

Tutum Ölçeği Hazırlamada Nitel ve Nicel Adımlar

Mahire Okcay ÇAM*

Leyla BAYSAN-ARABACI**

* Prof. Dr., Ege Üniversitesi
Hemşirelik Yüksekokulu Psikiyatri
Hemşireliği Anabilim Dalı,
Bornova-İzmir

* Dr., Üniversitesi Hemşirelik
Yüksekokulu Psikiyatri Hemşireliği
Anabilim Dalı, Bornova-İzmir
e-mail:leyla.baysan@ege.edu.tr

Özet

Hemşirelik araştırmalarında, tutum belirlemede sıklıkla yurt dışında geliştirilen ve Türkçe uyarlaması yapılan araçlar kullanılmaktadır. Bu sonucun oluşmasında, ölçme aracı geliştirme konusunda literatürde çok fazla bilgi yer almaması, bu konudaki bilgilerin henüz çok yeni olması ve hemşire araştırmacıların bu konuda yeterli ve kapsamlı bilgi sahibi olmamaları etkili olmaktadır. Ancak, farklı kültürlerde geliştirilen araçlar, toplumsal norm, değer ve bakış açımızı tam olarak yansıtamamaktadır. Özellikle de kültürden yoğun bir şekilde etkilenen tutumu belirlemeye yönelik araçlar değerlendirmelerde sınırlılıklar yaşanmasına neden olabilmektedir. Bu nedenle, kültürel yapıya uygun ölçme araçları geliştirme gerekliliği doğmaktadır. Bir ölçme aracı geliştirme, belirli süreçleri gerektirir ve ölçme aracının belirli özelliklere sahip olması gerekir. Ölçme aracı geliştirme çalışmaları, genellikle niteliksel veya niceliksel olmak üzere iki farklı şekilde gerçekleştirilir. Bir ölçeğin hem niteliksel hem de niceliksel adımlar kullanılarak oluşturulması, daha kuvvetli psikometrik özelliklere sahip ölçme araçlarının geliştirilmesine olanak sağlamaktadır. Bir ölçme aracında olması gereken özelliklerin belli başlıları ise; geçerlik ve güvenilirliktir. Bu makalenin amacı, bir tutum ölçeği hazırlamada niteliksel ve niceliksel adımlar konusunda bilgi vermektir.

Anahtar Sözcükler: Tutum ölçeği hazırlama, geçerlik ve güvenilirlik, niteliksel ve niceliksel adımlar

Qualitative and Quantitative Steps on Attitude Scale Construction

In nursing researches, the materials frequently developed abroad and adapted to Turkish are used for defining attitude. In having this result that there isn't so much data in literature about developing measurement material, the data about that subject are just new and nurse researches don't have enough and extensive data is effective. However, the materials developed in different cultures can not reflect our social norm, value and point of view completely. Especially materials towards defining the attitude intentionally affected by the culture can cause having limitedness in assessments. For this reason, ne-

ed to developing measurement materials suitable with our cultural structure arise. Developing a measurement material requires particular processes and the measurement material must have particular characteristics. The studies of developing measurement materials are carried out by two different ways as usually being qualitative or quantitative. A scale's being used by both qualitative and quantitative steps enables developing measurement materials having stronger psychometric characteristics. Main characteristics that a measurement material must have are; validity and reliability. The purpose of this article is to give information about the qualitative and quantitative steps on attitude scale construction.

Keywords: Attitude scale construction, validity and reliability, qualitative and quantitative steps.

Giriş

Psikoloji ve sosyal bilim araştırmalarının önemli bir bölümünde, veriler ölçme sonucunda oluşmaktadır. Ölçme, araştırma değişkenleriyle ilgili toplanan bilgileri belli ilkeler çerçevesinde sembollere dönüştürmektir. Değişkenler, sayısal ya da sınıfsal sembollere dönüştürülebilir (Kırcaali-İftar 1999).

Psikoloji ve sosyal bilim araştırmalarında yapılan ölçümlerin büyük çoğunluğunda, araştırma verilerini toplamak amacıyla ölçme araçları kullanılır. Kullanılan ölçme araçlarının, belli özelliklere sahip olması gerekir. Bu özelliklerin belli başlıları; geçerlik ve güvenilirliktir. Gerek önceden geliştirilmiş bir ölçme aracını kullanmaya karar verirken, gerekse yeni bir ölçme aracı geliştirirken, geçerlik ve güvenilirlik özellikleri dikkate alınmalıdır (Kırcaali-İftar 1999).

Ölçek geliştirme çalışmaları, genellikle niteliksel veya niceliksel olmak üzere iki farklı şekilde gerçekleştirilir. Niteliksel süreçte, büyük örneklem gruplarına ulaşamama durumlarında aday ölçek formundaki maddelere ilişkin uzman görüşleri alınarak nitel çalışma yapılır. Nitel çalışmada uzman görüşleri arasındaki uyumluluklar

test edilir. Niteliksel süreçte kapsam geçerlik oranları ve kapsam geçerlik indeksleri hesaplanırken, nitel süreç istatistiksel nicel bir sürece dönüştürülür (Yurdugül ve Aşkar 2008).

Niteliksel süreçte ise; literatür ya da uzman yaklaşımlarıyla aday ölçek formu elde edilir. Bu aday form hedef kitle ile benzer özellikler taşıyan bir örneklem grubuna deneme uygulaması yapılarak ölçek maddelerine ilişkin psikometrik özellikler belirlenip, ideal maddelerden nihai form elde edilir. Nicel bir çalışma olan bu süreçte, genellikle faktör analizleri kullanılır ve büyük örneklemle çalışmak gerekir (Yurdugül ve Aşkar 2008).

Ölçek geliştirme çalışmalarının, zorlu bir süreci gerektirmekle birlikte, hem niteliksel hem de niceliksel adımlarla gerçekleştirilmesi, daha kuvvetli psikometrik özelliklere sahip ölçme araçlarının geliştirilmesine olanak sağlamaktadır. Bu makalenin amacı, bir tutum ölçeği hazırlamada niteliksel ve niceliksel adımlar konusunda bilgi vermektir.

Tutum Ölçeği Hazırlamada Niteliksel ve Niceliksel Adımlar

Tutum psikolojik bir değişkendir ve diğer psikolojik değişkenler gibi kompleks, bileşik bir değişkendir. Bu sebeple, tutumu ölçmek, fiziksel değişkenleri ölçmekten çok daha güçtür; ama olanaksız değildir (Erkuş 2003).

Bireylerin bir konudaki görüş, düşünce ya da tutumlarını belirlemeye yarayan ölçme araçlarına tutum ölçeği adı verilmektedir. Bir tutum ölçeği, hedeflenen bir konudaki olumlu ve olumsuz görüş, düşünce ya da tutumları yansıtan bir dizi maddeden oluşur. Maddeler, ölçülmek istenen özelliğin kuramsal tanımından yola çıkarak gözlenebilir ve ölçülebilir bir biçimde eyleme vuruk olarak tanımlanmalıdır. Tutum ifadeleri, tutum nesnesi ile ilgili olabilecek yaşantılarda yer alan düşünsel, duygusal ve eyleme yönelik öğelerin tümünü veya ölçülmek istenen boyutu kapsamalıdır (Karasar 1995, Tezbaşaran 1997, Kırcaali-İftar 1999).

