



## YARGICI KARARLARINA DAYALI ÖLÇEKLEME YÖNTEMLERİNİN KARŞILAŞTIRILMASI ÜZERİNE AMPİRİK BİR ÇALIŞMA

### A COMPARISON OF SCALING METHODS BASED ON JUDGE DECISIONS: AN EMPIRICAL STUDY

Adnan KAN\*

**ÖZET:** Bu çalışma ile sosyal bilimlerde kullanılan ölçekleme modellerine bir bakış açısı sağlamak ve yargıcı kararlarına dayalı iki ölçekleme yönteminin benzer sonuçlar üretip üretmediğini belirlemek amaçlanmıştır. Çalışmada veri toplama aracı olarak Bağımlılık Yapıcı Maddelere (BYM) yönelik denemelik tutum ölçeği kullanılmıştır. Çalışma, yargıcı konumunda olan 84 akademisyen ve psikolog üzerinde yürütülmüştür. Toplanan veriler üzerinde sınıflama ve sıralama kararlarına dayalı ölçekleme işlemleri uygulanmıştır. Bu ölçekleme işlemleri sonucunda elde edilen ölçek değerleri arasındaki tutarlılığı incelemek için Spearman rho (rs) korelasyon katsayısı hesaplanmıştır. Araştırma sonuçları yargıcı kararlarına dayalı iki farklı ölçekleme prosedürüne dayalı olarak elde edilen ölçek değerleri arasında 0,20 düzeyinde düşük bir korelasyon olduğunu ve bu iki ölçekleme yaklaşımının benzer sonuçlar üretmediğini göstermiştir.

**Anahtar Sözcükler:** ölçekleme, sınıflama yargılarına dayalı ölçekleme, sıralama yargılarına dayalı ölçekleme

**ABSTRACT:** This study was conducted to provide a point of view scaling methods using in social sciences and to determine the consistency between the scaling value obtained from two different scaling procedure which were based judgment decision. An attitude scale toward using addictive substances (BYM) have been used as data collection instrument. 84 academician and psychology have been used as judge in this study. A series scaling procedure have been conducted to obtain scaling values for two scaling approaches. In order to determine relationship between scaling values obtained from two different scaling procedures, the Spearman Brown rho correlation coefficient has been calculated. The study revealed that there was low relationship between scale value obtained from two different scaling procedure based on judgement decision and this findings could be used as a evidence that these two scaling procedure have not been produced similar results.

**Keywords:** scaling, scaling with classfying judgment, scaling with ranking judgment.

## 1. GİRİŞ

Davranış bilimleri alanına giren ölçme işlemleri, ölçülecek olan özellikler ve ölçme işlemi için kullanılan araçlar göz önüne alındığında fiziksel ölçmelere nazaran daha titiz çalışmayı gerektirir. Kişilik, ilgi, tutum, motivasyon, yetenek, özyeterlik (self-efficacy), kendilik algısı (self-esteem) vb. gibi birçok psikoloji ve eğitim biliminin alanına giren, doğrudan gözlenemeyen bu sebeple de fiziksel büyüklükleri bilinmeyen bir çok değişken vardır. İnsan davranışlarına ve davranışlar arasındaki neden-sonuç ilişkilerine ışık tutabilmek için psikoloji biliminin temel konuları olan bu yapıların ölçülebilir ya da gözlenebilir kılınması son derece önemlidir. Psikolojik değişkenleri ölçme yollarının bulunması, ölçme araçlarının geliştirilmesi ve ölçeklenmesi psikometri bilim dalının temel konularını teşkil etmektedir. Bu çalışma ile öncelikle, sosyal bilimlerde çok fazla kullanılmayan yargıcı kararlarına dayalı çeşitli ölçekleme tekniklerine bir bakış açısı sağlamak amaçlanmıştır.

Ölçekleme tekniklerinin geliştirilmesi, deneysel psikolojinin bir alt dalı olan psiko-fizik alanıyla başlamıştır. Deneysel psikolojinin bu özel disiplini ölçülmek istenen özelliğin algılanan büyüklükleri ile gerçek büyüklükleri arasında bir bağıntı kurmaya odaklanır (Guilford, 1954; Crocker & Algina, 1986; Dunn-Runkin, 1982). Ölçeklemede kullanılan yaklaşımlar iki grupta toplanabilir. Bunlar; 1) Denek tepkilerine dayalı yaklaşımlar ve 2) Yargıcı kararlarına dayalı yaklaşımlardır.

Yargıcı kararlarına dayalı ölçekleme yaklaşımı, uyarıcıları uzman yada bilirkişi yargılarına dayalı olarak belirli bir boyutta ölçeklemeyi içerir. Bu tür modeller gözlemcilerin her bir uyarıcının uyarıcılık derecesini belli bir yöntemle (sıralama, sınıflama, ikili karşılaştırma vb. gibi) belirlemelerine dayalıdır. Burada gözlemcilerden beklenen her bir uyarıcının ölçekleme boyutundaki yerini diğer uyarıcılarla karşılaştırarak belirlemesidir. Bunun sonucunda gözlemci yargılarının ortalama değeri

\* G.Ü. Gazi Eğitim Fakültesi, Eğitim Bilimleri Bölümü Öğretim Üyesi, adnankan@gazi.edu.tr

uyarıcının ölçek değerini belirler. Yargıcı kararlarına dayalı ölçekleme işlemleri uyarıcı merkezlidir. Uyarıcı merkezli (Stimulus centered) yaklaşımların temelinde her bir uyarıcının psikolojik ölçek üzerindeki yerini gözlemci yargılarına dayalı olarak belirlemek yer alır. Psiko-fizikçilerin, psikometri bilimine en büyük katkısı yargı yaklaşımı olarak adlandırılan bu yöntemi geliştirmeleridir (Stevens, 1966).

