



An Investigation of Item Order and Rating Differences in Likert-Type Scales in Terms of Psychometric Properties and Attitudes of Respondents*

Gizem UYUMAZ**

Ömay ÇOKLUK***

Received: 27 February 2015

Accepted: 29 July 2015

ABSTRACT: In social sciences, usability of findings obtained by administration of a scale, accuracy and coherence of decisions made are closely related to psychometric properties of a scale. This study aims to investigate whether psychometric properties and attitudes of respondents differ on the number of response category, verbal or numerical rating and item distribution. To this end, related properties of a scale were changed and different versions of the same scale were used. According to the results of the study, attitudes of respondents did not significantly differ on the original, positive and negative forms of the scale, but significantly differed on the verbal and numerical forms of the scale. Gathering all the items, which are either positive or negative, under the same scale did not cause any response setting. The Cronbach alpha reliability coefficients of each version of the same scale were compared and it was found out that the increase in the number of response category caused increase in reliability of the scale. But, this increase was not statistically significant.

Keywords: Rating differences, Likert-type scales, item order, number of response categories, response setting

Extended Abstract

Purpose and Significance: In social sciences, usability of findings obtained by administration of a scale, accuracy and coherence of decisions made are closely related to psychometric properties of a scale. In this study, the factors that affect the reliability and validity of Likert scales were investigated. These factors can be dealt with under four headings: 1) Response set (number of categories, the numerical or verbal expression of ratings, extremity of response, etc.), 2) Carelessness, 3) Social desirability, 4) The wording of items (trueness, clearness, simplicity and clarity, moderate expressions, positive-negative balance, the layout of items, common stem, etc.) (Oskamp, 1977; Tezbaşaran, 1997; Tavşancıl, 2010). This study aims to investigate whether psychometric properties and attitudes of respondents differ on the number of response category, verbal or numerical rating and item distribution. To this end, related properties of a scale were changed and different versions of the same scale were used.

* This study is a part of master thesis conducted at 2013 at Institute of Educational Sciences, Ankara University.

** Corresponding Author: Res. Assist., Akdeniz University, Antalya, Turkey, gizemuyumaz@akdeniz.edu.tr

*** Assoc. Prof. Dr., Ankara University, Ankara, Turkey, cokluk@education.ankara.edu.tr

Citation Information

Uyumaz, G., & Çokluk, Ö. (2016). Likert tipi ölçeklerde madde düzeni ve derecelendirme farklılıklarının psikometrik özellikler ve yanıtlayıcı tutumları açısından incelenmesi. *Kuramsal Eğitim Bilim Dergisi [Journal of Theoretical Educational Science]*, 9(3), 400-425.

Methods: In the examinations, which will be made for this purpose, the forms of “The Attitude Scale toward School” developed by Gülleroğlu (2008), which were prepared by differentiating its related properties, were used. The Attitude Scale toward School consisting of 33 items in order to measure the attitudes of the students toward their universities was used in the research.

In the point of this general goal, some changes were made the Attitude Scale Toward School and different forms of these scale were prepared. In addition to the original form of the scale in which all the items were ordered randomly, two forms were prepared by changing the order of the items. In the first prepared form, all the positive items were taken the first part of the scale. In the second form, the vice versa happened and all the negative items were taken the first part and the positive ones were put after the negative ones. Then two extra forms of the scales were prepared. In the third form, the phrases were used in scaling. In the last form, the numbers were used in scaling. Also, the number based scaling forms were prepared as four different ways. The number of categories was changed to four, five, eight and nine.

After composing of these forms, the descriptive statistics, reliability analysis, validity proves, the students' total scores gained from the applications, the changes in the attitudes of the students and the response setting of the students were examined.

The study has been conducted as a fundamental research. The study group comprised of 547 students studying at different departments of Faculty of Educational Sciences of Ankara University in the 2nd Term of 2012-2013 academic years. There were 350 female and 157 male students and 318 junior and 229 senior students in the study group.

In the data collecting process, this scale was applied three times within the period of 15 days. The descriptive statistics and reliability coefficients of the results obtained through the application of the different forms of the scale were calculated. Significance of the difference between reliability coefficients of the different forms were calculated by using Fisher's Z Test. Also, item total correlation coefficients of the three forms were estimated and confirmatory factor analysis was done for the three forms. The relationship between the students' total scores calculated from the different applications of the scale was analyzed by using Pearson Correlation Coefficient.

One-way ANOVA for repeated measures and paired sample t-test were used to determine the significance of difference between the means of different forms. Paired Sample t-test was used to define whether response setting is a questioning the performances of the students.

Results: According to the results of the study, attitudes of respondents did not significantly differ on the original, positive and negative forms of the scale, but significantly differed on the verbal and numerical forms of the scale. Gathering all the items, which are either positive or negative, under the same scale did not cause any response setting. The Cronbach alpha reliability coefficients of each version of the same scale were compared and it was found out that the increase in the number of response

category caused increase in reliability of the scale. But, this increase was not statistically significant.

Discussion and Conclusions: Within the framework of the limitation of the study, according to the results, the obtained findings were as follows:

1. The differences made in the number of response categories, the type of scaling, using numbers or phrases, and the order of items affects the reliability and validity of Likert scales. But the differences in the coefficients alpha is not statistically significant.
2. The Cronbach Alpha internal consistency coefficient which were calculated for the original form of the scale in which all the items were ordered randomly is bigger than the Cronbach Alpha coefficient which were calculated for the two forms were prepared by changing the order of the items.
3. The Cronbach Alpha internal consistency coefficient which were calculated for the form of the scale in which the phrases were used in scaling is bigger than the Cronbach Alpha coefficient which were calculated for the form in which the numbers were used in scaling.
4. Increasing the number of the categories of the scale enables to obtain higher reliability coefficients.
5. The difference between the means of original scale form and positive and negative scale is not significant. But the difference between the means of original scale form and verbal and numerical scale form is statistically significant.
6. The difference between the means of verbal scale form and numerical scale form is statistically significant. It is due to confounding effect of the meanings of numbers attributed by the individuals.

Gathering all the items which are negative or positive together does not cause response setting.

Likert Tipi Ölçeklerde Madde Düzeni ve Derecelendirme Farklılıklarının Psikometrik Özellikler ve Yanıtlayıcı Tutumları Açısından İncelenmesi*

Gizem UYUMAZ**

Ömay ÇOKLUK***

Makale Gönderme Tarihi: 27 Şubat 2015 Makale Kabul Tarihi: 29 Temmuz 2015

ÖZ: Sosyal bilimlerde, uygulanan ölçekle elde edilen sonuçların kullanılabilirliği, verilen kararların doğruluğu ve anlamlılığı ölçme aracının psikometrik özellikleri ile yakından ilgilidir. Bu çalışmada Likert tipi ölçeklerin geçerlik ve güvenilirliğini etkileyen etmenlerden kategori sayısı, derecelendirmenin sayısal ya da sözel ifade edilmesi ve maddelerin ölçek içindeki dağılımının ölçeğin psikometrik niteliklerini ve yanıtlayıcıların tutumlarını nasıl değiştirdiğinin incelenmesi amaçlanmıştır. Bu amaçla yapılacak incelemelerde seçilen bir ölçeğin ilgili özellikleri farklılaştırılarak oluşturulan formları kullanılmıştır. Bu çalışma grubundan elde edilen verilerde ölçeğin orijinal formu ile olumlu ve olumsuz formları arasında yanıtlayıcı tutumları bakımından manidar bir fark bulunmazken, sözel ve sayısal formlar arasındaki fark manidardır. Olumlu ya da olumsuz maddelerin ölçek içerisinde bir araya toplanmasının yanıtlayıcılarda tepki kurulumuna neden olmadığı görülmektedir. Ayrıca çalışmada ele alınan formlarda Cronbach alfa iç tutarlılık katsayıları karşılaştırılmış ve yapılan analizler sonucunda tüm karşılaştırmalarda kategori sayısı arttıkça ölçeğin güvenirliliğinin arttığı sonucuna ulaşılmıştır. Fakat bu artış istatistiksel olarak manidar değildir.

