

Pazarlama Alanında Yapılan Araştırmalarda Elde Edilen Aritmetik Ortalama Ölçek Orta Noktasına Eşit Midir?

Ahmet BARDAKCI
abardakci@pamukkale.edu.tr

Duygu KOÇOĞLU
dkocoglu@pamukkale.edu.tr

Selçuk Burak HAŞILOĞLU
selcukburak@hasiloglu.com

Is The Mean Equal to Scale Neutral in Marketing Researches?

Abstract

This paper aims (1) to appraise means of measured concepts in marketing related fields in terms of number of scale points employed and (2) to investigate relationship between the number of scale points and response styles. In line with this aim papers, which were published in proceedings of national marketing congress (UPK) and journal of Pazarlama Dunyasi between 2000 and 2007, presenting the findings as means and as frequency tables have been recovered. In conjunction with the aim of our research 2443 variables from 76 articles were extracted. Findings indicate that means are not equal to neutral point of the scales that were utilized for individual studies; moreover, respondents indicate their responses above the scale neutral. Both DARS and ERS are higher at 3 response categories than 4 or 5 response categories. MPR is higher at 3 response categories than 5 response categories.

Key Words : Marketing Research, Neutral Point of Attitude Scales, Response Style.

JEL Classification Codes : M39.

Özet

Bu çalışma (1) Türkiye’de pazarlama alanında yapılan araştırmalarda elde edilen ortalamaların, kullanılan ölçeklerin nötr noktasına eşit olduğu savını sınamayı ve (2) cevap alternatifli sayısına göre cevaplama tarzında ne tür değişiklikler meydana geldiğini incelemeyi amaçlamaktadır. Bu doğrultuda 1999–2007 yılları arasında gerçekleştirilen Ulusal Pazarlama Kongreleri Kitaplarında ve 2000–2007 yılları arasında yayınlanan Pazarlama Dünyası Dergilerinde, tutum ölçeği kullanarak ölçülen değişkene ilişkin aritmetik ortalamaları sunan veya bulgularını frekans tablosu olarak ortaya koyan makaleler incelenmiştir. Araştırma amacına uygun kabul edilen 76 makedeki 2443 aritmetik ortalama araştırmanın kapsamına dâhil edilmiştir. Bulgular ortalamaların nötr noktada toplanmadığını ve cevaplayıcıların cevaplarını çoğunlukla ölçek orta noktalarından daha yüksek belirttiklerini ortaya koymaktadır. 3 cevap alternatifli ölçekte DARS ve ERS, hem 4 hem de 5 cevap alternatifli ölçekten daha yüksek düzeyde tespit edilmiştir. 5 cevap alternatifinde MPR, 3 cevap alternatifinden daha düşük düzeyde bulunmuştur.

Anahtar Sözcükler : Pazarlama Araştırmaları, Tutum Ölçeklerinde Ölçek Orta Noktası, Cevaplama Tarzı.

1. Giriş

Symonds, 1924 yılında ölçeklerdeki cevap alternatifleri sayısının 7 olması gerektiğini söylediğinden bu yana, yaklaşık 90 yıldır sosyal bilimciler tutum ölçeklerinde kullanılacak optimum cevap alternatifleri sayısını belirlemeye dönük araştırmalar gerçekleştirmektedirler. Sosyal bilimler alanında farklı disiplinlerden araştırmacıların optimum cevap alternatifleri sayısı konusundaki arayışını göstermesi açısından, Internet üzerinde sadece akademik makale taraması gerçekleştiren “google scholar” arama motorunun “*number of scale points*” taramasında 950 sonuç ve “*number of response categories*” taraması için de 2230 sonuç bulması, örnek olarak verilebilir¹. Ancak, tutum ölçeklerinde kullanılacak optimum cevap alternatifleri sayısı konusunda farklı disiplinlerden sosyal bilimcinin araştırmasına rağmen kesin bir sonuca ulaşmak mümkün olmamıştır (örn. Cox, 1980; Preston & Colman, 2000). Tutum ölçeklerinde kullanılan cevap alternatifleri sayısı ile bir güvenilirlik ölçütü olan Cronbach Alpha değeri arasındaki ilişkiyi analiz eden çalışmaların ortaya koyduğu sonuçlar arasında da farklılıklar bulunmaktadır. Bu konuda yapılan bazı çalışmalar güvenilirlik ile tercih edilen cevap alternatifleri sayısı arasında bir ilişki ortaya çıkarmazken, bazıları güvenilirliğin 7 cevap alternatifli ölçekle, bazıları 5 cevap alternatifli ölçekle ve bir kısmı da 4 cevap alternatifli ölçekle maksimuma ulaştığını ortaya çıkarmaktadır (Chang, 1994:205). Hatta maksimum bilgi elde edebilmek için yirmiden fazla cevap alternatifleri olması gerektiğini tespit eden çalışmalara dahi rastlanması olasıdır (Preston & Colman, 2000).

