

Uluslararası Öğrencilerin Türkçe Akademik Yazma Becerilerinin Değerlendirilmesine Yönelik Analitik Dereceli Puanlama Anahtarının Geliştirilmesi*

Development of an Analytical Rubric for the Assessment of International Students' Turkish Academic Writing Skills

Emre Bayraktar¹  Şükran Dilidüzgün² 

¹ Öğr. Gör. Dr., İstanbul Üniversitesi-Cerrahpaşa, Yabancı Diller Yüksekokulu, İstanbul, Türkiye

² Prof. Dr., İstanbul Üniversitesi-Cerrahpaşa, Hasan Âli Yücel Eğitim Fakültesi, İstanbul, Türkiye

Makale Bilgileri

Geliş Tarihi (Received Date)

05.08.2024

Kabul Tarihi (Accepted Date)

04.12.2024

*Sorumlu Yazar

Emre Bayraktar

İstanbul Üniversitesi-
Cerrahpaşa Yabancı Diller
Yüksekokulu Alkent 2000
Mah. Yiğittürk Cad. No:5
Büyükkçekmece/İstanbul

emre.bayraktar@iuc.edu.tr

Öz: Bu çalışmanın amacı, uluslararası öğrencilerin Türkçe akademik yazma ihtiyaçlarını temele alarak ilgili becerinin değerlendirilmesine yönelik bir analitik dereceli puanlama anahtarı geliştirmektir. Araştırmada, nitel ve nicel araştırma yöntemlerinin bir arada kullanıldığı karma araştırma yaklaşımlarından keşfedici sıralı desen kullanılmıştır. Çalışmanın nitel boyutunda, geliştirilen taslak dereceli puanlama anahtarının ölçüt ve alt ölçütleri; uluslararası öğrencilerle, dil öğretim merkezlerinde görevli öğretim görevlileriyle, uluslararası öğrencilerin bölüm derslerini yürüten öğretim üyeleriyle yapılan görüşmeler, uluslararası öğrencilerin ürettikleri bilimsel metinlerin çözümlemeleri ve alanyazın taraması bağlamında belirlenmiştir. Ayrıca çalışmada, dereceli puanlama anahtarının geçerliğinin ve güvenilirliğinin sınanması amacıyla tanımlama, görev, yönerge ve puanlama yöntemi olmak üzere dört bölümden oluşan bir performans görevi oluşturulmuştur. Bilimsel metin üretmeyi önceleyen performans görevini, amaçlı örnekleme yöntemleri içerisinde yer alan ölçüt örnekleme kullanılarak belirlenen 157 uluslararası öğrenci tamamlamıştır. Uluslararası öğrencilerin oluşturdukları bu metinler, taslak dereceli puanlama anahtarı kullanılarak değerlendirilmiştir. Bu çalışmanın odaklandığı nicel boyutta ise söz konusu ölçme aracının geçerliği ve güvenilirliği sınanmıştır. Taslak dereceli puanlama anahtarının geçerlik çalışması kapsamında Davis (1992) tekniği ile 10 uzmanın görüşü bağlamında kapsam ve açıklayıcı faktör analizi ile yapı geçerliği incelenmiştir. Güvenirlik çalışması kapsamında ise Cronbach Alpha katsayısı, 6 puanlayıcının değerlendirmeleri üzerinden Sınıf İçi Korelasyon ile Fleiss Kappa uyum katsayısı ve madde geçerliğine kanıt olarak madde-toplam korelasyonları hesaplanmıştır. Çalışma sonucunda, son hâli 5 ana ölçüt ve 24 alt ölçütten oluşan dereceli puanlama anahtarının, uluslararası öğrencilerin Türkçe akademik yazma becerilerini ölçebilecek nitelikte geçerli ve güvenilir bir ölçme aracı olduğu belirlenmiştir.

Anahtar Kelimeler: Yabancı/ikinci dil olarak Türkçe öğretimi, uluslararası öğrenciler, akademik yazma becerileri, analitik dereceli puanlama anahtarı geliştirme

Abstract: The purpose of this study is to develop an analytical rubric for the assessment of the relevant skill based on the Turkish academic writing needs of international students. In the study, exploratory sequential design, which is one of the mixed research methods in which qualitative and quantitative research methods are used together, was used. In the qualitative dimension of the study, the criteria and sub-criteria of the developed draft rubric were determined in the context of interviews with international students, lecturers working in language teaching centers, lecturers conducting departmental courses of international students, analysis of scientific texts produced by international students and literature review. In addition, in order to calculate the validity and reliability of the rubric, a performance task consisting of four parts: description, task, instruction, and scoring method was created. The performance task, which prioritized the production of scientific texts, was completed by 157 international students who were selected using criterion sampling, which is one of the purposive sampling methods. These texts created by international students were evaluated using the draft DPA. In the quantitative dimension, which is the focus of this study, the validity and reliability of the instrument were tested. Within the scope of the validity study of the draft DPA, content validity was examined in the context of the opinions of 10 experts with Davis (1992) technique, and construct validity was examined with exploratory factor analysis. As part of the reliability study, the Cronbach's Alpha coefficient, Intraclass Correlation and Fleiss Kappa agreement coefficient over the evaluations of six raters, and item-total correlations were calculated as evidence of item validity. As a result of the study, it was determined that the rubric, the final version of which consists of five criteria and 24 sub-criteria, is a valid and reliable measurement tool that can measure international students' Turkish academic writing skills.

Keywords: Teaching Turkish as a foreign/second language, international students, academic writing skills, analytical rubric development

Bayraktar, E. ve Dilidüzgün, Ş. (2024). Uluslararası öğrencilerin Türkçe akademik yazma becerilerinin değerlendirilmesine yönelik analitik dereceli puanlama anahtarının geliştirilmesi. *Erzincan Üniversitesi Eğitim Fakültesi Dergisi*, 26(4), 636-659. <https://doi.org/10.17556/erziefd.1528118>

Giriş

Dil becerileri bir dilin anlaşılması, kullanılması ve iletişim kurulması için gereken yeterlikleri ifade etmektedir. Bu beceriler anlama ve anlatmaya dayalı beceriler olmak üzere iki bölüme ayrılmaktadır. İnsanlar, birbirleriyle etkileşim kurmak ve fikirlerini iletmek için hem dinleyerek ve okuyarak anlama hem de konuşarak ve yazarak anlatma becerilerine ihtiyaç duymaktadır. Gerek ana dilin edinilmesi gerekse yabancı/ikinci bir dilin öğrenilmesi için tüm bu becerilerin geliştirilmesi gerekmektedir.

Dil becerilerini geliştirme sürecinde öğrenciler çeşitli düzey ve türde zorluklarla karşılaşabilmektedir. Genel olarak değerlendirildiğinde eğitim-öğretim sürecinde öğrencilerin ve hatta öğretim elemanlarının yaşadığı bu zorlukların daha çok yazma becerisi ile ilişkili olduğu görülmektedir. Yazma, içinde okuma, düşünme ve düzgün ifade etme becerileri içerdiği, ayrıca uygulaması zor bir eylem olduğu için öğrenciler bu öğrenme alanından uzak durma eğilimi içindedir (Ungan, 2007, s. 462). Yorgancı ve Baş'a (2021, s. 71) göre gerek ana dili gerekse yabancı dil öğretiminde üretici/anlatmaya dayalı dil becerilerinin gelişimi diğer

* Bu makale, Prof. Dr. Şükran DİLİDÜZGÜN'ün danışmanlığında Emre BAYRAKDAR tarafından hazırlanan "Uluslararası Öğrencilerin Türkçe Akademik Yazma Becerilerine Yönelik İhtiyaç Analizine Dayalı Ölçme Aracı Geliştirilmesi" başlıklı doktora tezinden üretilmiştir.

becerilere göre hem daha çok zaman almakta hem de daha zor gelişmektedir. Bir üretici dil becerisi olarak yazma; genelde yabancı dil öğretiminde, özelde Türkçenin yabancı/ikinci dil olarak öğretiminde geliştirilmesi en zor ve en son kazanılan beceri olarak kabul edilmektedir. Benzer şekilde Yıldırım ve Nurlu (2016, s. 832) öğrencilerin, dil eğitiminde özellikle üretici/anlatmaya dayalı dil becerilerinde daha çok zorlandığını, bu zorluğun konuşma becerisine oranla yazma becerisinde daha çok yaşandığını ifade etmektedir. Ayrıca yabancı/ikinci dil öğretiminde yazma becerisinde karşılaşılan zorlukların ana dili eğitimine göre daha çok görüldüğünü ve yaşanan bu zorlukların; yazma becerisinin sözcük, dil bilgisi, düşünce üretme, plan, yazım ve noktalama gibi birçok unsuru birlikte ve doğru kullanmayı gerektirmesinden kaynaklandığını belirtmektedir.

Yazma eyleminin akademik amaçlar için kullanımıyla beraber akademik yazma kavramı ortaya çıkmıştır. Söz konusu bu kavram ulusal ve uluslararası alanyazında birçok araştırmacı tarafından farklı şekillerde tanımlanmıştır (Hogue, 2008; Irvin, 2010; Swales & Feak, 2012; Tok, 2013; Bahar, 2014; Aydın, 2015; Kan, 2017; Karagöl, 2018; Güngör 2020). İlgili kavrama ilişkin yapılan tanım ve açıklamalardan hareketle; akademik yazma, akademik bir bağlamda gerçekleştirilen araştırmaların sonuçlarının; biçim, söylem, metin yapısı, plan ve düzenleme ile kaynak kullanımı yönlerinden kendine özgü ilkelere dikkat edilerek raporlandığı ayrıca üst düzey düşünme becerilerinin etkin bir şekilde kullanılmasını gerektiren bir yazma türü şeklinde ifade edilebilir. Türkçenin yabancı/ikinci dil olarak öğretimi bağlamında düşünüldüğünde öğrencilerin daha çok okul yaşantılarında kullanmaya ihtiyaç duyduğu bu kavram, yazma becerisi ile doğrudan bağlantılıdır. Güngör'e (2020, s. 167) göre "yabancı dil öğretim sürecinde, temel düzeyden itibaren yapılan yazma eğitimi çalışmaları, öğrencilerin hedef dili kullanarak akademik hayattaki başarılarını", dolayısıyla akademik yazma becerilerini doğrudan etkilemektedir. Yücelşen'e (2021, s. 15) göre "akademik yazma, yetenekten çok teknik bilgi gerektiren, uğraşsal bir zorunluluktur ve öğrenilebilir, geliştirilebilir bir beceridir." Akademik yazmanın teknik bilgi gerektirmesi yabancı/ikinci dil olarak Türkçe öğretiminde ilgili beceriyi geliştirmeye yönelik ders işleme sürecinin alana özel bileşenlerin dikkate alınarak tasarlanmasını gerektirmektedir. Demirci'ye (2014) göre öğrencilere akademik yazma konusunda deneyim yaşatmak ilgili becerinin gelişmesinde önemli bir etkidir. Özellikle üniversite öğrencilerinin gereksinim duyduğu bu beceriye yönelik deneyim yaşamalarını sağlamak amacıyla çeşitli kurslar ve dersler planlanmıştır. Genelde akademik Türkçe çatısı altında birleşen bu kurs ve derslerin, akademik yazma bağlamında uluslararası öğrencilere; akademik dilin özelliklerini ve kurallarını kavratmayı, alanyazın taraması yaparak geçerli ve güvenilir kaynaklara ulaşmayı, bilimsel metin oluşturmayı, akademik yazılardaki biçimsel özelliklere dikkat etmeyi, alıntı yapmayı ve kaynakça düzenlemeyi, etik ilkelere uymayı vb. farklı uygulamalar üzerinden öğretmeyi hedeflemesi gerekmektedir.

Akademik yazma becerilerinin geliştirilmesinde, deneyim yaşamak kadar düzenli olarak yapıcı geri bildirimlerin verilmesi ve hatalardan ders çıkarılması da son derece önemlidir (Pineteh, 2014; Demirci, 2014). Hatalardan ders çıkarabilmek, yanlışları ve eksiklikleri fark edebilmektir. Bireylerin güçlü ve zayıf yönlerini belirleyerek eksikliklerini gidermek, yanlışları düzeltmek ve gelişimlerine katkı sağlamak için ölçme ve değerlendirme büyük önem

taşımaktadır. Bireylerin yeterliklerini ve yetersizliklerini saptayabilmek için ölçülen niteliğe uygun ölçme araçları kullanılmalıdır (Polat-Demir, 2020, s. 321).

Üst düzey düşünme becerilerinin kullanılmasını gerektiren ve biçim, söylem, metin, plan, kaynak kullanımı, etik gibi birçok boyutu içeren akademik yazma becerilerinin ölçülmesinde, değerlendirilmesinde ve geliştirilmesinde kullanılabilecek en etkili ölçme araçlarından biri ise analitik (çözümleyici) dereceli puanlama anahtarıdır. Dereceli puanlama anahtarı, diğer alan ve konularda olduğu gibi yabancı/ikinci dil olarak Türkçe öğretimi alanında ve akademik yazma konusunda da büyük bir öneme sahiptir. Dereceli puanlama anahtarının yabancı/ikinci dil olarak Türkçe öğretiminde öğrencilerin akademik yazma becerilerinin gelişmesine yönelik tüm paydaşlara sağladığı faydalar aşağıda özetlenmiştir (Erman-Aslanoglu, 2003; Moskal, 2000; Goodrich-Andrade, 2005; Parlak ve Doğan, 2014; Çepni, 2015; Kutlu vd., 2017):

- Dereceli puanlama anahtarları ile öğrencilerin akademik yazma bağlamında düzeyleri ve eksiklikleri detaylı bir biçimde belirlenmektedir. Bir geri bildirim kaynağı özelliği göstermesi nedeniyle dereceli puanlama anahtarı, üst düzey zihinsel süreçlerin işletilmesini önceleyen akademik yazma becerisi noktasında öğrencilerin ve öğretim elemanlarının gözden geçirme ve iyileştirmeye yönelik etkili kararlar almasına olanak sağlamaktadır.
- Öğrencilerin oluşturduğu akademik metinlerin geçerli ve güvenilir bir şekilde puanlanabilmesine imkân sunmaktadır. Bu yöntem, öğretim elemanlarına, öğrencilerin çalışmalarını hangi ölçütlere göre puanlayacaklarını önceden bilme imkânı vermekte ve böylelikle değerlendirme ve puanlama güvenilirliği sağlamaktadır. Ayrıca, değerlendirme sürecinin kısalmasına ve basitleşmesine olanak tanımaktadır.
- Üzerinde uzlaşa sağlanamamış ve kapsamı tam olarak belirlenememiş Türkçe akademik yazma becerilerinin çerçevesini çizdiği, dikkat edilmesi gereken ölçütleri tanımladığı için nesnel değerlendirmelere olanak sağlarken; ilgili beceriden beklenen yeterlikleri tam olarak belirleyemeyen, konu hakkında yeterli bilgi ve deneyimi olmayan, pedagojik eksiklikleri bulunan öğretim elemanlarına yol gösterici olma niteliği taşımaktadır.
- Öğrenciler, hangi ölçütlere göre puanlanacaklarını ve en iyi performansın neleri içerdiğini bildikleri için çalışmalarını ona göre bilinçli bir şekilde düzenlemeye ve ilgili beceriyi geliştirmeye yönelik çaba göstereceklerdir.
- Öğrencilerin değerlendirilecekleri ölçütleri önceden bilmeleri, kendilerinden beklenen performansın boyutlarından, bu boyutların tanımlarından ve düzeylerinden haberdar olmaları, akademik başarılarına katkı sağlamanın yanı sıra öğretim sürecine ve başta öğretim elemanına olmak üzere bu süreçteki tüm paydaşlara olan tutumlarını olumlu etkilemektedir. Ayrıca öğrencilerin sözü edilen boyutlardan haberdar olmaları, kendilerini rahat ve özgüvenli hissetmelerini sağlamaktadır. Bu durum da ilgili performansı göstermeye daha istekli olmalarına ve kaygı düzeylerinin düşmesine yardımcı olmaktadır. Böylece dereceli puanlama anahtarının kullanımı, dil becerileri içerisinde en çok zorlanılan ve kaygı duyulan yazma

becerisi ile doğrudan ilişkili olan akademik yazma konusunda öğrencilerin tutumlarının değişmesine ve ilgili çalışmalarda başarı göstermelerine katkı sağlamaktadır.

Alanyazın incelendiğinde yabancı/ikinci dil olarak Türkçe öğretimi alanında daha çok yazma becerisinin değerlendirilmesine yönelik dereceli puanlama anahtarı geliştirme çalışmalarının yer aldığı (Yorgancı ve Baş, 2021; Durmuş, 2022; Kahveci, 2022; Harmankaya vd., 2022; Tekin ve Sallabaş, 2022) görülmekle birlikte farklı hedef kitlelerin akademik yazma becerilerini ölçmek amacıyla çeşitli türlerde ölçme araçları geliştirildiği de belirlenmiştir (Aydın, 2015; Izawa et al., 2017; Shekarabi, 2017; Al-Abdullatif et al., 2020; Polat-Demir, 2020; Baz-Bolluk ve Batur, 2023). Ancak uluslararası öğrencilerin Türkçe akademik yazma ihtiyaçlarını temele alarak ilgili becerinin değerlendirilmesine yönelik göreve dayalı geçerliği ve güvenilirliği çok boyutlu çözümleme teknikleri ile sınanmış analitik dereceli puanlama anahtarı geliştirme çalışmasına rastlanmamıştır. Bu nedenle çalışmada, uluslararası öğrencilerin Türkçe akademik yazma ihtiyaçlarını temele alarak ilgili becerinin değerlendirilmesine yönelik bir analitik dereceli puanlama anahtarının geliştirilmesi amaçlanmıştır. Bu temel amaç doğrultusunda çalışmada yanıtlanması hedeflenen sorular şöyle sıralanabilir:

1. Uluslararası öğrencilerin Türkçe akademik yazma becerilerine yönelik geliştirilen analitik dereceli puanlama anahtarının geçerlik analizi sonuçları nasıldır?
 - a. Uluslararası öğrencilerin Türkçe akademik yazma becerilerine yönelik geliştirilen analitik dereceli puanlama anahtarının kapsam geçerliği nasıldır?
 - b. Uluslararası öğrencilerin Türkçe akademik yazma becerilerine yönelik geliştirilen analitik dereceli puanlama anahtarının yapı geçerliği nasıldır?
2. Uluslararası öğrencilerin Türkçe akademik yazma becerilerine yönelik geliştirilen analitik dereceli puanlama anahtarının güvenilirlik analizi sonuçları nasıldır?
 - a. Uluslararası öğrencilerin Türkçe akademik yazma becerilerine yönelik geliştirilen analitik dereceli puanlama anahtarının Croanbach Alfa güvenilirlik katsayısı nasıldır?
 - b. Uluslararası öğrencilerin Türkçe akademik yazma becerilerine yönelik geliştirilen analitik dereceli puanlama anahtarına göre yapılan değerlendirmelerde puanlayıcılar arasındaki uyum nasıldır?
 - c. Uluslararası öğrencilerin Türkçe akademik yazma becerilerine yönelik geliştirilen analitik dereceli puanlama anahtarındaki alt ölçütlerin madde geçerlik düzeyleri nasıldır?

Geliştirilen analitik dereceli puanlama anahtarının nesnel değerlendirme yapmaya yardımcı olması, öğrencilerin Türkçe akademik yazma becerilerinin gelişimine katkı sağlaması, hem öğretim görevlilerine hem de öğrencilere ilgili becerinin başarımına ilişkin geri bildirim olanağı sunarak yol gösterici olması bağlamında alana katkı sağlayacağı düşünülmektedir.

Yöntem

Araştırmanın Deseni

Uluslararası öğrencilerin Türkçe akademik yazma becerilerinin değerlendirilmesine yönelik analitik dereceli puanlama anahtarının geliştirilmesinin amaçlandığı bu

çalışmada, karma desen stratejisi içerisinde yer alan keşfedici sıralı desen kullanılmıştır. Keşfedici sıralı desende araştırmacı nitel sonuçlara bağlı olarak oluşturulan aşamaların ortasında bir araç geliştirir ve bu aracı, nicel verileri toplarken kullanır (Creswell & Plano-Clark, 2018, s. 94). “Keşfedici sıralı desen teorik olarak küçük bir çalışma grubundan toplanan nitel veriler ile başlamaktadır. Nitel verilerin analiziyle elde edilen bulgu ve sonuçların ışığında nicel bir ölçme aracı geliştirilmektedir ve bu nicel aşama keşfedici desenin ikinci ve son aşamasını oluşturmaktadır” (Teddle & Tashakkori, 2006, s. 21 akt. Özdemir, 2018, s. 495). “Bu aşamada geliştirilen ölçme aracı, daha geniş bir örneklem üzerinde uygulanarak istatistiksel olarak test edilmektedir. Ölçülmek istenen değişkeni ölçebilecek nitelikte ve yeterlikte bir ölçme aracına ulaşılmaması ile araştırma sonlandırılmaktadır” (Özdemir, 2018, s. 495).

