




Zindelik Ölçeği Madde Havuzunun Yinelemeli Bilişsel Görüşmeler Yoluyla Test Edilmesi ve Geliştirilmesi



Hakan Kuru¹  Aykut Bulut²  Koray Kılıç³  Ece İbanoğlu⁴ 

MAKALE BİLGİLERİ

DOI: 10.29299/kefad.1697627

Yükleme: 13.05.2025

Düzeltilme: 10.10.2025

Kabul: 03.11.2025

Anahtar Kelimeler:

Zindelik Ölçümü,
Yapı Geçerliliği,
Ölçek Geliştirme,
Anlaşılabilirlik,
Ölçüm Hatası

ÖZ

Zindeliğin çok boyutlu bir yapı olarak ölçülmesinin gerekliliği kavramsal açıdan sağlam temellere dayanan ve pratikte anlaşılabilir ölçme araçlarının geliştirilmesini zorunlu kılmaktadır. Ancak, ölçek maddelerinin araştırmacılar tarafından kastedildiği biçimde algılanması, yapı geçerliliğinin sağlanması açısından kritik bir öneme sahiptir. Ayrıca yalnızca istatistiksel analizlere dayanarak madde geçerliliğinin sınanması, ölçme hatalarına yol açabilecek sınırlılıkları beraberinde getirebilmektedir. Söz konusu çalışma zindelik ölçeği madde havuzunun anlaşılabilirliğini yinelemeli bilişsel görüşmeler aracılığıyla test etmeyi ve iyileştirmeyi amaçlamaktadır. Bu doğrultuda, madde anlaşılabilirliği ve uygunluğunu artırmak amacıyla alan ve dil uzmanlarıyla (n=2) bilişsel görüşmeler gerçekleştirilmiştir. Devamında yinelemeli olarak düzenlenen madde havuzu üzerinden, amaçlı örnekleme yoluyla seçilen katılımcılarla (n=8) bilişsel görüşmeler yürütülmüştür. Katılımcıların maddeleri nasıl yorumladıklarını değerlendirmek amacıyla sesli düşünme ve sözel sondalama teknikleri uygulanmıştır. Elde edilen bulgular terimlerin kavramsal karmaşıklığı, madde uzunlukları, yanıt kategorilerindeki tutarsızlıklar ve hatırlama güçlükleri gibi çeşitli sorun alanlarını ortaya koymuştur. Tespit edilen sorunların giderilmesi amacıyla, ilk aşamayı takiben dil sadeleştirilmiş, uzun maddeler yeniden yapılandırılmış, yanıt seçenekleri katılımcıların yanıtlamalarını kolaylaştırmak amacıyla düzenlenmiştir. Yinelemeli iyileştirme süreci katılımcıların maddeyi anlama düzeylerinde ve yanıtlarının doğruluğunda belirgin gelişmeler sağlamıştır. Araştırma bulguları ölçeklerin geçerliliğinin test edilmesinde bilişsel görüşmeler arasında yinelemeli madde havuzu kullanımının önemine işaret etmektedir. Gelecekte yapılacak araştırmalar farklı alanlardaki farklı ölçek geliştirme çalışmalarında, ölçme geçerliliğini artırmak amacıyla bilişsel görüşmelerde yinelemeli yaklaşımın uygulanması dikkate alınabilir. Uygulanan yinelemeli yöntemin etkinliğinin anlaşılabilmesi için uzun vadede deneysel çalışmalara ihtiyaç duyulmaktadır.

1. Giriş

Zindelik bireyin fiziksel, zihinsel, duygusal, sosyal, mesleki ve ruhsal (spiritüel) yönlerini kapsayan bir esenlik halidir (National Wellness Institute; NWI, 2020). NWI zinde olma sürecini bireyin potansiyelini gerçekleştirmeyi amaçlayan bilinçli, öz-yönelimli ve sürekli gelişen bir süreç olarak tanımlamaktadır. Ortaya koyulan bu tanım sağlığın korunmasını ve bireyin yaşamının her alanında kendini geliştirmeye yönelik etkin katılımını içermekte; bunun sonucunda da tatmin edici ve anlamlı bir yaşamın

temellerini oluşturmaktadır. Zindelik, bireylerin yaşam kalitesini bütüncül bir biçimde artıran ve yaşamlarına dengeleyici bir çerçeve sunan bir sağlık modeli olarak kabul edilmektedir (Dunn, 1959). Bu yaklaşım sağlığı yalnızca hastalıkların önlenmesi olarak görmenin ötesine geçerek, sağlıklı, doyurucu ve etkin bir yaşam arayışına dayalı yeni bir sağlık anlayışı olarak yansıtmaktadır (Roscoe, 2009). Görülen bu paradigmatik değişim bireylerin yaşam kalitesini yükseltmeye yönelik çok

Sorumlu Yazar¹: Hakan Kuru, Dr., İstanbul Rumeli Üniversitesi, Türkiye, hakan.kuru@rumeli.edu.tr

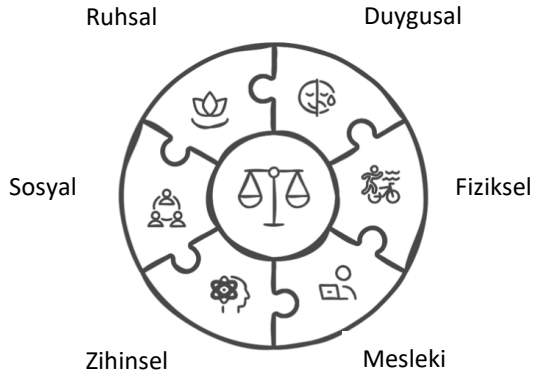
Yazar²: Aykut Bulut, Dr., Kırşehir Ahi Evran Üniversitesi, Türkiye, aykut.bulut@ahievran.edu.tr

Yazar³: Koray Kılıç, Dr., Kırşehir Ahi Evran Üniversitesi, Türkiye, koray.kilic@ahievran.edu.tr

Yazar⁴: Ece İbanoğlu, Doktora Öğrencisi, Utah State University, USA, ecce.ibanoğlu@gmail.com

boyutlu bir model olarak zindelik kavramının benimsenmesine zemin hazırlamıştır.

NWI'nin Altı Boyutlu Zindelik Modeli, bireylerin yaşam kalitesini sürdürülebilir biçimde artırmaya yönelik çok boyutlu bir çerçeve sunmaktadır. Hettler (1984) tarafından geliştirilen bu model, duygusal, fiziksel, mesleki, zihinsel, sosyal ve ruhsal (spiritüel) zindeliği desteklemeye yöneliktir. Zindeliği oluşturan boyutların her biri, bireyin yaşam kalitesine kendine özgü katkılar sunmakta ve sağlıklı, dengeli bir yaşam sürmelerine olanak tanımaktadır (Myers ve Sweeney, 2004; Roscoe, 2009). Zindelik kavramının her bir boyutuna odaklanmanın amacı bireyde bütüncül bir öz-gelişimi teşvik etmektir. Bu yönüyle Altı Boyutlu Model, yalnızca fiziksel sağlığı desteklemekle kalmamakta aynı zamanda kapsamlı bir zindelik hali durumunun oluşmasına da katkı sunmaktadır (bkz. Şekil 1).



Şekil 1. Altı boyutlu zindelik modeli (NWI, 2020)

Not. Altı Boyutlu Zindelik Modelinden uyarlanmıştır (NWI, 2020)

Altı Boyutlu Zindelik Modeline göre duygusal zindelik, bireyin duygularını tanıması, kabul etmesi, yönetmesi ve başkalarıyla sağlıklı ilişkiler kurması sürecini ifade etmektedir. Bununla beraber NWI (2020) duygusal zindeliği, bireyin kendi duygularını ve bu duygusal durumla ilişkili davranışlarını tanıma, kabul etme ve yönetme kapasitesi olarak tanımlamaktadır. Duygusal zindelik boyutunda birey, kendine ve yaşama yönelik olumlu bir bakış açısı geliştirirken, kendi duygularını fark etmekte ve ayrıca başkalarının duygularını da dikkate almaktadır. Stresle başa çıkmada özerkliğin desteklenmesi, iyimser bir bakış açısının benimsenmesi ve ilişkilerde açık, dürüst ve sağlıklı iletişimin sürdürülmesi duygusal zindeliğin temel unsurları arasında yer almaktadır.

Diğer taraftan fiziksel zindelik ise fiziksel öz-bakımı önceliklendirme ve sağlığı destekleyici fiziksel davranışlara sürekli olarak katılım sağlama sürecidir. Düzenli fiziksel aktivite, dengeli beslenme, yeterli uyku düzeni, dinlenme ve ihtiyaç duyulduğunda tıbbi desteğe erişim bu boyutun altında kapsamaktadır. Bireyler bu

boyutta, fiziksel sağlıklarını koruyarak sorunları tanımlamakta ve önleyici eylemlerde bulunmaktadır.

Zihinsel zindelik, bireyin öğrenme, kişisel gelişim ve yaratıcılık içeren uyarıcı etkinlikler aracılığıyla becerilerini paylaşma farkındalığı şeklinde ifade etmektedir. Yaşam boyu öğrenmeye yönelik bağlılık ve zihinsel gelişime duyulan istek bu boyutun temelini oluşturmaktadır. Zihinsel zindelik, bireylerin yeni hobiler edinmesine, ilgi alanlarını geliştirmesine, güncel olaylardan haberdar olmasına ve yeni düşünme biçimlerini keşfederek öğrenme süreçlerini sürekli olarak genişletmesi gibi durumları temel unsurları arasında bulundurmaktadır.

Mesleki zindelik, bireyin iş, akademik ya da gönüllü faaliyetler aracılığıyla tatmin ve zenginlik duygusu yaşamasını ifade etmektedir. Bireyler, değerleri ve ilgi alanlarıyla uyumlu iş ve etkinlikleri tercih ederek kendilerine olan güvenlerini artırmakta ve başarı duygusu kazanmaktadır. Mesleki zindelik bireylerin anlamlı ve doyum sağlayıcı faaliyetlere katılarak hedeflerini gerçekleştirmelerine zemin hazırlamaktadır.

Ruhsal (Spiritüel) zindelik, yaşamın daha derin anlamını fark etme, takdir etme ve evrenden gelen daha büyük doğa güçleriyle bağlantı kurma sürecini kapsamaktadır. Ruhsal zindelik boyutunda birey, yaşamın anlamını sorgulamakta ve kendi değerleriyle uyumlu bir yaşam amacı geliştirmektedir. Ruhsal (spiritüel) zindelik, bireylerin iç huzurla uyumlu bir yaşam sürmelerini, yaşam tarzlarını değerlerine göre şekillendirmelerini ve başkalarının inançlarına açık olmalarını mümkün kılmaktadır.

Sosyal zindelik boyutu ise bireyin içinde bulunduğu çevreye ve topluma katkı sunmasını, başkalarıyla ve doğayla olan karşılıklı bağı anlamasını ifade etmektedir. Sosyal yönden güçlü bireyler, topluma olumlu katkılarda bulunmakta, sosyal ilişkilerini güçlendirmekte ve çevresel kaynakların korunmasına yönelik bir sorumluluk duygusu taşımaktadır. Sosyal zindelik, bireyleri topluma entegre olmaya, sosyal farkındalıklarını artırmaya ve yaşadıkları ortamlarda etkin biçimde rol almaya teşvik etmektedir (NWI, 2020).

Zindeliğin çok boyutlu biçimde ölçülmesi, sağlık ve yaşam kalitesinin anlaşılabilmesi açısından büyük önem taşımaktadır (Kuru, 2022; Myers vd., 2017). Aynı zamanda zindelik ölçümleri, bireylerin güçlü ve gelişime açık yönlerini fark etmelerine olanak sağlayarak sağlık hedeflerini belirlemelerine yardımcı olmaktadır (Roscoe, 2009). Öznel bir biçimde gerçekleştirilen zindelik değerlendirmeleri, günlük yaşamda odaklanılması gereken hedeflerin daha sağlıklı biçimde belirlenmesini sağlamakta; böylece sağlıklı yaşam amaçlarına yönelik daha uygulanabilir bir yol haritasının oluşturulmasına ve sağlık stratejilerinin geliştirilmesine katkıda bulunmaktadır (Myers vd., 2000). Miller ve Foster'ın (2010) da belirttiği üzere, bireyin mevcut sağlık durumunu anlaması ve bu durumu iyileştirmeye yönelik stratejiler geliştirmesi açısından zindeliğin ölçülmesi temel bir rol

üstlenmektedir. Ayrıca, zindelik kavramının kapsamının anlaşılması, kamu sağlığı politikalarının etkinliğini artırmaya yönelik stratejilerin geliştirilmesi açısından da kritik bir öneme sahiptir (Dünya Sağlık Örgütü; WHO, 2018). Bireylerin sosyal, duygusal ve ruhsal zindeliklerinin ölçülmesi, toplumsal sağlığı geliştirmeye yönelik olası halk sağlığı müdahalelerine zemin hazırlayacak kanıtlara ulaşılmasını mümkün kılmaktadır. İyi oluş değerlendirmelerinden elde edilen bulgular, halk sağlığına yönelik müdahalelerin planlanması, uygulanması ve etkilerinin değerlendirilmesi süreçlerinde temel bir dayanak oluşturmaktadır (Haldane vd., 2019).