Anahtar kelimeler: Derecelendirme farklılıkları, Likert tipi ölçek, madde düzeni, kategori sayısı, tepki kurulumu.

Giriş

Sosyal bilimlerde değişkenlerin ölçülmesinde kullanılan ölçme araçlarının geliştirilmesi sırasında başvurulan ölçekleme teknikleri ile; araçların anlamlı ölçümler yapmasının sağlanması veya standart ölçümler vermeyen nicelendirmelerle toplanan verilere standart ölçüm nitelikleri kazandırılması amaçlanmaktadır (Turgut ve Baykul, 1992). Ölçekleme teknikleri Crocker ve Algina (2008) tarafından genel olarak iki grupta ele alınmaktadır. İlki psikolojik boyut üzerine madde yerleştirmedir; denek tepkilerine (Bogardus ve Guttman teknikleri) ve yargıcı kararlarına dayalı (Thurstone tipi ölçek geliştirme tekniği) olarak iki başlık altında incelenir. İkincisi ise psikolojik boyut üzerine birey yerleştirmedir ve yine denek tepkilerine dayalı (Likert tipi ölçekler, kontrol listeleri ve Osgood ölçekleri) ve yargıcı uyumlarına dayalı (gözlem yoluyla gerçekleşen ölçmeler, dereceleme ölçekleri) olmak üzere iki başlık altında incelenir.

Ölçekleme tekniklerinden en sık kullanılanı Likert'in 1932 yılında geliştirdiği toplamalı sıralama tekniği de denilen dereceleme toplamları (summated rating scale) ölçeğidir. Geliştirilmesinin diğer ölçeklere göre daha kolay olması, yüksek güvenilirlik ve geçerliğin sağlanabilmesi, kullanışlılığının yüksek olması ve bu ölçeklerin çeşitli psikolojik objelere uyum sağlayabilmesi sık kullanılmasının nedenlerindedir (Erkuş, 2003; Özgüven, 2011; Tavşancıl, 2010; Tekindal, 2009).

* Bu çalışma Ankara Üniversitesi Eğitim Bilimleri Enstitüsünde 2013 yılında yürütülen yüksek lisans tezinin bir bölümüdür.

** Sorumlu Yazar: Araş. Gör., Akdeniz Üniversitesi, Antalya, Türkiye, gizemuyumaz@akdeniz.edu.tr

*** Doç. Dr., Ankara Üniversitesi, Ankara, Türkiye, cokluk@education.ankara.edu.tr

Likert tipi ölçeklerde geçerlik ve güvenilirliği etkileyen etmenler dört başlık altında ele alınabilir: 1) Tepki seti (kategori sayısı, derecelendirmenin sayısal ya da sözel ifade edilmesi, derecelendirmenin aşırı uçluluğu vb.), 2) Dikkatsizlik, 3) Sosyal beğenirlik, 4) Madde yazımı (olgusallık, açıklık, sadelik ve anlaşılabilirlik, ılımlı ifadeler, olumlu-olumsuz dengesi, maddelerin dağılımı, ortak kök vb.) (Oskamp, 1977; Tavşancıl, 2010; Tezbaşaran, 1997). Bu çalışmada bu etmenlerden tepki setiyle ilişkili olan “kategori sayısı” ve “derecelendirmenin sayısal ya da sözel ifade edilmesi” ve madde yazımıyla ilişkili olan “maddelerin dağılımı” ile ilgili farklılıklar incelendiğinden aşağıda bu etmenlere ilişkin açıklamalara yer verilmiştir.

Tutum ölçeklerinin güvenilirlik ve geçerliğini etkileyen temel problemlerden biri tepki setidir (Oskamp, 1977). Tepki seti başlığı altında ilk olarak “kategori sayısı” ele alındığında; orijinal Likert tipi ölçek, “tamamen katılıyorum” ile “hiç katılmıyorum” arasında derecelendirilen beş tepki kategorisi içerir. Daha sonra geliştirilen Likert ölçeklerinde ikili, üçlü, dörtlü, altılı ve yedili tepki kategorileri de kullanılmıştır (Anderson, 1988b); ancak en ideal kategori sayısı beştir (Erkuş, 2003).

Tepki kategorilerinin sayısını artırmadaki temel amaçlardan biri, yanıtlayıcılara verilen tepki fırsatlarının sayısını artırarak ölçeğin iç tutarlığını yükseltmektir (Köklü, 1997). Dereceleme türü ölçeklerde (ölçülecek özelliğe bağlı olmakla birlikte) bir yere kadar kategori sayısı azaldıkça ölçmenin duyarlılığı azalırken, arttıkça ölçmenin duyarlılığı da artar. Bir yerden sonra ise kategoriler arası ayrımların güçleşmesi yada ölçeğin sınıflama düzeyine inmesi (bilgi kaybı) gibi sorunlar doğmaya başlar (Erkuş, 2003, 2012). Güvenirlikteki artış, yedi tepki kategorisine kadar önemli ölçüdedir ancak tepki kategorisi yedinin üzerine çıktıkça, güvenilirlikteki artış önemini kaybeder. Bu nedenle de yedinin üzerinde tepki kategorisi pek sık kullanılmaz çünkü hem anlamlı tepki kategorisi yazmak güçleşir hem de birey kendine uygun bir tepki kategorisi bulmakta zorlanır. Bunun yerine, güvenilirliği yükseltmek için madde sayısını artırmak daha kolaydır (Thorndike, 1997).

Likert tipi ölçeklerde kategorilere ilişkin bir diğer tartışma, kategori sayısının tek mi, çift mi olması gerektiğidir. Çift sayıda kategori kullanma düşüncesi, yanıtlayıcıların “kararsızım” seçeneğine yönelip, gerçek bir seçim yapmaktan kaçınabilecekleri düşüncesi ile ortaya çıkmıştır (Anderson, 1988b). Bireylerin “kararsızım” kategorisine yönelmesinin engellenmesi ve uçlara doğru bir tercihe zorlanmaları istenirse, kategori sayısı dört, altı, sekiz gibi çift sayıda düzenlenebilir (Erkuş, 2003; Erkuş, 2012) fakat bireylerde bu kez de maddeyi boş bırakma eğilimi gözlenebilir. Bu nedenle tepki kategorisi sayısının tek olmasının daha iyi sonuç vereceği belirtilmektedir (Thorndike, 1997).

Tepki seti başlığı altında ele alınacak ikinci durum ise “derecelendirmenin sayısal ya da sözel ifade edilmesi”dir. Likert tipi ölçeklerde sözel seçeneklerin yerine, karşılıkları belirtilerek sayısal seçenekler kullanılabilir. Ancak bu tür bir düzenlemede de bireylerin sayılara yüklemiş oldukları anlamların karıştırıcı etki yapma olasılığı bulunmaktadır (Tezbaşaran, 1997).