Keşfedici sıralı desenin adımları temele alınarak, çalışmada öncelikle akademik yazma bileşenlerini belirlemek için ulusal ve uluslararası alanyazın (Uzun, 2001; Goodric-Andrade, 2001; Bayat, 2009; Swales & Feak, 2012; Tok, 2012; Tok, 2013; Bahar, 2014; Şenöz-Ayata, 2014; The WIDA English Language Development Standards and Resource Guide [WIDA], 2014; Aydın, 2015; Razi, 2015; Yahşi-Cevher ve Güngör, 2015; Kan ve Gedik, 2016; Tok ve Gönülal, 2016; Ulutaş, 2016; Demir, 2017; Izawa et al., 2017; Kan, 2017; Shekarabi, 2017; Deniz ve Karagöl, 2018; Karagöl, 2018; Tüfekçioğlu, 2018; Azizoglu vd., 2019; Demiriz ve Okur, 2019; Konyar, 2019; Seyedi, 2019; Yurdakul-Güde, 2019; Al-Abdullatif et al., 2020; Bailey, 2020; Polat-Demir, 2020, Yücelşen, 2021; Global Scale of English Learning Objectives for Academic English [GSE], 2022) detaylı olarak taranmıştır. Ardından uluslararası öğrencilerin Türkçe akademik yazma ihtiyaçlarının belirlenmesi amacıyla uluslararası öğrencilerle, dil öğretim merkezlerinde görevli öğretim görevlileriyle, uluslararası öğrencilerin bölüm derslerini yürüten öğretim üyeleriyle yarı yapılandırılmış şekilde tasarlanan görüşmeler gerçekleştirilmiştir. Yapılan görüşme analizleri sonucunda belirlenen gereksinimlerin öğrencilerin yazma ürünlerindeki yansımaları tespit etmek, akademik yazma ihtiyaçlarını detaylandırmak ve taslak dereceli puanlama anahtarının ölçütlerini belirginleştirmek için uluslararası öğrencilerden verilen bir yazma görevini gerçekleştirmeleri istenmiştir. Toplanan tüm bu nitel verilerin çözümlenmesi sonucunda; uluslararası öğrencilerin akademik yazma konusunda yetersiz olduğu ve biçim, akademik söylem, metin yapısı, plan ve düzenleme, kaynaklardan yararlanma konularında becerilerini geliştirmelerine olanak sağlayan bir öğretime ve bu bağlamda geliştirilen ders materyalleri ile ölçme ve değerlendirme araçlarına gereksinim duydukları belirlenmiştir. Bütüncül bir bakış açısı ile değerlendirilen bu çözümlenmeler ve alanyazındaki bilgilerden hareketle taslak dereceli puanlama anahtarının ölçütleri, alt ölçütleri ile performans düzey ve tanımları belirlenmiştir. Nitel verilere dayalı olarak yapılandırılan ölçme aracı kullanılarak uluslararası öğrencilere verilen bilimsel metin yazmayı önceleyen performans görevleri puanlanmıştır. Son olarak, bu çalışmanın odaklandığı nicel boyutta ise söz konusu ölçme aracının geçerliği ve güvenilirliği sınanmıştır.

Çalışma Grubu

Taslak dereceli puanlama anahtarını geçerlik ve güvenilirlik açısından sınamak için araştırmacılar tarafından geliştirilen performans görevini, amaçlı örnekleme yöntemleri içerisinde yer alan ölçüt örnekleme kullanılarak belirlenen 157 uluslararası öğrenci tamamlamıştır. Nitel araştırma geleneği

içerisinde ortaya çıkan ve nitel araştırmalar için en uygun örnekleme yöntemi olarak tanımlanan amaçlı örnekleme çeşitlerinden olan ölçüt örneklemede temel yaklaşım, önceden belirlenmiş bir dizi ölçütü karşılayan durumların çalışılmasıdır (Yıldırım ve Şimşek, 2011). Çalışmada ölçüt örnekleminin seçilme nedeni, önceden tanımlanan ölçütlere göre katılımcıların belirlenme düşüncesidir. Çalışma kapsamındaki uluslararası öğrencilerin ülkemize yükseköğrenim görme amacıyla gelmiş ve C1 düzeyinden mezun olup akademik Türkçe dersi almış olması ölçütlerini karşılaması beklenmiştir.

Veri toplanan katılımcıların kişisel bilgilerine bakıldığında cinsiyet olarak 82'si kadın; 75'i erkektir. Yaş olarak ise 74'ü 17-20 yaş aralığında, 48'i 21-24 yaş aralığında, 23'ü 25-28 yaş aralığında, 9'u 29-32 yaş aralığında, 3'ü 33 yaş ve üzerindedir.

Dil değişkeni açısından bakıldığında katılımcıların 46'sı Arapça, 14'ü Endonezce, 13'ü Farsça, 9'u Fransızca, 8'i Urduca, 5'i Amharca, 5'i Arnavutça, 5'i İngilizce, 4'ü Rusça, 4'ü Yorubaca, 3'ü Boşnakça, 3'ü Kazakça, 3'ü Kürtçe, 3'ü Peştuca, 2'si Ahıska Türkçesi, 2'si Bengalce, 2'si Burmaca, 2'si İspanyolca, 2'si Malayca, 2'si Moğolca, 2'si Özbekçe, 2'si Sırpça, 2'si Somalice, 2'si Türkmençe, 1'i Berberice, 1'i Çeçence, 1'i Çevaca, 1'i Dinka dili, 1'i Ewece, 1'i Karadağca, 1'i Kaşkıyca, 1'i Rundice, 1'i Oromca, 1'i Portekizce, 1'i Svahilice, 1'i Tacikçe ana dillerine sahiptir.

Katılımcıların 95'i lisans, 49'u yüksek lisans ve 13'ü doktora düzeyinde öğrenim görmektedir/görecektir. Ayrıca 67 katılımcı Sakarya Üniversitesi Türk Dili Öğretimi Uygulama ve Araştırma Merkezinde (TÖMER), 26 katılımcı Ankara Üniversitesi Türkçe ve Yabancı Dil Uygulama ve Araştırma Merkezinde (TÖMER), 22 katılımcı Nevşehir Hacı Bektaş Veli Üniversitesi Türkçe ve Yabancı Dil Öğretimi Uygulama ve Araştırma Merkezinde (TÖMER), 16 katılımcı İstanbul Üniversitesi-Cerrahpaşa Yabancı Diller Yüksekokulunda, 10 katılımcı İstanbul Üniversitesi Dil Merkezinde (DİLMER), 9 katılımcı Sağlık Bilimleri Üniversitesi Türkçe ve Yabancı Dil Uygulama ve Araştırma Merkezinde (TÖMER), 5 katılımcı Bolu Abant İzzet Baysal Üniversitesi Türkçe Öğretimi Uygulama ve Araştırma Merkezinde (TÖMER), 1 katılımcı Ankara Yıldırım Beyazıt Üniversitesi Dil Eğitimi Uygulama ve Araştırma Merkezinde (DİLMER), 1 katılımcı Bursa Uludağ Üniversitesi Türkçe Öğretimi Uygulama ve Araştırma Merkezinde (ULUTÖMER) Türkçe öğrenimlerini tamamlayarak C1 düzeyine ulaşmıştır. Ankara Yıldırım Beyazıt Üniversitesi DİLMER'de öğrenim gören 1 katılımcı haricindeki tüm katılımcılar aynı kurumda akademik Türkçe derslerine katılmıştır. Sözü edilen 1 katılımcı Bolu Abant İzzet Baysal Üniversitesi TÖMER'de ilgili dersi tamamlamıştır.

Veri Toplama Araçları

Bu çalışmada veri toplamak amacıyla performans görevi ve taslak dereceli puanlama anahtarı kullanılmıştır. Söz konusu araçların geliştirilme süreçleri, ilgili başlıklar altında ayrıntılı olarak açıklanmıştır.

Performans Görevi

Bu çalışmada, araştırmacılar tarafından uluslararası öğrencilerin akademik yazma becerilerini yansıtmalarına olanak sağlayan bir performans görevi hazırlanmıştır. "Performans görevi, öğrencilerin performanslarını belirlemek amacıyla ders kazanımları ve becerileriyle ilişkili bireysel veya grup hâlinde yapılan uzun soluklu çalışmalardır" (Kutlu vd., 2008, s. 12). Kutlu vd.'e (2008) göre bir performans görevi tanımlama, görev, yönerge ve puanlama yöntemi olmak

üzere dört bölümden oluşmaktadır. Çalışma bağlamında geliştirilen performans görevi bu bölümler gözetilerek hazırlanmıştır ve geliştirilme süreci aşağıda ayrıntılı olarak açıklanmıştır:

Tanımlama bölümünde, performans görevinin ait olduğu ders, dil düzeyi ve kazanımları yer almaktadır. Bunun yanı sıra öğrencilerden beklenen performans tanımlanmış ve puanlama yöntemi belirtilmiştir. Beklenen performansa ilişkin kazanımlar Ankara Üniversitesi TÖMER Öğretim Programı'ndan (2015), Avrupa Ortak Başvuru Metni'nden (2021) ve Global Scale of English Learning Objectives for Academic English'ten (2022) alınmıştır.

Görev bölümünde, uluslararası öğrencilere beslenme ile ilgili on bir farklı kaynak bölümü verilmiştir. Öğrencilerden bu kaynaklardan yararlanarak beslenme kavramının farklı bir boyutunu ele alan bir derleme makalesi yazmaları istenmiştir. Belirlenen performans görevi için yabancılara Türkçe öğretimi alanında deneyimli 3 öğretim üyesinden ve uluslararası öğrencilere Türkçe öğreten kurumlarda görev yapan 4 öğretim görevlisinden uzman görüşü alınmıştır.

Yönerge bölümünde, performans görevinin istenilen düzeyde yerine getirilmesini sağlamak amacıyla görev oluşturulurken dikkat edilecek noktalar, teslim edilecek süre, ortaya konacak ürünün niteliği gibi bilgiler belirtilmiştir.

Puanlama Yöntemi bölümünde, tamamlanmış performans görevinin değerlendirileceği ölçüt ve alt ölçütlere yer verilmiştir.

Taslak Dereceli Puanlama Anahtarı

Bu çalışma kapsamında geliştirilen "Uluslararası Öğrencilere Yönelik Türkçe Akademik Yazma Analitik Dereceli Puanlama Anahtarı"nın geliştirilme süreci aşağıda aşamalar hâlinde anlatılmıştır. Bu aşamalar, çeşitli araştırmacılar tarafından açıklanan dereceli puanlama anahtarının geliştirilme adımları dikkate alınarak oluşturulmuştur (Popham, 1997; Goodrich-Andrade, 2000; Moskal, 2000; Mertler, 2001; Danielson & Hansen, 2016; Kutlu vd., 2017; Bıkmaz-Bilgen; 2023):

1. *Dereceli puanlama anahtarının kullanımının belirlenmesi*: Bu çalışmada uluslararası öğrencilerin Türkçe akademik yazma becerilerini değerlendirmek için dereceli puanlama anahtarının geliştirilmesi amaçlanmaktadır. Geliştirilen dereceli puanlama anahtarından, öğretim elemanlarına ve uluslararası öğrencilere geri bildirim kaynağı olması, söz konusu becerilerin gelişimini desteklemesi ve nesnel değerlendirmelere olanak sağlaması beklenmektedir.

2. *Kullanılacak dereceli puanlama anahtarının türüne karar verilmesi*: Çalışmada, araştırmacılar tarafından ölçütleri ve tanımları belirlenmiş bir görev bağlamında uluslararası öğrencilerin Türkçe akademik yazma becerilerinin ayrıntılı bir şekilde değerlendirilmesi istendiğinden "göreve özel analitik dereceli puanlama anahtarı" tercih edilmiştir.

3. *Performansı belirlemede kullanılacak ölçütlerin belirlenmesi*: Taslak dereceli puanlama anahtarı "Biçim" (Dil bilgisi, noktalama işaretleri, yazım kuralları, anlaşılabilirlik, sayfa düzeni), "Akademik Söylem" (Söz varlığı, cümle yapısı, üstsöylem belirleyicileri), "Metin Yapısı" (Konu, araştırmanın problem durumu ve amacı, bağdaşıklık, tutarlılık, paragraf yapısı, bilgileri sentezleme, görsel sunum), "Plan ve Düzenleme" (Başlık, özet, anahtar sözcükler, giriş, başlıklandırma, gelişme, tartışma, sonuç ve öneriler), "Kaynaklardan Yararlanma" (Kaynakça, doğrudan ve dolaylı alıntı, akademik etik ilkeler) olmak üzere 5 ana ölçüt ve 25 alt ölçütten oluşmaktadır.

Tablo 1. Performans düzeyi puan aralıkları

Aralık kat sayısı = (En Yüksek Puan-En Düşük Puan) / Ölçüt Sayısı			
(4-1)/4 = 3/4 = 0,75		(96-24)/4 = 72/4=18	
1 – 1,75	→ zayıf	24 – 41	→ zayıf
1,76 – 2,51	→ geliştirilebilir	42 – 59	→ geliştirilebilir
2,52 – 3,27	→ yeterli	60 – 77	→ yeterli
3,28 – 4,03 ~ 4,00	→ yetkin	78 – 95 ~ 96	→ yetkin

4. *Performans düzeylerinin belirlenmesi ve düzey tanımlamalarının yapılması:* Bu aşamada en yetersiz ve en yetkin performans düzeyleri saptanarak her alt ölçüt, 1 ile 4 arasında derecelendirilmiştir. Bu bağlamda, beklenen performansı çok az yerine getirerek “zayıf” düzeyde olan öğrencilerin alacağı puan 1, “geliştirilebilir” düzeyde olan öğrencilerin alacağı puan 2, “yeterli” düzeyde olan öğrencilerin alacağı puan 3, beklenen performansı tam olarak yerine getirerek “yetkin” düzeyde başarı gösteren öğrencilerin alacağı puan 4 olarak belirlenmiştir. Düzey tanımlamaları ise alanyazın taranarak, ilgili programların kazanımları (WIDA, 2014; TÖMER, 2015; AOBM, 2021; GSE, 2022) incelenerek, görüşme sonuçları değerlendirilerek ve uluslararası öğrencilerin yazma ürünlerinin çözümlenmesinden elde edilen veriler göz önüne alınarak yapılmıştır. Performans düzey tanımlamaları betimlenirken göreceli ifadelerden kaçınılmış, oldukça net ifadeler kullanılmaya çalışılmıştır. Bu amaçla ölçme aracının herhangi bir zamanda, herhangi bir kişi tarafından kullanıldığında yaklaşık aynı değerlendirmeye ulaşabilmesi adına tanımlamalarda genellikle nitel ölçütler yerine nicel ölçütler tercih edilmiştir. Diğer yandan özellikle uluslararası öğrencilerin akademik yazma ürünlerinin incelenmesinden elde edilen sonuçlar, performans düzeylerine karşılık gelen tanımlamalardaki nicel ifadelerin belirlenmesine ve ölçme aracında yer alan ilgili alt ölçütün en yüksek ve en düşük performansın sınırlarının çizilmesine yardımcı olmuştur.

5. *Uzman görüşlerinin alınması:* Bu aşamada, taslak dereceli puanlama anahtarının kapsam geçerliğini sağlamak, alt ölçütlerin ölçütlere uygunluğunu belirlemek ve düzey tanımlarının anlaşılır olup olmadığını ortaya koymak için uzman görüşü alınmıştır. Uzman grubu, Türkçe eğitimi ve Türkçenin yabancı/ikinci dil olarak öğretiminde akademik yazma konusunda deneyimli 6 öğretim üyesinden, akademik yazma ve dil becerileri konusunda çalışmaları olan 2 eğitimde ölçme ve değerlendirme uzmanından, akademik yazma konusunda deneyimli 1 İngilizce Öğretmenliği alanında görevli öğretim üyesinden ve uluslararası öğrencilere akademik Türkçe dâhil olmak üzere tüm düzeylerde ders verme deneyimine sahip 1 öğretim görevlisinden oluşmaktadır.

6. *Deneme uygulamasının yapılması:* Bu aşamada görevi en iyi ve en yetersiz şekilde yerine getiren 8 öğrencinin çalışması üzerinden deneme uygulaması yapılmıştır. Söz konusu çalışmaları 2 öğretim görevlisi taslak dereceli puanlama anahtarını kullanarak puanlamıştır. Deneme uygulamasının ardından yapılan değerlendirmelere göre taslak dereceli puanlama anahtarında gerekli düzenlemeler yapılmıştır.

7. *Geçerlik ve güvenilirlik çalışmasının yapılması:* 157 öğrenci tarafından tamamlanan performans görevleri taslak dereceli puanlama anahtarı kullanılarak puanlanmıştır. Bu puanlamalardan elde edilen sonuçlardan hareketle “Verilerin Çözülmesi” bölümünde açıklanan istatistiksel

çözümlenmeler aracılığıyla ilgili ölçme aracının geçerliği ve güvenilirliği sınanmıştır.

8. *Dereceli puanlama anahtarına son hâlinin verilmesi:* Geçerlik ve güvenilirlik analizlerinden elde edilen sonuçlara göre taslak dereceli puanlama anahtarına son hâli verilmiştir. Buna göre gözlenmek istenen performansı çok az yerine getiren ve önemli ölçüde eksiklikleri bulunan öğrencilerin alacağı en düşük puan 24, tüm görevlerde gözlenmek istenen performansı tam olarak yerine getiren öğrencilerin alacağı en yüksek puan 96 olarak belirlenmiştir. DPA'nın puan aralığı katsayısı 4 puan ve 96 puan üzerinden hesaplanmıştır.

Tablo 1’de gösterilen hesaplamalara bağlı olarak 4 puan üzerinden belirlenen aralık kat sayısı 0,75 olarak bulunmuştur. Bu değere bağlı olarak performans düzeyleri belirlenmiş ve yetkin düzeye karşılık gelen 4,03 değeri 4’e yuvarlanmıştır. 96 puan üzerinden hesaplanan aralık kat sayısı ise 18 olarak belirlenmiştir. Hesaplanan bu aralık katsayısına bağlı olarak performans düzeyleri dört eşit aralığa dağıldığında yetkin düzeye karşılık gelen 95 puan testten alınabilecek en yüksek puan olan 96’ya tamamlanmıştır.

Veri Toplama Süreci

Veri toplama süreci öncesinde çalışmanın etik kurallara uygun bir biçimde yapılandırılması için gerekli adımlar izlenmiştir. Bu bağlamda öncelikle, çalışma için gereksinim duyulan verilerin toplanabilmesi için İstanbul Üniversitesi-Cerrahpaşa Sosyal ve Beşerî Bilimler Araştırmaları Etik Kuruluna başvurulmuştur. İlgili kurumun 14.06.2022 tarihinde gerçekleştirmiş olduğu toplantıda, 2022/202 numaralı ve E-74555795-050.01.04-412030 sayılı kararı ile etik kurul onayı alınmıştır. Söz konusu onayın alınmasının ardından çalışma kapsamında yapılandırılan performans görevini akademik Türkçe öğrenimi gören ya da görmekte olan öğrencilerin tamamlaması beklendiği için söz konusu programı etkin olarak yürüten kurumlarla iletişime geçilmiştir. Kurumlardan çalışmanın gerçekleştirilebilmesi için gerekli izinler alınmıştır. Araştırmacı, ulaşabildiği çeşitli kurumlara bizzat giderek hem öğrenci hem de dersin öğretim görevlilerine performans görevinin bölümlerini, beklenen uygulamayı, önemli noktaları açıklamıştır. Ulaşma olanağı kısıtlı olan kurumlardan ise öğretim görevlileri aracılığıyla verileri toplamıştır. Araştırmacı tarafından veri toplama sürecinin başında gerekli bilgilendirmeler yapılmış ve uygulama boyunca iletilen tüm sorular yanıtlanmıştır. Çalışma kapsamında sadece elektronik ortamda oluşturulmuş metinler kabul edilmiştir. Araştırmacıya e-posta aracılığıyla toplamda 169 çalışma (derleme makalesi) ulaştırılmıştır. İlk okumaların ardından değerlendirmeye alınabilecek 157 çalışma belirlenmiştir. Öncelikle bu çalışmalardaki öğrenci kişisel bilgileri gizlenmiştir. Ardından hem öğrenci çalışmaları hem de değerlendirmede kullanılacak taslak dereceli puanlama anahtarı aynı klasörde, iki ayrı dosyada “Katılımcı 1”, “Katılımcı 2” şeklinde kodlanmıştır. Bu şekilde değerlendirilmeye hazır hâle gelen veriler, araştırmacı tarafından taslak dereceli puanlama anahtarına göre ayrı ayrı puanlanmıştır. Ayrıca performans görevini tam

olarak yerine getiren 25 öğrencinin çalışması, puanlayıcılar arası uyumu tespit etmek için araştırmacılar haricinde yabancılara Türkçe öğretimi alanında ve akademik yazma konusunda deneyimli, devlet ve vakıf üniversitelerinin dil öğretim merkezlerinde görev yapan 5 öğretim görevlisine değerlendirmeleri için aynı kodlama ve klasör düzeni içerisinde e-posta aracılığıyla gönderilmiştir. Puanlamalara başlamadan önce dereceli puanlama anahtarı araştırmacı tarafından 4 puanlayıcıya Zoom programı aracılığıyla çevrim içi ortamda; 1 puanlayıcıya ise yüz yüze tanıtılmıştır. Ayrıca değerlendirilmeye dâhil olmayan bir çalışma birlikte puanlanarak dereceli puanlama anahtarı ile ilgili sorular yanıtlanmıştır. Değerlendirmelerini tamamlayan puanlayıcılar, ilgili klasörü e-posta aracılığıyla araştırmacıya göndermiştir. Araştırmacı ve puanlayıcılar tarafından yapılan puanlama verilerinin tamamı, dereceli puanlama anahtarının geçerliğine ve güvenilirliğine yönelik istatistiksel çözümlenmeler için Excel’de oluşturulan şablona işlenmiştir.