Ölçeklerde kullanılan cevap alternatifleri sayısının tek olması durumunda (çoğunlukla 5 veya 7) olumlu ve olumsuz uç arasındaki denge noktası orta nokta (scale midpoint) veya nötr nokta (scale neutral) olarak adlandırılmaktadır. Bu bağlamda ülkemizde pazarlama alanında çalışan akademisyenlerin gerçekleştirmiş olduğu araştırmaların önemli bir kısmında tercih edilen cevap alternatifleri sayısının tek sayı olduğu, bu nedenle ölçeklerin çoğunda orta noktanın kullanıldığı görülmektedir. Ancak bu araştırmalarda elde edilen sonuçların çoğunlukla bu nokta etrafında yoğunlaşması nedeniyle çift sayıda ölçek maddesine sahip ölçeklerin kullanılmasının daha mı uygun olduğuna dair birtakım düşüncelerin ortaya çıktığı gözlemlenmiştir². Bu paralelde elinizdeki çalışmanın amacı, 1) Türkiye’de pazarlama alanında yapılan araştırmalarda elde edilen ortalamaların, kullanılan ölçeklerin nötr noktasına eşit olduğu savını sınamayı (2) cevap alternatifleri sayısına göre cevaplama tarzında ne tür değişiklikler meydana geldiğini

¹ Ne Google Scholar’da ne de Google’da Türkçe anahtar kelimelerle “cevap alternatifleri sayısı” ve “cevap kategorisi sayısı” aramalarında herhangi bir bilgiye ulaşılamamıştır. Bu noktada ülkemizdeki araştırmacıların bu konuya eğilmediği söylenebilir. Ülkemizdeki araştırmacıların konuya eğilmemiş olmasının bir sonucu olarak da ölçeklerdeki cevap alternatifleri sayısının önemsiz olduğunu düşünen özellikle genç araştırmacıların Google Scholar’a göre 244 defa alıntı yapılan Cox’un, (1980) detaylı literatür taramasını okuyarak konuya başlamasını tavsiye edebiliriz.

² Bu tür bir düşüncenin olduğu Sakarya Üniversitesi tarafından gerçekleştirilen 12. Ulusal Pazarlama Kongresinde dile getirilmiştir.

incelemeyi amaçlamaktadır. Çalışma, Türkiye’de gerçekleştirilmiş olan pazarlama alanındaki araştırmalarda ölçek kullanımı konusunda az sayıda çalışma yapılması nedeniyle önemli görülmektedir.

2. Teorik Altyapı

Bu çalışmanın amaçları arasında optimum ölçek büyüklüğüne ilişkin bir araştırma yer almamaktadır. Ancak, özellikle ülkemizdeki genç araştırmacıların kullandıkları cevap alternatifi sayısının ne olması gerektiği konusunda çok fazla araştırma yapmadan, daha önceki araştırmalarda kullanılan cevap alternatifi sayısından hareketle bu kritik kararı verdikleri düşünülmektedir. Bu nedenle teorik çerçeve bu konudaki yazından hareketle ortaya konulmaya çalışılmıştır.

Masters (1974)’ın çalışmasının bulguları; sonuçların varyansının büyük olduğu hallerde 2 cevap kategorili bir ölçek kullanımının cevaplayıcıları ayırt edebilmek için yeterli olacağını göstermiştir. Alwin ise (1992) ölçmenin amacı tutumun yönünü ölçmek olduğunda 2 cevap alternatifli ölçeklerin yeterli olacağını ifade etmektedir. Tutumun yönü değilde, yoğunluğu ölçülmek isteniyorsa ikiden fazla cevap alternatifi kullanılmalıdır.

Tutum ölçeklerinde optimum cevap alternatifi sayısı kullanılan ölçeğin amacına ve sistematik değişkenliğine bağlıdır. Bir ölçekte optimum cevap alternatifi sayısını; ölçeğin bilgiyi aktarma kapasitesi, cevaplayıcının uyarıcıyı cevap alternatiflerine aktarma kapasitesi, uyarıcıda aktarım için mevcut olan bilginin miktarı ve araştırmacının ihtiyaç duyduğu bilgi miktarı olmak üzere dört faktör tarafından belirlenmektedir. (Cox, 1980).

Lehmann ve Hulbert (1972) cevap alternatifi sayısının artırılmasının yuvarlama hatasını azaltırken, uygulamanın maliyetlerini arttıracığını, bunun yanı sıra cevaplayıcının tükenmesine neden olabileceği için cevaplamama hatasının artabileceğini belirtmektedir ki Weathers vd. (2005) cevap alternatifi sayısındaki artışın cevaplama süresinin arttırarak cevaplayıcının işini zorlaştırdığını tespit etmiştir. Benzer bulgulara Jensen ve Tendering (2005) de ulaşmıştır.

Jacoby ve Matell (1971), ise yukarıda bahsedilen çalışmalara aykırı olarak, cevap alternatifi sayısının az olmasının cevaplayıcının ayırt etme yeteneğini yok ederek ölçeği kısıtlı hale getireceğini aksine cevap alternatifi sayısının artırılmasının cevaplayıcının ayırt etme yeteneğini aşabileceğini belirtmektedir. Bu paralelde Cox (1980), cevap alternatifi sayısının belirli bir noktayı aştığında, cevaplayıcının artan sayıda alternatif arasındaki farkı anlamakta zorlanacağını buna bağlı olarak da gerçek değer ile cevaplayıcının belirttiği değer arasında farkın ortaya çıkmağa başlayacağını ifade etmektedir. Sonuç olarak cevap alternatifi sayısındaki artış bir yandan ölçeğin bilgi sağlama yeteneğini iyileştirirken öte yandan cevaplama hatasını artıracaktır.

