PAZARLAMA ARAŞTIRMALARINDA KULLANILAN TUTUM ÖLÇEKLERİNDEKİ CEVAP ALTERNATİFİ SAYISINA İLİŞKİN BİR LİTERATÜR TARAMASI

Ahmet BARDAKCI*

Özet

Pazarlama araştırmalarında kullanılan tutum ölçekleri genellikle 7 ± 2 kategori kullanılarak oluşturulmakta ancak bu kullanımın arkasındaki tartışmalar çoğunlukla incelenmemektedir. Bu makalede pazarlama araştırmalarında kullanılan tutum ölçeklerindeki cevap alternatiflerinin sayısına ilişkin tartışmalar ele alınmıştır. Bu sayede araştırmacıların kullandıkları ölçek büyüklüklerine ilişkin daha geniş bilgi sahibi olmaları amaçlanmaktadır.

Anahtar Kelimeler: Pazarlama Araştırmaları, Tutum Ölçeklerinde Cevap Alternatifi Sayısı

A LITERATURE REVIEW ABOUT NUMBER OF SCALE POINTS OF ATTITUDE SCALES FOR MARKETING RESEARCH

Abstract

Attitude scales used in marketing research are usually formed by using 7 ± 2 number of scale points without a clear understanding of the theoretical background. This study is concerned with the debates over the number of scale points of attitude scales used in marketing research. In this way, it is intended to provide the researchers with a greater insight into the magnitude of the scales used by them.

Key Words: Marketing Research, Number of Scale Points of Attitude Scales

1. GİRİŞ

1932 yılında Likert ölçeği olarak da bilinen ölçek[2] tanıtıldığından bu yana en uygun cevap alternatifi sayısını belirlemeye

* Doç. Dr., Pamukkale Üniversitesi, İİBF İşletme Bölümü, DENİZLİ, e-posta: abardakci@pau.edu.tr,

[2] İngilizce'de Summative Rating ölçeği olarak adlandırılan bu ölçek ülkemizde genellikle Likert ölçeği olarak bilinmektedir. Bu ölçeği Nakip (2006, 141) Likert derecelemesi olarak, Kurtuluş (2004; 285) ise Likert Toplama ölçeği olarak Türkçe'ye çevirmiştir.

Dutton & Blum (1968) summative rating'i şu şekilde tanımlamaktadır: "Likert türü ölçekler cevaplayıcının 5 alternatif arasından cevabını seçmek üzere oluşturulmuş üçüncü kişi önermelerini içermektedir. Buradaki her bir skor, rating (sıralama,

dönük çok sayıda araştırma gerçekleştirilmiştir. Ancak pazarlama araştırmalarında sıklıkla kullanılan tutum ölçeklerindeki cevap alternatifi sayısının ne kadar olması gerektiğini incelemek üzere gerçekleştirilen araştırmalar, maalesef problemi açık bir şekilde çözüme kavuşturamamıştır. Bu çalışmaların önemli bir kısmı güvenilirlik ile cevap alternatifi sayısı arasındaki ilişkiden hareketle en uygun cevap alternatifi sayısını bulmaya çalışmış, fakat farklı araştırmacılar güvenilirlikle cevap alternatifi sayısı arasında farklı sonuçlar ortaya koymuşlardır. Örneğin Chang'ın, (1994) ifadesine göre Symonds (1926) ve Nunnally (1967), 7 cevap alternatifli ölçekte güvenilirliğin maksimuma ulaştığını, Lissitz ve Gren (1975), ile Jenkins ve Teber (1977), 5 alternatifli ölçekte güvenilirliğin maksimuma ulaştığını, Bending (1954) 3 ve 4 cevap alternatifli ölçekte güvenilirliğin maksimuma ulaştığını ortaya koyan bulgulara ulaşmıştır. Bir kısım araştırmalar ise güvenilirlik dışındaki ölçütlerden hareketle en uygun cevap alternatifi sayısını tespit etmeğe çalışmış, ancak bunlardan da net sonuçlar almak mümkün olamamıştır.

2.TUTUM ÖLÇEKLERİNDEKİ CEVAP ALTERNATİFİ SAYISI

Cox (1980) en uygun sayıda cevap alternatifi konusunda yaptığı kapsamlı literatür taramasında her araştırma için uygun olacak tek tip cevap alternatifi sayısı olmadığını açık bir şekilde ortaya koymaktadır. Bu kapsamda ilgili literatürden hareketle en uygun cevap alternatifi sayısını belirlemek üzere birtakım faktörlerden söz edilebilir. İzleyen bölümde araştırmanın amacı, ilave alternatif sayısının faydası ve maliyeti, cevaplayıcının alternatifler arasındaki farkı ayırt edebilme yeteneği ve güvenilirlik alt başlıkları altında bu faktörlerin tanıtılmasına çalışılacaktır.

