Katılım Düzeyi Seçenek Sayısının İncelenmesi: Optimal Katılım Düzeyi Seçenek Sayısına İlişkin Bir Çıkarım¹

Analysis Of The Point Scale Range: An Inference On Optimal Point Scale Range

Volkan DOĞAN

Eskişehir Osmangazi Üniversitesi
Eskişehir, Türkiye
vodogan@ogu.edu.tr

Behçet Yalın ÖZKARA

Eskişehir Osmangazi Üniversitesi
Eskişehir, Türkiye
bozkara@ogu.edu.tr

Cengiz YILMAZ

Orta Doğu Teknik Üniversitesi
Ankara, Türkiye
ycengiz@metu.edu.tr

Ömer TORLAK

Rekabet Kurumu
Ankara, Türkiye
omer.torlak@rekabet.gov.tr

Özet

Bu araştırmanın amacı; veri karakteristiğinin (normal dağılım, çarpıklık, basıklık), içsel tutarlılık düzeyinin (Cronbach's Alpha), ölçekler arası korelasyon katsayılarının ve kovaryans matrislerinin yapılarının katılım düzeyi seçenek sayısına duyarlı olup olmadığını sistematik olarak incelemek ve istatistiksel olarak sınamaktır. Araştırma, Eskişehir Osmangazi Üniversitesi İktisadi ve İdari Bilimler Fakültesi öğrencilerinden kolayda örnekleme yöntemiyle elde edilmiş olan üç farklı veri seti kapsamında gerçekleştirilmiştir. Araştırma neticesinde katılım düzeyi seçenek sayısının arttıkça; içsel tutarlılık düzeyinin sistematik olarak arttığı ve 5'li ile 11'li katılım düzeyi seçenek sayısı arasındaki farkın istatistiksel olarak anlamlı olduğu görülmüştür. Ayrıca ölçekler arası korelasyon katsayısının ise sistematik olarak arttığı fakat artışın istatistiksel olarak anlamlı olmadığına işaret eden bulgulara ulaşılmıştır. Değişkenler arası kovaryans matrisleri yapılarının ise katılım düzeyi seçenek sayısına bağlı olarak anlamlı bir şekilde farklılaştığı fakat kovaryans matrisi temelli yapısal eşitlik modelindeki yol katsayılarının katılım düzeyi seçeneği sayısına bağlı olarak anlamlı bir şekilde farklılaşmadığı görülmüştür.

Anahtar kelimeler: katılım düzeyi seçenek sayısı, likert, veri karakteristiği, veri kalitesi.

Abstract

The objective of this study is to systematically analyze and statistically prove whether the internal consistency level of the data characteristics (normal distribution, skewness, and kurtosis) is sensitive to the point scale range of the correlation

¹ Bu makale 19. Ulusal Pazarlama Kongresinde sunulmuş olan bildirinin genişletilmiş halidir.

coefficients and covariance matrices. The research was carried out using three different data sets obtained by applying the convenience sampling method on students studying at Eskişehir Osmangazi University, Faculty of Economics and Administrative Sciences. According to the results, the internal consistency level increases when the point scale range increase; and, the difference between 5-point scale and 11-point scale was observed to be statistically significant. Furthermore, the inter-scales correlation coefficient was determined to increase systematically; however, this increase was found not to be statistically significant. It was also observed that the inter-variable covariance matrices change significantly based on the point scale range; yet the path coefficients in covariance-based structural equation modeling do not change significantly.

Keywords: point-scale range, likert, data characteristics, data quality.

1. Giriş

Bilimsel araştırmalarda yapılan ölçümlerin başlıca amacı, araştırma sorusunun cevaplanabilmesi ya da teorinin öne sürdüğü savların test edilebilmesi için gerekli olan bulguların elde edilmesidir (Alwin, 1997). Farklı veya benzer yöntemsel desenlere sahip çeşitli araştırmaların bulguları arasında belirli bir tutarlılık veya paralellik tespit edilmesi doğrultusunda ise elde edilen bulgulara ilişkin güvenilirlik oluşmaktadır. Tümevarımcı paradigmaya göre bilimsel ilerleme bu güvenilir bulguların kümülatif bir biçimde birikimi sayesinde gerçekleşmektedir. Öte yandan Popper’ci yanlışlamacı bilimsel yaklaşım ışığında odaklanılan nokta ise tutarlı birliktelik sergileyen bulguların kümülatif birikimden ziyade elde edilen tutarlılık sergilemeyen bulguların yaratmış olduğu paradigmatik (Kuhn, 1970) kırılmalardır. Bulguların kümülatif birikimi veya tutarsızlıkları doğrultusundaki ilerlemenin bilimsel meşruiyetinin, araştırmaların yöntemsel kurguları ve desenlemelerine bağlı bir husus olduğunu söylemek mümkündür. Araştırmanın deseni, bilimsel paradigmatik duruşu ve yöntemsel ayrıntıları, bulgular olarak kavramsallaştırılan somut delillerin soyutlaştırılarak teorik bilgi üretiminin ve teorik katkının sağlanması (Folger ve Turillo, 1999) için bilimsel bir gerekliliktir.

Pazarlama ve yönetim bilimleri açısından konuya yaklaşıldığında, bilimsel birikim oluşumuna en çok katkı (veri) sağlayan yöntemsel veri toplama aracının anket olduğu bilinmektedir (Scandura ve Williams, 2000). Anket; gerek düşük maliyeti, gerek sağladığı zaman kazancı, gerekse hızlı geri bildirim sağlaması nedeni ile pazarlama ve yönetim alanında sıklıkla kullanılan veri toplama aracıdır.

Psikometrik ölçeklerin pazarlama ve yönetim bilimleri alanında sıklıkla kullanılmaya başlanması ve kullanımda amaç ve yönetime uygun olmayan yaklaşımlara başvurulması, anket aracılığıyla elde edilen verilerin kalitelerinin sorgulanmaya ve tartışılmaya başlanmasına yol açmıştır. Bu tartışmalar; sistematik hata, sistematik olmayan hata, yanıt yanlılığı (response bias), yanıt dışı yanlılık (non-response bias) ve katılım düzeyi seçenek sayısı gibi yöntemsel detaylar üzerinden sürdürülmüştür (Armstrong ve Overton, 1977; Meade ve Craig, 2012; Barber vd., 2013). Özellikle bu tartışmaların katılım düzeyi seçenek sayısı üzerinden sürdürüldüğü araştırmalar incelendiğinde katılım düzeyi seçenek sayısının; katılımcıların ifadelerine katılım düzeyleri (Dawes, 2002), ölçeklerin güvenilirlik katsayıları (Cronbach, 1950; Bendig, 1953; Komorita, 1963; Jacoby ve Matell, 1971) ve ölçek ifadelerine cevap verme süreleri (Bevan ve Avant, 1968; Matell ve Jacoby, 1972) üzerindeki etkilere sahip olabildiği görülmektedir.

Yöntembilim kapsamında farklı ve gelişmiş yeni analiz tekniklerinin ortaya konulması (örn. Rasch analizi, partial least yapısal eşitlik modellemeleri, conditional process analysis, finite mixture partial least squares analysis gibi) pazarlama ve yönetim bilimleri alanında kabul görmekte ve araştırmacılar tarafından araştırma projeleri kapsamında uygulanmaktadır. Öte yandan bu yeni ve gelişmiş analiz tekniklerinin veri karakteristiğine duyarlı oldukları bilinmektedir (Dawes, 2008). Bu açıdan yöntem ve analiz tekniklerinin birlikteliğinin ve bu birlikteliğin bilimsel disiplini ve gereklilikleri sağlamasının önemi göz ardı edilemez. Bir diğer deyişle araştırmanın bulgularının elde edilmeye çalışıldığı analiz süreci, izlenen yöntem ve metodolojiye oldukça duyarlıdır. Nitekim bu çalışmanın amaçlarından ve katkılarından ilki yöntem ve metodolojiye ilişkin katılım düzeyi seçenek sayısını analiz sonuçları bağlamında incelemek, analiz sonuçlarının katılım düzeyi seçenek sayısına duyarlı olup olmadığını sınamak ve duyarlılığa ilişkin bir grafik ortaya koyabilmektir. Bu çalışmanın ikinci katkısı ise çalışılan ilgili evren ve kültürde, sistematik hatadan olabildiğince arınmış olan optimal katılım düzeyi seçenek sayısı (Cox, 1980) belirleyebilme adına önerilerde ve çıkarımlarda bulunabilmektir. Bu araştırmanın üçüncü katkısı ise veri karakteristiği ve kalitesi kapsamında yöntemsel ve metodolojik bir tercih olan katılım düzeyi seçenek sayısının etkisinin incelenmesi olacaktır. Bu araştırmanın son katkısı ise Türkiye bağlamındadır, ölçümlemenin kültüre oldukça duyarlı olduğu (Hui ve Triandis, 1989; Marin vd., 1992) göz önünde bulundurulduğunda konuya ilişkin olarak Türkçe literatürde, Türkiye örneklemini ele alan hiçbir çalışmaya rastlanılamamış olması, yapılan bu çalışmanın ilgili literatüre sunabileceği bir diğer anlamlı katkıyı ortaya çıkarmaktadır.