Tutum ölçeklerinden elde edilen bilginin güvenilirliğin 7 cevap alternatifi ile maksimize edilebileceği ilk olarak 1924 yılında Symond (Colman vd, 1997) tarafından ifade edilmiş olmasına rağmen bu konuda yapılan çalışmalarda çoğunlukla Miller (1956)'e atıf yapılmıştır. Miller'in (1956) tanımına göre "iki eşit alternatif arasında karar vermek için gerekli olan bilgi bir bit'lik bir bilgidir" ve çok sayıda alternatif arasından seçim yapabilmek için gerekli olan bilginin miktarı 2^{bit} şeklinde ifade edilebilir. Miller (1956), deneysel çalışmaların sonuçlarından hareketle tek boyutlu ölçekler için deneklerden alınan bilgi miktarının ortalamasını 2,6 bit, ve varyansını 0,36 bit hesaplamış ve buradan hareketle insan beyninin 7 ± 2 maddeyi anlamlı bir şekilde ayırt edebileceği sonucuna ulaşmıştır. Gren (1970) ve daha sonra da Peterson (1994), 7 cevap alternatifinin kullanılmasının uygun olacağına dair bulgular ortaya çıkarmıştır. Güvenilirlikten hareket eden Birkett (1986), ve Preston ve Colman (2000) cevap alternatifi sayısının 7'nin ötesinde belirlenmesi halinde güvenilirlikteki artışın oldukça küçük olacağı tespit etmiştir. Ancak bu çalışmaların aksine Bendig (Cox, 1980: 414'den alınmıştır) güvenilirliğin 3 ve 11 kategori arasında değişmediğini bulmuş ve 11'de karar kılmıştır. Aynı doğrultuda Scherpenzeel ve Saris (1995; Scherpenzeel, 2002'den alınmıştır) on ülkede yapılan bir tatmin çalışmasında 11 alternatifli olan ölçeklerin en geçerli ve en güvenilir ölçek olduğunu tespit etmiştir.

Ülkemiz dışında pazarlama alanındaki araştırmalarda tutum ölçekleri için çoğunlukla 5 veya 7 cevap alternatifi kullanıldığı gözlenmektedir. Tek sayıda cevap alternatifi kullanmanın sonuçlarından birisi de cevaplayıcıya nötr bir orta nokta sunmaktır. Ancak bu nötr nokta "farksızlık" mı yoksa "farkında olmamayı" mı ifade etmelidir? Likert türü tutum ölçeklerinde nötr nokta teknik olarak "ne katılıyorum, ne katılmıyorum" şeklinde ifade edilmelidir. Ancak ülkemizde gerçekleştirilen ve yayınlanan birtakım çalışmalarda bu noktanın "kararsızım", "fikrim yok" veya "bilmiyorum" şeklinde adlandırılarak analizlere dâhil edildiği görülmüştür. "Ne katılıyorum, ne katılmıyorum" şeklinde ifade edilen bir orta nokta farksızlık belirtirken, "bilmiyorum" veya "fikrim yok" vb. şeklinde ifade edilen bir nötr nokta "farkında olmamayı" ifade etmektedir.

Tutum ölçeklerinde nötr noktanın kullanımı cevaplayıcıyı tembelliğe iterek, bu konuda net bir tutumu olmayan cevaplayıcıların veya bu konunun farkında olmayan veya bilgi sahibi olmayan cevaplayıcıların nötr noktayı belirtmesi ile sonuçlanmaktadır (örn. Hughes, 1969). Belki bu tespitin de etkisiyle, tutum ölçeklerinin aritmetik ortalamaları nötr nokta etrafında yoğunlaşmaktadır. Ancak nötr noktanın fazlaca işaretlemesinin bir sonucu olarak, küçük bir varyans değeri tespit edilmektedir. Çalışmamızın amaçlarından birisi pazarlama alanındaki araştırmalarda elde edilen ortalamaların, ölçeğin nötr noktası etrafında yoğunlaşıp yoğunlaşmadığını tespit etmek olduğundan, birinci hipotezimiz;

H1: Bu araştırma kapsamında incelenen çalışmalarda, elde edilen ortalamalar ilgili araştırmalarda kullanılan ölçeklerin nötr noktasına eşittir,

şeklinde oluşturulmuştur. Ancak H1 verilerin değişkenliğini elimine edecek bir süreci kapsamamaktadır. Verilerin değişkenliğini elimine edebilmek amacıyla her bir ortalamanın kendi varyansı kullanılarak standardize edilmesi daha uygun olacaktır. Bu doğrultuda, H1'den hareketle ortaya atılabilecek bir başka hipotez

H2: Bu araştırma kapsamında incelenen çalışmalarda ölçekten bağımsız olarak, sürekli aynı standardize edilmiş skor elde edilmiştir

şeklinde oluşturulmuştur. Aritmetik ortalamanın nötr nokta etrafında ortaya çıkmasının sadece bu noktanın fazlaca tercih edilmesi yanında, uç noktaların simetrik şekilde seçilmesi sonucunda da ortaya çıkabileceği unutulmamalıdır. Bu doğrultuda ölçek noktalarının cevaplayıcılar tarafından simetrik olarak tercih edilip edilmediğinin ele alınması gerekir ki bu nokta araştırma sorusu olarak bırakılmıştır.