2.1. Araştırmacının Amacı

Ölçekteki cevap alternatifi sayısı konusunda verilecek kararı etkileyen faktörlerden birincisi araştırmacının amacıdır. Bu noktada Lehman ve Hulbert (1972), araştırmacının amacının cevaplayıcıların ortalamasını bulmak olduğunda veya yeni bir ölçek geliştirmek üzere birkaç ölçeğin toplamına ihtiyaç duyulduğunda 2 veya 3 cevap alternatifinin bu amaç için yeterli olacağını belirtmektedir. Ancak

---

dereceleme) olarak adlandırılır. Bu derecelemeler bütün skorlar için özetlendiğinden yöntem summated rating olarak bilinir."

 ISSN: 1308-2922

araştırmacı, bireysel bir davranışa odaklanıyor ise 5 veya 7 cevap alternatifi uygun görülmüştür. Alwin ise (1992) araştırmacının amacının sadece tutumun yönünü ölçmek olduğunda en uygun cevap alternatifinin iki olduğu tutumun yoğunluğunun ölçülmesi amaçlandığında ise daha fazla cevap alternatifinin bulunmasının uygun olacağı sonucuna varmıştır.

2.2. İlave Cevap Alternatiflerin Faydası ve Maliyeti

Cevap alternatifi sayısı konusundaki kararı etkileyen bir başka faktör, alternatif sayısının faydası ve maliyeti arasındaki ilişkidir. Cevap alternatifi ayısının artması cevaplama süresini arttıracaktır. Weathers vd. (2005) 3, 5, ve 7 cevap alternatiflerine sahip ölçekler için ortalama cevaplama sürelerini sırasıyla 190,20sn; 209,43sn ve 223,03sn olarak tespit etmişlerdir ve bu süreler istatistiksel olarak birbirlerinden farklıdır. Bu bulgu cevap alternatifi sayısındaki artışın, cevaplayıcının işini zorlaştırdığını göstermektedir. Jensen ve Tendering (2005) benzer şekilde daha fazla cevap alternatifi kullanmanın cevaplayıcının işini zorlaştıracak bir şekilde maliyetleri arttırdığı sonucuna ulaşmışlardır. Jensen ve Tendering (2005) ayrıca cevap alternatifi sayısını arttırmanın ayırt etme gücü açısından katlanılacak ilave maliyeti dengeleyemediğini tespit ederken[3], buna karşılık iki cevap alternatifli bir ölçeğin çoğu amaç için uygun olduğunu ve çok hızlı bir şekilde cevaplanabildiğini de tespit etmişlerdir.

Preston ve Colman (2000) 2, 3, 4, 5, 6, 7, 8, 9, 10 ve 101 cevap alternatifli ölçekleri kullanarak yaptıkları araştırmalarında 2, 3 ve 4 cevap kategorisine sahip ölçeklerin en hızlı cevap alınanlar olduğunu bulmuşlardır. Bu araştırmada cevaplayıcılar 5,7 ve 10 cevap alternatifli ölçeklerin kullanımını kolay bulurken, 10, 11 ve 101 cevap alternatifine sahip ölçeklerin duygularını daha rahat ifade edebilmelerine olanak verdiklerini belirtmişlerdir. Kullanım hızı, kullanım kolaylığı ve duyguları yansıtmaya olanak sağlama ölçütleri bir bütün olarak kullanıldığında 2, 3 ve 4 cevap alternatifine sahip ölçekler en az tercih edilen ölçeklerken, 7, 9 ve 10 cevap alternatifine sahip olanların cevaplayıcılar tarafından en fazla tercih edilen ölçekler olduğu bulunmuştur.

[3]Bu çalışmada cevap alternatifi sayısı 31e kadar yükseltilmiştir.

Lehmann ve Hulbert (1972) daha fazla cevap alternatifi kullanmanın bir fayda olarak yuvarlama hatasını[4] azalttığını ancak buna karşılık uygulama maliyetini yükselttiğini cevaplayıcının tükenmesine neden olarak cevaplamama hatasını arttırabileceğini belirtmektedir. Bu nedenle değişkenin doğru ölçümü açısından beş veya altı cevap alternatifi kullanmanın faydasının maliyetinden daha fazla olacağı belirtilmektedir.

2.3. Cevaplayıcının Cevap Alternatifleri Arasındaki Farkı Ayırt Edebilme Yeteneği

Ölçekteki cevap alternatifi sayısı konusunda verilecek kararı etkileyen belki de en önemli faktör cevaplayıcının ayırt edebilme yeteneğidir. Tuorangeau ve Rasinski (1988, Weathers, ve diğ., 2005'den alınmıştır) bir ölçeğe cevap verebilmek için cevaplayıcının 4 aşamalı bir süreci kullandığını ifade etmektedir. Bunlar sırasıyla; (1) ölçek maddesini yorumlamak, (2) bu maddeyle ilgili inançlarını ve tutumlarını hatırlamak (3) bu inanç ve tutumları, cevabı şekillendirecek şekilde mütalaa etmek ve (4) uygun cevap alternatifini seçmek olarak belirtilmiştir. Üçüncü aşamada cevaplayıcının yapacağı mütalaa sonucunu kendisine sunulacak cevap alternatiflerine aktarabilmesi için, yapılacak yargılamanın sunulacak cevap alternatifleri sayısı ile ilişkilendirilmesi gerekmektedir. En uygun sayıda cevap alternatifi sayısı arayışları da çoğunlukla bu noktaya odaklanmaktadır. Bir taraftan alternatif sayısının arttırılması cevaplayıcının kısıtlı ayırt edebilme yeteneğini aşıp, cevaplayıcının tükenmesine neden olabiliyor iken diğer taraftan az sayıda cevap alternatifi cevaplayıcının ayırt etme gücünü yok edip, cevaplayıcıyı kısıtlayarak yuvarlama hatalarına neden olabilmektedir (Jacoby ve Matell, 1971).