2. Teorik Çerçeve ve Literatür Taraması

Veri toplama aracı olan anketlerde yer alan ölçek ifadelerinin katılım düzeyi seçenek sayısı, literatürde bilgi teorisi kapsamında ele alınmaktadır. Bilgi teorisi, ölçek ifadelerine katılım düzeyinin belirtildiği katılım düzeyi seçenek sayısının arttıkça araştırılan olguya ilişkin elde edilecek bilgi düzeyinin de artacağını ileri sürmektedir (Alwin, 1997). Nitekim Garner (1960), Garner ve Hake (1951) ile Shannon ve Weaver (1949) tarafından gerçekleştirilmiş olan araştırmalarda bilgi teorisinin ileri sürdüğü savı destekleyen sonuçlara ulaşılmıştır. Benzer şekilde Dawes (2008) de yapmış olduğu çalışmada cevaplayıcıya daha çok katılım düzeyi seçenek sayısı sunulması ile ölçülmek istenen olguya ilişkin daha fazla bilgi elde edilebildiğine işaret eden sonuçlara ulaşmıştır. Daha fazla katılım düzeyi seçeneğinin olmasının (5'li yerine 7'li gibi), araştırmacının araştırmakta olduğu olguya ilişkin bilgi elde edimini hususunda daha kapsamlı bilgi sağlayabileceği ve cevaplayıcının kendini daha hassas bir şekilde yansıtabileceği ileri sürülmektedir (Alwin, 1997).

Bilgi teorisinin aksine motivasyon teorisi ise katılım düzeyi seçenek sayısının arttıkça bireylerin karmaşık görünen anketlerde yer alan ölçek ifadelerine kendilerini doğru ve tam olarak yansıtabilecek şekilde katılım düzeylerini belirtmeleri hususunda motivasyonlarının azalacağını ileri sürmektedir. Bu savın bir sonucu olarak, elde edilecek olan veri kalitesinin araştırma sonuçlarını etkileyebilecek düzeyde düşebileceği belirtilmektedir (Alwin, 1991; Krosnick, 1991).

Peabody (1962), Lunney (1970) ve Jacoby ve Matell (1971) yapmış oldukları araştırmalarda 2 ya da 3 katılım düzeyi seçenek sayısının yer aldığı ölçek ifadelerinin, araştırılan konuyla ilgili soyut olguya ilişkin gerekli bilgiyi sağlayabildiğini, daha fazla

katılım düzeyi seçenek sayısının olmasının bilgi artışı üzerinde anlamlı bir etkiye sahip olmayacağını belirtmişlerdir. Alwin (1992) ise konuya ilişkin daha detaylı bir ayırım yaparak; eğer amaç tutumun sadece yönelimini ölçmek ise 2 katılım düzeyi seçeneğinin yeterli olduğunu ancak aynı zamanda tutumun yoğunluğu da ölçülmek isteniyorsa ikiden fazla katılım düzeyi seçeneğine ihtiyaç duyulacağını ileri sürmüştür.

Revilla ve arkadaşları (2014), 50 bin katılımcı üzerinde yapmış oldukları geniş örneklemli çalışmada, optimal katılım düzeyi seçenek sayısının 5 olduğunu, 5’den fazla katılım düzeyi seçeneğinin olması durumunda verinin kalitesinin düştüğünü ifade etmişlerdir. Guilford (1954), katılım düzeyi seçenek sayısının en fazla 25 olması gerektiğini; Garner (1960) ise 20 katılım düzeyi seçeneğinin yeterli olduğunu daha fazla cevap alternatifinin elde edilen bilginin artışında anlamlı bir etkisinin olmayacağını belirtmiştir. Öte yandan Miller (1956) gerçekleştirdiği araştırmada, insanların ayırımı yapabilecekleri katılım düzeyi seçeneği sayısının 5 ila 9 arasında olduğunu ve 9’dan fazla katılım düzeyi seçeneğinin bulunmasının elde edilecek bilgi düzeyi üzerinde anlamlı bir katkısının olmadığını ifade etmişlerdir.

Literatürde üzerinde durulan bir diğer önemli konu ise, katılım düzeyi seçenek sayısının ölçek güvenilirlikleri üzerindeki etkisi olarak karşımıza çıkmaktadır. Katılım düzeyi seçenek sayısının ölçek güvenilirlikleri üzerindeki etkilerinin araştırıldığı çalışmalar incelendiğinde bulguların, genellikle katılım düzeyi seçenek sayısının artması ile birlikte güvenilirliğin de arttığı yönünde olduğu görülmüştür. Alwin (1997) yapmış olduğu çalışmada, güvenilirlik düzeyi açısından 11’li likertin 7’li likerte oranla daha iyi sonuçlar ortaya koyduğunu ifade etmiştir. Murphy ve Likert (1983) ise ölçeklere ilişkin güvenilirliğin, katılım düzeyi seçenek sayısının artırılarak iyileştirilebileceğini belirtmişlerdir. Preston ve Colman (2000) da farklı katılım düzeyi seçenek sayılarının bulunduğu karşılaştırmada benzer bir sonuca ulaşmışlar ve katılım düzeyi seçenek sayısındaki artış ile güvenilirlik katsayısındaki artışın sistematik olarak birlikte hareket ettiğini ileri sürmüşlerdir. Ayrıca Preston ve Colman (2000), katılım düzeyi seçenek sayısının fazla olmasının cevaplayıcılar tarafından pratiklik ve kullanım açısından olumsuz, tutum ve eğilimlerin ifade edilebilmesi açısından olumlu karşılandığını ortaya koymuşlardır. Optimal katılım düzeyi seçenek sayısına ilişkin olarak üzerinde uzun yıllardır çalışma yapılan bir diğer konu ise, katılım düzeyi seçenek sayısının verinin normal dağılımı yani basıklığı ve çarpıklığı üzerinde anlamlı bir etkisinin olup olmadığına ilişkindir. Bu konuya ilişkin literatürde yer alan araştırmalar incelendiğinde bir tutarlılık olduğunu söylemek zordur. Dawes (2002), çarpıklık ve basıklık ile katılım düzeyi seçenek sayısı arasında herhangi bir sistematik ilişki bulamazken, Peterson ve Wilson (1992) konuya ilişkin olarak katılım düzeyi seçenek sayısının çarpıklık ve basıklık üzerinden anlamlı etkileri olabileceğini ileri sürmüşlerdir.

Daha önce değinilenlere ilave olarak, ölçek ifadelerinin yer aldığı anketin uygulandığı kültürel yapının katılım düzeyi seçeneği sayısına yönelik yaklaşım üzerinde etkili olabileceğine araştırmacılar tarafından dikkat çekilmiştir (Bachman ve O'Malley, 1984; Marin vd., 1992). Chen ve arkadaşları (1995) yaptıkları araştırmada katılım düzeyi seçeneği sayısına yönelik cevaplayıcılar tarafından verilen tepkilerin kültürel olarak farklılaşabildiğini, batı kültürü cevaplayıcıları ile doğu kültürü cevaplayıcılarının cevap verme stillerinin farklı olduğuna işaret eden sonuçlara ulaşmışlardır. Dolayısıyla her kültürel yapının cevap alternatifi aralığı sayısını algılayış biçimleri ve cevap alternatifi aralığı sayısının katılımcılar üzerindeki etkisinin farklılaşabildiğini söylemek mümkündür.

Bu noktaya kadar değinilen çalışmalar çerçevesinde bu araştırmanın ilerleyen kısmında aşağıdaki araştırma sorularına cevap aranacak ve bu araştırma sorularına ilişkin hipotezler test edilecektir.

I.Katılım düzeyi seçeneği sayısı, ölçeğin güvenilirlik düzeyi (katsayısı) üzerinde sistematik ve istatistiksel düzeyde bir değişime sebep olmakta mıdır?

H₁: Katılımcı düzeyi seçenek sayısına bağlı olarak ölçek güvenilirlik düzeyi anlamlı bir şekilde farklılaşmaktadır.

II.Psikometrik ölçek ifadeleri aracılığıyla ölçümlenmeye çalışılan tutum ve eğilim gibi soyut olgular arasındaki korelasyon katsayıları, korelasyon matrisleri ve kovaryans matrisleri, katılım düzeyi seçeneği sayısına bağlı olarak sistematik ve istatistiksel olarak anlamlı bir şekilde farklılaşmakta mıdır?

H_{2a}: Katılım düzeyi seçenek sayısına bağlı olarak korelasyon matrisi anlamlı bir şekilde farklılaşmaktadır.

H_{2b}: Katılım düzeyi seçenek sayısına bağlı olarak kovaryans matrisi anlamlı bir şekilde farklılaşmaktadır.

III.Katılım düzeyi seçeneği sayısı, parametrik analizler için bir gereklilik olan (verinin kalitesinin ve karakteristiğinin) bir göstergesi olarak ele alınabilen basıklık ve çarpıklık değerlerini ne yönde etkilemektedir? Eğer bir etki söz konusu ise sistematik bir eğilim söz konusu mudur?

H_{3a}: Katılım düzeyi seçenek sayısına bağlı olarak basıklık değeri farklılaşmaktadır.

H_{3b}: Katılım düzeyi seçenek sayısına bağlı olarak basıklık değeri farklılaşmaktadır.