Mevcut çalışma Zindelik Ölçeği'nin geliştirilmesine yönelik madde havuzunun anlaşılabilirliğini test etmeyi ve iyileştirmeyi amaçlamaktadır. Bu doğrultuda madde havuzunun anlaşılabilirliğini test etmek ve geliştirmek amacıyla yinelemeli (iteratif) bilişsel görüşme süreci uygulanmıştır. Willis (2015) ve Collins (2003), bilişsel görüşme yönteminin, ölçek ve anket geliştirme süreçlerinde dilsel açıklığı artırma ve kavramsal uyumu güçlendirme açısından etkili bir araç olduğunu vurgulamaktadır. Daha önce yapılan araştırmalarda bilişsel görüşmelerin özellikle sözcük ve deyimlerin anlaşılmasında zorluk yaşayan genç bireylerle uygulanan ölçme araçlarının işlevselliğini artırmada etkili olduğunu göstermiştir (Kılıç ve İnce, 2016). Bu nedenle mevcut çalışma, bilişsel görüşme yöntemini yinelemeli bir yaklaşımla uygulayarak, İyilik Hali Ölçeği geliştirme sürecini sistematik biçimde iyileştirmeyi ve ölçek maddelerinin kültürel ve kavramsal uyumunu güçlendirmeyi amaçlamaktadır. Özellikle ölçek maddelerinin katılımcılar tarafından nasıl anlaşıldığını ve yanıt sürecinde karşılaştıkları güçlükleri daha iyi anlamaya yönelik olarak önerilen bilişsel görüşme yaklaşımına yinelemeli bir boyut entegre etmektedir (Beatty ve Willis, 2007). Uygulanan bu yöntemle, ölçek geçerliğini artırmaya yönelik ek iyileştirmeler yapılması hedeflenmektedir.

Mevcut çalışmanın özgün metodolojik yönü, zindelik ölçeği madde havuzunun geliştirilmesi sürecinde yinelemeli bilişsel görüşme yönteminin kullanılmış olmasıdır. Uygulanan metodolojinin önemi zindelik gibi çok boyutlu yapıları ölçmeyi amaçlayan ölçeklerin kalitesini artırarak elde edilen puanların geçerlik ve güvenilirliğini iyileştirilmesi katkı vermesidir. Bununla birlikte mevcut çalışma bilişsel görüşme yönteminin geçerli veriler elde etmede etkili bir araç olduğunu vurgulayan alanyazın (Rubin, 2004) ek olarak, veri geçerliğini sağlamak amacıyla söz konusu yöntemi birden fazla aşamada ve yinelemeli bir süreç dâhilinde uygulaması açısından özgün bir yaklaşım ortaya koymaktadır.

2. Yöntem

2.1. Araştırmanın Deseni

Mevcut çalışmada zindelik halini ölçmeye yönelik geliştirilen soru havuzunun anlaşılabilirliğini incelemek için nitel yöntem benimsenerek yinelemeli bilişsel görüşmeler gerçekleştirilmiştir. Bilişsel görüşmelerin ilk aşamasında soru maddelerinin anlaşılabilirliği uzman bir grup tarafından değerlendirilmiş ve elde edilen bulgular doğrultusunda maddeler revize edilmiştir. İkinci aşamada ise gözden geçirilen sorular farklı bir katılımcı grubuyla yeniden değerlendirilerek nihai hâline getirilmiştir. Yinelemeli yaklaşımın temel amacı, her aşamada elde edilen geri bildirimlere dayanarak soru formunu kademeli olarak geliştirmek ve anlaşılabilirliği en üst düzeye çıkarmaktır.

2.2. Çalışma Grubu

Çalışma toplam 10 katılımcı ile gerçekleştirilmiştir. Birinci aşamada bir dil ve bir alan uzmanı ($n=2$) yer almış, ikinci aşamada ise yinelemeli madde havuzu kullanılarak gerçekleştirilen bilişsel görüşmelere sekiz katılımcı katılmıştır. Sekiz kişilik örneklem, bilişsel görüşmelerde veri doygunluğuna ulaşılması için yeterli bulunmuş; görüşmeler sonucunda yeni tema veya sorun alanı ortaya çıkmadığından. Katılımcılar; yaş, cinsiyet, eğitim düzeyi ve meslek çeşitliliğini sağlamak amacıyla ölçüt örnekleme yöntemiyle seçilmiştir (Patton, 2002). Örneklem; dil ve alan uzmanları, öğrenciler ve çeşitli meslek gruplarından bireyleri içerecek şekilde farklı geçmişlere sahip bireyleri temsil edecek biçimde oluşturulmuştur. Tüm katılımcılar 18 yaş ve üzerindedir (bkz. Tablo 1).

Tablo 1.

Katılımcıların özellikleri

	Katılımcı No	Yaş	Cinsiyet	Eğitim Düzeyi
Adım 1	1	48	Kadın	Yüksek Lisans
	2	39	Erkek	Doktora
	3	37	Erkek	Yüksek Lisans
	4	19	Kadın	Lise
Adım 2	5	21	Erkek	Lise
	6	52	Kadın	Lisans
	7	32	Kadın	Doktora
	8	42	Erkek	Lisans
	9	33	Kadın	Lisans
	10	44	Erkek	Yüksek Lisans

2.3. Zindelik Ölçeği Soru Havuzu

Ölçek madde havuzunun geliştirilme süreci çok aşamalı bir yöntem izlenerek yürütülmüştür. İlk aşamada, Altı Boyutlu Zindelik Modeli temel alınarak geniş kapsamlı bir madde havuzu oluşturulmuştur. Literatürde yer alan ölçeklerden ve kuramsal tanımlardan yararlanılarak duygusal zindelik boyutunu temsil eden 8, zihinsel zindelik için 7, fiziksel zindelik için 7, sosyal zindelik için 8, ruhsal zindelik için 7

ve mesleki zindelik için 8 maddeden toplam 45 maddelik bir havuz ortaya çıkmıştır.

İkinci aşamada, madde havuzu davranışsal tasarım, eğitim bilimleri, spor bilimleri ve psikoloji alanlarından üç uzman tarafından bağımsız olarak değerlendirilmiştir. Uzmanlar maddeleri dilsel açıklık, kavramsal uygunluk ve kültürel bağlama uyum ölçütleri doğrultusunda incelemiştir. Yapılan değerlendirmeler sonucunda 12 madde dilsel açıdan yeniden yazılmış, 6 madde kavramsal belirsizlik nedeniyle elenmiş ve kalan maddeler üzerinde küçük revizyonlar yapılmıştır. Örneğin, “Güçlü ve zayıf taraflarımın farkındayım” ifadesi birden fazla boyutu aynı anda kapsadığı gerekçesiyle eleştirilmiş ve “Duygusal olarak güçlü olduğumu düşünüyorum” biçiminde yeniden düzenlenmiştir.

Uzmanların disiplinler arası katkıları, maddelerin çok boyutlu olarak test edilmesini sağlamıştır. Davranışsal tasarım bakış açısı ifadelerin davranışsal geçerliliğini güçlendirmiş, spor bilimleri bakış açısı fiziksel boyutun ölçme hassasiyetini artırmış, psikoloji bakış açısı ise duygusal ve bilişsel boyutlarda içerik geçerliliğini güvence altına almıştır.

Son aşamada, uzman görüşleri birleştirilerek revizyonlar tamamlanmış ve ölçeğin 39 maddeden oluşan nihai formu oluşturulmuştur. Takip edilen süreç ölçeğin içerik geçerliliğini pekiştirmiş, kavramsal ve dilsel tutarlılığı artırmış ve kültürel bağlama uygunluğunu güvence altına almıştır.

2.4. Veri Toplama Araçları

Bilişsel görüşmeler bire bir biçimde, yüz yüze ya da çevrim içi platformlar aracılığıyla gerçekleştirilmiştir. Görüşmeler sırasında ses kaydı alınmış ve ardından görüşmelerin yazılı şekilde transkripti yapılmıştır. Her bir görüşme yaklaşık 32 ila 53 dakika arasında sürmüştür. Görüşmeler sırasında katılımcıların soruları yanıtlarken yaşadıkları güçlükler, soruların anlaşılabilirliğine ilişkin değerlendirmeleri ve önerileri ayrıntılı biçimde not edilmiştir.

2.4.1. Bilişsel görüşmeler

Bilişsel görüşmeler, katılımcılardan sözel veri toplama sürecini içermektedir. Süreçte katılımcılardan, araştırmacının amaçladığı biçimde anlaşılıp anlaşılmadığını değerlendirmek ya da yanıtların kalitesini incelemek amacıyla anket sorularını yanıtlamaları istenir ve düşünme sürecine yönelik sorular yöneltilir (Beatty ve Willis, 2007). Psikososyal olguları anlama amacına hizmet etmesi nedeniyle bilişsel görüşmeler psikolojik bir yöntem ya da nitel görüşme türü olarak kabul edilmektedir (Collins, 2003). Anketlerde uygulanan geleneksel psikometrik testler genellikle güvenilirlik ve geçerliği ölçer; ancak bu testler, katılımcıların a) soruları doğru biçimde anladıkları, b) tüm katılımcıların soruları aynı şekilde yorumladıkları ve c) katılımcıların bu soruları doğru biçimde yanıtlayacak istek ve yeterliliğe sahip oldukları varsayımlarına dayanır (Collins, 2003). Bu nedenle

psikometrik testler, yanıt sürecini engelleyebilecek bazı sorunları tespit etse de, bu sorunların altında yatan nedenleri açıklamada ya da gizil sorunlara ilişkin kanıt sunmada yetersiz kalabilmektedir. Fakat bilişsel görüşmeler, bireylerin anket sorularını neden belirli biçimlerde yanıtladıklarını anlamaya yardımcı olurken aynı zamanda gözden kaçabilecek önemli yapıları ve konuya ilişkin hatalı varsayımları ortaya çıkarmaya olanak tanımaktadır (Desimone ve Le Floch, 2004).

Willis (2015) bilişsel görüşme tasarımlarında kullanılan iki temel sözel raporlama tekniğini “sesli düşünme (think-aloud)” ve “sözel sorgulama (verbal probing)” olarak tanımlamaktadır. Sesli düşünme tekniğinde katılımcılar, anket sorularını okurken ve yanıtlarken zihinsel süreçlerini sesli biçimde ifade eder. Bu süreçte görüşmeci katılımcının düşüncelerine müdahale etmeden, tarafsız kalmayı sürdürerek önyargısız ve etkilenmemiş düşünceleri yakalamayı hedefler. Sözel sondalama tekniği ise katılımcının anket sorularına nasıl yaklaştığını anlamak amacıyla önceden hazırlanmış sorguların yöneltilmesini içerir. Sesli düşünme katılımcı merkezli bir teknik olup zaman zaman katılımcıyı zorlayabilirken; sözel sondalama görüşmeci merkezli bir tekniktir ve süreci katılımcı açısından daha kolay hâle getirebilir.

Mevcut çalışmada sesli düşünme ve sözel sondalama odaklı bilişsel görüşmelerde yinelemeli (iteratif) yaklaşımın tercih edilmesinin temel nedeni, ölçme araçlarını sistematik ve katılımcı odaklı bir perspektifle geliştirme potansiyelidir. Geleneksel doğrulama yöntemlerinden farklı olarak yalnızca tek seferlik bilişsel görüşmeye dayanmayan yinelemeli tasarım; birden fazla veri toplama ve iyileştirme döngüsünü içermekte, böylece belirsizlikler, yanlış anlamalar ve istenmeyen soru yapıları sürekli olarak gözden geçirilip giderilebilmektedir. Mevcut çalışmada kullanılan yinelemeli süreç, her görüşme aşamasında elde edilen katılımcı geri bildirimlerine göre ölçek maddelerinin revize edilmesini ve güncellenen ölçeğin yeni bir katılımcı grubuyla tekrar test edilmesini içermektedir. Tekrar eden yineleme süreci araştırmacılara ölçeği her aşamada katılımcı geri bildirimlerine dayalı olarak iyileştirme olanağı tanımaktadır. Rubin (2004), test etme ve yeniden düzenleme sürecinin yinelemeli olarak yürütülmesinin, anket ve ölçek tasarımlarında dilsel ve kavramsal nüansların daha hassas biçimde yakalanmasına katkı sağladığını ileri sürmektedir. Böylelikle süreci yeniden biçimlendirerek yalnızca ilk görüşmeyle sınırlı kalmayan döngüsel bir gözden geçirme mekanizması oluşturur; böylece her adımda ölçeğin geçerlik ve güvenilirliği daha da güçlenmektedir.

Adım 1: Bilişsel Görüşmeler

Birinci aşamada iki uzman ile “sesli düşünme” ve “sözel sondalama” teknikleri kullanılarak bilişsel görüşmeler gerçekleştirilmiştir. Görüşme sırasında katılımcılardan soruları yanıtlarken yüksek sesle okumaları ve zihinsel süreçlerini ifade etmeleri istenmiştir. Ayrıca, soruların

anlaşılabilirliğini, yeterince açık olup olmadıklarını ve yanıtlamada herhangi bir güçlük yaşanıp yaşanmadığını anlamak amacıyla sözel sondalama tekniklerine uygun ek sorular yöneltilmiştir. Bu aşamada elde edilen veriler doğrultusunda sorular gözden geçirilmiş ve gerekli düzenlemeler yapılmıştır.

Adım 2: Yenilenen Madde Havuzuyla Bilişsel Görüşmeler

İkinci bir katılımcı grubuyla ($n=8$) bilişsel görüşmeler gerçekleştirilmiştir. Mevcut aşamada da uzmanlarla yapılan bilişsel görüşmelerde kullanılan, sesli düşünme ve sözel sorgulama teknikleri aynen uygulanmıştır. Ancak bu kez odak noktası, gözden geçirilmiş madde havuzunun anlaşılabilirliğinin artıp artmadığını ve önceki geri bildirimlere dayanarak yapılan değişikliklerin etkisini değerlendirmek olmuştur.

Bilişsel Görüşme Prosedürü

Sesli Düşünme/Genel: Katılımcıların her bir soruya yanıt verirken sesli düşünme tekniği aracılığıyla ifade ettikleri genel zihinsel süreçleri incelenmiştir. Katılımcıların soruları okurken karşılaştıkları güçlükler, sorunun amacını anlamadaki belirsizlikler ya da yanıt verirken yaptıkları çıkarımlar üzerinde durulmuştur. Willis (2015) ise sesli düşünme tekniğinin katılımcıların bilişsel süreçlerinin doğrudan gözlemlenmesine imkân tanıdığını vurgulamaktadır.

Sözel Sondalama/Kavrama: Katılımcıların soruların yapı ve içeriğini doğru anlayıp anlamadıklarını değerlendirmek amacıyla, anlama güçlüğü yaşadıkları bölümler her bir soru için tematik olarak kodlanmıştır. Collins (2003) kavrama sürecinin bilişsel işlemin temel bir bileşeni olduğunu ve katılımcıların soruyu nasıl yorumladıklarının anlaşılmasının kritik önem taşıdığını vurgulamaktadır.

