Yazıcı, T. ve Kartal, O.Y. (2026). İçsel konuşma ölçeği Türkçe dil uyarlaması: Geçerlilik ve güvenirlik çalışması. *Ana Dili Eğitimi Dergisi*, *14*(4), 990-1011. https://doi.org/10.16916/aded.2007291

Ana Dili Eğitimi Dergisi
Journal of Mother Tongue Education
www.anadiliegitimi.com

Geliş/Received: 30.07.2026 Kabul/Accepted: 08.10.2026
Araştırma Makalesi / Research Article

İçsel Konuşma Ölçeği Türkçe Dil Uyarlaması: Geçerlilik ve Güvenirlik Çalışması

Tuğçe YAZICI[*]
Osman Yılmaz KARTAL[**]

Özet

Bu araştırmanın amacı, öğrenenlerin bilişsel ve metabilişsel süreçlerinde önemli bir rol oynayan içsel konuşmayı ölçmek amacıyla Alderson-Day vd. (2018) tarafından geliştirilen İçsel Konuşma ölçeğinin "Varieties of Inner Speech Questionnaire- Revised (VISQ-R)" Türkçeye uyarlanmasıdır. Çalışma, Uluslararası Test Komisyonu (ITC) yönergeleri temel alınarak 9 aşamada (çeviri, uzman görüşü, dilsel eşdeğerlik vb.) gerçekleştirilmiştir. Araştırmanın ana saha uygulaması için ölçek eğitim fakültesinde öğrenim gören 323 öğretmen adayı ile yürütülmüştür. Ölçeğin yapı geçerliliği Doğrulayıcı Faktör Analizi (DFA) ile, güvenirliği ise Cronbach Alfa (α) ve McDonald's Omega (ω) iç tutarlılık katsayıları ile incelenmiştir. DFA sonuçları, ölçeğin orijinal formundaki beş faktörlü ve 25 maddelik nihai yapısının (Diyalojik, Değerlendirici/Eleştirel, Diğer Kişiler, Sıkıştırılmış, Pozitif/Düzenleyici) Türk kültüründe doğrulandığını, modelin kabul edilebilir uyum değerlerine ($\chi^2/df = 2.87$, RMSEA = .076) sahip olduğunu göstermektedir. Ölçeğin genel iç tutarlılık katsayıları α = .836 ve ω = .839 olarak hesaplanmış olup, alt boyutlara ait katsayılar .691 ile .865 (α: .691-.863; ω: .695-.865) arasında değişmektedir. Sonuç olarak, İçsel Konuşma Ölçeği Türkçe Formu'nun eğitim araştırmalarında eğitimciler üzerinde uygulanılabilecek geçerli, güvenilir ve pedagojik bir ölçme aracı olduğu tespit edilmiştir.

Anahtar Kelimeler: İçsel konuşma, ölçek uyarlama, geçerlik ve güvenirlik, doğrulayıcı faktör analizi, öğretmen adayları

Turkish Adaptation of The Inner Speech Scale: Validity and Reliability Analysis

Abstract

This study aimed to adapt the Varieties of Inner Speech Questionnaire-Revised (VISQ-R), developed by Alderson-Day et al. (2018) to measure learners' cognitive and metacognitive inner speech, into Turkish. Following International Test Commission (ITC) guidelines, the study was conducted in nine phases (translation, expert view, linguistic equivalence, etc.). The main field implementation included 323 preservice teachers in a faculty of education. Construct validity was examined using Confirmatory Factor Analysis (CFA), and reliability was assessed via Cronbach's alpha (α) and McDonald's Omega (ω) internal consistency coefficients. CFA results supported the five-factor, 25-item structure (Dialogic, Evaluative/Critical, Other people, Condensed, and Positive/Regulatory) in Turkish culture, showing acceptable model fit ($\chi^2/df = 2.87$, RMSEA = .076). Overall internal consistency was α = .836 and ω = .839, with subscales ranging from .691 to .865 (α: .691-.863; ω: .695-.865). In conclusion, the Turkish VISQ-R was proven to be a valid, reliable, and pedagogical measurement tool that can be administered to educators in educational research.

Keywords: Inner speech, scale adaptation, validity and reliability, confirmation factor analysis, preservice teachers

[*] Dr., Çanakkale, yazicituce@windowslive.com
[**] Prof. Dr., Çanakkale Onsekiz Mart Üniversitesi, Eğitim Fakültesi, Eğitim Bilimleri Bölümü, Çanakkale, osmanykartal@comu.edu.tr

Giriş

İçsel konuşma (inner speech), en genel tanımıyla öğrenenin içsel olarak, kendi kendine gerçekleştirdiği zihinsel dil kullanım sürecidir. Vygotsky'nin (1962) sosyokültürel teorisinden bu yana dil ve düşünce arasındaki etkileşimin temel mekanizması olarak kabul edilen bu süreç yalnızca pasif bir düşünme eylemiyle sınırlı değildir. Aksine içsel konuşma, öğrenme, problem çözme, öz düzenleme ve motivasyon gibi temel eğitimsel işlevlerin yerine getirilmesinde etkin bir rol oynamaktadır (Alderson-Day ve Fernyhough, 2015; Chella ve Pipitone, 2020; Morin, 2005). Bilişsel yükü azaltarak üst düzey bilişsel ve metabilişsel faaliyetlerin yürütülmesini destekleyen bu mekanizma (Kompa ve Mueller, 2024) ile öğrenen zihnindeki sesleri dinlemekte, kelimeleri sessizce tekrar etmekte ve bu yolla bilişsel süreçlerini yönetmektedir (de Guerrero, 2018). Özellikle muhakeme, problem çözme ve yaratıcı düşünme süreçlerinin yanı sıra öğrenenin keşfedici yönünün gelişiminde de içsel konuşmanın diyalojik yapısı belirleyici bir bilişsel araç işlevi görmektedir (Kompa ve Mueller, 2024).

Eğitimsel bağlamda içsel konuşma, öğretim etkinlikleri aracılığıyla edinilen bilginin içselleştirilmesi ve dışsallaştırılmasında köprü görevi görmektedir (Fossa, 2022). Öğrenenin kendi kendine konuşması öğrenme görevlerine katılımını artırmakta ve sosyal süreçleri düzenleme işlevi görerek bilişsel kapasitenin gelişimine araçsal bir katkı sunmaktadır (Flanagan ve Symonds, 2022). Nitekim alan yazın içsel konuşmayı destekleyen öğretimsel tekniklerin, öğrenenin performansını ve zihinsel becerilerini pekiştirdiğini ortaya koymaktadır (Hidayat vd., 2023). Bu durum nitelikli bir öğrenimin gerçekleşmesi adına eğitim bilimleri literatürünün önemle üzerinde durduğu metabilişsel farkındalık ve öz düzenleme gibi parametrelerin gelişiminde içsel konuşmanın kritik bir değişken olduğunu göstermektedir.

Eğitim bilimleri ve pedagojik açıdan içsel konuşmanın sınıf içi öğrenme ortamlarında motor öğrenme süreçlerinde ve problem çözme etkinliklerinde üst düzey bilişsel işlevleri harekete geçirdiği bilinmektedir (Dana ve Brinthaupt, 2025; Dana vd., 2024; Landin ve Hebert, 1999). Pan vd.'nin (2025) teknoloji destekli etkileşimli sınıflarda yaptıkları güncel çalışmasında görüldüğü üzere içsel konuşma öğrencilerin metabilişsel ve öz-düzenleme yeterliliklerinin geliştirilmesinde güçlü bir düzenleyici rol oynamaktadır. Sınıf içi doğal ortamlarda çocukların ve gençlerin kendi kendilerine konuşma stratejilerini hem motivasyonel hem de bilişsel bir düzenleyici olarak kullandıkları, özellikle çaba odaklı olumlu içsel konuşmanın nitelikli öğrenmeye ve akademik başarıya katkı sunduğu; buna karşın olumsuz içsel konuşmanın öğrenenin kaygı düzeyini artırdığı vurgulanmaktadır (Flanagan ve Symonds, 2022; Thomaes vd., 2020). Bu doğrultuda Türkiye eğitim sisteminde görev yapacak olan öğretmenlerin hem kendi hem de öğrencilerinin içsel konuşma çeşitliliklerinin ve seviyelerinin belirlenmesi onların pedagojik gelişimleri ve gelecekte öğrencilerine aktaracakları öz-düzenleme stratejileri açısından stratejik bir öneme sahiptir (Xiang vd., 2019).

Eğitim süreçlerindeki bu teorik önemine rağmen içsel konuşmanın doğasını ve işlevlerini eğitim bilimleri perspektifiyle derinlemesine inceleyecek geçerli ve güvenilir ölçme araçlarının eksikliği dikkat çekmektedir. Mevcut literatür incelendiğinde içsel konuşmaya yönelik geliştirilen ölçme araçlarının ve yapılan araştırmaların ağırlıklı olarak halüsinasyon eğilimi ve anksiyete gibi değişkenlerin incelendiği klinik psikoloji (Alderson-Day, Mitrenga, Wilkinson, McCarthy-Jones ve Fernyhough, 2018; McCarthy-Jones ve Fernyhough, 2011; Racy, Morin ve Duhnych, 2020; Racy, Morin ve Hagerty, 2024; Uttl, Morin ve Hamper, 2011), bilişsel nörobilim (Alderson-Day ve Fernyhough, 2015; Brinthaupt, Hein ve Kramer, 2009) ve spor bilimleri (Dana, Ghorbani ve Brinthaupt, 2024; Dana ve Brinthaupt, 2025; Latinjak, Viladrich, Alcaraz ve Torregrosa, 2016; Zervas, Stavrou ve Psychountaki, 2007; Zourbanos, Hatzigeorgiadis, Chroni, Theodorakis ve Papaioannou, 2009) alanlarında yoğunlaştığı görülmektedir. Bu bağlamda öğrenme süreçlerini merkeze alan doğrudan çalışmaların sınırlı olması (Xiang, Brinthaupt, Sun ve Ren, 2019) bu alana özgü parametrelerde içsel konuşmanın rolünün anlaşılmasını zorlaştırmaktadır. Ayrıca yurtiçinde yapılan araştırmaların da ağırlıklı olarak spor bilimleri ve performans odaklı olduğu öğretmen eğitimi ve yetiştirme süreçlerini temel alan ampirik çalışmaların gerçekleştirilmediği görülmektedir (Ekizoğlu, Acet ve Toros, 2024; Orhan, 2024). Oysaki eğitimsel bağlamda içsel konuşma, yalnızca bilginin içselleştirilmesi değil, aynı zamanda öğretmenin sınıf içi pratiklerini izlediği ve düzenlediği yansıtıcı bir diyalog işlevi görmektedir (Fossa, 2022). Nitekim öğretmen adaylarının karmaşık sınıf dinamikleri

içerisinde sergiledikleri içsel konuşma stratejileri öğretimsel kararlarını ve öğrenciye yönelik tutumlarını doğrudan etkileyen metabilişsel bir süreçtir (Flanagan ve Symonds, 2022; Fossa, 2022; Werani, 2018). Bu nedenle söz konusu sürecin eğitimsel bağlamda geçerli ve güvenilir araçlarla ölçülmesi öğretmen eğitimi programlarında öz düzenleme becerilerinin geliştirilmesi adına önemli bir gereksinimdir.

Sonuç olarak bu çalışma öğretmen yetiştirme literatüründe içsel konuşmaya yönelik araştırmaların sınırlılığını aşmaya ve bu olgunun öğretmen eğitimindeki rolünü somutlaştırmaya odaklanmaktadır. Bu bağlamda araştırma Alderson-Day vd. (2018) tarafından psikometrik özellikleri güçlendirilen "Varieties of Inner Speech Questionnaire - Revised (VISQ-R)" ölçme aracını Türkçeye uyarlamayı amaçlamaktadır. Çalışma içsel konuşmayı yalnızca klinik bir tanı aracı olmaktan çıkarıp öğretmen adaylarının öğrenme ve bilişsel gelişim süreçlerinde kullanılabilir pedagojik bir göstergeye dönüştürmesi ve eğitim bilimleri alan yazınına psikometrik nitelikleri kanıtlanmış bir ölçme aracı kazandırması açısından önem taşımaktadır.