Verilerin Çözülmesi

Bu çalışmada, uluslararası öğrencilerin Türkçe akademik yazma becerilerini değerlendirmek amacıyla geliştirilen dereceli puanlama anahtarının geçerliği ve güvenilirliği sınanmıştır. Bu amaçla öncelikle geçerlik bağlamında kapsam ve yapı geçerliği değerlendirilmiştir. Kapsam geçerliği, ölçme aracındaki maddelerin ölçülmek istenen davranışı (özelliği) ölçmede nicelik ve nitelik olarak yeterli olup olmadığının göstergesi olarak ifade edilmektedir (Büyüköztürk, 2018, s. 180). Dereceli puanlama anahtarındaki ölçüt ve alt ölçütlerin kapsam geçerlik indekslerinin belirlenmesi için Davis (1992) tekniği tercih edilmiştir. Davis (1992) tekniğine göre uzman görüşleri:

- “Madde ilgili özelliği yansıtıyor.”
- “Madde biraz düzeltilmeye gereksinim duyuyor.”
- “Madde oldukça düzeltilmeye gereksinim duyuyor.”
- “Madde ilgili özelliği yansıtıyor.” biçiminde dörtlü derecelendirilmektedir.

Uzmanlardan çalışma bağlamında oluşturulan uzman değerlendirme formunda, her bir alt ölçütü amacına uygunluk durumuna göre değerlendirmeleri istenmiştir. Bu kapsamda uzmanlar, madde (alt ölçüt) ilgili özelliği yansıtıyorsa “A”, biraz düzeltilmeye gereksinim duyuyorsa “B”, oldukça düzeltilmeye gereksinim duyuyorsa “C” ve ilgili özelliği yansıtıyorsa “D” seçeneğini işaretlemişlerdir. Davis tekniğinde maddelere (ölçütlere) ilişkin kapsam geçerlik indeksi (KGİ), A ve B seçeneğini işaretleyen uzmanların sayısının toplam uzman sayısına bölünmesi ile elde edilmektedir. Bu değer 0,80 ve üzerinde olması beklenmektedir.

Bu çalışmada geliştirilen dereceli puanlama anahtarının yapı geçerliği de sınanmıştır. “Yapı geçerliği, ölçme aracıyla ölçülmek istenen niteliğin başka özellikler karışmadan ölçebilme derecesidir.” (Güler, 2019, s. 49). Yapı geçerliğini incelemek için açılımlı ve doğrulayıcı faktör analizi yaygın olarak kullanılan iki temel analiz türüdür. Bu çalışmada yapı geçerliği, açılımlı faktör analizi (AFA) ile değerlendirilmiştir. Başol’a (2019, s. 180) göre “yeni geliştirilen bir ölçeğin ölçtüğü temel yapıları ortaya koymak için açılımlı faktör analizi yapılmaktadır.” Çokluk vd. (2021, s. 205), açılımlı faktör analizini uygulamadan önce, meydana gelebilecek sorunları en aza indirebilmek amacıyla; örneklem büyüklüğü, normallik, doğrusallık ve çoklu bağlantı gibi kavramların sorgulanması gerektiğini ifade etmektedir. Bu bağlamda çalışmada açılımlı faktör analizinden önce

veri seti üzerinde; çoklu bağlantı sorunu olup olmadığı sınanmış, normallik kontrolü için Basıklık ve Çarpıklık katsayıları hesaplanmıştır. Diğer yandan çok değişkenli normallik varsayımını sınamak ve veri yapısının örneklem büyüklüğü açısından uygunluğunu belirlemek için her ölçüt için ayrı ayrı Kaiser-Meyer-Olkin (KMO) değeri ile Bartlett küresellik testi sonuçları incelenmiştir. Söz konusu analizlerin sonuçlarına göre veri setinin faktör analizi uygulamaya uygun hâline temel bileşenler analizi kullanılarak AFA yapılmıştır. Dereceli puanlama anahtarındaki alt ölçütlerin, bulunduğu ana ölçütü karşılama durumlarını belirlemek için faktör yük değerleri hesaplanmıştır. Bunun yanı sıra dereceli puanlama anahtarının faktör yapısını belirlemek için özdeğer bulguları da incelenmiştir. Ek olarak, faktör sayısına karar vermeden önce yardımcı bir yöntem olarak çizgi grafiğine de bakılmıştır. Uygulanan faktör analizleri sonucunda her ölçüt tek bileşenden oluştuğu için döndürme işlemine gereksinim duyulmamıştır.

Dereceli puanlama anahtarının güvenilirliğini belirlemek için ise iç tutarlılık anlamında güvenilirliğe ilişkin Cronbach Alpha katsayısı hesaplanmış ve puanlayıcı güvenilirliği incelenmiştir. Cronbach Alpha güvenilirlik katsayısı her bir ölçüt ile alt ölçütlerin tamamı için hesaplanmıştır. Kutlu vd. (2017, s. 82) tarafından dereceli puanlama anahtarı açısından ele alınan güvenilirlik kavramı, puanlamanın bir puanlayıcıdan diğerine değişmemesi, yani tutarlılık olarak tanımlanmaktadır. Geliştirilen bir ölçme aracıda puanların güvenilirliklerinin belirlenmesi için puanlayıcılar arası güvenilirlik katsayısının hesaplanması önerilmektedir (Moskal & Leydens, 2000). Bu çalışmada, puanlayıcılar arası güvenilirliğe kanıt sağlamak için 6 puanlayıcının değerlendirmelerindeki uyum hesaplanırken alt ölçütler likert yapıda ve ikiden fazla puanlayıcı olduğundan her bir alt ölçüte ilişkin Fleiss Kappa değeri; ana ölçütler, alt ölçütlerin toplanması ile elde edildiği için Sınıf İçi Korelasyon (Intra-Class Correlation) katsayısı kullanılmıştır. Dereceli puanlama anahtarının güvenilirliğine kanıt oluşturma bağlamında ayrıca madde-toplam puan korelasyonları da incelenmiştir. Dereceli puanlama anahtarındaki alt ölçütlerin, ölçütlerle ve toplam puanla olan korelasyonları değerlendirilmiştir. “Madde-toplam puan korelasyonu, test maddelerinden alınan puan ile testin toplam puanı arasındaki ilişkiyi açıklamaktadır.” (Büyüköztürk, 2018, s. 183).

Uluslararası öğrencilerin akademik yazma becerilerinin değerlendirilmesine yönelik olarak geliştirilen taslak dereceli puanlama anahtarının yapı geçerliğine ilişkin açılımlı faktör analizi (AFA), güvenilirliğine ilişkin Cronbach Alpha katsayısı, puanlayıcılar arası uyumu belirleyebilmek için Sınıf İçi Korelasyon, madde-toplam puan korelasyon katsayıları SPSS 29 istatistik paket programı ve Fleiss Kappa katsayısı R programı kullanılarak çözümlenmiştir.

Bulgular

Dereceli Puanlama Anahtarının Geçerlik Analizi Sonuçlarına İlişkin Bulgular

Bu bölümde, uluslararası öğrencilerin Türkçe akademik yazma becerilerine yönelik geliştirilen analitik dereceli puanlama anahtarından kapsam ve yapı geçerliği bağlamında elde edilen bulgular ayrı başlıklarda gösterilmiştir.

Dereceli Puanlama Anahtarının Kapsam Geçerliğine İlişkin Bulguları

Davis (1992) tekniği kullanılarak hesaplanan taslak dereceli puanlama anahtarının (DPA) kapsam geçerlik indekslerine

(KGİ) ilişkin bulgularının yer aldığı Tablo 2’de incelendiğinde taslak DPA’nın ana ölçütlerinin ve alt ölçütlerinin 0,80 ila 1 arasında kapsam geçerlik indeksine sahip olduğu görülmektedir. Bu oranlar taslak DPA’da yer alan tüm ölçüt ve alt ölçütlerin kapsam geçerliği bakımından yeterli düzeyde olduğunu göstermektedir. (Davis Sayısı $\geq 0,80$). Uzmanların görüşleri doğrultusunda; “Yazım ve Noktalama” alt ölçütü, “Yazım Kuralları” ve “Noktalama İşaretleri” olarak iki ayrı ölçüte ayrılmıştır. “Sayfa Düzeni” tek bir ölçütte birleştirilerek uluslararası öğrencilerden performans görevini sadece elektronik ortamda tamamlamaları istenmiştir. “Etkileşimsel Üstsöylem Belirleyicileri” ve “Etkileşimli Üstsöylem Belirleyicileri” alt ölçütleri, “Üstsöylem Belirleyicileri” ölçütünde bütünleştirilerek “Akademik Söylem” ölçütünün altında değerlendirilmiştir. “Araştırma Problemi ve Amacı” alt ölçütü, “Araştırmanın Problem Durumu ve Amacı” olarak; “Bilgileri Özetleme” alt ölçütü, “Bilgileri Sentezleme” olarak; “Sonuç ve Öneriler” alt ölçütü, “Tartışma, Sonuç ve Öneriler” olarak yeniden tanımlanmıştır. Ayrıca uzmanların önerileri doğrultusunda ölçüt tanımlamaları yeniden gözden geçirilmiştir. Düzenlenmiş taslak DPA; 1’i ölçme ve değerlendirme, 3’ü Türkçe eğitimi alanından olmak üzere 4 öğretim üyesinin görüşüne tekrar sunulmuştur. Öğretim üyelerinin değerlendirmeleri göz önüne alınarak taslak DPA’da düzenlemeler yapılmıştır.

Dereceli Puanlama Anahtarının Yapı Geçerliğine İlişkin Bulguları

Uluslararası öğrencilerin Türkçe akademik yazma becerilerine yönelik geliştirilen analitik dereceli puanlama anahtarına yapı geçerliği çalışması bağlamında açımlayıcı faktör analizi (AFA) uygulanmıştır. AFA, söz konusu aracın faktör yapısını ortaya koymak amacıyla 157 uluslararası öğrenciden elde edilen veriler (performans görevi) üzerinden yapılmıştır.

Tablo 2. Taslak DPA kapsam geçerlik indeksi

		A	B	C	D	KGİ	KGİ
Biçim	Dil Bilgisi	3	6	1	0	0,90	
	Yazım ve Noktalama	6	3	1	0	0,90	
	Anlaşılabilirlik	5	5	0	0	1	0,92
	Sayfa Düzeni (Bilgisayar)	7	2	1	0	0,90	
	Sayfa Düzeni (El yazısı)	5	4	0	1	0,90	
Akademik Söylem	Söz Varlığı	8	2	0	0	1	0,95
	Cümle Yapısı	8	1	1	0	0,90	
Metin Yapısı	Konu	7	3	0	0	1	
	Araştırma Problemi ve Amacı	9	1	0	0	1	
	Bağdaşıklık	9	1	0	0	1	
	Tutarlılık	9	1	0	0	1	0,99
	Paragraf Yapısı	8	2	0	0	1	
	Bilgileri Özetleme	7	2	1	0	0,90	
	Görsel Sunum	8	2	0	0	1	
Üstsöylem Belirleyicileri	Etkileşimsel Üstsöylem Belirleyicileri	7	3	0	0	1	1
	Etkileşimli Üstsöylem Belirleyicileri	8	2	0	0	1	
Plan ve Düzenleme	Başlık	7	3	0	0	1	
	Özet	7	3	0	0	1	
	Anahtar Sözcükler	7	3	0	0	1	
	Giriş	6	4	0	0	1	
	Başlıklandırma	7	2	1	0	0,90	0,94
	Gelişme	7	1	2	0	0,80	
	Sonuç ve Öneriler	8	1	1	0	0,90	
	Kaynakça	8	1	1	0	0,90	
Kaynaklardan Yararlanma	Doğrudan ve Dolaylı Alıntı	9	1	0	0	1	1
	Akademik Etik İlkeler	7	3	0	0	1	

Öncelikle veri setinin AFA’nın uygulanmasına uygun olup olmadığının belirlenmesi için çeşitli varsayımlar sınanmıştır. Bu kapsamda ilk olarak DPA’nın alt ölçütleri arasında çoklu bağlantı sorunu olup olmadığı kontrol edilmiş ve ardından alt ölçütlere ait normallik kontrolü için Basıklık ve Çarpıklık katsayıları hesaplanmıştır. Çoklu bağlantı durumunu sınamak için alt ölçütlerin birbiriyle olan ilişkileri Pearson Korelasyon katsayısı ile değerlendirilmiştir ve 0,90’dan daha yüksek hesaplanan bir korelasyon katsayısı bulunmadığından alt ölçütler arasında çoklu bağlantı sorununa rastlanmamıştır. Diğer yandan, alt ölçütlerin Basıklık ve Çarpıklık katsayılarının -2 ve +2 arasında olması normal dağılıma uygun olduğunu ifade etmektedir. “Görsel Sunum” alt ölçütünün Çarpıklık değeri 3,501; Basıklık değeri ise 11,759 olarak hesaplandığından normal dağılıma uygunluk göstermemektedir. Bu nedenle “Görsel Sunum” alt ölçütünün, AFA yapılmasına uygun olmadığı sonucuna varılmış ve ölçme aracından çıkarılmıştır. DPA’yı oluşturan her bir ölçüte ilişkin yapılan AFA’ya yönelik bulgular ayrı başlıklar hâlinde sunulmuştur.

Biçim Ölçütü İçin Açımlayıcı Faktör Analizi Bulguları

AFA yapılmadan önce “Biçim” ölçütü için örneklem büyüklüğünün yeterliğini değerlendirmek amacıyla Kaiser-Meyer-Olkin (KMO) ve alt ölçütler arasındaki ilişkinin uygunluğu için Bartlett küresellik testi uygulanmıştır. “Biçim” ölçütü için KMO değeri 0,762 olarak hesaplanmış ve bu değer veri yapısının faktör analizi için yeterli olduğu sonucuna ulaşılmıştır. Bartlett küresellik testi sonuçları incelendiğinde ise ki-kare değeri ($X^2=517,818$, $p<0,001$) istatistiksel olarak anlamlı bulunmuş ve dolayısıyla verilerin çok değişkenli normal dağılıma uygun olduğu tespit edilmiştir.

Tablo 3. Biçim ölçütünde yer alan alt ölçütlere ait faktör yüklerine ilişkin bulgular

	Faktör Yüğü (Communalities)	
	Başlangıç	Ekstraksiyon
Dil Bilgisi	1,000	0,677
Noktalama İşaretleri	1,000	0,800
Yazım Kuralları	1,000	0,814
Anlaşılabilirlik	1,000	0,711
Sayfa Düzeni	1,000	0,327

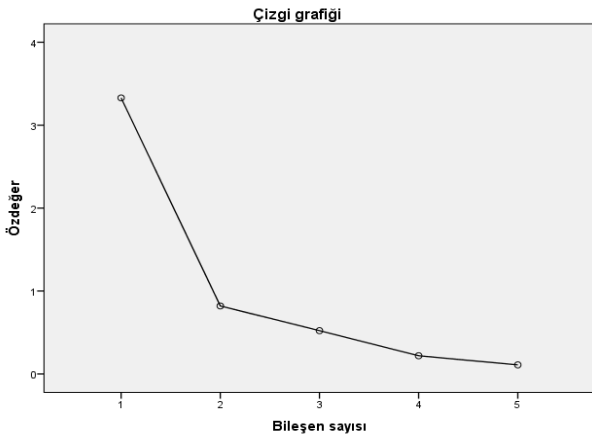
Ekstraksiyon Yöntemi: Temel Bileşenler Analizi

Tablo 3'te "Biçim" ölçütünde yer alan alt ölçütlere ait faktör yüklerinin 0,30'dan yüksek olduğu görülmektedir. Söz konusu alt ölçütlerin "Biçim" ölçütü için AFA'da kullanılabileceği söylenebilir.

Tablo 4. Biçim ölçütü için AFA sonucunda faktörlere ilişkin bulgular

Bileşen	Açıklanan Toplam Varyans			
	Başlangıç	Özdeğerleri	Kare Yüklerin Çıkarma Toplamları	
	Toplam Varyans	Kümülatif	Toplam Varyans	Kümülatif
	%	%	%	%
1	3,329	66,578	66,578	3,329
2	0,820	16,402	82,980	66,578
3	0,522	10,440	93,420	
4	0,219	4,387	97,807	
5	0,110	2,193	100,000	

Ekstraksiyon Yöntemi: Temel Bileşenler Analizi

**Şekil 1.** Biçim ölçütü için özdeğer bileşen çizgi grafiği

"Biçim" ölçütüne ilişkin AFA sonucunda elden edilen faktörlere ilişkin bulguların gösterildiği Tablo 4 incelendiğinde başlangıç özdeğerleri 1'den büyük olan sadece bir bileşen olduğu görülmüş ve bu bileşen ile "Biçim" ölçütüne ait varyansın toplam %66,578'inin açıklandığı gözlenmiştir.

"Biçim" ölçütü için özdeğer bileşen çizgi grafiği incelendiğinde (Şekil 1) 1. Bileşene ait özdeğerin 1'den yüksek olduğu, diğer bileşenlere ait özdeğerlerin ise 1'in altında olduğu görülmüş, bu alt ölçütler ile "Biçim" ölçütünün tek bir faktör ile açıklandığı belirlenmiştir.

Tablo 5 incelendiğinde "Biçim" ölçütünü oluşturan alt ölçütlerin tek bir bileşen ile açıklandığı görülmektedir. Bu nedenle bir döndürme yöntemi kullanılarak rotasyon matrisinin oluşturulmasına gereksinim duyulmamıştır.

Tablo 5. Biçim ölçütüne ait bileşen matrisine ilişkin bulgular

	Bileşen Matrisi	
	Bileşen 1	
Dil Bilgisi	0,823	
Noktalama İşaretleri	0,894	
Yazım Kuralları	0,902	
Anlaşılabilirlik	0,843	
Sayfa Düzeni	0,572	

Ekstraksiyon Yöntemi: Temel Bileşenler Analizi

Akademik Söylem Ölçütü İçin Açıklayıcı Faktör Analizi Bulguları

"Akademik Söylem" ölçütü için AFA yapılmadan önce örneklem büyüklüğünün yeterliğini değerlendirmek amacıyla Kaiser-Meyer-Olkin (KMO) ve alt ölçütler arasındaki ilişkinin uygunluğu için Bartlett küresellik testi uygulanmıştır. "Akademik Söylem" ölçütü için KMO değeri 0,690 olarak hesaplanmış ve bu değer veri yapısının faktör analizi için yeterli olduğu sonucuna ulaşılmıştır. Bartlett küresellik testi sonuçları incelendiğinde ise ki-kare değeri ($X^2=247,616$, $p<0,001$) istatistiksel olarak anlamlı bulunmuş ve dolayısıyla verilerin çok değişkenli normal dağılıma uygun olduğu tespit edilmiştir.

Tablo 6. Akademik söylem ölçütünde yer alan alt ölçütlere ait faktör yüklerine ilişkin bulgular

	Faktör Yüğü (Communalities)	
	Başlangıç	Ekstraksiyon
Söz Varlığı	1,000	0,857
Cümle Yapısı	1,000	0,822
Üstsöylem Belirleyicileri	1,000	0,679

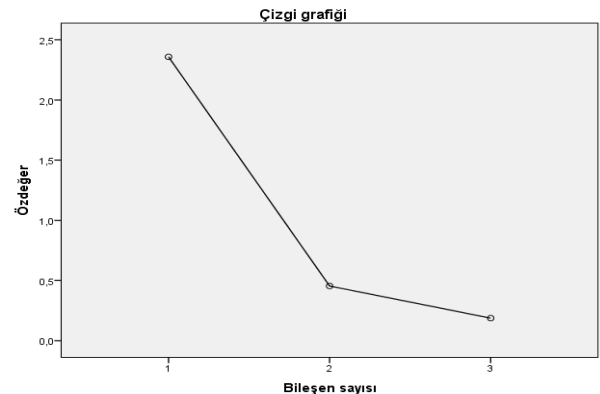
Ekstraksiyon Yöntemi: Temel Bileşenler Analizi

"Akademik Söylem" ölçütünde yer alan alt ölçütlere ait faktör yüklerinin 0,30'dan yüksek olduğu ve bu alt ölçütlerin "Akademik Söylem" ölçütü için AFA'da kullanılabileceği görülmüştür (Tablo 6).

Tablo 7. Akademik söylem ölçütü için AFA sonucunda faktörlere ilişkin bulgular

Bileşen	Açıklanan Toplam Varyans			
	Başlangıç	Özdeğerleri	Kare Yüklerin Çıkarma Toplamları	
	Toplam	Varyans Kümülatif	Toplam	Varyans Kümülatif
	%	%	%	%
1	2,358	78,613	78,613	2,358
2	0,454	15,145	93,758	
3	0,187	6,242	100,000	

Ekstraksiyon Yöntemi: Temel Bileşenler Analizi

**Şekil 2.** Akademik söylem ölçütü için özdeğer bileşen çizgi grafiği

“Akademik Söylem” ölçütüne ilişkin AFA sonucunda elden edilen faktörlere ilişkin bulguların gösterildiği Tablo 7 incelendiğinde başlangıç özdeğerleri 1’den büyük olan sadece bir bileşen olduğu görülmüş ve bu bileşen ile “Akademik Söylem” ölçütüne ait varyansın toplam %78,613’ünün açıklandığı gözlenmiştir.