Sözel Sondalama/Hatırlama: Katılımcıların geçmiş deneyimlerine dayalı olarak belleklerinden ilgili bilgileri hatırlama süreçleri analiz edilmiştir. Tourangeau'ya (1984) göre, bu aşama anket yanıtlarının güvenilirliği açısından

kritik öneme sahiptir. Bazı katılımcıların belirli sorularda özel bilgileri hatırlamakta zorlandıkları gözlemlenmiş, bu durum bazı sorularda yeterli açıklığın bulunmadığını göstermiştir. Özellikle zaman ifadesi içeren sorular (örneğin, "Son altı ay içinde ne sıklıkla...?") katılımcılar için güçlük oluşturmuş ve bu sorular, hatırlamayı kolaylaştıracak şekilde yeniden yapılandırılmıştır.

Sözel Sondalama/Değerlendirme: Bu aşamada, katılımcıların hatırladıkları bilgileri nasıl değerlendirdikleri ve soruya uygun yanıtlar verip vermedikleri incelenmiştir. Odak noktası, katılımcıların yanıtlarını nasıl değerlendirdiklerini ve bu yanıtların sorunun amacıyla ne derece örtüştüğünü belirleme süreçlerini anlamak olmuştur. Beatty ve Willis (2007) değerlendirme aşamasının anket sonuçlarının doğruluğu üzerinde önemli bir etkiye sahip olduğunu belirtmiştir.

Sözel Sondalama/Yanıt: Katılımcıların verdikleri cevapların uygunluğu ve cevaplandırma sürecinde karşılaştıkları güçlükler analiz edilmiştir. Bu aşamada özellikle, katılımcıların sunulan yanıt seçeneklerini yeterince kapsayıcı bulup bulmadıkları değerlendirilmiştir. Desimone ve Le Floch (2004), açık ve kapsayıcı yanıt seçeneklerinin anketin genel geçerliliği üzerinde doğrudan etkili olduğunu vurgulamaktadır.

Sözel Sondalama/Kendini İfade Etme: Katılımcılardan, bir maddeyi yanlış anlamaları ya da yanıtlamada güçlük yaşamaları durumunda, ilgili soruyu nasıl anladıklarını ve nasıl yanıtladıklarını kendi ifadeleriyle açıklamaları istenmiştir. Bu adım, katılımcıların soruyu araştırmacının amaçladığı biçimde anlayıp anlamadıklarını belirlemek amacıyla uygulanmıştır. Rubin (2004), sürecin katılımcıların soruları farklı şekillerde yorumlama potansiyelini ortaya koyduğunu belirtmektedir. Özellikle karmaşık ya da uzun soruların katılımcılar tarafından kendi ifadeleriyle yeniden yapılandırılması, bu soruların daha anlaşılır hâle getirilmesi açısından yapıcı olmuştur (bkz. Tablo 2).

Tablo 2.

Bilişsel görüşme prosedürü

Bilişsel Görüşme	Amaç	Örnek Soru
Sesli düşünme	Genel	Bu soruyu yanıtlarken aklınızdan neler geçtiğini anlatabilir misiniz? Bu soruyu ne derece kolay ya da zor buldunuz? Neden?
Sözel Sondalama	Kavrama	"X" terimi sizin için ne ifade ediyor? Bu sorunun amacını nasıl yorumladınız?
Sözel Sondalama	Hatırlama	Bu soruyu yanıtlarken aklınıza belirli bir zaman ya da dönem geldi mi? Bu soruyla ilgili olarak hatırladığınız belirli bir örnek var mı?
Sözel Sondalama	Değerlendirme	Verdiğiniz yanıttan ne derece eminsiniz? Yanıtınızın, soruda belirtilen durumu ne ölçüde yansıttığını düşünüyorsunuz?
Sözel Sondalama	Yanıt	İlk yanıtınızı en iyi yansıtan seçeneği bulmak sizin için kolay oldu mu? Bu soruyu yanıtlarken kendinizi nasıl hissettiniz?
Sözel Sondalama	Kendini İfade Etme	Bu soruyu kendi ifadelerinizle nasıl yeniden ifade ederdiniz? Bu sorunun anlamını farklı bir şekilde açıklayabilir misiniz?

Her bir bilişsel görüşme aşamasından elde edilen geri bildirimler, davranışsal tasarımcı, spor bilimci ve spor psikoloğundan oluşan üç araştırmacının yer aldığı bir panelde tartışılmıştır. Bu panel, katılımcıların yaşadığı sorunları ayrıntılı bir şekilde değerlendirmiş ve soruların revize edilmesine yönelik oybirliğiyle kararlar almıştır. Willis (2015), bu tür panel tartışmalarının soru geliştirme sürecinde etkili bir yöntem olduğunu ve farklı bakış açılarının ölçek geliştirme sürecine önemli katkılar sağladığını belirtmiştir. Panelde ele alınan her bir konu, katılımcı geri bildirimlerine dayanarak soruların ifadelerinde ya da yanıt seçeneklerinde gerekli düzenlemelerin yapılmasına yol açmıştır.

2.5. Veri Analizi

Bilişsel görüşmelerden elde edilen veriler, her bir soru özelinde tematik analiz yöntemiyle ayrıntılı olarak incelenmiştir (bkz. Tablo 3). Tematik analiz, verilerdeki örüntüleri tanımlamak, analiz etmek ve yorumlamak

amacıyla kullanılan nitel bir analiz yöntemidir ve altı temel aşamadan oluşur. İlk aşamada, tüm görüşme kayıtları yazılı metne dönüştürülmüş, ardından veriler tekrar tekrar incelenerek derinlemesine bir anlayış geliştirilmiştir. İkinci aşamada, tümevarımsal bir yaklaşımla görüşme dökümleri kodlanarak ilk kodlar oluşturulmuştur. Üçüncü aşamada, ilişkili kodlar gruplanarak verilerde anlamlı örüntüler yakalanmış ve temalar belirlenmiştir. Dördüncü aşamada, temaların iç tutarlılık ve dış ayrışma özelliklerini sağlayacak şekilde yinelemeli bir süreçle verilerle uyumu test edilmiştir. Beşinci aşamada, temalar açık biçimde tanımlanmış ve adlandırılmıştır. Son aşamada ise, araştırma amaçlarıyla temalar arasında bağ kuran anlatı temelli bir sentez oluşturulmuştur. Bu süreç, verilerin karmaşıklığına uyum sağlayabilecek esnekliğe sahip olmakla birlikte sistematik ve anlamlı bir analiz yürütülmesini sağlamıştır (Braun ve Clarke, 2006; Braun vd., 2016).

Tablo 3.

Tematik analizin altı aşaması (Braun ve Clarke, 2006)

Aşama	Açıklama	Uygulama
Veriye aşinalık	Verilerin transkripti yapılır, tekrar tekrar okunarak ilk izlenimler kaydedilir. Araştırmacı veriye derinlemesine hâkim olur.	Görüşme kayıtları yazılı hale getirilmiş, her bir transkript birden fazla kez incelenerek katılımcıların anlama güçlükleri, kavramsal sorunlar ve yanıt tutarsızlıkları hakkında ilk notlar çıkarılmıştır.
İlk kodların oluşturulması	Verilerden anlamlı birimleri temsil eden başlangıç kodları çıkarılır. Bu aşamada olabildiğince kapsayıcı bir kodlama yapılır.	Katılımcı ifadeleri ayrıntılı şekilde kodlanmış, özellikle "farkındalık", "duygusal dayanıklılık" gibi belirsiz kavramlara yönelik yorumlar ayrı kodlar olarak belirlenmiştir.
Temaların aranması	Kodlar arasında örüntüler aranır, benzer kodlar bir araya getirilerek temalar oluşturulur.	Benzer katılımcı yorumları gruplandırılmış, dilsel açıklık, hatırlama güçlüğü, yanıt seçeneklerinin belirsizliği ve yapısal karmaşıklık gibi üst temalar ortaya çıkarılmıştır.
Temaların gözden geçirilmesi	Belirlenen temaların verilerle uyumu test edilir; temaların iç tutarlılığı ve farklı temalardan ayrışabilirliği kontrol edilir.	Belirlenen temalar araştırmacı paneli tarafından tartışılarak revize edilmiş, temaların kapsamı yeniden sınırlandırılmıştır.
Temaların adlandırılması ve tanımlanması	Her tema açık biçimde tanımlanır, isimlendirilir ve bulgularla ilişkilendirilir.	"Dilsel açıklık", "yanıt ölçeği sorunları", "hatırlama güçlüğü" gibi temalar net şekilde tanımlanarak bulgular kısmında doğrudan katılımcı alıntılar ile desteklenmiştir.
Raporlaştırma	Temalar araştırma amaçlarıyla ilişkilendirilir ve bulgular sistematik biçimde yazılır.	Sonuç raporunda temalar bulgular bölümünde yapılandırılmış, yinelemeli bilişsel görüşmelerin ölçek geliştirme sürecine katkısı tartışılmıştır.

3. Bulgular

Bilişsel görüşmeler, zindelik ölçeği madde havuzundaki bazı ifadelerin anlaşılabilirliği konusunda yorum farklılıklarına ve kavrama güçlüklerine işaret etmiştir. Özellikle bazı terimlerin ve ifadelerin fazla genel ve belirsiz olduğu tespit edilmiştir. Katılımcılar, "farkındalık" (mindfulness) ve "duygusal dayanıklılık" (emotional resilience) kavramlarını tanımlamakta ortak şekilde zorlandıklarını ifade etmiş ve bu durum, yanıtların değişkenlik göstermesine neden olmuştur. Bir katılımcı bu durumu şu şekilde ifade etmiştir: "Bu sorudaki 'duygusal

dayanıklılık' kavramı hakkında hiçbir fikrim yok. Çok genel bir ifade." Bir başka katılımcı ise, "Bu soruyu tam olarak anlayabilmek için iki kez okumak zorunda kaldım; çok uzun." şeklinde görüş belirtmiştir.

Terimlerin ötesinde, yapısal karmaşıklığın da ifadelerin okunabilirliğini olumsuz etkilediği gözlemlenmiştir. Birçok madde, fazla uzun ve dilbilgisel açıdan yoğun olarak değerlendirilmiş; bu durum da anlaşılmasını zorlaştırmıştır. Bir katılımcı bu durumu şöyle ifade etmiştir: "Cümle çok uzun; anlamını kavrayabilmek için durup düşünmem gerekiyor." Bir diğer katılımcı ise, "Soru daha kısa

olsaydı, zihnimde bu kadar parçalamama gerek kalmazdı." şeklinde yorum yapmıştır.

Gerekli düzeltmelerin yapılabilmesi amacıyla, maddelerin yeniden düzenlenmesinde cümle yapıları sadeleştirilmiş ve ifadelerin açıklığı artırılmıştır. Anlamayı zorlaştıran tekrar eden ve birleşik yapıli sorular, anlaşılabilirliği artırmak ve özgün amaçlarını korumak üzere yeniden ifade edilmiştir. Katılımcılar arasında anlam birliğini sağlamak amacıyla, gerekli görülen yerlerde bazı temel kavramlara yönelik tanımlayıcı açıklamalara da yer verilmiştir. Örneğin, "farkındalık" kavramı, ölçeğin revize edilmiş hâlinde kısa bir açıklayıcı notla desteklenmiştir. Tüm bu düzenlemeler, katılımcıların sorulara daha doğru yanıtlar verebilmelerini kolaylaştırmayı amaçlamaktadır.

Tutarlı Bilişsel İşlem ve Yanıt

Bilişsel görüşmeler, altı aşamalı ve sıralı bir format doğrultusunda yürütülmüştür: Sesli düşünme/genel, sözel sondalama/kavrama, sözel sondalama/hatırlama, sözel sondalama/değerlendirme, sözel sondalama/yanıtlama ve sözel sondalama/kendini ifade etme. Her bir aşamada katılımcı geri bildirimlerinin analizi, ölçek maddelerine yanıt verme sürecindeki bilişsel zorluklara ilişkin yeni bilgiler sunmuştur.

Sesli Düşünme/Genel: Katılımcılar ölçek maddelerini okurken izledikleri bilişsel süreçleri sözlü olarak ifade etmişlerdir. Bazı katılımcılar özellikle bazı teknik terimlerin başlangıçta kafa karıştırıcı olduğunu belirtmiş, bazıları ise belirli ifadelerde anlam karmaşası yaşadıklarını ifade etmiştir. Bir katılımcı, "*Sesli okurken bazı terimlerin tam anlamını kavrayıp kavrayamadığımdan emin olamadım. Sürekli durup anlamını düşündüm ve ardından devam ettim.*" şeklinde yorum yapmıştır. Bir diğer katılımcı ise, "*Soruyu okumadan önce anlamını tahmin etmeye çalışıyordum; sesli okurken bazen maddeleri başlangıçta yanlış yorumladığım oldu.*" ifadesini kullanmıştır.

Katılımcılar uzun ve çok bölümlü cümlelerde zorlandıklarını belirtmişlerdir. Bir katılımcı "*İki kez okudum ama anlayamadım,*" diyerek, "*çünkü cümle çok uzundu ve fazla parçalıydı.*" ifadesini kullanmıştır. Bir diğer katılımcı ise, "*Bazı kelimeler beni şaşırttı ve aslında soruya yanıt verip vermediğimden bile emin olamadım.*" şeklinde yorum yapmıştır.

Sözel Sondalama/Kavrama: Kavrama sürecinde sözdizimsel karmaşıklık ve kavramsal ifadelerin yeterince açık olmaması temel sorunlar olarak belirlenmiştir. Katılımcılar, sıklıkla açık olmayan ve anlamı kişiden kişiye değişebilen belirli terim ve ifadelere maruz kaldıklarını belirtmişlerdir. Bir katılımcı, "*Bazı sorularda yorumlaması kolay olmayan terimler var. Ben kendi kişisel deneyimimle anlamını çıkardım ama başka biri farklı şekilde yorumlayabilir.*" demiştir. Bir başka katılımcı ise, "*Öz-farkındalık terimi oldukça genel bir ifadeye sahip. Duygularımın genel olarak mı farkındayım, yoksa belli bir durumdaki farkındalıktan mı bahsediliyor?*" şeklinde açıklama yapmıştır.