Cevaplama tarzı; sorunun içeriği ile doğrudan ilişkili olmayıp cevaplayıcının genel davranışsal özelliklerinden kaynaklanan sistematik cevaplama şekli olarak tanımlanmaktadır (Bachman ve O'Malley, 1984). Bu doğrultuda cevaplama tarzındaki farklılıkların sayısallaştırılması adına ilgili literatürde tanımlanan ölçütlerin bir kısmı Tablo 1'de sunulmuştur. Tablodaki ARS, DARS, ERS ve NARS ölçütleri Bachman ve O'Malley (1984), MRP ölçütünü ise Chen vd. (1995) geliştirilmiştir. Orta nokta kullanımını ifade eden MRP'nin yüksek olması cevaplayıcıların çoğunlukla orta noktayı tercih ettiklerinin bir göstergesi olarak düşünülmesi gerekirken, NARS'nin sifıra yakın olması uç noktaların simetrik bir tercih edildiğinin bir göstergesi olarak düşünülmesi gerekmektedir. Cevaplama tarzından hareketle ve H1 doğrultusunda

H3: Ülkemizde yapılan pazarlama alanında gerçekleştirilen araştırmalarda, MRP oldukça yüksektir.

şeklinde oluşturulmuştur.

Tablo: 1
Literatürde Tanımlanan Cevaplama Tarzı Ölçütleri

Ölçüt	Tanım	Kaynak
ARS (Acquiescence Response Style)	Tüm cevaplar içinde "tamamen katılıyorum" cevabının oranıdır.	Bachman ve O'Malley (1984)
DARS (Disacquiescence Response Style)	Tüm cevaplar içinde "hiç katılmıyorum" cevabının oranıdır.	Bachman ve O'Malley (1984)
ERS (Extreme Response Style)	"Tamamen katılıyorum" ve "hiç katılmıyorum" cevabının tüm cevaplara oranıdır.	Bachman ve O'Malley (1984)
NARS (Net Agreement Response Style)	Tamamen katılıyorum" cevabının oranı ile hiç katılmıyorum" cevabının oranı arasındaki farktır.	Bachman ve O'Malley (1984)
MRP (Midpoint Responding)	Tüm cevaplar içinde kullanılan orta noktaların oranıdır.	Chen, Lee ve Stevenson (1995)

3. Metodoloji, Veri ve Bulgular

Araştırmada kullanılan veriler 2000–2007 yılları arasında Pazarlama Dünyası Dergisi'nde yayınlanan makaleler ve 1999–2007 yılları arasında düzenlenen Ulusal Pazarlama Kongreleri Kitaplarında yayınlanmış tebliğlerden elde edilmiştir. Araştırmanın amacına uygun olarak tutum ölçeği kullanan ve ölçülen değişkene ilişkin aritmetik ortalama sunan veya bulgularını frekans tablosu olarak sunan toplam 76 makale tespit edilmiş ve bunlardan temin edilen 2443 aritmetik ortalamaların tamamı bu araştırmanın verisini oluşturmuştur.

Araştırmada kullanılan 76 çalışma verilerini farklı boyutlardaki ölçekleri toplamışlardır, bunlarda kullanılan ölçek boyutlarına ilişkin özet Tablo 2'de görülmektedir. Bazı makalelerde veriler birden fazla boyutta ölçek kullanılarak toplanmış olduğu için Tablo 2'de incelenen toplam makale sayısı 76 yerine, 102 olarak görülmektedir.

Tablo: 2
Araştırma Kapsamındaki Makalelerin Ölçekteki Cevap Alternatifi Sayısına Göre Dağılımları

Ölçekteki cevap alternatifi sayısı	Makale sayısı	İlgili çalışmaların örnek büyüklüğü
3	9	2573
4	8	2755
5	76	24107
7	9	2379
Toplam	102	31814

Hipotezlerimizi test edebilmek için öncelikle bütün ölçeklerde ölçek boyutuna bağlı olarak 3, 4, 5 veya 7 olumlu ucu ve bütün çalışmalarda 1 olumsuz ucu temsil etmek üzere, Bardakci ve Whitelock (2000) tarafından kullanılan ve Şekil 1'de örnek olarak 5 cevap alternatifine ilişkin dönüşümün gösterildiği yöntemle, tüm ortalamalar yeniden düzenlenmiştir.

Şekil: 1
1–5 Ölçeğindeki Bir Ortalamanın 5–1 Ölçeğindeki Karşılığı ve Dönüşüm

A tamamen katılıyorum	1	2	3	4	5	hiç katılmıyorum
B hiç katılmıyorum	5	4	3	2	1	tamamen katılıyorum
	6	6	6	6	6	

burada A ve B ölçek büyüklüğünü göstermek üzere örneğin B'den alınan bir cevabın (X olsun) A'daki karşılığı ((B+1)-X) olacaktır.

Araştırma kapsamına dâhil edilen 76 makaledeki ortalamaların uç noktaları standart hale getirildikten sonra, her bir ölçek büyüklüğü için hesaplanan aritmetik ortalamalar ve %95 güven aralıkları Tablo 3’de sunulmuştur. Ölçekte kullanılan cevap alternatifi sayısındaki artışa paralel olarak, ortalamaların ortalaması da Tablo 3’den de takip edilebileceği gibi artmaktadır. Güven aralığı sütununda görülebileceği gibi aslında bütün ölçeklerde ortalamaların aritmetik ortalaması, ölçek nötr noktasından büyüktür. Bu sonuç Tablo 4’de sunulan her bir ölçek büyüklüğünde orta nokta ile ortalamaların ortalamasının karşılaştırıldığı t testi sonuçlarından da açıkça görülmektedir.