“Akademik Söylem” ölçütü için özdeğer bileşen çizgi grafiği incelendiğinde (Şekil 2) 1. Bileşene ait özdeğerin 1’den yüksek olduğu, diğer bileşenlere ait özdeğerlerin ise 1’in altında olduğu görülmüş, bu alt ölçütler ile “Akademik Söylem” ölçütünün tek bir faktör ile açıklandığı tespit edilmiştir.

Tablo 8. Akademik söylem ölçütüne ait bileşen matrisine ilişkin bulgular

Bileşen Matrisi	
	Bileşen
	1
Söz Varlığı	0,926
Cümle Yapısı	0,906
Üstsöylem Belirleyicileri	0,824
Ekstraksiyon Yöntemi: Temel Bileşenler Analizi	

Tablo 8 incelendiğinde “Akademik Söylem” ölçütünü oluşturan alt ölçütlerin tek bir bileşen ile açıklandığı görülmektedir. Bu nedenle bir döndürme yöntemi kullanılarak rotasyon matrisinin oluşturulmasına gereksinim duyulmamıştır.

Metin Yapısı Ölçütü İçin Açıklayıcı Faktör Analizi Bulguları

“Metin Yapısı” ölçütü için AFA yapılmadan önce örneklem büyüklüğünün yeterliğini değerlendirmek amacıyla Kaiser-Meyer-Olkin (KMO) ve alt ölçütler arasındaki ilişkinin uygunluğu için Bartlett küresellik testi uygulanmıştır. “Metin Yapısı” ölçütü için KMO değeri 0,868 olarak hesaplanmış ve bu değer veri yapısının faktör analizi için yeterli olduğu sonucuna ulaşılmıştır. Bartlett küresellik testi sonuçları incelendiğinde ise ki-kare değeri ($X^2=579,990$, $p<0,001$) istatistiksel olarak anlamlı bulunmuş ve dolayısıyla verilerin çok değişkenli normal dağılıma uygun olduğu tespit edilmiştir.

Tablo 9. Metin yapısı ölçütünde yer alan alt ölçütlere ait faktör yüklerine ilişkin bulgular

Faktör Yüğü (Communalities)			
	Başlangıç	Ekstraksiyon	
Konu	1,000	0,728	
Araştırmanın Problem Durumu ve Amacı	1,000	0,363	
Bağdaşıklık	1,000	0,656	
Tutarlılık	1,000	0,786	
Paragraf Yapısı	1,000	0,723	
Bilgileri Sentezleme	1,000	0,738	
Ekstraksiyon Yöntemi: Temel Bileşenler Analizi			

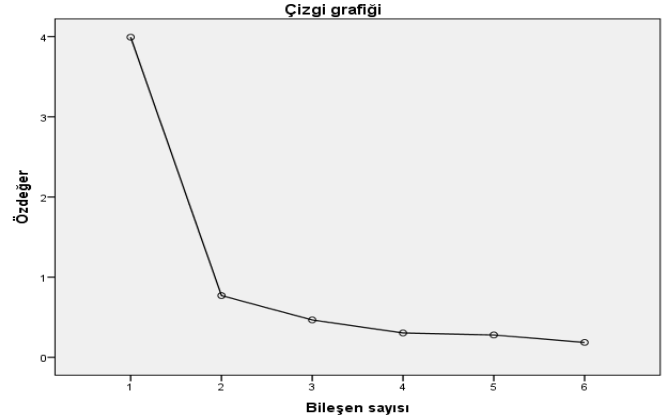
“Metin Yapısı” ölçütünde yer alan alt ölçütlere ait faktör yüklerinin 0,30’dan yüksek olduğu ve bu alt ölçütlerin “Metin Yapısı” ölçütü için AFA’da kullanılabilceği görülmüştür (Tablo 9).

“Metin Yapısı” ölçütüne ilişkin AFA sonucunda elden edilen faktörlere ilişkin bulguların gösterildiği Tablo 10 incelendiğinde başlangıç özdeğerleri 1’den büyük olan sadece bir bileşen olduğu görülmüş ve bu bileşen ile “Metin Yapısı” ölçütüne ait varyansın toplam %66,560’ını açıklandığı gözlenmiştir.

Tablo 10. Metin yapısı ölçütü için AFA sonucunda faktörlere ilişkin bulgular

Açıklanan Toplam Varyans					
Bileşen	Başlangıç	Özdeğerleri	Kare Yüklerin	Çıkarma	Toplamları
	Toplam Varyans	Kümülatif	Toplam Varyans	Kümülatif	
	%	%	%	%	
1	3,994	66,560	66,560	3,994	66,560
2	0,770	12,842	79,402		
3	0,467	7,778	87,179		
4	0,304	5,070	92,250		
5	0,279	4,651	96,900		
6	0,186	3,100	100,000		

Ekstraksiyon Yöntemi: Temel Bileşenler Analizi



Şekil 3. Metin yapısı ölçütü için özdeğer bileşen çizgi grafiği

“Metin Yapısı” ölçütü için özdeğer bileşen çizgi grafiği incelendiğinde (Şekil 3) 1. Bileşene ait özdeğerin 1’den yüksek olduğu, diğer bileşenlere ait özdeğerlerin ise 1’in altında olduğu görülmüş, bu alt ölçütler ile “Metin Yapısı” ölçütünün tek bir faktör ile açıklandığı tespit edilmiştir.

Tablo 11. Metin yapısı ölçütüne ait bileşen matrisine ilişkin bulgular

Bileşen Matrisi	
	Bileşen
	1
Konu	0,853
Araştırmanın Problem Durumu ve Amacı	0,602
Bağdaşıklık	0,810
Tutarlılık	0,887
Paragraf Yapısı	0,850
Bilgileri Sentezleme	0,859
Ekstraksiyon Yöntemi: Temel Bileşenler Analizi	

Tablo 11 incelendiğinde “Metin Yapısı” ölçütünü oluşturan alt ölçütlerin tek bir bileşen ile açıklandığı görülmektedir. Bu nedenle bir döndürme yöntemi kullanılarak rotasyon matrisinin oluşturulmasına gereksinim duyulmamıştır.

Plan ve Düzenleme Ölçütü İçin Açıklayıcı Faktör Analizi Bulguları

“Plan ve Düzenleme” ölçütü için AFA yapılmadan önce örneklem büyüklüğünün yeterliğini değerlendirmek amacıyla Kaiser-Meyer-Olkin (KMO) ve alt ölçütler arasındaki ilişkinin uygunluğu için Bartlett küresellik testi uygulanmıştır. “Plan ve Düzenleme” ölçütü için KMO değeri 0,878 olarak hesaplanmış ve bu değer veri yapısının faktör analizi için

yeterli olduğu sonucuna ulaşılmıştır. Bartlett küresellik testi sonuçları incelendiğinde ise ki-kare değeri ($X^2=555,924$, $p<0,001$) istatistiksel olarak anlamlı bulunmuş ve dolayısıyla verilerin çok değişkenli normal dağılıma uygun olduğu tespit edilmiştir.

Tablo 12. Plan ve düzenleme ölçütünde yer alan alt ölçütlere ait faktör yüklerine ilişkin bulgular

	Faktör Yüğü (Communalities)	
	Başlangıç	Ekstraksiyon
Başlık	1,000	0,507
Özet	1,000	0,652
Anahtar Sözcükler	1,000	0,516
Giriş	1,000	0,639
Başlıklandırma	1,000	0,667
Gelişme	1,000	0,661
Tartışma, Sonuç ve Öneriler	1,000	0,634

Ekstraksiyon Yöntemi: Temel Bileşenler Analizi

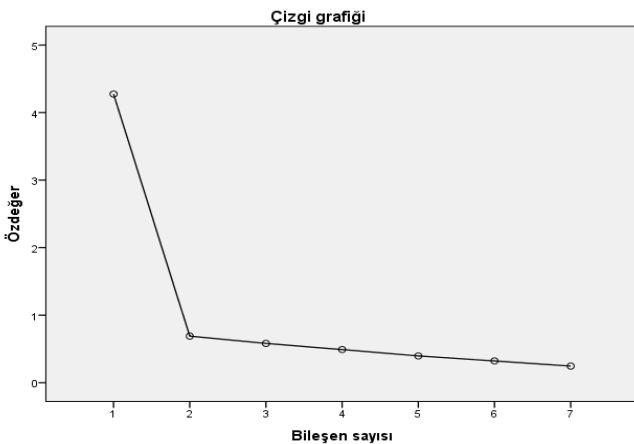
“Plan ve Düzenleme” ölçütünde yer alan alt ölçütlere ait faktör yüklerinin 0,30’dan yüksek olduğu ve bu alt ölçütlerin “Plan ve Düzenleme” ölçütü için AFA’da kullanılabileceği görülmüştür (Tablo 12).

Tablo 13. Plan ve düzenleme ölçütü için AFA sonucunda faktörlere ilişkin bulgular

Bileşen	Açıklanan Toplam Varyans					
	Başlangıç Özdeğerleri			Kare Yüklerin Çıkarma Toplamları		
	Toplam	Varyans %	Kümülatif %	Toplam	Varyans %	Kümülatif %
1	4,275	61,067	61,067	4,275	61,067	61,067
2	0,690	9,857	70,923			
3	0,581	8,303	79,226			
4	0,491	7,012	86,239			
5	0,396	5,654	91,893			
6	0,322	4,604	96,497			
7	0,245	3,503	100,000			

Ekstraksiyon Yöntemi: Temel Bileşenler Analizi

“Plan ve Düzenleme” ölçütüne ilişkin AFA sonucunda elden edilen faktörlere ilişkin bulguların gösterildiği Tablo 13 incelendiğinde başlangıç özdeğerleri 1’den büyük olan sadece bir bileşen olduğu görülmüş ve bu bileşen ile “Plan ve Düzenleme” ölçütüne ait varyansın toplam %61,067’sinin açıklandığı gözlenmiştir.



Şekil 4. Plan ve düzenleme ölçütü için özdeğer bileşen çizgi grafiği

“Plan ve Düzenleme” ölçütü için özdeğer bileşen çizgi grafiği incelendiğinde (Şekil 4) 1. Bileşene ait özdeğerin 1’den

yüksek olduğu, diğer bileşenlere ait özdeğerlerin ise 1’in altında olduğu görülmüş, bu alt ölçütler ile “Plan ve Düzenleme” ölçütünün tek bir faktör ile açıklandığı tespit edilmiştir.

Tablo 14. Plan ve düzenleme ölçütüne ait bileşen matrisine ilişkin bulgular

	Bileşen Matrisi	
	Bileşen 1	
Başlık	0,712	
Özet	0,807	
Anahtar Sözcükler	0,718	
Giriş	0,800	
Başlıklandırma	0,817	
Gelişme	0,813	
Tartışma, Sonuç ve Öneriler	0,796	

Ekstraksiyon Yöntemi: Temel Bileşenler Analizi

Tablo 14 incelendiğinde “Plan ve Düzenleme” ölçütünü oluşturan alt ölçütlerin tek bir bileşen ile açıklandığı görülmektedir. Bu nedenle bir döndürme yöntemi kullanılarak rotasyon matrisinin oluşturulmasına gereksinim duyulmamıştır.

Kaynaklardan Yararlanma Ölçütü İçin Açıklayıcı Faktör Analizi Bulguları

“Kaynaklardan Yararlanma” ölçütü için AFA yapılmadan önce örneklem büyüklüğünün yeterliliğini değerlendirmek amacıyla Kaiser-Meyer-Olkin (KMO) ve alt ölçütler arasındaki ilişkinin uygunluğu için Bartlett küresellik testi uygulanmıştır. “Kaynaklardan Yararlanma” ölçütü için KMO değeri 0,724 olarak hesaplanmış ve bu değer veri yapısının faktör analizi için yeterli olduğu sonucuna ulaşılmıştır. Bartlett küresellik testi sonuçları incelendiğinde ise ki-kare değeri ($X^2=232,163$, $p<0,05$) istatistiksel olarak anlamlı bulunmuş ve dolayısıyla verilerin çok değişkenli normal dağılıma uygun olduğu tespit edilmiştir.

Tablo 15. Kaynaklardan yararlanma ölçütünde yer alan alt ölçütlere ait faktör yüklerine ilişkin bulgular

	Faktör Yüğü (Communalities)	
	Başlangıç	Ekstraksiyon
Kaynakça	1,000	0,804
Doğrudan ve Dolaylı Alıntı	1,000	0,831
Akademik Etik İlkeler	1,000	0,738

Ekstraksiyon Yöntemi: Temel Bileşenler Analizi

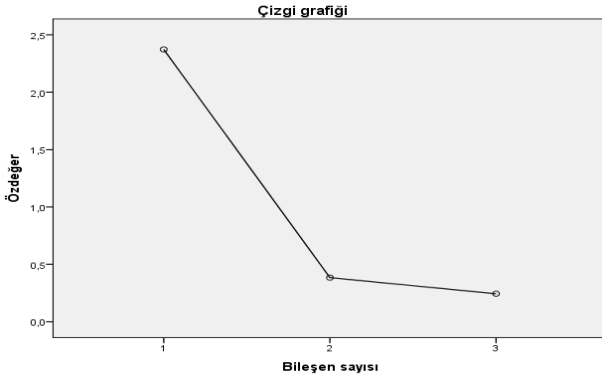
“Kaynaklardan Yararlanma” ölçütünde yer alan alt ölçütlere ait faktör yüklerinin 0,30’dan yüksek olduğu ve bu alt ölçütlerin “Kaynaklardan Yararlanma” ölçütü için AFA’da kullanılabileceği görülmüştür (Tablo 15).

Tablo 16. Kaynaklardan yararlanma ölçütü için AFA sonucunda faktörlere ilişkin bulgular

Bileşen	Açıklanan Toplam Varyans					
	Başlangıç Özdeğerleri			Kare Yüklerin Çıkarma Toplamları		
	Toplam	Varyans %	Kümülatif %	Toplam	Varyans %	Kümülatif %
1	2,373	79,085	79,085	2,373	79,085	79,085
2	0,384	12,800	91,884			
3	0,243	8,116	100,000			

Ekstraksiyon Yöntemi: Temel Bileşenler Analizi

“Kaynaklardan Yararlanma” ölçütüne ilişkin AFA sonucunda elden edilen faktörlere ilişkin bulguların gösterildiği Tablo 16 incelendiğinde başlangıç özdeğerleri 1’den büyük olan sadece bir bileşen olduğu görülmüş ve bu bileşen ile “Kaynaklardan Yararlanma” ölçütüne ait varyansın toplam %79,085’inin açıklandığı gözlenmiştir.



Şekil 5. Kaynaklardan yararlanma ölçütü için özdeğer bileşen çizgi grafiği

“Kaynaklardan Yararlanma” ölçütü için özdeğer bileşen çizgi grafiği incelendiğinde (Şekil 5) 1. Bileşene ait özdeğerin 1’den yüksek olduğu, diğer bileşenlere ait özdeğerlerin ise 1’in altında olduğu görülmüş, bu alt ölçütler ile “Plan ve Düzenleme” ölçütünün tek bir faktör ile açıklandığı tespit edilmiştir.

Tablo 17. Kaynaklardan yararlanma ölçütüne ait bileşen matrisine ilişkin bulgular

Bileşen	Bileşen Matrisi	
	Bileşen	
	1	
Kaynakça	0,897	
Doğrudan ve Dolaylı Alıntı	0,912	
Akademik Etik İlkeler	0,859	

Ekstraksiyon Yöntemi: Temel Bileşenler Analizi

Tablo 17 incelendiğinde “Plan ve Düzenleme” ölçütünü oluşturan alt ölçütlerin tek bir bileşen ile açıklandığı görülmektedir. Bu nedenle bir döndürme yöntemi kullanılarak rotasyon matrisinin oluşturulmasına gereksinim duyulmamıştır.

DPA’nın “Plan ve Düzenleme” ölçütü faktör analizi öncesinde “Başlık”, “Özet”, “Anahtar sözcükler”, “Giriş”, “Başlıklandırma”, “Gelişme”, “Tartışma, Sonuç ve Öneriler”, “Kaynakça” alt ölçütlerinden oluşmaktayken; faktör analizi sonrasında “Kaynakça” alt ölçütü “Plan ve Düzenleme” ölçütünden çıkartılarak “Kaynaklardan Yararlanma” ölçütü kapsamına dâhil edilmiştir. Bu sayede “Plan ve Düzenleme” ana ölçütünün açıklanan varyansı 58,806’dan 61,067’ye yükselmiş; “Kaynaklardan Yararlanma” ölçütünün KMO değeri ise 0,500’den 0,720’ye çıkmıştır. Böylelikle “Kaynaklardan Yararlanma” ölçütü “Doğrudan ve Dolaylı Alıntı” ve “Akademik Etik İlkeler” alt ölçütleri ile faktör analizi uygulanabilir yeterli örneklem genişliği varsayımlarını sağlamıştır.

Dereceli Puanlama Anahtarının Güvenirlik Analizi Sonuçlarına İlişkin Bulguları

Bu bölümde uluslararası öğrencilerin Türkçe akademik yazma becerilerine yönelik geliştirilen analitik dereceli puanlama anahtarından; Cronbach Alfa güvenirlik katsayı, puanlayıcılar arası uyum ve madde geçerliğine kanıt olarak hesaplanan

madde-toplam korelasyonu bağlamında elde edilen bulgular ayrı başlıklarda gösterilmiştir.

Dereceli Puanlama Anahtarının Cronbach Alfa Güvenirlik Katsayına İlişkin Bulguları

Uluslararası öğrencilerin Türkçe akademik yazma becerilerine yönelik geliştirilen analitik dereceli puanlama anahtarının güvenirlik çalışması bağlamında iç tutarlılığa dayalı yöntemlerden biri olan Cronbach Alfa güvenirlik katsayı hesaplanmıştır. 5 ana ölçüt ve tüm alt ölçütlerin toplamı için AFA öncesi ve sonrası elde edilen çözümlene sonuçları bütüncül olarak değerlendirildiğinde Cronbach Alpha güvenirlik katsayısının “Biçim” ölçütü için 0,851; “Akademik Söylem” ölçütü için 0,863; AFA’dan önceki “Metin Yapısı” ölçütü için 0,846 iken, AFA’dan sonra 0,891; AFA’dan önceki “Plan ve Düzenleme ölçütü” için 0,892 iken, AFA’dan sonra 0,885; AFA’dan önceki “Kaynaklardan Yararlanma” ölçütü için 0,797 iken, AFA’dan sonra 0,867 olduğu belirlenmiştir. Sözü edilen ölçme aracının toplam Cronbach Alpha güvenirlik katsayısı AFA’dan önce 0,953 iken, AFA’dan sonra 0,956 olarak hesaplanmıştır. Bu hesaplamalar sonucunda DPA’nın beş ana ölçütü ve tüm alt ölçütlerin toplamı için güvenirliğinin iyi ve mükemmel düzeyde olduğu söylenebilir.

Puanlayıcılar Arasındaki Uyuma İlişkin Bulgular

Puanlayıcılar arasındaki uyum için 6 puanlayıcının yaptığı puanlamalar üzerinden alt ölçütlerin her bir için Fleiss Kappa katsayısı (Tablo 18); ana ölçütler için Sınıf İçi Korelasyon katsayısı (Tablo 19) hesaplanmıştır.

Tablo 18. Alt ölçütlerin puanlayıcılar arası uyumuna ilişkin bulguları

Alt ölçütler	Fleiss Kappa	p
Dil Bilgisi	0,752	<0,001
Noktalama İşaretleri	0,707	<0,001
Yazım Kuralları	0,680	<0,001
Anlaşılabilirlik	0,637	<0,001
Sayfa Düzeni	0,816	<0,001
Söz Varlığı	0,700	<0,001
Cümle Yapısı	0,675	<0,001
Üstsöylem Belirleyicileri	0,788	<0,001
Konu	0,645	<0,001
Araştırmanın Problem Durumu ve Amacı	0,718	<0,001
Bağdaşlılık	0,768	<0,001
Tutarlılık	0,761	<0,001
Paragraf Yapısı	0,660	<0,001
Bilgileri Sentezleme	0,819	<0,001
Görsel Sunum	0,871	<0,001
Başlık	0,710	<0,001
Özet	0,837	<0,001
Anahtar Sözcükler	0,909	<0,001
Giriş	0,661	<0,001
Başlıklandırma	0,700	<0,001
Gelişme	0,728	<0,001
Tartışma, Sonuç ve Öneriler	0,832	<0,001
Kaynakça	0,765	<0,001
Doğrudan ve Dolaylı Alıntı	0,679	<0,001
Akademik Etik İlkeler	0,822	<0,001

Tablo 18’e göre alt ölçütler için puanlayıcılar arası uyum düzeyleri; en düşük “Anlaşılabilirlik”, en yüksek “Anahtar Sözcükler” olmak üzere 0,637-0,909 arasında değişim göstermektedir. Alt ölçütlere ait bütün uyum katsayıları istatistiksel olarak anlamlı olarak tespit edilmiştir (p<0,05). Hesaplanan Fleiss Kappa değerleri, puanlayıcılar arası uyumun yüksek ve çok yüksek düzeyde olduğunu göstermektedir.