Çalışmada ele alınan bir diğer durum Likert tipi ölçeklerde madde yazımıyla ilişkili olarak “maddelerin ölçek içerisindeki dağılımı”dır. Likert tipi ölçeklerde yer alan ifadeler, olumlu ve olumsuz olmak üzere iki türlü yapılandırılır. Olumlu ifade istenen bir durumu, olumsuz ifade ise istenmeyen bir durumu belirtir. Olumlu ve olumsuz durumların, eşit sayıda madde ile ifade edilmesi önerilir (Anderson, 1988b). Boş bırakma tepkisinden veya kalıp yargılara dayalı tepkilerden kaçınmak için, ölçekte yer alan olumlu ve olumsuz madde sayısı eşit olacak şekilde düzenlenmelidir (Tezbaşaran, 1997; Turgut & Baykul, 1992).

Olumlu ve olumsuz ifadelerle ilişkin vurgulanması gereken bir diğer nokta, kategori etiketleri aynı yönde olan ifadelerin alt alta gelmesinden kaynaklanan “tepki kurulumu”dur (Erkuş, 2012). Ölçekteki ifadelerin çoğunun yada tümünün olumlu olmasının, yanıtlayıcıyı olumlu yanıtlamaya, olumsuz olmasının ise olumsuz yanıtlamaya yönlendirici etki yapma olasılığı vardır. İfadelerin anlam yükünün yanıtlayıcıyı yönlendirici bir etkide bulunmasının önüne geçmenin bir yolu, ölçekteki olumlu ve olumsuz ifade sayısını birbirine denk tutmaktır (Tezbaşaran, 1997). Buradaki amaç, tüm maddelerin aynı şekilde (bütün maddelere “katılıyorum” gibi) yanıtlanma eğilimini (acquiescence) kontrol etmektir (Thorndike, 1997; akt: Sünbül, 2006). Bu nedenle, olumlu ve olumsuz anlam yüküne sahip maddeler ölçek içinde seçkisiz bir biçimde dağıtılmalı, bir düzen ya da sıra izlememeli ve gruplandırılmamalıdır (Tezbaşaran, 1997).

Yukarıdaki tartışmalar ve bu çalışmanın da genel amacı ile ilişkili olarak, alanyazında yer alan Likert tipi ölçeklerde derecelendirme farklılıklarının ölçeğin psikometrik özelliklerini nasıl etkilediğinin araştırıldığı çalışmalar incelenmiş ve farklı bulgulara rastlanmıştır. Geçerliğin incelendiği çalışmalardan; Jacoby ve Matell (1971) ve Masters (1974)’ın çalışmalarında derecelendirme farklılıklarının ölçeğin geçerliğini etkilediği, Erkuş, Sanlı, Bağlı ve Güven (2000) ve Leung (2011)’un çalışmalarında ise, derecelendirme farklılıklarında geçerliğin benzer olduğu belirlenmiştir. Kan (2009)’ın çalışmasında kategori sayısı arttıkça geçerlik kanıtlarının daha iyi olduğu görülürken, Atılğan ve Saçkes (2004)’in çalışmasında ise durum tam tersidir.

Güvenirliğin incelendiği çalışmalardan; Garland (1990)’ın çalışmasında, beşli derecelendirmenin sözel ifadelendirilmiş (labelled) ve sayısal ifadelendirilmiş (numbered) formlarının güvenilirlik katsayıları arasında manidar bir fark bulunmamıştır. Bazı çalışmalarda kategori sayısı arttıkça ölçeğin güvenilirliğinin arttığı sonucuna ulaşılmış (Atılğan & Saçkes, 2004; Masters, 1974; Tate, Simpson, Soo, & Lane-Brown, 2011; Weng, 2004), diğer bazı çalışmalarda ise, kategori sayısındaki farkın güvenilirlikteki etkisinin manidar olmadığı saptanmıştır (Brown, Wilding, & Coulter, 1991; Erkuş, Sanlı, Bağlı, & Güven, 2000; Jacoby & Matell, 1971; Leung, 2011).

Likert tipi ölçekler, ifadelerin bir boyut üstünde yerleştirildiği (Anderson, 1988a), bireylerin bir dizi maddeye verdikleri tepkilere dayalı çıkarımlar yapmaya olanak sağlayan (Anderson, 1988a), doğrudan çıkarımsal bir ölçme yapan (Ajzen, 2005), psikolojik boyut üzerine birey yerleştirmeye (Crocker & Algina, 2008) ve denek tepkilerine dayanan, sıralama düzeyinde ölçeklerdir (Erkuş, 2003). Bireylerin ölçülen

psikolojik özelliğe ne derece sahip oldukları ile ilgili bilgi toplamak ve elde edilen bilgiler doğrultusunda bireyler hakkında çeşitli kararlar vermek amacıyla oldukça sık kullanılan ölçme araçlarıdır. Uygulanan ölçkle elde edilen sonuçların kullanılabilirliği, verilen kararların doğruluğu ve anlamlılığı söz konusu ölçme aracının psikometrik özellikleri ile yakından ilgilidir. Geçerlik ve güvenilirliği düşük ölçkler, bireyler hakkında varılan kararların yanlış olmasına sebep olmaktadır.

Ölçeklerde geçerlik ve güvenilirliğin sağlanması için ölçülecek özelliğin tanımlanmasından başlayarak madde yazımı, nihai forma madde seçimi, maddelerin ölçek içindeki düzeni ve son olarak derecelendirmenin belirlenmesine kadar dikkat edilmesi gereken noktalar vardır. Ancak alanyazında farklı amaçlarla geliştirilen ölçklerin birçoğunda, sadece istatistiksel kanıtların (faktör analizi sonucu, madde-toplam korelasyonları vb.) ortaya konduğu ve bunların yeterli görüldüğü, oysa ölçklerin daha başlangıçta uygun olmayan ifadeler içerecek şekilde hazırlanması, maddelerin ölçek içindeki düzeninin uygun olmaması ya da buna hiç dikkat edilmemesi, belirlenen derecelendirme türü ve sayısının maddeler ve yanıtlayıcı özelliklerine uygun olmaması sıklıkla rastlanan sorunlar arasında yer almaktadır. Bu tür hatalar ölçklerin geçerlik ve güvenilirliklerini düşürmekte, bu da doğal olarak bireyler hakkında yanlış kararların verilmesine neden olmaktadır.

Bu çalışmada Likert tipi ölçklerin geçerlik ve güvenilirliğini etkileyen etmenler araştırılmıştır. Bu etmenlerden kategori sayısı, derecelendirmenin sayısal ya da sözel ifade edilmesi ve maddelerin ölçek içerisindeki dağılımının, ölçeğin psikometrik niteliklerini nasıl değiştirdiği Okula Yönelik Tutum Ölçeği'nin ilgili özellikleri farklılaştırılarak oluşturulan formları üzerinde karşılaştırmalı olarak incelenmiştir. Bir diğer deyişle araştırmanın problemini Likert tipi ölçklerde maddelerin ölçek içerisindeki dağılımı ve derecelendirmenin farklı şekillerde yapılandırılmasının, ölçeğin psikometrik özelliklerini ve yanıtlayıcıların ilgili tutum nesnesine ilişkin tutumlarını nasıl değiştirdiğinin incelenmesi oluşturmaktadır.