Tutum Ölçeği Hazırlamada Niteliksel Adımlar

Tutum ölçeği hazırlamada niteliksel adımda veri toplamak için, bireylerin ölçülmek istenen konu ile ilgili tutumlarını sözlü olarak ifade etmelerine olanak tanıyan Odak Grubu Araştırma Deseni kullanılabilir. Nitel araştırma desenlerinden odak grup araştırmaları, bireylerin belirli bir konuyla ilgili görüş, inanç, tutum, duygu, deneyim ve değerlerini ortaya koymaya olanak tanır (Debus 2003, Şencan 2005). Odak grup araştırmalarında, veri toplama aracı olarak görüşme tekniklerinden Odak Grup Görüşmesi kullanılabilir ve veri kaynakları, araştırmanın odaklandığı konuyla ilgili az sayıda kişinin bir araya getirilmesiyle oluşturulur (Şencan 2005). Odak grup görüşmesine katılacak bireylerden oluşan çalışma grubunun belirlenmesinde, nitel araştırma örneklem yöntemlerinden Amaçlı Örneklem Yöntemleri kullanılabilir. Birden çok yöntem vardır. Tutum ölçeği hazırlamak amacıyla yapılan odak grup görüşmelerinde, araştırma amacı doğrultusunda değişebilmekle birlikte Benzeşik (Homojen) Örneklem yöntemi kullanılır. Bu yöntem, benzeşik özelliklere sahip bireyleri bir araya getirerek küçük belirgin bir alt grup oluşturmaya olanak tanır (Yıldırım ve Şimşek 2006). Bir odak grupta olması gereken optimum kişi sayısı 6-8 veya 6-12'dir (Şencan 2005, Yıldırım ve Şimşek 2006). Ancak, Yıldırım ve Şimşek (2006) tarafından, nitel araştırmalarda geçerli olabilecek ve her araştırmaya uyabilen örneklem belirleme yöntemleri sunmanın mümkün olmadığı ve bazen bir bireyin araştırmanın örneklemini oluşturabileceği ifade edilmektedir. Odak grup görüşmeleri öncesi, odak gruba katılması planlanan kişilerle bir ön görüşme yapılır ve bu görüşmede bireylere odak grubun ne olduğu, görüşmeler sırasında ses kayıt cihazı veya raportör kullanılacağı ve ölçülmek istenen konu diğer bir deyişle araştırma açıklanır. Bundan sonra odak grup görüşmelerine katılmayı kabul eden kişilerle ölçülmek istenen tutum yönünden benzeşik özellikler gösteren bir alt grup

oluşturulur (Yıldırım ve Şimşek 2006). Oluşturulan bu gruptaki bireylerle, iş yoğunluğunun en az olduğu gün ve saatlerde odak grup görüşmeleri yürütülür. Odak grup görüşmeleri, her bir görüşmenin ayrıntısı ve derinliği dikkate alınmakla birlikte, en az yarım saat olmak üzere (Yıldırım ve Şimşek 2006) minimum üç-beş oturumdan oluşur. Bu görüşmeler, grup dinamikleri konusunda eğitim almış, grubu idare etmesini bilen ve grup psikolojisinden anlayan kişiler tarafından yürütülür (Şencan 2005).

Odak grup görüşmeleri bir rehber doğrultusunda yürütülür. Bu rehberin konu hakkında bir uzman tarafından geliştirilmesi veya değerlendirilmesi önerilir (Debus 2003). Bu rehberde yer alan soruların açık uçlu olmasına özen gösterilir. Odak grup araştırmalarında, açık uçlu ve birden fazla yanıtı olan sorular sorulması ve "evet-hayır" yanıtının verildiği kısa sorulardan kaçınılması gerektiği belirtilmektedir (Şencan 2005).

Niteliksel adımda, araştırmaya katılan bireylerin konu ile ilgili tutumlarını yazılı olarak ifade etmelerine olanak tanıyan açık uçlu sorulardan oluşan bir kompozisyon yazdırma, veri toplamak için bir başka yoldur. Kompozisyonlar, odak grup görüşmesine kıyasla aynı anda daha çok sayıda bireye ulaşmaya olanak sağlar. Böylece konu ile ilgili daha çok sayıda bireyin görüşleri elde edilmiş olur. Ancak, araştırmaya katılan bireyler tarafından kompozisyonların yeterli ölçüde veya hiçbir şekilde yazılmama olasılığı da vardır. Bu da kompozisyon yazdırmanın zayıf yönünü oluşturmaktadır.

Ölçek hazırlamanın niteliksel adımında veri toplamak için kullanılacak bir başka yol araştırmacı tarafından literatürde yer alan ölçekleri incelemektir. Burada araştırmacı literatürde yer alan, konu ile ilgili geliştirilmiş çeşitli tutum ölçeklerini inceleyerek bir madde havuzu oluşturur. Her bir veri toplama yönteminin güçlü ve zayıf yönleri dikkate alınarak her üçü veya herhangi ikisi birlikte kullanılarak daha nitelikli ölçme araçları geliştirilebilir.

Niteliksel adımda, herhangi iki veya her üç yöntem kullanılarak veri toplandıktan sonra bu veriler araştırmacı tarafından analiz edilir. Öncelikle odak grup görüşmeleri sonucunda elde edilen görüşme kayıtları araştırmacı tarafından çözümlenmek üzere tek tek yazılır. Bu şekilde sayfalarca kayıt elde edilmiş olur. Daha sonra hem odak grup görüşmelerinden hem de kompozisyonlardan elde edilen veriler üzerinde, onları açıklayabilecek kavramlara ve ilişkilere ulaşmak amacıyla araştırmacı tarafından içerik analizi yapılır (Yıldırım ve Şimşek 2006). İçerik analizinde tümevarımcı analiz yaklaşımı tercih edilir.

İçerik analizi için aşağıda yer alan adımlar atılmalıdır.

- Öncelikle, genel bir çerçevede içinde kodlama yapılır. Bu kodlama türüne göre, verilerin analizinden önce araştırılan tutum konusunda genel bir kavramsal yapı oluşturulur. Daha sonra, araştırmanın kavramsal çerçevesi dikkate alınarak, kaydedilen veriler, araştırmacı tarafından satır satır okunarak anlamlı bölümlere (sözcük ve cümleler gibi) ayrılır ve her bölüme tanımlayıcı isimler (kodlar) verilir. Böylece kodlar, bir cümleye, bir sözcüğe ya da bir sözcük grubuna ilişkin olarak oluşturulur. Bu işlem sırasında ortaya çıkan kodlar, doğrudan verilerin yan tarafına (kağıdın sağ kenarına) yazılır. Veri kodlama süreci boyunca, veri seti araştırmacı tarafından tekrar tekrar okunarak, ortaya çıkan kodlar üzerinde tekrar tekrar çalışılır (Yıldırım ve Şimşek 2006).

- İkinci olarak kavramlar oluşturulur. Kodlanan veriler arasında yer alan anlamlı bölümlere ve olaylara o bölümdeki anlamı en iyi yansıtabilecek tanımlamalar yapılır (Yıldırım ve Şimşek 2006).

