

Uzaktan Eğitim ile Yürütülen İngilizce Yabancı Dil Dersi Öğretim Programını Değerlendirme Ölçeği: Bir Geçerlik ve Güvenirlik Çalışması

Assessment Scale of English Language Course Curriculum Conducted by Distance Education: A Validity and Reliability Study

Ali ORHAN, Şule ÇEVİKER AY

ÖZ

Bu araştırmada, öğrencilerin uzaktan eğitim ile yürütülen İngilizce dersi öğretim programı hakkındaki görüşlerini belirlemek için Stufflebeam'e ait CIPP modelinin - bağlam, girdi, süreç ve ürün değerlendirme - prensiplerinden hareketle hazırlanan ölçeğin geçerliğinin ve güvenilirliğinin belirlenmesi amaçlanmıştır. Kapsam ve görünüş geçerliğini sağlamak için uzman görüşüne başvurulmuştur. Kapsam, görünüş ve dil geçerliği sağlanan ölçek, 2015-2016 akademik yılı güz döneminde, Bülent Ecevit Üniversitesi İlahiyat Fakültesi'nde ve Fen-Edebiyat Fakültesi Türk Dili ve Edebiyatı bölümünde öğrenim görmekte ve İngilizce Yabancı Dil I dersini almakta olan 250 öğrenciye uygulanmıştır. Ölçeğin yapı geçerliğini sağlamak için yapılan açımlayıcı faktör analizi sonucunda dört faktörlü 36 maddelik bir ölçek elde edilmiştir. Bu dört faktör, varyansın %62.69'unu açıklamaktadır. Dört boyuta CIPP modeline ait bağlam değerlendirme, girdi değerlendirme, süreç değerlendirme ve ürün değerlendirme isimleri verilmiştir. Ölçeğin güvenirlilik düzeyini saptamak için her boyuta ve toplam ölçeğe ait Cronbach alfa katsayılarına bakılmıştır. Ölçeği oluşturan alt boyutların güvenirlilik katsayıları 0.87 ile 0.94 arasında değişmektedir. Ayrıca bütün ölçeğin güvenirlilik katsayısı da 0.96 bulunmuştur. Ölçekte bulunan maddelerin madde toplam korelasyonları ise 0.50 ve 0.76 arasında değişmektedir. Bu bulgulara dayanarak ölçeğin geçerli ve güvenilir ölçümler üreten bir ölçme aracı olduğu ve öğrencilerin uzaktan eğitim ile yürütülen İngilizce dersi öğretim programına ilişkin görüşlerini saptamak için kullanılabileceği söylenebilir.

Anahtar Sözcükler: Uzaktan eğitim, İngilizce öğretim programı, Ölçek geliştirme, CIPP, Program değerlendirme

ABSTRACT

The aim of this study is to determine the validity and reliability of the evaluation scale developed by the researchers based on the principles of Stufflebeam's CIPP Evaluation Model to investigate the students' satisfaction levels of foreign language distance education curriculum. To explore the content and face validity of the scale, opinions of the experts were asked. The scale, whose content, face and language validity checks were performed, was applied to 250 students, whose were studying in Bülent Ecevit University Faculty of Theology and Faculty of Arts and Sciences Department of Turkish Language and Literature taking English Language I course during 2015-2016 fall term. As a result of the explanatory factor analysis, which was conducted to explore construct validity of the scale, it was found out that scale had four factors consisting of context, input, process and product components and had 36 items. The total variance explained by these four factors was %62.69. Examining whether each factor was measuring a single idea and whether the items that made up the factors were internally consistent, internal reliability data was obtained through Cronbach's alpha coefficient among the four factors. The reliability coefficients

Ali ORHAN (✉)

ORCID ID: 0000-0003-1234-3919

Bülent Ecevit Üniversitesi, Yabancı Diller Yüksekokulu, Temel İngilizce Bölümü, Zonguldak, Türkiye
Bülent Ecevit University, School of Foreign Languages, Department of Basic English, Zonguldak, Turkey
ali_orh_an@hotmail.com

Şule ÇEVİKER AY

ORCID ID: 0000-0002-9505-5105

Düzce Üniversitesi, Eğitim Fakültesi, Eğitim Bilimleri Bölümü, Düzce, Türkiye
Düzce University, Faculty of Education, Department of Educational Sciences, Duzce, Turkey

Geliş Tarihi/Received : 18.03.2017

Kabul Tarihi/Accepted : 02.05.2017

of four components ranged from 0.87 to 0.94 and the reliability coefficient of the whole scale was found as 0.96. Besides, item-total correlation of scale items ranged from 0.50 to 0.76. It was seen from the results that the scale is a valid and reliable curriculum evaluation instrument and can be used in the field of education to investigate the students' satisfaction levels on distance education curriculums of foreign language.

Keywords: Distance education, English curriculum, Scale development, CIPP, Curriculum evaluation

GİRİŞ

Birçok alanda meydana gelen teknolojik gelişmelerle birlikte artık insanların farklı kültür ve insanlarla iletişim kurma ihtiyacı yaşadığı ve bunun için de sadece ana dillerinin insanlara artık yetmiyor olduğu (Özkanal & Hakan, 2010) bir gerçektir. Bu küreselleşme sürecinde tüm dünyada, karşı konulamaz bir şekilde iletişim dili haline gelen (Mede & Uygun, 2014) İngilizcenin önemi artmaktadır. Devletler küresel dünyada var olabilecek, diğer kültür ve insanlarla iletişim kurabilecek yurttaşlara sahip olabilmek amacıyla İngilizcenin öğretimine bir hayli önem vermektedirler. Bu amaçla öğretim programlarını yenilemekte, farklı yöntem ve teknikler kullanmaktadırlar.

Günümüzde teknolojinin hızlı bir şekilde gelişmesine paralel birçok alanda görülen yenilikler eğitim dünyasına da yansımıştır. Teknolojiye paralel bir şekilde ülkelerin değişen eğitim anlayışı (Özer, 2011) ve daha geniş insan gruplarının daha maliyetsiz bir şekilde eğitim alması ve fırsat eşitliğini sağlama ihtiyacı (Cansu, 2010) uzaktan eğitimi, geleneksel eğitimin önemli bir alternatifi haline getirmiştir. Dünyada olduğu gibi Türkiye'de de ortaya çıkan bu yeni alternatifle birlikte bazı kurumlar öğretimlerini uzaktan eğitimle sağlama yoluna gitmeye başlamışlardır. Birçok alanda olduğu gibi İngilizce öğretiminde de uzaktan eğitim uygulamaları, Türkiye'de son zamanlarda yeni bir eğitim umudu olarak yaygınlaşmıştır. Uzun yıllar boyunca geleneksel olarak yüz yüze eğitimle sürdürülen İngilizce öğretimi konusunda Türkiye'nin başarılı bir geçmişi olmamıştır. Teknolojinin gelişmesi, artan öğrenci sayısı, öğretim elemanı sayısının yetersizliği vb. nedenlerle üniversitelerde ortak derslerde ve yabancı dil öğretiminde uzaktan eğitim denemeleri başlamıştır. Henüz tüm üniversitelere yayılmamış olan bu denemelerin başarısı yabancı dil öğrenme ve öğretme problemini aşma konusunda üniversitelere büyük katkı sağlayabilir. Bu nedenle uzaktan eğitimle yürütülen İngilizce öğretim programlarının ne derece etkili olduğu sorusunun cevaplanması ve gelecek yıllarda teknolojideki gelişmelerin daha farklı bir hal alacağı gerçeğinden hareketle eğitim anlayışını da etkileyebilecek bu eğitim programlarının değerlendirilmesi (Özer, 2011) gelişmiş öğrenme-öğretme süreçlerinin tasarlanması için önemlidir.

Teknolojinin bir hayli geliştiği bilgi çağında bilgiye ulaşmak, yaymak ve öğrenmek daha kolay bir hal almıştır. Bilgi çağı ile birlikte, bilginin coğrafi sınırlar tanımadan rahatça dolaşması sağlanmıştır (Demir, 2014). İçinde bulunduğumuz yüzyılda ortaya çıkan küresel iletişim ağı, dünyada yaşayan insanlara birbirleriyle yazılı, sözlü ve görüntülü iletişim kurma fırsatı verirken, eğitimcilere de küresel bir şekilde uzaktan eğitim sunma fırsatını vermektedir (İşman, 2008). Dolayısıyla eğitimi etkileyebilecek seviyede hızlı olan teknolojideki bu gelişmeler

toplumları bilginin yoğun olduğu bir yaşama sürüklemekte, eğitim programları yeniden düzenlenmekte, eğitime yapılan yatırımlar arttırılmakta, eğitim politikaları bu yönde oluşturulmakta, sanal üniversitelerin yaygınlaşacağı, eğitim ve öğretimin hızlı gelişen teknolojilerle küresel ağda uygulanacağı dile getirilmektedir (Çallı, İşman, & Torkul, 2001).