Bazı katılımcılar, tanımların ya da örneklerin daha derinlemesine bir anlamaya katkı sağlayabileceğini ifade etmiştir. Bir katılımcı, "*Eğer bir örnek olsaydı ya da daha kısa bir ifade, anlamakta zorlanmazdım. Hemen anlam kazandırır.*" diye belirtmiştir. Bir diğer katılımcı ise, "*Sürekli olarak, 'Benim yorumum araştırmacının zihnindeki yorumla örtüşüyor mu?' diye düşündüm.*" diyerek, "*çünkü bazı ifadeler yeterince özgül değildi,*" şeklinde eklemiştir.

Sözel Sondalama/Hatırlama: Katılımcılar, bir davranışın ne zaman gerçekleştiğine dair sorulara yanıt verirken yaşadıkları deneyimleri hatırlamakta zorlandıklarını belirtmişlerdir. Bir katılımcı, "*Son altı ay içinde bir davranışı kaç kez yaptığımı tam olarak hatırlamakta zorlanıyorum.*" diyerek, "*bu davranışları bilinçli olarak takip etmiyorum.*" ifadesini kullanmıştır. Aynı katılımcı, "*Kesin bir sayı hatırlamadığım için tahmine dayalı, yaklaşık bir cevap verdim. Ancak bir davranışı belli bir düzende yapmıyorsam, kaç kez uyguladığımı saymam.*" şeklinde devam etmiştir.

Bazı katılımcılar sorunun sunulma biçiminin deneyimlerini hatırlama şeklini etkilediğini ifade etmişlerdir. Bir katılımcı, "*Altı ay ifadesini görünce panikledim. Zihnimde bu şekilde bir zaman ölçeği yok. 'Geçen hafta' ya da 'geçen ay' şeklinde olsaydı, çok daha güvenle cevap verirdim.*" demiştir. Bir diğer katılımcı ise, "*Bireysel olayları mı saymam gerekiyor yoksa rutinlerime dair genel bir izlenim mi sunmalıyım, emin olamadım.*" şeklinde görüş belirtmiştir.

Sözel Sondalama/Güven Değerlendirmesi: Katılımcılar verdikleri yanıtlara ne düzeyde güvendiklerini derecelendirmişlerdir. Çoğu katılımcı, yanıtlarını seçerken bir belirsizlik sergilemiştir. Örneğin, bir katılımcı, "*Cevabımın doğru olup olmadığını bilmiyorum; mantıklı bir tahminde bulundum.*" demiştir. Bir diğer katılımcı ise, "*Cevabımdan tamamen emin değildim ama elimden gelen en uygun seçeneği işaretledim.*" şeklinde ifade etmiştir.

Bazı katılımcılar verdikleri yanıtı sorguladıklarını belirtmişlerdir. Bir katılımcı, "*İki seçenek arasında gidip geldim çünkü ikisinden de emin değildim.*" demiştir. Bir diğer katılımcı ise, "*Yanıt verirken tereddüt ettim çünkü cevabımı ne kadar spesifik vermem gerektiğinden emin değildim.*" şeklinde açıklamış ve "*Kendime şu soruyu sordum: 'Yanıtım net mi olmalı, yoksa yaklaşık bir değerlendirme yapabilir miyim?'*" diye eklemiştir.

Sözel Sondalama/Yanıt Seçimi: Bazı yanıt seçenekleri birbirine fazla benzer bulunmuş, bu nedenle katılımcılar arasında seçenekler arasında ayırım yapmakta güçlük yaşanmıştır. Bir katılımcı, "*Yanıt seçenekleri birbirine çok benziyordu ve iki seçenek arasında kararsız kaldım. 'Bazen' ile 'ara sıra' ifadelerinin belirli bir anlam farkı olup olmadığını gerçekten birbirinden farklı olup olmadıklarını anlayamadım.*" şeklinde belirtmiştir. Bir diğer katılımcı ise, "*'Sıklıkla' ve 'bazen' seçeneklerini ayırt edemedim ve bu davranışı sürekli yapan biri olmadığım için karar veremedim.*" demiştir.

Bazı katılımcılar ise yanıt aralıklarının orantısız olduğunu ifade etmişlerdir. Bir katılımcı, "*Olumlu ifadeler için çok fazla*

seçenek olduğunu ama olumsuzlar için yeterince seçenek olmadığını fark ettim." demiştir. Bir başka katılımcı, "Daha genel bir seçeneği işaretlemeye itildim." şeklinde ifade etmiştir. Bir diğer katılımcı "Benim gerçek cevabım iki seçenek arasında kaldı; bana uygun bir orta seçenek yoktu." diyerek yanıt aralığının sınırlılığını vurgulamıştır.

Sözel Sondalama/Açıklama: Katılımcılardan soruları kendi ifadeleriyle yeniden oluşturmaları istendiğinde, bazen araştırmacının amaçladığı anlamla örtüşmeyen alternatif anlamlar ürettikleri gözlemlenmiştir. Bu durum, ifadelerin daha az karmaşık bir şekilde yeniden düzenlenmesinin gerekliliğini ortaya koymuştur. Bir katılımcı, "Ben bu soruyu şöyle yeniden ifade ederdim: 'Hayatında duygularına bir haftada ne kadar dikkat edersin?' Çünkü ilk duyduğumda soru bana biraz karmaşık geldi." şeklinde önermede bulunmuştur. Bir diğer katılımcı, "Benim bir soruyu anlama şeklim başka bir bireyin anladığıyla örtüşmeyebilir." demiştir. "Günlük bir konuşmada olsaydı farklı cevap verirdim." şeklinde eklemiştir.

Bazı katılımcılar ise maddeleri zihinsel olarak çevirmekte zorlanmışlardır. Bir katılımcı, "Yaptığım çeviri denk midir bilmiyorum, ama bu soruya ilişkin aklıma gelen ilk şey buydu." demiştir. Bir başka katılımcı ise, "Kafamda çevirdim ama tam olarak ne anlama geldiğinden emin değilim." diyerek belirsizliğini ifade etmiştir.

Düzeltilmeler

Yinelenen bilişsel görüşme süreci maddelerin daha açık, anlaşılır ve yanıt tutarlılığı yüksek bir biçimde şekillenmesini sağlamıştır. Bu aşamada yapılan revizyonlar, maddelerin kullanım kolaylığını artırmak, yanıt yanlılığını azaltmak ve katılımcıların öz bildirimlerinin doğruluğunu güçlendirmek amacıyla gerçekleştirilmiştir. Özellikle karmaşık terimlerin sadeleştirilmesi sürecin en belirgin değişiklikleri arasında yer almıştır. "Dayanıklılık (resilience)" ve "farkındalık (mindfulness)" gibi terimlere, maddelerde daha anlaşılır tanımlar eklenerek yorum birliği sağlanmıştır. Ayrıca uzun ve karmaşık cümle yapıları, katılımcıların bilişsel yükünü azaltmak amacıyla kısa ve doğrudan ifadelerle dönüştürülmüştür. Geriye dönük anıları hatırlamayı kolaylaştırmak adına, "altı ay" gibi genel zaman referanslarına daha kısa ve somut zaman aralıkları (örneğin "geçen ay") eklenmiştir.

Bu değişikliklerin somut bir örneği aşağıda sunulmuştur:

İlk versiyon: "Son altı ay içinde stres yönetimi uygulamalarını ne sıklıkla gerçekleştirdiniz?"

Son versiyon: "Geçtiğimiz ay içerisinde rahatlatma amacıyla egzersiz, derin nefes alma ya da meditasyon gibi teknikleri kaç kez uyguladınız?" Bu yeniden yazılmış soru, katılımcıların hatırlama gücü ve detay eksikliği gibi geri bildirimlerine yanıt vermek üzere düzenlenmiştir. Soru yapısının netleştirilmesi ve katılımcıların ilişki kurabileceği örneklerin eklenmesiyle, yanıtların doğruluğu ve netliği artırılmıştır.

Buna ek olarak, yanıt ölçekleri daha dengeli ve okunabilir olacak şekilde değiştirilmiştir. Başlangıçta kullanılan "Hiçbir zaman, Bazen, Sıklıkla, Her zaman" ölçeği; "Hiçbir zaman, Nadiren, Ara sıra, Sık sık, Her zaman" şeklinde güncellenmiştir. Bu tür iyileştirmeler, maddelerin okunmasını kolaylaştırmış ve katılımcıların davranışlarını daha güvenle ve doğru şekilde bildirmelerini sağlamıştır. Ayrıca, "Neredeyse her zaman" ve "Neredeyse hiçbir zaman" gibi seçeneklerde yer alan "neredeyse" ifadesi, uzman önerileri doğrultusunda ölçekten çıkarılmıştır. Bu düzenlemelerle, ölçeğin son hâli yüksek düzeyde güvenilirlik ve geçerlilik sağlamış; öz-bildirim yoluyla yapılan zindelik değerlendirmelerinde yöntemsel bütünlük korunmuştur.

4. Tartışma

Mevcut çalışmanın bulguları çok boyutlu bir zindelik ölçeği geliştirme sürecinde dilin sade kullanımı ve kültürel uygunluğun önemini kanıtlamaktadır. Yinelemeli bilişsel görüşme süreci, temel terimlerin kullanımı, yapısal karmaşıklık ve yanıt doğruluğu gibi konuları belirlemiş ve bu temel konuların katılımcıların maddeyi anlama ve doğru yanıt verme süreçlerine önemli ölçüde etki ettiğini ortaya koymuştur. Bulgular öz-bildirim araçlarında açık ve sade bir dil kullanımının önemini vurgulayan önceki çalışmaların sonuçlarıyla örtüşmektedir (Beatty ve Willis, 2007; Collins, 2003; Willis, 2015).

Çalışmanın özgün katkılarından biri madde havuzu geliştirme ve ölçme geçerliliğini artırmak amacıyla sistematik bir araç olarak yinelemeli bilişsel görüşme sürecinin kullanılmasıdır. Geleneksel anket geçerlilik yaklaşımlarının çoğu nicel psikometrik testlere odaklanırken, bu çalışmada uygulanan yöntem, katılımcı geri bildirimini çoklu aşamalarda doğrudan dikkate alarak açık ifadelerin yinelemeli biçimde geliştirilmesine olanak sağlamıştır. Bulgular, belirsiz terimlerin ve karmaşık dil kullanımının katılımcı üzerinde bilişsel yük oluşturduğunu ve bu durumun yanıt doğruluğu olumsuz etkilediğini ortaya koymuştur (Desimone ve Le Floch, 2004). Psikometrik ölçme araçlarında açık ve doğrudan bir dil kullanımının gerekliliği, özellikle öz-bildirime dayalı sağlık davranışlarının ölçümünde yanlılığı azaltma bağlamında sıklıkla vurgulanmaktadır (Miller ve Foster, 2010; Tourangeau, 2000). Bu çalışma, yinelemeli bilişsel görüşmelerin maddelerin anlaşılabilirliğini ve de verilen yanıtların kalitesini nasıl artırdığını ortaya koyarak bu alana ampirik bir katkı sunmuştur.

Mevcut ölçme araçları katılımcı yorumlarını sınırlayan sabit ve durağan dil kullanımları nedeniyle eleştirilmiştir (Bolaños-Medina ve González-Ruiz, 2012; Eadie vd., 2006; Uher, 2023). Bu çalışmanın bulguları, söz konusu eleştirileri genişleterek; anlaşılabilirlik, ifade tarzı ve yanıt seçeneklerinin sistematik olarak değerlendirilmesi ve iyileştirilmesi gerektiğini ortaya koymuştur. "Neredeyse" gibi nicelik belirten sözcüklerin kullanımı çoğu durumda katılımcılar arasında kafa karışıklığına ve yanıt çeşitliliğine

yol açmış, bu durum ise test geliştirme çalışmalarını zayıflatmıştır (Rubin, 2004). Sudman vd. (1996) ifade ettiği şekilde anket yöntemleri üzerine yaptıkları çalışmalarda, yanıt seçeneklerinde nicel ifadelerin kullanılmasının Likert tipi ölçeklerin güvenilirliğini nasıl zedelediğini göstermiştir. Mevcut çalışmanın sonuçları zindelik ölçümlerinde bu düşünceleri doğrulamıştır.

Ayrıca çalışmada kullanılan yinelemeli yaklaşım, ölçek geliştirme sürecine yönetsel bir yenilik getirmektedir. Geleneksel psikometrik testlerde yalnızca nicel geçerlik ölçütleri kullanılırken (Krosnick, 1991), bu çalışmada nitel bir geri bildirim döngüsü, yinelemeli bir model çerçevesinde ölçeğin gözden geçirilmesine entegre edilmiştir. Bu tür bir yinelemeli modelin kullanılması, ölçeğin son hâlinin yalnızca istatistiksel olarak sağlam değil; aynı zamanda bağlamsal ve anlam açısından da tutarlı olmasını sağlamaktadır.

Anket metodolojisinde uzun süredir kullanılan bir teknik olan bilişsel görüşmeler, katılımcıların çalışmayı ve maddeleri araştırmacılar tarafından amaçlandığı şekilde anlayıp anlamadıklarını belirlemek amacıyla kullanılmaktadır (Willis, 2015). Bu yöntemde katılımcılar, “sesli - düşünme” protokolleri ve sözlü yoklama tekniklerini kullanarak anket sorularına yanıt verirken aynı zamanda bilişsel süreçlerini sözlü olarak ifade ederler (Beatty ve Willis, 2007). Bilişsel görüşmeler; madde formülasyonunun niteliği, açıklığı, yanıt seçeneklerinin uygunluğu ve bilişsel yük gibi çeşitli yönlerin belirlenmesine imkân sağlar (Collins, 2003). Ancak bilişsel görüşme tekniğinin önemli bir sınırlılığı, genellikle yalnızca tek seferliğine uygulanmasıdır. Ortaya çıkan durum madde düzeltmelerinin tek seferlik geri bildirim dayalı olarak gerçekleştirilmesi ve yinelemeli düzenlemeler için sistematik bir mekanizmanın bulunmaması anlamına gelmektedir (Desimone ve Le Floch, 2004).