3. Yöntem

3.1. Veri toplama süreci ve veri toplama araçları

Araştırmanın veri toplama süreci Eskişehir Osmangazi Üniversitesi İktisadi ve İdari Bilimler Fakültesi öğrencileri üzerinde, kolayda örnekleme yöntemi uygulanarak ve Ocak 2014 ayı içerisinde gerçekleşmiştir. Araştırmanın verileri bırak-al yöntemiyle toplanmıştır. Bu çalışma, yöntem-bilimsel araştırma sorularına sahip olup, belirli bir teoriyi test etme veya teoriye ilişkin revizyonda bulunma girişimi taşımamaktadır. Bu kapsamda araştırmanın evreni Eskişehir Osmangazi Üniversitesi İktisadi ve İdari Bilimler Fakültesi öğrencileri olarak tanımlanabilir. Daha önce değinildiği üzere bu araştırmanın veri toplama aracı anket olup, anket formu içerisinde sırasıyla; Salinas ve Perez (2009) tarafından Martin ve Brown (1990), Aaker (1996), Weiss vd. (1999) ve Villarejo (2002) çalışmalarından derlenerek geliştirilmiş olan ve toplam dört ifadeden oluşan marka imajı ölçeği, Severi ve Ling (2013) tarafından Kim ve Kim (2005) ile Yoo vd. (2000) çalışmalarından derlenerek geliştirilmiş olan ve toplam dört ifadeden oluşan marka sadakati ölçeği, Hsu (2008) tarafından geliştirilmiş olan ve toplam üç ifadeden oluşan müşteri memnuniyeti ölçeği ve son olarak katılımcıların demografik karakteristiklerini (cinsiyet ve yaş) belirlemeye yönelik ifadeler yer almaktadır. Araştırmanın veri toplama aracı olan anket formunda yer alan ölçek ifadelerine katılım düzeyi; “Kesinlikle Katılmıyorum (1),....., Kesinlikle Katılıyorum (5-7-11)” şeklinde 5’li, 7’li ve 11’li aralıkta sayısallaştırılmıştır. Ölçeklerin içsel geçerliliklerinin sağlanması bakımından uzman görüşleri alınmış ve tercüme-yeniden tercüme yöntemi uygulanmıştır. Ayrıca pilot anket uygulaması aracılığıyla ölçek ifadelerine son hali verilmiştir.

3.2. Araştırmanın yöntemsel tasarımı

Araştırmada, üç farklı anket formu aracılığıyla veri toplanmıştır. Bu üç farklı anket formu aracılığıyla veri toplama süreci araştırmanın yöntemsel tasarımından kaynaklanmaktadır. Anket formlarının yapıları ve katılım düzeyi seçenek sayıları Tablo 1'deki gibidir.

Tablo 1. Anket formlarının katılım düzeyi seçenek sayıları

	Marka imajı	Marka sadakati	Müşteri memnuniyeti
Anket-1	5	5	5
Anket-2	5	7	7
Anket-3	5	11	11

Marka imajı ölçeğinin katılım düzeyi seçenek sayısı üç farklı ankette de beş (5'li cevap aralığı) olarak sabit tutulmuştur. Her üç anket formunda da marka imajı ölçeğinin katılım düzeyi seçenek sayısının beş olarak sabit tutulmasının temel sebebi; yapılacak olan varyans analizi neticesinde aynı anakütleden çekilmiş olan üç farklı örneklemin araştırılan konu kapsamında kendi aralarında homojen olduğuna dair kanıt elde etmektir. Bu kanıt sayesinde farklı katılım düzeyi seçenek sayılarına ilişkin olarak ortaya çıkabilecek istatistiksel farklılıkların örneklem arası heterojenlikten kaynaklanmadığına ilişkin sav güçlendirilmiş olacaktır. Öte yandan marka sadakati ve müşteri memnuniyeti ölçeklerinin katılım düzeyi seçeneği sayıları ise üç anketin her birinde farklı kurgulanmıştır (5'li-7'li-11'li). Bu kurgunun sebebi katılım düzeyi seçenek sayısından kaynaklanan veri karakteristiği ve analiz sonuçları farklılıklarını ölçümleyebilme ve sınavabilmektir. Daha önce değinildiği üzere bu araştırmanın verileri üç farklı anket formu aracılığıyla, üç aşamalı veri toplama sürecinde elde edilmiştir. Üç aşamalı veri toplama sürecindeki üç farklı örneklem birbirinden farklı olup her bir örneklem Eskişehir Osmangazi Üniversitesi İktisadi ve İdari Bilimler Fakültesi öğrencileri anakütlesinden çekilmiştir.

4. Bulgular

Araştırmanın veri toplama süreci neticesinde anket-1 (5'li) aracılığıyla 158 veri, anket-2 (7'li) aracılığıyla 170 veri ve anket-3 (11'li) aracılığıyla 158 veri olmak üzere toplam 486 katılımcıdan veri toplanmıştır. Veri setinde yer alan kayıp değerlerin istatistiksel olarak tamamen rasgele dağılım sergileyip sergilemediği "Estimated Mean" yaklaşımı doğrultusunda "Missing Completely at Random" analiziyle sınanmış (Little's MCAR test: chi-square=0.001; sig.=0.001) ve kayıp verilerin yerine yeni veri atamama yani kayıp verilerin araştırma kapsamından çıkarılmasına karar verilmiştir. Bu kararın sebebi, kayıp verilerin rasgele (olasılıksal) olarak dağılmadığı yapılacak olan veri atamalarının araştırma sonuçlarında yanlışlık problemine yol açabilme şüphesinin var olmasıdır (MCAR>0.05). Kayıp veriler araştırmadan çıkarıldıktan sonra anket-1 aracılığıyla toplanmış 140 veri, anket-2 aracılığıyla toplanmış olan 164 veri ve anket-3 aracılığıyla toplanmış olan 145 veri geriye kalmıştır. İleride yapılacak olan analizlerin sonuçlarında alt gruplar arasındaki örneklem büyüklüğü farklılığından kaynaklanabilecek olan Tip-1 hatayı ortadan kaldırabilmek adına alt veri gruplarından en küçük olan 140'a tüm alt veri setleri (örneklem) eşitlenmiştir. Bu eşitleme sürecinde veri gruplarından (örneklem) 140'ar tane rassal atama yöntemiyle (SPSS) veri seçiminde bulunulmuştur. Sonuç olarak anket-1, anket-2 ve anket-3 aracılığıyla toplanmış 140'ar veri (toplam 420 veri) üzerinden araştırmanın ilerleyen sürecindeki analizler gerçekleştirilmiştir. Değinilen eşit örneklem büyüklüğüne sahip üç

farklı veri setinin bir diğer deyişle üç farklı örneklemin kendi aralarında homojenliğinin sınanması adına marka imajı ölçeği bağlamında üç farklı veri setinin (örneklemin) birbirinden istatistiksel olarak anlamlı farklılığa sahip olup olmadığı oneway-Anova (varyans analizi) testi ile sınanmıştır. Oneway-Anova testi öncesinde Levene istatistiği sonuçları (levene statistics=2.611; sig.=0.075) üzerinden veri setleri (örneklem) arası varyans homojenliği olduğu görülmüştür. Varyans homojenliği doğrultusunda scheffe post-hoc testi sonuçlarının yorumlanması yönünde karar kılınmıştır. Ayrıca marka imajı ölçeği ortalamasının üç farklı veri seti (örneklem) bağlamında anlamlı olarak farklılaşmadığı sonuçlarına oneway-Anova testi ile ulaşılmıştır ($F=2.754$; sig.=0.065). Scheffe post-hoc testi doğrultusunda yapılan veri setleri arası çoklu karşılaştırmaların istatistiksel anlamlılık değerlerinin de 0.066 ile 0.656 aralığında yer aldığı görülmüştür. Bir diğer deyişle üç farklı veri setinin (anket-1, anket-2, anket-3 örneklem) kendi aralarında homojen yapıya sahip olduğu ortaya konmuştur. Bu noktadan sonra veri setleri (örneklem) arasındaki farklılığın örneklem farklı karakteristik yapıya sahip olduğundan değil katılım düzeyi seçenek sayısından kaynaklandığı yönünde yorumlamalar yapabilme adına araştırmanın argümanları güçlendirilmiştir.

Tablo 2. Veri karakteristikleri

Ölçek	Cevap aralığı	Ort.	Varyans	std. Sapma	çarpıklık	Basıklık	K-S	S-W
Marka Sadakati	Anket-1 (5'li)	3.569	1.427	1.194	-.496	-.683	0.001	0.001
	Anket-2 (7'li)	4.738	2.740	1.655	-.184	-1.021	0.001	0.001
	Anket-3 (11'li)	7.248	7.993	2.827	-.389	-.763	0.001	0.001
Müşteri Memnuniyeti	Anket-1 (5'li)	3.689	1.319	1.149	-.785	-.191	0.001	0.001
	Anket-2 (7'li)	4.852	2.581	1.606	-.581	-.404	0.001	0.001
	Anket-3 (11'li)	7.424	7.825	2.797	-.501	-.795	0.001	0.001

İlk aşamada Tablo 2'de görüldüğü üzere üç farklı katılım düzeyi seçenek sayısına sahip ölçekler aracılığıyla toplanmış olan veri setlerine ilişkin temel veri karakteristikleri incelenmiştir. Cevap aralığı sayısına bağlı olarak çarpıklık, basıklık ve normallik değerleri ile ilgili dikkat çeken sistematik bir eğilim olduğunu söylemek oldukça zordur. Bu noktada sistematik eğilimden kasıt, katılım düzeyi seçenek sayısındaki farklılığa bağlı olarak belirli bir yönde olmak üzere çarpıklık, basıklık ve normallik değerlerinin matematiksel olarak artış veya azalış sergilemesidir. Dolayısıyla çalışmanın H3a ve H3b hipotezleri red edilmiştir. Çalışmanın ilerleyen kısımlarında da yer alan sistematik eğilim vurgusu, belirli bir yönde doğrusal olarak matematiksel artışı ifade etmektedir.