Cevaplayıcının anlamlı şekilde ayırt edebilme yeteneğinden hareketle yedi cevap alternatifinin en uygun olduğunu ilk olarak Symond (1924; Preston ve Colman; 2000'dan alınmıştır) belirtmiş olsa da Miller'in (1956) makalesi belki de bu konudaki en temel kaynak haline gelmiştir. Miller insan beyninin kaç farklı uyarıcıyı

[4] Cevaplayıcıların süreklilik gösteren duygularını kesikli bir ölçeğe aktarmaları istendiğinde (örn. Bir tutumunu 6 noktalı bir ölçeğe aktarma) yuvarlama hatası ortaya çıkar. Örneğin bir cevaplayıcıya bir partiye oy verip verilmeyeceği sorulmuş olsun. Cevaplayıcının gerçek tutumu %60 evet, %40 hayır olduğunda ve bu tutumunu evet/hayır şeklinde belirtmesi istendiğinde yuvarlama hatası ortaya çıkmış olacaktır (Lehmann ve Hulbert, 1972).

 ISSN: 1308-2922

anlamlı bir şekilde ayırt edebileceğini sorguladığı çalışmasında öncelikle "*bit*" kavramını tanımlamıştır. Miller'e göre iki eşit alternatif arasında seçim yapmak için gerekli olan bilgi bir bit'lik bir bilgidir[5]. Bu halde bir bitlik bilgi ile 2, iki bitlik bilgi ile 4, ve üç bitlik bilgi ile 8 alternatif arasından karar vermek mümkündür. Bu halde kaç alternatif arasından seçim yapılabileceği 2^{bit} şeklinde gösterilebilir.

Miller (1956) tek boyutlu uyarıcılar için Pollack, Gamer & Bebe Center, Rogers ve O'Connell'ın deneysel çalışmalarından hareketle tek boyutlu uyarıcılar için alınan bilginin ortalamasını 2,6 bit ($2^{2,6}$=6,0629) ve standart sapmasını da 0,6 bit ($2^{0,6}$=1,5157) olarak hesaplamıştır. Çok boyutlu uyarıcılar için de 7 cevap alternatifinin uygunluğunu işaret eden Miller aslında 7 rakamının sihrine inanır gibi bir görüntü sergilemektedir. Miller bu bağlamda dünyanın 7 harikası, 7 büyük günah, insanın 7 dönemi, cennetin 7 katı, 7 ana renk, haftanın 7 günü ve müziğin 7 notasının sihirli olduğunu düşündüğü 7 rakamı için örnek vererek "*belki de bütün bu yedilerin arkasında bizi kendisini keşfetmeye davet eden daha derin bir şeyler vardır*" diyerek, anlamlı bir ayırım için yedi cevap alternatifinin en uygun olduğunu belirtmektedir.

Miller'i izleyen çalışmaların bir kısmında da 7 cevap alternatifi için destek bulunmuştur. Mesela, Preston ve Colman (2000) en güvenilir skorları 7-10 cevap alternatiflerinde elde etmiştir. Ayırt edicilik 6 ve daha fazla cevap alternatiflerinde maksimuma ulaşmıştır. Sonuçta Preston ve Colman (2000) en çok 7 cevap alternatifi için destek bulurken, yaygın olarak kullanılan 5 cevap alternatifi için destek bulamamıştır.

Viswanathan *vd.* (2004) 7 cevap alternatifli ölçeklerin ilave varyansa sahip olarak daha iyi bir performansa sahip olması beklerken, 3 cevap alternatifli ölçeklerin de değişkenlikleri büyük ölçüde ortaya çıkardığını tespit etmişlerdir. Viswanathan *vd.* (2004) özellikle ürün özelliklerinde anlamlı ayırım için en uygun alternatif sayısının 3 olduğunu, bunun nedeni olarak belki de müşterilerin ürün özelliklerini "düşük", "orta" ve "yüksek" olarak değerlendirmeleri

---

[5] Miller (1956) bilginin miktarının varyans kavramı ile aynı olduğunu düşünmektedir. Varyans büyüdükçe ne olacağı konusunda bilgi azalmaktadır. Tekdüze dağılımda varyans $p \times (1-p)$ kadar olduğundan p=0,5 varyansın maksimum olduğu noktadır. Bilgi miktarı daha fazla gözlem yaparak arttırılabilir. Öte yandan küçük varyans gelecekte gözlemimizin hangi yöne gideceği hakkında çok şey bildiğimiz anlamına gelir ve ilave gözlemin sağlayacağı fayda azalır.

olabileceğini ifade ettikten sonra bulguların özellikle pazarlama araştırmaları için 5 veya 7 cevap alternatifi yerine 3 cevap alternatifinin uygunluğuna işaret ettiğini belirtmişlerdir.