İçsel Konuşmanın Kavramsal Çözümlenmesi

İçsel konuşma, öğrenenin açık ve duyulabilir bir şekilde kelimeleri telaffuz etmediği durumda öznel olarak dilsel süreçlerin gerçekleşme deneyimidir (Alderson-Day ve Fernyhough, 2015). Genel olarak öz konuşma olarak da ifade edilir. Öğrenenin bilişsel bir ürünü olmasının yanında içsel konuşma, öğrenenin kendisi ile çevresi hakkındaki tutum, inanç ve düşüncelerini yansıtmaktadır (Chen, 2012). Ayrıca içsel konuşma, öğrenenin kendisiyle diyalog kurması, kendisine bir şeyler söylemesi veya dile getirmesi (Langland-Hassan, 2014), sessizlik içerisinde kendi kendine konuşması şeklinde açıklanmaktadır (Morin, Duhnych ve Racy, 2018). Bu tanımlamalar içsel konuşmanın öğrenenin bilişsel faktörleri harekete geçirmesi, öğrenenin önemli bilişsel işlevlerini yerine getirmesi açısından önemini belirtmektedir.

Alderson-Day ve Fernyhough (2015), içsel konuşmaya yönelik gerçekleştirdikleri kapsamlı literatür derlemesinde bu sürecin çok boyutlu temel faktörlerden oluştuğunu kuramsal olarak ortaya koymuşlardır. Bu faktörler, daha sonra Alderson-Day vd. (2018) tarafından revize edilerek geliştirilen ölçme aracının boyutlarını oluşturmuştur. Bunlar sırasıyla; Diyalojik içsel konuşma, öğrenenin karşılıklı konuşma niteliğinde içsel konuşmaya girme eğilimini; Sıkıştırılmış (Condensed) içsel konuşma, konuşan kişinin zihinsel mesajı açıkça bildiği durumlarda, gramer kurallarını ve tam cümle yapılarını göz ardı ederek düşüncelerini tek sözcükler veya kestirme sembolik özetler hâlinde deneyimlemesi; Diğer insanlar (other people), içsel konuşmanın genellikle başkalarının seslerinin veya başka birinin söyleyeceği bir şeyi söylemesi şeklinde gerçekleşmesi; Değerlendirici içsel konuşma, öğrenenin içsel konuşmasının kendi davranışlarını yargılamak veya değerlendirmek amaçlı kullanılması; Pozitif\Düzenleyici iç konuşma, öğrenenin motivasyonunu artırmak, performansını desteklemek ya da sakinleşmek amacıyla kullandığı içsel konuşma türü olarak tanımlanmıştır.

Yöntem

Araştırmanın Modeli[1]

Ölçek uyarlama ve geçerlik-güvenirlik çalışmaları, doğası gereği tipik kesitsel (cross-sectional) araştırmalardan farklılaşan, kendine özgü istatistiksel ve prosedürel adımları olan metodolojik bir çalışma deseni gerektirmektedir (Arafat, 2016). Bu doğrultuda mevcut araştırma, bir ölçme aracının kültürlerarası uyarlamasında "altın standart" olarak kabul edilen Uluslararası Test Komisyonu (International Test Commission [ITC], 2005) yönergelerini temel alan metodolojik bir araştırma olarak tasarlanmıştır. Bu kapsamda sadece dilsel çeviri ile yetinilmemiştir. Aynı zamanda anlamsal, deyimsel ve kavramsal eşdeğerliği sağlamak adına kapsamlı bir kültürel uyarlama protokolü gerçekleştirilmiştir (Byrne, 2016; Gjersing, Caplehorn ve Clausen, 2010).

[1] Bu bir ölçek uyarlama çalışması olması nedeniyle kesitsel tarzdaki araştırma raporlaştırılmasından farklı olarak yöntem başlığı altında olması gereken ''Örneklem\ Araştırma grubu, veri toplama, veri analizi gibi alt başlıklar ölçek uyarlama protokolü aşamalarında detaylı olarak sunulmuştur.

İfade edilen bu metodolojik adımlar doğrultusunda, araştırmanın amacına bağlı olarak çalışmada veri toplama aracı olarak Alderson-Day ve arkadaşları (2018) tarafından geliştirilen "Varieties of Inner Speech Questionnaire – Revised (VISQ-R)" ölçeği kullanılmıştır. Psikometrik özellikleri güçlendirilmiş olan bu orijinal ölçek; diyalojik, sıkıştırılmış, diğer kişiler, değerlendirici/eleştirel ve pozitif/düzenleyici olmak üzere beş faktörlü bir yapı sergilemektedir. Ölçeğin orijinal formunun yapı geçerliliğine yönelik yapılan Açımlayıcı Faktör Analizi (AFA) sonucunda, toplam varyansın %68.43'ünü açıklayan 26 maddelik kararlı bir model elde edilmiştir. Bu yapıyı doğrulamak amacıyla uygulanan Doğrulayıcı Faktör Analizi (DFA) sonuçları ise kabul edilebilir uyum indeksleri (CMIN/df = 2.16, CFI = 0.925, RMSEA = 0.056, Hoelter indeksi= 200) ile modelin verilerle uyumlu olduğunu kanıtlamıştır (Alderson-Day vd., 2018). Ölçeğin alt boyutlarına ilişkin iç tutarlılık katsayıları; diyalojik faktör için .87, değerlendirici/eleştirel için .88, pozitif/düzenleyici için .80, sıkıştırılmış için .87 ve diğer kişiler faktörü için .91 şeklinde raporlanmıştır. Maddeler, deneyimlerin sıklığını belirlemek üzere "Hiçbir Zaman" (1), "Neredeyse Hiç" (2), "Nadiren" (3), "Bazen" (4), "Sık Sık" (5), "Oldukça Sık" (6) ve "Her zaman" (7) arasında derecelendirilen 7'li Likert tipi bir ölçek yapısına sahiptir.

Alan yazında geçerlik ve güvenirliği bu şekilde ortaya konmuş olan ölçme aracının orijinal İngilizce formundan Türkiye'nin eğitim anlayışı bağlamına taşınması sürecinde, bu araştırma kapsamında kültürlerarası uyarlamanın gerektirdiği psikometrik ve metodolojik standartlar bir dizi aşamalı protokol ile güvence altına alınmıştır (Arafat, 2016; Boateng vd., 2018). Bu protokollerin kurgulanmasında çeviri süreçlerinden başlayarak kavramsal, anlamsal ve deneyimsel eşdeğerliğin sağlanması uzman heyetince kültürel incelemelerin yapılması ve nihayetinde pilot çalışmalarla maddelerin bilişsel olarak anlaşılırlığının test edilmesi (Protokol 1-7) hedeflenmiştir (Byrne, 2016; Gjersing vd., 2010). Bu aşamalar, ölçeğin hedef kültürdeki yapısal geçerliğini korumak adına atılması gereken zorunlu adımları temsil etmektedir. Sürecin sonunda yer alan Protokol 8 ve 9 ise, elde edilen verilerin psikometrik analizlerle (Doğrulayıcı Faktör Analizi) doğrulanması, ölçeğin yapısal bütünlüğünün Türkiye'deki öğretmen adayı örnekleminde istatistiksel olarak kanıtlanması ve nihai formun alana kazandırılması için zorunlu bir saha ve analiz süreci teşkil etmektedir (Baharum vd., 2023; Heggestad vd., 2019).

Protokol 1 Aşaması: Hazırlık ve İzinler

Bu aşamada Alderson-Day vd. (2018) tarafından geliştirilen Inner Speech ölçme aracına yönelik dil uyarlama çalışması yapılacağına ilişkin e-posta aracılığı ile yazılı izin alınmıştır. Sosyal ve Beşerî Bilimler Etik Kurulu Komisyonu'nun almış olduğu 24.10.2025 tarih ve E-80110942-010.99-2500281178 sayılı kararı ile çalışmanın yürütülmesine ilişkin onay alınmıştır. Ayrıca Çanakkale Onsekiz Mart Üniversitesi Eğitim Fakültesi Dekanlığından 05.11.2025 tarihli ve E-34954581-903.99-2500294204 sayılı yazı ile araştırma izni alınmıştır.

Protokol 2 Aşaması: İleri Çeviri

Ölçeğin birbirlerinden bağımsız çalışan iki farklı çevirmen tarafından Türkçeye çevrildiği aşamadır. Çevirmen A, Psikoloji\PDR alanında lisans, Eğitim Programları ve Öğretim Bilim dalında yüksek lisans mezunu bir uzmandır. Kamusal ve özel sektörde en az 5 yıllık deneyime sahiptir. Çevirmen B, alan dışı dil uzmanıdır ve Lisansını İngiliz Dili Eğitimi bölümünde tamamlamıştır ve dilsel çeviriyi gerçekleştirmiştir. Çevirmen B alan dışı dil uzmanı olarak İngiliz Dili Eğitimi Bilim dalında doktorasını yapmış, Mili Eğitim Bakanlığında 15 senelik deneyime sahiptir. Çevirmen A ile Çevirmen B tarafından gerçekleştirilen çalışma sonucunda ölçme aracına ilişkin iki bağımsız Türkçe çeviri metin (T1 ve T2) elde edilmiştir.

Protokol 3 Aşaması: Sentez (Uzlaşı) Aşaması

Bu aşamada araştırma ekibi ve çevirmenler (Çevirmen 1: İngiliz Dili Eğitimi Doktorantı, Çevirmen 2: İngilizce Öğretmeni, EPÖ Bilim Uzmanı) bir araya gelerek T1 ve T2 formlarını karşılaştırmıştır. Çevirmen 2, en az 3 senelik kamusal alanda deneyime sahip dil uzmanıdır. Uzmanlar

kavramsal uygunluk ve dilsel akıcılık açısından en uygun ifadeleri seçerek ortaklaştırılmış tek bir Taslak Türkçe Form (T-1) oluşturmuştur.

Protokol 4 Aşaması: Geri Çeviri

T-1 formu, orijinal ölçeği hiç görmemiş ve anadili İngilizce olan veya dil hakimiyeti ana dil düzeyinde olan ve İngiliz Dili Eğitiminde Doktora derecesini almış, minimum 5 senelik mesleki deneyime sahip iki bağımsız uzman tarafından tekrar İngilizceye çevrilmiştir. Orijinal form ile geri çevrilen formlar karşılaştırılarak anlam kaymaları kontrol edilmiştir. Uzmanlar formlar arasında anlam kaymaları olmadığını belirterek bu aşamada T-1 formunun çeviri tutarlılığını doğrulamışlardır.

Protokol 5 Aşaması: Kapsam Geçerliliğine Yönelik Uzman Değerlendirmesi

Hazırlanan Türkçe form, kapsam geçerliliğinin sağlanması amacıyla metodologlar, dil uzmanları ve eğitim bilimleri alanından uzmanların yer aldığı disiplinler arası bir uzman heyetine sunulmuştur. Gjersing vd. (2010), kültürel uyum ve kapsam geçerliliği sürecinde metodologlar, dil uzmanları ve ilgili alan uzmanlarından oluşan bir panelin katılımını, ölçme aracının psikometrik ve kültürel bütünlüğünü korumak adına kritik bir zorunluluk olarak tanımlamaktadır. Uzmanlara ait özellikler şu şekildedir:

Uzman 1, Eğitim Bilimleri, Program Geliştirme ve özellikle Eğitimde Ölçme ve Değerlendirme alanlarında uzmanlaşmış kıdemli bir araştırmacıdır. Ölçek maddelerinin psikometrik nitelikleri, ölçme ilkelerine uygunluğu ve kültürel adaptasyon sürecindeki kavramsal eşdeğerliği boyutlarında metodolojik değerlendirmeler sunmuştur.

Uzman 2, Eğitim Programları ve Öğretim Ana Bilim Dalı'nda öğretim üyesi olup uzmanlık alanı öğretmen eğitimi, öğretim tasarımı ve eğitim bilimleridir. Ölçek maddelerinin pedagojik çerçevesi, maddelerin sosyolojik bağlamda kültürel uyumu ile öğretmen yetiştirme dinamiklerine kavramsal uygunluğunun değerlendirilmesine katkı sağlamıştır.