Tablo 3. Katılım Düzeyi Seçenek Sayısı Farklılığı Kapsamında Ölçeklerin İçsel Tutarlılıkları

Ölçek	Katılım Düzeyi Seçenek Sayısı	Cronbach's Alpha
Marka Sadakati	Anket-1 (5'li)	0.795
	Anket-2 (7'li)	0.832
	Anket-3 (11'li)	0.867
Müşteri Memnuniyeti	Anket-1 (5'li)	0.911
	Anket-2 (7'li)	0.928
	Anket-3 (11'li)	0.938

İkinci aşamada Tablo 3’de görüldüğü üzere üç farklı katılım düzeyi seçenek sayısına sahip ölçekler aracılığıyla toplanmış olan veri setlerine ilişkin içsel tutarlılık katsayıları incelenmiştir. İçsel tutarlılık katsayılarının katılım düzeyi seçenek sayısındaki artışla birlikte arttığı; marka sadakati ve müşteri memnuniyeti olmak üzere her iki ölçek kapsamında da gözlemlenmiştir. Bir diğer deyişle katılım düzeyi seçeneği sayısı arttıkça içsel tutarlılık katsayısının arttığı yönünde sistematik bir eğilim olduğu görülmüştür. Bu sistematik eğilimin (değişimin) istatistiksel olarak da anlamlı kırılmalara (değişimlere) sahip olup olmadığını sınamak üzere Feldt’s test (Feldt, 1969) ve Fisher-Bonett test (Kim ve Feldt, 2008) uygulanmıştır. Test sonuçlarına ilişkin bilgiler Tablo 4’de yer almaktadır.

Feldt’s test süreci aşağıdaki süreci kapsamaktadır:

$$W = \frac{1 - \hat{\alpha}_2}{1 - \hat{\alpha}_1} \quad (\hat{\alpha}_1 \text{ ve } \hat{\alpha}_2 = \text{içsel tutarlılık katsayıları; } W = \text{Hakstian Whalen test değeri})$$

formülü aracılığıyla elde edilen W değeri, $(N_1-1).(N_2-1)$ doğrultusunda hesaplanan serbestlik derecesi (df) ile birlikte F dağılımı (istatistiği) tablosunda değerlendirilip ulaşılan değer anlamlılık (sig.) değeri olarak belirlenir. Böylece iki farklı içsel tutarlılık katsayısı arasında anlamlı farklılık olup olmadığı sınanmış olur (Feldt, 1969).

Fisher Bonett test süreci ise Z dağılımını (istatistiğini) kullanan bir süreci kapsamaktadır. Ayrıca Fisher Bonett testini Feldt’in testine kıyasla ölçek ifadelerinin sayılarına ve ölçeklerin ifade sayıları arasındaki farklılıklara karşı daha hassas olduğu bilinmektedir. Fisher Bonett testi doğrultusunda elde edilen ki-kare değeri aşağıda yer alan formül ile elde edilmektedir (Kim ve Feldt, 2008):

$$M_{F-B} = \sum_{j=1}^m \frac{[\ln(1 - r_j) - \hat{\mu}^*]^2}{\sigma^2[\ln(1 - r_j)]}$$

(M_{F-B} =asimptotik olarak $m-1$ serbestlik derecesi için ki-kare değeri; $\hat{\mu}^* = \ln(1-p_j)$ ’nin ağırlıklandırılmış ortalaması; k = karşılaştırılan α sayısı; m = farklı örneklem sayısı; σ^2 = varyans; r_j = ağırlıklandırılmış ortalama varyansın bileşeni)

Ayrıca Feldt’s test ve Fisher-Bonett test, içsel tutarlılık katsayıları arasında istatistiksel olarak anlamlı farklılığın olup olmadığını sınarken sadece iki farklı içsel tutarlılık katsayısı arasındaki farkı sınavabilmektedir. Bu doğrultuda Tablo 4 yapılandırılmıştır.

Tablo 4. Cronbach’s alpha katsayıları arası farklılıklar

Ölçek	Karşılaştırma	Feldt’s test	Fisher-Bonett test
Marka Sadakati	Anket-1 (5’li) Anket-2 (7’li)	0.120	0.170
	Anket-2 (7’li) Anket-3 (11’li)	0.080	0.130
	Anket-1 (5’li) Anket-3 (11’li)	0.005*	0.019*
Müşteri Memnuniyeti	Anket-1 (5’li) Anket-2 (7’li)	0.106	0.154
	Anket-2 (7’li) Anket-3 (11’li)	0.189	0.236
	Anket-1 (5’li) Anket-3 (11’li)	0.016*	0.041*

Tablo 4’de ve aşağıda yer alan Şekil 1’de görüldüğü üzere marka sadakati ve müşteri memnuniyeti olmak üzere her iki ölçek için de katılım düzeyi seçenek sayısı arttıkça içsel tutarlılık katsayısı sistematik olarak artmaktadır. Fakat istatistiksel anlamlılık kapsamında bu eğilime bakıldığında her iki ölçek için de 5’li ve 11’li derecelendirme koşullarının (katılım düzeyi seçenek sayısı) içsel tutarlılık katsayıları arasında istatistiksel olarak anlamlı farklılık olduğu görülmektedir. Bir diğer deyişle 5’li ve 11’li katılım düzeyi seçenek sayısı bağlamında içsel tutarlılık katsayısı anlamlı olarak artmaktadır. Dolayısıyla çalışmanın H_1 hipotezi kısmen desteklenmiştir, daha detaylandırmak gerekirse 5’li ile 7’li katılım düzeyi seçenek sayısı arasında anlamlı bir farklılık bulunamazken, araştırma sonuçları 5’li ile 11’li katılım düzeyi seçenek sayısı arasında güvenilirlik katsayısı bazında anlamlı farklılık olduğuna işaret etmektedir. İçsel tutarlılık katsayısı olan Cronbach’s alpha katsayısı formülünde madde puanlarının ve ölçek puanlarının varyansının aynı anda ele alındığı göz önünde bulundurulduğunda bu araştırma kapsamında tespit edilen farklılığın katılım düzeyi seçeneği sayısı ve katılımcıların psikometrik düzeyde ölçülmesi bağlamında tartışılabileceğini söylemek mümkündür.

Şekil 1. Katılım Düzeyi Seçenek Sayısı Değişimi Doğrultusunda İçsel Tutarlılık Katsayıları

Üç farklı katılım düzeyi seçenek sayısı kapsamında marka sadakati ile müşteri memnuniyeti arasındaki korelasyon katsayısı ve anlamlılığına ilişkin pearson-korelasyon analizi çıktıları ise Tablo 5’de yer almaktadır. Ayrıca Tablo 5’de katılım düzeyi seçenek sayısı farklılığının marka sadakati ve müşteri memnuniyeti korelasyon ve kovaryans yapılarında istatistiksel olarak anlamlı farklılığa yol açıp açmadığı bir diğer deyişle kovaryans matrislerinde istatistiksel anlamlı farklılık oluşturup oluşturmadığının sınındığı Box’s-M test çıktıları da yer almaktadır.

Tablo 5. Farklı Katılım Düzeyi Seçeneği Sayıları Kapsamında Korelasyon Katsayılarının Ve Kovaryans Matrislerinin Karşılaştırılması

	<i>Cevap Alternatifi Sayısı</i>	<i>Korelasyon katsayısı</i>	<i>sig. (p)</i>			
Marka Sadakati & Müşteri Memnuniyeti	Anket-1 (5’li)	0.681	0.001 [*]			
	Anket-2 (7’li)	0.701	0.001 [*]			
	Anket-3 (11’li)	0.714	0.001 [*]			
Box’s M kovaryans matrislerinin eşitliği testi						
<i>Cevap Alternatifi Sayısı</i>	<i>Değişkenler</i>	<i>Marka Sadakati</i>	<i>Müşteri Memnuniyeti</i>	<i>Box’s M</i>	<i>F</i>	<i>Sig. (p)</i>
Anket-1 (5’li)	Marka Sadakati	1.427	0.935	207.944	34.417	0.001 [*]
	Müşteri Memnuniyeti	0.935	1.319			
Anket-2 (7’li)	Marka Sadakati	2.740	1.864			
	Müşteri Memnuniyeti	1.864	2.581			
Anket-3 (11’li)	Marka Sadakati	7.993	5.646			
	Müşteri Memnuniyeti	5.646	7.825			
Box’s M korelasyon matrislerinin eşitliği testi						
<i>Cevap Alternatifi Sayısı</i>	<i>Değişkenler</i>	<i>Marka Sadakati</i>	<i>Müşteri Memnuniyeti</i>	<i>Box’s M</i>	<i>F</i>	<i>Sig. (p)</i>
Anket-1 (5’li)	Marka Sadakati	1	0.681	250.954	41.569	0.001 [*]
	Müşteri Memnuniyeti	0.681	1			
Anket-2 (7’li)	Marka Sadakati	1	0.701			
	Müşteri Memnuniyeti	0.701	1			
Anket-3 (11’li)	Marka Sadakati	1	0. 714			
	Müşteri Memnuniyeti	0.714	1			