Uzaktan eğitim, hem açıklama, soru sorma, rehberlik etme gibi öğretimin etkileşimli kısmının hem de programın planlanması, hedeflerin seçilmesi veya öğretim stratejilerinin belirlenmesi gibi ileriye dönük planlama kısmının öğrenci ve öğretmenin fiziksel olarak ayrı ortamlarda bulunması sebebiyle basılı, mekanik ya da elektronik aletler vesilesiyle yapıldığı bir eğitim türüdür (Moore, 1973). Uzaktan eğitim, eğitime büyük ölçüde esneklik özelliği kazandırmakta, kendine özgü özellikleriyle öğretimin bireylerin kendi buldukları çevrede ve bireysel öğrenme yeterliklerine uygun biçimde oluşmasını (Kaya, 2012) sağlamaktadır. Uzaktan eğitim ve geleneksel eğitim arasındaki bazı farklılıklar Tablo 1'de verilmiştir.

Geleneksel eğitim ve uzaktan eğitimin karşılaştırıldığı Tablo 1 incelendiğinde temel farklılıkların fiziksel ortamdan ve zamandan bağımsızlık, ders sürelerindeki esneklik, geniş kitlelere ulaşabilme, maliyet ve tekrar olanakları gibi konularda olduğu görülmektedir.

Tarih boyunca uzaktan eğitim birkaç farklı dönem geçirerek bugünkü halini almıştır. İlk kuşak uzaktan eğitim iletişim ve öğretimin büyük ölçüde mektupla sağlandığı dönemdir. İkinci kuşak uzaktan eğitimde ise öğretim radyo ve televizyon yayınlarıyla sağlanmıştır. Üçüncü kuşak ise "açık üniversite" kavramının ortaya çıktığı yeni bir uzaktan öğretim yönteminin hüküm sürdüğü bir dönemdir. 1980'lere gelindiğinde ise uzaktan eğitim telefon, uydu veya bilgisayar ağlarıyla yapılan video ya da sesli konferans yöntemiyle yapılmaya başlanmıştır. Son olarak ise bu gün uzaktan eğitim, internet teknolojilerinden faydalanarak oluşturulmuş sanal sınıflar veya üniversiteler aracılığıyla "online" olarak gerçekleştirilmektedir (Moore & Kearsley, 2005).

1960 ve 1970'ler bilgisayar teknolojisinde büyük gelişmelerin meydana geldiği yıllardır. 1971'de Intel'in ilk mikro işlemcisi yapması, 1975'e gelindiğinde ise ilk kişisel bilgisayarın satışa sunulması bilgisayar destekli öğretimin gözle görülür bir şekilde artış meydana gelmiş ve bilgisayar teknolojisinden faydalanarak ilk uzaktan eğitim denemeleri yapılmaya başlanmıştır. İnternetin keşfi ve yaygınlaşmasıyla uzaktan eğitim için yeni bir dönemin başlangıcı olarak görülebilecek 1990'lu yıllarda bazı üniversitelerde internet temelli programlar açılmaya başlanmıştır (Moore & Kearsley, 2005). Doğuşundan bu yana beş farklı kuşak geçiren uzaktan eğitim internet teknolojisiyle birlikte bu

Tablo 1: Geleneksel Eğitim ve Uzaktan Eğitimin Karşılaştırılması

Geleneksel Eğitim	Uzaktan Eğitim
<ol style="list-style-type: none"> 1. Öğretim sürecinin tamamında öğretmen ve öğrenci aynı ortamda bulunmak zorundadır. 2. Eğitim belli bir mekânda ve belli bir zaman diliminde gerçekleşir. 3. Ders süreleri yetkililer tarafından belirlenir ve sabittir. 4. Öğretmen sayısı kısıtlıdır ve bu sebeple artan talepleri tam anlamıyla karşılamak mümkün değildir. 5. Öğretmen sayısında kısıtlılık sebebiyle yeni gelişmeler kısa zamanda büyük kitlelere aktarılamaz. 6. Eğitim kurumlarının azlığı ve mekân yetersizliği geleneksel eğitim için sorundur. 7. Bir sınıf ortamında bulunan bazı öğrenciler daha fazla tekrara ihtiyaç duyabilir. Ancak geleneksel eğitimdeki zaman sıkıntısı sebebiyle konunun tekrarlanması genellikle mümkün değildir. 8. Bazı kişiler eğitim almak istedikleri halde işi, ailesi, sorumlulukları veya yaşı yüzünden geleneksel eğitime katılmakta zorluk çekerler. 9. Birey başına düşen yatırım ve işletme giderlerine bakıldığında masraflıdır. 	<ol style="list-style-type: none"> 1. Öğretim sürecinin büyük çoğunluğunda öğretmen ve öğrenci fiziksel olarak ayrı ortamdadırlar. 2. Eğitim yer veya zamandan bağımsız olarak yapılır. 3. Ders süreleri kişiye göre değişebilir. 4. Uzaktan eğitim sayesinde daha geniş kitlelere rahat bir şekilde ulaşmak mümkündür. 5. Yeni gelişmeleri rahat bir şekilde coğrafik mesafe fark etmeksizin geniş gruplara aktarmak mümkündür. 6. Zaman ve mekân fark etmeksizin rahat bir şekilde yapılabilen uzaktan eğitim için eğitim kurumlarının azlığı ve mekân yetersizliği gibi sorunlar problem teşkil etmez. 7. Öğrenci bireysel olarak konuyu tamamıyla kavradığını hissedene kadar konuyu tekrar dinleyebilir ya da izleyebilir. 8. İş veya sorumlulukları sebebiyle yeterince vakti olmayan kişiler bile rahat bir şekilde uzaktan eğitime katılabilirler. 9. Birey başına düşen yatırım ve işletme giderlerine bakıldığında çok daha az maliyetlidir.

Holmberg (1995), Uşun (2006) ve Kaya (2012)'dan faydalanarak hazırlanmıştır.

gün kullanılan halini almış bulunmaktadır. Her bir kuşakta farklı bir yöntem ve teknik izleyen uzaktan eğitim sistemi internetin kullanımıyla birlikte de kendine özgü yeni yöntem ve teknikler geliştirmiştir.

Yukarıda değinildiği gibi, İngilizcenin artan önemi ve eğitimde uzaktan eğitim teknolojilerinin kolay erişilebilir oluşu üniversitelerde de uzaktan eğitim denemelerinin İngilizce dersleri üzerinde yapılmasını kaçınılmaz kılmıştır. Örgün ve yaygın eğitim kurumlarının temel işlevinin, hedef kitlelerine eğitim programlarında öngörülen davranış özelliklerini, uygun eğitim durumları yoluyla kazandırmak (Sağlam & Yüksel, 2007; Varış, 1988) olduğu göz önüne alındığında, eğitim kurumlarının bu işlevleri ne ölçüde yerine getirdiklerini, uyguladıkları eğitim programlarını değerlendirerek ölçmek gereklidir. Bu amaçla öğretim programları değerlendirilmektedir. Program değerlendirme, tasarlanan ve uygulanan bir eğitim programının etkililiği hakkındaki bilgilerin toplandığı, bu bilgilerin analiz edilip yorumlandığı ve sonuçta programın sürdürülmesi, geliştirilmesi ya da sonlandırılması kararının alındığı bir süreçtir (Sağlam & Yüksel, 2007).

Varış (1988)'a göre, eğitimde program geliştirme ve değerlendirme bir sürecin iç içe geçmiş öğeleridir. Program geliştirme süreci boyunca öğrenciyi biçimlendirmeye yönelik sınavlar uygulanır ve alınan sonuçlar programın geliştirilmesi için kullanılır. Program sonunda ise, programın türünü ve değerini belirlemek üzere değerlendirme yapılır. Program değerlendirme sürecinde programdaki obje, kişi ya da işlemlerin verimliliği (Posner, 2003), öğretim materyalleri, ders verimliliği ve öğretim

programının etkisi hakkında bilgi toplanabilir (Gredler, 1996). Böylece program geliştirme sürecinin son ve tamamlayıcı halkası olarak eğitim hedeflerinin ne ölçüde gerçekleştirildiği (Ertürk, 1994) belirlenir. Değerlendirme program geliştirme sürecinin, her temel aşaması sonrasında gerçekleştirilen bir çaba olarak düşünülmelidir (Olivia, 1997).

Program değerlendirme, programın tüm boyutlarının birbirinin etkileyicisi olmalarından dolayı karmaşık bir süreçtir. İyi belirlenmemiş program amaçları, amaçlara ulaşmayı sağlamayan içerik ve öğrenme yaşantıları, uygun ölçme araçlarının kullanılmaması ya da ölçütlerin iyi belirlenmemesi programın diğer boyutlarını da etkiler. Aynı şekilde program değerlendirme süreci sonucunda programın herhangi bir boyutunda yapılacak düzenlemenin programın diğer boyutlarını ve dolayısıyla programın başarısını etkileyeceği söylenebilir (Gözütok, 1999). Program değerlendirme, programın uygulanmasıyla ve sonuçlarıyla ilgili dönüt sağlayarak, programın geliştirilmesi, kabul edilmesi veya ortadan kaldırılması ya da programın öğeleri hakkında yapılabilecek değişiklikler hakkında karar vermeyi kolaylaştıran bir süreçtir. (Posavac & Carey, 2002, Langford, 2010; Demirel, 2003; Bilen, 2006).