Tablo 19. Ölçütlerin puanlayıcılar arası uyumuna ilişkin bulguları

		Sınıf Korelasyon	İçî %95 Güven Aralığı		F Testi			p
			Alt limit	Üst limit	Değer	Sd1	Sd2	
Biçim	Tek Ölçüm	0,909	0,845	0,954	71,429	24	120	<0,001
	Ortalama Ölçümler	0,984	0,970	0,992	71,429	24	120	<0,001
Akademik Söylem	Tek Ölçüm	0,932	0,886	0,965	83,052	24	120	<0,001
	Ortalama Ölçümler	0,988	0,979	0,994	83,052	24	120	<0,001
Metin Yapısı	Tek Ölçüm	0,965*	0,939*	0,982*	180,990*	24	120	<0,001
	Ortalama Ölçümler	0,966	0,941	0,983	190,376	24	120	<0,001
Plan ve Düzenleme	Tek Ölçüm	0,994*	0,989*	0,997*	180,990*	24	120	<0,001
	Ortalama Ölçümler	0,994	0,990	0,997	190,376	24	120	<0,001
Kaynaklardan Yararlanma	Tek Ölçüm	0,985*	0,973*	0,992*	402,285*	24	120	<0,001
	Ortalama Ölçümler	0,980	0,966	0,990	301,546	24	120	<0,001
Toplam	Tek Ölçüm	0,997*	0,995*	0,999*	402,285*	24	120	<0,001
	Ortalama Ölçümler	0,997	0,994	0,998	301,546	24	120	<0,001
Toplam	Tek Ölçüm	0,940*	0,898*	0,969*	97,331*	24	120	<0,001
	Ortalama Ölçümler	0,960	0,931	0,980	157,284	24	120	<0,001
Toplam	Tek Ölçüm	0,989*	0,981*	0,995*	97,331*	24	120	<0,001
	Ortalama Ölçümler	0,993	0,988	0,997	157,284	24	120	<0,001
Toplam	Tek Ölçüm	0,993*	0,988*	0,997*	903,709*	24	120	<0,001
	Ortalama Ölçümler	0,993	0,988	0,997	926,871	24	120	<0,001
Toplam	Tek Ölçüm	0,999*	0,998*	0,999*	930,709*	24	120	<0,001
	Ortalama Ölçümler	0,999	0,998	0,999	926,871	24	120	<0,001

*Faktör analizi öncesi

Tablo 19'a göre puanlayıcılar arası uyum düzeyleri tek ölçümlerde en düşük "Biçim" ölçütüne ait iken, en yüksek "Plan ve Düzenleme" olmak üzere 0,909-0,980 arasında değişim göstermektedir. Puanlayıcılar arası uyum düzeyleri ortalama ölçümlerde en düşük "Biçim" ölçütüne ait iken, en yüksek "Plan ve Düzenleme" olmak üzere 0,984-0,997 arasında değişim göstermektedir. Ölçütlerin toplamından elde

edilen "Toplam" puan değeri ise tek ölçümde 0,993 hesaplanırken, ortalama ölçümlerde 0,999 olarak hesaplanmıştır. Ölçütlere ve toplam puana ait bütün tek ve ortalama ölçüm uyum katsayıları istatistiksel olarak anlamlı olarak tespit edilmiştir ($p < 0,05$). Tüm ölçütler için hesaplanan Sınıf İçî Korelasyon değerleri tek ve ortalama ölçümlerde mükemmel uyum göstermektedir.

Tablo 20. Madde-toplam korelasyon katsayılarına ilişkin bulgular

	Biçim	Akademik Söylem	Metin Yapısı	Plan ve Düzenleme	Kaynaklardan Yararlanma	Toplam Puan
Dil Bilgisi	0,801					0,568
Noktalama İşaretleri	0,867					0,638
Yazım Kuralları	0,874					0,574
Anlaşılabilirlik	0,829					0,621
Sayfa Düzeni	0,666					0,716
Söz Varlığı		0,921				0,785
Cümle Yapısı		0,903				0,710
Üstsöylem Belirleyicileri		0,832				0,782
Konu			0,858			0,817
Araştırmanın Problem Durumu ve Amacı			0,662			0,659
Bağdaşıklık			0,787			0,698
Tutarlılık			0,867			0,759
Paragraf Yapısı			0,834			0,752
Bilgileri Sentezleme			0,856			0,825
Başlık				0,705		0,644
Özet				0,823		0,649
Anahtar Sözcükler				0,766		0,613
Giriş				0,778		0,677
Başlıklandırma				0,808		0,714
Gelişme				0,797		0,833
Tartışma, Sonuç ve Öneriler				0,774		0,750
Kaynakça					0,899	0,628
Doğrudan ve Dolaylı Alıntı					0,916	0,689
Akademik Etik İlkeler					0,851	0,701

Dereceli Puanlama Anahtarındaki Alt Ölçütlerin Madde Geçerlik Düzeylerine İlişkin Bulgular

Uluslararası öğrencilerin Türkçe akademik yazma becerilerine yönelik geliştirilen DPA’da alt ölçütlerin, ilgili olduğu ölçüt ve toplam puanla uyumlu olup olmadığını belirlemek amacıyla madde geçerlik düzeylerine ilişkin Pearson Korelasyon katsayısı kullanılarak hesaplanan bulguların yer aldığı Tablo 20 incelendiğinde “Biçim” ölçütündeki alt ölçütlerin madde-toplam korelasyon değerlerinin 0,666 ile 0,874 arasında; “Akademik Söylem” ölçütündeki alt ölçütlerin madde-toplam korelasyon değerlerinin 0,832 ile 0,921 arasında; “Metin Yapısı” ölçütündeki alt ölçütlerin madde-toplam korelasyon değerlerinin 0,662 ile 0,867 arasında; “Plan ve Düzenleme” ölçütündeki alt ölçütlerin madde-toplam korelasyon değerlerinin 0,705 ile 0,823 arasında; “Kaynaklardan Yararlanma” ölçütündeki alt ölçütlerin madde-toplam korelasyon değerlerinin 0,851 ile 0,916 arasında değiştiği görülmektedir. Ayrıca alt ölçütler ile toplam puan arasındaki madde-toplam korelasyon değerleri 0,568 ile 0,833 arasında değişmektedir. Madde-toplam korelasyon değerlerinin tamamı 0,05 düzeyinde anlamlıdır. Hesaplamalar, bütün madde-toplam korelasyon değerlerinin 0,30’dan yüksek ve pozitif olduğunu ortaya koymaktadır. Bu bulgular, DPA’daki alt ölçütlerin, ölçütlerle ve toplam puanla tutarlı olduğunu, öğrencileri iyi derecede ayırt ettiğini ve aynı davranışı ölçen maddeler olduğunu göstermektedir.

Tartışma, Sonuç ve Öneriler

Uluslararası öğrencilerin Türkçe akademik yazma becerilerine yönelik geliştirilen dereceli puanlama anahtarının geçerliğinin sınanması bağlamında kapsam ve yapı geçerliği değerlendirilmiştir. Kapsam geçerliği için taslak DPA’nın 6 ana ölçüt ve 26 alt ölçütten oluşan ilk hâli; alt ölçütlerin, ölçütlere uygunluğunu belirlemek ve düzey tanımlarının anlaşılır olup olmadığını ortaya koymak için 10 uzmandan görüş alınmıştır. Uzmanlardan gelen yanıtlara göre ölçüt ve alt ölçütlerin kapsam geçerlik indeksleri Davis (1992) tekniği kullanılarak belirlenmiştir. Alanyazında Lawshe (1975) tekniği ve Davis (1992) tekniği, kapsam geçerlik indeksi (KGI) kullanılarak ölçme aracı geliştirme sürecinde en yaygın olarak kullanılan iki teknik arasında yer almaktadır. Lynn (1986) kapsam geçerlik indeksini belirlemek için alınan uzman görüşlerinde; tarafsız ve kararsız bir orta noktadan kaçınılması gerektiği için en uygun dereceleminin 4’lü olduğunu ve uzman sayısının da en az 3 olması gerektiğini; fakat 10 uzmandan fazlasının da gereksiz olduğunu belirtmektedir. Çalışma kapsamında geliştirilen DPA’nın ana ölçütlerinin ve alt ölçütlerinin 0,80 ila 1 arasında kapsam geçerlik indeksine sahip olduğu görülmektedir. Bu oranlar, taslak DPA’da yer alan tüm ölçüt ve alt ölçütlerin kapsam geçerliği bakımından yeterli düzeyde olduğunu göstermektedir. (Davis Sayısı \geq 0,8). Bunun yanı sıra uzmanların görüşleri doğrultusunda DPA’nın ölçüt ve alt ölçütleri gözden geçirilmiştir. Bu bağlamda taslak DPA, 5 ana ölçüt (biçim, akademik söylem, metin yapısı, plan ve düzenleme, kaynaklardan yararlanma) ve 25 alt ölçüt (dil bilgisi, noktalama işaretleri, yazım kuralları, anlaşılabilirlik, sayfa düzeni, söz varlığı, cümle yapısı, üstsöylem belirleyicileri, konu, araştırmanın problem durumu ve amacı, bağdaşıklık, tutarlılık, paragraf yapısı, bilgileri sentezleme, görsel sunum, başlık, özet, anahtar sözcükler, giriş, başlıklandırma, gelişme, tartışma, sonuç ve öneriler, kaynakça, doğrudan ve dolaylı alıntı, akademik etik ilkeler) olmak üzere yeniden

düzenlenmiştir. Alanyazında farklı hedef kitleler ve öğrenme alanlarına ilişkin geliştirilen ölçme araçlarının kapsam geçerlik indekslerinin belirlenmesi için Davis (1992) tekniğinin kullanıldığı tespit edilmiştir (Bozkurt ve Arıca-Akkök, 2019; Arıcı, 2022; İltar ve Karataş, 2022). Söz konusu çalışmalarda kullanılan teknik ile bu çalışmada kapsam geçerliğinin belirlenmesi amacıyla kullanılan teknik koşutluk göstermektedir.

Uluslararası öğrencilerin Türkçe akademik yazma becerilerine yönelik geliştirilen analitik dereceli puanlama anahtarına yapı geçerliği çalışması bağlamında 157 uluslararası öğrenciden elde edilen veriler (performans görevi) üzerinden açımlayıcı faktör analizi (AFA) uygulanmıştır. Bu bağlamda ilk olarak verilerin AFA’ya uygunluğunu tespit etmek amacıyla DPA’nın alt ölçütleri arasında çoklu bağlantı sorunu olup olmadığı kontrol edilmiş ve ardından alt ölçütlere ait normallik kontrolü için Basıklık ve Çarpıklık katsayıları hesaplanmıştır. Basıklık ve Çarpıklık katsayılarının -2 ve +2 arasında olması verilerin normal dağılıma uygun olduğunu kanıtlamaktadır (George & Mallery, 2022, s. 115). Çoklu bağlantı, değişkenler arasındaki korelasyonların ($r > 0,90$ ve üzeri) yüksek olması durumunda ortaya çıkmaktadır” (Çokluk vd., 2021, s. 210). Yapılan hesaplamalarda; 0,90’dan daha yüksek bir korelasyon katsayısı bulunmadığı için çoklu bağlantı sorununun olmadığı ve “Görsel Sunum” alt ölçütü dışında diğer alt ölçütlerin Basıklık ve Çarpıklık katsayılarının -2 ve +2 arasında değiştiği için normal dağılıma uygun olduğu sonucuna ulaşılmıştır. “Görsel Sunum” alt ölçütünün Çarpıklık değeri 3,501; Basıklık değeri ise 11,759 hesaplandığından, normal dağılıma uygunluk göstermediği için AFA’nın yapılmasının uygun olmadığı sonucuna varılmıştır ve ölçme aracından çıkarılmıştır. Bu değerler ortaya çıkma sebebi olarak öğrencilerin çok büyük bir bölümünün performans görevlerinde tablo, şekil, grafik vb. görselleri kullanmamaları gösterilebilir. Öğrenciler tarafından kullanılmayan görseller, “Görsel Sunum” alt ölçütü için verilen puanın benzer şekilde tekrarlanmasına ve söz konusu ölçüte ilişkin verilerin normal dağılıma uygunluk göstermemesine neden olmuştur. Öğrencilerin uygulama öncesinde değerlendirme ölçütlerinden haberdar olmalarına rağmen görsel, tablo, şekil vb. araçlar ile akademik yazılarını desteklememeleri, akademik yazma bağlamında görsel kullanımını önemli görmedikleri sonucunu ortaya koymaktadır. Bunların yanı sıra AFA yapılmadan önce ölçütler için örneklem büyüklüğünün yeterliliğini tespit etmek amacıyla Kaiser-Meyer-Olkin (KMO) ve alt ölçütler arasındaki ilişkinin faktör analizine uygunluğunu belirlemek için Bartlett Küresellik testi uygulanmıştır. Çözümlemelerde KMO değeri “Biçim” ölçütü için 0,762; “Akademik Söylem” ölçütü için 0,690; “Metin Yapısı” ölçütü için 0,868; “Plan ve Düzenleme” ölçütü için 0,878 ve “Kaynaklardan Yararlanma” ölçütü için 0,724 olarak hesaplanmıştır. Ayrıca tüm alt ölçütler için Bartlett Küresellik testinde ki-kare değeri (“Biçim” için $X^2=517,818$, $p < 0,001$; “Akademik Söylem” için $X^2=247,616$, $p < 0,001$; “Metin Yapısı” için $X^2=579,990$, $p < 0,001$; “Plan ve Düzenleme” için $X^2=555,924$, $p < 0,001$; “Kaynaklardan Yararlanma” için $X^2=232,163$, $p < 0,001$) istatistiksel olarak anlamlı bulunmuştur. “Kaynaklardan Yararlanma” ölçütünün hesaplanan ilk KMO değeri 0,500 olarak hesaplandığı için “Plan ve Düzenleme” ölçütünde yer alan “Kaynakça” alt ölçütünün yeri değiştirilmiştir ve böylelikle “Kaynaklardan Yararlanma” ana ölçütü “Doğrudan ve Dolaylı Alıntı” ve “Akademik Etik İlkeler” alt ölçütleri ile faktör analizi uygulanabilir yeterli örneklem genişliği varsayımlarını

sağlamıştır. Tüm bu sonuçlardan hareketle; bütün ölçütlerin KMO değerinin 0,60'tan yüksek çıkması (Büyüköztürk, 2018) ve verilerin çok değişkenli normal dağılıma uygun olması (Çokluk vd., 2021) AFA'nın söz konusu ölçme aracına uygulanabilirliğini kanıtlamaktadır.

Veri setinin faktör analizi uygulamaya uygun hâli ile temel bileşenler analizi kullanılarak AFA yapılmıştır. DPA'nın alt ölçütlerinin bulunduğu ana ölçütü karşılama düzeylerini tespit etmek için faktör yük değerleri hesaplanmıştır. Ayrıca faktör yapısını belirlemek için özdeğer bulguları da incelenmiştir. Ek olarak faktör sayısına karar vermeden önce yardımcı bir yöntem olarak çizgi grafiği de gözlemlenmiştir. Çözümlemeler sonucunda DPA'nın tüm alt ölçütlerinin faktör yüklerinin 0,30'dan yüksek olduğu gözlemlendiği için bulunduğu ölçütü iyi derecede ölçtüğü belirlenmiştir. Kline (1994) 0,30 ve daha yüksek olan faktör yüklerini yeterli olarak görmektedir. Benzer şekilde Büyüköztürk (2002) de 0,30-0,59 arasındaki yük değerlerini orta; 0,60 ve üstü yük değerlerini yüksek büyüklükler olarak tanımlamaktadır. Ayrıca DPA'da her bir ölçütte özdeğeri 1'den büyük (Çokluk vd., 2021) olan tek bir bileşenin bulunduğu ve bu durumun çizgi grafiğinde de görüldüğü tespit edilmiştir. Dolayısıyla alt ölçütlerin, bağlantılı olduğu ölçütleri tek faktörlü bir yapıda açıkladığı sonucuna varılmıştır. Başlangıç özdeğerleri 1'den büyük olan tek bileşenin, bağlı olduğu ölçüte ait varyansları incelendiğinde; "Biçim" ölçütü için %66,578'ini; "Akademik Söylem" ölçütü için %78,613'ünü, "Metin Yapısı" ölçütü için %66,560'ını, "Plan ve Düzenleme" ölçütü için %61,067'sini ve "Kaynaklardan Yararlanma" ölçütü için %79,085'ini açıkladığı gözlenmiştir. Alpar'a (2017) göre sosyal bilimlerde açıklanan varyans değerinin %60 olması gerekli görülmektedir. Bu bağlamda mevcut çalışmada hesaplanan varyans değerlerinin yeterli olduğu sonucuna ulaşılmıştır. Ayrıca uygulanan faktör analizleri sonucunda her ölçüt, tek bileşenden oluştuğu için döndürme işlemine gereksinim duyulmamıştır. Alanyazında dereceli puanlama anahtarı geliştirme çalışmalarında (Öztürk ve Kaplan, 2017; Yılmaz ve İnceağaç, 2017; Aktaş ve Alıcı, 2018; Yılmaz ve Canpolat, 2020) söz konusu aracın yapı geçerliğini sınamak amacıyla gerçekleştirilen açımlayıcı faktör analizinde (AFA), bu çalışmada izlenen aşamalar takip edilmiş ve elde edilen değerler benzer şekilde yorumlanmıştır.

Uluslararası öğrencilerin Türkçe akademik yazma becerilerine yönelik geliştirilen analitik dereceli puanlama anahtarının güvenilirliği bağlamında Cronbach Alfa iç tutarlılık katsayısına, puanlayıcılar arası uyuma ve madde-toplam korelasyon değerlerine bakılmıştır. Geliştirilen analitik dereceli puanlama anahtarının Cronbach Alpha güvenilirlik katsayısı AFA'dan önce ve sonra olmak üzere iki ayrı şekilde her bir ölçüt ve alt ölçütlerin toplamı için hesaplanmıştır. Bu hesaplamalar sonucunda Cronbach Alpha güvenilirlik katsayısının "Biçim" ölçütü için 0,851, "Akademik Söylem" ölçütü için 0,863, "Metin Yapısı" ölçütü için AFA'dan önce 0,846 iken AFA'dan sonra 0,891, "Plan ve Düzenleme" ölçütü için AFA'dan önce 0,892 iken AFA'dan sonra 0,885, "Kaynaklardan Yararlanma" ölçütü için AFA'dan önce 0,797 iken AFA'dan sonra 0,867 olduğu belirlenmiştir. Sözü edilen ölçme aracının toplam Cronbach Alpha güvenilirlik katsayısı ise AFA'dan önceki hâli için 0,953; AFA'dan sonraki hâli için 0,956 olarak hesaplanmıştır. Alanyazında güvenilirlik katsayısının yorumlanmasına ilişkin farklı görüşler yer almaktadır. Ancak kaynakların çoğunda güvenilirlik katsayısının 0,70 ve daha yüksek olması ölçme aracının güvenilirliği için yeterli olarak kabul edilmektedir (Nunnally,

1978; Büyüköztürk, 2018; Sönmez ve Alacapınar, 2019). Benzer şekilde, George & Mallery (2022) her durum için kabul edilebilir bir Alfa değerine ilişkin kesin bir yorum yapılamayacağını belirterek, çoğu durumda kabul edilebilirlik ölçütlerini; 0,70'ten büyükse kabul edilebilir düzeyde, 0,80'den büyükse iyi düzeyde ve 0,90'dan büyükse mükemmel düzeyde olarak ele almaktadır. Bu çalışmada Cronbach Alfa güvenilirlik katsayısına ilişkin yapılan hesaplamalar sonucunda DPA'nın tamamı ve beş ana ölçütü için güvenilirliğinin iyi ve mükemmel düzeyde olduğu görülmektedir.