Uzman 3, Temel Eğitim Bölümü Sınıf Eğitimi Ana Bilim Dalı öğretim üyesi olup Eğitim Programları ve Öğretim ile ilköğretim alanlarında geniş bir araştırma ve yayın birikimine sahiptir. Uyarlanan formun hedef kitlenin gelişimsel ve öğretimsel düzeyine uygunluğu ile maddelerin içerik geçerliliğinin denetlenmesinde görev almıştır.

Uzman 4, Eğitim Programları ve Öğretim Ana Bilim Dalı öğretim üyesi olup düşünme becerileri, Sokratik sorgulama ve öğretmen rolleri üzerine çalışmalar yürütmektedir. İçsel konuşmanın bilişsel ve metabilişsel yapısını oluşturan maddelerin kuramsal derinliği ve ifade akıcılığı yönünden incelenmesinde uzman görüşü sağlamıştır.

Bu aşamada uzmanlar, formdaki maddeleri hedef kültüre uygunluk, kavramsal eşdeğerlik ve dilsel anlaşılırlık boyutlarında nitel olarak değerlendirmiş ve bağımsız geri bildirimler sunmuştur. Uzmanlardan gelen dönütler araştırmacılar tarafından bir araya getirilerek incelenmiştir. Maddeler üzerinde uzmanların beraber uzlaştığı noktalar belirlenmiştir. Genel olarak uzlaşı durumu formdaki maddelerin hedef kitleye ve kültürel yapıya uyumlu olduğu, ölçek maddelerinin kavramsal eşdeğerlik ve dilsel anlaşılırlık açısından geçerli olduğunu beyan etmişlerdir. Yapılan değerlendirmeler sonucunda uzman görüşleri yüksek oranda birbiriyle örtüşmüştür. Sunulan sözel ve niteleyici öneriler doğrultusunda gerekli semantik ve yapısal düzeltmeler yapılmıştır. Bu bağlamda kapsam geçerliliği güvence altına alınmış "Deneme Formu" (T-2) elde edilmiştir.

Protokol 6 Aşaması: Dilsel Eşdeğerlik Çalışması (Pilot A)

Araştırmanın bu aşamasında dilsel eşdeğerlik çalışması yürütülmüştür. Bu süreç üç aşamada gerçekleştirilmiştir.

Analiz Faz 1

Ölçeğin Türkçe formunun dilsel eşdeğerliğini test etmek amacıyla maddelerin ordinal yapısına uygun olan Wilcoxon İşaretli Sıralar Testi kullanılmıştır. Çalışmada "Sıra Etkisi"ni (Order Effect) kontrol altına almak için Dengeleme Deseni (Counterbalanced Design) uygulanmıştır. Örneklem grubu bir devlet üniversitesi eğitim fakültesi, İngiliz Eğitimi Anabilim dalında öğrenim gören 3. sınıf

öğrencilerinden oluşmaktadır. Bu örneklemin tercih edilme sebebi, içsel konuşma gibi soyut, metaforik ve fenomenolojik derinliğe sahip ifadelerin (örn. Madde 12'deki "gerçek ses duyma" veya Madde 21'deki "monolog/konuşma yapma" ayrımı) diller arası anlamsal kaymalara uğramadan test edilebilmesi, ancak iki dilli (bilingual) yetkinliğe sahip bireylerden oluşan bir dengeleme deseni (counterbalanced design) örneklemiyle mümkün hale gelmesidir. Örneklem grubu seçkisiz olarak ikiye ayrılmıştır. Grup A (N=26) için "önce İngilizce sonra Türkçe", Grup B (N=24) için "önce Türkçe sonra İngilizce" uygulama sırası izlenmiştir. Analiz tablosunda (Tablo 1), hem birleştirilmiş genel grubun betimsel ve istatistiksel sonuçlarına yer verilmiş hem de sıra etkisinin izlenebilmesi amacıyla alt grupların (Grup A ve Grup B) anlamlılık düzeylerine ilişkin bilgiler sunulmuştur.

Analiz Faz 2

Ölçeğin İngilizce ve Türkçe formları arasındaki dilsel eşdeğerliği ve paralel form güvenirliğini test etmek amacıyla korelasyon analizleri yapılmıştır. Analiz sürecinde; formlardan elde edilen toplam puanların normal dağılım göstermesi ve sürekli veri niteliğinde olması nedeniyle parametrik testlerden Pearson Momentler Çarpımı Korelasyon Katsayısı kullanılmıştır. Bununla birlikte, ölçek maddelerinin Likert tipi (sıralı/ordinal) yapıda olması dikkate alınarak, madde bazlı tutarlılığı incelemek amacıyla parametrik olmayan Spearman Sıra Farkları Korelasyon Katsayısı hesaplanmıştır. Yapılan analizler sonucunda formlar arası anlamlılık düzeyi .01 olarak kabul edilmiştir.

Analiz Faz 3

Araştırmanın dilsel eşdeğerlik aşaması İngilizce ve Türkçe formlardan elde edilen puan ortalamaları arasında istatistiksel olarak anlamlı bir fark olup olmadığını belirlemeyi amaçlamaktadır. Bu nedenle fark testleri uygulanmıştır. Analiz öncesinde verilerin normal dağılım varsayımını sağlayıp sağlamadığı Shapiro-Wilk testi ile incelenmiştir. Analiz sonuçlarına göre hem İngilizce form toplam puanlarının (SW=.976, p=.393) hem de Türkçe form toplam puanlarının (SW=.983, p=.666) normallik varsayımını karşıladığı görülmüştür (p> .05). Verilerin normal dağılım göstermesi ve ölçümlerin sürekli değişken yapısında olması nedeniyle iki form arasındaki farkın testi için parametrik bir test olan Bağımlı Örneklem t-Testi (Paired Samples t-test) uygulanmıştır.

Protokol 7 Aşaması: Bilişsel Görüşme (Pilot B)

Bu aşamada ölçeğin Türkçe formu, hedef kitleyi temsil eden küçük bir gruba uygulanmıştır. Bu grup içerisinde 1 Türkçe Dil Uzmanı, lisans düzeyinde farklı sınıf düzeylerinde öğrenim gören 10 öğrenci vardır. Türkçe Dil uzmanı en az 20 yıllık mesleki deneyime sahiptir ve bir devlet üniversitesinde Türkçe Dil Bilgisi dersini yürütmektedir. Diğer katılımcılar ise eğitim fakültesinde çeşitli bölümlerde 2. ve 3. Sınıf düzeyinde öğrenim gören öğretmen adaylarından oluşmaktadır. Bu grubun tercih edilme gerekçesi, maddelerin hedef kitle olan öğretmen adayları tarafından nasıl anlamlandırıldığını doğrudan nitel dönütlerle test etmek ve Türkçe morfolojik/sentaktik yapıdaki olası pürüzleri ana saha öncesinde gidermektir. Ölçek formu katılımcılara okutturularak maddelerin anlaşılırlığı, okunabilirliği ve kültürel uygunluğu üzerine görüş belirtmeleri istenmiştir. Katılımcılardan edinilen dönütler ile nihai Türkçe formun anlaşılırlığı test edilmiştir.

Protokol 8 Aşaması: Ana Uygulama-Saha Çalışması

Bu aşamada nihai form, evreni temsil eden büyük örneklem grubuna uygulanmıştır. Veri toplama süreci 8-26 Aralık 2025 tarihleri arasında bir devlet üniversitesinin eğitim fakültesinin Japon Dili Eğitimi (n=49), İngiliz Dili Eğitimi (n=136), Türkçe Eğitimi (n=79), Okul Öncesi Eğitimi (n=32) ve Özel Eğitim (n=27) anabilim dallarının 1. 2. 3. ve 4. sınıflarında öğrenim gören 323 öğrenci ile gerçekleşmiştir. Öğretmen adaylarının tercih edilmesinin temel nedeni, içsel konuşmanın sınıf içi karar alma, öğretimsel yansıtma ve öz-düzenleme süreçlerinde etkin bir pedagojik mekanizma olmasıdır. Öğrenenler Psikometrik uyarlama çalışmalarında örneklem büyüklüğünün yeterliliği için literatürde sıklıkla madde-örneklem oranı (item-sample ratio) kullanılmakta olup madde sayısının 10 katı büyüklüğündeki bir örneklem kabul edilebilir bir standart olarak görülmektedir (Boateng vd., 201). Ayrıca Doğrulayıcı Faktör Analizi (DFA) için parametre başına 5 ila 10 kişi düşmesinin yeterli

olduğu bilinmektedir (Arafat, 2016; Baharum vd., 2023). Bu doğrultuda 26 maddelik ölçme aracı için ulaşılan 323 kişilik örneklem (yaklaşık 12.4 oranı), faktör analizi varsayımlarını karşılamak için fazlasıyla yeterli kabul edilmiştir. Uygulama aşamasında katılımcılara araştırmanın amacına yönelik bilgiler verilmiş, bilimsel etik kuralları gereğince kendilerinden edinilecek bilgilerin hiçbir şekilde üçüncü kişiler ile paylaşılmayacağı konusunda bilgilendirme yapılmıştır. Araştırmaya katılım için sözlü onamları alındıktan sonra uygulama gerçekleştirilmiştir. Öğretmen adaylarından edinilen bilgiler analiz edilmek üzere araştırmacılar tarafından IBM SPSS-21 programına işlenmiştir. Veri setinin analizlere hazırlanması sürecinde, ölçeğin orijinal yapısında olumsuz/zıt ifade içeren maddeler (Madde 7, Madde 15 ve Madde 21) ters puanlanarak (1=7, 2=6, 3=5, 4=4, 5=3, 6=2, 7=1) veri setine işlenmiş ve Doğrulayıcı Faktör Analizi bu ters kodlanmış veriler üzerinden gerçekleştirilmiştir.

Protokol 9 Aşaması: Psikometrik Analizler

Oluşturulmuş data setindeki veriler üzerinde yapılan incelemeler üzerinden ölçeğin yapı geçerliliği ve güvenirlik testleri yapılmıştır. Yapı geçerliliği için Doğrulayıcı (DFA) Faktör Analizi gerçekleştirilmiştir. DFA tümdengelimsel bir yaklaşımla uygulanan, değişken kümeleri arasındaki ilişkiler konusunda hipotezleri test etmek için kullanılan istatistiksel bir stratejidir (Hoyle, 2000). Bu stratejide ölçme aracına ilişkin belirlenmiş alt boyutların hedeflenen örneklemin görüşleriyle ne kadar tutarlı olduğunu değerlendirme hususunda önemli, işlevsel bir tekniktir. Güvenirlik konusunda ölçme aracını meydana getiren maddelerin birbirleriyle ne derece uyumlu olduğunu incelemek adına Cronbach Alpha katsayıları ve madde-toplam korelasyonları analiz edilmesi önemlidir (Pallant,2020).