Garner ise (Cox, 1980'den alınmıştır) cevaplayıcının bilgi aktarımının 20 cevap alternatifinde maksimuma ulaştığını bulmuş ve sonuç olarak "cevap alternatifi sayısının arttırılması ile bilginin kaybedilmeyeceği açıktır" şeklinde bir ifade kullanmıştır.

Cox (1980) uygulama düzeyinde en uygun cevap alternatifi sayısının ölçeğin amacına ve sistematik değişkenliğine bağlı olduğunu belirtmektedir. Cox'un belirttiğine göre; uyarıcı merkezli (stimulus centered)[6] bir ölçekte en uygun cevap alternatifi sayısı dört faktör tarafından belirlenmektedir. (1) ölçeğin bilgiyi aktarma kapasitesi (2) cevaplayıcının uyarıcıyı cevap alternatiflerine aktarma yeteneği (3) uyarıcıdan aktarım için mevcut olan bilgi miktarı ve (4) araştırmacının ihtiyaç duyduğu bilgi miktarı.

Buradaki ilk üç maddeyi göz önünde bulundurarak Cox (1980) cevap alternatiflerinin sayısının arttırılması halinde, cevaplayıcının cevap alternatifleri arasındaki farkı anlamakta zorlanacağını, bu nedenle de cevaplayıcının gerçek değeri ile belirttiği değer arasında farklar ortaya çıkmaya başlayacağını belirtmektedir. Bu halde cevap alternatifi sayısındaki artış bir yandan ölçeğin bilgi aktarma kapasitesini iyileştirirken diğer yandan cevaplama hatasını arttıracak bir etki meydana getirecektir. Davis (1987) cevaplayıcının uyarıcıya aktarma yeteneği ile ilişkili olarak ölçek formatı seçerken genel kuralın basit olanı seçmek olduğunu fakat bu konuda eğitimli cevaplayıcılar için karmaşık yapıların daha ilgi çekici olabileceğini belirtmiştir.

Cevaplayıcının alternatifler arasındaki farkı anlamalarına yardımcı olabilmek amacıyla alternatifler adlandırılabilir. Fakat bu noktada karşımıza Türkiye'de gerçekleştirilen araştırmalarda 5'den fazla alternatif kullanılmak istendiğinde, alternatiflerin ne şekilde adlandırılacağı noktasında problem çıkmaktadır. Bu konuda Türkiye'de basılı belli başlı kullanılan kaynaklardan Altunışık ve diğ. (2001), Gegez (2005), Kurtuluş (2004) ve Nakip (2006) örnek olarak oluşturdukları ölçekleri, sadece 5 cevap alternatifi üzerine

[6] Uyarıcı merkezli ölçek; cevaplayıcılara genellikle bir küme ölçek maddesi sorulur, buradaki her bir madde ilgilenilen bir özelliğin derecesini (yoğunluğunu) ölçmek için kullanılır. Bu tür bir ölçekle elde edilen veride sütunlar uyarıcıların, satırlar cevaplayıcıların verilerini içerir.

 ISSN: 1308-2922

kurgulamışlar bunun dışında örnekler geliştirmemişlerdir. Sadece Kurtuluş (2004; 283) boyutsal ayırma ölçeği için; tamamen yeterli, yeterli, biraz yeterli, ne yeterli ne yetersiz, oldukça yetersiz, yetersiz, tamamen yetersiz şeklinde bir adlandırma kullanmıştır. Bu bağlamda belki de etiket oluşturmanın zorluğu ve araştırma kitaplarındaki boşluk nedeniyle Türkiye'de basılan yöntem kitaplarında ve pazarlama alanında gerçekleştirilen çalışmalarda nedeni sorgulanmadan büyük ölçüde 5 cevap alternatifi kullanılarak gerçekleştirildiği gözlenmektedir.

Churchil ve Peter (1984) ölçekteki bütün cevap alternatiflerini adlandırmakla sadece uç noktaları adlandırmak arasında güvenilirliği etkileyecek bir ilişkinin olmadığını bulmuştur. Belli başlı Pazarlama Araştırmaları kitapları çoğunlukla bu bilgiyi aktarmaktadır (örn. Churchil 1999). Fakat, Türkiye'de etiket kullanmak isteyen araştırmacıların 5 cevap alternatifinin ötesinde özellikle 7 cevap alternatifli ölçekler için bu etiketleri kullanma olanağı oldukça sınırlı gözükmektedir.

Alwin (1992) fazla sayıda cevap alternatifinin tutumun yönü yanında yoğunluğunu da ölçtüğünü belirtmektedir. Aynı doğrultuda Davis (1987) daha fazla cevap alternatifinden elde edilen verinin tutumun yönünü ölçmek üzere daha az sayıda cevap alternatifine indirgenilebiliyor iken[7], tutumun yönünü ölçmek üzere elde edilen verinin tutumun yoğunluğu hakkında bilgi sağlayamayacağını işaret etmektedir. Ancak unutmamak gerekir ki çok sayıda cevap alternatifi daha fazla ilave fayda sağlamaz (örn. Cox, 1980; Gren ve Rao, 1970).