Ölçme ve değerlendirmede bireyler hakkında bir karara varılması amaçlanır. Bu kararın doğru olması için öncelikle toplanan bilginin doğru olması gerekir. Toplanan bilginin doğruluğu ise ölçme aracının geçerli ve güvenilir olmasına bağlıdır. Ölçeklerden elde edilecek puanların geçerlik ve güvenilirliğin sağlanması için ölçülecek özelliğin tanımlanmasından başlayarak madde yazımı, nihai forma madde seçimi, maddelerin ölçek içindeki düzeni ve son olarak derecelendirmenin belirlenmesine kadar dikkat edilmesi gereken noktalar vardır. Araştırmacı tarafından yapılan taramalarda hem yurtiçi, hem de yurtdışında maddelerin ölçek içindeki düzeninin ölçeğin psikometrik özellikleri üzerindeki etkisinin incelendiği bir çalışmaya rastlanmamıştır. Bu özelliği ile çalışmanın alanyazına katkı sağlayacağı düşünülmektedir. Ayrıca derecelendirme farklılıklarının ölçeğin geçerlik ve güvenilirliğine etkisi üzerine yurtdışında yapılan pek çok araştırma olmasına rağmen (bazıları: Garland, 1990; Jacoby & Matell, 1971; Leung, 2011; Masters, 1974; Tate, Simpson, Soo, & Lane-Brown, 2011; Weng, 2004;) Türkiye'de Atılgan ve Saçkes (2004), Erkuş, Sanlı, Bağlı ve Güven (2000) ve Kan

(2009)'ın çalışmaları dışında bir çalışmaya rastlanmamıştır. Bu yönüyle de çalışma ülkemiz alanyazınına katkı sağlayacaktır.

Araştırma sonuçları, ileride geliştirilecek tutum ölçeklerinin madde düzeninin ve derecelendirme yapısının ölçeğin psikometrik niteliklerini arttırıcı yönde yapılandırılmasında yol gösterici olacaktır. Araştırma, geçerliği ve güvenilirliği yüksek Likert tipi tutum ölçeklerinin geliştirilebilmesi için nerelere dikkat edilmesi gerektiği konusunda da bilgi sağlaması bakımından önemlidir.

Amaç

Bu araştırmada Likert tipi tutum ölçeklerinde kategori sayısı ve maddelerin ölçek içerisindeki dağılımında yapılan değişiklikler ile derecelendirmenin sayısal ya da sözel ifade edilmesinin ölçeğin psikometrik özelliklerini ve yanıtlayıcıların tutumlarını nasıl değiştirdiğinin incelenmesi amaçlanmaktadır.

Bu genel amaç doğrultusunda aşağıdaki sorulara yanıt aranmıştır;

1. Likert tipi bir ölçeğin; olumlu ve olumsuz maddelerin seçkisiz dağıldığı (orijinal form), olumlu maddelerle başlayan (olumlu form) ve olumsuz maddelerle başlayan (olumsuz form) formlarının, her bir derecenin sözel (sözel form) ve sayısal ifadelendirildiği (sayısal form) formlarının, beşli-dörtlü, beşli-sekizli, beşli-dokuzlu ve dörtlü-sekizli Likert tipi derecelendirilmiş formlarının karşılaştırıldığı gruplarda;
 - a) Betimsel istatistikler nasıldır?
 - b) Güvenirlik kestirimleri nasıldır?
 - c) Geçerlik kanıtları nasıldır?
 - d) Alınan toplam puanlar arasındaki ilişki nasıldır?
2. Likert tipi bir ölçeğin;
 - a) Orijinal, olumlu ve olumsuz formlarının,
 - b) Sözel ve sayısal formlarının karşılaştırıldığı gruplarda yanıtlayıcı tutumları farklılaşmakta mıdır?
3. Likert tipi bir ölçeğin orijinal, olumlu ve olumsuz formlarında yanıtlayıcılarda tepki kurulumu oluşmakta mıdır?

Yöntem

Araştırma Modeli

Bu araştırma, bir ölçme aracının, maddelerin ölçek içindeki yeri ve derecelendirme yapısı bakımından farklı yapılandırılan formlarının psikometrik niteliklerini ve yanıtlayıcıların tutum farklılıklarını karşılaştırmayı amaçlayan bir temel araştırmadır.

Araştırma Grubu

Bu çalışmanın araştırma grubunu 2012-2013 Eğitim Öğretim Yılı Bahar Yarıyılında Ankara Üniversitesi Eğitim Bilimleri Fakültesi'nde Bilgisayar ve Öğretim Teknolojileri Öğretmenliği, Rehberlik ve Psikolojik Danışmanlık, Okul Öncesi Öğretmenliği, Sınıf Öğretmenliği, Sosyal Bilgiler Öğretmenliği ve Zihinsel Engelliler Öğretmenliği lisans programlarının üçüncü ve dördüncü sınıflarında öğrenim görmekte olan 547 öğrenci oluşturmaktadır. Öğrencilerin 350'si (%63.99) kadın, 157'si (%28.70) erkektir; 40 (%7.31) kişi ise cinsiyetini belirtmemiştir. Öğrencilerin 318'i (%58.14) üçüncü, 229'u (%41.86) dördüncü sınıfta öğrenim görmektedir.

Veri Toplama Araçları

Araştırmada veri toplamak amacıyla, Gülleroğlu (2008) tarafından geliştirilen "Okula Yönelik Tutum Ölçeği (OYTÖ)" kullanılmıştır. Ölçek beşli Likert tipi bir ölçektir ve derecelendirme "Tamamen Uygun", "Uygun", "Kararsızım", "Uygun Değil" ve "Hiç uygun değil" şeklindedir. OYTÖ, öğrencilerin üniversitelere ilişkin genel tutumlarının belirlenmesi amacıyla geliştirilmiştir. Ölçeğin 38 maddeden oluşan ön deneme formu 300 öğrenciye uygulanmıştır. Yapılan analizlerde ön deneme formunda üç madde dışındaki tüm maddelerin madde-toplam test korelasyonları oldukça yüksek ve 0.05 düzeyinde manidar bulunmuştur. Bu üç madde ölçekten çıkarılmıştır. Faktör analizi sonucunda aracın üç faktör altında toplanan 33 maddeden oluştuğu belirlenmiştir. Birinci faktör "okulu sevme" olarak adlandırılmıştır ve 15 maddeden (m1, m3, m6, m8, m10, m11, m13, m16, m17, m18, m19, m22, m26, m29, m30) oluşmaktadır. İkinci faktör "okulu benimseme" (m7, m12, m14, m15, m20, m21, m25, m28, m32), üçüncü faktör ise "okula ait olma hissini taşıma" (m2, m4, m5, m9, m23, m24, m27, m31, m33) olarak adlandırılmıştır ve 9'ar maddeden oluşmaktadır. Bu faktörlerin özdeğerleri sırasıyla 12.046, 1.952 ve 1.752'dir. Üç faktör birlikte toplam varyansın %48'ini açıklamaktadır. Ölçeğin tümüne ilişkin Cronbach-alfa iç tutarlılık katsayısı 0.94, alt boyutlara ilişkin katsayılar ise sırası ile 0.85, 0.84 ve 0.85'tir.