Bulgular

Bu çalışmada İçsel konuşma ölçeğinin Türkçeye uyarlama çalışması amaçlanmaktadır. Bu amaca bağlı olarak Protokol 6 ve Protokol 8 aşamasında toplanan veriler üzerinde yapılan analizler incelenecektir. Bu bağlamda başlangıç olarak İçsel Konuşma ölçme aracının Türkçe formuna ait Sıra etkisi kontrol edilmiştir. Yapılan analiz sonucunda elde edilen bulgular Tablo 1'deki gibidir:

Tablo 1
İngilizce ve Türkçe Form Maddelerinin Karşılaştırmalı Analizi ve Sıra Etkisi Kontrolü

Madde	İngilizce (X^{-}±Ss)	Türkçe (X^{-}±Ss)	Z (Genel)	*p* (Genel)	*p* (Grup A) (İng → Tr)	*p* (Grup B) (Tr → İng)
1	3.66 ± 1.53	3.90 ± 1.46	-1.198	.231	.376	.429
2	4.92 ± 1.37	4.88 ± 1.55	-0.608	.543	.020*	.032*
3	3.00 ± 1.75	2.86 ± 1.60	-0.964	.335	.431	.518
4	3.02 ± 1.79	2.74 ± 1.66	-0.927	.354	.720	.343
5	2.66 ± 1.57	2.58 ± 1.54	-0.390	.697	.474	.776
6	4.54 ± 1.46	4.50 ± 1.49	-0.182	.855	.557	.540
7	4.80 ± 1.31	4.84 ± 1.31	-0.091	.927	.145	.080
8	3.92 ± 1.52	3.88 ± 1.38	-0.424	.672	.961	.470
9	5.32 ± 1.46	5.32 ± 1.43	-0.104	.917	.592	.808
10	5.28 ± 1.33	5.22 ± 1.39	-0.551	.582	1.00	.405
11	5.26 ± 1.50	5.22 ± 1.61	-0.414	.679	.785	.271
12	2.76 ± 1.68	3.20 ± 1.95	-2.094	.036*	.021*	.803
13	4.90 ± 1.69	4.82 ± 1.66	-0.243	.808	.091	.150
14	3.46 ± 1.68	3.46 ± 1.62	-0.194	.846	.205	.085
15	4.42 ± 1.79	4.30 ± 1.67	-0.617	.537	.902	.356
16	3.74 ± 1.47	3.54 ± 1.39	-0.850	.395	.824	.164
17	4.68 ± 1.61	4.98 ± 1.32	-1.412	.158	.887	.012*
18	5.24 ± 1.61	5.22 ± 1.61	-0.163	.870	1.00	.763
19	4.56 ± 1.68	4.50 ± 1.61	-0.512	.609	.290	.564
20	5.08 ± 1.61	4.82 ± 1.52	-1.408	.159	.767	.048*
21	4.42 ± 1.49	4.04 ± 1.40	-2.147	.032*	.182	.075

22	4.40 ± 1.59	4.64 ± 1.61	-1.272	.203	.172	.618
23	3.96 ± 1.84	4.24 ± 1.78	-1.613	.107	.202	.353
24	4.00 ± 1.78	4.00 ± 1.76	-0.234	.815	.810	.454
25	3.78 ± 1.66	4.18 ± 1.72	-2.651	.008**	.014*	.148
26	3.88 ± 1.61	4.00 ± 1.86	-0.774	.439	.200	.637

Not: N=50 (Grup A n=26, Grup B n=24). Wilcoxon İşaretli Sıralar Testi kullanılmıştır. * p < .05, ** p < .01.

Tablo 1'deki bulgular genel olarak değerlendirildiğinde ölçekte yer alan 26 maddenin 23'ünde genel grup bazında İngilizce ve Türkçe formlar arasında istatistiksel olarak anlamlı bir fark bulunmadığını göstermektedir (p> .05). Bu sonuç ölçek maddelerinin %88.5 oranında diller arası eşdeğerliğe sahip olduğunu göstermektedir.

Araştırmada kullanılan dengeleme deseni dilsel eşdeğerlik çalışmalarında karşılaşılan "sıra etkisini" başarıyla kontrol altına almıştır. Özellikle 2. madde (*Zihnimdeki şeyler hakkında kendi kendime konuşurken...*) ve 17. madde (*İçimden konuşarak kendime bir şeyler yapmamamı söylerim...*) incelendiğinde; tekil gruplarda (Grup A ve Grup B) ortaya çıkan zıt yönlü anlamlı farklılıkların, birleştirilmiş analizde (p > .05) birbirlerini etkisiz hale getirdiği tespit edilmiştir. Bu durum, söz konusu maddelerdeki varyansın çeviri hatasından değil, uygulanan dillerin sırasından kaynaklandığını kanıtlamaktadır. Ayrıca istatistiksel analizler sonucunda anlamlı farklılığın devam ettiği 12, 21 ve 25. maddeler ölçek formları üzerinden dilbilimsel ve yapısal olarak irdelenmiştir:

Madde 12 (p=.036, TR> İNG)

İngilizce formda yer alan *"I hear other people's actual voices"* ifadesi, işitsel bir halüsinasyon vurgusu taşıyan "gerçek" kelimesiyle oldukça keskin bir sınır çizmektedir. Ayrıca Türkçe formdaki "başkalarının... gerçek seslerini duyarım" ifadesi, Türk kültüründe "zihinde canlandırma" (vivid imagery) bağlamında daha metaforik ve kabul edilebilir bir ifade olarak algılanmış olabilir. Türkçe formdaki ortalamanın (X_{Tr}}=3.20) İngilizce formdan (X_{Eng}=2.76) yüksek çıkması, katılımcıların Türkçe maddede "ses duyma" eylemini patolojik bir durumdan ziyade "zihinsel canlandırma" olarak yorumlamaya daha meyilli olduklarını düşündürmektedir.

Madde 21 (p=.032, İNG> TR)

İngilizce formda *"like a speech or a monologue" şeklinde* daha nötr ve genel bir ifade kullanılmıştır. Ayrıca Türkçe çeviride "speech" kelimesinin karşılığı olarak *"tek taraflı bir konuşma ya da monolog"* ifadesi tercih edilmiştir. "Tek taraflı" vurgusu, maddeye İngilizce orijinalinde olmayan bir "iletişimsizlik" veya "kısıtlılık" anlamı yüklemiş olabilir. Bu nüansın katılımcıların Türkçe maddeye katılım düzeyini (X_{Tr}=4.04), İngilizce forma (X_{Eng}=4.42) kıyasla düşürdüğü değerlendirilmiştir.

Madde 25 (p=.008, TR> İNG)

İngilizce formda *"When angry"* şeklinde edilgen/kişileştirilmemiş bir yapı kullanılmıştır. Türkçe formda ise *"Öfkelendiğimde"* şeklinde birinci tekil şahıs ekiyle doğrudan kişiselleştirilmiş bir yapı kullanılmıştır. Türkçe formda öznenin (ben) eylemi sahiplenmesinin anlamsal olarak daha güçlü olması ve *"sakinleşmeme yardımcı olabilir"* ifadesindeki yeterlilik vurgusu, katılımcıların bu maddeyi İngilizce forma (X_{Eng}=3.78) göre Türkçe formda (X_{Tr}=4.18) daha yüksek oranda onaylamalarını sağlamıştır.

Madde 12 ve 25'teki istatistiksel farklılaşmanın, kavramsal bir uyumsuzluktan ziyade dillerin yapısal ve kültürel algı farklılıklarından kaynaklandığı tespit edilmiştir. Ölçeklerin kültürlerarası uyarlamalarında bu tür istatistiksel farklılaşmalar bir çeviri hatasından ziyade, maddelerin farklı sosyokültürel bağlamlarda farklı anlamlar uyandırmasının doğal bir sonucu olan ''madde yanlılığı'' olarak açıklanmaktadır (Byrne, 2016). Örgütsel ve davranışsal bilimler literatüründe kültürel uyum için madde ifadelerinde yapılan bu tür değişikliklerin okuyucuya şeffaf bir şekilde raporlanması ve ardından DFA ile yapı geçerliliğinin kanıtlanması en iyi uygulama standardı olarak önerilmektedir (Heggestad vd., 2019). Nihai formda anlam bütünlüğünü korumak adına gerekli düzeltmeler yapılarak

eşdeğerlik sağlanmıştır. Araştırmada İngilizce ve Türkçe formların toplam puanları arasındaki ilişki de belirlemiştir. Bu nedenle yapılan Pearson korelasyon'a ait analizi sonuçları Tablo 2'de sunulmuştur.

Tablo 2
İngilizce ve Türkçe Form Toplam Puanlarına İlişkin Betimsel İstatistikler ve Pearson Korelasyon Sonuçları

Formlar	N	Ortalama ($\bar{X}$)	Ss	r	p
İngilizce Form	50	4.22	.74	.935**	.000
Türkçe Form	50	4.23	.74		

**Not:* **p < .01

Tablo 2 incelendiğinde, katılımcıların İngilizce formdan aldıkları toplam puan ortalaması (X=4.22) ile Türkçe formdan aldıkları puan ortalamasının (X=4.23) birbirine oldukça yakın olduğu görülmektedir. İki form arasındaki ilişki incelendiğinde pozitif yönde çok yüksek düzeyde ve istatistiksel olarak anlamlı bir korelasyon saptanmıştır (r = .935, p < .001) (Can, 2016; Pallant, 2020). Bu bulgu ölçeğin orijinal ve uyarlanan formlarının %87 oranında (r^2=.87) ortak varyansı paylaştığını ve eşdeğer ölçümler yaptığını göstermektedir.

Ölçeğin dilsel eşdeğerliğini madde bazında detaylandırmak amacıyla Spearman korelasyon analizi yapılmıştır. Analiz sonuçları ise Tablo 3'te verilmiştir.

Tablo 3
İngilizce ve Türkçe Form Maddeleri Arasındaki Spearman Sıra Farkları Korelasyon Katsayıları

Madde	Rs	Madde	rs
M1	.660**	M14	.680**
M2	.652**	M15	.763**
M3	.829**	M16	.567**
M4	.616**	M17	.513**
M5	.667**	M18	.841**
M6	.753**	M19	.867**
M7	.563**	M20	.645**
M8	.466**	M21	.653**
M9	.672**	M22	.599**
M10	.822**	M23	.757**
M11	.790**	M24	.732**
M12	.733**	M25	.791**
M13	.619**	M26	.770**

**Not:* N=50. ** p < .01 (Çift yönlü).

Tablo 3'teki bulgulara göre ölçekte yer alan 26 maddenin tamamında İngilizce ve Türkçe formlar arasında pozitif yönde, istatistiksel olarak anlamlı (p< .001) ilişkiler tespit edilmiştir. Ayrıca madde bazlı korelasyon katsayılarının .466 (Madde 8) ile .867 (Madde 19) arasında değiştiği görülmektedir. Maddelerin büyük çoğunluğunun .60 ve üzerinde yüksek korelasyon değerlerine sahip olması çevirinin sadece toplam puan bazında değil, madde içeriği ve anlamı bazında da tutarlı olduğunu kanıtlamaktadır.

Ölçeğin İngilizce ve Türkçe formlarından elde edilen puan ortalamalarının eşdeğerliğini test etmek amacıyla Bağımlı Örneklem t-Testi analizi yapılmıştır. Analiz sonuçları Tablo 4'te sunulmuştur.

Tablo 4
İngilizce ve Türkçe Form Toplam Puanlarının Karşılaştırılmasına İlişkin t-Testi Sonuçları

Ölçüm	N	Ortalama (X^-)	Ss	Sd	t	p
İngilizce Form	50	4.22	.74	49	-.225	.823
Türkçe Form	50	4.23	.74			

Not: p > .05.

Tablo 4 incelendiğinde katılımcıların İngilizce formdan aldıkları puan ortalaması (X=4.22) ile Türkçe formdan aldıkları puan ortalaması (X=4.23) arasında istatistiksel olarak anlamlı bir fark bulunmadığı görülmektedir (t(49) = -.225, p = .823). Puan ortalamaları arasındaki farkın ihmal edilebilir düzeyde (yaklaşık .01 puan) olması ve p değerinin anlamlılık sınırının (.05) oldukça üzerinde çıkmıştır. Bu durum Türkçe formun, orijinal İngilizce form ile aynı puanlama yapısına ve ölçüm gücüne sahip olduğunu kanıtlamaktadır. Daha önce yapılan korelasyon analizinde saptanan yüksek ilişki (r=.935) ile birlikte değerlendirildiğinde; uyarlanan Türkçe form puanları ile orijinal form puanları arasında oldukça güçlü ve pozitif yönde bir ilişki olduğu; ayrıca iki formun toplam puan ortalamaları arasında anlamlı bir farklılık bulunmadığı sonucuna ulaşılmıştır. Bu bulgular, Türkçe formun orijinal form ile yüksek düzeyde yapısal ve dilsel bir uyum sergilediğine işaret etmektedir.

Yapılan analizler sonucunda madde bazında (Wilcoxon) tespit edilen 12, 21 ve 25. maddelerdeki istatistiksel farklılaşmalar tespit edilmiştir. Buna rağmen ölçeğin toplam puanları üzerinden yapılan analizlerde (Pearson korelasyonu ve t-testi) formlar arasında yüksek düzeyde uyum ve farksızlık tespit edilmiştir. Bu durum, söz konusu üç maddedeki ayrışmanın ölçeğin genel ölçüm gücünü ve yapısal bütünlüğünü bozacak düzeyde olmadığını göstermektedir. Yani ölçek mevcut çevirisiyle dahi dilsel eşdeğerliği sağlamaktadır. Ancak, ölçeğin psikometrik kalitesini en üst düzeye çıkarmak ve kültürel/kavramsal nüansları tam olarak karşılamak amacıyla, ilgili maddelerde belirtilen nitel düzeltmeler yapılarak Nihai Türkçe Form oluşturulmuştur. Bu sayede, istatistiksel olarak 'ihmal edilebilir' düzeyde olan hatalar da kavramsal olarak giderilmiştir.