Tablo: 3
Farklı Ölçek Büyüklüklerinde Ortalamaların Tanımlayıcı İstatistikleri

Cevap Alternatifi Sayısı	N	Ortalamaların Arit. Ort.	Ortalamaların Std. Sap	Ortalamanın % 95 G.A.		Min	Max
				Alt	Üst		
3	140	2,1629	,44762	2,0881	2,2377	1,09	2,87
4	117	2,6590	,79424	2,5136	2,8045	1,38	3,96
5	1745	3,4690	,75253	3,4337	3,5044	1,00	4,96
7	441	4,7985	1,16651	4,6894	4,9077	1,70	6,92
Genel	2443	3,5954	1,05993	3,5533	3,6374	1,00	6,92

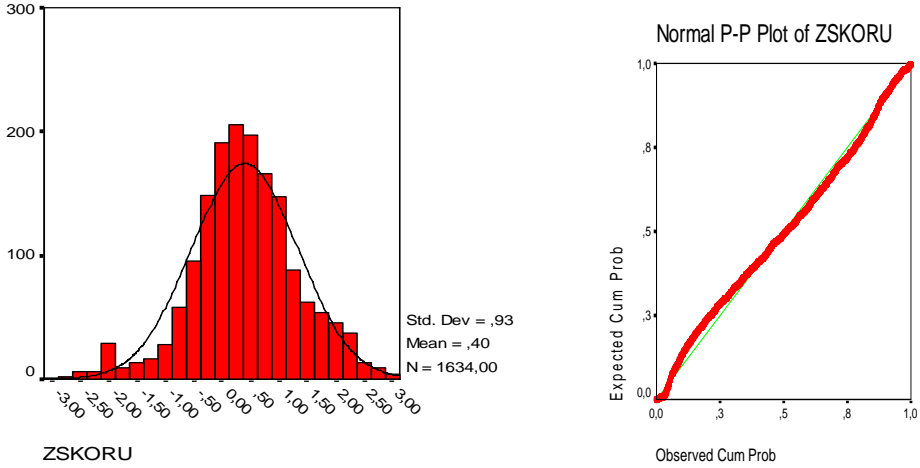
Tablo: 4
Ölçek Nötr Noktası İle Ortalamaların Ortalamasının Eşitliğini Sınayan Hipotez Testi Sonuçları

Cevap alternatifi sayısı	Test Değeri	t	df	p-Değeri
3	2	4,305	139	,000
4	2,5	2,166	116	,032
5	3	26,036	1744	,000
7	4	14,375	440	,000

Dikkat edilirse buraya kadarki analizlerde sadece aritmetik ortalamalar kullanılmış olup dolayısıyla verilerin değişkenliği dikkate alınmamıştır. Verilerin değişkenliğini dikkate alabilmek amacıyla aritmetik ortalama yerine standardize edilmiş skorlarla analiz yapmak gereği açıktır. Bu doğrultuda aritmetik ortalama yerine her bir ortalama kendi varyansı kullanılarak standardize edilmiştir³. Standardize edilen skorların büyük ölçüde ± 3 aralığında bulunması gerektiğinden ± 3 aralığı dışında kalan skorlar analiz dışında tutulmuş bu sayede 1634 standart skor elde edilmiştir. Şekil 3’ten görüleceği gibi standardize edilen skorlar (Z) oldukça yakın olsa da simetrik bir dağılım göstermemektedir.

³ Ortalamaları standardize edebilmek için, her bir ortalamanın bir ana küleden alındığı varsayılmıştır. Bilinmeyen ana kütle ortalaması yerine de, ana kütlelerin verilen ölçek üzerinde normal dağıldığını varsayarak ölçeğin orta noktasını kullanılmıştır.

Şekil: 2
Standardize Edilmiş Skorların Dağılımı ve P-P Plot



Z skorlarına ilişkin tanımlayıcı istatistikler Tablo 5’de ve standardize edilen skorlar için orta nokta 0’a eşitliğin sınındığı test sonuçları Tablo 6’da sunulmuştur. Tablo 5’ten görülebileceği gibi cevap alternatifleri sayısı 5 ve 7 olduğunda, ortalamaların ölçeğin nötr noktasına eşit olduğu iddiasını destekleyecek kanıt bulunamamıştır. Bunun yanında, cevap alternatifleri sayısı 3 ve 4 olduğunda sonuçların nötr noktaya eşit olduğu iddiasını reddedecek kanıt da bulunamamıştır. Buradaki bulgularımızdan hareketle ölçekteki cevap alternatifleri sayısını 4’ün üzerine taşımamızın elde edilen sonuçları nötr noktadan uzaklaştırdığını söylemek mümkün olabilir.

Tablo: 5
Z Skorlarına İlişkin Tanımlayıcı İstatistikler

Cevap alternatifleri sayısı	n	Z arit. Ort.	Std. Sap.	Z % 95 Güven Aralığı		Min	Max
				Alt	Üst		
3	44	-,1742	1,10143	-,5090	,1607	-2,57	2,81
4	112	,2371	1,53678	-,0507	,5248	-2,99	2,73
5	1160	,4423	,84036	,3939	,4907	-2,55	2,98
7	318	,3845	,92337	,2826	,4863	-2,65	2,76
Genel	1634	,4004	,93340	,3551	,4457	-2,99	2,98

⁴ Cevaplayıcıya 4 cevap alternatifleri sunulduğunda aslında cevaplayıcının seçebileceği orta nokta olmayacaktır. 1, 2 veya 3, 4 seçeneklerinin simetrik tercih edilmesi sonucunda aritmetik ortalama ölçek nötr noktası olarak belirtilen 2,5’e yakın çıkacaktır.