Bu çalışmada, yukarıda bahsedilen ölçekleme yaklaşımlarından sınıflama ve sıralama yargılarına dayalı ölçekleme yöntemleri ve bu ölçekleme yöntemlerinden elde edilen sonuçların tutarlılığının karşılaştırılması üzerinde durulmuştur. Sınıflama yargılarıyla ölçekleme yaklaşımı, uyarıcıların ardışık aralıklarla sınıflandığı durumlarda, aralık sınırlarıyla uyarıcıların ölçek değerleri arasındaki ilişkileri belirlemeye dönük istatistiksel bir modele dayalıdır. Sınıflama kararlarına dayalı yargıların toplanması için öncelikle gözlemciler, K tane uyarıcının tümü verilir ve her bir uyarıcının önceden tanımlanmış sıralı sınıflardan hangisine düştüğünü belirtmeleri istenir. Daha sonra elde edilen gözlemci yargılarına dayalı olarak uyarıcıların ölçek değerleri belirlenir.

Sıralama yargılarına dayalı yargıların toplanması için öncelikle gözlemciler, K tane uyarıcının (ölçek maddeleri) tümü verilir. Yargıcıların uyarıcıların tümünü birden düşündüğü ve her bir uyarıcıyı tüm uyarıcılarla karşılaştırarak bir sıra numarası verdiği temel sayıltısı altında, her bir uyarıcıya bir sıra sayısı tayin etmesi istenir. Bu şekilde uyarıcılar grubunun tümü, her bir uyarıcının karşılaştırıldığı birleşik bir standarda dönüştürülür ve uyarıcılara verilen sıralama yargılarından elde edilen oranlar bu bileşik standartla karşılaştırılarak ölçek değerleri elde edilir (Turgut ve Baykul, 1992; Guilford, 1954).

Sosyal bilimlerde özellikle psikolojik özelliklerin ölçülmesinde kullanılacak ölçeklerin geliştirilmesi aşamasında kullanılabilecek birçok yöntem geliştirilmiştir. Bu yöntemler temelde iki kaynağa dayalıdır. Bunlar yargıcı kararları veya denek tepkileridir. Her iki yöntemde, sosyal bilim alanındaki araştırmalarda ölçme aracı geliştirmek için başvurulan yöntemlerdir. Bu çalışmada yargıcı kararlarına dayalı farklı ölçekleme tekniklerinin benzer sonuçlar üretip üretmediğini belirlemek amaçlanmıştır. Bu çalışma ile sınıflama ve sıralama yargılarına dayalı geliştirilebilecek bir ölçeğin maddelerinin uyuşup uyuşmadığına veya ne derece uyuştuğuna ilişkin bilgi edinilmeye çalışılmıştır. Bu amaçla, çalışmada bu türden sorulara yanıt bulabilmek amacıyla aşağıdaki problemlere yanıt aranmıştır.

1. Sınıflama yargılarına dayalı ölçekleme yönteminden elde edilen madde ölçek değerleri nasıldır?
2. Sıralama yargılarına dayalı ölçekleme yönteminden elde edilen madde ölçek değerleri nasıldır?
3. Sınıflama yargılarına dayalı olarak elde edilen ölçek değerleri ile sıralama yargılarına dayalı olarak elde edilen ölçek değerleri arasında nasıl bir ilişki vardır?

### 1.1. Çalışmanın Amacı ve Önemi

Ölçekleme, ölçme sürecinde nitel ayrımları gösteren gözlemlerden, nicel ayrımları gösteren ölçülere geçişte çok önemli bir halkayı teşkil eden bir çalışma alanıdır (Anıl ve Güler, 2006). Fakat ülkemizde ve yurt dışında ölçekleme alanına dönük çalışmalar son derece kısıtlıdır. Tezbaşaran (2004) yaptığı bir çalışmada konuyla ilgili çeşitli veri tabanlarını taramış ve bu konudaki çalışmaların çok sınırlı olduğunu rapor etmiştir. Bu araştırmanın bu alanda sınırlı olan çalışmalara katkı sağlayacağı düşünülmektedir.

Bu çalışmada ayrıca farklı yaklaşımlara göre ölçekleme yapıldığında uyarıcıların psikolojik uzaydaki yerlerinin (ölçek değerleri) farklılaşıp farklılaşmadığı araştırılmıştır. Bu yolla eğer farklı ölçekleme yaklaşımları kullanılırsa oluşturacak ölçeğin aynı maddelere sahip olup olmayacağı, aynı uyarıcıların psikolojik boyutta aynı yere sahip olup olmadığı (aynı algılanıp algılanmadığı), yargıcı kararlarını alırken kullanılan yöntemin ulaşılabilecek sonuçlar üzerinde bir etkisi olup olmadığı gibi pratik bilgiler üretmek amaçlanmıştır. Buradan üretilen bilginin ise denekler üzerinde uygulanamayan, bu sebeple yargıcı kararlarına dayanarak yapılan ölçekleme sonuçlarına ne kadar güvenilebileceğine ilişkin bir görüş oluşturacağı umulmaktadır.

## 2.YÖNTEM

### 2.1. Çalışma Grubu

Bu çalışma, yargıcı kararlarına dayalı ölçek geliştirme tekniklerinin karşılaştırılmasını amaçladığı için araştırmanın örnekleme yaşları 23 ile 60 arasında değişen daha önceden tutum ölçeği geliştirme deneyimine sahip 84 akademisyen (Prof. Dr., Doç. Dr., Yrd. Doç. Dr., Arş. Gör.) ve psikolog'tan oluşmaktadır. Yargıcı kararlarına dayalı ölçek geliştirme uygulamasına dayalı örnekleme girecek bilirkişi (yargıcı) sayısı araştırmacının ulaşabildiği, çeşitli üniversitelerin ölçme ve değerlendirme (14), psikoloji (8), psikolojik danışma ve rehberlik (18), Eğitim yönetimi ve denetimi (6), Eğitim programları ve öğretim (10) bölümlerinde akademisyen olarak görev yapan ve ölçek geliştirme ve tutumların ölçülmesi konusuna hakim öğretim elemanları (56 adet) ile psikoloji bölümünden mezun, ölçek geliştirme ve tutumların ölçülmesi konusuna hakim psikologlarla (28 adet) sınırlıdır. Örnekleme de yer alan bireylerin tamamı eğitim süreci içinde ölçme değerlendirme, ölçek geliştirme, ölçme aracı geliştirme, istatistik dersleri almıştır ve çeşitli konularda ölçek geliştirme deneyimine sahiptir. Ayrıca örnekleme giren bireyler belirlenirken herhangi bir konuya yönelik en az bir tutum ölçeği geliştirmiş olanlar tercih edilmiştir. Bütün bunlar yani yargıcı konumundaki bireylerin aldıkları dersler ve ölçek geliştirmiş olmaları, bu konuda yetkin olduklarına ve yargıcı olabileceklerine ilişkin ölçü olarak kabul edilmiştir.