Yapı Geçerliliği: Doğrulayıcı Faktör Analizi (DFA)

İçsel Konuşma Ölçeği Türkçe Formu'nun yapı geçerliliğini test etmek amacıyla, orijinal ölçekte tanımlanan beş faktörlü yapı (Diyalojik, Değerlendirici/Eleştirel, Diğer Kişiler, Sıkıştırılmış ve Pozitif/Düzenleyici) üzerinden DFA gerçekleştirilmiştir. Analiz sürecinde Maksimum Olabilirlik (Maximum Likelihood) tahmin yöntemi kullanılmıştır.

Modelin verilerle uyumunu belirlemek amacıyla χ^2 / df (Ki-kare/Serbestlik Derecesi), RMSEA (Yaklaşık Hataların Ortalama Karekökü), CFI (Karşılaştırmalı Uyum İndeksi), TLI (Tucker-Lewis İndeksi), GFI (İyilik Uyum İndeksi) ve SRMR değerleri incelenmiştir. İlk analizler sonucunda, ölçeğin orijinal formunda ters kodlanan 21. maddenin ("Sözcüklerle düşünme tarzım, karşılıklı bir sohbetten çok bir konuşma yapma ya da monolog şeklindedir") standartlaştırılmış faktör yükünün zayıf ve istatistiksel olarak anlamsız olduğu tespit edilmiştir (p > .05). Bu durumun, maddenin taşıdığı anlamın Türk örnekleminde orijinalindeki gibi diyalojik bir yapıyı yansıtmamasından kaynaklandığı değerlendirilmiştir. Psikometrik sağlamlığı korumak amacıyla 21. madde modelden çıkarılmış ve yapı geçerliliği kalan 25 madde üzerinden yeniden sınanmıştır. Elde edilen ilk 25 maddelik modelin (İlk Model) uyum indekslerinin kabul edilebilir sınırlarda olduğu görülmüştür. Ancak model uyumunu daha da iyileştirmek amacıyla Modifikasyon İndeksleri (MI) incelenmiş ve aynı teorik yapı (Değerlendirici/Eleştirel) altında yer alan, kavramsal olarak birbirine çok yakın ifadeler barındıran e11 ve e12 hata terimleri arasında tek bir kovaryans tanımlanmıştır. 21. maddenin çıkarılmasıyla oluşturulan ilk modele (25 Madde) ve modifikasyon işlemi sonrasındaki modele ait uyum indeksleri Tablo 5'te ayrı ayrı sunulmuştur.

Tablo 5
İçsel Konuşma Ölçeği Türkçe Formu DFA Uyum İndeksi Sonuçları ve Referans Değerler

Uyum İndeksi	İlk Model Değeri (25 Madde)	Modifiye Model Değeri	Kabul Edilebilir / İyi Uyum Eşiği	Atıf (Referans)
χ^2 (CMIN)	882.708	757.579	-	
df	265	264	-	
χ^2 / df	3.331	2.870	≤ 2 (İyi); ≤3 (Kabul edilebilir)	Schermelleh-Engel vd. (2003)
RMSEA	.085	.076	≤.05 (İyi); ≤.08 (Kabul edilebilir)	Schermelleh-Engel vd. (2003)
CFI	.800	.840	≥ .95 (İyi)	Hu ve Bentler (1999)
GFI	.818	.840	≥ .95 (İyi); ≥ .90 (Kabul edilebilir)	Schermelleh-Engel vd. (2003)
TLI	.773	.818	≥ .95 (İyi)	Hu ve Bentler (1999)
SRMR	.097	.094	≤ .08 (Kabul edilebilir/İyi)	Hu ve Bentler (1999)

Tablo 5 incelendiğinde, 21. maddenin çıkarılması ve gerçekleştirilen modifikasyon işlemi sonrasında χ^2 / df oranının 2.87'ye düşerek 3'ün altında kaldığı ve modelin kabul edilebilir bir uyuma sahip olduğu tespit edilmiştir (Schermelleh-Engel vd., 2003). RMSEA değeri (.076) kabul edilebilir uyum sınırları içerisindedir. Bununla birlikte; CFI (.840), TLI (.818) ve hesaplanan SRMR (.094) değerleri, Hu ve Bentler (1999) tarafından önerilen katı standartların (CFI ve TLI için ≥ .95, SRMR için ≤ .08) bir miktar dışında kalmıştır. Ancak, ölçeğin çok boyutlu karmaşık yapısı, sıralı kategorik veri (Likert) kullanımının uyum indeksleri üzerindeki doğal baskılayıcı etkisi ve uyarlama çalışmalarındaki kültürel farklılıklar göz önüne alındığında modelin genel olarak doğrulanabilir bir yapı sergilediği söylenebilir. Ayrıca Hoelter indeksi, örneklem büyüklüğünün analiz için yeterli olduğunu desteklemektedir. Analiz sonucunda kalan 25 maddeye ait standartlaştırılmış faktör yükleri (λ) ve açıklanan varyans (R^2) değerleri Tablo 6'da sunulmuştur.

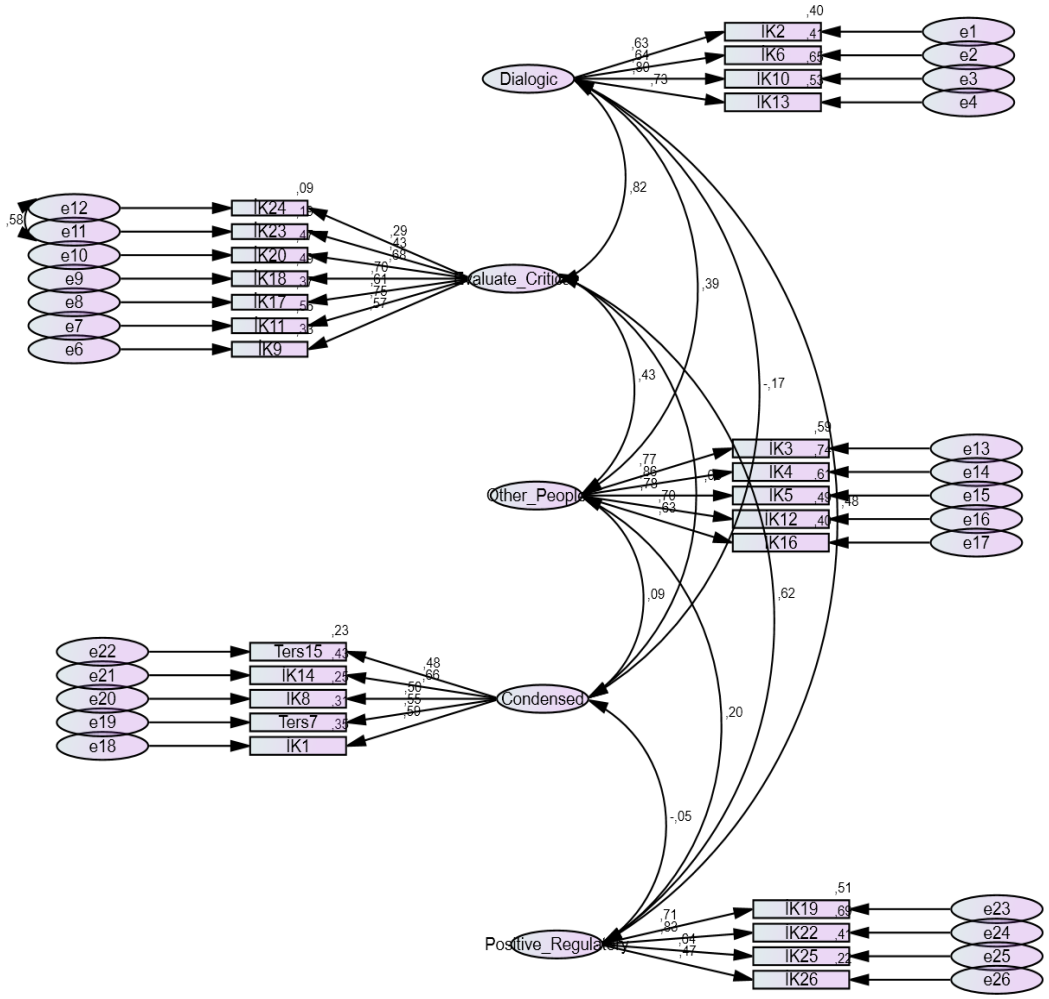
Tablo 6
Ölçek Maddelerine İlişkin Standartlaştırılmış Faktör Yükleri ve R^2 *Değerleri*

Faktör / Madde	Faktör Yükü (λ)	R2	p
Diyalojik			
İK2	0.631	0.398	*
İK6	0.637	0.405	*
İK10	0.805	0.648	*
İK13	0.729	0.532	*
Değerlendirici / Eleştirel			
İK9	0.573	0.328	*
İK11	0.749	0.561	*
İK17	0.608	0.369	*
İK18	0.699	0.489	*
İK20	0.683	0.467	*
İK23	0.433	0.188	*
İK24	0.293	0.086	*
Diğer Kişiler			
İK3	0,769	0,592	*
İK4	0,859	0,738	*
İK5	0,779	0,607	*
İK12	0,700	0,490	*
İK16	0,635	0,404	*
Sıkıştırılmış			

İK1	0.594	0.353	*
İKT7	0.552	0.305	*
İK8	0.497	0.247	*
İK14	0.657	0.431	*
İKT15	0.482	0.233	*
Pozitif / Düzenleyici			
İK19	0,713	0,508	*
İK22	0,829	0,687	*
İK25	0,641	0,411	*
İK26	0,471	0,222	*

** p < 0.001* İKT7 ve İKT15 maddeleri olumsuz/zıt ifade taşıyan maddeler olup analiz öncesinde ters puanlanmıştır. Tüm parametre tahminleri istatistiksel olarak anlamlıdır.

Tablo 6 incelendiğinde, 21. maddenin çıkarılmasıyla birlikte kalan 25 maddenin standartlaştırılmış faktör yüklerinin genel olarak 0.29 ile 0.86 arasında dağıldığı ve tüm parametre tahminlerinin istatistiksel olarak anlamlı ($p < .001$) olduğu görülmektedir. İK24 ($\lambda = 0.29$) maddesinin faktör yükü literatürde genellikle kabul gören .30 veya .40 eşik değerlerinin sınırında kalmasına rağmen, madde analiz dışında bırakılmamış ve modelde tutulmuştur. Bu durum, model kurulurken yalnızca istatistiksel faktör yüklerine bakılarak mekanik bir eleme yapılmadığını, ölçme aracının teorik bütünlüğünün ve olgusal çerçevesinin korunmasının esas alındığını göstermektedir (Kline, 2015). Faktör yükü kesme değerlerinin evrensel bir mutlaklığının olmadığı ve araştırmanın amacına bağlı olarak içerik geçerliliğini korumak adına esnetilebileceği bilinmektedir (Baharum vd., 2023). Dahası, 0.40 ile 0.60 arasındaki faktör yükleri geniş kapsamlı (broad) kuramsal yapılar için son derece normal kabul edilirken (Steenkamp ve Maydeu-Olivares, 2023), Maksimum Olabilirlik (ML) tahmin edicisinin Likert tipi gibi sıralı-kategorik veriler üzerinde kullanıldığında faktör yüklerini sistematik olarak olduğundan daha düşük hesaplama eğilimine sahip olduğu güncel çalışmalarla da kanıtlanmıştır (Brauer, Ranger, ve Ziegler, 2023). Elde edilen bulgular doğrultusunda, modifikasyon sonrası elde edilen nihai DFA path diyagramı Şekil 1'de verilmiştir.