Buna karşılık yinelemeli bilişsel görüşme çok aşamalı ve sürekli döngüsel bir yaklaşımı benimseyerek, ölçek maddelerinin birden fazla katılımcı grubu aracılığıyla yeniden test edilmesine ve düzenlenmesine olanak tanır (Willis, 2015). Bu yaklaşım, her bir ardıl ölçek versiyonunun bir öncekinden daha iyi olması gerektiği ilkesine dayanır ve süreç boyunca ortaya çıkan anlama sorunları, yanıt tutarsızlıkları ve istenmeyen yanlılıkları sistematik biçimde çözmeyi hedefler (Beatty ve Willis, 2007). Gözden geçirilmiş anket sorularının yeni katılımcılarla yeniden uygulanması sayesinde, yinelemeli bilişsel görüşmeler, klasik bilişsel görüşmelerin sağlayamadığı kapsam geçerliliği ve yanıt doğruluğu açısından önemli katkılar sunar.

Bir diğer temel bulgu ise yanıt doğruluğu etkileyen hatırlama güçlüğüdür. Katılımcıların geçmişteki davranışlarını uzun zaman dilimleri için (örneğin, 6 ay) hatırlamaları zor olmuştur. Görülen hatırlama güçlüğü, tutarsız öz-bildirimlere yol açmıştır. Hatırlama aralığının kısaltılması (örneğin, altı aydan bir aya) katılımcıların

güven düzeyini ve yanıtlarını açık ve net belirtmelerini anlamlı ölçüde artırmıştır. Bu durum, geri getirme kuramı ile de örtüşmektedir (Tourangeau, 1984). Bu bulgu, uzun hatırlama aralıklarının fazla ya da eksik bildirimlere neden olduğunu ileri süren alanyazındaki araştırmaları da desteklemektedir (Bradburn vd., 1987).

Çalışma ayrıca katılımcıların yanıt verme süreçlerini de incelemiştir. Bazı katılımcılar, birbirine çok yakın anlamlar taşıyan yanıt seçeneklerini (örneğin, bazen ile sık sık) ayırt etmekte zorlanmışlardır. Bu bulgu, yanıt ölçeklerinin net ve iyi kalibre edilmiş olmasının önemini ortaya koymaktadır. Çünkü belirsiz nicelik belirleyicilerin, istenmeyen yanıt yanlılıklarına neden olduğu da bilinmektedir (DeVellis ve Thorpe, 2021). Bu çalışmada, yanıt seçeneklerinin daha belirgin ve eşit aralıklı alternatiflerle yeniden yapılandırılması, yanıtların güvenilirliğini ve ölçeğin kullanılabilirliğini artırmıştır.

Bir diğer önemli bulgu, “neredeyse” gibi belirsiz sıklık belirleyicilerin (örneğin, neredeyse her zaman, neredeyse hiç) yanıt seçeneklerinde kullanılmasıdır. Bu tür belirleyiciler, sıklık düzeylerini aşamalı biçimde ayırt etmeyi amaçlasa da katılımcılar için net bir referans noktası sunmamış ve belirsizlik yaratmıştır. Bu belirsizlik, katılımcıların öznel yorumlar yapmasına ve yanıt tutarsızlıklarına yol açmıştır. Katılımcılar sıklıkla, “neredeyse her zaman” ifadesinin %90 oranında mı, yoksa güçlü ancak tutarsız bir eğilim mi ifade ettiğini sorgulamışlardır. Benzer şekilde, “neredeyse hiç” ifadesi bazı katılımcılar tarafından çok nadir durumlar olarak yorumlanırken, bazıları bunu “hiç”e çok yakın ama mutlak olmayan bir durum olarak algılamıştır.

Mevcut çalışmada, “neredeyse” gibi belirsiz belirleyiciler kaldırılarak ve yanıt seçenekleri daha belirgin ve kesin hâle getirilerek yanıt doğruluğu, güvenilirliği ve yorumlanabilirliği artırılmıştır. Bu yaklaşım ise, ölçeklerde yanıt seçeneklerini oluştururken öznel belirsizlik yaratmaksızın açık ayrımlar sağlaması gerektiğini vurgulayan ölçek geliştirme alanındaki en iyi çalışmalarla örtüşmektedir (Rubin, 2004). Katılımcıların kendi deneyimlerini doğru şekilde ölçekteki seçeneklerle eşleştirebilmeleri, ölçme hatasını en aza indirmek ve veri geçerliliğini artırmak açısından kritik öneme sahiptir.

Çalışmanın bazı sınırlılıkları bulunmaktadır. İlk olarak, örneklem büyüklüğü ($n = 10$) nispeten küçüktür ve elde edilen bulguların genellenebilirliğini sınırlandırabilir. Nitekim literatürde bilişsel görüşmeler için genellikle 5–15 katılımcının yeterli olabileceği ifade edilse de (Beatty ve Willis, 2007), ölçeğin farklı yaş gruplarında, mesleklerde ve kültürel bağlamlarda daha geniş ve çeşitlendirilmiş örneklemlemlerle test edilmesi, gerek durumlarda ise hedeflenen gruba yönelik uygulanması elde edilen bulguların güvenilirliğini ve geçerliliğini arttıracaktır.

İkinci olarak, yinelemeli bilişsel görüşmeler ölçek maddelerinin dilsel ve kavramsal uygunluğunu geliştirmede etkili olsa da, ölçeğin bilimsel geçerliliğini ve

güvenirliğinin sağlanabilmesi için tek başına yeterli değildir. Ölçek geliştirme sürecinde bilişsel görüşmelerin, sistematik psikometrik analizlerle (örneğin, madde tepki kuramı, faktör analizi) mutlaka desteklenmesi gerekmektedir. Bu tür analizlerin entegrasyonu ile ölçeğin ölçme hassasiyeti ve yapı geçerliliği daha da artırılabilir. Aynı zamanda farklı örneklemeler üzerinde tutarlı ve genellenebilir sonuçların elde edilmesine de katkı sağlar. Dolayısıyla, yinelemeli bilişsel görüşmeler ölçek geliştirme çalışmalarında başlangıç için kritik bir rol üstlenirse de, geçerliliği ve güvenilirliği yüksek bir ölçüm aracı oluşturmak için psikometrik yöntemlerle entegrasyonu zorunludur.

Mevcut çalışma zindelik ölçüm araçlarının iyileştirilmesinde yinelemeli bilişsel görüşmelerin etkililiğini ortaya koymaktadır. Anlama, hatırlama ve yanıt yanlılıkları sistematik bir şekilde ele alınarak; dilsel açıdan daha erişilebilir, kavramsal olarak daha sade ve yöntemsel açıdan daha sağlam bir ölçek geliştirilmiştir. Bu bağlamda kullanılan yinelemeli yaklaşım, katılımcı geribildirimine dayalı sürekli iyileştirme sürecini merkeze alan ve gelecek ölçek geliştirme çalışmalarına örnek teşkil edebilecek nitelikte tekrarlanabilir bir model sunmaktadır.

İleride yapılacak araştırmalarda revize edilen ölçek farklı katılımcılara uygulanmalı, boylamsal tutarlılığı incelemeli ve nicel doğrulama teknikleriyle (örneğin, madde tepki kuramı, faktör analizi) desteklenmelidir.

Sonuç olarak, mevcut çalışma yinelemeli bilişsel görüşmelerin ölçek geliştirme sürecinde güçlü bir metodolojik araç olduğunu ortaya koymaktadır. Ancak daha geniş örneklemeler, psikometrik doğrulama teknikleri ve kültürler arası çalışmalarla desteklendiğinde, geliştirilen ölçeğin bilimsel açıdan daha sağlam, genellenebilir ve güvenilir bir ölçüm aracı haline geleceği öngörülmektedir.

Etik Bildirim: Yapılan bu çalışmada "Yükseköğretim Kurumları Bilimsel Araştırma ve Yayın Etiği Yönergesi" kapsamında uyulması belirtilen tüm kurallara uyulmuş; yönergenin ikinci bölümü olan "Bilimsel Araştırma ve Yayın Etiğine Aykırı Eylemler" başlığı altında belirtilen eylemlerden hiçbiri gerçekleştirilmemiştir. Bu çalışma, Orta Doğu Teknik Üniversitesi İnsan Araştırmaları Etik Kurulu'nun 06/05/2024 tarihli ve 0266-ODTÜİAEK-2024 sayılı izni doğrultusunda gerçekleştirilmiştir.

Yazar Katkıları : Giriş: Birinci yazar, İkinci yazar. Yöntem: Birinci Yazar, Üçüncü yazar, Bulgular: Birinci, İkinci yazar, Üçüncü yazar, Dördüncü yazar, Tartışma: Birinci yazar, İkinci yazar, Üçüncü yazar, Dördüncü yazar.

Finansman : Çalışma için herhangi bir finansal destek alınmamıştır.

Çıkar Çatışması : Çalışmada herhangi bir çıkar çatışması bulunmamaktadır.

Veri Erişilebilirliği : Çalışmada kullanılan veriler isteğe bağlı olarak araştırmacılardan talep edilerek erişilebilir.

Kaynakça

- Beatty, P. C., & Willis, G. B. (2007). Research synthesis: The practice of cognitive interviewing. *Public Opinion Quarterly*, 71(2), 287–311. <https://doi.org/10.1093/poq/nfm006>
- Bolaños-Medina, A., & González-Ruiz, V. (2012). Deconstructing the translation of psychological tests. *Meta: Journal Des Traducteurs/Meta: Translators' Journal*, 57(3), 715–739. <https://doi.org/10.7202/1017077ar>
- Bradburn, N. M., Rips, L. J., & Shevell, S. K. (1987). Answering autobiographical questions: The impact of memory and inference on surveys. *Science*, 236(4798), 157–161. <https://doi.org/10.1126/science.236.4798.157>
- Braun, V., & Clarke, V., (2006). Using thematic analysis in psychology *Qualitative research in psychology* 3(2), 77-101.
- Braun, V., Clarke, V., & Weate, P. (2016). Using thematic analysis in sport and exercise research. In B. Smith ve A. C. Sparkes (Eds) *Routledge handbook of qualitative research in sport and exercise* (pp. 213–227). Routledge.
- Collins, D. (2003). Pretesting survey instruments: An overview of cognitive methods. *Quality of Life Research*, 12, 229–238. <https://doi.org/10.1023/A:1023254226592>
- Desimone, L. M., & Le Floch, K. C. (2004). Are we asking the right questions? Using cognitive interviews to improve surveys in education research. *Educational Evaluation and Policy Analysis*, 26(1), 1–22. <https://doi.org/10.3102/01623737026001001>
- DeVellis, R. F., & Thorpe, C. T. (2021). *Scale development: Theory and applications*. Sage publications.
- Dunn, H. L. (1959). What high-level wellness means. *Canadian Journal of Public Health/Revue Canadienne de Sante'e Publique*, 50(11), 447–457.
- Eadie, T. L., Yorkston, K. M., Klasner, E. R., Dudgeon, B. J., Deitz, J. C., Baylor, C. R., Miller, R. M., & Amtmann, D. (2006). Measuring communicative participation: A review of self-report instruments in speech-language pathology. *Journal of Speech, Language and Hearing Research*, 49(5), 904-920 [https://doi.org/10.1044/1058-0360\(2006\)030](https://doi.org/10.1044/1058-0360(2006)030)
- Haldane, V., Chuah, F. L., Srivastava, A., Singh, S. R., Koh, G. C., Seng, C. K., & Legido-Quigley, H. (2019). Community participation in health services development, implementation, and evaluation: A systematic review of empowerment, health, community, and process outcomes. *PloS One*, 14(5), e0216112. <https://doi.org/10.1371/journal.pone.0216112>
- Hettler, B. (1984). Wellness: Encouraging a lifetime pursuit of excellence. *Health Values*, 8(4), 13–17.
- Kılıç, K., ve İnce, M. L. (2016). Sporda pozitif gençlik gelişimi ölçeklerinin 8-14 yaş grubu çocuk ve genç sporcular tarafından anlaşılabilirliğinin incelenmesi. *Beden Eğitimi ve Spor Bilimleri Dergisi*, 10(2), 213–225.
- Krosnick, J. A. (1991). Response strategies for coping with the cognitive demands of attitude measures in surveys. *Applied Cognitive Psychology*, 5(3), 213–236. <https://doi.org/10.1002/acp.2350050305>
- Kuru, H. (2022). *Designing and Evaluating a Need-Based Employee Wellness Program*. [Doctoral dissertation, Middle East Technical University]
- Kuru, H. (2024). Sağlık, zindelik ve iyi oluş kavramlarının incelenmesi. *İstanbul Rumeli Üniversitesi Sağlık Bilimleri Dergisi*, 2(4), 1–11.
- Miller, G., & Foster, L. T. (2010). *Critical synthesis of wellness literature*. Wellness Research Institute.
- Myers, J. E., & Sweeney, T. J. (2004). The indivisible self: An evidence-based model of wellness. *Journal of Individual Psychology*, 60, 234–244.
- Myers, J. E., Sweeney, T. J., & Witmer, J. M. (2000). The wheel of wellness counseling for wellness: A holistic model for treatment planning. *Journal of Counseling ve Development*, 78(3), 251–266. <https://doi.org/10.1002/j.1556-6676.2000.tb01906.x>
- Myers, N. D., Prilleltensky, I., Prilleltensky, O., McMahon, A., Dietz, S., & Rubenstein, C. L. (2017). Efficacy of the fun for wellness online intervention to promote multidimensional well-being: A randomized controlled trial. *Prevention Science*, 18, 984–994. <https://doi.org/10.1007/s11121-017-0779-z>
- National Wellness Institute. (2020). *Six Dimensions of Wellness*. National Wellness Institute. <https://nationalwellness.org/resources/six-dimensions-of-wellness/>
- Roscoe, L. J. (2009). Wellness: A review of theory and measurement for counselors. *Journal of Counseling and Development: JCD*, 87(2), 216. <https://doi.org/10.1002/j.1556-6678.2009.tb00570.x>
- Rubin, D. B. (2004). *Multiple imputation for nonresponse in surveys* (Vol. 81). John Wiley & Sons.
- Sudman, S., Bradburn, N. M., & Schwarz, N. (1996). *Thinking about answers: The application of cognitive processes to survey methodology*. Jossey-Bass.
- Tourangeau, R. (1984). Cognitive sciences and survey methods. *Cognitive Aspects of Survey Methodology: Building a Bridge between Disciplines*, 15, 73–100.
- Tourangeau, R. (2000). *The psychology of survey response*. University of Cambridge.
- Uher, J. (2023). What's wrong with rating scales? Psychology's replication and confidence crisis cannot be solved without transparency in data generation. *Social and Personality Psychology Compass*, 17(5), e12740. <https://doi.org/10.1111/spc3.12740>
- Willis, G. B. (2015). *Analysis of the cognitive interview in questionnaire design*. Oxford University Press, USA.
- World Health Organization. (2018). *Health in All Policies as part of the primary health care agenda on multisectoral action*. World Health Organization.
- World Health Organization. (2021). *Towards developing WHO's agenda on well-being*. World Health Organization