- Son olarak kavramlar daha önceden belirlenmiş temalar kapsamında sınıflandırılır. Bunun için ortaya çıkan kodların benzerlik ve farklılıkları, yakınlık ve uzaklıkları değerlendirilir ve birbiri ile ilişkili kodlar önceden belirlenen veya yeni belirlenecek olan temalar altında bir araya getirilerek düzenlenir. Böylece birbiri ile anlamlı bütün oluş-

turan veriler, belirli temalar altında yer almış olur (Şencan 2005, Yıldırım ve Şimşek 2006).

Özetle, toplanan veriler önce kodlanmış ve kavramlaştırılmış, daha sonra ortaya çıkan kavramlara göre mantıklı bir biçimde düzenlendikten sonra, veriyi açıklayan homojen, objektif, ayırt edici niteliği olan ve bütünsellik taşıyan temalar altında sınıflandırılmış olur. Böylece içerik analiziyle, birbirine benzeyen veriler, belirli temalar çerçevesinde bir araya getirilerek düzenlenmiş ve betimlenmiş olur (Bilgin 1999, Şencan 2005, Yıldırım ve Şimşek 2006).

Ölçeğin hazırlanmasında niteliksel adımda son olarak, çalışma grubuyla yapılan odak grup görüşmelerinin ve kompozisyonların içerik analizinden elde edilen tüm veriler ve ilgili literatürün taranması sonucunda oluşturulan madde havuzundaki ifadeler, araştırmacı tarafından değerlendirilerek, tutum ifadesi niteliği taşıyan cümlelere dönüştürülür (Yıldırım ve Şimşek 2006). Tutum cümleleri geliştirilirken, daha fazla veri kaynağında (odak grup görüşmesi, kompozisyon ve madde havuzu) yer alan ve daha sık tekrar edilen kavramlar temel alınarak tutum cümleleri oluşturulur.

Tutum ifadeleri yazılırken göz önüne alınması gereken kurallar aşağıda yer almaktadır (Erkuş 2003, Tavşancıl 2006).

- “Geçmiş” değil, “şimdiyi” ifade eden cümleler kurulmalıdır.

- Olgusal ifadelerden ya da olgusal gibi yorumlanan ifadelerden kaçınılmalıdır.

- Değişik biçimde yorumlanabilen, muğlak ifadelerden kaçınılmalıdır.

- İncelenen psikolojik nesneyle ilişkili olmayan ifadelerden kaçınılmalıdır.

- “Hepsi” ya da “hiçbiri” gibi ifadelerden kaçınılmalıdır.

- İlgilenilen duyuşsal ölçeğin tüm ranjını kapsayan ifadeler seçilmelidir.

- Basit, açık ve doğrudan ifade edilen bir dil kullanılmalıdır.
- İfadeler kısa olmalıdır.
- Her bir ifade yalnızca bir tek düşünceyi kapsamalıdır.
- Belirsizlik yaratabilen genel ifadelerden (her zaman, hepsi, asla gibi) kaçınılmalıdır.
- “Sadece” gibi sözlükler dikkatle kullanılmalıdır.
- Cümleler karmaşık değil, basit olmalıdır.
- Ölçeği alanlar tarafından anlaşılmayacak sözcükleri kullanmaktan kaçınılmalıdır.

Tutum Ölçeği Hazırlamada Niceliksel Adımlar

Tutum ölçeğinin hazırlanmasında niceliksel adımda nicel araştırma tekniklerinden yararlanılarak ölçeğin geçerlilik ve güvenilirlik analizleri yapılır.

Ölçeğin Geçerliliği

Geçerlilik çalışmasıyla, bir tutum ölçeğinin geliştirildiği amaca hizmet derecesi; yani, ölçeğin ölçülmek istenen özelliğe uygun olup olmadığı belirlenir. Başka bir ifadeyle, ölçümün kurallara uygun olarak doğru yapıldığı ve ölçüm verilerinin gerçekten ölçülmek istenen özelliği yansıtmayı yansıtmadığı belirlenmiş olur (Akçul 1997, Şencan 2005). Bunun için aşağıda belirtilen geçerlik ölçüm yöntemlerine başvurulur.

1. Kapsam (İçerik) Geçerliliği: Ölçekte bulunan maddelerin ölçülmek istenen alanı temsil edip etmediğini belirlemek amacıyla kapsam geçerliliği değerlendirilir. Ölçeğin, kapsam geçerlik çözümlemesi için, uzman görüşlerine dayalı nitel çalışmaları, istatistiksel nicel çalışmalara dönüştüren iki süreç kullanılır: Lawshe Tekniği ve Davis Tekniği (Yurdugül ve Aşkar 2008).

Bu makalede, daha yaygın kullanılan Lawshe tekniğine göre kapsam geçerliliği anlatılacaktır. Lawshe tekniğine göre kapsam geçerliliği için hazırlanan aday ölçeğin, kapsam geçerlik oranları ve indeksleri hesaplanır. Bunun için aşağıdaki yol

izlenir (Lawshe 1975, Şencan 2005, Yurdugül ve Aşkar 2008)

a) Nitel araştırma teknikleri kullanılarak aday ölçek formu geliştirilir. Geliştirilmek istenen ölçekteki konu hakkında alanda yeterlilik sahibi olan en az 5 en fazla ise 40 uzman belirlenir (Yurdugül ve Aşkar 2008).

b) Belirlenen uzmanlara bir değerlendirme formu aracılığıyla görüşleri sorulur. Bu görüşlerin belirlenmesinde daha önceden hazırlanmış Uzman Değerlendirme Formu kullanılabilir. Bu form, uzman görüşü vermeyi kabul eden uzmanlara, elden ve e-mail yoluyla ulaştırılır.

Uzmanlardan formda yer alan her bir maddeyi, aşağıdaki kriterler doğrultusunda değerlendirmeleri istenir.

- Ölçülecek tutum özelliklerini temsil ediyor mu?
- Hedef kitle tarafından kolayca anlaşılabilir mi?
- Yeteri kadar açık ifade edilmiş mi?
- Önceden belirlenmiş boyutlarda yer alabilir mi?

Bu doğrultuda uzmanların her bir maddeyi, “gerekli” (belirtilen özelliği net olarak ölçmeye aday bir madde) “yararlı, ancak yetersiz” (madde konu kapsamında ama düzenlenmesi ya da değiştirilmesi gerekiyor) ve “gereksiz” (madde belirtilen özelliği temsil etmiyor) şeklinde derecelendirmeleri istenir (Şencan 2005, Yurdugül ve Aşkar 2008).

c) Uzman görüşleri, madde madde tek bir formda birleştirilir. Bu formda her bir maddenin olası seçeneklerine kaç uzman tarafından onay verildiği toplamsal olarak belirtilir. Bu formda, her bir madde için kapsam geçerlik oranları hesaplanır.