OYTÖ'nün seçilme nedeni üniversite öğrencilerine uygulanabilir olması ve deneme formunun oluşturulması sürecinde izlenen yolun ölçek geliştirme basamaklarına uygunluğudur. Öncelikle 84 kişilik bir gruptan okullarına ilişkin duygular ve düşüncelerini anlatan bir kompozisyon yazmaları istenmiştir. Bu kompozisyonlar analiz edilerek tutum ifadesi olabilecek olanlar tutum maddesi yazma kurallarına uygun olarak yazılmıştır. Bu maddeler; bilişsel, duyuşsal ve davranışsal boyutlar dikkate alınarak ve olumlu-olumsuz ifade sayısının eşit olmasına dikkat edilerek oluşturulmuştur. Hazırlanan maddeler beş ölçme değerlendirme, bir gelişim psikolojisi ve bir rehberlik ve psikolojik danışma alan uzmanı tarafından dil, kapsam ve psikometrik açıdan incelenmiştir. 19'u olumlu, 19'u olumsuz 38 tutum maddesi yazılmış, 311 kişiye uygulanmış ve gerekli analizler yapılarak ölçek oluşturulmuştur (Gülleroğlu, 2008).

Nihai ölçekte yer alan 33 maddenin 17 maddesi olumlu (m1, m2, m3, m4, m5, m6, m9, m11, m16, m17, m18, m23, m24, m26, m27, m30, m31), 16 maddesi ise olumsuzdur (m7, m8, m10, m12, m13, m14, m15, m19, m20, m21, m22, m25, m28,

m29, m32, m33). Ölçek maddelerinin 11'i bilişsel (m2, m7, m8, m13, m14, m16, m18, m20, m24, m28, m30), 17'si duyuşsal (m1, m3, m4, m5, m6, m9, m10, m11, m12, m15, m17, m19, m21, m26, m27, m29, m33) ve 5'i davranışsaldır (m22, m23, m25, m31, m32).

Verilerin Toplanması

Çalışmada verilerin toplanması amacıyla araştırma grubundaki öğrencilerden üç farklı grup oluşturulmuştur. Birinci gruba (155 kişi) OYTÖ'nün farklı şekillerde yapılandırılmış formlarından önce "orijinal form", on beş gün sonra "olumlu form" ve bundan on beş gün sonra da "olumsuz form" uygulanmıştır. İkinci gruba (159 kişi) önce "sözel form" (bu aynı zamanda formun orijinal halidir), on beş gün sonra "sayısal form/beşli Likert tipi form" ve bundan on beş gün sonra da "dokuzlu Likert tipi form" uygulanmıştır. Üçüncü gruba (233 kişi) önce beşli Likert tipi "orijinal form", on beş gün sonra "dörtlü Likert tipi form" ve bundan on beş gün sonra da "sekizli Likert tipi form" uygulanmıştır.

Verilerin Analizi

Çalışma kapsamında öncelikle derecelendirme farklılıklarından kaynaklanan puan ranjını standartlaştırmak için farklı formlardan alınan toplam puanlar yüzdelik puanlara çevrilmiştir. Sonra ilk alt amaç doğrultusunda OYTÖ'nün farklı şekilde düzenlenen formlarının betimsel özellikleri incelenmiş ve formlar arası karşılaştırmalar yapılmıştır. Her bir formun güvenilirliği için Cronbach-alfa iç tutarlılık katsayıları hesaplanmış ve güvenilirlik katsayıları arasında manidar farklılık olup olmadığı Fisher'in Z Testi ile incelenmiştir. Sonraki aşamada tüm formların madde-toplam korelasyonları hesaplanmış ve formlar arası karşılaştırmalar yapılmıştır.

Her bir formda aracın üç faktörlü orijinal yapısının doğrulanıp doğrulanmadığını belirlemek amacıyla doğrulayıcı faktör analizi (DFA) yapılmıştır. Karşılaştırılabilirliğin sağlanabilmesi için (yapıda herhangi bir değişiklik karşılaştırılabilirliği ortadan kaldıracığından), her bir formun DFA sonuçları, önerilen modifikasyonlardan hiçbirini yapılmaksızın aynen raporlanmıştır.

Bireylerin uygulamalardan aldıkları toplam puanlar arasındaki ilişkiyi belirlemek için ise Pearson korelasyon katsayısı kullanılmıştır.

Çalışmanın ikinci alt amacı doğrultusunda, ölçeğin orijinal formu ile oluşturulan formları arasında yanıtlayıcı tutumları açısından manidar farklılık olup olmadığını belirlemek amacıyla; orijinal, olumlu ve olumsuz formlardan elde edilen verilere ilişkili örneklem tek faktörlü varyans analizi, sözel ve sayısal formlardan elde edilen verilere ise ilişkili örneklem t-testi uygulanmıştır.

Çalışmanın üçüncü alt amacı doğrultusunda, ölçeğin olumlu ve olumsuz formlarının, olumlu ve olumsuz maddelerine ilişkin toplam puanlar ayrı ayrı hesaplanmıştır. Bir diğer deyişle bireylerin olumlu formun olumlu maddelerinden aldıkları toplam puan, olumlu formun olumsuz maddelerinden aldıkları toplam puan, olumsuz formun olumlu maddelerinden aldıkları toplam puan, olumsuz formun olumsuz maddelerinden aldıkları toplam puan hesaplanmıştır. Ölçek formlarına ait olumlu

maddelerden alınan toplam puanlar ve olumsuz maddelerden alınan toplam puanlar arasında manidar fark olup olmadığını belirlemek amacıyla ayrı ayrı ilişkili örneklem için t-testi uygulanmıştır.

Bulgular

OYTÖ'nün orijinal, olumlu ve olumsuz, sözel ve sayısal, beşli-dörtlü, beşli-sekizli, beşli-dokuzlu, dörtlü-sekizli Likert tipi derecelendirilmiş şekillerde düzenlenen formlarının tekrarlı uygulanması sonucunda elde edilen puanlara ilişkin bulgular ve yorumlar aşağıda sunulmuştur.

Araştırmanın İlk Alt Amacına İlişkin Bulgular

Çalışmanın ilk alt amacı Likert tipi bir ölçeğin; olumlu ve olumsuz maddelerin seçkisiz dağıldığı (orijinal form), olumlu maddelerle başlayan (olumlu form) ve olumsuz maddelerle başlayan (olumsuz form) formlarının, her bir derecenin sözel (sözel form) ve sayısal ifadelendirildiği (sayısal form) formlarının, beşli-dörtlü, beşli-sekizli, beşli-dokuzlu ve dörtlü-sekizli Likert tipi derecelendirilmiş formlarının karşılaştırıldığı gruplarda;

- a) Betimsel istatistikler nasıldır?
- b) Güvenirlilik kestirimleri nasıldır?
- c) Geçerlik kanıtları nasıldır?
- d) Alınan toplam puanlar arasındaki ilişki nasıldır? şeklindedir.

Formlara ilişkin betimsel istatistikler.

OYTÖ'nün farklı şekillerde düzenlenmiş formlarına ilişkin betimsel istatistikler Tablo 1'de sunulmaktadır.

Tablo 1'de sunulan betimsel istatistikler incelendiğinde; merkezi eğilim ölçülerini oluşturan aritmetik ortalama, medyan (ortanca) ve mod (tepedeğer) değerlerinin; olumsuz formda, beşli ve dokuzlu derecelendirilmiş formların karşılaştırıldığı gruplarda beşli formda, dörtlü ve sekizli derecelendirilmiş formların karşılaştırıldığı gruplarda sekizli formda kendi içinde birbirine yakın değerler aldığı görülmektedir.