2.4. Güvenilirlik

SSCI'e göre Cronbach'ın 1951 yılındaki makalesi 7000'den fazla makalede kaynak gösterilmiştir. Pazarlama konusundaki araştırmalarda da genel olarak bir güvenilirlik ölçütü olarak, orijinal makaleye atıf yapılmaksızın, Cronbach Alpha değeri kullanılmaktadır. Cronbach alpha çok maddeli ölçeklerin içsel tutarlılıklarını ölçmek üzere geliştirilmiş ve şu şekilde formüle edilmiştir:

$$\alpha = \left(\frac{k}{k-1}\right)\left(1-\sum_{i=1}^{k}\sigma_i^2/\sigma_s^2\right) \text{ veya}$$

[7] Bunu ilerleyen kısımda da görülebileceği gibi aslında ampirik olarak Jacoby ve Matell (1971) göstermiştir.

$$\alpha = \frac{k\bar{r}}{1+\bar{r}(k-1)}$$ burada;

k: ölçekteki madde sayısı

σ_i^2: maddenin varyansını

σ_s^2: ölçeğin varyansını

$\bar{r}$: maddeler arasındaki ortalama korelasyonu göstermektedir (Peterson, 1994).

Epistemoloji literatürünün çoğunlukla cevap alternatifi sayısının arttırılmasının güvenilirliği arttıracağını söylemektedir. Çünkü cevap alternatiflerinin sayısını arttırmak cevaplayıcıya kendi nesnel durumu daha iyi yansıtma olanağı sunacak, böylelikle bilgi kaybedilmeyecek ve sonuçta daha güvenilir bilgi alınabilecektir. Bu doğrultuda ölçekteki cevap alternatifi sayısı ile güvenilirlik arasındaki ilişkiyi sorgulayan çalışmalar da birbirleriyle oldukça çelişkili sonuçlar ortaya çıkarmıştır. Birtakım çalışmalar bu iki değişken arasında bir ilişkiye rastlamaz iken, (+) ve (-) yönlü ilişkiler ortaya çıkaran çalışmaların varlığı dikkat çekmektedir.

Jacoby ve Matell (1971) hem geçerliliğin hem de güvenilirliğin Likert türü maddeler için cevap alternatifi sayısından bağımsız olduğunu bulmuştur. Bending ise (Cox 1980'den alınmıştır) 3 ile 9 cevap alternatifi arasında güvenilirliğin değişmediğini, 11 cevap alternatifinde ise azaldığını tespit etmiştir. Scherpenzel ve Saris (1995; Scherpenzel'den alınmıştır) yuvarlama hatasını hatırlatarak nispeten çok seçeneğe sahip ölçeklerin az seçeneğe sahip olanlardan daha güvenilir olduğunu belirtmektedir. Scherpenzel cevap alternatifi sayısı ile güvenilirlik arasında önce artan sonra azalan oranda bir ilişki bulunduğunu, bu nedenle güvenilirliği artırmak adına ilave edilecek seçeneklerin bir sınırı olduğunun altını çizmekte ve sonuç olarak 11 cevap alternatifine sahip ölçeklerin en güvenilir ölçekler olduğunu belirtmektedir.

Preston ve Colman (2000) ise 7-10 aralığındaki cevap alternatiflerinde güvenilirliğin, bu aralığın dışındakilere oranla daha yüksek olduğunu tespit etmiştir.

Cicchetti ve diğ. (1985) Monte Carlo simülasyonu kullanarak gerçekleştirdikleri çalışmalarında cevap alternatifi ile güvenilirlik arasında bir doğru orantı tespit etmişlerdir. Cevap alternatifi 2'den 3'e çıktığında elde edilen güvenilirlik artışı yüksek olmuş, ancak 7 cevap

 ISSN: 1308-2922

alternatifinin ötesinde güvenilirlik artışı ihmal edilebilecek kadar küçük bulunmuştur.

Birkett (1986) aynı soruları 2, 6 ve 14 cevap kategorisinde sorduğu çalışmasında cevap alternatifi sayısına bağlı olarak güvenilirliğin de artmasını beklemiş, ancak bu beklentisi gerçekleşmemiştir. Birkett'in (1986) sonuçlarına göre kategori sayısı yediyi aştıktan sonra güvenilirlikteki artış oldukça küçük olmaktadır ve 14 cevap alternatifi kullanıldığında güvenilirlik azalmıştır. Buradan hareketle, Birkett (1986) 5 veya 7 cevap alternatifinin en uygun çözümler olduğu sonucuna varmıştır.

Masters (1974) 2, 3, 4, 5, 6 ve 7 cevap alternatifine sahip ölçekler ile yaptığı deneysel çalışmasında ölçülen kavram üzerine tutumların geniş bir dağılım gösterdiği hallerde, cevap alternatifi sayısının güvenilirlik üzerinde küçük bir etkiye sahip olduğunu, yani güvenilirliğin cevap alternatifi sayısından bağımsız olduğunu bulmuştur. Masters (1974) cevaplayıcıları ayırt etmek için [tutumun yönü önemli hale geldiğinde] iki cevap alternatifinin bile yeterli olabileceği sonucuna ulaşmıştır.