Kapsam geçerlik oranlarını belirlemek için öncelikle aday ölçek formundaki her bir madde için “gerekli”, “yararlı, ancak yetersiz” ve “gereksiz” görüşlerini belirten uzmanların sayısı hesaplanır. Daha sonra her bir madde için o maddeye ilişkin

gerekli görüşünü belirten uzman sayısı, maddeye ilişkin görüş belirten toplam uzman sayısının yarısına bölünür. Her bir madde için bu oranın 1 ekşiği alınarak maddelere ilişkin Kapsam Geçerlik Oranı (KGO) belirlenir (Şencan 2005).

$$KGO_{\text{madde}} = \frac{NG}{N/2} - 1$$

KGO : Kapsam geçerlik oranı

NG : Maddeye “Gerekli” diyen uzman sayısı

N : Maddeye ilişkin görüş belirten toplam uzman sayısı

Aday ölçek maddelerinden KGO değerleri negatif ya da “0” değer içeren maddeler elendikten sonra, KGO değerleri pozitif olan maddelerin, Kapsam Geçerlilik Ölçütü’ne (KGÖ) göre anlamlı olup olmadığı değerlendirilir.

KGÖ, maddenin istatistiksel olarak anlamlı olup olmadığını test etmek için geliştirilen bir ölçüttür. Bu ölçüt standart normal dağılım ilkelerinden yararlanılarak elde edilmektedir (Yurdugül ve Aşkar 2008). Ancak, Veneziano ve Hooper (1997) hesaplama kolaylığı açısından, $\alpha = 0.05$ anlamlılık düzeyinde minimum Kapsam Geçerlik Ölçütü değerlerini tabloya dönüştürmüşlerdir (Tablo 1). Bu tabloya göre, ölçeğin uzman sayısı

doğrultusunda, $\alpha = 0.05$ anlamlılık düzeyinde minimum KGÖ belirlenir (Yurdugül ve Aşkar 2008).

d) Aday ölçeğin her bir maddesi için elde edilen KGO’lardan istatistiksel olarak anlamsız bulunanlar elenir. Bundan sonra, KGO $\alpha = 0.05$ düzeyinde anlamlı olan maddeler son forma alınır. Bu maddelerin toplam KGO’ların ortalamaları alınarak Kapsam Geçerlik İndeksi (KGİ) hesaplanır.

$$KGİ = \sum KGO / \text{Madde Sayısı}$$

e) KGİ / KGÖ’ye göre son form oluşturulur. Ölçeğin son formuna alınan maddelerden oluşan ilk şekli için,

$$KGİ \geq KGÖ \text{ veya } KGİ / KGÖ \geq 0'1$$

sağladığında kapsam geçerliği istatistiksel olarak anlamlı bulunur (Yurdugül ve Aşkar 2008).

Kapsam geçerliği gerçekleştirilmiş olan ölçme aracındaki maddelerin yarısı tutum boyutunun bir tarafını (olumlu uzanımı), diğer yarısı da öteki tarafını (olumsuz uzanımı) kapsayacak şekilde tekrar düzenlenir. Bu düzenlemede, olumlu ve olumsuz ifadelerin, ölçme aracı içerisinde rastgele bir biçimde dağıtılmasına, düzenli bir sıra izlenmemesine ve gruplandırılmamasına dikkat edilir (Karasar 1995, Tezbaşaran 1997, Kırcaali-İftar 1999).

Tablo 1: Lawshe Minimum Kapsam Geçerlik Ölçütleri

Uzman Sayısı	Minimum KGÖ Değeri	Uzman Sayısı	Minimum KGÖ Değeri
5	0.99	13	0.54
6	0.99	14	0.51
7	0.99	15	0.49
8	0.78	20	0.42
9	0.75	25	0.37
10	0.62	30	0.33
11	0.59	35	0.31
12	0.56	40 ve ↑	0.29

$p < 0.05$

Veneziano L., Hooper J. (1997). A Method For Quantifying Content Validity of Health-Related Questionnaires, American Journal of Health Behavior, 21(1).

2. Yüzey Geçerliliği: Yüzey geçerliliği, bir ölçeğin, araştırılan yapıyı ölçüp ölçmediğini incelemek için yapılır (Şencan 2005, Tavşancıl 2006). Mantıksal geçerlik olarak da isimlendirilen yüzey geçerliliği, ilk aşamada araştırmacının kendisi tarafından yapılır (Şencan 2005).

Hazırlanan ölçeğin yüzey geçerliliği için, ilk önce tüm madde ifadelerine yönelik uzman önerileri dikkate alınır. Daha sonra araştırmacı bu maddeleri anlaşılabilirlik ve ifade yönünden genel olarak değerlendirir ve ilgili düzenlemeleri yapar. İkinci olarak, araştırma evreni dışında yer alan evrendeki birey özelliklerine benzer küçük bir grup tarafından, ölçek maddeleri düzgünlük ve anlamlılık, okunurluk, terimlerin anlaşılabilirliği, cümlelerin uzunluğu, anlamın açıklığı ve netliği yönünden değerlendirilir. Bu pilot çalışmayla, aynı zamanda ölçeğin cevaplayıcıların eğitim düzeylerine, kültürel yapılarına ve yetenek düzeylerine uygun olup olmadığı da incelenir.

Bundan sonra, ölçekteki her bir madde standart bir anahtar çerçevesinde (tümüyle katılıyorum, biraz katılıyorum, kararsızım, pek katılmıyorum, hiç katılmıyorum ya da tümüyle yanlış, biraz yanlış, biraz doğru, çok doğru gibi) düzenlenir. Bu tip standartlarda çeşitli sayıda kategori kullanılıyorsa da, en ideal kategori sayısı 5'tir. Kategori sayısı 5'ten aşağı düşükçe ölçek düzeyi açısından bilgi kaybı; 5'ten yukarı çıktıkça da kategoriler arasındaki fark ayırt edilemez olmaktadır. Eğer bireylerin, kararsızım kategorisinden kaçındırılması ve uçlara doğru tercihe zorlanmaları istenirse, kategori sayısı 4, 6, 8 gibi çift sayılı olabilir (Kırcaali-İftar 1999, Erkuş 2003). Bu son düzenleme sonrasında ölçek araştırma evrenine uygulanır.

3. Uyum Geçerliliği: Uyum geçerliliği, yeni geliştirilen bir ölçme aracı ile aynı amaca yönelik bir başka ölçme aracı arasındaki uyumun derecesini göstermektedir (Kırcaali-İftar 1999). Uyum geçerliliği, hazırlanan ölçekle aynı kavramsal yapıyı ölçen aynı nitelikte bir başka (eşdeğer) ölçekle yapılır. Karşılaştırma yapmak için aynı nitelikte bir dış kriter ölçek bulunmadığı durumlarda, ölçülmek

istenen kavramsal yapıya benzer ilgili (alternatif) ölçeklerden yararlanılır (Şencan 2005).

Yeni hazırlanan ölçme aracından alınan puanlar, diğer eşdeğer veya alternatif ölçekten alınan puanlarla uyuşursa, yeni geliştirilen aracın uyum geçerliliğine sahip olduğu sonucuna ulaşılır (Kırcaali-İftar 1999). Uyum geçerliliği için, aynı kavramsal yapıyı ölçen ölçeklerde korelasyon katsayısının .70 - .80 gibi güçlü değerler olması gerekir. İlgili ölçeklerle yapılan karşılaştırmalarda ise, orta derecede ilişkiyi gösteren .50 - .70 büyüklüğündeki korelasyon katsayıları geçerlik kanıtı olarak değerlendirilir. İlgili ölçeklerle yapılan karşılaştırmalarda, belirsizliği ifade ettiği için, korelasyon katsayısının .30'un altında olmaması gerekir (Şencan 2005).