### 2.2. Ölçme Aracı:

Bu çalışmada veri toplama aracı olarak, araştırmacının katkılarıyla Tansel (2006) tarafından geliştirilen bağımlılık yapıcı madde kullanan bireylere yönelik tutum ölçeği kullanılmıştır. Yargıcı kararlarına dayalı ölçekleme çalışması için ölçek bilirkişi konumundaki yargıcılara verilmiş ve her bir maddenin, ölçülmek istenen özelliği tanımlanan grupta (hedef kitle) ne derece ölçtüğü, diğer bir ifadeyle bağımlılık yapıcı maddelere yönelik tutuma sahip olanla olmayanı ne derece ayırabildiğini 1 (çok az) ve 5 (çok iyi) arasında sınıflamaları istenmiştir. Sıralama yargılarına dayalı ölçekleme çalışması için aynı ölçek maddeleri bilirkişi konumundaki yargıcılara verilmiş ve yargıcılardan tüm uyarıcıları (maddeler) göz önünde tutarak ölçülmek istenen özelliği en iyi ölçenden en az ölçene doğru bir sıra numarası vermek suretiyle sıralamaları istenmiştir. Her iki ölçekleme işlemi için ayrı ayrı yöntemlere sahip farklı formlar düzenlenmiştir.

### 2.3. Verilerin Analizi:

Ölçme aracı vasıtasıyla toplanan veriler üzerinde belirlenen alt problemlere çözüm bulmak amacıyla aşağıdaki istatistiksel işlemler yürütülmüştür.

1. Sınıflama yargılarıyla ölçekleme çalışmasına ilişkin olarak, öncelikle yargıcılardan elde edilen sınıflama yargılarına ilişkin frekans ve yığılmalı frekans matrisleri oluşturulmuş daha sonrada yığılmalı frekans matrislerinden yığılmalı oranlar matrisi oluşturulmuştur. Bu aşamadan sonra her bir yığılmalı orana karşılık gelen birim normal sapmalar Excell programı aracılığıyla hesaplanarak birim normal sapmalar (Z) matrisi oluşturulmuş ve bu matris üzerinden D haliyle tam veri matrisinden ölçekleme işlemleri yürütülmüştür. Bu matrisin sütun ortalamaları alınarak sınıf sınır değerleri kestirilmiştir. Daha sonra matrisin genel ortalaması hesaplanmış ve bu ortalamadan satır ortalamaları çıkarılarak uyarıcıların (maddelerin) ölçek değerleri kestirilmiştir.
2. Sıralama yargılarına dayalı ölçekleme çalışmasına ilişkin olarak, öncelikle ölçek maddelerine ait sıra frekansları matrisi oluşturulmuş, bu tablo yardımıyla her bir uyarıcının karşılaştırılacağı bileşik standart oluşturulmuştur. Daha sonra her bir uyarıcının bu bileşik standarttan büyük bulunma oranları hesaplanmış ve bu oranlara karşılık gelen Z değerleri Excell programı aracılığıyla belirlenmiştir. Bulunan bu Z değerleri ölçek değerleri olarak alınmıştır.
3. Sınıflama yargılarına dayalı ölçeklemeden elde edilen ölçek değerlerinin güvenilirliği A.D bağıntısı yardımıyla (Torgerson, 1958; Turgut ve Baykul, 1992), sıralama yargılarına dayalı ölçek değerlerinin tutarlılığını belirlemek için ise Turgut ve Baykul (1992) tarafından önerilen varyans analizi yöntemi kullanılmıştır.
4. Her iki yöntemden elde edilen ölçek değerleri arasındaki tutarlılığı incelemek için ise Spearman-Brown sıra farkları korelasyon katsayısı hesaplanmıştır.