Sözel ve sayısal formlarda, beşli ve dörtlü derecelendirilmiş formların karşılaştırıldığı gruplarda, beşli ve sekizli derecelendirilmiş formların karşılaştırıldığı gruplarda beşli formlarda, dörtlü ve sekizli derecelendirilmiş formların karşılaştırıldığı grupta sekizli formda ortalama ve medyanın birbirine oldukça yakın değerler aldığı, modun ise bu değerlerden düşük olduğu görülmektedir.

Olumlu formda ise ortalama ve medyan değerleri birbirine yakinken, modun bu değerlere göre yüksek olduğu görülmektedir. Her bir form için değerlerin birbirinden çok uzak olmadığı, dağılımın normal dağılımdan aşırı sapmadığı söylenebilir.

Değişim ölçülerinden ranj (dağılım genişliği), varyans ve standart sapma incelendiğinde; puanların ortalamadan olan ortalama uzaklıkları (standart sapma) beşli ve sekizli ile dörtlü ve sekizli Likert tipi formlarda bir miktar farklılaşmaktadır, sekizli Likert tipi formlardan elde edilen puanlar daha geniş bir dağılım göstermektedir. Diğer

karşılaştırmalarda standart sapma değerleri birbirine yakındır. Dağılım genişlikleri incelendiğinde, derecelendirme sayısı arttıkça formlardan elde edilen puanların dağılım genişliklerinin de arttığı görülmektedir.

Tablo 1

Okula Yönelik Tutum Ölçeği'nin Farklı Düzenlenmiş Formlarının Betimsel İstatistikleri

Betimsel İstatistikler	Orijinal	Olumlu	Olumsuz	Sözel	Sayısal	Beşli	Dörtlü	Beşli	Sekizli	Beşli	Dokuzlu	Dörtlü	Sekizli
N	74	74	74	139	139	77	77	78	78	113	113	65	65
Ortalama	77.99	79.67	78.59	70.17	67.52	67.23	72.22	68.88	64.48	76.09	61.19	70.24	63.41
Medyan	80.53	81.21	79.39	72.12	67.51	69.09	75.18	69.74	66.21	77.58	60.94	72.17	63.26
Mod	75.15	90.91	79.39	66.67	58.79	63.03	77.27	63.03	62.50	77.58	64.65	59.85	62.50
SS	13.22	11.99	12.00	15.62	16.12	17.92	16.54	16.78	19.13	11.72	13.35	17.17	20.07
Varyans	174.87	143.82	144.23	243.99	259.92	320.98	273.66	281.61	365.85	137.39	178.28	294.96	402.63
Ranj	59.39	53.33	46.84	70.91	77.58	74.55	65.91	72.73	87.12	55.76	60.48	74.24	87.50
Min	39.39	46.06	51.95	28.48	22.42	24.24	31.82	24.24	12.50	44.24	28.74	25.00	12.50
Max	98.79	99.39	98.79	99.39	100	98.79	97.73	96.97	99.62	100	89.23	99.24	100
Çarpıklık	-0.799	-0.528	-0.460	-0.564	-0.281	-0.570	-0.606	-0.651	-0.605	-0.288	-0.091	-0.506	-0.615
Basıklık	0.364	-0.101	-0.422	0.000	-0.117	-0.071	-0.315	-0.250	-0.275	-0.058	-0.324	-0.144	-0.536

En düşük maksimum puan dokuzlu derecelendirilmiş formdan elde edilmiştir. Maksimum puanlar genel olarak yüksek ve karşılaştırılan gruplarda çok büyük farklılık göstermezken, minimum puanların farklılaştığı görülmektedir. Derecelendirme sayısının az olduğu formdan elde edilen verilerde minimum puanlar daha yüksektir.

OYTÖ'nün farklı şekillerde düzenlenen formlarına ilişkin dağılım eğrileri incelenmiştir. Dağılım eğrileri betimsel istatistik değerleri de göz önünde bulundurularak yorumlandığında; tüm formların uygulanması sonucunda elde edilen puanlara ilişkin dağılımın normal dağılımdan aşırı sapmadığı, dağılımların genel olarak hafif sola çarpık olduğu, olumsuz formun ve sekizli formun dağılımının aynı zamanda hafif sivri olduğu söylenebilir (Howell, 2010; Howitt & Cramer, 1997; Mertler & Vannatta, 2005).

Formlara ilişkin güvenilirlik kestirimleri.

OYTÖ'nün farklı şekillerde düzenlenen formlarına ait Cronbach alfa güvenilirlik katsayıları Tablo 2'de verilmiştir.

Tablo 2

Okula Yönelik Tutum Ölçeği'nin Farklı Düzenlenmiş Formlarının Cronbach Alfa Güvenirlik Katsayıları

Karşılaştırma	Formlar	Toplam	OS	OB	OAHT
1.Grup	Orijinal Form	0.956	0.898	0.906	0.882
	Olumlu Form	0.954	0.892	0.857	0.844
	Olumsuz Form	0.945	0.891	0.813	0.800
2.Grup	Sözel Form	0.965	0.882	0.901	0.908
	Sayısal Form	0.962	0.896	0.910	0.900
3.Grup	Beşli Likert Tipi Form	0.965	0.929	0.883	0.933
	Dörtlü Likert Tipi Form	0.962	0.909	0.905	0.932
4.Grup	Beşli Likert Tipi Form	0.954	0.901	0.877	0.902
	Sekizli Likert Tipi Form	0.964	0.921	0.897	0.927
5.Grup	Beşli Likert Tipi Form	0.890	0.843	0.538	0.806
	Dokuzlu Likert Tipi Form	0.903	0.853	0.556	0.833
6.Grup	Dörtlü Likert Tipi Form	0.964	0.905	0.908	0.940
	Sekizli Likert Tipi Form	0.967	0.924	0.911	0.939

Hem toplam puanlara hem de alt faktörlere ait Cronbach alfa iç tutarlılık katsayıları incelendiğinde; orijinal, olumlu ve olumsuz, sözel ve sayısal, beşli ve dörtlü, beşli ve sekizli ile dörtlü ve sekizli ölçek formlarının karşılaştırıldığı gruplardan elde edilen verilerin güvenilirlik derecelerinin yüksek olduğu görülmektedir (Özdamar, 2004). Orijinal/sözel formlara ilişkin bu bulgular ölçeğin geliştirilme çalışması ile paralellik göstermektedir (Gülleroğlu, 2008).

Karşılaştırmalı gruplar ele alındığında, toplam ve alt faktörlere ilişkin güvenilirlik katsayıları; orijinal, olumlu ve olumsuz formların karşılaştırıldığı grupta orijinal formda en yüksek, olumsuz formda en düşük değerleri almaktadır. Sözel ve sayısal formların karşılaştırıldığı grupta toplam test için sözel forma ait güvenilirlik katsayısı daha yüksek olmasına rağmen iki alt faktörde sayısal forma ait güvenilirlik katsayısının daha yüksek olduğu görülmektedir.

Beşli ve dörtlü Likert tipi formlardan beşli Likert tipi forma ait güvenilirlik katsayıları genel olarak dörtlü Likert tipi forma göre daha yüksek bulunmuştur. Bu sonuç Erkuş, Sanlı, Bağlı ve Güven (2000), Leung (2011) ve Weng (2004)'in

çalışmalarında da desteklenmekte fakat Jacoby ve Matell (1971)'in çalışmasına ters düşmektedir.