4. Yapı (Kavram) Geçerliliği: Hazırlanan ölçeğin hem ölçtüğü niteliklerin neler olduğunu, hem de ölçeğin uygulandığı kişilerin aldığı puanların ne anlama geldiğini araştırmak için ölçeğin yapı geçerliliği değerlendirilir (Erkuş 2003). Yapı geçerliliği için daha sıklıkla, temelde birbiri ile bağlantılı değişkenleri belli bir kümede bir araya getirmeye yarayan Faktör Analizi Yöntemi kullanılır (Aksayan ve diğ. 2002).

Faktör analizi yöntemlerinden, Keşfedici Faktör Analizi yapılır. Bu analiz, yeni geliştirilen ölçek çalışmalarında, literatürde konuyla ilgili bir kuramsal bilgi bulunmaması ve konunun kaç faktörden oluştuğunun önceden bilinmemesi durumunda uygulanır. Keşfedici faktör analizi, ölçüm değişkenlerinin ne şekilde gruplaştığını, başka bir ifade ile ölçek maddelerinin arka planında hangi faktörlerin bulunduğunu görmek için yapılır (Şencan 2005). Veriler eşit aralıklı ölçek niteliğinde (Likert gibi) ve esas amaç ölçek geliştirmek olduğunda, ölçek maddelerinin hangi başlıklar altında gruplanabileceğini saptamak için keşfedici faktör analizi yöntemlerinden Temel Bileşenler Analizi yöntemi kullanılır (Şencan 2005, Tavşancıl 2006). Birden fazla faktör (alt boyut) ortaya çıkarmak amaçlanıyorsa, dik açılı döndürme biçimlerinden biri olan Varimax Rotasyonu yapılır (Şencan 2005).

Bir ölçekte faktör analizi yönteminin uygulanabilmesi için örneklemin belirli bir büyüklüğe sahip olması gerekir. Literatürde, örneklem büyüklüğü için değişik kurallar yer almaktadır. Bunlardan biri 10 kuralı olup buna göre madde başına en az 10 katılımcı bulunmalıdır. Bir diğeri 100 kuralı olup madde başına 5 katılımcı olmalı veya en az 100 kişiye ulaşılmalıdır. Bir diğer yöntem ise, Kaiser-Meyer-Olkin (KMO) Testi'dir. KMO test sonucunun .50'den büyük olması durumunda faktör analizi yapılabilir. Ancak bilim adamları, örneklemin 100 kişiden daha az olmasını faktör analizi için yetersiz ve güvenilmez bulmaktadır (Şencan 2005).

Faktör yapılandırılmada, her bir boyutun yüksek faktör yüküne sahip en az üç değişkene (maddeye) sahip olması gerekir. Faktörler, iki veya daha az değişken içermesi durumunda bilim adamı tarafından yorumlanmaya çalışılmamalıdır. Bir değişkenin bir faktör grubunda yer alabilmesi için faktör yükünün en az .40, bazı bilim adamlarına göre ise en az .30 olması gerekmektedir. Şencan (2005) tarafından belirtildiğine göre, Kim-Yin (2004) faktör yükü değerlerinin örneklem büyüklüğüyle ilişkili olduğunu belirtmiştir. Buna göre, faktör yükü .30 olan maddelerin ölçüğe alınması için örneklem büyüklüğünün en az 350; .40 faktör yükü için 200; .50 faktör yükü için 120; .60 faktör yükü için 85; .70 faktör yükü için 60 kişilik bir örneklemin yeterli olacağı ifade edilmiştir.

Faktör analizinde, bazen bir değişkene ait faktör yüklerinin birden fazla faktöre yaklaşık olarak eşit oranda dağılması durumuyla karşılaşılabilir. Bu şekilde, bir değişkenin birden fazla faktörde eşit veya benzer yük değerlerine sahip olması durumuna çapraz yükler (cross-loadings veya split loadings) denir.

Birden fazla faktörde anlamlı çapraz yüklere sahip değerler, aşağıda görüldüğü gibi üç şekilde değerlendirilebilir (Şencan 2005).

- Döndürme yöntemine devam edilir ve faktör sayısı azaltılmaya çalışılır.

- Çapraz yüklere sahip maddelerin ifadenendirme biçimi gözden geçirilir. İfadelerin yüzey geçerliği dikkate alınarak bu maddeler kendileri için en uygun olan faktörün altına yerleştirilir.

- Yüksek çapraz yüklü maddeler ölçekten çıkarılarak analiz yeniden yapılır.

Faktör analizi sonucunda ortaya çıkan faktör yüklerine bakılarak, faktörler altında toplanabilecek değişkenler belirlendikten sonra, her bir faktöre uygun isim bulunur. Bu işleme "etiketleme" adı verilir (Şencan 2005).

Ölçeğin Güvenirliği

Hazırlanan tutum ölçeğinin hatalardan arınık olarak ölçme yapabildiğini, verileri doğru topladığını ve yinelenebilir bir ölçek olduğunu göstermek amacıyla güvenirliliği incelenir (Aksayan ve Gözüm 2002, Erkuş 2003).

Eğitimde ve psikolojide, test geliştirme aşamasında madde ve test (güvenirlik) istatistiklerini kestirirken Klasik Test Teorisi (KTT) ve Modern Test Teorisi (MTT) (Madde Tepki Teorisi veya Örtük Özellikler Teorisi) olmak üzere iki teoriden yararlanılır. Bir ölçek geliştirilirken, hangi teorisinin daha avantajlı olacağı hala üzerinde tartışılan ve sınıranan bir konudur (Kan 2006). KTT, bir çok ölçüğe uygulanmasını kolaylaştıran zayıf teorik varsayımlara sahip olması, uygulama ve parametreleri kestirme kolaylığı olması sebebiyle daha avantajlı kabul edilir (Hambleton ve Jones 1993, Kelecioğlu 2001, Kan 2006). KTT'ye kıyasla MTT, daha büyük gruplarla çalışmayı ve daha güçlü ancak yerine getirilmesi daha güç varsayımları gerektirir. Kan (2006) tarafından, bu iki kuramın madde seçme ve test geliştirme prosedürü açısından benzer sonuçlar ortaya koyup koymadığını belirlemek amacıyla yapılan bir çalışmada, her iki teorisinin de madde seçme ve test geliştirme sürecine ilişkin olarak benzer sonuçlar ürettiği ve bu sebeple bir test geliştirme sürecine ilişkin koşullar sağlandıktan sonra, testi oluşturacak maddelerin seçilmesi aşamasında uygulayıcıların iki teoriden herhangi birini kullanabileceği belirtilmektedir. Ancak, ölçek geliştirme ve ölçek

uyarlama çalışmalarında güvenilirlik incelemeleri genellikle, daha kolay ve yaygın uygulanan KTT'ye göre yapılmaktadır (Kelecioğlu 2001, Kan 2006).