Alwin (1992) epistemoloji literatürü doğrultusunda cevap alternatifi sayısının güvenilirliği arttırdığı hipotezini test ettiği çalışmasında, bulguları bu hipotezi destekler nitelikte olmasına rağmen 2 cevap alternatifine sahip ölçeklerin beklenenden çok daha güvenilir olduğunu bulmuştur. Bu çalışma özellikle davranış ölçümlerinde 2 cevap alternatifli ölçeklerin, 3 veya 4 cevap alternatifli ölçeklere oranla daha güvenilir olduğunu ortaya çıkarmıştır. Bu nedenle ölçmenin amacının tutumun yönünü ölçmek olduğunda 2 cevap alternatifli ölçeklerin kullanımı tavsiye edilmektedir.

Komorita ve Graham (Jacoby ve Matell, 1971'den alınmıştır) "...eğer bu geçerli bir genelleme ise bunun en büyük etkisi basitlik ve uygulama kolaylığı nedeniyle bütün ölçekler iki noktalı [alternatifli] olmalıdır" şeklinde bir ifade kullanmıştır. Jacoby ve Matell (1971) bu doğrultuda 2 den 19 cevap alternatifine kadar 18 farklı cevap alternatifi kullanılarak hazırlanan anketlerini yirmişer öğrenciye uygulamışlardır. Verilerin toplanmasını izleyen aşamada 2, 4, 6, 8, 10 ,12 ,14, 16 ve 18 cevap alternatifine sahip ölçekler alınan veriyi dichotomous (katılıyorum, katılmıyorum) 3, 5, 7, 9, 11, 13, 15, 17 ve 19 cevap alternatiflerine sahip ölçekleri trichotomious (katılıyor, emin değil ve katılmıyor) şeklinde 2 veya 3 kategoriye indirgemişlerdir. Bu indirgeme öncesi ve sonrasındaki güvenilirlik arasında anlamlı olmayan bir düşme gözlenmiştir. Bu analizleri sonucunda Jacoby ve

Matell (1971) hem geçerliliğin ve hem de güvenilirliğin Likert türü ölçekler için cevap alternatifi sayısından bağımsız olduğu sonucuna ulaşmışlardır. Bu bulgu ayrıca, farklı cevap alternatiflerinden elde edilen sonuçların iki veya üç cevap alternatifi formatına dönüştürülerek karşılaştırılabileceğini, bu dönüşümlerin verilerin güvenilirliğini etkilemeyeceğini göstermektedir. Dahası bu bulgu; az sayıda seçenek kullanma stratejisinin, aksi halde cevap almanın zor olduğu koşullarda rahatlıkla kullanılabileceğini göstermektedir.

Ölçekteki cevap alternatifi sayısının artması hem epistemoloji hem de ölçme literatürüne göre belirli bir noktaya kadar güvenilirliği arttırmaktadır. Weathers *vd.* (2005) bu noktaya ilginç bir açıklama getirmektedir. Araştırmacılara göre cevap alternatifi sayısındaki artış önceden ifade edildiği gibi cevaplayıcının işini zorlaştıracaktır. Fakat cevaplayıcı bu karmaşık görevi kabullenmek yerine işini basitleştirme adına bir önceki soruda verdiği cevabı bir sonraki soru içinde kullanma eğilimine girebilecektir. Bu eğilim özellikle cevap alternatifleri arasındaki ayırımı yapmak üzere gerekli muhakemeyi yapmak istemeyen veya bu yetenekte olmayan cevaplayıcılar arasında daha fazla görülecektir. Bu sayede cevaplayıcı çoğu maddeye aynı cevabı vermiş olacaktır. Bunun sonucunda da toplam ölçek varyansı azalacak bir başka ifade ile maddeler arasındaki korelasyon artacak ve nihayetinde güvenilirlik yüksek gözükecektir. Ancak ölçeğin geçerliliği azalacaktır.