DPA'ya yönelik yapılacak güvenilirlik çalışmalarında uygulanabilecek bir diğer yöntem ise puanlar arasındaki tutarlılığa bakmaktır. Bu amaçla bir göreve dayalı geliştirilmiş DPA'nın birden çok değerlendirici tarafından puanlanması sağlanmalıdır (Kutlu vd., 2017). Bu çalışmada araştırmacı dâhil 6 puanlayıcı 25 öğrencinin performans görevini değerlendirmiştir. Puanlayıcılar arası uyumu değerlendirmek için her alt ölçüt ayrı ayrı ve ikiden fazla puanlayıcı tarafından değerlendirildiğinden dolayı Fleiss Kappa değerleri hesaplanmış; ana ölçütler, alt ölçütlerin toplanması ile elde edildiği için Sınıf İçi Korelasyon (Intra-Class Correlation) katsayısı kullanılmıştır. Kutlu vd. (2017) toplam puan üzerinden uyum düzeylerine bakılacağı zaman Kendall W veya Sınıf İçi Korelasyon Katsayısına; her ölçüt için ayrı ayrı uyum düzeyi değerlendirilmek istendiğinde iki puanlayıcı için Cohen's Kappa, ikiden fazla puanlayıcı için Fleiss Kappa formülünün kullanılması gerektiğini belirtmektedir. Uluslararası öğrencilerin Türkçe akademik yazma becerilerine yönelik geliştirilen analitik dereceli puanlama anahtarına ait alt ölçütler için puanlayıcılar arası uyum düzeyleri; en düşük "Anlaşılabilirlik", en yüksek "Anahtar Sözcükler" olmak üzere 0,637-0,909 arasında değişim göstermektedir. -1 ile +1 arasında değer alan Fleiss Kappa katsayısı alanyazında, 0,41-0,60 arasında orta düzeyde; 0,61-0,80 arasında yüksek düzeyde; 0,81-1 arasında çok yüksek düzeyde uyum biçiminde yorumlanmaktadır (Fleiss, 1971; Landis & Koch, 1977). Buna göre DPA'nın alt ölçütlerine ilişkin hesaplanan Fleiss Kappa değerleri, puanlayıcılar arası uyumun yüksek ve çok yüksek düzeyde olduğunu göstermektedir. Bunun yanı sıra DPA'daki beş ölçütün Sınıf İçi Korelasyon katsayısı sonuçları değerlendirildiğinde puanlayıcılar arası uyum düzeylerinin tek ölçümlerde en düşük "Biçim" ölçütüne, en yüksek "Plan ve Düzenleme" ölçütüne ait olmak üzere 0,909-0,980 arasında değişim gösterdiği görülmektedir. Ortalama ölçümlerde ise en düşük uyum düzeyi "Biçim" ölçütüne, en yüksek "Plan ve Düzenleme" ölçütüne ait olmak üzere 0,984-0,997 arasında değişim göstermektedir. DPA'dan alınan "Toplam" puan değeri ise tek ölçümlerde 0,993 hesaplanırken, ortalama ölçümlerde 0,999 olarak hesaplanmıştır. Alanyazında, Sınıf İçi Korelasyon (Intraclass Correlation Coefficient) katsayısının 0 ve 1 arasında değerler aldığı belirtilmiş olup; 0,50'den küçük değerler zayıf güvenilirliği, 0,50-0,75 arasındaki değerler orta düzeyde güvenilirliği, 0,75-0,90 arasındaki değerler iyi güvenilirliği ve 0,90'dan büyük değerler mükemmel güvenilirliği göstermektedir. (Koo & Li, 2016). Buna göre DPA'nın tüm ölçütleri ve toplam puanı için hesaplanan Sınıf İçi Korelasyon değerlerinin tek ve ortalama ölçümlerde mükemmel uyum gösterdiği tespit edilmiştir. Tüm bu bulgular bütüncül olarak değerlendirildiğinde, puanlayıcılar arasında yüksek ve çok yüksek uyum olduğu sonucuna ulaşılmıştır. Bu sonuç DPA kullanılarak yapılan değerlendirmelerde; puanlamanın puanlayıcılar arasında değişmediğini, puanlayıcıların birbiriyle tutarlı puanlar verdiğini ve nesnellüğün sağlandığını göstermektedir. Benzer şekilde

alanyazında DPA kullanılmasının nesnellığı artırdığına ilişkin sonuçları ortaya koyan çalışmalara rastlanmıştır (Erman-Aslanoğlu ve Kutlu, 2003; Bıkmaz-Bilgen ve Doğan, 2017; Bilican-Demir ve Yıldırım, 2019; Polat-Demir, 2020; Yorgancı ve Baş, 2021). Diğer yandan çeşitli araştırmalarda (Durmuş, 2022; Tarakcı, 2022) DPA geliştirme sürecinde güvenilirliğin sınanması bağlamında puanlayıcılar arası uyumu belirlemek için kullanılan yöntemler ile bu çalışmada kullanılan yöntemlerin koşutluk gösterdiği tespit edilmiştir.

Bu çalışma kapsamında madde geçerliğine kanıt olarak DPA'da alt ölçütlerin ilgili olduğu ana ölçüt ve toplam puanla uyumlu olup olmadığını belirlemek amacıyla madde-toplam korelasyonları hesaplanmıştır. Madde-toplam korelasyonunun pozitif ve yüksek olması, maddelerin benzer davranışları örneklediğini göstermektedir ve testin iç tutarlılığının yüksek olduğunu kanıtlamaktadır. Bu bağlamda madde-toplam korelasyonu 0,30 ve daha yüksek olan maddelerin, bireyleri iyi derecede ayırt edebildiği için doğrudan teste alınması önerilmektedir. 0,20 ile 0,30 arasında kalan maddelerin ise zorunlu olması durumunda alınabileceği veya düzeltilmesi gerektiği belirtilmektedir. Bunun yanı sıra, 0,20'den düşük maddelerin ise teste kesinlikle alınmaması gerektiği ifade edilmektedir (Büyüköztürk, 2018, s. 183). Madde-toplam korelasyon değerlerini belirlemeye yönelik yapılan hesaplamalar sonucunda ana ölçütlerdeki alt ölçütlerin madde-toplam korelasyonlarının tamamının 0,30'dan yüksek olduğu tespit edilmiştir. Ayrıca alt ölçütler ile toplam puan arasındaki madde-toplam korelasyon değerlerinin; en düşük "Dil Bilgisi", en yüksek "Gelişme" alt ölçütü olmak üzere 0,568 ile 0,833 arasında değiştiği görülmüştür. Tüm hesaplamalardan elde edilen madde-toplam korelasyon değerlerinin beklenen düzeyden yüksek ve pozitif olmasından dolayı DPA'daki alt ölçütlerin, ölçütlerle ve toplam puanla tutarlı olduğu sonucuna ulaşılmıştır. Alanyazında farklı amaçlarla geliştirilen dereceli puanlama anahtarı çalışmalarında (Polat-Demir, 2020; Bülbül ve Bülbül, 2021) madde geçerliğine kanıt olarak madde-toplam korelasyon değerlerinin incelendiği görülmektedir. Söz konusu çalışmalar ile mevcut çalışmanın bu bağlamda benzerlik gösterdiği söylenebilir.

Bu çalışmada geçerlik ve güvenilirlik çalışmaları sonucunda uluslararası öğrencilerin Türkçe akademik yazma becerilerine yönelik geliştirilen analitik dereceli puanlama anahtarının son hâli, 5 ana ölçüt ve 24 alt ölçütten oluşmuştur. Ana ölçütler, alt ölçütleriyle birlikte şöyle sıralanmaktadır: "Biçim" (Dil bilgisi, noktalama işaretleri, yazım kuralları, anlaşılabilirlik, sayfa düzeni), "Akademik Söylem" (Söz varlığı, cümle yapısı, üstsöylem belirleyicileri), "Metin Yapısı" (Konu, araştırmanın problem durumu ve amacı, bağdaşıklık, tutarlılık, paragraf yapısı, bilgileri sentezleme), "Plan ve Düzenleme" (Başlık, özet, anahtar sözcükler, giriş, başlıklandırma, gelişme, tartışma, sonuç ve öneriler), "Kaynaklardan Yararlanma" (Kaynakça, doğrudan ve dolaylı alıntı, akademik etik ilkeler). Alanyazın incelendiğinde uluslararası öğrencilerin Türkçe akademik yazma gereksinimlerini temele alarak geliştirilen bir ölçme aracı tespit edilemese de farklı hedef kitlelerin ilgili becerilerini ölçmek amacıyla kullanılan çeşitli türlerde ölçme araçları olduğu belirlenmiştir (Aydın, 2015; Izawa et al., 2017; Shekarabi, 2017; Al-Abdullatif et al., 2020; Polat-Demir, 2020; Baz-Bolluk ve Batur, 2023). Al-Abdullatif et al. (2020) çalışmasında; "Doğruluk", "Açıklık", "Tutarlılık" ve "Sözcük Kullanımı" olmak üzere 4 boyuttan oluşan bir ölçekte lisans öğrencilerinin akademik yazma becerilerini ölçmüştür. Aydın'ın (2015), Razi'dan (2015) uyarladığı Akademik Makale Değerlendirme Rubriği ise "Genel Özellikler",

"Özet", "Giriş", "Yöntem", "Bulgular ve Yorum", "Sonuç ve Öneriler", "Kaynakça" olmak üzere 7 bölümden ve 50 maddeden oluşmaktadır. Izawa et al. (2017) tarafından kullanılan akademik yazma analitik dereceli puanlama anahtarı; "Tez ve İçerik", "Organizasyon", "Dil Bilgisi ve Mekanik (Noktalama işaretleri, yazım kuralları, sözcük kullanımı, cümle yapısı vb.)", "Sözcük Bilgisi", "Ton" ve "Uzunluk" olmak üzere 6 ölçütten oluşmaktadır. Shekarabi (2017) ikinci dil olarak Japonca öğrenen öğrencilerin eleştirel düşünme becerilerinin akademik yazmaya etkisini araştırdığı çalışmasında; "İçerik", "Organizasyon", "Tutarlılık", "Dil Doğruluğu" ve "Mekanik" olmak üzere 5 ölçütten oluşan dereceli puanlama anahtarını kullanmıştır. Polat-Demir'in (2020) lisans öğrencilerinin akademik yazma becerilerini değerlendirmek için göreve dayalı (derleme makalesi) olarak geliştirdiği dereceli puanlama anahtarı; "Görev Tanımı", "İçerik", "Kaynak Kullanımı", "Anlatım" ve "Biçimsel Özellikler" olmak üzere 5 ana ölçüt ve 18 alt ölçütten oluşmaktadır. Baz-Bolluk ve Batur (2023) akademik yazma çalışmalarının yeterliğini değerlendirmeye yönelik "Başlık", "Özet", "Giriş", "Yöntem", "Bulgu ve Yorumlar", "Sonuç ve Tartışma", "Öneri", "Kaynakça", "Dil ve Anlatım" ve son olarak "Diğer" olmak üzere 61 maddeden ve 10 alt boyuttan oluşan bir ölçme aracı geliştirmiştir. Mevcut çalışmada geliştirilen DPA'daki ölçüt ve alt ölçütler, söz konusu çalışmalarda yer alan ölçüt ve alt ölçütleri büyük oranda kapsamakla birlikte çalışmanın hedef kitlesine özgü özellikleri de göz önünde bulundurularak oluşturulmuştur. Özetle, bu çalışma kapsamında geliştirilen "Uluslararası Öğrencilere Yönelik Türkçe Akademik Yazma Analitik Dereceli Puanlama Anahtarı"nın söz konusu öğrencilerin ilgili becerilerini ölçebilecek nitelikte geçerli ve güvenilir bir araç olduğu sonucuna varılmıştır.

Araştırmadan elde edilen sonuçlar göz önünde bulundurularak program tasarımcılarına, uygulayıcılara ve araştırmacılara katkı sağlayacak öneriler aşağıda sıralanmıştır:

1. Bu çalışma kapsamında geliştirilen performansa dayalı analitik dereceli puanlama anahtarının bazı alt ölçütlerinde değişiklikler yapılarak farklı akademik metin türleri, dil düzeyleri ve öğretim kademeleri için uyarlama çalışmaları yapıp, tekrar geçerlik ve güvenilirlik durumları sınanabilir. Benzer şekilde ana dillerinde öğrenim gören öğrencilerin Türkçe akademik yazma becerilerinin değerlendirilmesinde de söz konusu DPA'nın kullanılıp kullanılmayacağına ilişkin mevcut ölçme aracı üzerinde geçerlik ve güvenilirlik çalışmaları yapılarak sonuçları yorumlanabilir.
2. Çalışmada geliştirilen DPA, yapı geçerliği bağlamında açımlayıcı faktör analizi (AFA) ile değerlendirilmiştir. AFA ile elde edilen verilerin doğrulanıp doğrulanmadığını belirlemek amacıyla yeni bir öğrenci grubundan elde edilen veriler ile doğrulayıcı faktör analizi (DFA) gerçekleştirilebilir.
3. Bu çalışma bağlamında geliştirilen analitik dereceli puanlama anahtarındaki ana ölçüt, alt ölçüt ve düzey tanımlamalarının, hem hazırlanacak akademik yazma kazanımları ve ders içeriklerine hem de öğretim görevlilerine yol gösterici olacağı düşünülmektedir. Bu bağlamda tüm paydaşlar, bu çalışmadan elde edilen verilerden faydalanabilir.

Yazar Katkı Oranı

Bu makale, ikinci yazarın danışmanlığında birinci yazar tarafından hazırlanan doktora tezinden üretilmiştir. Tüm yazarlar makalenin yazımına eşit oranda katkı sağlamış, çalışmanın son hâlini okumuş ve onaylamıştır.

Etik Kurul Beyanı

Bu çalışma, İstanbul Üniversitesi-Cerrahpaşa Sosyal ve Beşerî Bilimleri Araştırmaları Etik Kurulunda 14.06.2022 tarihli 2022/202 toplantısında alınan onay kararı ile yürütülmüştür.

Çatışma Beyanı

Yazarlar, bu çalışma kapsamında herhangi bir kurum veya kişi ile çıkar çatışması bulunmadığını beyan etmektedir.

Kaynakça

- Aktaş, M. ve Alıcı, D. (2018). Yazılan hikâyeyi değerlendirmeye yönelik analitik rubrik geliştirme: Geçerlik ve güvenilirlik çalışması. *Mersin Üniversitesi Eğitim Fakültesi Dergisi*, 14(2), 597-610. <https://doi.org/10.17860/mersinefd.424198>
- Al-Abdullatif, A. M., Alsubaie, M. A. & Aldoughan, E. A. (2020). Exploring the effects of excessive texting through mobile applications on students' technostress and academic writing skills in the Arabic language. *IEEE Access*, 8, 166940-166950.
- Alpar, R. (2017). *Uygulamalı çok değişkenli istatistiksel yöntemler* (5. Baskı). Detay Yayıncılık.
- Ankara Üniversitesi TÖMER (2015). *Yabancı dil olarak Türkçe öğretim programı*. M. E. Gökmen (Ed.), Ankara Üniversitesi Basımevi.
- AOBM (2021). *Diller için Avrupa ortak başvuru metni: Öğrenme, öğretme ve değerlendirme, tamamlayıcı cilt* (MEB Komisyonu, Çev.). Pozitif Matbaa (Eserin orijinal yayın tarihi: 2020).
- Arıcı, M. A. (2022). *Yansıtıcı düşünmeye dayalı öğretimin Türkçe öğrenen yabancı öğrencilerin konuşma becerilerine ve konuşma özyeterlik algılarına etkisi* [Yayımlanmamış doktora tezi]. İstanbul Üniversitesi-Cerrahpaşa.
- Aydın, G. (2015). *Türk dili ve edebiyatı hizmet öncesi öğretmen eğitiminde yansıtıcı öğretim uygulamalarının akademik yazma becerilerine etkisi* [Yayımlanmamış doktora tezi]. Atatürk Üniversitesi.
- Azizoğlu, N. İ., Demirtaş-Tolaman, T. ve İdi-Tulumcu, F. (2019). Yabancı dil olarak Türkçe öğretiminde akademik yazma becerisi: Karşılaşılan sorunlar ve çözüm önerileri. *Uluslararası Yabancı Dil Olarak Türkçe Öğretimi Dergisi*, 2(1), 7-22.
- Bahar, M. A. (2014). Lisansüstü eğitimde akademik yazma ve önemi. *International Journal of Language Academy*, 2(4), 209-233.
- Bailey, S. (2020). *Akademik yazma: Uluslararası öğrenciler için el kitabı* (M. Karataş, Çev. Ed.). Nobel.
- Başol, G. (2019). *Eğitimde ölçme ve değerlendirme* (6. Baskı). Pegem Akademi.
- Bayat, N. (2009). Türkçe ve İngilizce öğretmen adaylarının akademik yazılarında yapısal özellikler. *Dil Dergisi*, (145), 48-63. https://doi.org/10.1501/Dilder_0000000115
- Baz-Bolluk, D. ve Batur, Z. (2023). Akademik yazma çalışmalarının yeterliliğini değerlendirmeye yönelik bir ölçme aracı geliştirme önerisi. *Turkish Studies-*

- Educational Sciences*, 18(1), 15-26. <http://dx.doi.org/10.7827/TurkishStudies.67322>
- Bıkmaz-Bilgen, Ö. (2023). Tamamlayıcı ölçme ve değerlendirme teknikleri I: Performans değerlendirme. N. Doğan (Ed.), *Eğitimde ölçme ve değerlendirme* (4. Baskı) içinde (s. 181-216). Pegem Akademi.
- Bıkmaz-Bilgen, Ö. ve Doğan, N. (2017). Puanlayıcılar arası güvenilirlik belirleme tekniklerinin karşılaştırılması. *Eğitimde ve Psikolojide Ölçme ve Değerlendirme Dergisi*, 8(1), 63-78. <https://doi.org/10.21031/epod.294847>
- Bilican-Demir, S. ve Yıldırım, Ö. (2019). Yazılı anlatım becerilerinin değerlendirilmesi için dereceli puanlama anahtarı geliştirme çalışması. *Pamukkale Üniversitesi Eğitim Fakültesi Dergisi*, 47, 457-473. <https://doi.org/10.9779/pauefd.588565>
- Bozkurt, B. Ü. ve Arıca-Akkök, E. (2019). Anadili Türkçe olan konuşucular için konuşma becerisi derecelendirme ölçeğinin geliştirilmesi. *İlköğretim Online*, 18(1), 416-436. <https://doi.org/10.17051/ilkonline.2019.527649>
- Bülbül, S. ve Bülbül, B. Ö. (2021). Ortaokul ders kitaplarının değerlendirilmesi: Bir dereceli puanlama anahtarı geliştirme çalışması. *Van Yüzcü Yıl Üniversitesi Eğitim Fakültesi Dergisi*, 18(1), 882-911. <https://doi.org/10.33711/yyuefd.938727>
- Büyüköztürk, Ş. (2002). Faktör analizi: Temel kavramlar ve ölçek geliştirmede kullanımı. *Kuram ve Uygulamada Eğitim Yönetimi*, 32(32), 470-483.
- Büyüköztürk, Ş. (2018). *Sosyal bilimler için veri analizi el kitabı* (24. Baskı). Pegem Akademi.
- Creswell, J. W. & Plano-Clark, V. L. (2018). *Karma yöntem araştırmaları tasarımı ve yürütülmesi* (Y. Dede ve S. B. Demir, Çev. Ed.; 3. Baskı). Anı Yayıncılık.
- Çepni, S. (2015). Performansların değerlendirilmesi. E. Karip (Ed.), *Ölçme ve değerlendirme* (7. Baskı) içinde (s. 231-290). Pegem Akademi.
- Çokluk, Ö., Şekercioğlu, G. ve Büyüköztürk, Ş. (2021). *Sosyal bilimler için çok değişkenli istatistik: SPSS ve LISREL uygulamaları* (6. Baskı). Pegem Akademi.
- Danielson, C. & Hansen, P. (2016). *Performance tasks and rubrics for early elementary mathematics* (2nd Edition). Routledge Taylor & Francis Group.
- Davis, L. L. (1992). Instrument review: Getting the most from a panel of experts. *Applied Nursing Research*, 5(4), 194-197. [https://doi.org/10.1016/S0897-1897\(05\)80008-4](https://doi.org/10.1016/S0897-1897(05)80008-4)
- Demir, D. (2017). *Uluslararası öğrencilerin akademik Türkçe ihtiyaçları* [Yayımlanmamış doktora tezi]. Hacettepe Üniversitesi.
- Demirci, A. (2014). Akademik araştırmalarda yazım aşaması. Y. Arı ve İ. Kaya (Ed.), *Coğrafya araştırma yöntemleri* içinde (s. 109-154). Coğrafyacılar Derneği.
- Demiriz, H. N. ve Okur, A. (2019). Türkçe öğretiminde yazma öğretimine akademik Türkçe aşamasında yabancı öğrenciler üzerinden bir bakış. *Ana Dili Eğitimi Dergisi*, 7(2), 436-449. <https://doi.org/10.16916/aded.525728>
- Deniz, K. ve Karagöl, E. (2018). Akademik yazma açısından dergi yazar rehberleri. *Ana Dili Eğitimi Dergisi*, 6(4), 1210-1225. <https://doi.org/10.16916/aded.458116>
- Durmuş, A. O. (2022). *Dereceli puanlama anahtarına dayalı geribildirim Türkçenin yabancı dil olarak öğretiminde yazma becerisinin geliştirilmesine etkisi* [Yayımlanmamış doktora tezi]. Ankara Üniversitesi.
- Erman-Aslanoglu, A. (2003). *Öğretimde sunu becerilerinin değerlendirilmesinde dereceli puanlama anahtarı (rubric)*