Yukarıda yer alan Tablo 5'de görüldüğü üzere marka sadakati ile müşteri memnuniyeti arasındaki korelasyon katsayıları; 5'li katılım düzeyi seçenek sayısı aracılığıyla elde edilen veriler kapsamında (0.681), 7'li katılım düzeyi seçenek sayısı aracılığıyla elde edilen veriler kapsamında (0.701) ve 11'li katılım düzeyi seçenek sayısı aracılığıyla elde edilen veriler kapsamında (0.714) olmak üzere istatistiksel olarak anlamlı şekilde tespit edilmiştir. Tespit edilen bu korelasyon katsayıları arasında istatistiksel olarak anlamlı farklılık olup olmadığının sınanması aşağıda yer alan formüller (Fisher, 1921) aracılığıyla gerçekleştirilmiştir.

$$r' = (0.5) \log_e \left[\frac{1+r}{1-r} \right] \quad z = \frac{r'_1 - r'_2}{\sqrt{\frac{1}{n_1 - 3} + \frac{1}{n_2 - 3}}}$$

(formülde geçen r_1 ve r_2 , korelasyon katsayıları; n_1 ve n_2 , örneklem büyüklükleridir.)

Fischer (1921)'in z dönüşüm formülü aracılığıyla yapılmış olan korelasyon katsayıları arasındaki istatistiksel karşılaşmalar neticesinde; 5'li katılım düzeyi seçenek sayısı kapsamında elde edilen korelasyon katsayısı (0.681) ile 7'li katılım düzeyi seçenek sayısı kapsamında elde edilen korelasyon katsayısı (0.701) arasında istatistiksel olarak anlamlı farklılık olmadığı ($z=-0.32$; $p=0.375$), 7'li katılım düzeyi seçenek sayısı kapsamında elde edilen korelasyon katsayısı (0.701) ile 11'li katılım düzeyi seçenek sayısı kapsamında elde edilen korelasyon katsayısı (0.714) arasında istatistiksel olarak anlamlı farklılık olmadığı ($z=-0.22$; $p=0.413$), 5'li katılım düzeyi seçenek sayısı

kapsamında elde edilen korelasyon katsayısı (0.681) ile 11’li katılım düzeyi seçeneği sayısı kapsamında elde edilen korelasyon katsayısı (0.714) arasında istatistiksel olarak anlamlı farklılık olmadığı ($z=-0.53$; $p=0.298$) görülmüştür. Özetlemek gerekirse, katılım düzeyi seçeneği sayısındaki farklılığa bağlı olarak marka sadakati ile müşteri memnuniyeti arasındaki korelasyon katsayısının sistematik bir şekilde arttığı fakat bu artışın istatistiksel olarak anlamlı olmadığı görülmüştür.

Öte yandan Tablo 5’de yer alan Box’s M test sonuçları incelendiğinde katılım düzeyi seçeneği sayısına bağlı olarak marka sadakati ile müşteri memnuniyeti değişkenlerinin kovaryans matrisinin yapılarının farklılaştığı görülmüş olup bu farklılığın istatistiksel olarak anlamlı bir farklılık olduğu tespit edilmiştir ($F=34.417$; $p=0.001$). Dolayısıyla çalışmanın H_{2b} hipotezinin desteklendiği görülmektedir. Marka sadakati ile müşteri memnuniyeti değişkenleri arası korelasyon matrisinin katılım düzeyi seçeneği sayısına bağlı olarak farklılaşıp farklılaşmadığı ise değişkenlerin standardize edilmiş z skorları üzerinden yapılan Box’s M test analizi ile test edilmiştir. Box’s M test sonuçları marka sadakati ile müşteri memnuniyeti değişkenleri arası korelasyon matrisinin katılım düzeyi seçeneği sayısına bağlı olarak anlamlı bir şekilde farklılaştığı tespit edilmiştir ($F=41.569$; $p=0.001$). Dolayısıyla çalışmanın H_{2a} hipotezinin desteklendiği görülmektedir.

Son olarak müşteri memnuniyetinin marka sadakati üzerindeki etkisinin (yol katsayısının) 5’li, 7’li ve 11’li katılım düzeyi seçeneği sayıları kapsamında farklılaşıp farklılaşmadığı incelemek adına kovaryans algoritması temelli yapısal eşitlik modellemesi (Amos 16.0) ve varyans algoritması temelli (en küçük kareler yöntemi temelli) yapısal eşitlik modellemesi (SmartPLS 2.0) yol katsayıları incelenmiştir. İlgili analizlerin çıktıları aşağıda yer alan Tablo 6’da yer almaktadır.

Tablo 6. Kovaryans Algoritması Temelli Ve Varyans Algoritması Temelli Yapısal Eşitlik Modellemesi Çıktıları

Kovaryans algoritması temelli yapısal eşitlik modellemesi				
	<i>Katılım düzeyi Seçenek Sayısı</i>	<i>Yol Katsayısı</i>	<i>Standart Hata</i>	<i>sig. (p)</i>
Müşteri Memnuniyeti → Marka Sadakati	Anket-1 (5’li)	0.535	0.078	0.001*
	Anket-2 (7’li)	0.640	0.079	0.001*
	Anket-3 (11’li)	0.694	0.082	0.001*
Varyans algoritması temelli yapısal eşitlik modellemesi				
	<i>Cevap Aralığı</i>	<i>Yol Katsayısı</i>	<i>Standart Hata</i>	<i>t değeri</i>
Müşteri Memnuniyeti → Marka Sadakati	Anket-1 (5’li)	0.696	0.053	12.796*
	Anket-2 (7’li)	0.702	0.060	11.660*
	Anket-3 (11’li)	0.717	0.063	11.032*

Yukarıda yer alan Tablo 6’da yer alan yol katsayılarının birbirinden istatistiksel olarak anlamlı bir şekilde farklı olup olmadıklarını sınamak adına aşağıda yer alan formülden (Chin, 2009; Wilson, 2010) faydalanılmıştır.

$$t = \frac{Path_{sample_1} - Path_{sample_2}}{\sqrt{\frac{(m-1)^2}{(m+n-2)} * S.E.^2_{sample1} + \frac{(n-1)^2}{(m+n-2)} * S.E.^2_{sample2}}} * \left[\sqrt{\frac{1}{m} + \frac{1}{n}} \right]$$

(m= ilk yolun bulunduğu modelin serbestlik derecesi ; n= ikinci yolun bulunduğu modelin serbestlik derecesi)

Yapılmış olan hesaplamalar doğrultusunda 5’li, 7’li ve 11’li katılım düzeyi seçeneği sayısına bağlı olarak, kovaryans algoritması temelli yapısal eşitlik modellemesi ve varyans algoritması temelli (en küçük kareler yöntemi temelli) yapısal eşitlik modellemesi kapsamında yol katsayıları arasında istatistiksel olarak anlamlı farklılık olmadığı tespit edilmiştir. Bir diğer deyişle 5’li, 7’li ve 11’li katılım düzeyi seçeneği sayısına bağlı olarak, müşteri memnuniyetinin marka sadakati üzerindeki etkisinin sınındığı yol katsayısının hem kovaryans algoritması temelli yapısal eşitlik modellemesinde hem de varyans algoritması temelli (en küçük kareler yöntemi temelli) yapısal eşitlik modellemesinde istatistiksel olarak anlamlı bir şekilde farklılaşmadığı görülmüştür.

5. Sonuç ve Tartışma

Hemen hemen herkes sosyal ve davranışsal bilimlerde ölçüm kalitesinin araştırmanın temel taşı olduğu konusunda hem fikirdir (Alwin, 2014). Ayrıca ölçümlemeye ilişkin konular bilimsel araştırmalar adına kritik öneme sahiptir (Alwin, 2007). Metodolojik ve yöntemsel tasarım tercihlerinden kaynaklanan ölçümleme farklılıklarına odaklanılan bu araştırma kapsamında pazarlama araştırmaları için önemli bir sayısallaştırma kaynağı olan point-scale tipi (Weijters vd., 2010) katılım düzeyi seçeneği sayısı kapsamında veri karakteristiği, içsel tutarlılık katsayıları, korelasyon katsayıları, kovaryans matrislerinin yapıları ve path katsayılarına ilişkin bazı sonuçlara ulaşılmıştır.