Program değerlendirme sürecinin doğru bir şekilde işleyebilmesi için değerlendirme türlerinin, program değerlendirme yaklaşımlarının ve modellerinin bilinmesi gerekmektedir. Pek çok program değerlendirme modeli bulunmaktadır. Hangi modelin hangi program değerlendirme çalışmasında kullanılacağı, program değerlendirme modellerinin kullanılış amacına bağlı olarak karşılaştırılmalıdır. Eğer programı hazırlayan, uygu-

layan ve yöneten kişilere öğretim programının geliştirilmesi amacıyla bilgi sağlanmak isteniyorsa Stufflebeam tarafından geliştirilen “Bağlam, Girdi, Süreç ve Ürün (CIPP) değerlendirme modeli” tercih edilebilir.

Stufflebeam’e göre değerlendirme, karar seçeneklerini belirlemek amacıyla işe yarayabilecek bilgileri betimleme, elde etme ve yorumlama sürecini içerir (Oliva, 1997). Stufflebeam, değerlendirmenin idarecilerin program hakkında doğru ve yerinde kararlar almasında yardımcı olması gerektiğini savunur (Worthern et al., 1997). Değerlendirmeyi devam eden bir süreç olarak gören Ornstein ve Hunkins (2004)’e göre CIPP modeli proje, program, personel, ürün, kurum ve sistemlerin değerlendirilmelerini yönlendirmek için kapsamlı bir çerçeve programıdır. Bu modelin çekirdek kavramları bağlam, girdi, süreç ve üründür (Stufflebeam, 2000). Modeli oluşturan bu kavramlar ayrı ayrı değerlendirme boyutu olarak ele alınabileceği gibi bu kavramların birbiri ile ilişkisi ve bağı aracılığıyla toplam değerlendirme de yapılabilir (Stufflebeam, 2003).

CIPP modeline göre program değerlendirmede dört ana unsur göz önüne alınmaktadır (Brown, 1994):

- Değerlendirme karar verme amacıyla yapılır. Çünkü değerlendirme karar vericilere bilgi sağlamalıdır.
- Değerlendirme döngüsel ve sürekli olan bir süreçtir ve bundan dolayı da sistematik bir program aracılığı ile uygulanmak zorundadır.
- Değerlendirme süreci planlama, bilgi alma ve sağlama şeklinde üç ana aşamayı içerir. Bu aşamalar değerlendirmenin yöntemi için temel oluşturur.
- Değerlendirme sürecindeki planlama ve bilgi sağlama aşamaları işbirliği gerektiren birbirine dönük aşamalardır.

Çok kapsamlı olan bu modelde planlama, yapılandırma, uygulama ve yeniden düzenlemeyle ilgili olarak bağlam, girdi, süreç ve ürüne ilişkin dört alanda karar vermek gerekmektedir (Ornstein & Hunkins, 2004). Bu dört alan hakkında karar vermek için, “ne yapılması gerekiyor?” “nasıl yapılmalı?” “planlandığı gibi yapılıyor mu?” ve “başarılı oldu mu?” sorularının cevaplanması gerekir (Stufflebeam, 2007).

Bağlam değerlendirme boyutu, hedeflerin, öğrenen ihtiyaçlarının ve çözülmesi gereken problemin gözden geçirildiği ilk basamaktır (Gilchrist & Roberts, 1974). Bu aşama, karar vericilere yardım etmek için ihtiyaçların belirlenmesi, hedef ve önceliklerin tanımlanması ve geniş kullanıcı kitlelerine yardımcı olmak için hedefler, öncelikler ve sonuçlar hakkında karar verilmesini içerir (Stufflebeam, 2000). Bağlamın değerlendirilmesinin nedeni hedeflerin belirlenmesi için ölçütler belirlemektir. Bağlamın değerlendirilmesine bir anlamda durum analizi de denilebilir. Bağlamın değerlendirilmesi esnasında öğretimin gerçekleşeceği ortam betimlenir ve bu ortam hakkında arzu edilen şartlar belirlenir. Sorulan bazı sorular çerçevesinde bağlamın değerlendirilmesi yapılabilir:

- Programın amaçları öğrencilerin, toplumun ve iş yaşamının gereksinimlerini karşılamakta mıdır?

- Program iş yaşamıyla iş birliğini içermekte midir?
- Programın amaçları öğrencilerin gelecekteki öğrenmelerine hazırlık sağlamakta mıdır?
- Programın amaçları güncel midir? (Reece & Walker, 1993; Akt: Gözütok, 1999)

Girdi değerlendirme boyutunda ise, amaçlara ulaşmak ve hedeflenenleri karşılamak için çeşitli yaklaşımları, eylem planlarını ve bunların yapılabirlik düzeyleri ile bütçeleri belirlenir ve değerlendirilir (Stufflebeam, 2000). Öğrencilerin programla ilgili hazırbulunuşluk düzeyleri, öğretmenlerin bu programı uygulamaya ilişkin yeterlikleri, kaynakların yeterliği, öğretim materyallerinin yeterliği bu aşamada ortaya konması gereken verilerdir (Gözütok, 1999). Girdi değerlendirmede özellikle şu sorulara cevap aranır:

- Hedefler doğru ve uygun bir şekilde hazırlanmış mıdır?
- Hedefler kurumun ya da okulun genel hedefleri ile uyumlu mudur?
- Seçilen öğretim yöntemleri hedeflere uygun mudur?
- Program içinde belirlenen diğer stratejiler programın hedeflerine uygun mudur? (Ornstein & Hunkins, 2004)

Süreç değerlendirme boyutu ise çalışanların etkinliklerine yardımcı olacak ve işlerini kolaylaştıracak olan planların uygulanmasını değerlendirerek programın performansı hakkında karar verilmesini ve sonuçların yorumlanmasını içerir (Stufflebeam, 2000). Yani planlanan etkinliklerin uygulamaya konması ile başlanan aşamadır. Büyük çaplı programlar uygulamaya konmadan önce pilot uygulamalarda süreç değerlendirme yapılabilir. “Program uygulanmaya başlamadan önce hazırlanan plan ne kadar iyi uygulanıyor?”, “uygulama esnasında hangi problemlerle karşılaşılıyor?”, “başarıya engel olabilecek durumlar nelerdir?” ve “hangi düzenlemelere ihtiyaç duyulmaktadır?” gibi sorulara yanıt bulunduğu, programın etkililiği hakkında bir fikir sahibi olunabilir ve programa müdahale edilebilir (Worthern et al., 1997). Ayrıca süreç değerlendirme aşamasında programı uygulamanın maliyeti ve katılımcıların programın uygulanışı hakkındaki görüşleri de belirlenir (Stufflebeam & Shinkfield, 2007). Sürecin değerlendirilmesinde de bazı sorulara yanıt aramak gerekir:

- Öğretmenler program dokümanlarını nasıl kullanıyorlar?
- Hazırlanan değerlendirme dokümanları öğretmenler tarafından anlaşılabilir mi?
- Karşılaşılan problemlerin çözümü için öğretmenlere ve öğrencilere yardım sağlanmakta mıdır?
- Önerilen öğretim stratejileri kullanılmakta mıdır?
- Belirlenen içerik tam olarak verilmekte midir? (Gözütok, 1999)

Yukarıda bahsedilen sorulara yanıt aramada gözlem formları, anketler, başarı testleri, performans testleri, öğretmen davranışlarını ölçmeye yarayan envanterler ve uzman görüşlerinden yararlanılabilir (Gözütok, 1999).

Ürün değerlendirme, programın belirlenen amaçlarına ulaşip ulaşmadığı ve program sonunda elde edilen ürünün ne ölçüde başarılı olduğu konusunda karar vermeye yardımcı olur (Gilchrist & Roberts, 1974; Ornstein & Hunkins, 2004). Diğer bir deyişle program ile belirlenen hedeflere ne ölçüde ulaşıldı sorusuna cevap aranır. Bu değerlendirme sonunda programın değiştirilip düzeltilmeye ihtiyaç olup olmadığına ilişkin kararlar alınır. Ürünün değerlendirilmesinde de bazı sorulara yanıt aramak gerekir:

- Başarı yüzdesi nedir?
- Hedeflere ne ölçüde ulaşılabildi?
- Programın başında belirlenen ihtiyaçlar hangi oranda karşılanabildi?
- Mezunlar iş bulabilmekte midir?
- Mezunların programa ilişkin görüşleri nelerdir?
- İş yaşamının mezunlar ve program hakkındaki görüşleri nelerdir? (Gözütok, 1999; Worthen et al., 1997)

Yukarıda bahsedilen sorulara yanıt aramada performans testleri, gözlem formları, anketler, başarı testleri, iş doyumu ölçekleri kullanılabilir. Ürünün değerlendirilmesi aşamasından elde edilen veriler ışığında programın sürdürülmesi, değiştirilmesi ya da kaldırılması konusunda karara varılabilir (Gözütok, 1999).

