisan). *Dayanıklılık için tekrarlı ölçümleri içeren boylamsal bir ölçekleme modeli üzerine bir değerlendirme*. 26. Uluslararası Eğitim Bilimleri Kongresi'nde sunulmuş bildiri. Pagem, Antalya.
- Funder, D. C. (2006). Towards a resolution of the personality triad: Persons, situations, and behaviors. *Journal of Research in Personality*, 40, 21-34.
- Gutman, L. M., & Schoon, I. (2013). The impact of non-cognitive skills on outcomes for young people. *Education Endowment Foundation*, 59. 22.2.2019 tarihinde <https://pdfs.semanticscholar.org/f4a5/2db3001fb6fb22eef5dc20267b5b807fd8ff.pdf> adresinden alınmıştır.
- Jackson, C. K. (2012). Non-cognitive ability, test scores, and teacher quality: Evidence from 9th grade teachers in North Carolina (No. w18624). National Bureau of Economic Research. 22.1.2019 tarihinde <https://www.nber.org/papers/w18624> adresinden alınmıştır.

- Kautz, T., Heckman, J. J., Diris, R., Ter Weel, B., & Borghans, L. (2014). *Fostering and measuring skills: Improving cognitive and non-cognitive skills to promote lifetime success* (No. w20749). National Bureau of Economic Research. 22.2.2019 tarihinde <https://www.nber.org/papers/w20749.pdf> adresinden alınmıştır.
- Masten, A. S., Cutuli, J. J., Herbers, J. E., & Reed, M. G. J. (2009). Resilience in Development. In *The Oxford Handbook of Positive Psychology*, (2 Ed.) Oxford University Press.
- Messick, S. (1989). Validity. In R. L. Linn (Eds.), *Educational measurement* (3. Ed.) New York: American Council on Education and Macmillan.
- Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist*, 50, 741-749.
- Muthén, B. O., & Curran, P. J. 1997. General longitudinal modeling of individual differences in experimental designs: A latent variable framework for analysis and power estimation. *Psychological Methods*, 2, 371.
- Norris, F. H., Tracy, M. ve Galea, S. (2009). Looking for resilience: Understanding the longitudinal trajectories of responses to stress. *Social Science & Medicine*, 68, 2190-2198.
- Salkind, N. J. (Ed.). (2010). *Encyclopedia of research design* (Vol. 1). Sage.
- Sayer, A.G., & Cumsille, P. E. 2001. Second-order latent growth models. L.M. Collins & A.G. Sayer (Ed.), *New methods for the analysis of change* (ss. 179-200). Washington, DC: American Psychological Association
- Scriven, M. (1967). The methodology of evaluation. R. Tyler, R. Gagne, & M. Scriven (Ed.), *Perspectives of curriculum evaluation* (ss. 39-83). Chicago: Rand McNally
- OECD, PISA, The Organization for Economic Co-operation and Development. 24.9.2018 tarihinde <http://www.oecd.org/pisa/> adresinden alınmıştır.
- Sliwinski, M. J., Almeida, D. M., Smyth, J., & Stawski, R. S. (2009). Intra-individual change and variability in daily stress processes: Findings from two measurement-burst diary studies. *Psychology and Aging*, 24, 828-840.
- Smith, B.W., Dalen, J., Wiggins, K., Tooley, E., Christopher, P., & Bernard, J. (2008). The brief resilience scale: assessing the ability to bounce back. *International Journal of Behavioral Medicine*, 15, 194-200.
- Sözer, E., Çavdar, D. & Kahraman, N. (2017, Nisan). *Bireylerin uyku kaliteleri ve duygu durumları ile ilgili duyuşsal becerileri arasındaki ilişkilerin boylamsal analizi*. 26. Uluslararası Eğitim Bilimleri Kongresi'nde sunulmuş bildiri. Pagem, Antalya.
- Rutter, M. (1993). Resilience: Some conceptual considerations. *Journal of Adolescent Health*, 14(8), 626-631.
- Werner, E. E., & Smith, R. S. (1992). *Overcoming the odds: High risk children from birth to adulthood*. New York: Cornell University Press.
- Werner, E. E. (2013). *What can we learn about resilience from large-scale longitudinal studies?* Handbook of resilience in children (ss. 87-102). Springer US.
- Widaman, K. F., Ferrer, E. & Conger, R. D. (2010). Factorial invariance within longitudinal structural equation models: measuring the same construct across time. *Child Development Perspectives*, 4, 10-18.

## EXTENDED ABSTRACT

### 1. Introduction

This paper proposes a longitudinal assessment approach to be used in the assessment of non-cognitive constructs. Focusing on how to evaluate “state” versus “trait” properties of propensities or skills from the affective domain, a comparative narrative is provided attesting to the utility of employing a longitudinal perspective. Illustrations on how to collect, analyze and interpret time-series data are provided using some of the results from three separate studies. The data were from a larger longitudinal study, referred to as the Emotion Ruler Field Study, aiming to investigate emotional aspects of learning and social-emotional skill repertoire of college of education students. First application shows daily self-ratings data of three students on their stress and tenseness levels collected over thirteen days illustrating how the differences inferred across days might differ when based on assessments taken only once versus over time. The second and the third studies use weekly self-assessment data from two different studies investigating the amount of variance over time for “sleep quality” and “psychological resilience” of college students, respectively. The applications illustrate how inferences made about the relationships between the measures of interest might evolve to different heights when over time (within-persons) variances are taken into account when investigating individual differences (between-persons). After a brief description of the Emotion Ruler Field Study, the following text summarizes the assessment designs and some of the comparative findings from the three illustrative studies.