Tablo: 6
Z Skorlarının 0'a Eşitliğinin Sınandığı Hipotez Testleri

Cevap Alternatifi Sayısı	Test değeri = 0					
	t	df	Anlamlılık. (çift kuyruk)	Arit. Ort. farkı	Farkın % 95 güven aralığı	
					Alt	Üst
7	7,425	317	,000	,3845	,2826	,4863
5	17,925	1159	,000	,4423	,3939	,4907
4	1,633	111	,105	,2371	-,0507	,5248
3	-1,049	43	,300	-,1742	-,5090	,1607

Tüm ortalamaların standardize edilmesi H1'i ölçekte kullanılan cevap alternatifi sayısından bağımsız olarak test edilmesini mümkün hale getirmiştir. Tablo 7'den izlenebileceği gibi Türkiye'de pazarlama alanında gerçekleştirilen ve araştırmamız kapsamına alınan çalışmalarda, elde edilen ortalamalar ölçek orta noktasına eşittir, savını destekleyecek bir kanıt ortaya çıkarılamamıştır.

Tablo: 7
Tüm Çalışmalar İçin Z Skorlarının 0'a Eşitliğinin Sınaması Sonucu

	n	Arit Ort.	Std. Sap.	Ort. Std. Hata.
Tüm çalışmalar için Z	1634	0,4004	0,93340	0,02309

	Test değeri = 0					
	t	df	Anlamlılık (çift kuyruk)	Ort. Farkı	Farkın 95% güven aralığı	
					Alt	Üst
Tüm çalışmalar için Z	17,339	1633	0,000	0,4004	0,3551	0,4457

H2 ölçekten bağımsız olarak sürekli aynı ortalamanın elde edildiğini ifade etmektedir. Bu hipotezimiz test edebilmek için 1634 Z değerinin, ölçekteki cevap alternatifine bağlı olup olmadığını kontrol edebilmek üzere gerçekleştirilen ANOVA analizi farklı cevap alternatifine sahip ölçeklerden alınan Z skorlarının aynı olmadığını göstermektedir (Tablo 8).

Tablo: 8
Farklı Ölçeklerden Elde Edilen Z Skorlarının Birbiriyle Eşitliğinin Sınaması

	Kareler toplamı	df	Kareler ortalaması	F	p-değeri
Gruplararası	19,627	3	6,542	7,601	0,000
Grupiçi	1403,085	1630	0,861		
<i>Toplam</i>	<i>1422,713</i>	<i>1633</i>			

Arařtırma sorumuz (AS₁) ölçek noktalarının cevaplayıcılar tarafından simetrik olarak tercih edilip edilmediđini sorgulamaktadır. Bu soruya cevap bulabilmek için bulgularını tanımlayıcı istatistikler yerine frekans tabloları olarak özetleyen çalıřmalarda, sunulan frekans tablolarının birleřtirilmesi sonucunda elde edilen çapraz tablolar ve uyumun iyiliđi analizleri sonuçları Tablo 9'da sunulmuřtur. 3, 4 veya 5'in olumlu uçlar olduđu hatırlanarak, 4 ve 5 cevap alternatiflerine sahip ölçekler için cevaplayıcıların çođunlukla cevap seeneklerini olumlu alternatiften yana kullandıkları görülebilmektedir.

Tablo: 9
Birleřtirilen Frekans Tabloları Özeti ve Uyumun İyiliđi Testi Bulguları

Cevap Alternatifi sayısı=3	1	2	3	Toplam
N	6719	1961	3884	12564
χ^2	2735,89			
p-deđeri:	0,0000			

Cevap Alternatifi sayısı =4	1	2	3	4	Toplam
N	1272	2060	3079	3011	9422
χ^2	940,1083				
p-deđeri:	0,0000				

Cevap Alternatifi sayısı=5	1	2	3	4	5	Toplam
N	15617	26522	18951	37881	28542	127513
χ^2	11926,25					
p-deđeri:	0,0000					

Ölçekteki cevap alternatifi sayısının cevap tarzında bir fark oluřturup oluřturmadıđını görebilmek için, sonuçlarını frekans tabloları halinde sunan çalıřmalardan elde edilen bulgulardan hareketle hesaplanan ARS, DARS, ERS, NARS ve MRP oranları Tablo 10'da karřılařtırılmalı olarak sunulmuřtur.

Tablo: 10
Cevaplama Tarzına İlişkin ANOVA Testi Sonuçları

Ölçekteki cevap alternatifleri sayısı	N	Arit. Ort.	Std. Sapma.	95% güven aralığı		F	p-değeri	
				Alt sınır	Üst sınır			
ARS	3	46	0,3253	0,20795	0,2635	0,3870	7,974	0,000
	4	34	0,4020	0,26987	0,3078	0,4961		
	5	346	0,2655	0,19700	0,2447	0,2864		
	Toplam	426	0,2829	0,20811	0,2631	0,3027		
DARS	3	46	0,4748	0,25267	0,3998	0,5499	127,932	0,000
	4	34	0,1568	0,15126	0,1040	0,2096		
	5	346	0,1027	0,12821	0,0892	0,1163		
	Toplam	426	0,1472	0,18739	0,1294	0,1651		
ERS	3	46	0,8001	0,13439	0,7602	0,8400	142,229	0,000
	4	34	0,5588	0,21269	0,4846	0,6330		
	5	346	0,3683	0,16907	0,3504	0,3862		
	Toplam	426	0,4301	0,21882	0,4093	0,4510		
NARS	3	46	-0,1496	0,44285	-0,2811	-0,0180	22,229	0,000
	4	34	0,2451	0,38234	0,1117	0,3786		
	5	346	0,1628	0,28619	0,1325	0,1931		
	Toplam	426	0,1356	0,33010	0,1042	0,1671		
MPR	3	46	0,1999	0,13439	0,1600	0,2398	4,783	0,029
	5	346	0,1580	0,12044	0,1452	0,1707		
	Toplam	392	0,1629	0,12272	0,1507	0,1751		

Tablodan da görüleceği üzere;

ARS: Tutum ölçeğindeki ifadeye katılma açısından 4 cevap alternatifli ölçekte katılım 5 cevap alternatifli ölçekten yüksek düzeydedir. Ancak 3 cevap alternatifli ölçek ile 4 veya 5 cevap alternatifli ölçek arasında anlamlı bir fark yoktur.