Görüldüğü gibi bu model, değerlendirmenin en önemli amacının kanıtlamak değil, geliştirmek olduğu görüşü üzerine temellendirilmiş ve çeşitli alanlardan eğitimcilerin, uzmanların ve ilgililerin kullanımı için geliştirilmiştir (Stufflebeam, 2000). Bu nedenle bu çalışmada Bülent Ecevit Üniversitesi'nde uzaktan eğitimle gerçekleştirilen İngilizce öğretimini CIPP modeli ile değerlendirmede kullanılacak bir ölçek geliştirmek amaçlanmıştır. Değerlendirmenin uzaktan eğitimle yürütülen İngilizce dersi öğretim programının hedef kitlesi olan üniversite öğrencilerinin görüşlerine dayalı olarak yapılabilmesi amacıyla ölçek öğrenciler için geliştirilmiştir.

YÖNTEM

Çalışma Grubu

Taslak ölçme aracının geçerlik ve güvenilirlik çalışmalarının yapılabilmesi için pilot uygulama aşamasında kaç kişilik bir gruba

uygulanması gerektiği konusunda farklı görüşler mevcuttur. Gorsuch (1974), Floyd ve Widaman (1995) ve Anderson, (1988) gibi araştırmacılar taslak ölçme aracında bulunan madde sayısına bağlı olarak gözlem sayısına karar verilmesi gerektiğini belirtmişlerdir. Gorsuch (1974) taslak ölçme aracının, madde sayısının 5 ya da 10 katı kadar kişiye uygulanması gerektiğini belirtirken, Floyd ve Widaman (1995) ve Anderson (1988) ise 5 ya da 4 katının yeterli olduğunu söylemektedir. Diğer taraftan Comrey ve Lee'ye (1973) göre geçerlik çalışmasının yapılabilmesi için pilot uygulamanın en azından 250 kişiyle yapılması gerekmektedir. Tabachnick ve Fidell (2012) ise 150 deneyin yeterli olduğunu söylemektedir. Kline (2010) ise gözlem sayısının 100-200 arasında olması gerektiğini söylemektedir. Bu görüşlerden hareketle hazırlanan taslak ölçme aracı İngilizce dersi öğretim programının uzaktan eğitimle yürütüldüğü Bülent Ecevit Üniversitesi ve İlahiyat ve Türk Dili ve Edebiyatı bölümlerinde öğrenim görmekte olan 250 öğrenciye uygulanmıştır. Ölçekler arasından titizlikle cevaplanmayan ve eksikleri bulunan ölçekler elendikten sonra geriye kalan 163 ölçek analize dâhil edilmiştir.

Ölçeğin pilot çalışmasında görüşleri analize dâhil edilen öğrencilerin %35'i Edebiyat Fakültesi'nde %65'i ise İlahiyat Fakültesi'nde öğrenim görmektedir. 113 öğrenci kadın (%69.3) 50 öğrenci ise erkektir (%30.7). 84 öğrenci örgün öğretimde (%51.5) eğitim alırken 79 öğrenci ikinci öğretimde (%48.5) eğitim almaktadır. Ayrıca öğrencilerin büyük çoğunluğu 18 ve 21 yaşları arasındadır (%90.1).

Süreç

İngilizce Dersi Öğretim Programını Değerlendirme Ölçeği (İÖP-DÖ) hazırlanırken öncelikle CIPP modeli ve ölçek geliştirme hakkında ilgili literatür ayrıntılı bir şekilde incelenmiştir. İnceleme sonunda ölçek geliştirme aşamalarını farklı başlıklar altında sınıflayan araştırmacıların olduğu görülmüştür. Bu sınıflandırmalar Tablo 2'de görülmektedir.

Araştırmacılar ölçek geliştirme aşamalarına farklı isimler vermiş olsalar da araştırmacıların görüşleri birbiriyle paralellik göstermektedir. Bu çalışmada literatür ile uyumlu olarak öncelikle madde havuzu oluşturulmuş, ardından taslak ölçme aracı hazırlanmış, taslak ölçme aracına ilişkin uzman görüşü alınmış, pilot uygulama yapılmış, sonrasında da geçerlik ve güvenilirlik çalışmaları yapılarak ölçeğe son hali verilmiştir.

Tablo 2: Literatürde Bulunan Ölçek Geliştirme Aşamaları

McMillan ve Schumacher (2006), Baş (2013)	Tavşancıl (2005), Dunn Rankin (2004), DeVellis (2003)	Karasar (2011), Balcı (2015)
1. Ölçeğin kuramsal çerçevesinin belirlenmesi	1. Madde havuzunun oluşturulması	1. Madde havuzu aşaması
2. Ölçekte kullanılacak maddeleri içeren madde havuzunun oluşturulması	2. Uzman görüşünün alınması	2. Uzman görüşü aşaması
3. Uzman görüşüne başvurulması	3. Ön deneme uygulamasının yapılması	3. Ön deneme aşaması
4. Ön uygulamanın yapılması ve ölçeğin düzenlenmesi	4. Ölçek taslağının çalışma grubuna uygulanması ve faktör analizi	4. Faktör analizi aşaması
5. Geçerlik ve güvenilirlik analizlerinin yapılması	5. Ölçeğin güvenilirliğinin hesaplanması	5. Güvenirlik hesaplama aşaması
6. Ölçeğe son halinin verilmesi		

Madde Havuzunun Oluşturulması

CIPP değerlendirme modelinin prensiplerinden hareketle hazırlanan ölçeğe ait madde havuzu oluşturulmadan önce uzaktan eğitim ve CIPP değerlendirme modeli üzerine kapsamlı bir literatür taraması yapılmıştır. Ayrıca uzaktan eğitim programlarını değerlendiren çalışmaların (Cansu, 2010; Özer, 2011; Çakır & Yurtsever, 2012; Ekmekçi, 2015; Yarmohammadian ve ark., 2011; Tokmak ve ark., 2013) ve CIPP değerlendirme modelini kullanmış farklı program değerlendirme çalışmalarının ölçme araçları, detaylı bir şekilde incelenmiştir (Karataş, 2007; Tunç, 2010; Virlan, 2014; Ünal, 2011; Manasneh, 2013; Burke Sinclair, 2012; Hanchell, 2014). Sonrasında uzaktan eğitimle yürütülen Yabancı Dil I dersinin bağlam, girdi, süreç ve ürün boyutlarının değerlendirilmesi amacıyla her boyutla ilgili çok sayıda madde hazırlanmıştır. Madde havuzu oluşturulurken, maddelerin yalın ve anlaşılır bir dille yazılmasına, maddeler arasında olgusal ifadelerin olmamasına, bir maddenin tek bir yargı ya da düşüncüyü ölçmeyi hedeflemesine, olumlu maddeler kadar olumsuz maddelerin de havuzda bulunmasına (Tavşancıl, 2005) dikkat edilmiştir.

Taslak Ölçme Aracının Hazırlanması

Yukarıda belirtilen kapsamda 98 madde oluşturulmuş ve maddelerin her biri için 5'li Likert formunda ("tamamen katılıyorum", "katılıyorum", "ne katılıyorum ne katılmıyorum", "katılmıyorum", "hiç katılmıyorum") cevap seçenekleri eklenmiştir. Dolayısıyla katılımcıların kişisel bilgilerini toplamak amacıyla hazırlanmış ilk bölüm ve CIPP modelinin dört boyutuyla ilgili verilerin toplanması amacıyla hazırlanan 98 maddelik ikinci bölümden oluşan taslak ölçme aracı oluşturulmuştur.

Uzman Görüşüne Başvurulması

Taslak ölçme aracının kapsam ve görünüş geçerliliğini belirlemek için uzman görüşü alınmıştır. Kapsam geçerliliği, testi oluşturan maddelerin ölçülmek istenilen tanımlanmış davranışlar bütününe ölçmede ne derece temsil ettiğine ilişkindir (Büyüköztürk et al., 2014). Başka bir deyişle bir ölçme aracında bulunan maddelerin beklenen davranışları ne ölçüde ölçtüğünün belirlenmesi işidir (Balci, 2015). Bu ölçeğin kapsam geçerliliğini incelemek için, kullanılabilir mantıksal yollardan biri olan uzman görüşüne başvurulmuştur (Büyüköztürk et al., 2014). Taslak ölçme aracı, program geliştirme alanında uzman üç öğretim üyesine ulaştırılmıştır. Öğretim üyelerinden maddelerin ölçekte kalması veya değiştirilerek kalması ya da ölçekten çıkarılması ve CIPP modelinin hangi boyutuna ait olduğu konusunda görüş bildirmeleri istenmiştir. Uzmanlardan gelen dönütler sonrasında taslak ölçekte bulunan bazı maddelerde değişiklikler yapılmıştır.

Geçerlik ve Güvenirlik Çalışmaları

Bir ölçme aracının soyut bir olguyu ne derece doğru ölçebildiğini gösteren (Tavşancıl, 2005) yapı geçerliliğinin incelenmesinde en sık kullanılan yöntem faktör analizidir (Büyüköztürk, 2014). Dolayısıyla kapsam ve görünüş geçerliliği uzman görüşüyle sağlanan ölçme aracının yapı geçerliliğini kontrol etmek amacıyla faktör analizi yapılmıştır. Faktör analizi birbiriyle ilişkili çok sayıda değişkeni bir araya getirerek, kavramsal olarak anlamlı

daha az sayıda yeni değişkenler bulmayı amaçlayan bir istatistik olarak tanımlanabilir (Büyüköztürk, 2002). Ölçme aracının yapı geçerliliği dört aşamada incelenmiştir. İlk aşamada verilerin faktör analizine uygun olup olmadığına bakılmıştır. İkinci aşamada ölçeğin faktörleri ortaya çıkartılmıştır. Faktörlerin belirlenmesinin ardından üçüncü aşama olarak hangi maddenin hangi faktörde yer alacağına karar vermek için faktör yük değerlerine ve maddelerin madde-toplam korelasyonlarına bakılmıştır. Geçerlik analizinin son aşaması olarak ortaya çıkan boyutlara isimler verilmiştir.