### 2. Method

Illustration data sets used in this study are from a larger data set collected in the years of 2016-2017 using an on-line longitudinal research and development platform called the Emotion Ruler Field Study. In addition to be used for developing and testing new item formats with visual and audial aids that only an on-line platform would permit, the Emotion Ruler Field Study initiative was launched for developing, applying and testing new and advanced longitudinal assessment designs and to address some of the psychometric challenges encountered when measuring constructs with a dynamic factorial structure (i.e., show construct or component variance over time). Utilizing either a new item format or a new longitudinal assessment design, or both, all the research projects of the research and development initiative were designed to tackle a psychometric problem and look for means of improvement. Focusing on the issues in teacher education in particular, the research projects of the platform to date have been recruiting college of education students as volunteer subjects.

The first illustration provided shows how the inferences made about individual differences may differ when based on a single-take versus on a time-series assessment data. Graphical representations are provided showing the daily ratings of three college students that were randomly sampled from a group of 64. All the participants in the study were asked to rate themselves on a list of emotional states, which included how stressed and tense they felt every day, using a rating scale of 1 to 5 (1 indicating “none or very little” and 5 indicating “a lot”). The students’ daily ratings varied greatly over time. Subject 1 seemed to be going through a couple of ups and downs over the days with a day “lag” when transitioning from feeling stressed to feeling tense or vice versa. The data suggest that this might be due to the student having had some rough days around the 5th and 11th days of the study. Subject 2 was as stressed and tense as Subject 1, however, it seemed s/he was steadily feeling stressed and tense (i.e., every day over the 13 days without showing much fluctuations). Subject 3 was not as stressed or tense as subject 1 and 2 or as stable as the Subject 2, yet s/he seemed to have an upward trend starting from the 4th day of the study and her/his upward trend was a bit more linear for stress ratings than that of tenseness ratings. The visuals provided nicely illustrate that the inferences about how stressed and how tense these three subject have felt and how they would compare among themselves would be interpreted quite differently, if and when the data were collected by a single-take administration, for example, using any given day out of the 13.

The second and the third studies utilizing data sets with weekly assessments are used to discuss the advantages of using longitudinal assessment designs when estimating the influence of extraneous or relevant factors on growth or change over time. The second study models the influence of sleep quality on college students’ perceived sense of well-being and utilizes a four-week longitudinal measurement design, while the third study models the variance over time for psychological resilience ratings and utilizes an eight-week assessment design. The study groups of these two studies consisted of college of education students who volunteered to do weekly self-ratings on several 1 to 5 rating scales. The study groups and the scales used in each study were specific to the study.

The data set of Study 2 including 69 students’ weekly ratings on their sleep quality and sense of well-being was analyzed using a Multi Trait Multi Method matrix and the Latent Growth Curve analysis to investigate the degree to which these ratings were correlated over the four weeks of the study and to model the degree to which “sleep quality” influenced students’ sense of “well-being” over time. The findings were supportive of the usefulness a

longitudinal assessment design in that students' sleep quality were subject to within-person differences over the weeks. When taken into account, they were helpful in distinguishing students who were likely to have a consistent sleep issue interfering with their overall sense of well-being. When they were ignored, the results seemed to have shrunk to being rough summaries of what was being observed over the four weeks. The data set of the Study 3 including 80 students was selected from a larger group of individuals having indicated that they had a difficult life experience in the fourth week of the study. The data set then was analyzed using a cluster analysis and the Latent Growth Curve Modeling. The findings were supportive of the use of a longitudinal assessment design in that most students' resilience ratings varied over the eight weeks. Moreover, there were before and after "the fourth week" differences for the majority of students. The analysis of the variations observed in the data set as before and after the 4th week suggested there were four distinct sub-groups of students. The illustration shows that, if not observed longitudinally, an individual's ratings on a given week might have reflected the effect of a time lag (i.e., how long it was since a recent life challenge was experienced), rather than the individual's capacity of bouncing back. The findings overall support that psychological resilience might be one of those constructs for which the assessments must be taken longitudinally and along with a precursor variable helping the researchers properly place individuals' life experience into the process being observed.

### **3. Findings, Discussion and Results**

In spite of the recent trends in educational research and policy development studies around the world highlighting the importance of non-cognitive traits, such as, emotional balance, psychological resilience and well-being as priority learning outcomes, longitudinal measurement models that would be the most suited to study such constructs and their transitions over time have not yet been fully explored. The current study takes on the issue from a psychometric point of view and explores the usefulness of a longitudinal assessment designs.

Potential validity threats with respect to the issue of unaccounted "change over time" in the assessments often go unchecked in most educational and psychological measurement studies that use cross-sectional (single-administration) measurement models, where, by design, within-person differences cannot be observed. Hence, questions about "how long the meaning of any of the collected measures at a particular point of time will generalize over time?" often go unanswered, as well as, those about "trait versus state properties" of various competencies. This paper draws attention to some of the potential advantages of using a longitudinal perspective and provides tools and tips that might be useful in practice. Illustrations are provided to underline the importance of using assessment designs with repeated measurement and how such assessment plans may help address some of the important validity issues encountered in practice, for example, when trying to predict construct-relevant and construct-irrelevant components of variance when observed over time (Messick, 1989, 1995).

It is argued that a longitudinal measurement perspective might be superior to a cross-sectional one when state versus trait nature of constructs are to be questioned, especially for those that have a complex factorial structure and show a non-trivial amount of variance over time. Albeit limited, the illustrations provided in this study support that it pays off to develop and use longitudinal assessment designs if there is growth or change in the constructs of interest over time. Overall, a longitudinal perspective is recommended for the researchers in the field of educational and psychological assessment who are working with constructs with a complex and/or dynamic structure and are aware of the potential validity threats when using strictly cross-sectional assessment designs to measure them.