Şekil 1. İçsel Konuşma Ölçeği Türkçe Formu DFA Path Diyagramı

Şekil 1'deki DFA sonuçlarına bağlı standartlaştırılmış katsayılar incelendiğinde, gizil değişkenlerin gözlenen maddeleri açıklama oranlarının, sınırda kalan İK24 maddesi de dâhil olmak üzere, literatürde öngörülen sınırlar dâhilinde kümelendiği görülmektedir (Alderson-Day vd., 2018). Alt boyutların kendi aralarındaki korelasyon katsayılarının yönü ve şiddeti, ölçeğin teorik altyapısıyla uyumludur. Ayrıca modelin uyum indekslerini iyileştirmek adına Modifikasyon İndeksleri (MI) temel alınarak, semantik açıdan birbirine yakın olan e11 ve e12 hata terimleri arasında kurulan çift yönlü kovaryansın diyagram üzerinde yapılandırılması, model uyumunda anlamlı bir iyileşme sağlamıştır.

Ölçeğin Güvenirliğine İlişkin Bulgular

İçsel Konuşma Ölçeği Türkçe Formu'nun genelinin ve alt boyutlarının iç tutarlılığını belirlemek amacıyla Cronbach Alpha ve McDonald's Omega güvenirlik katsayıları hesaplanmıştır. Yapılan analiz sonucunda elde edilen değerler Tablo 7'de sunulmuştur.

Tablo 7
İçsel Konuşma Ölçeği Türkçe Formu Güvenirlik Analizi Sonuçları

Faktör	Madde Sayısı	Cronbach Alpha (α)	McDonald's Omega (ω)
Diyalojik (Dialogic)	4	.789	.793
Değerlendirici / Eleştirel (Evaluate/Critical)	7	.787	.791
Diğer Kişiler (Other People)	5	.863	.865
Sıkıştırılmış (Condensed)	5	.691	.695
Pozitif / Düzenleyici (Positive/Regulatory)	4	.753	.757
Ölçek Toplamı	25	.836	.839

Tablo 7 incelendiğinde, 25 maddeden oluşan ölçeğin bütününe ilişkin genel iç tutarlılık katsayıları Cronbach Alfa (α) = .836 ve McDonald's Omega (ω) = .839 olarak bulunmuştur. Psikometrik ölçümlerde iç tutarlılık katsayılarının .70 ile 1.00 arasında olması, ölçeğin yüksek düzeyde güvenilir olduğunu kanıtlamaktadır.

Alt faktörler bazında değerlendirildiğinde; Diğer Kişiler (α = .863, ω = .865), Diyalojik (α = .789, ω = .793), Değerlendirici/Eleştirel (α = .787, ω = .791) ve Pozitif/Düzenleyici (α = .753, ω = .757) faktörlerinin literatürde genel kabul gören .70 sınırının oldukça üzerinde ve tatminkâr bir güvenilirliğe sahip olduğu saptanmıştır. Sıkıştırılmış (Condensed) faktörünün güvenirlik katsayısı (α = .691, ω = .695) geleneksel .70 sınırının hemen kıyısında görünse de, sosyal bilimler literatüründe az maddeli (4-5 madde) alt faktörlerin daha düşük iç tutarlılık katsayıları üretme eğiliminde olduğu bilinmekte (Heggestad vd., 2019) ve kültürlerarası uyarlama çalışmalarında .60 ile .70 arasındaki değerler sınırda ancak "kabul edilebilir" iç tutarlılık düzeyleri olarak değerlendirilmektedir (Ursachi vd., 2015). Nitekim McDonald's Omega katsayısının geleneksel alfa katsayısına kıyasla faktör yüklerindeki farklılıkları daha iyi tolere ederek daha isabetli bir güvenirlik tahmini sunduğu dikkate alındığında (ω = .695), söz konusu alt boyutun yapısal tutarlılığının yeterli düzeyde olduğu doğrulanmıştır.

Ölçek maddelerinin ait oldukları alt faktörler içerisindeki ayırt ediciliklerini belirlemek amacıyla incelenen alt boyutlara ait düzeltilmiş madde-toplam korelasyonu (corrected item-total correlation) katsayılarının .388 ile .770 aralığında değiştiği tespit edilmiştir. Literatürde madde-toplam korelasyonunun .30 ve üzerinde olması, maddelerin ölçülmek istenen özelliği ayırt etmede yeterli güce sahip olduğunu göstermektedir (Pallant, 2020). Elde edilen bu sonuçlar, yapı geçerliliği aşamasında 21. maddenin modelden çıkarılmasıyla ulaşılan 25 maddelik nihai formun psikometrik açıdan güçlü olduğunu, ölçeğin hem genelinin hem de alt faktörlerinin hedef örneklem üzerinde oldukça tutarlı ve kararlı ölçümler yapabildiğini göstermektedir.

Tartışma ve Sonuç

Bu çalışmada, öğrenenlerin bilişsel ve öz-düzenleme süreçlerini yönetmede kritik bir role sahip olan içsel konuşma mekanizmasını ölçmek amacıyla geliştirilen ölçeğin Türkçe formunun geçerlilik ve güvenirlik kanıtları ortaya konmuştur. Yapılan DFA sonuçları, ölçeğin orijinal formunda yer alan beş faktörlü yapısının Türk kültürü ve öğretmen adayları örnekleminde genel olarak doğrulandığını göstermektedir.

Modelin uyum iyiliği indeksleri bütüncül olarak incelendiğinde, yapısal uyumsuzluk gösteren 21. maddenin çıkarılmasıyla ulaşılan 25 maddelik nihai yapıda χ^2 /df (2.87) oranının 3'ün altında, RMSEA (.076) değerinin ise önerilen sınırların altında olduğu saptanmıştır. Bu durum Schermelleh-Engel vd. (2003) açısından modelin kabul edilebilir bir genel uyum sergilediğini göstermektedir. Bununla birlikte, CFI (.840) ve TLI (.818) gibi karşılaştırmalı uyum indeksleri, Hu ve Bentler (1999) tarafından sıklıkla vurgulanan geleneksel .90 veya .95 katı eşik değerlerinin altında kalmıştır. Alan yazında ise benzer içsel konuşma ölçeklerinin geliştirme ve uyarlama çalışmalarında (Alderson-Day vd., 2018; Brinthaupt vd.,2009; Racy vd.,2020; Zourbanos vd.,2009) psikometrik sınır değerlerin ve faktör yapılarının zaman zaman örneklem özelliklerine göre hassasiyet gösterdiği bildirilmektedir. Racy, Morin ve Hagerty (2024) ile Uttl, Morin ve Hamper (2011) tarafından gerçekleştirilen eş zamanlı

geçerlilik çalışmalarında da vurgulandığı üzere, içsel konuşma gibi bireyin tamamen öznel ve örtük fenomolojik deneyimlerine dayanan yapıların öz-bildirim ölçekleriyle ölçülmesi, değişken bazlı varyans dağılımlarını doğrudan etkileyebilmektedir.

Söz konusu karşılaştırmalı uyum indekslerinin (CFI, TLI) beklenen seviyenin altında kalması metodolojik, dilsel ve örnekleme bağlı gerekçelerle açıklanabilir. Metodolojik açıdan Kline (2015) özellikle çok maddeli, gizil değişkenler arası korelasyonların yüksek olduğunu ve çok boyutlu karmaşık modellerde simülasyon çalışmalarına dayanan katı ve mutlak eşik değerlerin (Hu ve Bentler, 1999) karşılanmasının oldukça zor olabileceğini belirtmektedir. Bu tür durumlarda modelin kuramsal ve teorik anlamlılığının ön planda tutulması gerekmektedir. Benzer şekilde nicel veri analizi literatüründe, örneklem grubunun homojen yapısının veya verilerin çoklu normallik eğilimlerinin uyum indekslerini baskılayabileceği bilinmektedir (Can, 2016; Pallant, 2020). Buna ek olarak, uyum indekslerinin hesaplanmasında kullanılan tahmin edicinin veri tipiyle etkileşimi de önemli bir faktördür. Likert tipi sıralı (ordinal) ve kesikli (discrete) verilerin analizinde standart Maksimum Olabilirlik (ML) yönteminin kullanılması, sürekli veri varsayımını tam karşılamadığı için CFI ve TLI gibi indeksleri matematiksel olarak doğal bir şekilde baskılamaktadır (Brauer vd., 2023). Tüm bunlara rağmen, χ^2/df oranının kabul edilebilir sınırlar içinde kalması (Baharum vd., 2023; Schermelleh-Engel vd., 2003) ve RMSEA değerinin 0.08 sınırının altında kalarak evrendeki makul yaklaşım hatalarını (reasonable errors of approximation) temsil etmesi (Gjersing vd., 2010), modelin genel geçerliliğini sağlamada yeterli bir kanıt olarak kabul edilmektedir.

Dilsel ve kültürel bağlamda ise, Türkçe'nin sondan eklemeli morfolojik yapısı ve ifadelerin kültürel şemalardaki algılanma biçimleri, maddelerin faktör yükleri üzerinde belirleyici olmuştur. Özellikle Madde 12, 21 ve 25'te gözlenen durum, dil uyarlama çalışmalarında sıklıkla karşılaşılan semantik ve kültürel nüansların somut bir göstergesidir (Latinjak vd.,2016; Zervas vd.,2007). Nitekim 21. maddedeki bu kültürel/semantik farklılaşma, maddenin yapı geçerliliği aşamasında (DFA) yetersiz faktör yükü vererek modelden düşmesiyle de kendini kanıtlamıştır. Fossa (2022) tarafından belirtildiği üzere, içsel konuşma kültür ve eğitim süreçleriyle iç içe geçmiş, dilin içselleştirilmesiyle şekillenen bir olgudur. Dolayısıyla orijinal ölçek maddelerinde yer alan ifadelerin Türk kültüründeki ve dilindeki karşılıkları, öğretmen adaylarının zihinsel süreçlerinde batı dillerindeki yapılardan daha farklı bir varyans dağılımı üretmiştir.

Çalışmanın güvenirlik bulguları değerlendirildiğinde, 25 maddelik Türkçe formun genel iç tutarlılık katsayısının yüksek düzeyde çıkması (α = .836, ω = .839), ölçme aracının maddelerinin kendi içinde tutarlı olduğunu ve homojen bir yapı sergilediğini doğrulamaktadır. Faktör bazındaki güvenirlik katsayıları incelendiğinde Diğer Kişiler (α = .863, ω = .865), Diyalojik (α = .789, ω = .793), Değerlendirici/Eleştirel (α = .787, ω = .791) ve Pozitif/Düzenleyici (α = .753, ω = .757) alt boyutlarına ilişkin iç tutarlılıkların literatürde genel kabul gören 0.70 sınırının oldukça üzerinde ve tatminkâr düzeyde olduğu görülmüştür. Yalnızca Sıkıştırılmış (α = .691, ω = .695) faktörünün güvenirlik katsayısının .70 sınırının hemen kıyısında kaldığı saptanmıştır. Literatürde Cronbach Alfa iç tutarlılık katsayısı için genellikle 0.70 değeri anlamlı kabul edilse de, kültürlerarası uyarlama çalışmalarında .60 ile .69 arasındaki değerler sınırda ancak kabul edilebilir düzeyler olarak değerlendirilebilmektedir (Ursachi vd., 2015). Ayrıca Heggestad vd. (2019) ölçeklerde madde sayısının azaltılmasının (ilgili faktör 5 maddeden oluşmaktadır) güvenirlik katsayılarını doğal olarak düşüreceğini, bu nedenle kısa ölçeklerin iç tutarlılıklarının salt istatistiksel olarak değil, teorik kapsam geçerliliği (content validity) bağlamında değerlendirilmesi gerektiğini vurgulamaktadır. Bu durum aynı zamanda kültürlerarası aktarım sürecinde varyans dağılımlarının farklılaşmasıyla da (Gjersing vd., 2010) açıklanabilmektedir. Elde edilen bu güçlü güvenirlik kanıtları ve faktör yapılarının istikrarı, Türkçe formun pedagojik araştırmalarda kullanılabilecek geçerli ve güvenilir bir ölçme aracı olduğunu teyit etmektedir.