Beşli ve sekizli Likert tipi formlardan sekizli Likert tipi forma ait güvenilirlik katsayıları ölçeğin toplamında ve tüm alt faktörlerde beşli Likert tipi forma göre daha yüksek bulunmuştur. Bu sonuç Jacoby ve Matell (1971) ile Weng (2004)'in çalışmalarında da desteklenmektedir.

Dörtlü ve sekizli Likert tipi formlardan sekizli Likert tipi forma ait güvenilirlik katsayıları genel olarak dörtlü Likert tipi forma göre daha yüksek bulunmuştur. Bu sonuç Weng (2004)'in çalışmasında da desteklenmektedir. Fakat Fisher'ın Z Testi ile yapılan analizde formların güvenilirlik katsayıları arasında manidar bir fark bulunmamıştır. Bu bulgu Jacoby ve Matell (1971)'in çalışmasında da desteklenmektedir.

Beşli ve dokuzlu Likert tipi formlarda ikinci alt faktörde güvenirliliğin düşük olduğu görülmektedir (Özdamar, 2004). Dokuzlu Likert tipi forma ait güvenilirlik katsayıları ölçeğin toplamında ve tüm alt faktörlerde beşli Likert tipi forma göre daha yüksek bulunmuştur. Bu sonuç Weng (2004)'in çalışmasında da desteklenmektedir.

Likert tipi ölçeklerde fazla sayıda tepki seçeneğinin kullanımı yanıtlayıcılara verilen toplam tepki fırsatlarının sayısını artırarak ölçeğin iç tutarlılığını yükseltmeyi amaçlar (Köklü, 1997). Bu çalışmada farklı kategori sayıları kullanılıyorsa da, en ideal kategori sayısı beştir. Dereceleme türü ölçeklerde bir yere kadar kategori sayısı arttıkça ölçenin duyarlılığı da artar; bir yerden sonra kategoriler arası ayırtedilemezlik ya da ölçeğin sınıflama düzeyine inmesi yani bilgi kaybı gibi sorunlar oluşmaya başlar (Erkuş, 2003; 2012). Güvenirlikteki artış, yeditepki kategorisine kadar önemli ölçüdedir. Tepki kategorisi sayısı yedinin üzerine çıktıkça, güvenirlikteki artış önemini kaybederçünkü birey kendine uygun tepki kategorisi bulmakta zorlanır (Thorndike, 1997). Fakat bu çalışmada Likert tipi ölçeklerde beşli ve dörtlü, beşli ve sekizli, beşli ve dokuzlu ile dörtlü ve sekizli derecelendirilmiş formlarda hem toplam puanlara hem de alt faktörlere ait Cronbach alfa iç tutarlılık katsayıları karşılaştırılmış ve yapılan analizler sonucunda tüm karşılaştırmalarda kategori sayısı arttıkça, ölçeğin güvenirliliğinin arttığı sonucuna ulaşılmıştır. Bu sonuç daha önce yapılan bazı çalışmalarda da desteklenmektedir (Atılğan & Saçkes, 2004; Tate, Simpson, Soo, & Lane-Brown, 2011; Weng, 2004).

İki güvenilirlik katsayısı arasındaki farkın manidarlığının test edilmesi, iki aritmetik ortalama, iki yüzde ya da diğer iki istatistik arasındaki farkın manidarlığının test edilmesinde izlenen yaklaşımla aynıdır; önce iki katsayı arasındaki fark bulunur, sonra farkın standart hatası hesaplanır ve farka bölünür (Akhun, 1984). Güvenirlik katsayılarının birbirinden manidar bir biçimde farklılaşıp farklılaşmadıkları Fisher'ın Z Testi ile belirlenmiştir. Güvenirlik katsayılarına ait Fisher'ın Z testi sonuçları Tablo 3'de verilmiştir.

Tablo 3

Okula Yönelik Tutum Ölçeği'nin Farklı Düzenlenmiş Formlarının Güvenirlik Katsayılarına Ait Fisher'in Z Değerleri

	Fisher Z değeri				Kritik değer
	Toplam	OS	OB	OAoHT	0.05
Orijinal-Olumlu	0.000	0.298	1.331	0.698	1.96
Orijinal-Olumsuz	0.614	0.296	2.112	1.639	1.96
Olumlu-Olumsuz	0.614	0.000	0.781	0.941	1.96
Sözel-Sayısal	0.561	0.587	0.463	0.463	1.96
Beşli Likert -Dörtlü Likert	0.414	0.579	0.616	0.238	1.96
Beşli Likert -Sekizli Likert	0.784	0.717	0.435	0.925	1.96
Beşli Likert -Dokuzlu Likert	0.135	0.119	0.164	0.678	1.96
Dörtlü Likert- Sekizli Likert	0.000	0.696	0.000	0.000	1.96

Tablo 3'de görüldüğü üzere yalnızca orijinal ve olumsuz formların karşılaştırıldığı formlarda ikinci alt boyuta ait Z değerleri kritik Z değerinden büyük olduğundan güvenirlilik katsayıları arasındaki fark 0.05 düzeyinde manidardır. Diğer karşılaştırmalarda güvenirlilik katsayıları arasında manidar fark yoktur.

Cronbach alfa güvenirlilik katsayılarının belirlenmesinin ardından OYTÖ'nün farklı şekillerde düzenlenen formlarına korelasyona dayalı madde analizi yapılmıştır. Formların Pearson korelasyon katsayısı ile hesaplanan madde-toplam test korelasyonları ve manidarlık düzeyleri belirlenmiştir.

Madde-toplam korelasyonları orijinal formda 0.371-0.863, olumlu formda 0.153-0.782 ve olumsuz formda ise 0.291-0.830 arasında değişmektedir. Olumlu ve olumsuz formlarda birer maddenin (sırasıyla M22 ve M32) madde-toplam korelasyonu 0.30'un altındadır. Bu maddeler dışındaki tüm korelasyonlar 0.001 düzeyinde manidardır ve olumlu ve olumsuz tutuma sahip bireyleri iyi ayırt edebilmektedir.

Sayısal formda, dörtlü, beşli ve sekizli Likert tipi formlarda madde-toplam korelasyonu 0.30'un altında olan madde bulunmamakta ve tüm korelasyonlar 0.001 düzeyinde manidar çıkmaktadır. Sözel formdaki (M19: Mezun olacağım günü iple çekiyorum.) ve dokuzlu Likert tipi formdaki (M20: Bu üniversitede okumaktansa lise mezunu olmayı tercih ederim.) birer madde dışında tüm maddeler olumlu ve olumsuz tutuma sahip bireyleri iyi ayırt edebilmektedir. Maddeler farklı derecelendirmeye sahip formlarda benzer ayırıcılık değerlerine sahiptir. Bu bulgu Erkuş, Sanlı, Bağlı ve Güven (2000) tarafından yapılan çalışma ile tutarlıdır.

"Mezun olacağım günü iple çekiyorum (M19)." maddesi uygulamalar sırasında da yön bakımından çok anlaşılır bulunmamıştır. Yapılan incelemelerde hem M19 hem de M20'de bireylerin büyük çoğunluğunun ilk kategoriye tercih ettiği görülmüştür. Bu

bulgu maddenin olumlu ve olumsuz tutuma sahip bireyleri iyi ayırt edemediğini göstermektedir.

Formlara ilişkin geçerlik kanıtları.