İlk olarak, katılım düzeyi seçenek sayısının arttıkça içsel tutarlılık katsayısının sistematik olarak arttığı fakat bu artışın sadece 5’li ve 11’li katılım düzeyi seçenek sayısı arasında istatistiksel olarak anlamlı bir değişim sergilediği görülmüştür. Bir anlamda katılım düzeyi seçenek sayısının psikometrik ölçüm niteliğinde bazı farklılıklara yol açabildiği çıkarımında bulunmak mümkündür. Fakat bu farklılığın sadece 5’li ve 11’li katılım düzeyi seçenek sayısı arasında istatistiksel olarak anlamlı olduğu göz ardı edilmemelidir. Bu bulgular katılım düzeyi seçenek sayısı tercihi bağlamında yapılacak olan manipülasyonların bu araştırmanın kısıtlılıkları ve şartları kapsamında istatistiksel olarak farklılıklara sadece 5’li yerine 11’li katılım düzeyi seçenek sayısı tercih edildiğinde neden olabileceği sonucunu ortaya koymaktadır. 5’li katılım düzeyi seçenek sayısı yerine 7’li katılım düzeyi seçenek sayısı tercihinde bulunmanın farklılığı sadece sistematik düzeyde kılacağı yönünde sonuçlara ulaşılmıştır. Bu sonuçlar, psikometrik ve içsel geçerliliği farklı kültürlerde tutarlı bir grafik sergilemeyen ve daha önce Türk kültüründe yeterli içsel tutarlılık düzeyi ortaya koyamamış ölçeklerin Türkiye örnekleminde uygulanırken 11’li katılım düzeyi seçenek sayısı aracılığıyla ölçümlenmesinin araştırmacıların elini güçlendirebileceği yönünde çıkarım sağlamaktadır. Öte yandan orijinal olarak 11’li katılım düzeyi seçenek sayısı aracılığıyla geliştirilmiş olan ölçeklerin 5’li katılım düzeyi seçenek sayısı ile Türkiye örnekleminde uygulandığında içsel tutarlılığa ilişkin göreceli olarak düşük içsel tutarlılık katsayılarının elde edilmesi bu araştırmanın sonuçları ışığında olağan karşılanmalıdır. Elbette ki kültürel içeriğe bağlı olarak farklı kültürlerde cevap yanlılığı (response bias) düzeyinin farklılaştığından (Hui ve Triandis, 1989; Marin vd., 1992) dolayı farklı sonuçlar elde edilebilir. Bu araştırma neticesinde elde edilen bulgular doğrultusunda çalışmanın ilk araştırma sorusuna ilişkin elde edilen sonuç şu şekilde ifade edilebilir: katılım düzeyi seçenek sayısı arttıkça içsel tutarlılık katsayısı sistematik olarak artarken, istatistiksel anlamlı artış ise sadece 5’li ile 11’li katılım düzeyi seçenek sayısı arasında gerçekleşmektedir. Bu sonuç Arce-Ferrer (2006) tarafından ileri sürülen

farklı cevaplama formatlarının cevaplayıcılar tarafından farklı algılamalara neden olduğu savını sistematik olarak tam anlamıyla istatistiksel olarak ise kısmi şekilde desteklemektedir. Ayrıca Alwin (1997), Murphy ve Likert (1983) ve Preston ve Colman (2000) tarafından tespit edilen ölçeklere ilişkin güvenilirliğin, katılım düzeyi seçenek sayısının artırılarak iyileştirilebileceği yaklaşımı bu araştırmanın sonuçlarıyla paralellik sergilemektedir. Bu araştırmanın literatüre katkılarında bir tanesi de katılım düzeyi seçenek sayısının içsel tutarlılıkta yarattığı farklılıkların görece olarak karşılaştırılmasının (Alwin, 1997; Murph ve Likert, 1983; Preston ve Colman, 2000) ötesine geçip istatistiksel olarak karşılaştırılmasının gerçekleştirilmesidir.

İkinci olarak, psikometrik ölçekler aracılığıyla ölçümlenmeye çalışılan tutum ve eğilim gibi soyut olgular (değişkenler) arası kovaryans matrislerinin katılım düzeyi seçeneği sayısına bağlı olarak sistematik ve istatistiksel olarak anlamlı bir şekilde farklılaştığı, korelasyon katsayılarının ise katılım düzeyi seçeneği sayısına bağlı olarak sistematik olarak farklılaştığı fakat istatistiksel olarak anlamlı bir şekilde farklılaşmadığı yönünde bulgulara bu araştırmanın kısıtlılıkları kapsamında ulaşılmıştır. Bir diğer deyişle 5’li, 7’li veya 11’li katılım düzeyi seçeneği sayısı arasından araştırmacı tarafından yapılacak olan tercih araştırmaya konu olan değişkenler arası kovaryans matrislerinde istatistiksel olarak anlamlı farklılık oluşturmaktadır. Araştırma neticesinde elde edilen bu sonuç kovaryans matrisi temelli istatistiksel analiz yöntemleri sonuçlarının katılım düzeyi seçeneği sayısının manipüle edilerek farklılaştırılabileceği şüphesini akıllara getirmektedir. Fakat araştırma kapsamında yapılmış olan kovaryans temelli yapısal eşitlik modellerinde yer alan yol katsayılarının katılım düzeyi seçeneği sayısına bağlı olarak farklılaşmadığı bulgusu bu şüpheyi ortadan kaldırmaktadır. Bir diğer deyişle katılım düzeyi seçeneği sayısı kovaryans matrisi temelli bir istatistiksel sonuç manipülasyonuna imkan sunmamaktadır. Öte yandan değişkenler arası korelasyon katsayısının da katılım düzeyi seçeneği sayısı arttıkça sistematik olarak arttığı fakat bu artışın istatistiksel olarak anlamlı olmadığı yönünde bulgulara ulaşılmıştır. Elde edilen bu bulguyu katılım düzeyi seçeneği sayısı tercihi ile korelatif ilişkilerin incelendiği araştırmaların sonuçların manipüle edilemeyeceği sonucunu çıkarmak mümkündür. Bu sonuçlar ışığında çalışmanın ikinci ve üçüncü araştırma sorularına ilişkin ulaşılan cevaplar şu şekilde ifade edilebilir: katılım düzeyi seçeneği sayısı tercihiyle bağlı olarak değişkenler arası kovaryans matrisleri istatistiksel olarak anlamlı bir şekilde farklılaşmakta fakat bu farklılık kovaryans temelli yapısal eşitlik modellerindeki yol katsayılarında istatistiksel olarak anlamlı farklılığa neden olmamaktadır. Ayrıca katılım düzeyi seçeneği sayısı arttıkça değişkenler arası korelasyon katsayısı artmakta fakat bu artış istatistiksel olarak anlamlı değildir. Sağladığı bu sonuçlar itibarıyla bu çalışmanın literatüre bir diğer katkısı da katılım düzeyi seçenek sayısı ile korelasyon katsayısı, kovaryans matrisi ve korelasyon matrisi arasındaki ilişkiler üzerinedir.

Son olarak, katılım düzeyi seçeneği sayısı ile verinin normal dağılım sergilemesi arasında sistematik olarak bir ilişki görülemediği. Bir diğer deyişle üç farklı katılım düzeyi seçeneği sayısında da (5’li-7’li-11’li) veriler normal dağılım yapısını sağlayamamışlardır. Çarpıklık ve basıklık değerleri kapsamında da verilerin katılım düzeyi seçeneği sayısına bağlı olarak bir yönelim sergilemediği görülmüştür. Bu sonuçlar ışığında çalışmanın dördüncü araştırma sorusuna ilişkin ulaşılan cevap şu şekilde ifade edilebilir: katılım düzeyi seçeneği sayısına bağlı olarak veriler, normal dağılım sergilemeye ilişkin bir yönelim sergilememekte ve çarpıklık ile basıklık

değerleri sistematik olarak herhangi bir yönde eğilim sergilememektedir. Bir diğer deyişle bu araştırmanın sınırlılıkları içerisinde; 5'li, 7'li veya 11'li katılım düzeyi seçenek sayısı tercihinde bulunmanın, verilerin normal dağılım sergilemesi adına bir avantaj sağlamadığı görülmüştür. Araştırma ulaştığı bu sonuç itibarıyla, Dawes (2002)'in çarpıklık- basıklık ile katılım düzeyi seçenek sayısı arasında herhangi bir sistematik ilişki olmadığı savını desteklemektedir.

5.1. Sınırlılıklar ve gelecek araştırmalara yönelik çıkarımlar

Öğrenci örnekleme sınırlılıkları çerçevesinde gerçekleştirilen bu araştırmanın farklı örneklem kapsamlarında tekrarlanması ölçümün doğasına ilişkin daha genel çıkarım sağlayacaktır. Çalışma kapsamında kullanılan kolayda örnekleme ile seçilmiş olan öğrenci örnekleme dolayısıyla araştırmanın sonuçları genellenebilirliğinin kısıtlı olduğu göz önünde bulundurulmalıdır.

Ayrıca ölçümlemenin kültüre oldukça duyarlı olduğu (Hui ve Triandis, 1989; Marin vd., 1992) göz önünde bulundurulduğunda, gelecekte farklı kültürler üzerinden 5'li, 7'li ve 11'li katılım düzeyi seçenek sayısı karşılaştırılması yapılmasının literatüre anlamlı katkı sağlayacağı düşünülmektedir.