En temel anlamıyla ölçme sonuçlarının kararlılık derecesi olarak veya ölçme sonuçlarının hatalardan arınık olma derecesi olarak tanımlanan (Seçer, 2013) güvenilirlik bir ölçme aracının en önemli teknik özelliklerinden biridir (Tekin, 2000). Başka bir deyişle güvenilirlik bir ölçme aracının ölçmek istediği özelliği ne derecede bir kararlılıkla ölçmekte olduğunu göstergesidir (Tavşancıl, 2005). Likert tipi bir tutum ölçeğinde güvenilirlik düzeyini saptamak için iç tutarlılığın bir ölçütü olan Cronbach alfa katsayısı (Tavşancıl, 2005) ölçek geliştirme ve uyarlama çalışmalarında güvenilirliğin belirlenmesinde sıklıkla başvurulan bir yöntemdir (Seçer, 2013). Bu nedenle ölçeğin güvenilirlik çalışması yapılırken Cronbach alfa katsayısı ve madde toplam korelasyonlarına bakılmıştır.

BULGULAR

Geçerlik Çalışmasına İlişkin Bulgular

Verilerin faktör analizine uygun olduğunu söyleyebilmek için literatürde Barlett's testinin anlamlı olması gerektiği belirtilirken, KMO testinin ideal değerinin kaç olması gerektiği konusunda farklı görüşler mevcuttur. Tavşancıl (2005) ve Sharma (1996) KMO değerinin 1'e yaklaştıkça mükemmel, 0.60 değerinin ise vasat olmakla birlikte kabul edilebilir olduğunu belirtirken, Kalaycı (2008) ise 0.50 değerinin üzerinin yeterli olduğunu söylemektedir. Büyüköztürk (2002), Tabachnick ve Fidel (2012) ile Kaiser (1974) ise bu değer en az 0.60 ve üzerinde olması gerektiğini söylemektedir. Alpar (2000) ise daha yüksek bir değer biçerek KMO değerinin 0.80'nin üzerinde olması gerektiğini belirtmektedir. Bryman ve Cramer (2003) ise 0.60 değerini orta bir değer olarak sınıflarken 0.90 değerini mükemmel olarak kabul etmektedir. Ölçeğe uygulanan KMO testi sonucunda KMO değeri 0.905 olarak bulunmuştur. Dolayısıyla örneklem büyüklüğünün faktör analizi için mükemmel bir değer taşıdığı söylenebilir. Ayrıca faktör analizi için başka bir ön koşul olan örneklemin normal dağılıma zorunluluğunu kontrol etmek amacıyla Barlett's testi yapılmış ve dağılımın normal olduğu görülmüştür (p: .000). Bu sonuçlara göre veriler faktör analizi için uygundur.

Bir maddenin ölçekte yer alması için faktör yük değeri Tabachnick ve Fidel'e (2012) göre en az 0.45 olmalıdır. Büyüköztürk (2014) de yine aynı şekilde faktör yük değeri için alt sınırı 0.45 belirlerken, Kalaycı (2008) bu değer en az 0.50 olması gerektiğini söylemektedir. Tekin (2000) ise bu değer 0.35'e kadar, Seçer (2013) ise 0.30'a kadar inebileceğini belirtmektedir. Ayrıca maddelerin sadece tek bir faktörde yüksek yük değerine sahip olması gerekmektedir (Büyüköztürk, 1997). Farklı faktörlerdeki yüksek iki yük değeri arasındaki farkın en az 0.10 olması

(Büyüköztürk, 2014) gerekmede ve aradaki farkın bu değerden az olması durumunda bu maddeler binşik madde olarak adlandırılmakta (Yavuz, 2005) ve ölçekten çıkarılması gerekmektedir. Ölçeğin geçerlik çalışması esnasında bir maddenin bir faktörde yer alıp almaması konusunda karar verilirken bu değerlerden faydalanılmıştır.

98 maddelik taslak ölçme aracına faktör analizi yapılmasının ardından ölçeğin toplam varyansın yüzde 75.40'ını açıklayan ve öz değeri 1'den büyük 20 boyutta toplandığı görülmüştür. Ancak bu boyutlara anlamlı bir şekilde isim verilememiştir. Ayrıca ölçek CIPP değerlendirme modelinin prensiplerine göre hazırlandığından maddelerin dört boyutta toplanması düşünülmüştür. Dolayısıyla faktör yükleri 0.45'in altında olan veya binşik madde durumunda olan ya da yeni boyutlar oluşmasına sebebiyet veren 62 sorunlu madde ölçekten çıkartılmıştır. Sorunlu maddeler çıkartıldıkça faktör analizi tekrar edilmiş ve bu işlem ölçek dört boyutlu halini alana kadar yinelenmiştir. Ölçeğe dört boyutlu haliyle bakıldığında ölçek toplam varyansın %62.69'unu açıklamaktadır. Birinci boyutun öz değeri 16.20, açıkladığı varyans %45.05; ikinci boyutun öz değeri 2.58, açıkladığı varyans %7.18; üçüncü boyutun öz değeri 2.17, açıkladığı varyans %6.04 ve dördüncü boyutun öz değeri 1.60 ve açıkladığı varyans %4.45'dir.

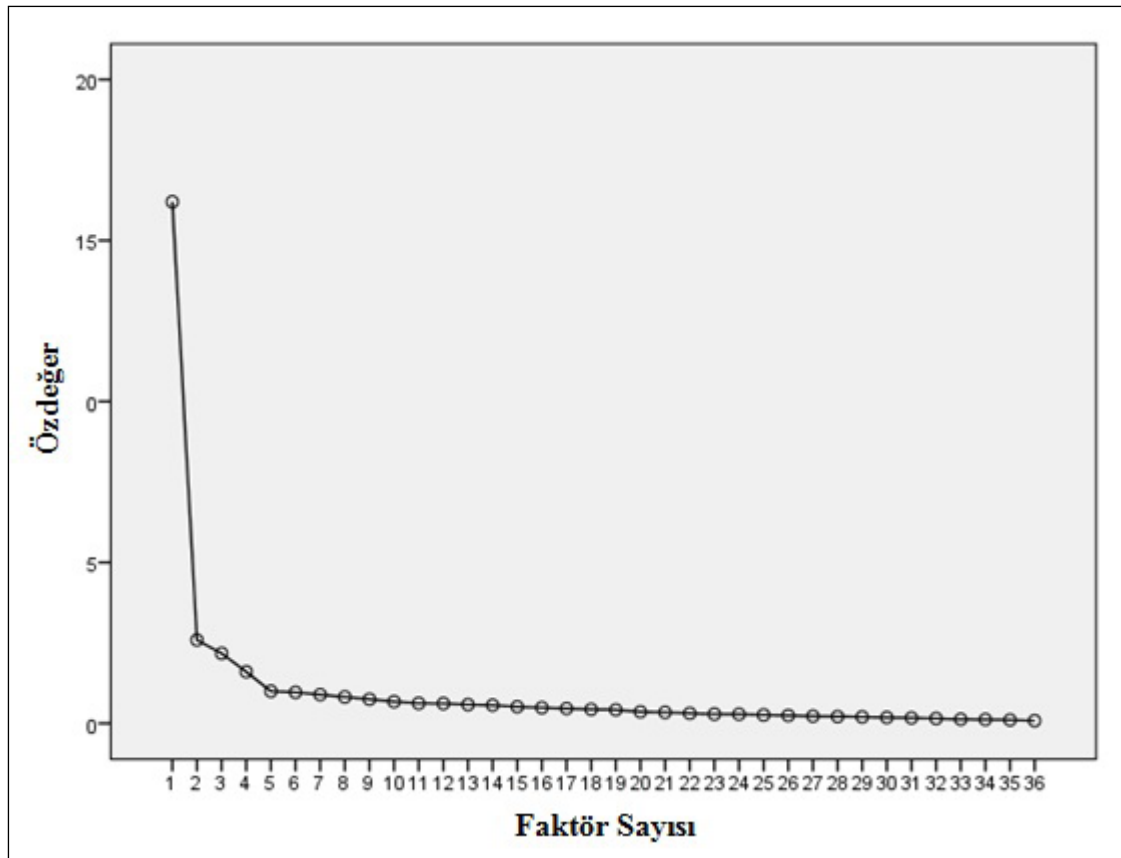
Bryman ve Cramer (2003) öz değeri 1 veya 1'den büyük olan faktörlerin önemli faktör olarak nitelendirilmesi gerektiğini belirtmektedir. Aynı şekilde Seçer (2013) de her bir alt boyuta ait öz değerin en az 1 olması gerektiğini belirtmektedir. Dola-

yısıyla dört faktörün öz değerleri ve açıkladıkları varyanslar literatürde belirtilen değerlere uygundur. Ayrıca faktör sayısına daha sağlıklı karar verebilmek amacıyla faktörlerin öz değerlerine dayanan yamaç-yığıntı grafiğine de bakmak gerekmektedir (Büyüköztürk, 2002). Bu amaçla incelenen yamaç-yığıntı grafiği Şekil 1'de görülmektedir.