KKT'ye göre güvenilirlik, gerçek değerlerle gözlemlenen değerler arasındaki ilişkinin yüksek olmasıdır ve tesadüfi hata düşük olduğu oranda bu ilişki yüksek çıkar. KKT'ye göre güvenilirlik analizi, araştırmacı tarafından kullanılan form sayısı ve seans sayısına göre değişen birden çok yöntemle yapılabilmesine karşın temelde dört grupta ele alınmaktadır: İç tutarlılık güvenirliliği, test-tekrar test güvenirliliği, paralel formlar güvenirliliği, gözlemciler arası güvenirlilik (Şencan 2005).

Yeni hazırlanan bir ölçek için, iç tutarlılık güvenirliliğinin gerekli ancak yeterli olmadığı ve ölçüm çalışmasının niteliğine göre, aynı zamanda diğer güvenirlilik analizlerinin de yapılması gerektiği belirtilmektedir (Şencan 2005). Bu makalede, iç tutarlılık güvenirliliğinin yanı sıra test-tekrar test güvenirlilik analizinden söz edilmektedir.

1. İç Tutarlılık Güvenirliliği: İç tutarlılık güvenirliliğinde, tek bir ölçüm aracı kullanılarak ve tek bir seansta ölçüm yapılarak maddelerin belirli bir kavramsal yapıyı tutarlı bir şekilde ölçüp ölçmediği araştırılır. Maddeler arasındaki iç tutarlılığı yüksek olan araçların güvenilir olduğu kabul edilir (Şencan 2005).

Geliştirilen bir ölçeğin iç tutarlılık güvenirliliğini sınamak için Cronbach Alpha iç tutarlılık katsayısı (r) hesaplanır (Erkuş 2003, Şencan 2005, Tavşancıl 2006). Alpha, arka planda yatan gizli kavramsal yapıdaki değişkenlik hakkında bilgi veren, matematiksel hesaplamalara dayanan güvenirlilik indeksi değeridir ve maddelerin birbiri ile tutarlı olup olmadığını ve maddelerin hipotetik bir değişkeni ölçüp ölçmediğini belirler. Bu sebeple, ölçeğin kullanıldığı her bir farklı örneklem için yeniden hesaplanmalıdır (Şencan 2005).

Alpha değeri tek boyutlu ölçekler için iyi bir güvenirlilik katsayısı iken, çok boyutlu ölçekler için, her bir faktör bazında alpha katsayısı ile bir-

likte test-tekrar test veya paralel formlar güvenirlilik yöntemlerinden birinin daha uygulanması önerilir (Şencan 2005).

Cronbach alpha ve diğer güvenirlilik katsayılarının ne olması gerektiği konusu bilim adamları, bilim disiplinleri ve araştırma alanlarına göre farklılık göstermektedir. Şencan (2005) ve Tavşancıl (2006) tarafından belirtildiği üzere Nunnally'e göre (1998), alpha güvenirlilik değeri .70'den büyük olmalıdır. Şencan'ın (2005) belirttiğine göre, George ve Mallery'e (2003) göre ise, alpha değerinin,

- >.90 olması "mükemmel"
- .80-.90 arasında olması "iyi"
- .70-.80 arasında olması "kabul edilebilir"
- .60-.70 arasında olması "kuşku"lu"
- .50-.60 arasında olması "zayıf"
- < .50 olması ise "kabul edilemez" olarak değerlendirilir.

Cronbach alpha değerinin, psikolojik kavramsal yapıları ortaya çıkarmayı amaçlayan ölçeklerde, yetenek ve beceri ölçen testlerde en az .70 olması gerektiği belirtilirken, sosyal bilimlerde de alt düzey .70 olarak kabul edilir. Öte yandan ölçek geliştirmeye yönelik olarak yapılan pilot araştırmalar için alpha değerinin .60, temel araştırmalar için .80, uygulamalı araştırmalar için .90-.95 olması gerektiği belirtilir. Alpha değerinin .95'in üzerinde olması tesadüfi bir tutarlılığı göstermiş olabileceği için istendik bir durum değildir (Şencan 2005).

Geliştirilen bir ölçeğin iç tutarlılık güvenirliliğini sınamanın bir başka yolu da korelasyon analizidir. KTT kapsamında güvenirliliği saptamak için değişik korelasyon analizleri yapılır: Maddeler arasındaki korelasyon, toplam puanla maddeler arasındaki korelasyon, toplam puanlar arasındaki korelasyon, gözlemci değerlendirme puanları arasındaki korelasyon (Şencan 2005).

Bu makalede, geliştirilen bir ölçeğin, iç tutarlılık güvenirliliğini sınamak için "maddeler arası korelasyon" ve "toplam puanla-maddeler arası

Tablo 2: Kaiser-Meyer-Olkin Test Sonuçlarının ve Korelasyon Katsayılarının Yorumu

Değer	Güvenirlilik
KMO < .50 olması	<i>Yetersiz</i>
KMO = .50 - .60 arasında olması	<i>Zayıf</i>
KMO = .60 - .70 arasında olması	<i>Kötü</i>
KMO = .70 - .80 arasında olması	<i>Orta</i>
KMO = .80 - .90 arasında olması	<i>İyi</i>
KMO > .90 olması	<i>Mükemmel</i>
Değer	Güvenirlilik
r > .80	<i>Yüksek</i>
r = .60 - .80	<i>Güçlü ilişki</i>
r = .40 - .59	<i>Orta derecede ilişki</i>
r = .20 - .39	<i>Düşük ilişki</i>
r < .20	<i>Zayıf ilişki (Tesadüfi olabilir)</i>

Şencan, H. (2005). Sosyal ve Davranışsal Ölçümlerde Güvenirlilik ve Geçerlilik, Birinci Baskı, Seçkin Yayıncılık Sanayi ve Ticaret AŞ., Ankara, 499-559.

korelasyon” analizlerinden söz edilmiştir. Bu analizler sonucunda elde edilen korelasyon katsayıları doğrultusunda değişkenler arasındaki ilişkinin gücü belirlenir (Şencan 2005) (Tablo 2).

Maddeler arasındaki ilişkilerin güçlü veya zayıf olup olmadığını görmek, maddeler arasındaki tutarlılığı belirlemek veya maddelerin arka planındaki gizli değişkeni ortaya çıkarmak için maddeler arası korelasyon analizi yapılır. Maddeler arası korelasyon analizi sonucunda, iki değişken arasındaki ilişki negatif değerli çıkmışsa, bu maddeler arasında ters bir ilişki olduğunu gösterir ve bu maddelerden biri veya duruma göre her ikisi de ölçekten çıkarılabilir. Bunun için her bir maddenin diğer maddelerle olan ilişkisine bakılarak karar verilir. Bu konuda diğer bir yöntem maddelerin toplam puanla olan korelasyonuna bakmaktır. Ancak madde-toplam puan korelasyonunun yapılabilmesi için 100 ile 200 arasında ya da madde sayısının en az beş katı kadar cevaplayıcının olması öngörülmektedir (Şencan 2005). Madde-toplam korelasyon katsayısının negatif ve düşük olmaması gerekir (Tavşancıl 2006). Madde-toplam puan

korelasyon katsayısının .30 un altında olması, bu maddelerde ciddi bir sorun olduğunu gösterir (Şencan 2005). Ayrıca, madde-toplam korelasyonunun düşük olması güvenirliliği düşürdüğü için, bu maddelerin ölçekten çıkarılması gerekir (Tavşancıl 2006). Ancak, bu korelasyon katsayı sınırı örneklem büyüklüğüne bağlı olarak değişebilmektedir; dört yüz ve daha fazla katılımcının bulunduğu büyük örneklemelerde .20 gibi daha düşük bir korelasyon katsayıları da kabul edilebilir (Şencan 2005, Tavşancıl 2006). Bazı araştırmacılar tarafından da, madde seçiminde kabul edilebilir korelasyon katsayısının .25’den büyük olması önerilmektedir (Tezbaşaran 1997, Çimen ve diğ. 2005).