- kullanılmasına ilişkin bir araştırma [Yayımlanmamış yüksek lisans tezi]. Ankara Üniversitesi.
- Erman-Aslanoglu, A. ve Kutlu, Ö. (2003). Öğretimde sunu becerilerinin değerlendirilmesinde dereceli puanlama anahtarı (rubric) kullanılmasına ilişkin bir araştırma, *Ankara Üniversitesi Eğitim Bilimleri Fakültesi Dergisi*, 36(1-2), 25-36.
- Fleiss, J. L. (1971). Measuring nominal scale agreement among many raters. *Psychological Bulletin*, 76(5), 378-382. <https://doi.org/10.1037/h0031619>
- George, D. & Mallery, P. (2022). *IBM SPSS statistics 27 step by step: A simple guide and reference* (17nd Edition). Routledge.
- Goodrich-Andrade, H. (2000). Using rubrics to promote thinking and learning. *Educational Leadership*, 57(5), 13-18.
- Goodrich-Andrade, H. (2001). The effects of instructional rubrics on learning to write. *Current Issues in Education*, 4(4), 1-22.
- Goodrich-Andrade, H. (2005). Teaching with rubrics: The good, the bad, and the ugly. *College Teaching*, 53(1), 27-31. <https://doi.org/10.3200/CTCH.53.1.27-31>
- GSE (2022). *Global Scale of English Learning Objectives for Academic English*. Pearson Education. 15 Ocak 2023 tarihinde <https://www.pearson.com/languages/why-pearson/the-global-scale-of-english/resources.html> adresinden edinilmiştir.
- Güler, N. (2019). *Eğitimde ölçme ve değerlendirme* (14. Baskı). Pegem Akademi.
- Güngör, H. (2020). Akademik Türkçe öğretiminde yazma eğitimi. B. Tüfekçioğlu (Ed.), *Akademik amaçlar için Türkçe öğretimi: Kuram ve uygulama* içinde (s. 165-190). Pegem Akademi.
- Harmankaya, M. Ö., Sallabaş, M. E. ve Toker, T. (2022). Türkçeyi yabancı dil olarak öğreten ve öğrenenler için A2 düzeyi yazma becerisi dereceli puanlama anahtarı geliştirme çalışması. *IBAD Sosyal Bilimler Dergisi*, (13), 511-524. <https://doi.org/10.21733/ibad.1212818>
- Hogue, A. (2008). *First steps in academic writing*. Pearson Longman.
- Irvin, L., L. (2010). What is “academic” writing?. C. Lowe & P. Zemliansky (Eds.), *Writing spaces: Readings on writing* içinde (s. 3-17). Parlor Press.
- Izawa, Y., Fujimori, C., Godfrey, C. & Oida, Y. (2017). Relationships among self-efficacy, willingness, and writing performance in an academic writing program. *Fora*, 1, 12-22.
- İltar, L. ve Karataş, A. G. (2022). Türkçenin yabancı dil olarak öğretiminde anlatmaya/göstermeye dayalı metinler için yazma becerisi dereceli puanlama anahtarı. *Okuma Yazma Eğitimi Araştırmaları*, 10(2), 194-213. <https://doi.org/10.35233/oyea.1177730>
- Kahveci, N. (2022). *Yabancı dil olarak Türkçe eğitiminde yazma becerisini değerlendirmeye yönelik dereceli puanlama anahtarı geliştirme çalışması* [Yayımlanmamış yüksek lisans tezi]. Bartın Üniversitesi.
- Kan, M. O. (2017). Türkiye’de akademik yazma alanında yapılan lisansüstü tezler. *İnsan ve Toplum Bilimleri Araştırmaları Dergisi*, 6(2), 1037-1048.
- Kan, M. O. ve Gedik, F. N. (2016). Türkçe eğitimi alanında yüksek lisans tezi tamamlayanların akademik yazma ve tez oluşturma sürecine ilişkin görüşleri. *Yükseköğretim ve Bilim Dergisi*, 6(3), 402-408.
- Karagöl, E. (2018). *Akademik yazma açısından lisansüstü tezler* [Yayımlanmamış doktora tezi]. Gazi Üniversitesi.
- Kline, P. (1994). *An easy guide to factor analysis*. Routledge.
- Konyar, M. (2019). *Uluslararası öğrencilerin akademik Türkçe ihtiyaç analizi ve örnek ders içeriği* [Yayımlanmamış yüksek lisans tezi]. Tokat Gaziosmanpaşa Üniversitesi.
- Koo, T. K. & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine*, 15(2), 155-163.
- Kutlu, Ö., Doğan, C. D. ve Karakaya, İ. (2017). *Ölçme ve değerlendirme: Performansa ve portfolyoya dayalı durum belirleme* (5. Baskı). Pegem Akademi.
- Kutlu, Ö., Karakaya, İ. ve Doğan, C. (2008). Üst düzey zihinsel becerilerin belirlenmesi: Performans görevi yazma. *İlköğretmen Dergisi*, (18), 10-15.
- Landis, J. R. & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33(1), 159-174. <https://doi.org/10.2307/2529310>
- Lawshe, C. H. (1975). A quantitative approach to content validity. *Personnel Psychology*, 28(4), 563-575. <https://doi.org/10.1111/j.1744-6570.1975.tb01393.x>
- Lynn, M. R. (1986). Determination and quantification of content validity. *Nursing Research*, 35(6), 382-385. <https://doi.org/10.1097/00006199-198611000-00017>
- Mertler, C. A. (2001). Designing scoring rubrics for your classroom. *Practical Assessment, Research & Evaluation*, 7(25), 1-8. <https://doi.org/10.7275/gcy8-0w24>
- Moskal, B. M. & Leydens, J. A. (2000). Scoring rubric development: Validity and reliability. *Practical Assessment, Research & Evaluation*, 7(10), 1-6. <https://doi.org/10.7275/q7rm-gg74>
- Moskal, B. M. (2000). Scoring rubrics: What, when and how? *Practical Assessment, Research & Evaluation*, 7(3), 1-5. <https://doi.org/10.7275/a5vq-7q66>
- Nunnally, J. C. (1978). *Psychometric theory* (2nd Edition). McGraw-Hill.
- Özdemir, T. Y. (2018). Karma yöntem araştırmalarında keşfedici desen. K. Beycioğlu, N. Özer ve Y. Kondakçı (Ed.), *Eğitim yönetiminde araştırma* içinde (s. 493-502). Pegem Akademi.
- Öztürk, M. ve Kaplan, A. (2017). Matematik öğretmenlerine yönelik ispat yapma teşhis testi ve teste yönelik dereceli puanlama anahtarı geliştirilmesi. *Van Yüzüncü Yıl Üniversitesi Eğitim Fakültesi Dergisi*, 14(1), 360-381. <http://dx.doi.org/10.23891/efdyyu.2017.14>
- Parlak, B. ve Doğan, N. (2014). Dereceli puanlama anahtarı ve puanlama anahtarından elde edilen puanların uyum düzeyleri. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*, 29(2), 189-197.
- Pineteh, E. A. (2014). The academic writing challenges of undergraduate students: A South African case study. *International Journal of Higher Education*, 3(1), 12-22. <https://doi.org/10.5430/ijhe.v3n1p12>
- Polat-Demir, B. (2020). Akademik yazma becerilerinin değerlendirilmesi için analitik dereceli puanlama anahtarının geliştirilmesi. *Amasya Üniversitesi Eğitim Fakültesi Dergisi*, 9(2), 318-327.
- Popham, W. J. (1997). What’s wrong and what’s right with rubric. *Educational Leadership*, 55(2), 72-75.
- Razi, S. (2015). Development of a rubric to assess academic writing incorporating plagiarism detectors. *SAGE Open*, 5(2), 1-13. <https://doi.org/10.1177/2158244015590162>

- Seyedi, G. (2019). *Yabancı dil olarak Türkçe öğretiminde akademik yazma öğretimi* [Yayımlanmamış doktora tezi]. Hacettepe Üniversitesi.
- Shekarabi, Z. (2017). Cognitive skills for critical thinking and second language academic writing, *Bull. Grad. School Educ. Hiroshima Univ.*, 2(66), 155-164.
- Sönmez, V. ve Alacapınar, F. G. (2019). *Örneklendirilmiş bilimsel araştırma yöntemleri* (7. Baskı). Anı Yayıncılık.
- Swales, J. M. & Feak C. B. (2012). *Academic writing for graduate students essentials and skills*. University of Michigan Press.
- Şenöz-Ayata, C. (2014). *Bilimsel metin üretimi*. Papatya.
- Tarakcı, R. (2022). *Ortaokul öğrencilerinin öyküleyici metin yazma becerileri* [Yayımlanmamış doktora tezi]. Gazi Üniversitesi.
- Tekin, E. ve Sallabaş, M. E. (2022). Yabancı dil olarak Türkçe öğretiminde A2 seviyesi yazma becerisine yönelik analitik rubrik geliştirme. *Anasay*, 6(22), 81-95. <https://doi.org/10.33404/anasay.1181729>
- Tok, M. (2012). *Türkçenin yabancı dil olarak öğretiminde akademik yazma becerilerinin geliştirilmesine yönelik uygulamalı bir çalışma* [Yayımlanmamış doktora tezi]. Onsekiz Mart Üniversitesi.
- Tok, M. (2013). Türkçenin yabancı dil olarak öğretiminde akademik yazma ihtiyacı. *Mustafa Kemal Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, 10(23), 1-25.
- Tok, M. ve Gönülal, N. (2016). Akademik yazma sürecine ilişkin bir ihtiyaç analizi. *Uluslararası Türkçe Eğitimi ve Öğretimi Dergisi*, 1(1), 142-161.
- Tüfekçioğlu, B. (2018). *Yabancı dil olarak akademik Türkçe: Sosyal bilimlerde akademik ve teknik söz varlığı* [Yayımlanmamış doktora tezi]. Hacettepe Üniversitesi.
- Ulutaş, M. (2016). *Türkiye’de yükseköğrenim görecekt yabancılar için özel amaçlı Türkçe öğretimine yönelik program geliştirme: Bir model önerisi* [Yayımlanmamış doktora tezi]. Gazi Üniversitesi.
- Ungan, S. (2007). Yazma becerisinin geliştirilmesi ve önemi. *Erciyes Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, 23(2), 461-472.
- Uzun, L. (2001). Bilimsel metne özgü önbiçimlenişler ve bilimsel metin yazma edimi. *Anatolia: Turizm Araştırmaları Dergisi*, 12(2), 197-204.
- WIDA (2014). *The WIDA English Language Development Standards and Resource Guide, International Edition*.13 Nisan 2023 tarihinde <https://wida.wisc.edu/resources/wida-english-language-development-standards-and-resource-guide-international-edition> adresinden edinilmiştir.
- Yahşi-Cevher, Ö. ve Güngör, C. (2015). Yabancı dil olarak Türkçe öğretiminde akademik Türkçenin önemine ilişkin uygulamalı bir araştırma: Türk dili öğretimi uygulama ve araştırma merkezi örneği. *International Journal of Languages' Education and Teaching (UDES 2015)*, 2267-2274.
- Yıldırım, A. ve Şimşek, H. (2011). *Sosyal bilimlerde nitel araştırma yöntemleri* (8.baskı). Seçkin Yayıncılık.
- Yıldırım, M. ve Nurlu, M. (2016). Yazma gelişim dosyası uygulamasının yabancı dil olarak Türkçe öğrenen öğrencilerin yazma başarılarına etkisi. *Turkish Studies*, 11(14), 827-846. <http://dx.doi.org/10.7827/TurkishStudies.9786>
- Yılmaz, M. ve Canbolat, C. (2020). Masal kitabı tasarımını değerlendirmeye yönelik dereceli puanlama anahtarının geliştirilmesi. *JRES*, 7(2), 543-560.
- Yılmaz, M. ve İnceağaç, M. (2017). Görsel sanatlar eğitimi uygulamalarında edebiyat ürünlerinin etkisini belirlemeye yönelik dereceli puanlama anahtarının (rubrik) geliştirilmesi. *Erzincan Üniversitesi Eğitim Fakültesi Dergisi*, 19(3), 1-16. <https://doi.org/10.17556/erziefd.341456>
- Yorgancı, O. K. ve Baş, B. (2021). Yabancı dil olarak Türkçe öğrenenler için B1 düzeyi yazma becerisi dereceli puanlama anahtarı geliştirme çalışması. *Okuma Yazma Eğitimi Araştırmaları*, 9(1), 67-80. <https://doi.org/10.35233/oyea.934684>
- Yurdakul-Güde, C. (2019). *Yabancı dil olarak Türkçenin öğretiminde akademik yazmaya yönelik bir içerik önerisi* [Yayımlanmamış yüksek lisans tezi]. Gazi Üniversitesi.
- Yücelşen, N. (2021). *Türkçe öğretmeni adaylarının akademik yazma becerilerinin geliştirilmesine yönelik bir eylem araştırması* [Yayımlanmamış doktora tezi]. İstanbul Üniversitesi- Cerrahpaşa.

Extended Summary

Introduction

Language skills refer to the competencies required for understanding, using, and communicating in a language. These skills are divided into two parts: comprehension and expressive skills. People need both comprehension and expression skills to interact with each other and communicate their ideas. While reading and listening constitute the comprehension dimension of language, speaking and writing constitute the expression dimension. All of these language skills are interrelated, and all these skills need to be developed both for the acquisition of the mother tongue and for the learning of a foreign/second language.

In the process of developing language skills, learners may encounter various levels and types of difficulties. In general, it is seen that these difficulties experienced by learners and even instructors in the educational process are related to writing skills. The concept of academic writing has emerged with the use of writing for academic purposes. This concept has been defined in different ways by many researchers in national and international literature (Hogue, 2008; Irvin, 2010; Swales & Feak, 2012; Tok, 2013; Bahar, 2014; Aydın, 2015; Kan, 2017; Karagöl, 2018; Güngör 2020). Based on the definitions and explanations of the related concept, academic writing can be defined as a type of writing in which the results of research conducted in an academic context are reported by paying attention to specific principles in terms of form, discourse, text structure, plan and organization, and resource use, and which requires the effective use of higher-order thinking skills. According to Demirci (2014), it is important to provide students with experience in academic writing to develop these skills. Various courses and lessons have been planned in order to provide experience in this skill, which is especially needed by university students. When considered in the context of teaching Turkish as a foreign/second language, these courses and lectures, which are generally united under the umbrella of academic Turkish, should aim to teach international students the features and rules of academic language, reaching valid and reliable sources by conducting a literature review, creating scientific texts, paying attention to formal features in academic writing, citing and organizing bibliography, complying with ethical principles, etc. through different applications.

In the development of academic writing skills, giving regular constructive feedback and learning from mistakes are as important as having experiences (Pineteh, 2014; Demirci, 2014). Learning from mistakes is made possible by recognizing mistakes and deficiencies. Measurement and evaluation are of great importance in order to determine the strengths and weaknesses of individuals, to eliminate their deficiencies, to correct mistakes, and to contribute to their development. In order to determine the strengths and weaknesses of individuals, appropriate measurement tools should be used (Polat-Demir, 2020, p. 321).

One of the most effective measurement tools that can be used in the measurement, evaluation, and development of academic writing skills, which require the use of high-level thinking skills and include many dimensions such as form, discourse, text, plan, resource use, and ethics, is the analytical rubric. Rubrics have great importance in the field of teaching Turkish as a foreign/second language and academic writing, as in other fields and subjects.

When the literature is examined, it is seen that in the field of teaching Turkish as a foreign/second language, there are

mostly studies on the development of rubrics for the evaluation of writing skills (Yorgancı & Baş, 2021; Durmuş, 2022; Kahveci, 2022; Harmanakaya et al., 2022; Tekin & Sallabaş, 2022), it was also determined that various types of measurement tools were developed to measure the academic writing skills of different target groups (Aydın, 2015; Izawa et al., 2017; Shekarabi, 2017; Al-Abdullatif et al., 2020; Polat-Demir, 2020; Baz-Bolluk & Batur, 2023). However, there is no study on the development of an analytical rubric whose task-based validity and reliability have been tested with multidimensional analysis techniques for the assessment of the relevant skill based on the Turkish academic writing needs of international students. Therefore, this study aimed to develop an analytical rubric for the assessment of the relevant skill based on the Turkish academic writing needs of international students. In line with this main purpose, the questions aimed to be answered in the study can be listed as follows:

1. What are the validity analysis results of the analytical rubric developed for international students' Turkish academic writing skills?
 - a. What is the content validity of the analytical rubric developed for international students' Turkish academic writing skills?
 - b. What is the construct validity of the analytical rubric developed for international students' Turkish academic writing skills?
2. What are the reliability analysis results of the analytical rubric developed for international students' Turkish academic writing skills?
 - a. What is the Croanbach's Alpha reliability coefficient of the analytical rubric developed for international students' Turkish academic writing skills?
 - b. How is the agreement between the raters in the evaluations made according to the analytical rubric developed for international students' Turkish academic writing skills?
 - c. What is the item validity level of the sub-criteria in the analytical rubric developed for international students' Turkish academic writing skills?

It is thought that the developed analytical rubric will contribute to the field in terms of helping to make objective evaluations, contributing to the development of students' Turkish academic writing skills, and guiding both instructors/teachers and students/learners by providing feedback on the achievement of the relevant skill.

Method

In this study, which aims to develop an analytical rubric for the evaluation of international students' Turkish academic writing skills, exploratory sequential design, which is a mixed design strategy, was used. In the exploratory sequential design, "the researcher develops a tool as a middle step between the stages based on qualitative results and uses this tool while collecting quantitative data" (Creswell & Plano-Clark, 2018, p. 94). "The exploratory sequential design theoretically starts with qualitative data collected from a small study group. In the light of the findings and results obtained by analyzing qualitative data, a quantitative measurement tool is developed, and this quantitative phase constitutes the second and final phase of the exploratory design" (Teddlie & Tashakkori, 2006, p. 21 cited in Özdemir, 2018, p. 495).

In this study, a performance task and a draft rubric were used to collect data. First of all, a performance task consisting

of four parts: description, task, instructions, and scoring was created to reflect the academic writing skills of international students and to test the validity and reliability of the draft DPA. The performance task was completed by 157 international students who were selected using criterion sampling, which is one of the purposive sampling methods. The basic understanding of criterion sampling, which is one of the types of purposive sampling that emerged within the qualitative research tradition and is defined as the most appropriate sampling method for qualitative research, is to study situations that meet a set of predetermined criteria (Yıldırım & Şimşek, 2011). The reason for choosing criterion sampling in the study is the idea of determining the participants according to predefined criteria. The international students within the scope of the study were expected to meet the criteria of having come to Turkey for higher education, having graduated from C1 level, and having taken an academic Turkish course. The other data collection tool, the draft "Analytical Rubric for International Students' Turkish Academic Writing Skills" was created by considering the steps suggested by various researchers (Popham, 1997; Goodrich-Andrade, 2000; Moskal, 2000; Mertler, 2001; Danielson & Hansen, 2016; Kutlu et al., 2017; Bıkmaz-Bilgen; 2023). These steps can be listed as follows:

- Determining the purpose of using the rubric
- Deciding on the type of rubric to be used
- Determining the criteria to be used in determining performance
- Determination of performance levels and level definitions
- Obtaining expert opinions
- Conducting a trial application
- Validity and reliability study
- Finalizing the rubric

Within the scope of the validity study of the draft DPA, construct validity was examined with the Davis (1992) technique in the context of the opinions of 10 experts, and construct validity was examined with exploratory factor analysis. As part of the reliability study, the Cronbach's Alpha coefficient, Intraclass Correlation and Fleiss Kappa agreement coefficient over the evaluations of six raters, and item-total correlations were calculated as evidence of item validity.

Discussion, Conclusion, and Suggestions

It is seen that the main criteria and sub-criteria of the DCI developed within the scope of the study have content validity indices between 0.80 and 1. These ratios indicate that all criteria and sub-criteria in the draft DPA are at a sufficient level in terms of content validity. (Davis Number ≥ 0.8). In the literature, it was found that Davis (1992) technique was used to determine the content validity indices of measurement tools developed for different target groups and learning domains (Bozkurt & Arıca-Akkök, 2019; Arıcı, 2022; İltar & Karataş, 2022).

In the context of a construct validity study, exploratory factor analysis (EFA) was applied to the analytical rubric developed for international students' Turkish academic writing skills over the data (performance task) obtained from 157 international students. In this context, firstly, in order to determine the suitability of the data for EFA, it was checked whether there was a multicollinearity problem among the sub-criteria of the DPA, and then the Skewness and Kurtosis coefficients were calculated for the normality control of the sub-criteria. The Skewness and Kurtosis coefficients between

-2 and +2 prove that the data are normally distributed (George & Mallery, 2022, p. 115). Multicollinearity occurs when the correlations between variables ($r > 0.90$ and above) are high (Çokluk et al., 2021, p. 210). In the calculations made, it was concluded that there was no multicollinearity problem since there was no correlation coefficient higher than 0.90 and that the Skewness and Kurtosis coefficients of the other sub-criteria, except for the "Visual Presentation" sub-criterion (Skewness value: 3.501; Kurtosis value: 11.759), were in accordance with the normal distribution since they varied between -2 and +2. Therefore, it was concluded that the "Visual Presentation" sub-criterion was not suitable for EFA, and it was removed from the measurement tool. In addition, before conducting EFA, Kaiser-Meyer-Olkin (KMO) was applied to determine the adequacy of the sample size for the criteria and the Bartlett Sphericity test was applied to determine the suitability of the relationship between the sub-criteria for factor analysis. Based on all these measurements, the fact that the KMO value of all criteria was higher than 0.60 (Büyüköztürk, 2018) and the data were suitable for multivariate normal distribution (Çokluk et al., 2021) proves the applicability of EFA to the measurement tool in question.

EFA was conducted using principal component analysis on the version of the data set suitable for applying the analysis in question. Factor loadings were calculated to determine the level of fulfillment of the main criterion in which the sub-criteria of the DPA are located. Eigenvalue findings were also examined to determine the factor structure. In addition, a line graph was also observed as an auxiliary method before deciding on the number of factors. As a result of the analysis, factor loadings of all sub-criteria of the DPA were found to be higher than 0.30, and it was determined that they measured the criterion well. Kline (1994) considers factor loadings of 0.30 and higher as sufficient. Similarly, Büyüköztürk (2002) defines the loadings between 0.30-0.59 as medium and 0.60 and above as high magnitudes. In addition, it was determined that there is only one component with an eigenvalue greater than 1 (Çokluk et al., 2021) in each criterion in the DPA, and this is also seen in the line graph. Therefore, it was concluded that the sub-criteria explain the related criteria in a single-factor structure. When the variances of the single component with initial eigenvalues greater than 1 were examined, it was observed that it explained 66.578% for the "Format" criterion, 78.613% for the "Academic Discourse" criterion, 66.560% for the "Text Structure" criterion, 61.067% for the "Planning and Organization" criterion, and 79.085% for the "Utilization of Resources" criterion. According to Alpar (2017), the variance explained in social sciences should be 60%. In this context, it was concluded that the variance values calculated in the current study were sufficient. In addition, since each criterion consisted of a single component as a result of the factor analysis, rotation was not required. In the exploratory factor analysis (EFA) conducted to test the construct validity of the instrument in the studies on rubric development in the literature (Öztürk & Kaplan, 2017; Yılmaz & İnceağaç, 2017; Aktaş & Alıcı, 2018; Yılmaz & Canpolat, 2020) similar steps were followed and values were examined.