Diğer yandan gelecekte gerçekleştirilecek olan çalışmalarda katılım düzeyi seçenek sayısı yerine semantik farklılıklar veya stapel tipi katılım düzeyi seçenek sayısı kapsamında veri karakteristiği, ölçeklerin psikometrik ölçütleri, kovaryans yapıları ve korelasyon katsayıları karşılaştırması gerçekleştirilebilir. Ayrıca ölçek ifadelerine katılım düzeyi seçeneklerinin yanına bilmiyorum seçeneği de eklenerek 5'li, 7'li ve 11'li katılım düzeyi seçenek sayısı kapsamında veri karakteristiği farklılıkları incelenebilir ve bilmiyorum seçeneği tercih sıklığı karşılaştırması yapılabilir. Bir diğer potansiyel araştırma önerisi olarak, farklı yöntemlerle (elektronik, telefon, yüz-yüze, bırak-al gibi) anket formu aracılığıyla ulaşılan cevaplayıcıların kendi aralarında katılım düzeyi seçenek sayısına yönelik tepkileri, kayıp veri sayısının sağlıklı veri sayısına oranlanması yoluyla karşılaştırılabilir. Ayrıca gelecek çalışmalarda, katılım düzeyi seçenek sayısı ile örneklem hacmi arasındaki ilişkinin incelenmesinin literatüre anlamlı katkı sağlayabileceği düşünülmektedir.

5.2. Sonuç

Bu araştırma neticesinde katılım düzeyi seçenek sayısındaki artışın özellikle psikometrik bir yapı olan ölçeğin içsel tutarlılık katsayısı kapsamında istatistiksel olarak anlamlı bir artışa neden olduğu; psikometrik olmaktan daha çok istatistiksel ve analize ilişkin olan korelasyon katsayıları kapsamında, kovaryans matrisleri ve kovaryans matrisleri temelli yapısal eşitlik modelleri yol katsayıları kapsamında istatistiksel olarak anlamlı bir artışa neden olmadığına işaret eden sonuçlara ulaşılmıştır. Bu sonuç, katılım düzeyi seçenek sayısı tercihinin araştırma sonucunda elde edilecek olan sonuçlarda yapaylık oluşturmayacağı (spurious korelasyon gibi) ve hatta tersine sosyal bilimlerde yapılan araştırmalar için kritik öneme sahip olan psikometrik ölçeklerin içsel tutarlılıkları (güvenilirlikleri) adına bir fayda sağlayacağını göstermektedir. Bir diğer deyişle katılım düzeyi seçenek sayısı tercihinin, araştırmanın yöntemsel kalitesi adına beraberinde yanlılıkları değil kazanımları getirdiğine işaret eden sonuçlara bu çalışma kapsamında ulaşılmıştır. Ve bu sonuçlar bilgi teorisinin savlarına ilişkin (Shannon ve Weaver, 1949; Garner ve Hake, 1951; Garner, 1960; Alwin 1997; Dawes, 2008) bir kanıt sunmaktadır.

Sonuç olarak bu araştırma, optimal katılım düzeyi seçenek sayısı olarak araştırmacılara 11’li katılım düzeyi seçenek sayısını önermektedir. Weijters ve arkadaşları (2010) tarafından öğrenci örnekleme için fazla sayıda katılım düzeyi seçenek sayısına sahip cevap formatı, genel örneklem için ise 5’li katılım düzeyi seçenek sayısına sahip cevap formatı önerilmiştir. Bu kapsamda bu araştırmanın sonuçları ile Weijters ve arkadaşları (2010) tarafından gerçekleştirilen araştırmanın sonuçları tutarlılık sergilemektedir.

Kaynakça

- Aaker, D. A. (1996), “Measuring brand equity across products and markets”, *California Management Review*, 38(3), 102-120.
- Alwin, D. F. (1991), “Research on Survey Quality”, *Sociological Methods & Research*, 20, 3-29.
- Alwin, D. F. (1992), “Information Transmission in the Survey Interview: Number of Response Categories and the Reliability of Attitude Measurement”, *Sociological Methodology* (1992) kitabının içinde 83-118, editör: Marsden, P. V., American Sociological Association: Washington, DC.
- Alwin, D. F. (1997), “Feeling Thermometers vs. Seven-point Scales: Which are Better?”, *Sociological Methods and Research*, 25, 318-340.
- Alwin, D. F. (2007), *Margins of Error-A Study of Reliability in Survey Measurement*, John Wiley: NJ, Hoboken.
- Alwin, D. F. (2014), “Investigating Response Errors in Survey Data”, *Sociological Methods & Research*, 43(1), 3-14.
- Arce-Ferrer, A. J. (2006), “An investigation into the factors influencing extreme-response style”, *Educational and Psychological Measurement*, 66(3), 374-392.
- Armstrong, J. S. ve Overton, T. S. (1977), “Estimating Nonresponse Bias in Mail Surveys”, *Journal of Marketing Research*, 14, 396-402.
- Bachman, J. G. ve O’Malley, P. M. (1984), “Yea-saying, Nay-saying, and Going to Extremes: Are Black-White Differences in Survey Results Due to Response Styles?”, *Public Opinion Quarterly*, 48, 409-427.
- Barber, L. K., Barnes, C. M. ve Carlson, K. D. (2013), “Random and Systematic Error Effects of Insomnia on Survey Behavior”, *Organizational Research Methods*, 16(4), 616-649.
- Bendig, A. W. (1953), “The reliability of self-ratings as a function of the amount of verbal anchoring and of the number of categories on the scale”, *Journal of Applied Psychology*, 37, 38-41.
- Bevan, W. ve Avant, L. (1968), “Response latency, response uncertainty, information transmitted and the number of available judgemental categories”, *Journal of Experimental Psychology*, 76, 394-397.
- Chen, C., Lee, S.-Y. ve Stevenson, H. W. (1995), “Response style and cross-cultural comparisons of rating scales amongst East Asian and North American students”, *Psychological Science*, 6, 170-175.

- Chin, W. W. (2009), “Frequently asked questions–partial least squares and PLS-graph”, available at: <http://disc-nt.cba.uh.edu/chin/plsfaq.htm> (erişim tarihi: 3 Şubat 2014).
- Cox, E. P. (1980), “The optimal number of response alternatives for a scale: A review”, *Journal of Marketing Research*, 17, 407-422.
- Cronbach, L. J. (1950), “Response set and test validity”, *Educational and Psychological Measurement*, 6, 475-494.
- Dawes J. G. (2002), “Five point vs. eleven point scales: Does it make a difference to data characteristics?”, *Australasian Journal of Market Research*, 10(1), 39-47.
- Dawes, J. (2008), “Do data characteristics change according to the number of scale points used?”, *International Journal of Market Research*, 50(1), 61-77.
- Feldt, L. S. (1969), “A test of the hypothesis that Cronbach's alpha reliability coefficient is the same for two tests administered to the same sample”, *Psychometrika*, 45(1), 99-105.
- Fisher, R. A. (1921), “On the probable error of a coefficient of correlation deduced from a small sample”, *Metron*, 1, 3-32.
- Folger, R. ve Turillo, C. J. (1999), “Theorizing As The Thickness Of Thin Abstraction”, *Academy of Management Review*, 24(4), 742-758.
- Garner, W. R. ve Hake, H. W. (1951), “The amount of information in absolute judgments”, *Psychological Review*, 58, 446-459.
- Garner, W. R. (1960), “Rating scales, discriminability, and information transmission”, *Psychological Review*, 67(6), 343-352.
- Guilford, J. P. (1954), *Psychometric methods*, (2nd edition), McGraw-Hill: New York.
- Hsu, S.-H. (2008), “Developing an index for online customer satisfaction: Adaptation of American Customer Satisfaction Index”, *Expert Systems with Applications*, 34(4), 3033-3042.
- Hui, H. C. ve Triandis, H. C. (1989), “Effects of culture and response format on extreme response style”, *Journal of Cross-Cultural Psychology*, 20(3), 296-309.
- Jacoby, J. ve Matell, M. (1971), “Three point Likert scales are good enough”, *Journal of Marketing Research*, 8, 495-500.
- Kim, H. B. ve Kim, W. G. (2005), “The relationship between brand equity and firm's performance in luxury hotels and chain restaurants”, *Tourism Management*, 26, 549-560.
- Kim, S. ve Feldt, L. S. (2008), “A Comparison of Tests for Equality of Two or More Independent Alpha Coefficients”, *Journal of Educational Measurement*, 45(2), 179-193.
- Komorita, S. S. (1963), “Attitude content, intensity, and the neutral point on a Likert scale”, *Journal of Social Psychology*, 61, 327-334.
- Krosnick, J. A. (1991), “Response strategies for coping with the cognitive demands of attitude measures in surveys”, *Applied Cognitive Psychology*, 5, 213-236.