Ölçekteki maddelerin aynı niteliği daha iyi bir şekilde ölçüp ölçmediği ve maddelerin kavramsal yapıya yaptıkları katkı konusunda fikir sahibi olabilmek için geliştirilen ölçeğin düzeltilmiş madde-toplam puan korelasyon katsayıları (madde-kalan korelasyon katsayısı) hesaplanır (Şencan 2005, Tavşancıl 2006). Düzeltilmiş madde-toplam puan korelasyon katsayısı, toplam puandan söz konusu maddenin puanı çıkarılarak, kalan top-

lam ile madde arasındaki korelasyon hesaplanarak elde edilir. Bu korelasyon katsayısının da negatif olmaması ve .20'den hatta .25'den küçük olmaması gerekir (Tavşancıl 2006). Şencan (2005) ise, bu katsayının .30'un altında olmaması gerektiğini belirtir. Ayrıca, madde ölçekten çıkarıldığında alfa katsayısı yükseliyorsa, o maddenin güvenilirliği azaltıldığına ve ölçekten çıkarılması gerektiğine karar verilir. Böylece ölçeğin homojenliği artırılmış ve dolayısıyla güvenilirliği yükseltilmiş olur (Tavşancıl 2006).

Düzeltilmiş madde-toplam puan korelasyon katsayısı .40'ın üzerinde olan maddeler ayırtehdicilik özelliği "çok iyi", .21 ile .40 arasında olan maddeler ayırtehdicilik özelliği "iyi" ve .20'nin altında olan maddeler ayırtehdicilik özelliği "kötü" olarak değerlendirilir. Ayırtehdicilik özelliği kötü olan maddeler ölçeğin güvenilirliğini düşürdüğünden, bu maddeler ölçeğin son haline alınmamalıdır (Şencan 2005).

2. Test-Tekrar Test Güvenirliđi: Bu güvenilirlik analizi ile hazırlanan ölçeğin, uygulamadan uygulamaya tutarlı sonuçlar verebilme, zamana göre deđişmezlik gösterebilme gücü incelenir. Bunun için, hazırlanan ölçek ilk uygulamadan bir süre sonra, araştırma kapsamına alınan bireylerden istatistiksel yöntemlerle (tabaka, sistematik örnekleme yöntemi gibi) seçilen, belirli sayıda kişiye tekrar uygulanır. İki ölçüm arasında geçen sürenin ne olması gerektiđi konusunda bir kesinlik yoktur. Aksayan ve diđ. (2002) tarafından, bir aracın zamana karşı deđişmezliğini kestirmek için kullanılan test-tekrar test yönteminde, ölçümlerarası en uzun aralıđın dört-sekiz haftadan fazla olmaması önerilir. Tavşancıl tarafından ise; iki ölçüm arasındaki sürenin ölçülen özelliđe göre deđişmekle birlikte, genellikle iki-üç ile dört-altı hafta arasında bir sürenin yeterli olabileceđi ifade edilir (Tavşancıl 2006). Bununla birlikte, test-tekrar test yönteminde, uygulamalar arasında geçen sürenin uzunluđu test alan kişilerin deđişme olasılıđını artırarak, güvenilirlik katsayısının düşmesi-

ne ve ölçümün standart hatasının artmasına neden olacađı belirtilmektedir (Şencan 2005).

Test-tekrar test güvenilirliđi analizinde, iki test sonucu (test toplam puan ile tekrar test toplam puan) arasında istatistiksel olarak anlamlı bir fark olup olmadıđı eşleştireilmiş iki grup arasındaki farkların testi (paired sample test) ile incelenir. İstatistiksel olarak anlamlı bir fark olmadıđı belirlendikten sonra iki test sonucu arasındaki tutarlılık için korelasyona bakılarak "kararlılık/güvenirlik katsayısı" belirlenir (Karasar 1995, Tezbaşaran 1997, Kırcaali-İftar 1999, Erkuş 2003, Şencan 2005). Bu katsayının en az .80 olması gerektiđi belirtilirken, aynı zamanda bazı bilim adamları tarafından .70 güvenilirlik katsayısının da yeterli bulunduđunu ifade edilmektedir (Şencan 2005).

Hazırlanan ölçeğin her bir maddesinin test ile tekrar test puanları arasındaki uyuşmayı incelemek için ise Cohen tarafından geliştirilen Kappa istatistiđi uygulanır. Bu istatistik sonucunda, şans faktörünün etkisini ortadan kaldırması bakımından güvenilir bir istatistiksel veri olan Kappa Katsayısı elde edilir. Kappa katsayısı -1 ile +1 arasında deđişir. Sıfır deđeri tesadüfi uyuşmayı, negatif deđerler tesadüfi olmaktan daha kötü bir uyuşmayı ve +1 ise mükemmel uyuşmayı gösterir (Şencan 2005) (Tablo 3). Böylece tutum ölçeđi hazırlamada nitel ve nicel adımlar tamamlanmış olur (Şekil 1).