Testing and Development of the Wellness Scale Item Pool Through Iterative Cognitive Interviews

Hakan Kuru¹ Aykut Bulut² Koray Kılıç³ Ece İbanoğlu⁴

ARTICLE INFO

DOI: 10.29299/kefad.1697627

Received: 13.05.2025

Revised: 10.10.2025

Accepted: 03.11.2025

Keywords:

Wellness Measurement,
Construct Validity,
Scale Development,
Comprehensibility,
Measurement Error

ABSTRACT

The necessity of measuring wellness as a multidimensional construct is grounded in solid conceptual foundations and requires the development of measurement tools that are practically comprehensible. However, it is critically essential for construct validity that the scale items are perceived by respondents as intended by the researchers. Moreover, testing item validity solely based on statistical analyses may bring limitations that could lead to measurement errors. This study aims to test and improve the comprehensibility of the wellness scale item pool through iterative cognitive interviews. Accordingly, to enhance item clarity and appropriateness, cognitive interviews were conducted with field and language experts (n=2). Subsequently, based on the iteratively revised item pool, cognitive interviews were carried out with purposively selected participants (n=8). To evaluate participants' interpretations of the items, think-aloud and verbal probing techniques were used. The findings revealed several problem areas including the conceptual complexity of terms, item length, inconsistencies in response categories, and recall difficulties. To address the identified issues, after the initial phase, the language was simplified, lengthy items were restructured, and response options were adjusted to facilitate participants' responses. The iterative refinement process led to notable improvements in participants' understanding of the items and the accuracy of their responses. The findings indicate the importance of using an iterative item-pool approach in cognitive interviews when testing scale validity. Future research in different fields of scale development may consider applying an iterative approach in cognitive interviews to enhance measurement validity. Long-term experimental studies are needed better to understand the effectiveness of the applied iterative method.

1. Introduction

Wellness is a state of well-being encompassing an individual's physical, mental, emotional, social, occupational, and spiritual aspects (National Wellness Institute [NWI], 2020). The NWI describes wellness as a conscious, self-directed, and evolving process aimed at realizing an individual's full potential. This definition involves maintaining health and actively engaging in personal growth across all areas of life, thereby forming the foundation for a satisfying and meaningful existence. Wellness is regarded as a health model that holistically enhances individuals' quality of life and provides a balancing framework for their lives (Dunn, 1959). This

approach goes beyond viewing health merely as disease prevention, reflecting instead a new understanding of health based on the pursuit of a healthy, fulfilling, and active life (Roscoe, 2009). This paradigmatic shift has paved the way for the adoption of wellness as a multidimensional model aimed at improving individuals' quality of life.

NWI's Six-Dimensional Model of Wellness offers a multidimensional framework to enhance an individual's quality of life sustainably. This model, developed by Hettler (1984), is oriented towards supporting emotional, physical, occupational, intellectual, social, and spiritual wellness. Each dimension contributes uniquely to an individual's quality of life, enabling them to live a healthy,

Corresponding Author¹: Hakan Kuru, Dr., İstanbul Rumeli University, Türkiye, hakan.kuru@rumeli.edu.tr

Author²: Aykut Bulut, Dr., Kırşehir Ahi Evran University, Türkiye, aykut.bulut@ahievran.edu.tr

Author³: Koray Kılıç, Dr., Kırşehir Ahi Evran University, Türkiye, koray.kilic@ahievran.edu.tr

Author⁴: Ece İbanoğlu, Phd Student, Utah State University, USA, ecee.ibanoglu@gmail.com

balanced lives (Myers & Sweeney, 2004; Roscoe, 2009). The goal of focusing on each dimension of wellness is to encourage holistic self-development. In this sense, the Six-Dimensional Model not only supports physical health but also contributes to the creation of a comprehensive state of well-being (see Figure 1).



Figure 1. Six-dimensional model of wellness (NWI, 2020)

Note. Adapted from the Six Dimensions of Wellness Model (NWI, 2020).

According to the six-dimensional model, emotional wellness refers to recognizing, accepting, managing emotions, and building healthy relationships with others. Further, NWI (2020) defines emotional wellness as the capacity to recognize, accept, and manage one's emotions and the behaviors associated with them. Within this dimension, individuals develop a positive outlook on themselves and their lives, become aware of their own emotions, and consider the emotions of others. Supporting autonomy in coping with stress, maintaining an optimistic perspective, and sustaining open, honest, and healthy communication in relationships are among fundamental elements of emotional wellness.

Physical wellness, on the other hand, refers to prioritizing physical self-care and consistently engaging in health-promoting physical behaviors. Regular physical activity, balanced nutrition, adequate sleep, rest, and access to medical support when needed are included within this dimension. In this dimension, individuals maintain their physical health by identifying potential problems and taking preventive actions.

Intellectual wellness represents awareness of sharing one's skills through stimulating activities that involve learning, personal development, and creativity. A commitment to lifelong learning and a desire for mental growth form the foundation of this dimension. Intellectual wellness involves acquiring new hobbies, expanding areas of interest, staying informed about current events, and continually broadening learning processes by discovering new ways of thinking.

Occupational wellness refers to the sense of satisfaction and enrichment an individual experiences through work, academic, or volunteer activities. By choosing jobs and activities that align with their values and interests, individuals enhance their self-confidence and gain a sense of accomplishment. Occupational wellness enables individuals to participate in meaningful and fulfilling activities that help them achieve their goals.

Spiritual wellness encompasses the process of perceiving and appreciating the deeper meaning of life and connecting with greater natural or universal forces. Within this dimension, individuals question the meaning of life and develop a sense of purpose aligned with their personal values. Spiritual wellness enables individuals to live in harmony with inner peace, shape their lifestyles in line with their values, and remain open to others' beliefs.

Social wellness refers to contributing to one's community and environment and understanding the reciprocal connection between oneself, others, and nature. Socially strong individuals make positive contributions to society, strengthen social relationships, and maintain a sense of responsibility toward preserving environmental resources. Social wellness encourages individuals to integrate into society, increase their social awareness, and take an active role in their environments (NWI, 2020).

The multidimensional measurement of wellness is essential for understanding health and quality of life (Kuru, 2022; Myers et al., 2017). Concurrently, wellness assessments help individuals recognize their strengths and areas for improvement, guiding them in setting health goals (Roscoe, 2009). Subjective wellness assessments enable individuals to identify daily priorities more effectively, thereby contributing to the creation of more practical roadmaps for achieving healthy living goals and developing health strategies (Myers et al., 2000). As Miller and Foster (2010) note, measuring wellness is fundamental to helping individuals understand their current health status and develop strategies to improve it. Moreover, understanding the scope of wellness is critically essential for developing strategies that enhance the effectiveness of public health policies (World Health Organization [WHO], 2018). Measuring individuals' social, emotional, and spiritual wellness can provide evidence for potential public health interventions aimed at improving community well-being. Findings from well-being assessments serve as a fundamental basis for the planning, implementation, and evaluation of public health interventions (Haldane et al., 2019).

The present study aims to test and improve the comprehensibility of the item pool developed for the Wellness Scale. Accordingly, an iterative cognitive interviewing process was employed to test and enhance the clarity of the item pool. Willis (2015) and Collins (2003) emphasize that the cognitive interview method is an effective tool for improving linguistic clarity and

conceptual alignment in scale and questionnaire development processes. Previous research has shown that cognitive interviews are particularly effective in enhancing the functionality of measurement tools for younger individuals who have difficulty understanding words and phrases (Kılıç & İnce, 2016). Therefore, the present study applies the cognitive interviewing method in an iterative manner to systematically enhance the development of the Wellness Scale and strengthen its cultural and conceptual alignment. Specifically, this study integrates an iterative dimension into the cognitive interview approach to better understand how participants interpret the items and the difficulties they encounter during the response process (Beatty & Willis, 2007a). Through this method, additional refinements were made to enhance measurement validity.

The unique methodological aspect of the current study is the use of the iterative cognitive interviewing method in the development process of the wellness scale item pool. The significance of the applied methodology is its contribution to improving the validity and reliability of scores obtained from scales designed to measure complex, multidimensional constructs such as wellness. Furthermore, consistent with the literature emphasizing the effectiveness of cognitive interviews in producing valid data (Rubin, 2004), this study presents a novel approach by applying the method across multiple stages within an iterative process to ensure data validity.

2. Method

2.1. Design

This study examines the comprehensibility of the question pool developed to measure wellness. A two-stage cognitive interview process with an iterative approach was conducted. In the first stage, a group of experts reviewed the comprehensibility of the questions, and the questions were revised based on the findings. In the second stage, the revised questions were reevaluated with a new participant group to finalize the questions. The primary objective of the iterative approach used in this study is to progressively improve the question form based on feedback at each step and enhance comprehensibility.

2.2. Study Group

The study included 10 participants. In the first stage, two participants (one language expert and one field expert) participated, while in the second stage, eight participants participated in cognitive interviews using the iterative item pool. The sample of eight participants was deemed sufficient to achieve data saturation in the cognitive interviews, as no new themes or problem areas emerged. Participants were selected using criterion sampling to ensure diversity in age, gender, education level, and occupation (Patton, 2002). The sample was chosen to represent individuals from diverse backgrounds, including language and field experts, students, and members of

various professional groups. All participants were 18 years of age or older (see Table 1).

Table 1.

Participants demographics

	Participant ID	Age	Sex	Education Level
Step 1	1	48	Woman	Masters
	2	39	Man	Doctorate
	3	37	Man	Masters
	4	19	Woman	H. School
Step 2	5	21	Man	H. School
	6	52	Woman	Bachelors
	7	32	Woman	Doctorate
	8	42	Man	Bachelors
	9	33	Woman	Bachelors
	10	44	Man	Masters

2.3. Wellness Scale Item Pool

The development process of the scale item pool followed a multi-stage method. In the first stage, a comprehensive item pool was created based on the Six-Dimensional Wellness Model. Drawing on existing scales and theoretical definitions in the literature, the pool included 8 items for emotional wellness, 7 for intellectual wellness, 7 for physical wellness, 8 for social wellness, 7 for spiritual wellness, and 8 for occupational wellness, for a total of 45 items.

In the second stage, the item pool was independently evaluated by three experts from the fields of behavioral design, sport sciences, and psychology. The experts examined the items based on criteria of linguistic clarity, conceptual relevance, and cultural appropriateness. As a result of these evaluations, 12 items were rewritten for linguistic reasons, 6 items were removed due to conceptual ambiguity, and minor revisions were made to the remaining items. For example, the statement "I am aware of my strengths and weaknesses" was criticized for covering multiple dimensions simultaneously and was revised as "I consider myself emotionally strong."

The interdisciplinary contributions of the experts ensured that the items were tested from multiple perspectives. The behavioral design perspective enhanced the behavioral validity of the items, the sport sciences perspective increased measurement precision in the physical domain, and the psychological perspective ensured content validity in the emotional and cognitive domains.

In the final stage, the experts' feedback was synthesized, revisions were completed, and a final form of 39 items was established. The process reinforced the scale's content validity, improved conceptual and linguistic consistency, and ensured cultural relevance.

2.4. Measures

The cognitive interviews were conducted individually, either face-to-face or via online platforms. Audio recordings were taken during the interviews, and each session was subsequently transcribed verbatim. Each interview lasted approximately 32 to 53 minutes. During the interviews, participants' difficulties in answering the questions, their assessments of question clarity, and their suggestions were carefully noted.

2.4.1. Cognitive interviews

Cognitive interviews involve collecting verbal data from participants. In this process, participants are asked to respond to survey questions. They are probed about their thought processes to assess whether the questions are understood as intended by the researcher or to examine the quality of responses (Beatty & Willis, 2007). As cognitive interviews aim to understand psychosocial phenomena, they are considered a psychological method or a type of qualitative interview (Collins, 2003). Traditional psychometric tests applied in surveys generally assess reliability and validity; however, they rely on assumptions that: a) participants correctly understand the questions, b) all participants interpret the questions in the same way, and c) participants have the willingness and ability to respond accurately (Collins, 2003).

Therefore, although psychometric tests can detect apparent issues that hinder the response process, they may fall short in explaining the underlying causes of these problems or providing evidence of latent issues. Cognitive interviews, however, help explain why individuals respond to survey questions in specific ways and identify overlooked constructs or mistaken assumptions related to the topic (Desimone & Le Floch, 2004).

Willis (2015) defines two main verbal reporting techniques used in cognitive interview designs as "think-aloud" and "verbal probing." In the think-aloud technique, participants express their mental processes aloud as they read and answer survey questions. During this process, the interviewer avoids interfering with the participant's thoughts and remains neutral to capture unbiased and uninfluenced thinking. The verbal probing technique, on the other hand, involves asking pre-prepared follow-up questions to understand how participants approach the survey questions. The think-aloud technique is participant-centered and can sometimes be challenging, whereas verbal probing is interviewer-centered and can make the process easier for participants.