Yamaç-yığıntı grafiğindeki yüksek ivmeli, hızlı düşüşler önemli faktör sayısını verirken yatay çizgiler ise varyansı açıklama katsayısının birbirine çok yakın olduğunu göstermektedir (Büyüköztürk, 2014). Dolayısıyla grafik incelendiğinde dördüncü boyuttan sonra çizgilerin yatay bir hal alması ölçeğin dört boyutta kullanılabilceğini desteklemektedir.

Maddelerin hangi faktör altında kullanılacağına karar verirken maddelerin o faktör altında en az 0.45 yük değerine sahip olmasına ve birden fazla faktör altında değere sahip olan maddeler arasında en az 0.10 değerinde fark olmasına dikkat edilmiştir. Dolayısıyla bu koşullara uymayan 62 madde ölçekten atılmıştır. Ölçekten atılan 62 maddenin arasında ters işleyen olumsuz maddeler de bulunmaktadır. Kalan maddeler Varimax yöntemiyle döndürüldükten sonra maddelerin faktörleri belirlenmiştir.

Varimax döndürme işlemi sonrasında oluşturulan yukarıdaki tablo incelendiğinde ölçeğin açıkladığı toplam varyansın %62.69 olduğu görülmüştür. Stevens (1996) bir ölçme aracında açıklanan varyans oranının en az %75 ve üzeri olması gerektiğini, Henson ve Roberts (2006) ise %52 ve üzerinde bir değerin ölçek



Şekil 1: İngilizce dersi öğretim programını değerlendirme ölçeğine ait yamaç yığıntı grafiği.

çalışmalarında yeterli olduğunu belirtmektedir (Akt: Seçer, 2013). Dolayısıyla ölçeğin açıkladığı toplam varyans yeterlidir. Bu varyansın %21.213'ünü birinci faktör, %16.200'ünü ikinci faktör, %14.326'sını üçüncü faktör, %10.951'ini ise dördüncü faktör oluşturmaktadır. Ayrıca birinci faktörde maddelerin

faktör yükleri 0.49 ve 0.79 arasında, ikinci faktördeki maddelerin faktör yükleri 0.58 ve 0.79 arasında, üçüncü faktördeki maddelerin faktör yükleri, 0.49 ve 0.82 arasında, dördüncü faktördeki maddelerin faktör yükleri ise 0.60 ve 0.78 arasında değişmektedir. Dolayısıyla maddeler literatürde belirtilen

Tablo 3: İÖPDÖ Maddelerinin Döndürme Sonrası Faktör Yükleri, Faktörlerin Açıkladıkları Varyanslar ve Madde-Toplam Korelasyonları

Madde	Ölçek Faktörleri				Madde-Toplam Korelasyonları
	1. Faktör	2. Faktör	3. Faktör	4. Faktör	
M92	.773				.707
M96	.768				.713
M97	.754				.657
M94	.737				.731
M98	.734				.640
M90	.725				.711
M95	.721				.758
M91	.708				.713
M93	.690				.767
M89	.657				.706
M88	.646				.648
M86	.527				.673
M83	.490				.661
M67		.796			.634
M68		.767			.708
M66		.762			.684
M64		.669			.504
M77		.668			.603
M70		.619			.589
M69		.617			.637
M78		.606			.676
M74		.584			.604
M26			.823		.612
M25			.806		.635
M24			.792		.579
M27			.710		.616
M23			.658		.643
M28			.609		.624
M22			.534		.607
M30			.493		.568
M2				.787	.619
M1				.700	.628
M4				.698	.598
M5				.688	.658
M3				.615	.564
M8				.609	.522
Toplam Varyans	%21,21	%16,20	%14,32	%10,95	
	Toplam Varyans: %62,69				

en az 0.45 yük değerine sahip olma koşulunu sağlamaktadır. Analizler sonucunda nihai ölçekte maddelerin hangi faktörler altında toplandığı ve bu faktörlere verilen isimler Tablo 4’de görülmektedir.

Tablo 4 incelendiğinde birinci faktörde 13 madde, ikinci faktörde 9 madde, üçüncü faktörde 8 madde ve dördüncü faktörde 6 madde olduğu görülebilir. Maddeler ölçek içinde incelenmiş ve CIPP değerlendirme modelinin değerlendirme boyutları göz önüne alınarak bu faktörlere isim vermeye çalışılmıştır. Dolayısıyla 13 maddelik ilk faktöre “ürün değerlendirme”, 9 maddelik ikinci faktöre “süreç değerlendirme”, 8 maddelik üçüncü faktöre “girdi değerlendirme” ve 6 maddelik dördüncü faktöre “bağlam değerlendirme” ismi verilmiştir.

Tablo 4: Maddelerin Faktörlere Göre Dağılımı ve Faktör İsimleri

Faktörler	CIPP Boyutu	Maddeler
1. Faktör	Ürün Değerlendirme	83, 86, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98
2. Faktör	Süreç Değerlendirme	64, 66, 67, 68, 69, 70, 74, 77, 78
3. Faktör	Girdi Değerlendirme	22, 23, 24, 25, 26, 27, 28, 30
4. Faktör	Bağlam Değerlendirme	1, 2, 3, 4, 5, 8

Tablo 5: Ölçek Alt Boyutlarının Güvenirlik Katsayıları ve Madde Toplam Korelasyonları

Boyutlar	Madde No	α	r
Bağlam	1, 2, 3, 4, 5, 8	0.87	0.52 - 0.65
Girdi	22, 23, 24, 25, 26, 27, 28, 30	0.90	0.56 - 0.64
Süreç	64, 66, 67, 68, 69, 70, 74, 77, 78	0.91	0.50 - 0.70
Ürün	83, 86, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98	0.94	0.64 - 0.76
Toplam		0.96	0.50 - 0.76

Tablo 6: Ölçekte Bulunan Maddelerin Ortalamaları, Standart Sapma Değerleri ve Madde-Toplam Korelasyon Değerleri

Madde	\bar{X}	ss	r	Madde	\bar{X}	ss	r
1	2.77	1.23	0.62	69	3.04	1.12	0.63
2	2.74	1.18	0.61	70	3.04	1.17	0.58
3	3.07	1.27	0.56	74	3.19	1.19	0.60
4	2.94	1.26	0.59	77	3.00	1.12	0.60
5	2.51	1.24	0.65	78	3.17	1.18	0.67
8	2.53	1.35	0.52	83	2.81	1.25	0.66
22	2.93	1.07	0.60	86	2.79	1.20	0.67
23	2.79	1.23	0.64	88	2.77	1.19	0.64
24	2.55	1.21	0.57	89	2.85	1.25	0.70
25	2.58	1.15	0.63	90	2.86	1.27	0.71
26	2.69	1.10	0.61	91	2.90	1.32	0.71
27	2.77	1.14	0.61	92	2.79	1.20	0.70
28	2.89	1.18	0.62	93	2.76	1.16	0.76
30	3.04	1.19	0.56	94	2.63	1.25	0.73
64	2.85	1.37	0.50	95	2.78	1.23	0.75
66	3.04	1.23	0.68	96	2.73	1.28	0.71
67	3.04	1.16	0.63	97	2.73	1.24	0.65
68	3.06	1.19	0.70	98	2.78	1.35	0.64

Güvenirlik Çalışmasına İlişkin Bulgular

Bir ölçeğin güvenilir kabul edilebilmesi için Cronbach alfa katsayısının en az 0.70 ve üzeri olması gerekmektedir (Seçer, 2013; Büyüköztürk, 2014; Şencan, 2005). Bu nedenle ölçeğin Cronbach alfa katsayısı ve madde toplam korelasyonlarına bakılmıştır. Ayrıca ölçekte yer alan her bir boyutun ayrı ayrı güvenilirlik katsayıları ve madde toplam korelasyonları Tablo 5’de görülmektedir.

Tablo 5’den de görülebileceği gibi ölçeği oluşturan alt boyutların güvenilirlik katsayıları 0.87 ile 0.94 arasında değişmektedir. Ayrıca bütün ölçeğin güvenilirlik katsayısı da 0.96 bulunmuştur. Ölçekte bulunan maddelerin madde toplam korelasyonları ise 0.50 ve 0.76 arasında değişmektedir. Dolayısıyla literatürde

belirtilen değerlere göre ölçeğin güvenirlik katsayısı yüksek düzeydedir.

Ayrıca her bir maddenin madde-toplam korelasyon değerleri, ortalaması ve standart sapma değerleri hesaplanmış ve Tablo 6'da verilmiştir.