### 3.BULGULAR VE YORUM

Araştırmanın birinci alt problemini çözümlmek için öncelikle, gözlemciler tarafından her bir uyarıcının (madde) herhangi bir sınıfa kaç defa konulduğu sayılmak suretiyle frekans matrisi (F) oluşturulmuştur. Bu frekans matrisi ek-1'de verilmiştir. Daha sonra Hazırlanan bu F matrisinin satır elemanları toplanarak yığılmalı frekanslar ( $\Phi$ ) matrisi oluşturulmuştur. Bu matrisin son sütunundaki eleman gözlemci sayısına eşittir.  $\Phi$  matrisinin sütunlarında yer alan her bir elemanı gözlemci sayısına bölmek suretiyle oranlar matrisi (P) elde edilmiştir.  $\Phi$  ve P matrisleri ek-2 ve ek-3'te verilmiştir. Oranlar matrisinin her bir elemanına karşılık gelen birim normal sapmalar Excell paket programı vasıtasıyla hesaplanarak birim normal sapmalar matrisi (Z) oluşturulmuştur. Bu matrisi oluşturmanın diğer bir yolu da bu iş için hazırlanmış tablolardan yararlanmaktır. Bu matris oluşturulduktan sonra, her bir maddenin ölçek değeri matris üzerinden belirlenmiştir. Öncelikle Bu matrisin satır ( $Z_j$ ) ve sütun ortalamaları ( $t_g$ ) hesaplanmıştır. Sütun ortalamaları sınıfların üst sınır değerleridir. Bu sınıf sınır üst değerleri sınıf sayısına bölünmek suretiyle matrisin genel ortalaması ( $Z_{..}$ ) elde edilmiştir. Daha sonra matrisin satır ortalamaları genel ortalamadan çıkarılmak suretiyle her bir maddeye ait ölçek değeri ( $S_j$ ) elde edilmiştir. En küçük ölçek değeri başlangıç (0.00) olarak alınmış ve diğer ölçek değeri üzerine en küçük ölçek değerinin mutlak değeri eklenmek suretiyle, başlangıç noktası 0.00 olan yeni ölçek değerleri ( $S_c$ ) elde edilmiştir. Tüm bu bulgular Tablo 1'de verilmiştir Elde edilen bu ölçek değerlerinin tutarlığına ilişkin A.D değeri, 0,073 olarak bulunmuştur. Bu bağıntı yardımıyla, gözlem sonuçlarına dayalı kurulan modelin, ampirik verilere uygun olup olmadığı kontrol edilir. Modelden hareketle işlemler sondan başa doğru gidilerek elde edilen teorik verilerle gözlem sonuçları arasındaki farklar karşılaştırılır. (Turgut ve Baykul, 1992) A.D değerinin küçüklüğü teorik verilerle ampirik veriler arasındaki tutarlığın ölçüsüdür ve ölçekleme sonuçlarının tutarlığını gösterirken, katsayının büyüklüğü ise ölçek değerlerine güvenilmeyeceği anlamına gelir. Bu çalışmaya ait elde edilen katsayı oldukça düşüktür. Bu durumda, ölçekleme sonucunda elde edilen ölçek değerlerinin güvenilir olduğunu göstermektedir.

**Tablo: 1 Sınıflama Yargılarına Dayalı Ölçekleme Sonuçları**

	1	2	3	4	Top.	$Z_j$	$S_j$	$S_c$	Sıra No
1	-1,383	-0,833	-0,272	0,566	-1,922	-0,481	0,237	0,654	10
2	-1,383	-0,792	-0,531	0,060	-2,646	-0,662	0,418	0,835	6
3	-1,122	-0,751	-0,398	0,180	-2,092	-0,523	0,280	0,697	8
4	-1,668	-1,180	-0,876	-0,150	-3,874	-0,969	0,725	1,142	3
5	-1,180	-0,751	-0,180	0,792	-1,320	-0,330	0,087	0,504	13
6	-1,465	-0,792	-0,241	0,497	-2,001	-0,500	0,257	0,674	9
7	-1,383	-0,833	-0,090	1,068	-1,238	-0,310	0,066	0,483	14
8	-1,180	-0,180	0,566	1,309	0,515	0,129	-0,372	0,045	19
9	-1,559	-1,068	-0,833	-0,241	-3,701	-0,925	0,682	1,099	4
10	-1,559	-0,751	0,120	0,751	-1,439	-0,360	0,117	0,534	12
11	-1,242	-0,674	-0,150	0,833	-1,233	-0,308	0,065	0,482	15
12	-1,180	-0,497	0,241	0,921	-0,515	-0,129	-0,114	0,303	17
13	-1,981	-0,967	-0,180	0,967	-2,161	-0,540	0,297	0,714	7
14	-1,180	-0,751	-0,180	0,921	-1,190	-0,298	0,055	0,472	16
15	-1,668	-0,751	0,120	0,792	-1,509	-0,377	0,134	0,551	11
16	-0,637	-0,030	0,241	1,122	0,696	0,174	-0,417	0,000	20
17	-1,180	-0,150	0,366	0,674	-0,289	-0,072	-0,171	0,246	18
18	-1,668	-1,016	-0,366	0,150	-2,901	-0,725	0,482	0,899	5
19	-1,668	-1,309	-1,242	-0,464	-4,683	-1,171	0,928	1,345	1
20	-1,981	-1,559	-0,967	0,090	-4,417	-1,104	0,861	1,278	2
toplam	-28,267	-15,637	-4,853	10,838	-37,919				
$t_g$	-0,725	-0,401	-0,124	0,278	-0,972				
						Z.. -0,243			

Tablo 1 incelendiğinde en yüksek ölçek değerine sahip olan maddenin 19. madde, en düşük ölçek değerine sahip maddenin ise 16. madde olduğu gözlenmektedir. Bir diğer ifade ile sınıflama yargılarına göre ölçekleme yapıldığında, yargıcılar en az ayırıcılık gücüne sahip maddenin 16. madde en çok ayırıcılık gücüne sahip maddenin ise 19. maddenin olduğunu belirtmişlerdir.

Araştırmanın ikinci alt problemini çözmek için öncelikle, hangi uyarıcının yargıcılar tarafından hangi sıraya kaç kez konulduğunu gösteren sıra frekansları matrisi oluşturulmuştur. Bu frekans matrisi ek-4'de verilmiştir Bu matris üzerinden bir Uj uyarıcısı için tüm seçimlerin sayısı (Cj>cs) belirlenmiştir. Bulunan bu sayı her bir uyarıcı için toplam karşılaştırmaların sayısına (K.N) bölünerek, bir Uj uyarıcısı için bileşik standarttan büyük bulunma oranları (Pj>cs) bulunmuştur. Daha sonra bu oranlara karşılık gelen birim normal sapmalar (z değerleri) Excell paket programı vasıtasıyla belirlenerek ölçek değerleri elde edildikten sonra, en küçük ölçek değeri başlangıç (0.00) olarak alınmış ve diğer ölçek değerleri üzerine en küçük ölçek değerinin mutlak değeri eklenmek suretiyle, başlangıç noktası 0.00 olan yeni ölçek değerleri (Sc) elde edilmiştir (Turgut ve Baykul, 1992; Torgerson, 1958). Tüm bu bulgular Tablo 2 'de verilmiştir. Sıralama yargılarına dayalı ölçek değerlerinin tutarlığını belirlemek üzere varyans analizi uygulanmış ve sonuçlar Tablo 3'te verilmiştir. Ölçek değerlerinin tutarlı olup olmadığının göstergesi olarak,  $KO_{kal}=KO_{top}$  olup olmadığı kontrol edilmiştir. Eğer bu eşitlik sağlanıyorsa ölçekleme değerlerinin tutarsız olduğu ve yargıcıların uyarıcıları gelişigüzel sıraya koydukları anlamına gelmektedir.