DARS: Tutum ölçeğindeki ifadeye katılmama açısından bakıldığında 3 cevap alternatifli ölçek hem 4 cevap alternatifli hem de 5 cevap alternatifli ölçekten daha fazla düzeydedir. Ancak, 4 cevap alternatifli ölçek ile 5 cevap alternatifli ölçek arasında anlamlı bir fark yoktur.

ERS: Aşırı uç kullanımı açısından 3 cevap alternatifli ölçek doğal olarak hem 4 cevap alternatifli hem de 5 cevap alternatifli ölçekten daha yüksektir. Ayrıca ERS 4 cevap alternatifli ölçek de 5 cevap alternatifli ölçekten daha yüksek düzeydedir. Buna bağlı olarak cevap alternatifleri sayısı arttıkça aşırı uç kullanımının azaldığı söylenebilir.

MRP: 3 cevap alternatifli ölçekte orta nokta kullanımını ortalama %19,99 iken 5 cevap alternatifli ölçekte ortalama kullanım %15,80 olarak bulunmuřtur. Ancak bu fark istatistiksel olarak anlamlı deęildir. Bu oranlar H_3 'ün reddedilmesini gerektirmektedir.

NARS: Net katılım oranları da DARS ile paralellik göstermektedir. 3 cevap alternatifli ölçeklerde NARS oranı, hem 4 cevap alternatifli hem de 5 cevap alternatifli ölçekten daha düşüktür. Fakat NARS oranı açısından 4 cevap alternatifli ölçek ile 5 cevap alternatifli ölçek arasında anlamlı bir fark yoktur. 3 cevap alternatifli ölçekte verilen ifadelere net katılmama ortaya çıkmıřtır. Ancak tüm ölçek boyutları için NARS oranları cevap alternatiflerinin simetrik olarak kullanılmadığını göstermektedir.

4. Sonuç

Bu arařtırmanın verilerini 1999–2007 yılları arasında gerçekleştirilen Ulusal Pazarlama Kongreleri Kitaplarında ve 2000–2007 yılları arasında yayınlanan Pazarlama Dünyası Dergilerinde, tutum ölçeęi kullanarak ölçülen deęişkene ilişkin aritmetik ortalamaları sunan veya bulgularını frekans tablosu olarak ortaya koyan 76 makaleden elde edilen 2443 aritmetik ortalama oluřturmuřtur. İlk olarak arařtırma kapsamındaki çalışmaların bulgularının ölçek nötr noktasına eřit olduęu savı test edilmiřtir. Arařtırmamıza dahil edilen çalışmalardaki ortalamaların, nötr noktaya eřit olmadığı ortaya çıkarılarak H_1 hipotezi reddedilmiřtir.

Ancak salt ortalamaları test etmek yerine, kendi varyansları ile standardize edilmiş skorlar kullanılarak test gerçekleştirildiğinde 3 ve 4 cevap alternatifli kullanılarak gerçekleştirilen arařtırmalar için, arařtırmacıların elde ettikleri sonuçların, ölçek nötr noktasına eřit olduęu iddiasını reddedecek kanıt ortaya çıkmazken 5 ve 7 cevap alternatifli kullanılarak gerçekleştirilen arařtırmalarda sonuçların ölçek nötr noktasında eřit olmadığı bulunmuřtur. Bu sonuçlar H_2 'nin ölçeklerde kullanılan cevap alternatifleri sayısı ile ilişkilendirilmesi halinde 3 ve 4 cevap alternatifli ölçekler için reddedilemediğini fakat 5 ve 7 cevap alternatifli ölçekler için reddedildiğini göstermektedir. Eęer bizim arařtırmamıza dâhil edilen tüm çalışmalar bir bütün olarak ele alınırsa, H_2 'nin reddedilmesi gerektięi ortaya çıkmıřtır.

Arařtırma kapsamında alınan bütün ortalamaların standardize edilen skorlarının ortalamasının sıfırdan daha büyük olması, cevaplayıcıların tutumlarını çoęunlukla ölçek orta noktalarından daha yüksek olarak belirttikleri sonucunu ortaya çıkarmaktadır. Aynı bulgu, frekans tablosu řeklinde sunulan bulguların birleřtirilmesi sonucunda da ortaya çıkmıřtır.

Arařtırmamız kapsamına alınan ve sonuçlarını frekans tablosu olarak sunan çalışmaların incelenmesi sonucunda, 3 cevap alternatifine sahip ölçekte hiç katılmama oranı DARS hem 4'li hem de 5'li ölçekten fazla olduęu bulunmuřtur. Üç cevap

alternatifinde NARS oranı ile 5 cevap alternatifi NARS oranı arasında yaklaşık %30 fark ortaya çıkmıştır.