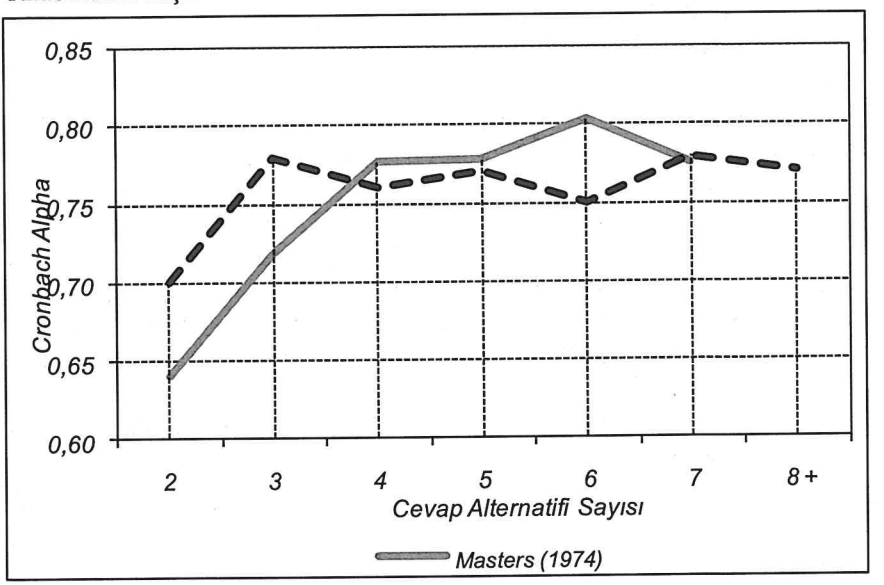
Ölçeklerdeki cevap alternatifleri sayısı ile güvenilirlik arasındaki ilişkiyi basılı araştırmalardan hareketle inceleyen meta-analizlerden birisi Churchill ve Peter (1984) tarafından gerçekleştirilmiştir. 2-20 arasından cevap alternatiflerine sahip 131 çalışmanın incelendiği bu meta-analiz cevap alternatifi sayısı ile güvenilirlik arasında bir ilişki ortaya çıkarmıştır[8]. Bu konudaki ikinci meta analiz Peterson (1994) tarafından gerçekleştirilmiştir. Peterson güvenilirlik ile cevap alternatifi sayısı konusunda Şekil 1'de görülen sonuçları elde etmiş ve sonuç olarak iki cevap alternatifine sahip ölçeklerin güvenilirliklerinin düşük olduğunu ancak daha fazla cevap alternatiflerinde güvenilirliğin yüksek olduğunu tespit etmiştir. Her ne kadar bu bulgu Master (1974) ile birebir örtüşmese de; bu iki çalışmanın bir arada değerlendirilmesi sonucunda ilave cevap alternatifi ilave etmenin, ilgili literatürde ifade edildiği gibi,

[8] Yazarlar bu ilişkinin normal aralıkta geçerli olduğunu belirtmişler, ancak normal aralıktan neyin kastedildiğini belirtmemişlerdir.

 ISSN: 1308-2922

güvenilirliği çok fazla arttırmadığı görülmektedir. Masters'ın (1974) kendi deneysel çalışmalarından Peterson'un (1994) meta analizinden elde ettiği cevap alternatifi sayısı ile Cronbach Alpha arasındaki ilişki Şekil'de görülmektedir.

Şekil 1: Cevap Alternatifi Sayısı İle Cronbach Alpha Arasındaki İlişki



Kaynak: Masters 1974 ve Peterson (1994) uyarlanmıştır

3. ÖZET VE SONUÇLAR

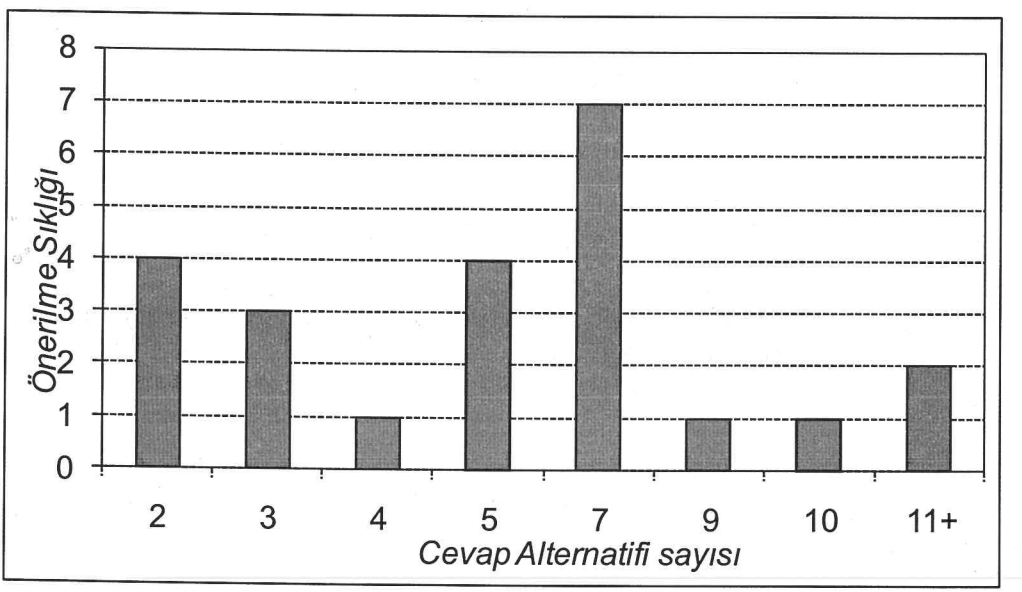
Araştırma tasarımları hazırlanırken, ölçekteki cevap alternatifi sayısı konusunda verilecek kararın önemsiz olduğunu söylemek mümkün değildir. Ancak kaç adet cevap alternatifi kullanılması gerektiği konusunda yapılan araştırmalar bu probleme net bir çözüm getirememiş, buna paralel olarak farklı araştırmalar, uygulamalar ve amaçlar için farklı sayıda cevap alternatifinin uygun olacağına dair ipuçları ortaya çıkarmıştır. Bu konuda literatürde en fazla destek 7 cevap alternatifi için bulunmuştur. Ancak 5, 6, 8 ve 9 cevap alternatiflerinin, kullanım sıklıkları kadar destek bulamadıkları görülmektedir. İlginç bir şekilde, araştırmacının amacına da bağlı

olarak, 2 ve 3 kategorili ölçeklerin daha fazla destek bulduğu anlaşılmaktadır.