**Tablo 2: Sıralama Yargılarına Dayalı Ölçekleme Sonuçları**

Madde/Sıra	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	40	180	80	180	100	40	120	80	120	100	40	100	40	40	0	20	100	80	100	40
2	171	171	190	133	76	76	38	76	57	76	114	38	57	57	0	19	0	38	57	57
3	36	180	270	72	108	90	108	72	0	54	54	126	0	36	36	18	18	36	36	72
4	323	102	136	187	119	119	34	51	34	51	0	17	17	17	51	17	17	17	17	17
5	32	48	16	48	96	48	64	64	64	16	80	48	96	96	96	16	80	48	48	128
6	30	45	105	75	90	135	45	90	120	45	30	45	75	45	30	15	15	60	60	30
7	0	14	28	56	56	70	0	70	126	126	84	42	98	70	56	42	56	84	0	28
8	26	0	26	78	52	13	91	78	78	13	39	65	78	91	26	52	13	26	78	104
9	60	180	48	60	132	96	96	48	24	48	12	24	24	12	12	48	12	0	0	12
10	11	0	66	0	55	11	22	22	33	77	55	99	44	33	66	66	55	55	33	66
11	10	0	20	20	0	40	20	20	40	40	30	50	60	70	60	140	40	30	60	40
12	9	36	0	18	27	0	36	18	9	63	90	90	27	27	54	36	72	46	27	27
13	0	24	16	24	8	48	56	16	0	48	64	40	72	24	48	32	56	24	8	24
14	14	7	21	7	0	0	0	35	14	21	21	14	28	70	63	70	70	28	42	28
15	6	12	6	12	12	0	30	30	42	18	12	18	36	48	42	42	42	24	30	12
16	0	0	0	10	5	5	5	10	5	15	20	10	0	35	25	30	25	70	60	65
17	12	0	0	4	12	8	16	16	4	12	16	12	8	24	28	36	24	32	36	36
18	0	15	6	3	9	24	12	15	15	12	6	24	12	12	12	3	21	12	9	9
19	54	10	16	6	10	6	10	4	12	8	4	6	0	2	4	0	0	2	4	0
20	0	3	2	8	3	10	7	9	6	3	7	3	4	1	2	0	4	3	0	4
$\bar{f}_i$	834	1027	1052	1001	970	839	810	804	815	841	780	857	792	794	707	703	714	715	784	751
$0,5^*N$	39,5	39,5	39,5	39,5	39,5	39,5	39,5	39,5	39,5	39,5	39,5	39,5	39,5	39,5	39,5	39,5	39,5	39,5	39,5	39,5
$Q_j>cs$	794,5	987,5	1012,5	961,5	930,5	799,5	770,5	764,5	775,5	801,5	740,5	817,5	752,5	754,5	667,5	663,5	674,5	675,5	744,5	711,5
$F_j>cs$	0,503	0,625	0,6408	0,6085	0,589	0,506	0,488	0,494	0,491	0,507	0,469	0,5174	0,476	0,478	0,422	0,42	0,427	0,428	0,471	0,46
$Z_j>cs$	0,008	0,319	0,3607	0,2755	0,225	0,015	-0,03	-0,04	-0,023	0,018	-0,079	0,0436	-0,06	-0,056	-0,196	-0,2	-0,184	-0,183	-0,07	-0,12
$S_c$	0,208	0,519	0,5607	0,4755	0,425	0,215	0,169	0,16	0,177	0,218	0,121	0,2436	0,14	0,144	0,004	0,00	0,016	0,017	0,128	0,075
Sıra Nb	8	2	1	3	4	7	10	11	9	6	15	5	13	12	19	20	18	17	14	16

Tablo 2 incelendiğinde en yüksek ölçek değerine sahip maddenin 3. madde en düşük ölçek değerine sahip maddenin ise 16. madde olduğu görülebilir. Bir diğer ifadeyle, sıralama yargılarına dayalı ölçekleme yapıldığında, yargıcılar ölçülmek istenen özelliği en iyi ölçen maddenin 3. madde, en az ölçen maddenin ise 16. madde olduğunu belirtmişlerdir. Bu bulgular sınıflama yargılarına dayalı olarak belirlenen ölçek değerleriyle karşılaştırıldığında, bazı maddelerin ( 4, 11,16 ve 17. maddeler) hem sınıflama yargılarına dayalı hem de denek tepkilerine dayalı ölçek değerlerinin aynı sıralara sahip iken, bazılarının (1 ve 6. maddeler) birbirine yakın, bazılarının (18, 19, 20. maddeler ) ise çok farklı sıralara sahip olduğu gözlenmiştir. Sıralama yargılarına ve sınıflama yargılarına dayalı ölçeklemeden elde edilen ölçek değerleri (uyarıcıların algılanan büyüklükleri) arasındaki tutarlığın derecesini belirlemek ve böylece üçüncü alt problemine cevap bulabilmek için bu iki yaklaşıma göre elde edilen ölçek değerleri arasında Spearman rho korelasyon katsayısı hesaplanmış ve 0,20 olarak bulunmuştur. Bu bulguya dayanarak iki farklı bilirkişi yargılarına dayalı metodun benzer sonuçlar üretmediği, diğer

bir ifadeyle iki farklı yargıcı kararlarına dayalı olarak kestirilen ölçek değerleri büyük ölçüde benzerlik göstermediği söylenebilir.