Diğer taraftan, aşırı uçların tercih edilmesi olasılığı ölçek büyüklüğüne bağlı olarak azalmaktadır. Yani, cevaplayıcıya daha fazla alternatif sunuldukça aşırı uç kullanımında bir azalma söz konusudur. Yine, aşırı uçların kullanılma olasılığına bağlı olarak verilen cevapların çeşitliliği farklılaşmaktadır. Örneğin 3'lü ölçekte tüm cevap alternatifleri kullanılırken 4'lü ölçekte alternatiflerin %92,08'si, 5'li ölçekte ise %88,24'ü kullanılmıştır. Son olarak, orta noktanın kullanımı açısından da 3'lü ve 5'li ölçekler arasında bir fark bulunamamıştır.

Gerçekleştirilen çalışmanın amaçları içerisinde ülkemizde pazarlama alanında yapılan çalışmalar için optimum cevap alternatifi sayısına ilişkin doğrudan bir analiz bulunmamaktadır. Fakat bu çalışmanın sonucunda, araştırmacının elde etmeyi umduğu bilginin, sadece tutumun yönünü ölçmek olduğunda 3 cevap alternatifi, yaklaşık %80 yön belirtmektedir. Cevap alternatifi sayısının artması halinde cevapların tutumun yönünden çok yoğunluğu ölçer bir hal aldığı ERS ölçütü sonuçlarından izlenebilmektedir. Bu doğrultuda amaç yoğunluk tespiti olduğunda -bu araştırmaya dâhil edilen çalışmalarda 5 cevap alternatifli çalışmaların yoğunluğu da unutulmamalıdır- 5 cevap alternatifi daha uygun görülmektedir.

Son olarak sosyal bilimler alanında araştırma yapmayı düşünen genç arkadaşlarımıza, likert tipi ölçeklerde nötr noktanın “ne katılıyorum, ne katılmıyorum” şeklinde farksızlık ifadesi olması gerekirken, araştırmamıza dahil edilen çalışmaların önemli bir kısmında bu noktanın “fikrim yok”/“bilmiyorum” şeklinde farkında olmamayı işaret ettiğinin tespitini hatırlatmak isteriz.

Kaynakça

- Alwin, D.F. (1992), “Information Transmission in the Survey Interview: Number of Response Categories and the Reliability of Attitude Measurement”, *Sociological Methodology*, 22, 83-118.
- Bachman, J.G. ve P.M. O'Malley (1984), “Yea-Saying, Nay-Saying, and Going to Extremes: Black-White Differences in Response Styles”, *Public Opinion Quarterly*, 48, 491-509.
- Bardakci, A. ve J. Whitelock (2000), “The Standardization Of Marketing Towards A Meta Analysis Of Empirical Studies”, içinde, Cooley, F.V.G (Ed) (2000), *AMA Marketing In A Global Economy Conference*, June 28–July 1, Buenos Aires, Argentina, Proceedings CD, s: 168–176.

- Birkett, N.J. (1986), "Selecting the Number of Response Categories for a Likert-Type Scale", *Proceedings of the American Statistical Association. Section on Survey Research Methods*, Washington, DC: American Statistical Association.
- Chang, L. (1994), "A Psychometric Evaluation of 4-Point And 6-Point Likert-Type Scales In Relation To Reliability And Validity", *Applied Psychological Measurement*, 18(3), 205–215.
- Chen, C., S.Y. Lee, H.W. Stevenson (1995), "Response Style And Cross-Cultural Comparisons Of Rating Scales Among East Asian and North American Students", *American Psychological Society*, 6(3), 170-175.
- Colman, A.M., C.E. Norris, C.C. Preston (1997), "Comparing Rating Scales of Different Lengths: Equivalence of Scores From 5-Point And 7-Point Scales", *Psychological Reports*, 80, 355-362.
- Cox, E.P.III (1980), "The Optimal Number of Response Alternatives", *Journal of Marketing Research*, 17, 407–422.
- Gren, P.E. (1970), "Rating Scales And Information Recovery –How Many Scales And Response Categories to Use", *Journal of Marketing Research*, 34, 33–39.
- Jacoby, J. ve M.S. Matell (1971), "Three-Point Likert Scales Are Good Enough", *Journal of Marketing Research*, 8, 495–500.
- Jensen, C., J. Tendering (2005), "Perceived Prominence and Scale Types", *Fonetic 2005*, Proceedings, Göteborg University, 111-114.
- Masters, J.R. (1974), "The Relationship Between Number Of Response Categories And Reliability of Likert-Type Questionnaires", *Journal of Educational Measurement*, 11(1), Spring.
- Peterson, R.A. (1994), "A Meta-Analysis of Cronbach's Coefficient Alpha", *Journal of Consumer Research*, 21, 381–391.
- Preston C.C., A.M. Colman (2000), "Optimal Number of Response Categories in Rating Scales: Reliability, Validity and Discriminating Power and Respondent Preferences", *Acta Psychologica*, 104, 1-15.
- Scherpenzeel, A. (2002), "Why use 11-point scales?" *Swiss Household Panel (SHP)* <<http://www.swisspanel.ch/file/doc/faq/11pointscs.pdf>>, 09.04.2008.
- Weathers, D., S. Sharma, R.W. Niedrich (2005), "The Impact of The Number of Scale Points, Dispositional Factors, And The Status Quo Decision Heuristic on Scale Reliability and Response Accuracy", *Journal of Business Research*, 58, 1516-1524.