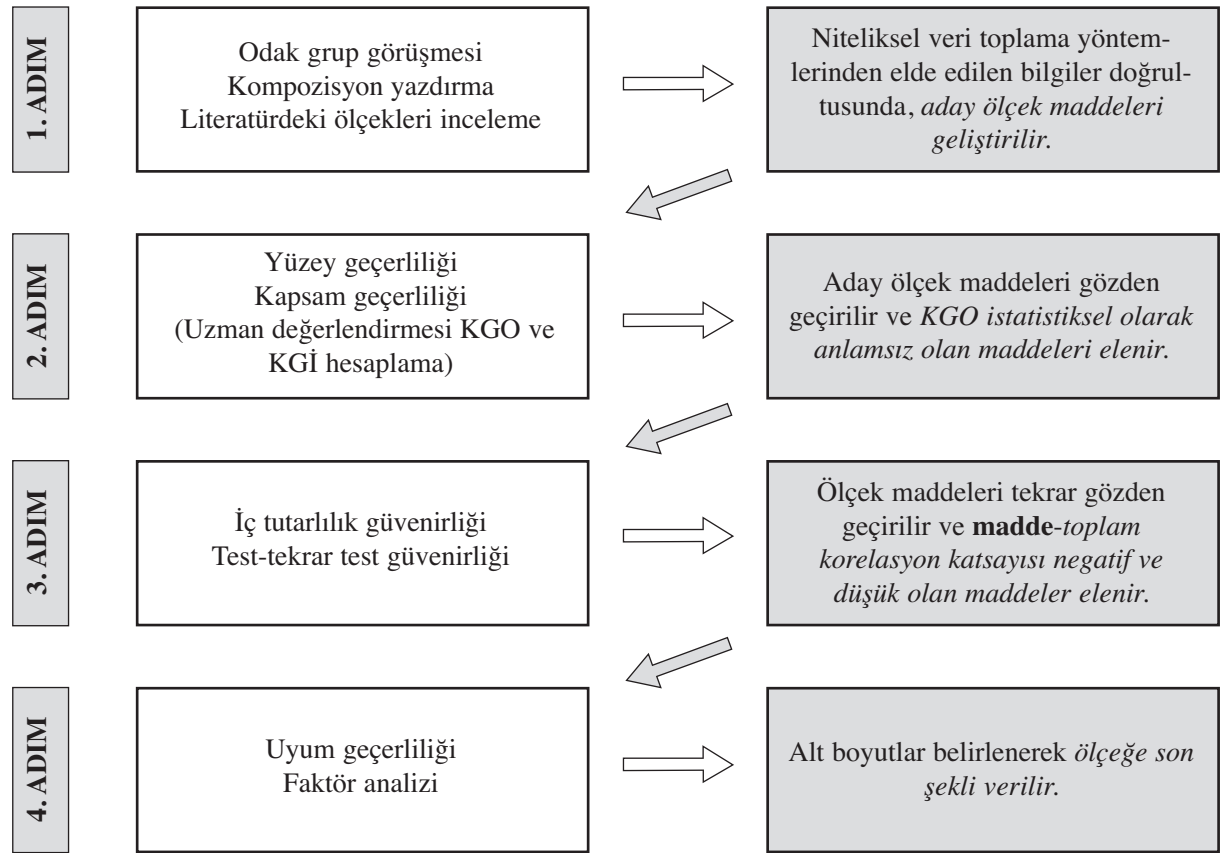
Sonuç olarak, ülkemizde hemşirelik alanında tutumu belirlemede, uyarılama çalışmaları sonucunda elde edilen ölçeklerin faktör yapılarının batı kültüründekine tam benzer olmaması ve standardize edilmiş ölçüm araçlarının hazırlanmasına ilişkin bilgi eksikliđi, çeşitli örneklemler ve ortamlarda yapılan çalışmalarda kavram karmaşaları yaşanmasına ve yorumlamada sınırlılıklara neden olmaktadır. Bu makale, hemşire araştırmacıların, tutum niteliđini dođru olarak ölçebilen, kültürel yapımıza uygun, standardize geçerli ve güvenilir ölçekler geliştirmesine katkı sağlayacaktır.

Tablo 3: Kappa Katsayılarının Yorumu

Değer	Güvenirlilik
< .20	<i>Yetersiz</i>
.21 - .40	<i>Düşük (zayıf) ilişki</i>
.41 - .60	<i>Orta derecede ilişki</i>
.61 - .80	<i>Güçlü ilişki</i>
.81 - 1.00	<i>Mükemmel ilişki</i>

Altman DG (1991). Practical Statistics For Medical Research, Chapman and Hal, London.

Şekil 1: Tutum Ölçeği Hazırlamada Nitel ve Nicel Adımlar



Kaynaklar

Akgül A (1997) *Tıbbi araştırmalarda istatistiksel analiz teknikleri*. Yükseköğretim Kurulu Matbaası, Ankara, 440-446.

Aksayan S, Bahar Z, Bayık A, Emiroğlu ON, Erefe İ, Görak G, Karataş N, Kocaman G, Kubilay

G, Seviğ Ü (2002) *Hemşirelikte araştırma ilke süreç ve yöntemleri*. Erefe İ (Ed.) (1. Baskı), Odak Ofset, İstanbul, 169-187.

Aksayan S, Gözüm S (2002) Kültürlerarası ölçek uyarlaması için rehber I: Ölçek uyarlama aşamaları ve dil uyarlaması. *Hemşirelik Araştırma Dergisi* 4,1, 9-14.

Altman DG (1991) *Practical Statistics For Medical Research*, Chapman and Hall, London.

Bilgin N (1999) *Sosyal psikolojide yöntem ve pratik çalışmalar*. Ege Üniversitesi Basımevi, İzmir, 127-142.

Çimen S, Bahar Z, Öztürk C, Bektaş M (2005) AIDS tutum ölçeğinin geçerlik ve güvenilirlik çalışması. *Zonguldak Sağlık Yüksekokulu Sağlık Eğitim Araştırma Dergisi* **1**, 1, 1-12.

Debus M (2003) *Odak grup araştırmalarında mükemmellik için el kitabı*. (Harmancı H, Çev.), Kurtiş Matbaacılık, İstanbul.

Erkuş A (2003) *Psikometri üzerine yazılar*. (1. Basım), Türk Psikologlar Derneği Yayınları, Ankara, 34-158.

Hambleton RK, Jones RW (1993) Comparison of classical test theory and item response theory and their applications to test development, educational measurement. *Issues and Practice* **12**, 3, 38-47.

Kan A (2006) Klasik test teorisine ve örtük özellikler teorisine göre kestirilen madde parametrelerinin karşılaştırılması üzerine ampirik bir çalışma. *Mersin Üniversitesi Eğitim Fakültesi Dergisi* **2**, 2, 227-235.

Karasar N (1995) *Bilimsel araştırma yöntemi* (7. Basım) Sim Matbaası, Ankara.

Kelecioğlu H (2001) Örtük özellikler teorisindeki “b” ve “a” parametreleri ile klasik test teorisindeki “p” ve “r” istatistikleri arasındaki ilişki. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi* **20**, 104-110.

Kırcaali-İftar G (1999) *Ölçme, sosyal bilimlerde araştırma yöntemleri*. (Bir AA, Ed.) Anadolu Üniversitesi Yayınları, 11-22.

Lawshe CH (1975) A quantitative approach to content validity. *Personnel Psychology* **28**, 563-575.

Şencan H (2005) *Sosyal ve davranışsal ölçümlerde güvenilirlik ve geçerlilik* (1. Basım) Seçkin Yayıncılık Sanayi ve Ticaret AŞ, Ankara, 499-559.

Tavşancıl E (2006) *Tutumların ölçülmesi ve SPSS ile veri analizi* (3. Basım) Nobel Yayın Dağıtım, Ankara, 16-156.

Tezbaşaran A (1997) *Likert tipi ölçek geliştirme kılavuzu* (2. Basım) Türk Psikologlar Derneği Yayınları, Ankara, 5-51.

Veneziano L, Hooper J (1997) A method for quantifying content validity of health-related questionnaires. *American Journal of Health Behavior* **21**, 1, 67-70.

Yıldırım A, Şimşek H (2006) *Sosyal bilimlerde nitel araştırma yöntemleri* (6. Basım) Seçkin Yayıncılık, Ankara, 99-252.

Yurdugül H, Aşkar P (2008) An investigation of the factorial structures of pupils' attitude towards technology (PATT): A Turkish sample. [Electronic version] *Elementary Education* **7**, 2, 288-309.