Sonuç olarak, elde edilen bazı karşılaştırmalı uyum indekslerindeki düşüklüklere rağmen kuramsal bütünlüğün korunması, parametre tahminlerinin tamamının anlamlı çıkması ve maddelerin eğitimsel geçerliliği göz önüne alınarak ölçeğin Türkçe formunun mevcut yapısıyla kullanılabilir olduğu tespit edilmiştir. Gelecek çalışmalarda ölçeğin psikometrik özelliklerinin daha geniş, heterojen ve farklı eğitim kademelerindeki örneklem gruplarında (motor beceri öğrenimi bağlamında veya farklı sınıf düzeylerindeki öğrencilerde) yeniden test edilmesi ve test-tekrar test gibi zamana karşı tutarlılık

kanıtlarının incelenmesi önerilmektedir. Ayrıca bu çalışmada verilerin normallik varsayımını karşılaması nedeniyle geleneksel ve yaygın kabul gören ML (Maksimum Olabilirlik) tahmin yöntemi kullanılmış olsa da; gelecekteki uyarlama çalışmalarında sıralı kategorik veriler için alternatif olarak geliştirilmiş WLSMV (Weighted Least Squares Mean and Variance adjusted) gibi tahmin edicilerin de sürece dahil edilerek (Brauer vd., 2023), farklı tahmin yöntemlerinin faktör yükleri ve uyum indeksleri üzerindeki etkilerinin kültürlerarası bağlamda karşılaştırmalı olarak incelenmesi önerilmektedir.

Araştırma ve Yayın Etiği

Bu çalışmada "Yükseköğretim Kurumları Bilimsel Araştırma ve Yayın Etiği Yönergesi" kapsamında uyulması belirtilen tüm kurallara uyulmuştur. Yönergenin ikinci bölümü olan "Bilimsel Araştırma ve Yayın Etiğine Aykırı Eylemler" başlığı altında belirtilen eylemlerden hiçbiri gerçekleştirilmemiştir.

Etik Kurul İzni

Kurul adı = Çanakkale Onsekiz Mart Üniversitesi Sosyal ve Beşerî Bilimler Etik Kurulu
Karar tarihi=24.10.2025
Belge sayı numarası**=** E-80110942-010.99-2500281178

Yazar Katkı Oranları

Yazar 1: Kavramsallaştırma, araştırma problemlerini oluşturma, verileri toplama, veri analizi.
Yazar 2: Literatür taraması, veri analizi, araştırma sonuçlarını raporlaştırma, metni düzenleme.

Çıkar Çatışması

Yazarlar arasında çıkar çatışması yoktur.

Yapay Zekâ Kullanım Beyanı

İngilizce akademik çeviri desteği dışında araştırma sürecinde yapay zekâ araçlarından yararlanılmamıştır.

Kaynaklar

Alderson-Day, B., & Fernyhough, C. (2015). Inner speech: Development, cognitive functions, phenomenology, and neurobiology. *Psychological Bulletin*, *141*(5), 931-965. https://doi.org/10.1037/bul0000021

Alderson-Day, B., Mitrengaa, K., Wilkinson, S., McCarthy-Jones, S., & Fernyhough, C., (2018). The varieties of inner speech questionnaire - Revised (VISQ-R): Replicating and refining links between inner speech and psychopathology. *Consciousness and Cognition 65,*48-58. https://doi.org/10.1016/j.concog.2018.07.001

Arafat, S. M. Y. (2016). Cross cultural adaptation & psychometric validation of instruments: Step-wise. *International Journal of Psychiatry, 1*(1), 1-4. https://doi.org/10.33140/ijp/01/01/00001

Baharum, H., Ismail, A., Awang, Z., McKenna, L., Ibrahim, R., Mohamed, Z., & Hassan, N. H. (2023). Validating an instrument for measuring newly graduated nurses' adaptation. *International Journal of Environmental Research and Public Health, 20*(4), 2860. https://doi.org/10.3390/ijerph20042860

Boateng, G. O., Neilands, T. B., Frongillo, E. A., Melgar-Quiñonez, H. R., & Young, S. L. (2018). Best practices for developing and validating scales for health, social, and behavioral research: A primer. *Frontiers in Public Health, 6*. https://doi.org/10.3389/fpubh.2018.00149

Brauer, K., Ranger, J., & Ziegler, M. (2023). Confirmatory factor analyses in psychological test adaptation and development: A nontechnical discussion of the WLSMV estimator. *Psychological Test Adaptation and Development, 4*(1), 4-12. https://doi.org/10.1027/2698-1866/a000034

Brinthaupt, T. M., Hein, M. B., & Kramer, T. E. (2009). The self-talk scale: Development, factor analysis, and validation. *Journal of Personality Assessment*, *91*(1), 82-92. https://doi.org/10.1080/00223890802484498

Byrne, B. M. (2016). Adaptation of assessment scales in cross-national research: Issues, guidelines, and caveats. *International Perspectives in Psychology: Research, Practice, Consultation, 5*(1), 51-65. https://doi.org/10.1037/ipp0000042

Can, A. (2016). *SPSS ile bilimsel araştırma süreçlerinde nicel veri analizi* (4.Baskı). Ankara: Pegem Akademi.

Chella, A., & Pipitone, A. (2020). A cognitive architecture for inner speech. *Cognitive Systems Research, 59*, 287–292. https://doi.org/10.1016/j.cogsys.2019.09.010

Chen, W. J. (2012). Immediate effects of positive self-talk on stress and speech performance [Yüksek lisans tezi / Master's thesis, San José State University]. San José State University ScholarWorks.

Dana, A., & Brinthaupt, T. M. (2025). Application of inner Speech in the motor learning context: A validation study. *Health Nexus, 3*(4), 1-9. https://doi.org/10.61838/kman.hn.3.4.4

Dana A, Ghorbani S, M. & Brinthaupt T. (2024). Adaptation and validation of the learning-specific inner speech scale (LISS) to the motor learning context. *Phys. Act. Child, 1*(1), 58-64. https://doi.org/10.61186/pach.2024.465761.1015

de Guerrero, M. C. M. (2018). Private and inner speech in L2 learning: The impact of Vygotskyan sociocultural theory. In J. P. Lantolf, M. E. Poehner, & M. Swain (Eds.), *The Routledge handbook of sociocultural theory and second language development* (pp. 152–164). Routledge. https://doi.org/10.4324/9781315624747-10

Ekizoğlu, Ö., Acet, M. & Toros, C. (2024). The effect of professional volleyball players' mental endurance, imagination, and internal speech levels on the way they cope with stress. *Research in Sports Science (Online), 14*(2), 99 - 106. https://doi.org/10.5152/rss.2024.23024

Flanagan, R. M., & Symonds, J. E. (2022). Children's self-talk in naturalistic classroom settings in middle childhood: A systematic literature review. *Educational Research Review*, *35*, 100432. https://doi.org/10.1016/j.edurev.2022.100432

Fossa, P. (Ed.). (2022). *Inner speech, culture & education*. Springer.

Gjersing, L., Caplehorn, J. R. M., & Clausen, T. (2010). Cross-cultural adaptation of research instruments: Language, setting, time and statistical considerations. *BMC Medical Research Methodology, 10*(13). https://doi.org/10.1186/1471-2288-10-13

Heggestad, E. D., Scheaf, D. J., Banks, G. C., Monroe Hausfeld, M., Tonidandel, S., & Williams, E. B. (2019). Scale adaptation in organizational science research: A review and best-practice recommendations. *Journal of Management, 45*(6), 2596-2627. https://doi.org/10.1177/0149206319850280

Hidayat, Y., Yudiana, Y., Hambali, B., Sultoni, K., Ustun, U. D., & Singnoy, C. (2023). The effect of the combined self-talk and mental imagery program on the badminton motor skills and self-confidence of youth beginner student-athletes. *BMC Psychology, 11*(1). https://doi.org/10.1186/s40359-023-01073-x

Hoyle, R. H. (2000). Confirmatory factor analysis. In *Handbook of applied multivariate statistics and mathematical modeling* (pp. 465-497). Academic Press.

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, 6(1), 1-55. https://doi.org/10.1080/10705519909540118

International Test Commission. (2005). ITC guidelines for translating and adapting tests (Version 1.0). https://www.intestcom.org

Kline, R. B. (2015). *Principles and practice of structural equation modeling*. Guilford publications.

Kompa, N. A., & Mueller, J. L. (2024). Inner speech as a cognitive tool-or what is the point of talking to oneself? *Philosophical Psychology, 37*(8), 1971–1994. https://doi.org/10.1080/09515089.2022.2112164

Landin, D., & Hebert, E. P. (1999). The influence of self-talk on the performance of skilled female tennis players. *Journal of Applied Sport Psychology*, *11*(2), 263-282. https://doi.org/10.1080/10413209908404204

Langland-Hassan, P. (2014). Inner speech and metacognition: In search of a connection. *Mind & Language*, *29*(5), 511-533. https://doi.org/10.1111/mila.12064

Latinjak, A. T., Viladrich, C., Alcaraz, S., & Torregrosa, M. (2016). Spanish adaptation and validation of the Automatic Self-Talk Questionnaire for Sports. *International Journal of Sport and Exercise Psychology*, *14*(4), 402-413. https://doi.org/10.1080/1612197X.2015.1055287

McCarthy-Jones, S., & Fernyhough, C. (2011). The varieties of inner speech: Links between quality of inner speech and psychopathological variables in a sample of young adults. *Consciousness and Cognition*, *20*(4), 1586-1593. https://doi.org/10.1016/j.concog.2011.08.005

Morin, A. (2005). Possible links between self-awareness and inner speech: Theoretical background, underlying mechanisms, and empirical evidence. *Journal of Consciousness Studies, 12*(4-5), 115-134.

Morin, A., Duhnych, C., & Racy, F. (2018). Self-reported inner speech use in university students. *Applied Cognitive Psychology*, *32*(3), 376-382. https://doi.org/10.1002/acp.3404

Orhan, S. (2024). Futbolcularda sürekli sportif kendine güven ve içsel konuşmanın zihinsel dayanıklılık ile ilişkisi ve zihinsel dayanıklılığa etkisi: Kesitsel araştırma. *Türkiye Klinikleri Spor Bilimleri Dergisi , 16*(1), 81 - 88. https://doi.org/10.5336/sportsci.2023-97970

Pallant, J. (2020). *SPSS survival manual: A step by step guide to data analysis using IBM SPSS*. Routledge.

Pan, Y., Kim, H., Huang, Y., Hwang, G. J., & Tu, Y. F. (2025). Promoting students' metacognition and self-regulatory efficacy through metacognitive scaffolding in mobile technology-enhanced interactive classrooms: The moderating role of inner speech. *Education and Information Technologies*, *30*(8), 10811-10836. https://doi.org/10.1007/s10639-024-13284-9

Racy F., Morin, A. & Duhnych, C. (2020). Using a thought listing procedure to construct the general inner speech questionnaire: An ecological approach. *Journal of Constructivist Psychology*, *33*(4), 385–405. https://doi.org/10.1080/10720537.2019.1633572

Racy, F., Morin, A., & Hagerty, J. (2024). Concurrent validity and reliability of representative inner speech questionnaires. *Psychological Reports*, *127*(2), 994-1002. https://doi.org/10.1177/00332941221109109

Schermelleh-Engel, K., Moosbrugger, H., & Müller, H. (2003). Evaluating the fit of structural equation models: Tests of significance and descriptive goodness-of-fit measures. *Methods of Psychological Research, 8*(2), 23–74.