**Tablo 3: Sıra sayılarıyla varyans analizi**

V. kaynağı	Kareler Toplamı	Serbestlik derecesi	Kareler ortalaması	F
Gözlemci	,000	78	,000	,000
Uyarıcı	36122,382	19	1901,178	170,531
Kalan	16435,36	1478	11,120	,301
Toplam	52557,742	1579	33,28	

Tablo 3'te görüldüğü gibi  $KO_{top}$  ile  $KO_{kal}$  birbirine eşit değildir. Bu iki kareler ortalaması arasındaki farkın büyüklüğü nispetinde ölçekleme sonuçlarının tutarlı olacağı gözönüne alındığında, bu çalışmada elde edilen ölçekleme sonuçlarının güvenilir olduğu söylenebilir.

#### 4. SONUÇ VE ÖNERİLER

Ölçekleme teknikleri, Psikofizik alanına ait bir disiplindir. Psikofizik fiziksel büyüklüklerin insan duyularınca nasıl algılandığını ortaya çıkarma, uyarıcıların ölçülen büyüklükleri ile algılanan büyüklükleri arasındaki bağıntıyı belirlemeyi amaçlayan bir bilim dalıdır. Ölçeklemede kullanılan yaklaşımlar temelde ikiye ayrılır. Bunlar yargı yaklaşımı ve tepki yaklaşımıdır. Tepki yaklaşımının en bilinen yöntemi ve en tipik örneği Likert tipi ölçeklerdir. Birçok uygulamada çeşitli sebeplerle (sınav güvenliği, ölçme aracının uygulanacağı popülasyonun tanımlanamaması, bulunamaması veya ulaşılmasının güç olması vb.) denek tepkilerine başvurma imkânı olmayabilir. Bu durumda uyarıcıların ilgilenilen boyuttaki ölçek değerini belirlemek için yargıcı kararlarına başvurulabilir. Fakat merak edilen konu acaba yargıcı kararlarının elde edilmiş biçimi (sınıflama ve sıralama) uyarıcıların psikofiziksel boyuttaki yerlerini etkiler mi? Bu önemli ve aydınlatılması gereken bir konudur.

Bu çalışmada, psikolojik özelliklerin bir başka deyişle uyarıcıların ölçekleme boyutundaki değerlerini belirlemek için kullanılan yargıcı kararlarına dayalı yaklaşımlardan ikisi (sınıflama ve sıralama yargılarına dayalı) üzerinde durulmuş ve bu iki yaklaşıma dayalı olarak belirlenen ölçek değerlerinin birbiriyle tutarlı sonuçlar üretip üretmediği belirlenmeye çalışılmıştır. İki farklı yaklaşıma göre (sınıflama ve sıralama) toplanan yargıcı kararları üzerinde yürütülen tutarlık analizleri elde edilen ölçekleme sonuçlarının yeterince güvenilir olduğunu göstermektedir. Çalışmanın bulguları, iki farklı yargıcı kararlarına dayalı geliştirilen ölçek değerleri arasında düşük düzeyde ( $r_s=0,20$ ) bir ilişki olduğunu gösterirken, bu ilişki istatistiksel olarak manidar değildir ( $p>0,05$ ). Bu sonuçlar ışığında, çalışma yargıcı kararlarının elde edilmiş biçiminin uyarıcıların psikofiziksel boyuttaki yerleri üzerinde etkili olduğunu göstermiştir. Bu durumda, eğer belli bir uyarıcı takımına ilişkin yargıcıların algılarına dayalı bilgi toplanmak istenirse, yargıcı kararlarının elde edilmiş biçimine dikkat edilmesi gerektiğini göstermektedir. Turgut ve Baykul (1992) nin bu iki yaklaşıma ait değerlendirmeleri dikkate alınırca, sonuçların böyle çıkmasında sıralama yargılarına ait ölçekleme yöntemlerinin sebep olduğu düşünülebilir. Nitekim Turgut ve Baykul (1992) sıralama yargılarına dayalı yöntemlerin bazı sayıtlarının karşılanmasının çok zor olduğunu belirtmişler ve sıralama sayılarına dayalı prosedürler içinde en tutarlı ölçek değerlerini veren yöntemin ikili karşılaştırmalara dayalı yöntem olabileceğini öne sürmüşlerdir. Bu çalışmada, Guilford (1954) tarafından açıklanan birleşik standartla sıralama yargılarına dayalı ölçekleme yapılmıştır. Bu bağlamda, farklı yöntemlere dayalı sıralama yargılarından elde edilecek ölçekleme sonuçları sınıflama yargılarından elde edilecek ölçekleme sonuçlarıyla daha fazla tutarlık gösterebilir.

Bu konuda çalışma yapacak araştırmacılara özellikle Turgut ve Baykul (1992) nin önerilerini de dikkate alarak sıralama yargılarına dayalı diğer ölçekleme yönteminden elde edilen ölçek değerleri ile sınıflama yargılarına dayalı elde edilecek ölçek değerlerinin tutarlığını incelemeleri önerilebilir. Bu konuda yapılan sınırlı sayıda çalışmaların çoğu denek tepkilerine dayalı ölçekleme tekniklerine dayalı madde seçme tekniklerini karşılaştırmak üzerinedir (Tezbaşaran, 2004; Bindak, 2005). Literatürde, denek tepkileri ve yargıcı kararlarına dayalı çeşitli ölçekleme yaklaşımlarını birbiriyle karşılaştıran bir makaleye rastlanmıştır. O'Neil ve Chissom (1993) üç ölçekleme modelini (ikili karşılaştırma, sıralama yargıları ve Likert tipi) birbiriyle karşılaştırmış ve yargıcı kararlarına dayalı olan ikili karşılaştırma ve

sıralama yargılarına dayalı geliştirilen ölçek maddeleri arasında, bu modellerle denek tepkilerine dayalı geliştirilen ölçek (Likert tipi) maddeleri arasındakinden daha yüksek korelasyon olduğunu bulgulamışlardır. Yargıcı kararlarına dayalı iki ölçekleme yöntemine dayalı geliştirilen ölçek maddeleri arasında düşük ilişki olduğunu belirtmişlerdir. Bu çalışmanın bulguları, O'Neil ve Chissom (1993) 'un bulgularıyla benzerlik göstermektedir.