- Kuhn, T. S. (1970), *The Structure of Scientific Revolutions*, (2nd edition), University of Chicago Press: Chicago, IL.
- Lunney, G. H. (1970), “Using analysis of variance with a dichotomous dependent variable: An empirical study”, *Journal of Educational Measurement*, 7, 263-269.
- Marín, G., Gamba, R. J. ve Marín, B. V. (1992), “Extreme response style and Acquiescence among Hispanics: The role of acculturation and education”, *Journal of Cross-Cultural Psychology*, 23(4), 498-509.
- Martin G. S. ve Brown T. J. (1990), “In search of brand equity: the conceptualization and measurement of the brand impression construct”, In: Childers ML, et al, editor. *Marketing theory and applications*, American Marketing Association: Chicago, 2, 431-438.
- Matell, M. S. ve Jacoby, J. (1972), “Is there an optimal number of alternatives for Likert scale items? Effects of testing time and scale properties”, *Journal of Applied Psychology*, 56(6), 506-509.
- Meade, A. W. ve Craig, S. B. (2012), “Identifying Careless Responses in Survey Data”, *Psychological Methods*, 17(3), 437-455.
- Miller, G. A. (1956), “The magical number seven, plus or minus two: some limits on our capacity for processing information”, *Psychological Review*, 63, 81-97.
- Murphy, G. ve Likert, R. (1983), *Public opinion and the individual*, Harper: New York.
- Peabody, D. (1962), “Two components in bipolar scales: Direction and extremeness”, *Psychological Review*, 69(2), 65-73.
- Peterson, R. A. ve Wilson, W. R. (1992), “Measuring Customer Satisfaction: Fact and Artifact”, *Journal of the Academy of Marketing Science*, 20(1), 61-71.
- Preston, C. C. ve Colman, A. M. (2000), “Optimal number of response categories in rating scales: reliability, validity, discriminating power, and respondent preferences”, *Acta Psychologica*, 104, 1-15.
- Revilla, M. A., Saris, W. E. ve Krosnick, J. A. (2014), “Choosing the Number of Categories in Agree-Disagree Scales”, *Sociological Methods & Research*, 43(1), 73-97.
- Salinas, E. A. ve Perez, J. M. P. (2009), “Modeling the brand extensions' influence on brand image”, *Journal of Business Research*, 62, 50-60.
- Scandura, T. A. ve Williams, E. A. (2000), “Research methodology in management: Current practices, trends, and implications for future research”, *Academy of Management Journal*, 43, 1248-1264.
- Severi, E. ve Ling, K. C. (2013), “The Mediating Effects of Brand Association, Brand Loyalty, Brand Image and Perceived Quality on Brand Equity”, *Asian Social Science*, 9(3), 125-137.
- Shannon, C. E. ve Weaver, W. (1949), *The Mathematical Theory of Information*, University of Illinois Press: Illinois, Urbana.
- Villarejo, A. (2002), *La Medición del Valor de Marca en el Ámbito de la Gestión de Marketing*, Sevilla: CEADE.

- Weijters, B., Cabooter, E. ve Schillewaert, N. (2010), “The effect of rating scale format on response styles: The number of response categories and response category labels”, *International Journal of Research in Marketing*, 27(3), 236-247.
- Weiss A. M., Anderson, E. ve MacInnis, D. J. (1999), “Reputation management as a motivation for sales structure decisions”, *Journal of Marketing*, 63(October), 74–89.
- Wilson, B. (2010), “Using PLS to investigate interaction effects between higher order branding constructs”, Vinzi, V. E., Chin, W. W., Henseler, J. ve Wang, H. (Eds), *Handbook of Partial Least Squares Concepts, Methods and Applications*, Springer-Verlag, Berlin, 621-652.
- Yoo, B., Donthu, N. ve Lee, S. (2000), “An examination of selected marketing mix elements and brand equity”, *Journal of the Academy of Marketing Science*, 28(2), 195-211.

Analysis of the Point Scale Range within the Scope of Data Characteristics and Quality: An Inference on Optimal Point Scale Range

Volkan DOĞAN

Eskişehir Osmangazi University
Eskişehir, Turkey
vodogan@ogu.edu.tr

Behçet Yalın ÖZKARA

Eskişehir Osmangazi University
Eskişehir, Turkey
bozkara@ogu.edu.tr

Cengiz YILMAZ

Middle East Technical University
Ankara, Turkey
ycengiz@metu.edu.tr

Ömer TORLAK

Turkish Competition Authority
Ankara, Turkey
omer.torlak@rekabet.gov.tr

Extensive Summary

Introduction

The fact that the psychometric scales are now widely used in marketing and management sciences; that the approaches that are not appropriate for the purpose and method are applied and the quality of the data obtained via the questionnaire are now questioned and discussed. These discussions have been continued over methodological details such as systematic error, non-systematic error, response bias, non-response bias and number of participation level options (Armstrong and Overton, 1977; Meade and Craig, 2012; Barber et al., 2013). Especially when the researches where these discussions are continued over the number of options of the participation level, it can be observed that the number of options of the participation level may have effects on the participation levels of the participants to the items (Dawes, 2002), the reliability coefficients of the scales (Cronbach, 1950; Bendig, 1953; Komorita, 1963; Jacoby and Matell, 1971) and response time to scale items (Bevan and Avant, 1968; Matell and Jacoby, 1972).

Putting forth different and improved new analysis techniques (i.e. Rasch analysis, partial least structural equality modelings, conditional process analysis, finite mixture partial least squares analysis) is accepted in marketing and management sciences and applied within the scope of research projects by the researchers. On the other hand, this new and improved analysis techniques are known to be sensitive towards the data characteristics (Dawes, 2008). In this sense, the importance of association of the method and analysis techniques and this association fulfilling the scientific discipline and requirements cannot be ignored. In other words, the analysis period where the research findings are in the process of obtaining is very sensitive towards the method and methodology followed. Hence the first objective and contribution of this research is to examine the number of options of the participation level regarding the method and methodology, to test whether the analysis results are sensitive towards the number of options of the participation level and to reveal a graphic regarding the sensitivity. The

second contribution of this research is to be able to make suggestions and inferences regarding the determination of optimal number of options of the participation level which has been purified from systematic error within the universe and culture (Cox, 1980). The final contribution of this research is to examine the effect of the number of options of the participation level, which is methodological choice within the scope of data characteristics and quality.

Method

The data collection phase of the research has been carried out on the students of Faculty of Economics and Administrative Sciences of Eskişehir Osmangazi University by means of convenience sampling method in January 2014. The research data has been collected with drop and collect method. This study has methodological research questions and does not attempt testing a specific theory or revising such theory. In this sense, the population of the research can be defined as the students of Faculty of Economics and Administrative Sciences of Eskişehir Osmangazi University. As mentioned before, the questionnaire is the data collection instrument of this research and the questionnaire form includes the brand image scale developed from the works of Martin and Brown (1990), Aaker (1996), Weiss et. al. (1999) and Villarejo (2002) by Salinas and Perez (2009) and consisting of four items; the brand loyalty scale developed from the works of Kim and Kim (2005) and Yoo et al. (2000) by Severi and Ling (2013) consisting of four items; customer satisfaction scale developed by Hsu (2008) and consisting of three items and finally, the items determined to the demographical characteristics of the participants (gender and age). The participation level for the scale items included in the questionnaire form that is the data collection instrument of the research is numbered with 5, 7 and 11 intervals as "I definitely agree,....., I definitely do not agree". The expert opinions have been taken in order for the scales to fulfill the internal validity and translation/back-translation method has been applied. In addition, the scale items have been finalized by means of pilot questionnaire administration.

In the research, the data has been collected by means of three different questionnaire forms. The data collection period with these three different questionnaire forms occurs due to the methodological design of the research. The structure of questionnaire forms and the numbers of participation level options of questionnaire forms are as follows in Table 1.

Table 1. Numbers of participation level options of questionnaire forms

	Brand image	Brand loyalty	Customer satisfaction
Questionnaire-1	5	5	5
Questionnaire-2	5	7	7
Questionnaire-3	5	11	11

The number of options of the participation level for the brand image scale has been kept stable as 5 (5 response range) in 3 different questionnaires. The main reason for the fact that number of options of the participation level for the brand image scale has been kept stable as 5 in 3 different questionnaire is to obtain evidence regarding the fact that three different samples taken from the same universe as a result of the variance analysis to be carried out is homogeneous among themselves. Thanks to this evidence, the argument regarding the fact that the statistical differences that might arise

concerning the different number of options of the participation level do not occur due to heterogeneity among the samples. On the other hand, the participation level numbers of the brand loyalty and customer satisfaction scales have been framed differently in each of 3 questionnaires (5, 7, and 11). The reason for this framework is to be able to measure and test the data characteristic and the differences in analysis results due to the number of options of the participation level. As stated above, the data of this research is obtained via three different questionnaire forms during three-phased data collection period. The three different samples in three-phased data collection period is different from each other and each sample is taken from the universe consisting of the students of Eskişehir Osmangazi University Faculty of Economics and Administrative Sciences.

Results

First of all, it has been observed that the more the number of options of the participation level, the more the increase in the internal consistency coefficient; however, that this increase indicated a statistically significant change in 5 and 11 number of options of the participation level. It is possible to infer in a matter of speaking that the number of options of the participation level may cause some differences in psychometric scale.

Second of all, the findings regarding the fact that the covariance matrices among the latent phenomenons (variables) such as attitude and tendency which has been the aim to be tested via psychometric scales have differentiated in a systematically and statistically significant way depending on the number of options in participation level and that the correlation coefficients have also differentiated systematically but has not differentiated in a statistically significant way has been reached within the scope of the limitations of this research. In other words, the choice to be made by the researcher between 5, 7 or 11 participation level option count poses a statistically significant difference in the covariance matrices among the researched variables.

Finally, no systematic relation has been observed between the number of participation level options and the normal distribution of data. In other words, the data has not been able to procure normal distribution structure in three different counts of participation level (5, 7, and 11). It has also been observed that the data does not have a tendency depending on the count of participation level within the scope of skewness and kurtosis levels.

Conclusion

In conclusion, this research suggests 11 as the number of options of the participation level for the researchers as optimal number of options of the participation level. Weijters et al. (2010) suggested an increased number for the number of options of the participation level for student participation and 5 as the number of options of the participation level for general sampling. In this sense, the results of this research and the research conducted by Weijters et al. (2010) display coherence.