In the present study, the reason for adopting an iterative (repetitive) approach focused on think-aloud and verbal-probing cognitive interviews is its potential to develop measurement tools systematically and, in a participant-centered manner. Unlike traditional validation methods that rely on a single cognitive interview session, the iterative design includes multiple data-collection and

refinement cycles, allowing ambiguities, misunderstandings, and undesirable question structures to be continuously reviewed and resolved. The iterative process used in this study involved revising the scale items based on participant feedback at each interview stage and retesting the updated scale with a new participant group. The repeated iterative process allowed researchers to refine the scale at every stage based on participant feedback. Rubin (2004) argues that conducting testing and revision iteratively helps capture linguistic and conceptual nuances more precisely in questionnaire and scale designs. Thus, it establishes a cyclical review mechanism that reshapes the process beyond a single interview, strengthening the validity and reliability of the scale with each step.

Step 1: Cognitive Interviews

In the first stage, cognitive interviews were conducted with two experts using the "think-aloud" and "verbal probing" techniques. During the interviews, participants were asked to read the questions aloud as they answered and to verbalize their thought processes. Additionally, supplemental questions were asked following verbal probing techniques to assess the clarity of the questions, whether they were sufficiently explicit, and whether participants encountered any difficulties while responding. Based on the data obtained at this stage, the questions were reviewed, and necessary revisions were made.

Step 2: Cognitive Interviews with the Revised Item Pool

Cognitive interviews were conducted with a second participant group (n=8). In this stage, the same think-aloud and verbal probing techniques used with experts were applied. However, this time the focus was on assessing whether the comprehensibility of the revised item pool had improved and evaluating the effects of the changes made based on previous feedback.

Cognitive Interview Procedure

Think-Aloud / General: The general cognitive processes participants expressed through the think-aloud technique while answering each question were examined. Difficulties encountered while reading the questions, ambiguities in understanding their purpose, and inferences made while responding were noted. Willis (2015) emphasizes that the think-aloud technique allows direct observation of participants' cognitive processes.

Verbal Probing / Comprehension: To evaluate whether participants correctly understood the structure and content of the questions, difficult-to-comprehend sections were coded thematically for each question. Collins (2003) stresses that comprehension is a fundamental component of cognitive processing and that understanding how participants interpret a question is critically important.

Verbal Probing / Recall: Participants' recall processes based on past experiences were analyzed. According to Tourangeau (1984), this stage is crucial for the reliability of

survey responses. Some participants had difficulty recalling specific information for certain questions, indicating insufficient clarity in those items. Especially questions involving time expressions (e.g., "How often in the past six months...?") posed challenges for participants and were restructured to facilitate recall.

Verbal Probing / Judgment: At this stage, how participants evaluated recalled information and whether their responses aligned with the intent of the question were examined. The focus was on understanding how participants assessed their responses and the extent to which those responses aligned with the question's purpose. Beatty and Willis (2007) note that the judgment stage significantly affects the accuracy of survey results.

Verbal Probing / Response: The appropriateness of participants' responses and the difficulties they encountered during the response process were analyzed.

This stage particularly assessed whether participants found the provided response options sufficiently inclusive. Desimone and Le Floch (2004) emphasize that transparent and inclusive response options directly influence the overall validity of a survey.

Verbal Probing / Self-Expression: Participants who misunderstood or had difficulty answering an item were asked to explain, in their own words, how they understood the question and how they responded. This step aimed to determine whether participants comprehended the question as intended by the researcher. Rubin (2004) states that this process reveals participants' potential to interpret questions differently. Especially for complex or lengthy questions, reconstructing them in participants' own words proved effective in making them more comprehensible (see Table 2).

Table 2.

Cognitive interview procedure

Cognitivite Interview	Purpose	Sample Question
Think-aloud	General	Can you describe your thoughts as you answered this question? How easy or difficult did you find this question? Why?
Verbal Probing	Comprehension	What does the term "X" mean to you? How did you interpret the purpose of this question?
Verbal Probing	Retreival	Did you recall a specific time while answering this question? Can you remember a particular example related to this question?
Verbal Probing	Judgement	How confident are you in your answer? How well do you think your response reflects the situation described in the question?
Verbal Probing	Response	Was it easy to find the option that best fits your initial answer? How did you feel while answering this question?
Verbal Probing	Self-Expression	How would you restate this question in your own words? Can you explain the meaning of this question differently?

Feedback from each stage of the cognitive interviews was discussed in a panel comprising three researchers: a behavioral designer, an educational scientist, a sport scientist, and a sport psychologist. This panel examined the issues participants experienced in detail and reached unanimous decisions on revisions to the questions. Willis (2015) notes that such panel discussions are an effective method in the question development process and that diverse perspectives make significant contributions to scale development. Each topic addressed in the panel prompted necessary adjustments to question wording or response options based on participant feedback.

2.5. Data Analysis

The data obtained from the cognitive interviews were analyzed in detail for each question using thematic analysis (see Table 3). Thematic analysis is a qualitative method used to identify, analyze, and interpret patterns within data and consists of six main phases. In the first phase, all interview recordings were transcribed, and the data were repeatedly reviewed to develop an in-depth understanding. In the second phase, using an inductive approach, the interview transcripts were coded, and initial codes were created. In the third phase, related codes were grouped to identify meaningful patterns in the data, and themes were defined. In the fourth phase, the alignment of themes with the data was tested through an iterative

process to ensure internal consistency and external distinction. In the fifth phase, themes were clearly defined and named. Finally, in the sixth phase, a narrative synthesis linking the themes to the research objectives was

developed. This process, while flexible enough to accommodate the complexity of the data, enabled systematic and meaningful analysis (Braun & Clarke, 2006; Braun et al., 2016).

Table 3.

Six stages of thematic analysis (Braun & Clarke, 2006)

Stage	Description	Implementation
Familiarization with the data	The data are transcribed, read repeatedly, and initial impressions are recorded. The researcher becomes deeply familiar with the data.	Interview recordings were transcribed, and each transcript was examined multiple times to extract initial notes on participants' comprehension difficulties, conceptual problems, and response inconsistencies.
Generating initial codes	Initial codes representing meaningful units from the data are generated. At this stage, coding is conducted as inclusively as possible.	Participant statements were coded in detail, and comments on ambiguous concepts such as "mindfulness" and "emotional resilience" were identified as separate codes.
Searching for themes	Patterns among the codes are sought, and similar codes are grouped together to form themes.	Similar participant comments were grouped, and overarching themes such as linguistic clarity, recall difficulty, ambiguity in response options, and structural complexity were identified.
Reviewing themes	The consistency of the identified themes with the data is tested; internal coherence and distinction from other themes are checked.	The identified themes were discussed by the researcher panel and revised; the scope of the themes was redefined.
Defining and naming themes	Each theme is clearly defined, named, and linked to the findings.	Themes such as "linguistic clarity," "response scale issues," and "recall difficulty" were clearly defined and supported with direct participant quotations in the findings section.
Reporting	Themes are related to the research objectives and the findings are written systematically.	In the final report, themes were structured in the findings section, and the contribution of iterative cognitive interviews to the scale development process was discussed.

3. Results

Cognitive interviews revealed differences in interpretation and comprehension difficulties regarding the clarity of certain statements in the wellness scale item pool. Some terms and expressions were found to be overly general and ambiguous. Participants reported that they commonly struggled to define the concepts of "awareness (mindfulness)" and "emotional resilience," which led to variability in responses. One participant expressed this as follows: *"I have no idea what the concept of 'emotional resilience' in this question means. It's a very general expression."* Another participant stated, *"I had to read this question twice to fully understand it; it's too long."*

Beyond terminology, structural complexity also adversely affected the readability of the statements. Many items were considered overly long and grammatically dense, which made them difficult to understand. One participant described this as follows: *"The sentence is too long; I have to pause and think to grasp its meaning."* Another participant commented, *"If the question were shorter, I wouldn't have to break it down so much in my mind."*

To make the necessary corrections, sentence structures were simplified, and the clarity of the expressions was improved during the revision of the items. Repetitive and

compound questions that made understanding difficult were rephrased to enhance comprehensibility while preserving their original intent. To ensure shared understanding among participants, brief explanatory notes were included for certain core concepts where deemed necessary. For example, the concept of "mindfulness" was supported with a short descriptive note in the revised version of the scale. All these adjustments were intended to make it easier for participants to provide more accurate answers to the questions.

Consistent Cognitive Processing and Response

Cognitive interviews were conducted in accordance with a six-stage, sequential format: think-aloud/general, verbal probing/comprehension, verbal probing/recall, verbal probing/judgment, verbal probing/response, and verbal probing/self-expression. At each stage, the analysis of participant feedback provided new insights into the cognitive challenges involved in responding to the scale items.

Think-Aloud/General: Participants verbally expressed the cognitive processes they followed while reading the scale items. Some participants noted that specific technical terms were particularly confusing at first, while others reported semantic confusion with certain expressions. One

participant commented, *"While reading aloud, I wasn't sure whether I fully grasped the exact meaning of some terms. I kept pausing to think about their meaning and then continued."* Another participant said, *"Before reading the question, I tried to guess its meaning; while reading aloud, I sometimes initially misinterpreted the items."*

Participants reported difficulty with long and multi-clause sentences. One participant said, *"I read it twice but couldn't understand,"* adding, *"because the sentence was too long and too fragmented."* Another participant commented, *"Some words threw me off, and I wasn't even sure whether I actually answered the question."*

Verbal Probing/Comprehension: Syntactic complexity and insufficient clarity in conceptual expressions were identified as the main issues during comprehension. Participants reported frequently encountering specific terms and expressions that were unclear and could vary in meaning from person to person. One participant said, *"Some questions include terms that aren't easy to interpret. I inferred their meaning from my personal experience, but someone else might interpret them differently."* Another participant explained, *"The term self-awareness is quite general. Is it asking whether I am generally aware of my emotions, or is it referring to awareness in a specific situation?"*

Some participants stated that definitions or examples could deepen understanding. One participant said, *"If there were an example or a shorter expression, I wouldn't have trouble understanding it. It would immediately make sense."* Another participant added, *"I kept wondering, 'Does my interpretation match what the researcher has in mind?' because some expressions were not specific enough."*

Verbal Probing/Recall: Participants reported having difficulty recalling their experiences when answering questions about when a behavior occurred. One participant said, *"I find it hard to remember exactly how many times I performed a behavior in the past six months,"* adding, *"I don't consciously track these behaviors."* The same participant continued, *"Since I couldn't recall an exact number, I gave an approximate, guess-based answer. But if I don't perform a behavior regularly, I don't count how many times I've done it."*

Some participants stated that the way the question was presented affected how they recalled their experiences. One participant said, *"When I saw the 'six months' phrase, I panicked. I don't have that kind of time scale in my mind. If it had been phrased as 'last week' or 'last month,' I would have answered with much more confidence."* Another participant commented, *"I wasn't sure whether I should count individual events or provide a general impression of my routines."*

Verbal Probing/Confidence Assessment: Participants rated their confidence in the answers they provided. Most participants exhibited uncertainty when selecting their responses. For example, one participant said, *"I don't know whether my answer is correct; I made a reasonable guess."*

Another participant stated, *"I wasn't completely sure about my answer, but I selected the most suitable option I could."*

Some participants reported questioning the answers they gave. One participant said, *"I went back and forth between two options because I wasn't sure about either."* Another participant explained, *"I hesitated while responding because I wasn't sure how specific my answer needed to be,"* and added, *"I asked myself: 'Should my answer be precise, or can I provide an approximate assessment?'"*

Verbal Probing/Response Option Selection: Some response options were found to be too similar, making it difficult for participants to distinguish among them. One participant said, *"The response options were very similar, and I was undecided between two. I couldn't tell whether there was any real difference in meaning between 'sometimes' and 'occasionally,' and whether they were truly distinct."* Another participant said, *"I couldn't differentiate between 'often' and 'sometimes,' and since I'm not someone who does this behavior regularly, I couldn't decide."*

Some participants also stated that the response ranges were disproportionate. One participant noted, *"I noticed there were too many options for positive statements but not enough for negative ones."* Another participant said, *"I felt pushed to choose a more general option."* Another participant highlighted the limitation of the response range by saying, *"My real answer fell between two options; there wasn't an intermediate option that fit me."*

Verbal Probing/Elaboration: When participants were asked to rephrase the questions in their own words, they sometimes generated alternative meanings that did not align with the researcher's intended meaning. This revealed the need to restructure the statements to be less complex. One participant suggested, *"I would rephrase this question as: 'In a week, how much attention do you pay to your emotions in your life?' Because when I first heard it, the question felt a bit complex to me."* Another participant said, *"The way I understand a question may not match how another person understands it,"* and added, *"If it were in an everyday conversation, I would answer differently."*

Some participants had difficulty mentally translating the items. One participant said, *"I don't know if the translation I made is equivalent, but this was the first thing that came to mind regarding this question."* Another participant expressed uncertainty by saying, *"I translated it in my head, but I'm not exactly sure what it means."*

Revisions

The repeated cognitive interview process enabled the items to be shaped in a more straightforward, more understandable way with higher response consistency. The revisions made at this stage aimed to increase ease of use, reduce response bias, and strengthen the accuracy of participants' self-reports. Simplifying complex terms was among the most prominent changes in the process. To ensure consistency of interpretation, more understandable

definitions were added within the items for terms such as “resilience” and “mindfulness.” In addition, long, complex sentence structures were simplified to short, direct expressions to reduce participants’ cognitive load. To facilitate the recall of retrospective memories, briefer and more concrete time intervals (e.g., “last month”) were added to general time references such as “six months.”

A concrete example of these changes is presented below:

Initial version: “How often have you practiced stress-management techniques in the past six months?”

Final version: “In the past month, how many times have you used techniques such as exercise, deep breathing, or meditation for relaxation?”

This rewritten question was adjusted to address participant feedback, including recall difficulty and a lack of detail. By clarifying the question structure and adding examples that participants could relate to, the accuracy and clarity of responses were improved.