Büyüköztürk'e (2014) göre madde toplam korelasyonu 0.30'dan yüksek maddeler bireyleri iyi derecede ayırt edebilmektedir. Maddelerin madde-toplam korelasyonu değerlerinin 0.50 ile 0.76 arasında değiştiği görülmüştür. Dolayısıyla madde toplam korelasyonun yüksek düzeyde olduğu ve maddelerin ölçek tam puanları ile yüksek düzeyde ve pozitif yönde anlamlı bir ilişki içinde olduğu söylenebilir. Ayrıca maddelerin ayırt ediciliğinin de yüksek olduğu söylenebilir.

TARTIŞMA ve SONUÇ

Bu araştırmada, uzaktan eğitimle yürütülen İngilizce dersi öğretim programının değerlendirilmesinde kullanılmak amacıyla geçerli ve güvenilir bir ölçme aracı geliştirilmesi amaçlanmıştır. İÖPDÖ, Stufflebeam'e ait CIPP değerlendirme modelinin prensiplerinden hareketle hazırlanmıştır. Bu sebeple ölçeğin bağlam değerlendirme, girdi değerlendirme, süreç değerlendirme ve ürün değerlendirme boyutları olmak üzere dört boyutlu bir yapıda olması planlanmıştır. Ölçek beşli Likert ("tamamen katılıyorum", "katılıyorum", "ne katılıyorum ne katılmıyorum", "katılmıyorum", "hiç katılmıyorum") türünde hazırlanmış ve ölçekte en olumsuz ifadeye 1 puan ve en olumlu ifadeye 5 puan verilmiştir.

Kapsam ve görünüş geçerliliği uzman görüşüyle sağlanan ölçme aracının yapı geçerliliğini kontrol etmek amacıyla faktör analizi yapılmıştır. Verilerin faktör analizi için uygun olup olmadığını belirlemek için KMO ve Barlett's testlerine bakılmıştır. Ölçeğe uygulanan KMO testi sonucunda KMO değeri 0.90 olarak bulunmuş ve yapılan Barlett's testi sonucunda da dağılımın normal olduğu görülmüştür (p: .000). Bu sonuçlara göre veriler faktör analizi için uygundur. 98 maddelik taslak ölçme aracına faktör analizi yapılmasının ardından ölçeğin toplam varyansın yüzde 75.40'ını açıklayan ve öz değeri 1'den büyük 20 boyutta toplandığı görülmüştür. Ancak bu boyutlara anlamlı bir şekilde isim verilememiştir. Dolayısıyla ölçeği dört boyuta indirmek için sorunlu maddeleri ölçekten çıkarma yoluna gidilmiştir. Sorunlu maddeler çıkartıldıkça faktör analizi tekrar edilmiş ve bu işlem ölçek dört boyutlu halini alana kadar yinelenmiştir. Ölçeğe dört boyutlu haliyle bakıldığında ölçek toplam varyansın %62.69'unu açıklamaktadır. Birinci boyutun öz değeri 16.20, açıkladığı varyans %45.05; ikinci boyutun öz değeri 2.58, açıkladığı varyans %7.18; üçüncü boyutun öz değeri 2.17, açıkladığı varyans %6.04 ve dördüncü boyutun öz değeri 1.60 ve açıkladığı varyans %4.45'dir. Ölçeğin dört boyutlu halini almasının ardından, maddelerin hangi boyutlarda yer alacağına ilişkin işlemler yapılmış ve boyutlara isim verilerek faktör analizi sonlandırılmıştır.

Ölçeğin güvenirlik düzeyini saptamak için ölçekte yer alan her bir boyutun ayrı ayrı güvenirlik katsayıları ve madde toplam korelasyonlarına bakılmıştır. Yapılan analizler sonucunda, ölçe-

ği oluşturan alt boyutların güvenirlik katsayıları 0.87 ile 0.94 arasında değişmektedir. Ayrıca bütün ölçeğin güvenirlik katsayısı da 0.96 bulunmuştur. Ölçekte bulunan maddelerin madde toplam korelasyonları ise 0.50 ve 0.76 arasında değişmektedir. Dolayısıyla yapılan geçerlik ve güvenirlik analizleri sonucunda varyansın %62.69'unu açıklayan ve güvenirlik katsayısı 0.96 olan dört boyutlu 36 maddelik bir ölçek elde edilmiştir.

Özellikle teknolojik gelişmelerle birlikte önemi artarak geleneksel eğitime alternatif oluşturmaya başlayan uzaktan eğitime ait programların değerlendirilmesi önem arz etmektedir. Ülkelerinde sundukları eğitimin kalitesini, öğrenme ve öğretimin niteliğini artırmak amacıyla bilgi ve iletişim teknolojilerini daha iyi kullanabilmede ülkelerin birbiriyle yarış halinde olduğunu söyleyen Yarmohammadian ve ark., (2011)'na göre uzaktan eğitimin etkililiğini değerlendiren çalışmalar bu süreçte oldukça gereklidir.

İlgili literatür incelendiğinde bir hayli yaygınlaşmış olan uzaktan eğitim programlarını değerlendiren çalışmaların sayıca azlığı görülmektedir. Uzaktan eğitimle yürütülen yabancı dil derslerine ait programların değerlendirildiği çalışmaların sayısı ise çok daha azdır. Literatürde uzaktan eğitim programlarını farklı program öğeleri açısından ve daha geniş örneklerle değerlendirilen çalışmaların yapılmasını öneren araştırmalar olduğu gibi farklı ölçme araçları ve farklı örneklerle başka değerlendirme çalışmalarının yapılmasını öneren araştırmalar da mevcuttur (Özer, 2011; Ekmekçi, 2015). Dolayısıyla uzaktan eğitimle yürütülen İngilizce dersi öğretim programlarını kapsamlı bir şekilde değerlendirmede kullanılacak bir ölçeğin literatüre kazandırılması açısından bu çalışma önemlidir. Ölçeğin CIPP değerlendirme modelini temel alarak hazırlanmış olması, ölçeğin kapsamlı bir program değerlendirmesine imkân sağlamaktadır. Ayrıca CIPP modelini oluşturan bağlam, girdi, süreç ve ürün kavramları ayrı ayrı değerlendirme boyutu olarak ele alınabileceği gibi bu kavramların birbiri ile ilişkisi ve bağı aracılığıyla toplam değerlendirme de yapılabileceğinden dolayı (Stufflebeam, 2003) ölçek farklı şekillerde kullanılabilir.

KAYNAKLAR

- Alpar, R. (2000). *Spor, sağlık ve eğitim bilimlerinde uygulamalı istatistik ve geçerlik-güvenirlik*. Ankara: Detay Yayıncılık.
- Anderson, L. M. (1988). *"Likert scales" educational research methodology and measurement. An international handbook*. New York: Pergamon Press.
- Balcı, A. (2015). *Sosyal bilimlerde araştırma*. Ankara: Pegem Akademi Yayıncılık.
- Baş, T. (2013). *Anket nasıl hazırlanır, uygulanır, değerlendirilir?* (7. Baskı). Ankara: Seçkin Yayıncılık.
- Bilen, M. (2006). *Plandan uygulamaya öğretim*. (7. Baskı). Ankara: Anı Yayıncılık.
- Brown, J. D. (1994). *The Elements of Language Curriculum: a systematic approach to program development*. Boston: Heinle And Heinle Publishers.
- Bryman, A., & Cramer, D. (2003). *Quantitative data analysis with SPSS release 8 for Windows*. London: Routledge.