Bu çalışmanın bulgularına dayanarak, yargıcı kararlarına dayalı ölçek geliştirirken veya herhangi bir özelliğe ilişkin ölçekleme yapılırken yargıcı kararlarının ne şekilde alındığının ölçek değerleri üzerinde etkili olduğu ve benzer sonuçlar üretmediği söylenebilir. Bu sebeple yargıcı kararlarına dayalı ölçekleme yöntemlerine dayalı ölçek geliştirecek veya ölçekleme yapacak (uyarıcı ölçek değerlerini belirleme) uygulayıcıların bu durumu göz önünde tutmaları önerilebilir. Ayrıca bu konuda çalışma yapacak araştırmacılara yargıcı kararlarına dayalı ölçekleme yaklaşımları içinde yer alan farklı yöntemleri birbirleriyle karşılaştırarak bu yöntemlerin benzer sonuçlar üretip üretmediğini araştırmaları önerilebilir. Örneğin, sıralama yargılarına dayalı ölçekleme yöntemleri birbiriyle deneysel olarak karşılaştırmaları bu çalışmanın devamı olarak önerilebilir.

#### KAYNAKÇA

- Anıl, D. ve Güler, N. (2006). İkili karşılaştırma yöntemi ile ölçekleme çalışmasına bir örnek. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*, (30), 30–36
- Bindak, R. (2005). Tutum ölçeklerinde madde seçmede kullanılan tekniklerin karşılaştırılması. *İnönü Üniversitesi Eğitim Fakültesi Dergisi*, 6(10), 11–17
- Crocker, L., & Algina, J. (1986). *Introduction to Classical & Modern Test Theory*. Florida: Holt, Rinehart and Winston Inc.
- Dunn-Runkin, (1982). *Scaling methods*. New Jersey: Lawrence Erlbaum Associates Publishers.
- Guilford, J. P. (1954). *Psychometrics methods*. New York: Mc Graw-Hill Book Co.
- O'Neil, M. R., & Chissom, B. S. (1993). *A Comparison three methods for assessing attitudes*. Paper Presented at the Annual Meeting of the Educational Research Association. New Orleans, USA.
- Stevens, S. S. (1966). *Handbook of experimental psychology*. New York: John Wiley and Sons. Inc.
- Tansel, B. (2006). *Mersin Üniversitesi öğrencilerinin bağımlılık yapıcı maddelere yönelik tutumlarının incelenmesi*. Yayınlanmamış Yüksek Lisans Tezi, Mersin Üniversitesi, Mersin.
- Tezbaşaran, A. A. (2004). Likert tipi ölçeklere madde seçmede geleneksel madde analizi tekniklerinin karşılaştırılması. *Türk Psikoloji Dergisi*, 19, 77-87.
- Torgerson, W. S. (1958). *Theory and methods of scaling*. Newyork: John Wiley & Sons Inc.
- Turgut, M. F. ve Baykul, Y. (1992). *Ölçekleme teknikleri*. Ankara: ÖSYM Yayınları.

#### EXTENDED ABSTRACT

There are several methods to measure a psychological variables like interview, observation, scale, etc. When a social scientist intended to measure psychological or educational variables such as attitudes, motivation, self-efficacy, self esteem, success, ability, etc. especially uses a scale. Because of it's economic and practical features. For this reason, scale developing procedures are very important for psychology and education. Establishing measurement methods and tools like scales etc. for psychological variables. Scaling methods are basic concept for psychometri science.

There are many unobservable variables in psychology and education. For this reason, their physical magnitude is unknown. Making these construct observable and measurable is very important issue for psychology and education in order to make clear people behaviour and cause effect between behaviours. Developing scaling procedures focuse on a functional relationship between percieved magnitude and psychical magnitude of trait. Fechner expressed this functional relationship as;

$$\text{Percieved magnitude} = f(\text{true (psychical) magnitude})$$

If making measurement without error is possible, percieved magnitude and true magnitude will be identical. Scaling procedure are based on subject centered or stimulus centered. All of the scale development procedure are especially based on this two approaches. Scaling based on expert judgement (stimulus centered) includes scaling stimulus with respect to expert judgements. There are many methods for determining the value of stimuli on psychological continuum, when we need to have experts judgement about a subject. These methods aims to rank or determine quantity of stimuli on psychological continuum. At that point a great cruosity arised about whether the same results could

be obtain or not, when different methods are used for scaling stimuli. All of these kinds of scaling methods require experts determining the degree of stimuli for every stimulus with a procedure like ranking, classifying, pair-wise comparison. In this procedure, observer (expert) determine stimulus place on scaling space by comparing with other stimulus (items). Scaling procedures based on expert judgements are stimulus-centered approaches. Scaling based on subject centered approaches focus on primarily on locating individuals at different points on a continuum. Systematic changes given for stimulus are attributed to differences between responders. These two procedure could be used when a scale developing in order to measure psychological variables. Scaling procedure based on expert judgements are especially used when the focus group are not accessible or very difficult to access. Another reason to use this procedure could be applying measurement tool on focus group are not possible because of test security (e.g selection and replacement examination).

There are several studies which compare some item selection procedures for developing scales (Tezbaşaran, 2004; Bindak, 2005). All of these comparison including scaling procedures based on subject centered scaling. There are a few study comparing scaling approaches eachother.(O'Neil & Chissom, 1993). As a results of these study, it has been found that there were higher correlation between scale values obtained stimulus-centered procedures than between between scale values obtained from subject-centered and stimulus-centered procedures.

The purpose of this study is twofold. The first is to provide a framework for stimulus-centered scaling procedure used for developing scale. The second is to compare these scaling procedure eachother with respect to scale value.

In this study attitudes scale towards using addictive substances (BYM) has been used as a data collection instrument. This instrument consist of 20 items that intended to measure towards BYM. For developing a scale based on experts judgement. All the items have been inspected by 84 experts whose age change between 23 to 60. All the experts have had experience about developing attitude scale. By that way, experts opinion about whether the items measures intended trait have been taken.