Steenkamp, J. B. E. M., & Maydeu-Olivares, A. (2023). Unrestricted factor analysis: A powerful alternative to confirmatory factor analysis. *Journal of the Academy of Marketing Science, 51*(1), 86-113. https://doi.org/10.1007/s11747-022-00888-1

Thomaes, S., Tjaarda, I. C., Brummelman, E., & Sedikides, C. (2020). Effort self-talk benefits the mathematics performance of children with negative competence beliefs. *Child Development, 91*(6), 2211–2220. https://doi.org/10.1111/cdev.13347

Ursachi, G., Horodnic, I. A., & Zait, A. (2015). How reliable are measurement scales? External factors with indirect influence on reliability estimators. *Procedia Economics and Finance, 20*, 679-686. https://doi.org/10.1016/S2212-5671(15)00123-9

Uttl, B., Morin, A., & Hamper, B. (2011). Are inner speech self-report questionnaires reliable and valid? *Procedia - Social and Behavioral Sciences, 30,* 1719-1723. https://doi.org/10.1016/j.sbspro.2011.10.332

Vygotsky, L. S. (1962). *Thought and language* (E. Hanfmann & G. Vakar, Eds. & Trans.). MIT Press. https://doi.org/10.1037/11193-000

Werani, A. (2018). Inner speech and its impact on teaching and learning. In *The Routledge handbook of sociocultural theory and second language development* (pp. 136-151). Routledge. https://doi.org/10.4324/9781315624747-9

Xiang, X., Brinthaupt, T. M., Sun, S., & Ren, X. (2019). The learning-specific inner speech scale (LISS). *European Journal of Psychological Assessment*, *36*(4). https://doi.org/10.1027/1015-5759/a000544

Zervas, Y., Stavrou, N. A., & Psychountaki, M. (2007). Development and validation of the Self-Talk Questionnaire (S-TQ) for Sports. *Journal of Applied Sport Psychology*, *19*(2), 142-159. https://doi.org/10.1080/10413200601185156

Zourbanos, N., Hatzigeorgiadis, A., Chroni, S., Theodorakis, Y., & Papaioannou, A. (2009). Automatic self-talk questionnaire for sports (ASTQS): Development and preliminary validation of a measure identifying the structure of athletes' self-talk. *The Sport Psychologist*, *23*(2), 233-251. https://doi.org/10.1123/tsp.23.2.233

Extended Abstract

Introduction

Inner speech is fundamentally defined as the subjective experience of language in the absence of overt articulation (Alderson-Day & Fernyhough, 2015). Rooted in Vygotsky's (1962) sociocultural theory, inner speech is recognized not merely as a passive thought process, but as an active mechanism driving cognition, problem-solving, self-regulation, metacognition, and motivation (Chella & Pipitone, 2020; Morin, 2005). Within educational contexts, inner speech serves as a crucial bridge between internalizing instruction and externalizing knowledge, facilitating metacognitive awareness and active learner engagement (Flanagan & Symonds, 2022; Fossa, 2022; Kompa & Mueller, 2024).

Despite its profound pedagogical relevance, measurement tools evaluating inner speech have predominantly focused on clinical psychology (e.g., investigating hallucination proneness or anxiety; Alderson-Day et al., 2018; McCarthy-Jones & Fernyhough, 2011; Racy et al., 2020, 2024; Uttl et al., 2011), cognitive neuroscience (Brinthaupt et al., 2009), and sports sciences (Dana & Brinthaupt, 2025; Latinjak et al., 2016; Zourbanos et al., 2009). Direct applications within educational sciences, particularly in teacher training, remain surprisingly scarce (Xiang et al., 2019). For preservice teachers, inner speech functions as a reflective self-dialogue that directly influences instructional decision-making, classroom management strategies, and affective domain control (Fossa, 2022; Werani, 2018). Therefore, adapting a robust, multidimensional instrument to measure inner speech in educational settings is essential. This study aimed to adapt the Varieties of Inner Speech Questionnaire-Revised (VISQ-R), originally developed by Alderson-Day et al. (2018), into Turkish, validating its psychometric properties on a sample of preservice teachers.

Method

This study employed a methodological research design following the cross-cultural adaptation guidelines established by the International Test Commission (ITC, 2005) and recommended best practices (Arafat, 2016; Byrne, 2016; Gjersing et al., 2010). The adaptation process was systematically executed across nine structured protocols:

1. Permissions & Ethics: Institutional ethical approval was obtained from Çanakkale Onsekiz Mart University Social and Human Sciences Ethics Committee (Date: 24.10.2025, No: E-80110942-010.99-2500281178), alongside permission from the scale developers.
2. Forward Translation: Two independent translators (one domain expert in Psychology/Guidance Counseling and one naive language specialist) translated the 26-item VISQ-R from English to Turkish (Forms T1 and T2).
3. Synthesis: The research team synthesized T1 and T2 into a single consensus draft (Form T-1).
4. Back Translation: Two bilingual experts, native or fluent in English, back-translated Form T-1 to English without prior sight of the original scale.

5. Expert Review: An interdisciplinary panel of methodologists, linguists, and educational scientists evaluated item content, semantic, conceptual, and experiential equivalence, yielding the experimental Turkish form (T-2).
6. Linguistic Equivalence (Pilot A): Administered to 50 bilingual participants using a counterbalanced design (Group A: English-to-Turkish, N=26; Group B: Turkish-to-English, N=24). Analyses included Wilcoxon Signed-Rank Test for order effect, Pearson correlation for total scores, Spearman rank correlation for item-level analysis, and Paired Samples t-test.
7. Cognitive Debriefing (Pilot B): Conducted with 10 undergraduate students and 1 Turkish language specialist to verify readability, clarity, and cultural appropriateness (Form T-3).
8. Main Application: Data were collected from 323 preservice teachers enrolled in the Faculty of Education at Çanakkale Onsekiz Mart University. The item-to-sample ratio was approximately 12.4 (323 participants / 26 items. Reverse-worded items (Items 7, 15, and 21) were recoded prior to statistical analyses.
9. Psychometric Analyses: Construct validity was evaluated via Confirmatory Factor Analysis (CFA) using Maximum Likelihood estimation. Reliability was examined using Cronbach's alpha (α) and McDonald's Omega (ω) internal consistency coefficients, alongside item-total correlations.

Result and Discussion

Linguistic equivalence analyses revealed that 23 out of 26 items showed no statistically significant differences between English and Turkish forms ($p > .05$), indicating an 88.5% baseline linguistic parity across languages. Minor variations observed in Items 12, 21, and 25 were attributed to cross-cultural linguistic nuances and item wording rather than translation errors (Byrne, 2016; Heggestad et al., 2019). Total score comparisons demonstrated an exceptionally strong, positive, and significant correlation between English and Turkish forms ($r = .935$, $p < .001$), sharing 87% common variance ($r^2 = .87$). Furthermore, a paired-samples t-test indicated no significant difference between the total mean scores of the English ($X = 4.22$, $SD = .74$) and Turkish ($X = 4.23$, $SD = .74$) versions ($t_{(49)} = -.225$, $p = .823$), indicating a strong and positive relationship between the linguistic structures of the two forms. Spearman item-level correlations ranged from .466 (Item8) to .867 (Item19) ($p < .001$).

To test construct validity, a five-factor CFA model (Dialogic, Evaluative/Critical, Other People, Condensed, and Positive/Regulatory) was executed. Due to insufficient factor loadings, Item 21 was removed from the model, and the final construct validity was confirmed over the remaining 25 items. Following theoretical considerations and modification indices, an error covariance was specified between a semantically similar item pair under the same latent factor (e11-e12) (Byrne, 2016; Gjersing et al., 2010; Kline, 2015). The final model fit indices demonstrated satisfactory parameter fit: χ^2 / df = 2.870, RMSEA = .076, CFI = .840, GFI = .840, TLI = .818, and SRMR = .094. Although comparative fit indices (CFI, TLI) fell below the conventional .90 and stringent .95 thresholds (Hu & Bentler, 1999), this is common when applying Maximum Likelihood estimation to ordinal Likert data with complex latent dimensions (Brauer et al., 2023; Kline, 2015). Crucially, the $\chi^2/df < 3.0$ and RMSEA < .08 values confirm acceptable model fit within reasonable approximation errors (Gjersing et al., 2010; Schermelleh-Engel et al., 2003). Standardized factor loadings ranged from .29 to .86 across items, with all parameters achieving statistical significance ($p < .001$).

Reliability analysis revealed strong internal consistency for the overall Turkish VISQ-R (α = .836, ω = .839). Subscale reliability coefficients were recorded as follows: Other People (α = .863, ω = .865), Dialogic (α = .789, ω = .793), Evaluative/Critical (α = .787, ω = .791), Positive/Regulatory (α = .753, ω = .757), and Condensed (α = .691, ω = .695). Although one subscale (Condensed) fell slightly below the .70 threshold (α = .691, ω = .695), shorter scales naturally tend to yield lower reliability

estimates (Heggestad et al., 2019), and values in the .60-.70 range are considered acceptable in cross-cultural scale adaptations (Ursachi et al., 2015).

In conclusion, the Turkish adaptation of the Varieties of Inner Speech Questionnaire-Revised (VISQ-R) provides a valid, reliable, and pedagogically sound instrument for measuring inner speech dynamics in Turkish educational contexts. Incorporating this tool into teacher education programs can facilitate deeper insights into preservice teachers' metacognitive self-regulation, instructional self-talk, and reflective practice (Dana & Brinthaupt, 2025; Pan et al., 2025). Future studies are recommended to test the instrument across more diverse student populations and evaluate alternative estimators, such as WLSMV, for ordinal data modeling (Brauer et al., 2023).

Ekler

Ek-1. İçsel Konuşma Ölçeği Türkçe Formu

	İÇSEL KONUŞMA ÖLÇEĞİ	**Hiçbir zaman**	**Neredeyse hiç**	**Nadiren**	**Bazen**	**Sık sık**	**Oldukça sık**	**Her zaman**
1	Kendi kendime düşünürken, tam cümleler yerine kısa ifadeler ve tek kelimeler kullanırım.							
2	Zihnimdeki şeyler/Aklıma takılanlar hakkında kendi kendime konuşurken, sanki kendime sorular sorup bunları cevaplıyormuşum gibi karşılıklı bir süreç yaşarım.							
3	Zihnimde başka bir kişinin sesini duyarım. Örneğin; belli bir şekilde davrandığımda, zihnimde annemin sesini duyarım.							
4	Zihnimde, bana sorular soran başka insanların seslerini duydum.							
5	Zihnimde, bana sürekli söylenen (dırdır eden) başka insanların seslerini duyarım.							
6	Sözcüklerle düşünme biçimim, bir monologdan ziyade kendimle bir diyalog gibidir."							
7	İçimden düşünürken/konuşurken tam cümleler kurarım.							
8	Kendi kendime sözcüklerle düşünme biçimim; tam, kurallı ve dil bilgisine uygun cümlelerden ziyade, kısaltılmış notlar gibidir.							
9	Yaptıklarım ve bunların doğru olup olmadığı hakkında içimden konuşarak düşünürüm.							
10	Zihnimdeki/ Aklımdaki şeyler hakkında konuşurken, sanki kendimle sohbet ediyormuşum gibi olur.							
11	Zihnimde sessizce konuşarak kendime bir şeyler yapmamı söylerim.							
12	Zihnimde, başkalarının bana daha önce hiç söylemedikleri şeyleri, bizzat kendi sesleriyle söylediklerini duyarım.							
13	Zihnimde bir şeyler hakkında kendimle karşılıklı konuşurum.							
14	Sözcüklerle düşünme biçimim, normal sesli konuşmama kıyasla kısaltılmıştır. Örneğin; içimden 'markete gitmem gerek' gibi şeyler söylemek yerine, sadece 'market' derim.							
15	Düşüncelerimi kağıda yazsaydım, dil bilgisine uygun, normal cümleler gibi okunurdu.							
16	Zihnimde, başkalarının bana daha önce bizzat söyledikleri							

	şeyleri söyleyen gerçek seslerini duyarım.							
17	İçimden konuşarak kendime bir şeyler yapmamamı söylerim.							
18	Davranışlarımı içimden konuşarak değerlendiririm. Örneğin; kendime 'bu iyiydi' veya 'bu saçmaydı' derim.							
19	İçimden, kendime cesaret verici bir şekilde konuşurum.							
20	Zihnimde, kendimle eleştirel bir şekilde konuşurum.							
22	İçimden kendimle konuşarak kendimi sakinleştiririm.							
23	İçimden söylediklerim beni kaygılandırır.							
24	İçsel konuşmam, kendimi mutsuz ve depresif hissetmemde rol oynar.							
25	Öfke anında, içsel konuşmam sakinleşmeme yardımcı olabilir.							
26	Üzücü şeyler hakkında düşünürken; zihnimde konuyu kolayca değiştirip kendimle başka şeyler hakkında konuşabilirim.							

Not: Orijinal İngilizce formda ters kodlu olan 21. madde, yapı geçerliliğini (DFA) sağlayamadığı için Türkçe nihai formdan çıkarılmıştır.

Ters Maddeler: 7,15
Alt Boyutlar:
Diyalojik: 2,6,10,13 **Sıkıştırılmış:** 1,7,8,14,15
Değerlendirici\Eleştirel: 9,11,17,18,20,23,24 **Pozitif\Düzenleyici:** 19,22,25,26
Diğer Kişiler: 3,4,5,12,16