In terms of the reliability of the analytical rubric developed for international students' Turkish academic writing skills, Cronbach's Alpha internal consistency reliability coefficient, inter-rater agreement, and item-total correlation values were examined. The Cronbach Alpha reliability coefficient of the developed analytical rubric was calculated in two separate ways, before and after EFA. As a result of these calculations,

it was determined that the total Cronbach Alpha reliability coefficient of the measurement tool was 0.953 for the version before EFA and 0.956 for the version after EFA. There are different opinions on the interpretation of the reliability coefficient in the literature. However, in most of the sources, a reliability coefficient of 0.70 and higher is accepted as sufficient for the reliability of the measurement tool (Nunnally, 1978; Büyüköztürk, 2018; Sönmez & Alacapınar, 2019). In this study, as a result of the calculations made regarding the Cronbach's Alpha reliability coefficient, it is seen that the reliability of the DPA is sufficient. In addition, in this study, six raters, including the researcher, evaluated 25 students' performance tasks in order to determine the inter-rater agreement in terms of reliability. In the analyses, Fleiss Kappa values were calculated since each sub-criterion was evaluated separately and by more than two raters; the Intra-Class Correlation coefficient was used since the main criteria were obtained by summing the sub-criteria. As a result of the analysis, inter-rater agreement levels vary between 0.637-0.909, with the lowest being "Comprehensibility" and the highest being "Keywords". Fleiss Kappa coefficient, which takes a value between -1 and +1, is interpreted as moderate agreement between 0.41-0.60, high agreement between 0.61-0.80, and very high agreement between 0.81-1 (Fleiss, 1971; Landis & Koch, 1977). Accordingly, the Fleiss Kappa values calculated for the sub-criteria of the DPA show that the inter-rater agreement is at high and very high levels. In addition, when the results of the intraclass correlation coefficient of the five criteria in the DPA are evaluated, it is seen that the inter-rater agreement levels vary between 0.909-0.980 in single measurements, the lowest belonging to the "Form" criterion and the highest belonging to the "Plan and Organization" criterion. In the average measurements, the lowest level of agreement belongs to the "Form" criterion and the highest to the "Plan and Organization" criterion, ranging between 0.984-0.997. The "Total" score value obtained from the DPA was calculated as 0.993 in single measurements and 0.999 in average measurements. In the literature, it is stated that the Intraclass Correlation Coefficient takes values between 0 and 1; values less than 0.5 indicate poor reliability, values between 0.5-0.75 indicate moderate reliability, values between 0.75-0.9 indicate good reliability, and values greater than 0.90 indicate excellent reliability (Koo & Li, 2016). Accordingly, it was determined that the intraclass correlation values calculated for all criteria and the total score of the DPA showed excellent agreement in single and average measurements. When all these findings were evaluated holistically, it was concluded that there was high and very high agreement between the raters. Similarly, it was determined that the methods used in various studies in the literature (Durmuş, 2022; Tarakcı, 2022) to determine the inter-rater agreement in the context of testing the reliability in the process of developing the IPA and the methods used in this study were parallel. Within the scope of this study, item-total correlations were calculated in order to determine whether the sub-criteria were compatible with the main criterion and total score as evidence of item validity. In this context, items with an item-total correlation of 0.30 and higher are recommended to be taken directly into the test because they can distinguish individuals well (Büyüköztürk, 2018, p. 183). As a result of the calculations made to determine the item-total correlation values, it was determined that all of the item-total correlations of the sub-criteria in the main criteria were higher than 0.30. In addition, the item-total

correlation values between the sub-criteria and the total score ranged between 0.568 and 0.833, with the lowest being "Language Knowledge" and the highest being "Development" sub-criteria. In rubric studies developed for different purposes in the literature (Polat-Demir, 2020; Bülbül & Bülbül, 2021) item-total correlation values were examined as evidence of item validity.

In this study, the final version of the analytical rubric developed for international students' Turkish academic writing skills as a result of validity and reliability studies consisted of 5 criteria and 24 sub-criteria. The criteria are listed as follows with their sub-criteria: "Style" (grammar, punctuation, spelling rules, comprehensibility, page layout), "Academic Discourse" (vocabulary, sentence structure, metadiscourse markers), "Text Structure" (topic, problem situation and purpose of the research, coherence, coherence, paragraph structure, synthesizing information), "Planning and Organizing" (title, summary, keywords, introduction, titling, development, discussion, conclusion and suggestions), "Making Use of Sources" (bibliography, direct and indirect quotation, academic ethical principles). When the literature was examined, it was determined that although there was no measurement tool developed based on the Turkish academic writing needs of international students, there are various types of measurement tools used to measure the relevant skills of different target groups (Aydın, 2015; Izawa et al., 2017; Shekarabi, 2017; Al-Abdullatif et al., 2020; Polat-Demir, 2020; Baz-Bolluk & Batur, 2023). While the criteria and sub-criteria in the DPA developed in the current study largely cover the criteria and sub-criteria in the aforementioned studies, they were also created by taking into account the specific characteristics of the target group of the study. In summary, it was concluded that the "Analytical Rubric for International Students' Academic Writing Skills" developed within the scope of this study is a valid and reliable tool that can measure the relevant skills of these students.

The performance-based "Analytical Rubric for International Students' Turkish Academic Writing Skills" developed within the scope of this study can be adapted for different academic text types, language levels, and teaching levels by making changes in some sub-criteria, and its validity and reliability can be tested again. Similarly, validity and reliability studies can be conducted on the existing measurement tool, and the results can be interpreted as to whether the mentioned RAT can be used in the evaluation of the academic writing skills of students studying in their mother tongue.

Author Contributions

This article is derived from a doctoral dissertation prepared by the first author under the supervision of the second author. All authors contributed equally to the writing of the article, read and approved the final version of the manuscript.

Ethical Declaration

This study was conducted with the approval decision of Istanbul University-Cerrahpaşa Social and Human Sciences Research Ethics Committee at its 2022/202 meeting dated 14.06.2022.

Conflict of Interest

The authors declare that there is no conflict of interest with any institution or person within the scope of this study.

EK: Uluslararası Öğrencilere Yönelik Türkçe Akademik Yazma Analitik Dereceli Puanlama Anahtarı

Ölçütler	Alt Ölçütler	Zayıf (1)	Geliştirilebilir (2)	Yeterli (3)	Yetkin (4)	Puan
Biçim (20)	Dil Bilgisi	Metinde en az 11 tür dil bilgil hata yapılmıştır.	Metinde 7 ila 10 tür dil bilgil hata yapılmıştır.	Metinde 3 ila 6 tür dil bilgil hata yapılmıştır.	Metinde en fazla 2 tür dil bilgil hata yapılmıştır.	
	Noktalama İşaretleri	Noktalama işaretleri kullanımında en az 8 tür hata yapılmıştır.	Noktalama işaretleri kullanımında 5 ila 7 tür hata yapılmıştır.	Noktalama işaretleri kullanımında 2 ila 4 tür hata yapılmıştır.	Noktalama işaretleri kullanımında en fazla 1 tür hata yapılmıştır.	
	Yazım Kuralları	Yazım kurallarında en az 8 tür hata yapılmıştır.	Yazım kurallarında 5 ila 7 tür hata yapılmıştır.	Yazım kurallarında 2 ila 4 tür hata yapılmıştır.	Yazım kurallarında en fazla 1 tür hata yapılmıştır.	
	Anlaşılrlık	Metnin en az 3/4'ünde anlam belirsizliğine yol açan dil kullanımları bulunmaktadır.	Metnin 2/4'ünde anlam belirsizliğine yol açan dil kullanımları bulunmaktadır.	Metnin 1/4'ünde anlam belirsizliğine yol açan dil kullanımları bulunmaktadır.	Metinde anlam belirsizliğine yol açan dil kullanımları bulunmamaktadır.	
	Sayfa Düzeni	Belirtilen formata uygun sayfa düzeni [kenar boşlukları, harf (punto, karakter), satır (hizalama, aralık), paragraf, sayfa numarası] ölçütlerinden en fazla 2'sine uyulmuştur.	Belirtilen formata uygun sayfa düzeni [kenar boşlukları, harf (punto, karakter), satır (hizalama, aralık), paragraf, sayfa numarası] ölçütlerinden 3 ya da 4'üne uyulmuştur.	Belirtilen formata uygun sayfa düzeni [kenar boşlukları, harf (punto, karakter), satır (hizalama, aralık), paragraf, sayfa numarası] ölçütlerinden 5 ya da 6'sına uyulmuştur.	Belirtilen formata uygun sayfa düzeni [kenar boşlukları, harf (punto, karakter), satır (hizalama, aralık), paragraf, sayfa numarası] ölçütlerinden tamamına uyulmuştur.	
Akademik Söylem (12)	Söz Varlığı	Sözcüklerin anlamı (somut, temel, terim) ve türü (eylem, isim sıfat vb.) en fazla 1/4 oranında akademik yazıya uygun bir biçimde kullanılmıştır.	Sözcüklerin anlamı (somut, temel, terim) ve türü (eylem, isim, sıfat vb.) 2/4 oranında akademik yazıya uygun bir biçimde kullanılmıştır.	Sözcüklerin anlamı (somut, temel, terim) ve türü (eylem, isim, sıfat vb.) 3/4 oranında akademik yazıya uygun bir biçimde kullanılmıştır.	Sözcüklerin anlamı (somut, temel, terim) ve türü (eylem, isim, sıfat vb.) metnin tamamında akademik yazıya uygun bir biçimde kullanılmıştır.	
	Cümle Yapısı	Eylem çatısı, eylem zamanı ve cümle türü maddeleri akademik yazıya uygun bir biçimde kullanılmamıştır.	Eylem çatısı, eylem zamanı ve cümle türü maddelerinden 1 tanesi akademik yazıya uygun bir biçimde kullanılmıştır.	Eylem çatısı, eylem zamanı ve cümle türü maddelerinden 2 tanesi akademik yazıya uygun bir biçimde kullanılmıştır.	Eylem çatısı, eylem zamanı ve cümle türü maddelerinden tamamı akademik yazıya uygun bir biçimde kullanılmıştır.	
	Üstsöylem Belirleyicileri*	Metnin algılanması bağlamında alıcıya gerekli olan gönderim ve vurgulama türlerinden en fazla 1'i kullanılmıştır.	Metnin algılanması bağlamında alıcıya gerekli olan gönderim ve vurgulama türlerinden 2'si kullanılmıştır.	Metnin algılanması bağlamında alıcıya gerekli olan gönderim ve vurgulama türlerinden 3'ü kullanılmıştır.	Metnin algılanması bağlamında alıcıya gerekli olan gönderim ve vurgulama türlerinden en az 4'ü kullanılmıştır.	
Metin Yapısı (24)	Konu	Konunun seçimi (alıcı, desteklenebilirlik, özgünlük, süre) ve sınırlandırılması bağlamındaki ölçütlerin en fazla 1/4'ü uygulanmıştır.	Konunun seçimi (alıcı, desteklenebilirlik, özgünlük, süre) ve sınırlandırılması bağlamındaki ölçütlerin 2/4'ü uygulanmıştır.	Konunun seçimi (alıcı, desteklenebilirlik, özgünlük, süre) ve sınırlandırılması bağlamındaki ölçütlerin 3/4'ü uygulanmıştır.	Konunun seçimi (alıcı, desteklenebilirlik, özgünlük, süre) ve sınırlandırılması bağlamındaki ölçütlerin tamamı uygulanmıştır.	

	Araştırmanın Problem Durumu ve Amacı	Araştırmanın problem durumu ve amacından en fazla 1'i ifade edilmiştir.	Araştırmanın problem durumu ve amacından her 2'si de ifade edilmiştir; ancak en az 1'i tam olarak ifade edilmemiştir.	Araştırmanın problem durumu ve amacından her ikisi de tam olarak ifade edilmiştir; ancak birbiriyle ilişkilendirilmemiştir.	Araştırmanın problem durumu ve amacı tam olarak ifade edilmiştir ve birbiriyle ilişkilendirilmiştir.	
	Bağdaşıklık	Dilsel, dil bilgisel uyum metnin en fazla 1/4'ünde sağlanmıştır.	Dilsel, dil bilgisel uyum metnin 2/4'ünde sağlanmıştır.	Dilsel, dil bilgisel uyum metnin 3/4'ünde sağlanmıştır.	Dilsel, dil bilgisel uyum metnin tamamında sağlanmıştır.	
	Tutarlılık	Metin, içerik ve düzen bağlamında en fazla 1/4 oranında tutarlıdır.	Metin, içerik ve düzen bağlamında 2/4 oranında tutarlıdır.	Metin, içerik ve düzen bağlamında 3/4 oranında tutarlıdır.	Metnin tamamı içerik ve düzen bağlamında tutarlıdır.	
	Paragraf Yapısı	Paragraflarda konu, destekleme, bitiş cümlelerinden hiçbiri doğru ve yerinde kullanılmamıştır.	Paragraflarda konu, destekleme, bitiş cümlelerinden sadece 1'i doğru ve yerinde kullanılmıştır.	Paragraflarda konu, destekleme, bitiş cümlelerinden 2'si doğru ve yerinde kullanılmıştır.	Paragraflarda konu, destekleme, bitiş cümlelerinin tamamı doğru ve yerinde kullanılmıştır.	
	Bilgileri Sentezleme	Derlenen bilgilerin en fazla 1/4'ü araştırma konusunu destekleyici şekilde bütünleştirilmiştir.	Derlenen bilgilerin 2/4'ü araştırma konusunu destekleyici şekilde bütünleştirilmiştir.	Derlenen bilgilerin 3/4'ü araştırma konusunu destekleyici şekilde bütünleştirilmiştir.	Derlenen bilgilerin tamamı araştırma konusunu destekleyici şekilde bütünleştirilmiştir.	
	Görsel Sunum**	Gerekli yerlerde grafik, tablo, şekil vb. gibi görseller kullanılmamıştır.	Grafik, tablo, şekil vb. gibi görseller amaca uygun olmayan/gereksiz bir şekilde kullanılmıştır.	Grafik, tablo, şekil vb. gibi görseller amaca uygun olarak kullanılmıştır; ancak kolayca anlaşılır değildir.	Grafik, tablo, şekil vb. gibi görseller amaca uygun ve kolayca anlaşılır bir şekilde kullanılmıştır.	
Plan ve Düzenleme (28)	Başlık	Çalışmaya başlık konulmamıştır.	Çalışmaya başlık konulmuştur, ama başlıkta biçim [sözcük (tür, sayı) vb.] ve içerik (metne uygunluk, ifade vb.) özelliklerine uyulmamıştır.	Çalışmanın başlığında biçim [sözcük (tür, sayı) vb.] ve içerik (metne uygunluk, ifade vb.) özelliklerinden yalnızca 1'ine uyulmuştur.	Çalışmanın başlığında biçim [sözcük (tür, sayı) vb.] ve içerik (metne uygunluk, ifade vb.) özelliklerinin tamamına uyulmuştur.	
	Özet	Çalışmada özet yoktur.	Araştırmanın amacı, süreci (yöntem), sonucu gibi bilgilerden en fazla 1 tanesine yer verilmiştir.	Araştırmanın amacı, süreci (yöntem), sonucu gibi bilgilerden 2 tanesine yer verilmiştir.	Araştırmanın amacı, süreci (yöntem), sonucu gibi bilgilerden tamamına yer verilmiştir.	
	Anahtar Sözcükler	Araştırmanın amacına ve problem durumuna uygun en fazla 1 anahtar sözcük belirlenmiştir.	Araştırmanın amacına ve problem durumuna uygun 2 anahtar sözcük belirlenmiştir.	Araştırmanın amacına ve problem durumuna uygun 3 anahtar sözcük belirlenmiştir.	Araştırmanın amacına ve problem durumuna uygun en az 4 anahtar sözcük belirlenmiştir.	
	Giriş	Problem durumu, alanyazın taraması ve çalışma içeriği bileşenlerinden - konu, önem ve tanıtım- en fazla 1'i yansıtılmıştır.	Problem durumu, alanyazın taraması ve çalışma içeriği bileşenlerinden - konu, önem ve tanıtım- 2'si yansıtılmıştır.	Problem durumu, alanyazın taraması ve çalışma içeriği bileşenlerinden - konu, önem ve tanıtım- 3'ü yansıtılmıştır.	Problem durumu, alanyazın taraması ve çalışma içeriği bileşenlerinden - konu, önem ve tanıtım- tamamı yansıtılmıştır.	
	Başlıklandırma	Ana başlık ya da alt başlıkların en fazla 1/4'ü kurallara uygun olarak yazılmıştır.	Ana başlık ve alt başlıkların 2/4'ü kurallara uygun olarak yazılmıştır.	Ana başlık ve alt başlıkların 3/4'ü kurallara uygun olarak yazılmıştır.	Ana başlık ve alt başlıkların tamamı kurallara uygun olarak yazılmıştır.	
	Gelişme	Bütünlük, destek, detay ve tutarlılık boyutlarından en fazla 1'i gerçekleştirilmiştir.	Bütünlük, destek, detay ve tutarlılık boyutlarından 2'si gerçekleştirilmiştir.	Bütünlük, destek, detay ve tutarlılık boyutlarından 3'ü gerçekleştirilmiştir.	Bütünlük, destek, detay ve tutarlılık boyutlarından tamamı gerçekleştirilmiştir.	

	Tartışma, Sonuç ve Öneriler	Araştırma konusunu özetleme, çalışmanın sonuçlarını alanyazındaki benzer çalışmalarla yorumlama, yeni çalışma önerileri sunma maddelerinden hiçbirine yer verilmemiştir.	Araştırma konusunu özetleme, çalışmanın sonuçlarını alanyazındaki benzer çalışmalarla yorumlama, yeni çalışma önerileri sunma maddelerinden 1'ine yer verilmiştir.	Araştırma konusunu özetleme, çalışmanın sonuçlarını alanyazındaki benzer çalışmalarla yorumlama, yeni çalışma önerileri sunma maddelerinden 2'sine yer verilmiştir.	Araştırma konusunu özetleme, çalışmanın sonuçlarını alanyazındaki benzer çalışmalarla yorumlama, yeni çalışma önerileri sunma maddelerinden tamamına yer verilmiştir.		
Kaynaklardan Yararlanma (12)	Kaynakça	Alıntılanan kaynaklar belirtilmemiştir ya da en fazla 1/4'ü belirtilen formata uygun şekilde kaynakçada doğru gösterilmiştir.	Alıntılanan kaynakların 2/4'ü belirtilen formata uygun şekilde kaynakçada doğru gösterilmiştir.	Alıntılanan kaynakların 3/4'ü belirtilen formata uygun şekilde kaynakçada doğru gösterilmiştir.	Alıntılanan kaynakların tamamı belirtilen formata uygun şekilde kaynakçada doğru gösterilmiştir.		
	Doğrudan ve Dolaylı Alıntı	Alıntılar hiç gösterilmemiştir ya da kaynakların en fazla 1/4'ü belirtilen format ve alıntılama türüne uygun şekilde metin içinde gösterilmiştir.	Kaynakların 2/4'ü belirtilen format ve alıntılama türüne uygun şekilde metin içinde gösterilmiştir.	Kaynakların 3/4'ü belirtilen format ve alıntılama türüne uygun şekilde metin içinde gösterilmiştir.	Kaynakların tamamı belirtilen format ve alıntılama türüne uygun şekilde metin içinde gösterilmiştir.		
	Akademik Etik İlkeler	İntihal yapma, gerçekte olmayan verileri kullanma, verileri değiştirme vb. gibi etik dışı davranışlardan en az 3'ü gözlenmiştir.	İntihal yapma, gerçekte olmayan verileri kullanma, verileri değiştirme gibi etik dışı davranışlardan 2'si gözlenmiştir.	İntihal yapma, gerçekte olmayan verileri kullanma, verileri değiştirme gibi etik dışı davranışlardan 1'i gözlenmiştir.	İntihal yapma, gerçekte olmayan verileri kullanma, verileri değiştirme gibi etik dışı davranışların tamamından kaçınılmıştır.		
TOPLAM BAŞARI PUANI ***							

***Kaçınmalar:** Örnek: Belki, gibi görünmektedir, söylenebilir / **Vurgulayıcılar:** Örnek: Şüphesiz, kesinlikle / **Tutum Belirleyicileri:** Örnek: Dikkat çekicidir, önemlidir / **İlişki/Bağlantı Belirleyicileri:** Örnek: Tablo 3'e bakınız, görebiliriz / **Kendini Anma:** Örnek: Çalışmamda, gözlemledim **Bağlayıcılar:** Örnek: Ama, bu yüzden, buna ek olarak / **Çerçeve Belirleyicileri:** Örnek: Sonuç olarak, bu bölümde ... ele alınacaktır / **Metiniçi Belirleyiciler:** Örnek: Tablo 3'te olduğu gibi, 25. örnek / **Tanıtlayıcılar:** Örnek: Hyland (2005) / **Kod Çözümleyicileri:** Örnek: Örneğin, bir başka deyişle.

**Görsel sunum alt ölçütü puanlamaya dâhil edilmemiştir.

*** Puan aralığı; 24-41 (zayıf), 42-59 (geliştirilebilir), 60-77 (yeterli), 78-96 (yetkin) olarak hesaplanmıştır.