- Burke Sinclair, M. (2012). Utilizing Stufflebeam's Cipp model to evaluate an adult degree completion program. The Degree of Doctor of Education, School of Education in Partial Fulfillment of the Requirements Gardner-Webb University, United Kingdom.
- Büyükoztürk, Ş. (1997). Araştırmaya yönelik kaygı ölçeğinin geliştirilmesi. *Eğitim Yönetimi Dergisi*, 3(4), 453- 464.
- Büyükoztürk, Ş. (2002). Faktör analizi: Temel kavramlar ve ölçek geliştirmede kullanımı. *Kuram ve Uygulamada Eğitim Yönetimi Dergisi*, 32, 470-483.
- Büyükoztürk, Ş., Çakmak, E. K., Akgün, Ö. E., Karadeniz, Ş., & Demirel, F. (2014). *Bilimsel araştırma yöntemleri (16. baskı)*. Ankara: Pegem Akademi Yayıncılık.
- Büyükoztürk, Ş. (2014). *Sosyal bilimler için veri analizi el kitabı*. Ankara: Pegem Yayıncılık.
- Cansu, T. (2010). *Anadolu Üniversitesi ilköğretimde teknoloji uygulamaları e-sertifika programının öğrenen görüşüne göre bağlam, girdi, süreç ve ürün (CIPP) modeli ile değerlendirilmesi*. (Yayınlanmamış Yüksek Lisans Tezi). Anadolu Üniversitesi Sosyal Bilimler Enstitüsü, Eskişehir.
- Comrey, A. L., & Lee, B. H. (1973). *A first course in factor analysis*. New York: Academic Press.
- Çakır, G., & Yurtsever, B. (2012). An implementation of distance education program for teaching common necessary courses in formal education: Karabük University sample. *Procedia-Social and Behavioral Sciences*, 55, 187-195.
- Çallı, İ., İşman, A., & Torkul, O. (2001). Sakarya Üniversitesi'nde uzaktan eğitimin dünü bugünü ve geleceği. *Sakarya Üniversitesi Eğitim Fakültesi Dergisi*, 3, 1-7.
- Demir E. (2014). Uzaktan eğitime genel bir bakış. *Dumlupınar Üniversitesi Sosyal Bilimler Dergisi*, 39, 1-1.
- Demirel, Ö. (2003). *Kuramdan uygulamaya program geliştirme*. Ankara: Pegem Akademi Yayınları.
- DeVellis, R. (2003). *Scale development theory and applications*. London: SAGE Publications.
- Dunn Rankin, P. (2004). *Scaling methods*. London: Routledge.
- Ekmekçi, E. (2015). Distance education in foreign language teaching: evaluations from the perspectives of freshman students. *Procedia-Social and Behavioral Sciences*, 176, 390-397.
- Ertürk, S. (1994). *Eğitimde program geliştirme*. Ankara: Meteksan.
- Floyd, F. J., & Widaman, K. F. (1995). Factor analysis in the development and refinement of clinical assessment instruments. *Psychological Assessment*, 7(3), 286-299.
- Gilchrist, R. S., & Roberts B. R. (1974). *Curriculum development (a humanized systems approach)*. California: Lear Siegler, Inc./ Fearon Publishers.
- Gredler, M. (1996). *Program evaluation*. London: Prentice-Hall.
- Gorsuch, R. L. (1974). *Factor analysis*. Philadelphia: Sounders.
- Gözütok, F. D. (1999). *Cumhuriyet döneminde eğitim. Program değerlendirme*. Ankara: MEB Yayınları.
- Hanchell, V. F. (2014). A program evaluation of a Christian College Baccalaureate Program utilizing Stufflebeam's CIPP model. Education Theses, Dissertations and Projects. Paper 13. Retrieved from https://digitalcommons.gardner-webb.edu/cgi/viewcontent.cgi?article=1011&context=education_etd
- Holmberg, B. (1995). *Theory and practice of distance education*. (2nd ed.) London, Newyork: Routledge.
- İşman, A. (2008). *Uzaktan eğitim*. Ankara: Pegem Akademi.
- Kaiser, H. F. (1974). An index of factorial simplicity. *Psychometrika*, 39(1), 31-36.
- Kalaycı, Ş. (2008). *"Faktör analizi" SPSS uygulamalı çok değişkenli istatistik teknikleri*. Ankara: Asil Yayın Dağıtım Ltd. Şti.
- Karasar, N. (2011). *Bilimsel araştırma yöntemi, kavramlar, ilkeler*. İstanbul: Nobel Akademik Yayıncılık.
- Karataş, H. (2007). *Yıldız Teknik Üniversitesi Modern Diller Bölümü İngilizce II Dersi öğretim programının öğretmen ve öğrenci görüşlerine göre bağlam, girdi, süreç ve ürün (CIPP) modeli ile değerlendirilmesi*. (Yayınlanmamış Yüksek Lisans Tezi). Yıldız Teknik Üniversitesi Sosyal Bilimler Enstitüsü, İstanbul.
- Kaya, M. (2012). *Uzaktan eğitimde öğrenenlerin yabancı dil öğreniminde özerk öğrenme becerileri: Uzaktan İÖLP örneği*. (Yayınlanmamış Yüksek Lisans Tezi), Anadolu Üniversitesi, Eskişehir.
- Kline, R. B. (2010). *Principles and practice of structural equation modeling (3rd ed.)*. New York: Guilford Press.
- Langford, L. L. (2010). *The development and testing of an evaluation model for special education*. Austin: The University Of Texas. Retrieved from <http://gradworks.umi.com/34/17/3417501.html>
- Manasneh, A. (2013). *Translation training in the jordanian context: curriculum evaluation in translator education*. Binghamton University, New York. Retrieved from <https://eric.ed.gov/?id=ED566304>
- McMillan, J. H., & Schumacher, S. (2006). *Research in education: evidence based inquiry (6th ed.)*. Boston: Pearson.
- Mede, E., & Uygun, S. (2014). Evaluation of a language preparatory program: a case study. *ELT Research Journal*, 3(4), 201-221.
- Moore, M. (1973). Toward a theory of independent learning and teaching. *Journal of Higher Education*, 44(9), 661-679.
- Moore, M., & Kearsley, G. (2005). *Distance education: a systems view (2nd ed.)* Belmont: Wadsworth.
- Oliva, P. F. (1997). *Developing the curriculum*. New York: Longman.
- Ornstein, A. C., & Hunkins, F.P. (2004). *Curriculum foundations: principles and theory*. Boston: Allyn And Bacon.
- Özer, B. (2011). *Uzaktan eğitim programlarının öğrenci ve öğretim üyesi görüşleri açısından değerlendirilmesi (Sakarya Üniversitesi örneği)*. (Yayınlanmamış Yüksek Lisans Tezi). Abant İzzet Baysal Üniversitesi, Eğitim Bilimleri Enstitüsü, Bolu.
- Özkanal, Ü., & Hakan, A. G. (2010). Effectiveness of university English preparatory programs: Eskisehir Osmangazi University Foreign Languages Department English Preparatory Program. *Journal of Language Teaching and Research*, 1(3), 295-305.
- Posner, A. (2003). *Analyzing the curriculum*. (2nd ed.) New York: McGraw-Hill Inc.
- Sharma, S. (1996). *Applied multivariate techniques*. USA: John Wiley & Sons.
- Sağlam, M., & Yüksel, İ. (2007). Program değerlendirmede meta-analiz ve meta-değerlendirme yöntemleri. *Dumlupınar Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, 18(2), 175-189.
- Stufflebeam, D. L. (2000). CIPP model (context, input, process, product). Retrieved from <http://srmo.sagepub.com/view/encyclopedia-of-evaluation/n82.xml>

- Stufflebeam D. L. (2003). The CIPP model for evaluation. Annual Conference of the Oregon Program Evaluators Network (OPEN). Portland, Oregon.
- Stufflebeam, D. L. (2007). CIPP evaluation model checklist (2nd ed.). Retrieved from https://www.wmich.edu/sites/default/files/attachments/u350/2014/cippchecklist_mar07.pdf
- Stufflebeam, D. L., & Shinkfield, A. J. (2007), *Evaluation theory, models & applications*. San Francisco: Jossey-Bass.
- Seçer, İ. (2013). *SPSS ve Lisrel ile pratik veri analizi*. Ankara: Anı Yayıncılık.
- Şencan, H. (2005). *Sosyal ve davranışsal ölçümlerde güvenilirlik ve geçerlilik*. Ankara: Seçkin Yayıncılık.
- Tabachnick, B. G., & Fidell L. S. (2012). *Using multivariate statistics*. (6th ed.). USA: Pearson.
- Tavşancıl, E. (2005). *Tutumların ölçülmesi ve SPSS ile veri analizi*. Ankara: Nobel Akademik Yayıncılık.
- Tekin, H. (2000). *Eğitimde ölçme ve değerlendirme*. Ankara: Yargı Yayınları.
- Tokmak, H. C., Baturay, H. M., & Fadde, P. (2013). Applying the context, input, process and product evaluation model for evaluation, research and redesign of an online master's program. *The International Review Research in Open and Distance Learning*, 14(3), 273-293.
- Tunç, F. (2010). *Evaluation of an English language teaching program at a public university using CIPP model*. (Yayınlanmamış Yüksek Lisans Tezi). Orta Doğu Teknik Üniversitesi Sosyal Bilimler Enstitüsü, Ankara.
- Uşun, S. (2006). *Uzaktan eğitim*. Ankara: Nobel Yayın Dağıtım.
- Ünal, M. (2011). Avrupa Birliği erasmus öğrenci öğrenim hareketliliği programının CIPP (bağlam, girdi, süreç, ürün) modeline göre değerlendirilmesi. (Yayınlanmamış Doktora Tezi). Gazi Üniversitesi Eğitim Bilimleri Enstitüsü, Ankara.
- Variş, F. (1988). *Eğitimde program geliştirme teori ve teknikler*. Ankara: Ankara Üniversitesi Eğitim Fakültesi Yayınları.
- Virlan, A. Y. (2014). *Örnek durum incelemesi: Devlet üniversitesi İngilizce hazırlık okulu konuşma becerisi programının CIPP modeliyle değerlendirilmesi*. (Yayınlanmamış Yüksek Lisans Tezi). Yeditepe Üniversitesi Eğitim Bilimleri Enstitüsü, İstanbul.
- Worthern, B. R., Sanders J. R., & Fitzpatrick J. L. (1997). *Program evaluation-alternative approaches and practical guidelines*. New York: Longman.
- Yarmohammadian, M. H., Ahmadi, A., Sadrian, M., & Fooladvand, M. (2011). Evaluation of distance education programs based on the NADE-TDCE 2009-2010. *Procedia-Social and Behavioral Sciences*, 28(1), 117-119.
- Yavuz, S. (2005). Developing a technology attitude scale for pre-service chemistry teachers. *The Turkish Online Journal of Educational Technology*, 4(1), 17-25.