With this study, comparing two stimulus-centered scaling procedure used for developing scale or scaling stimuli for measuring psychological variables have been aimed. Also exploring whether choosing one of these two scaling procedure makes differences or not with respect to scale value. Results revealed that the spearman rank correlation coefficient between these two sets of scaling value obtained by two different stimulus-centered scaling procedure was 0,20. This findings shows that some items or stimuli have different rank or place in psychological continuum, when different scaling procedure used for scaling stimuli (items). As a results of this study, it could be said that, the methods of taking experts judgement is very important and effective factor on scaling procedure, when the stimuli are scaled according to experts judgements. For this reason, the authors should consider this condition, if they will develop a scale based on expert judgements.



## EKLER

Ek 1: Gözlemci yargılarına ait frekans (F) matrisi

U/SINIF	1	2	3	4	5
1	7	10	16	27	24
2	7	11	7	19	40
3	11	8	10	19	36
4	4	6	6	21	47
5	10	9	17	30	18
7	6	12	16	24	26
8	7	10	22	33	12
9	10	26	24	16	8
17	5	7	5	17	50
24	5	14	27	19	19
25	9	12	16	30	17
26	10	16	24	19	15
27	2	12	22	34	14
29	10	9	17	33	15
30	4	15	27	20	18
31	22	19	9	23	11
32	10	27	17	9	21
33	4	9	17	17	37
36	4	4	1	18	57
37	2	3	9	31	39

Ek 2: Gözlemci yargılarına ait yığılmış frekans ( $\Phi$ ) matrisi

U/SINIF	1	2	3	4	5
1	7	17	33	60	84
2	7	18	25	44	84
3	11	19	29	48	84
4	4	10	16	37	84
5	10	19	36	66	84
7	6	18	34	58	84
8	7	17	39	72	84
9	10	36	60	76	84
17	5	12	17	34	84
24	5	19	46	65	84
25	9	21	37	67	84
26	10	26	50	69	84
27	2	14	36	70	84
29	10	19	36	69	84
30	4	19	46	66	84
31	22	41	50	73	84
32	10	37	54	63	84
33	4	13	30	47	84
36	4	8	9	27	84
37	2	5	14	45	84

Ek 3: Gözlemci yargılarına ait yığılmış oranlar (p) matrisi

U/SINIF	1	2	3	4
1	0,083	0,202	0,393	0,714
2	0,083	0,214	0,298	0,524
3	0,131	0,226	0,345	0,571
4	0,048	0,119	0,190	0,440
5	0,119	0,226	0,429	0,786
7	0,071	0,214	0,405	0,690
8	0,083	0,202	0,464	0,857
9	0,119	0,429	0,714	0,905
17	0,060	0,143	0,202	0,405
24	0,060	0,226	0,548	0,774
25	0,107	0,250	0,440	0,798
26	0,119	0,310	0,595	0,821
27	0,024	0,167	0,429	0,833
29	0,119	0,226	0,429	0,821
30	0,048	0,226	0,548	0,786
31	0,262	0,488	0,595	0,869
32	0,119	0,440	0,643	0,750
33	0,048	0,155	0,357	0,560
36	0,048	0,095	0,107	0,321
37	0,024	0,060	0,167	0,536

Ek 4: Gözlemci yargılarına ait sıra frekansları matrisi

mad/sıra	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	sıra	Ri	ri	
1	2	9	4	9	5	2	6	3	6	5	2	5	2	2	0	1	5	4	5	2	79	1	20	
2	9	9	10	7	4	4	2	4	3	4	6	2	3	3	0	1	0	2	3	3	79	2	19	
3	2	10	15	4	6	5	6	4	0	3	3	7	0	2	2	1	1	2	2	4	79	3	18	
4	19	6	8	11	7	7	2	3	2	3	0	1	1	1	3	1	1	1	1	1	79	4	17	
5	2	3	1	3	6	3	4	4	4	1	5	3	6	6	6	1	5	3	8	5	79	5	16	
6	2	3	7	5	6	9	3	6	8	3	2	3	5	3	2	1	1	4	4	2	79	6	15	
7	0	1	2	4	4	5	0	5	9	9	6	3	7	5	4	3	4	6	0	2	79	7	14	
8	2	0	2	6	4	1	7	6	6	1	3	5	6	7	2	4	1	2	6	8	79	8	13	
9	5	15	4	5	11	8	8	4	2	4	1	2	2	1	1	4	1	0	0	1	79	9	12	
10	1	0	6	0	5	1	2	2	3	7	5	9	4	3	6	6	5	5	3	6	79	10	11	
11	1	0	2	2	0	4	2	2	4	4	3	5	6	7	6	14	4	3	6	4	79	11	10	
12	1	4	0	2	3	0	4	2	1	7	10	10	3	3	6	4	8	5	3	3	79	12	9	
13	0	3	2	3	1	6	7	2	0	6	8	5	9	3	6	4	7	3	1	3	79	13	8	
14	2	1	3	1	0	0	0	5	2	3	3	2	4	10	9	10	10	4	6	4	79	14	7	
15	1	2	1	2	2	0	5	5	7	3	2	3	6	8	7	7	7	4	5	2	79	15	6	
16	0	0	0	2	1	1	1	2	1	3	4	2	0	7	5	6	5	14	12	13	79	16	5	
17	3	0	0	1	3	2	4	4	4	1	3	4	3	2	6	7	9	6	8	9	79	17	4	
18	0	5	2	1	3	8	4	5	5	5	4	2	8	4	4	4	1	7	4	3	79	18	3	
19	27	5	8	3	5	3	5	2	6	4	2	3	0	1	2	0	0	1	2	0	79	19	2	
20	0	3	2	8	3	10	7	9	6	3	7	3	4	1	2	0	4	3	0	4	79	20	1	
madde	79	79	79	79	79	79	79	79	79	79	79	79	79	79	79	79	79	79	79	79	79	1580		