Cevap alternatifi sayısı ile güvenilirlik arasındaki ilişki önce artan, sonra azalan bir niteliğe sahiptir. İlave cevap alternatifi eklemek cevaplayıcının işini zorlaştırarak farklı hatalara neden olurken, ilave seçenekler belirli bir noktaya kadar güvenilirlik üzerinde önemli bir katkı sağlamakta, bunun ötesinde sağlanan katkılar önemli olmamaktadır. İlgili literatürden anlaşıldığına göre yedinin ötesindeki ilave cevap alternatifleri güvenilirlik üzerinde önemli bir artış sağlayamamaktadır.

En uygun cevap alternatifi sayısı hakkında bu çalışma için kullanılan literatürde demokratik bir oylama yapılsa, bu taramada ele alınan çalışmaların sonuçlara göre; Şekil 2'de görüleceği gibi, seçimin galibi 7 cevap alternatifi olacaktır. Ancak çok sık kullanılmayan 2 kategori ilginç bir şekilde özellikle Türkiye'de en sık kullanılan 5 cevap alternatifi kadar destek bulmuştur. Son dönemlerde kullanımı artan 9 ve 11 cevap alternatifi ise bu demokratik seçimde oldukça başarısız görülmektedir.

Şekil 2: En Uygun Cevap Alternatifi Sayılarının Bu Çalışmadaki Önerilme Sıklıkları



ISSN: 1308-2922

KAYNAKÇA

Alwin D. F. (1992), "Information Transmission in the Survey Interview: Number of Response Categories and the Reliability of Attitude Measurement", **Sociological Methodology**, Vol. 22. pp. 83-118.

Birkett. Nicholas 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**, v.18, n.3, ss.205-215

Churchill, G.A. and Peter, J.P(1984), "Research Design Effects On The Reliability Of Rating Scales: A Meta Analysis", **Journal Of Marketing Research**, November, ss.360-375

Cicchetti, D.V. Showalter, D. and Tyrer, P.J. (1985), "The Effect Of Number Of Rating Scale Categories On Level Of Interrater Reliability: A Monte Carlo Investigation", **Applied Psyhological Measurement**, v.9, n.1, March, ss.31-36

Cox, E.P.III (1980), "The Optimal Number Of Response Alternatives", **Journal Of Marketing Research**, Vol.17, November, ss.407-422

Dawis, R.V.,(1987) "Scale Construction", **Journal Of Counseling Psychology**, v.34(4), ss.481-489

Dutton & Blum (1968) "The Measurement Of Attitudes Toward Arithmetic With A Likert-Type Test", **The Elementary School Journal**, February, ss.259-264

Friedman, L.W. & Friedman H.H.(1986), "Comparison Of Itemized vs Graphic Rating Scales: A Validity Approach", **Journal Of The Market Research Society**, v.28, n.3, ss.285-289

Gren, P.E.(1970), "Rating Scales And Informatio Recovery –How Many Scales And Response Categories To Use", **Journal Of Marketing**, v.34, ss.33-39

Jacoby, J. ve Matell, M.S.(1971) "Three-Point Scales Are Good Enough", **Journal Of Marketing Research**, n.8, November, ss.495-500

Jensen C. and Tendering, J.(2005) "Perceived Prominence And Scale Types", **FONETIC 2005**, proceedings, Göteborg University, ss.111-114

Lehmann, D.R., & Hulbert, J.(1972), "Are Three - Point Scales Always Good Enough", Journal **Of Marketing Research**, v.9, November, ss. 444-446

Masters, J.R.(1974), "The Relationship Between Number Of Response Categories And Reliablity Of Likert-Type Questionnaires", **Journal Of Educational Mesasurement**, v.11, n.1, Spring

Miller, G.A.(1956), "The Magical Number Seven, Plus Or Minus Two: Some Limits On Our Capacity For Processing Information", **The Pychological Review**, v.63, ss:81-97

Peterson, R.A.(1994) "A Meta-Analysis Of Cronbach's Coefficient Alpha", **Journal Of Consumer Research**, v.21, September, ss.381-391

Preston C.C and Colman A.M.(2000) "Optimal Number Of Response Categories In Rating Scales: Reliability, Validity And Discriminating Power And Respondent Preferences", **Acta Psyhologica**, 104, ss.1-15

Scherpenzeel, A., **"Why Use 11-Point Scale"** Retrieved 03/11/2008 from: http://www.swisspanel.ch/file/doc/faq/11pointscales.pdf

Viswanathan, M., Sudman, S. and Johnson, M. (2004) "Maximum Versus Meaningful Discrimination In Scale Response: Implications For Validity Of Measurement Of Consumer Perceptions About Products", **Journal Of Business Research**, 57, ss 108-124

Weathers, D., Sharma, S. and Niedrich R.W.(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**, v.58, ss.1516-1524

 ISSN: 1308-2922