Additionally, the response scales were modified to be more balanced and easier to read. The initial scale of “Never, Sometimes, Often, Always” was updated to “Never, Rarely, Occasionally, Frequently, Always.” Such improvements made the items easier to read and enabled participants to report their behaviors with greater confidence and accuracy. Furthermore, the qualifier “almost” in options such as “Almost always” and “Almost never” was removed from the scale, in line with expert recommendations. With these adjustments, the final version of the scale achieved high reliability and validity, and methodological integrity was maintained in self-report wellness assessments.

4. Discussion

The findings of the present study demonstrate the importance of using plain language and cultural appropriateness in developing a multidimensional wellness scale. The iterative cognitive interview process identified issues related to key terms, structural complexity, and response accuracy. It revealed that these core issues significantly affect participants’ understanding of items and their ability to respond accurately. The findings are consistent with previous studies that emphasize the importance of clear and straightforward language in self-report instruments (Beatty & Willis, 2007; Collins, 2003; Willis, 2015).

One of the study’s original contributions is the use of an iterative cognitive interview process as a systematic tool to develop the item pool and enhance measurement validity. While most traditional approaches to survey validity focus on quantitative psychometric tests, the method employed in this study enabled the iterative development of explicit statements by directly incorporating participant feedback across multiple stages. The findings revealed that ambiguous terms and complex language impose a

cognitive burden on participants and negatively affect response accuracy (Desimone & Le Floch, 2004). The importance of using clear, direct language in psychometric instruments is frequently emphasized, particularly to reduce bias in the measurement of self-reported health behaviors (Miller & Foster, 2010; Tourangeau, 2000). By demonstrating how iterative cognitive interviews improve item comprehensibility and response quality, this study makes an empirical contribution to the field.

Existing measurement tools have been criticized for using fixed, static language that limits participants’ interpretation (Bolaños-Medina & González-Ruiz, 2012; Eadie et al., 2006; Uher, 2023). The findings of this study extend these critiques by showing that comprehensibility, phrasing, and response options need to be systematically evaluated and improved. The use of quantitative qualifiers such as “almost” has, in many cases, led to confusion among participants and variability in responses, thereby weakening test development efforts (Rubin, 2004). As Sudman et al. (1996) noted in their work on survey methods, the use of quantitative expressions in response options can undermine the reliability of Likert-type scales. The results of the present study corroborate these concerns in the context of wellness measurement.

Moreover, the iterative approach used in the study introduces a methodological innovation to the scale development process. Whereas traditional psychometric tests rely solely on quantitative validity criteria (Krosnick, 1991), this study integrates a qualitative feedback loop into an iterative model for revising the scale. The use of such an iterative model ensures that the final version of the scale is not only statistically sound but also consistent across contexts and meanings.

Cognitive interviews, a long-standing technique in survey methodology, are used to determine whether participants understand the study and items as intended by researchers (Willis, 2015). In this method, participants respond to survey questions while simultaneously verbalizing their cognitive processes, using “think-aloud” protocols and verbal probing techniques (Beatty & Willis, 2007). Cognitive interviews enable the identification of various aspects such as the quality and clarity of item formulation, the suitability of response options, and cognitive load (Collins, 2003). However, a key limitation of the cognitive interview technique is that it is typically applied only once. This means that item revisions are made based on one-time feedback and that there is no systematic mechanism for iterative adjustments (Desimone & Le Floch, 2004).

In contrast, iterative cognitive interviewing adopts a multi-stage, continuously cyclical approach, allowing scale items to be retested and revised across multiple participant groups (Willis, 2015). This approach is based on the principle that each subsequent version of the scale should be better than the previous one, and it aims to systematically address comprehension problems, response

inconsistencies, and undesired biases that arise during the process (Beatty & Willis, 2007). By re-administering revised survey questions to new participants, iterative cognitive interviews offer significant benefits in terms of content validity and response accuracy that classical cognitive interviews cannot.

Another key finding concerns recall difficulty, which affects response accuracy. Participants found it difficult to remember their past behaviors over extended time frames (e.g., six months). This recall difficulty led to inconsistent self-reports. Shortening the recall period (e.g., from six months to one month) significantly increased participants' confidence and the clarity and precision of their responses. This aligns with retrieval theory (Tourangeau, 1984). This finding also supports research in the literature suggesting that long recall periods cause over- or under-reporting (Bradburn et al., 1987).

The study also examined participants' response processes. Some participants had difficulty distinguishing between response options with very similar meanings (e.g., 'sometimes' versus 'often'). This finding highlights the importance of response scales being precise and well-calibrated. It is also known that ambiguous quantitative qualifiers lead to undesired response biases (DeVellis & Thorpe, 2021). In this study, restructuring response options with more distinct and equidistant alternatives increased the reliability of responses and the usability of the scale.

Another important finding is the use of ambiguous frequency qualifiers in response options, such as "almost" (e.g., "almost always," "almost never"). Although such qualifiers aim to distinguish frequency levels in a graded manner, they did not provide a precise reference point for participants, leading to ambiguity. This ambiguity led participants to make subjective interpretations, resulting in inconsistent responses. Participants frequently questioned whether "almost always" indicated 90% or a strong but inconsistent tendency. Similarly, "almost never" was interpreted by some participants as very rare occurrences, whereas others perceived it as a state very close to "never" but not absolute.

In the present study, by removing ambiguous qualifiers such as "almost" and making response options more explicit and precise, response accuracy, reliability, and interpretability were improved. This approach aligns with best practices in scale development, which emphasize that response options should provide clear distinctions without creating subjective ambiguity (Rubin, 2004). Participants' ability to accurately match their experiences to the options provided in the scale is critical for minimizing measurement error and improving data validity.

The study has some limitations. First, the sample size ($n = 10$) is relatively small and may limit the generalizability of the findings. Although 5–15 participants are often considered sufficient for cognitive interviews in the literature (Beatty & Willis, 2007), testing the scale with

larger and more diverse samples across different age groups, occupations, and cultural contexts—and, where necessary, administering it to the targeted group—will increase the reliability and validity of the findings.

Second, although iterative cognitive interviews are effective at improving the linguistic and conceptual appropriateness of scale items, they are not sufficient on their own to ensure the scale's scientific validity and reliability. In the scale development process, cognitive interviews must be supported by systematic psychometric analyses (e.g., item response theory, factor analysis). Integrating such analyses can further enhance the measurement precision and construct validity of the scale, while also contributing to consistent, generalizable results across different samples. Therefore, while iterative cognitive interviews play a critical role at the outset of scale development, their integration with psychometric methods is essential for creating a measurement tool with high validity and reliability.

The present study demonstrates the effectiveness of iterative cognitive interviews in improving the measurement of wellness. By systematically addressing comprehension, recall, and response biases, a scale has been developed that is linguistically more accessible, conceptually more straightforward, and methodologically more robust. In this context, the iterative approach used provides a replicable model centered on a continuous improvement process informed by participant feedback. It can serve as an example for future scale development studies.

In future research, the revised scale should be administered to different participants, its longitudinal consistency examined, and supported with quantitative validation techniques (e.g., item response theory, factor analysis).

In conclusion, the present study shows that iterative cognitive interviews are a powerful methodological tool in the scale development process. However, when supported by larger samples, psychometric validation techniques, and cross-cultural studies, the developed scale is expected to become a more scientifically robust, generalizable, and reliable measurement instrument.

Ethical Considerations: In this study, all rules specified within the "Directive on Scientific Research and Publication Ethics of the Higher Education Institutions" were followed, and none of the actions listed under Section Two, titled "Actions Contrary to Scientific Research and Publication Ethics," were committed. This study was conducted with the approval of the Middle East Technical University Human Research Ethics Committee, dated 06/05/2024 and numbered 0266-ODTÜIAEK-2024.

Author Contributions : Introduction: First author, Second author. Method: First author, Third author. Findings: First, Second, Third, and Fourth authors. Discussion: First, Second, Third, and Fourth authors.

Funding : No financial support was received for this study.

Conflict of Interest : There is no conflict of interest in this study.

Data Availability : The data used in this study are available from the authors upon reasonable request.

References

- Beatty, P. C., & Willis, G. B. (2007). Research synthesis: The practice of cognitive interviewing. *Public Opinion Quarterly*, 71(2), 287–311. <https://doi.org/10.1093/poq/nfm006>
- Bolaños-Medina, A., & González-Ruiz, V. (2012). Deconstructing the translation of psychological tests. *Meta: Journal des Traducteurs/Meta: Translators' Journal*, 57(3), 715–739. <https://doi.org/10.7202/1017077ar>
- Bradburn, N. M., Rips, L. J., & Shevell, S. K. (1987). Answering autobiographical questions: The impact of memory and inference on surveys. *Science*, 236(4798), 157–161. <https://doi.org/10.1126/science.236.4798.157>
- Braun, V., Clarke, V., & Weate, P. (2016). Using thematic analysis in sport and exercise research. In B. Smith & A. C. Sparkes (Eds.), *Routledge handbook of qualitative research in sport and exercise* (pp. 213–227). Routledge.
- Collins, D. (2003). Pretesting survey instruments: An overview of cognitive methods. *Quality of Life Research*, 12, 229–238. <https://doi.org/10.1023/A:1023254226592>
- Desimone, L. M., & Le Floch, K. C. (2004). Are we asking the right questions? Using cognitive interviews to improve surveys in education research. *Educational Evaluation and Policy Analysis*, 26(1), 1–22. <https://doi.org/10.3102/01623737026001001>
- DeVellis, R. F., & Thorpe, C. T. (2021). *Scale development: Theory and applications*. Sage Publications.
- Dunn, H. L. (1959). What high-level wellness means. *Canadian Journal of Public Health/Revue Canadienne de Santé Publique*, 50(11), 447–457.
- Eadie, T. L., Yorkston, K. M., Klasner, E. R., Dudgeon, B. J., Deitz, J. C., Baylor, C. R., Miller, R. M., & Amtmann, D. (2006). Measuring communicative participation: A review of self-report instruments in speech-language pathology. *Journal of Speech, Language, and Hearing Research*, 49(5), 904–920. [https://doi.org/10.1044/1058-0360\(2006\)030](https://doi.org/10.1044/1058-0360(2006)030)
- Haldane, V., Chuah, F. L., Srivastava, A., Singh, S. R., Koh, G. C., Seng, C. K., & Legido-Quigley, H. (2019). Community participation in health services development, implementation, and evaluation: A systematic review of empowerment, health, community, and process outcomes. *PLoS ONE*, 14(5), e0216112. <https://doi.org/10.1371/journal.pone.0216112>
- Hettler, B. (1984). Wellness: Encouraging a lifetime pursuit of excellence. *Health Values*, 8(4), 13–17.
- Kılıç, K., ve İnce, M. L. (2016). Sporda pozitif gençlik gelişimi ölçeklerinin 8-14 yaş grubu çocuk ve genç sporcular tarafından anlaşılabilirliğinin incelenmesi. *Beden Eğitimi ve Spor Bilimleri Dergisi*, 10(2), 213–225.
- Krosnick, J. A. (1991). Response strategies for coping with the cognitive demands of attitude measures in surveys. *Applied Cognitive Psychology*, 5(3), 213–236. <https://doi.org/10.1002/acp.2350050305>
- Kuru, H. (2022). *Designing and evaluating a need-based employee wellness program* [Doctoral dissertation, Middle East Technical University].
- Kuru, H. (2024). Sağlık, zindelik ve iyi oluş kavramlarının incelenmesi. *İstanbul Rumeli Üniversitesi Sağlık Bilimleri Dergisi*, 2(4), 1–11.
- Miller, G., & Foster, L. T. (2010). *Critical synthesis of wellness literature*. Wellness Research Institute.
- Myers, J. E., & Sweeney, T. J. (2004). The indivisible self: An evidence-based model of wellness. *Journal of Individual Psychology*, 60, 234–244.
- Myers, J. E., Sweeney, T. J., & Witmer, J. M. (2000). The wheel of wellness counseling for wellness: A holistic model for treatment planning. *Journal of Counseling & Development*, 78(3), 251–266. <https://doi.org/10.1002/j.1556-6676.2000.tb01906.x>
- Myers, N. D., Prilleltensky, I., Prilleltensky, O., McMahon, A., Dietz, S., & Rubenstein, C. L. (2017). Efficacy of the fun for wellness online intervention to promote multidimensional well-being: A randomized controlled trial. *Prevention Science*, 18, 984–994. <https://doi.org/10.1007/s11121-017-0779-z>
- National Wellness Institute. (2020). *Six dimensions of wellness*. <https://nationalwellness.org/resources/six-dimensions-of-wellness/>
- Roscoe, L. J. (2009). Wellness: A review of theory and measurement for counselors. *Journal of Counseling & Development*, 87(2), 216–226. <https://doi.org/10.1002/j.1556-6678.2009.tb00570.x>
- Rubin, D. B. (2004). *Multiple imputation for nonresponse in surveys* (Vol. 81). John Wiley & Sons.
- Sudman, S., Bradburn, N. M., & Schwarz, N. (1996). *Thinking about answers: The application of cognitive processes to survey methodology*. Jossey-Bass.
- Tourangeau, R. (1984). Cognitive sciences and survey methods. In T. B. Jabine, M. L. Straf, J. M. Tanur, & R. Tourangeau (Eds.), *Cognitive aspects of survey methodology: Building a bridge between disciplines* (pp. 73–100). National Academy Press.
- Tourangeau, R. (2000). *The psychology of survey response*. Cambridge University Press.
- Uher, J. (2023). What's wrong with rating scales? Psychology's replication and confidence crisis cannot be solved without transparency in data generation. *Social and Personality Psychology Compass*, 17(5), e12740. <https://doi.org/10.1111/spc3.12740>
- Willis, G. B. (2015). *Analysis of the cognitive interview in questionnaire design*. Oxford University Press.
- World Health Organization. (2018). *Health in all policies as part of the primary health care agenda on multisectoral action*. World Health Organization.
- World Health Organization. (2021). *Towards developing WHO's agenda on well-being*. World Health Organization.