OYTÖ'nün farklı şekillerde düzenlenen formlarının geçerlik kanıtları Doğrulayıcı Faktör Analizi (DFA) yapılarak incelenmiştir. DFA önceden belirlenmiş ya da kurgulanmış bir yapının doğrulanması amacını taşımaktadır. Gizil değişkenler arasındaki ilişkileri betimleyen model ile elde edilen verinin ne oranda uyduğuna ilişkin ayrıntılı istatistikler sunmaktadır (Sümer, 2000).

Sonuçlar ele alındığında; olumlu ve olumsuz formlardaki birer maddenin (sırasıyla M25= "Elimde olsa bir gün bile bu üniversitede durmam." ve M31= "Üniversitemin konu olduğu herhangi bir tartışmada her zaman üniversitemi savunurum.") t değerleri manidar değildir. Bu iki madde dışında OYTÖ'nün farklı şekillerde düzenlenen formlarının tümünde bütün maddelere ait t değerleri 0.01 düzeyinde, p değeri 0.01 düzeyinde manidardır.

Standartlaştırılmış yük değerleri incelendiğinde; hem olumlu hem de olumsuz formlardaki birer maddenin (sırasıyla M25 ve M31), sözel ve sayısal formlarda aynı maddenin (M15= "Üniversitemin adını duymak bile beni huzursuz ediyor."), beşli ve sekizli Likert tipi formlardan beşli Likert tipi formdaki bir maddenin (M26= Benim için üniversitem çok değerlidir.), beşli ve dokuzlu Likert tipi formlardan dokuzlu Likert tipi formdaki bir maddenin (M20= Bu üniversitede okumaktansa lise mezunu olmayı tercih ederim.) standartlaştırılmış yük değeri 0.30'un altındadır.

Olumlu ve olumsuz formlara ilişkin karşılaştırmada birer madde (sırasıyla M25, M31), sözel ve sayısal formların her ikisinde de aynı madde (M15), beşli ve dördümlü Likert tipi formlara ilişkin karşılaştırmada beşli Likert tipi formdaki bir madde (M7= "Burada zamanımı boşa harcıyorum."), beşli ve sekizli Likert tipi formlara ilişkin karşılaştırmada beşli Likert tipi formdaki bir madde (M26), beşli ve dokuzlu Likert tipi formlara ilişkin karşılaştırmada dokuzlu Likert tipi formdaki bir madde (M20) 0.10'dan düşük varyans açıklamaktadır.

Gözlenen değişkenlerin hata varyansları incelendiğinde; olumlu formda bir (M25), olumsuz formda ise iki maddenin (M16, M31), sözel formda iki (M7, M15), sayısal formdaki bir maddenin (M15), beşli ve dördümlü Likert tipi formdaki birer maddenin (sırasıyla M7 ve M14), beşli ve sekizli Likert tipi formlardan beşli Likert tipi formdaki bir maddenin (M26), beşli ve dokuzlu Likert tipi formlardan dokuzlu Likert tipi formdaki iki maddenin (M15 ve M20) hata varyanslarının çok yüksek (>0.90) olduğu görülmektedir.

Formlarının geçerlik kanıtları incelendiğinde; orijinal, olumlu ve olumsuz formlar karşılaştırıldığında genel olarak en yüksek standartlaştırılmış yük değerleri olumlu formda, en düşük yük değerleri ise orijinal formda yer almaktadır. Hata varyans değerleri ise orijinal formda en yüksek, olumlu formda en düşük değerleri almıştır. Sözel formla karşılaştırıldığında sayısal formda, dördümlü Likert tipi formla karşılaştırıldığında beşli Likert tipi formda, beşli Likert tipi formla karşılaştırıldığında

sekizli Likert tipi formda, beşli Likert tipi formla karşılaştırıldığında dokuzlu Likert tipi formda, sekizli Likert tipi formla karşılaştırıldığında dördümlü Likert tipi formda standartlaştırılmış yük değerleri genel olarak daha yüksek, hata varyans değerleri ise daha düşüktür.

OYTÖ'nün farklı şekillerde düzenlenen formlarına ait uyum indeksi değerleri Tablo 4'de verilmiştir.

Tablo 4

Okula Yönelik Tutum Ölçeği'nin Farklı Şekillerde Düzenlenen Formlarının Uyum İndeksi Değerleri

İndeks	Uyum İndeks Değerleri												
	Orijinal	Olumlu	Olumsuz	Sözel	Sayısal	Beşli	Dördümlü	Beşli	Sekizli	Beşli	Dokuzlu	Dördümlü	Sekizli
χ^2/sd	2.44	2.37	2.13	3.98	6.25	2.12	2.15	3.21	3.66	2.44	4.99	1.87	3.66
RMSEA	0.14	0.14	0.12	0.15	0.20	0.12	0.12	0.17	0.20	0.14	0.19	0.12	0.20
GFI	0.50	0.51	0.53	0.54	0.43	0.55	0.54	0.45	0.37	0.50	0.43	0.53	0.37
AGFI	0.43	0.44	0.47	0.47	0.34	0.48	0.48	0.37	0.28	0.43	0.35	0.47	0.28
SRMR	0.10	0.11	0.10	0.10	0.12	0.08	0.09	0.11	0.13	0.10	0.12	0.09	0.13
NNFI	0.91	0.90	0.91	0.92	0.89	0.94	0.94	0.87	0.89	0.91	0.89	0.93	0.89
CFI	0.91	0.90	0.91	0.93	0.90	0.94	0.94	0.87	0.90	0.91	0.89	0.93	0.90

Tablo 4'de yer alan uyum indeksi değerleri literatürde yer alan kabul aralıklarına göre değerlendirilerek, formlara ait uyum düzeyleri oluşturulmuş ve Tablo 5'de verilmiştir.

Tablo 4 ve 5'de sunulan uyum indeks değerleri ve bunların karşılık geldiği uyum düzeyleri incelendiğinde; χ^2/sd oranı orijinal, olumlu ve olumsuz formların karşılaştırıldığı grupta her üçünde de, beşli ve dördümlü formların karşılaştırıldığı grupta ikisinde de, beşli ve dokuzlu formların karşılaştırıldığı grupta beşli formda, dördümlü ve sekizli formların karşılaştırıldığı grupta dördümlü formda mükemmel uyumu göstermektedir (Kline, 2005). Sözel formda ve beşli ve sekizli formların karşılaştırıldığı grupta iki formda da, beşli ve dokuzlu formların karşılaştırıldığı grupta dokuzlu formda, dördümlü ve sekizli formların karşılaştırıldığı grupta sekizli formda orta düzeyde uyumu göstermektedir (Sümer, 2000). Sayısal formda uyum orta düzey sınırının altındadır.

RMSEA, GFI ve AGFI değerleri tüm formlarda zayıf uyum/kabul edilebilir uyum kriterinin de altındadır (Tabachnick & Fidell, 2001; Yılmaz & Çelik, 2009). GFI ve AGFI değerleri örneklem büyüklüğüne çok duyarlıdır ve örneklem büyüklüğü

arttıkça daha uygun değerler verir (Tabachnick & Fidell, 2001). Değerlerin bu kadar düşük olmasının bir sebebi olarak örneklem büyüklüğünün küçük olması düşünülebilir.

Tablo 5

Okula Yönelik Tutum Ölçeği'nin Farklı Şekillerde Düzenlenen Formlarının Uyum Düzeyleri

İndeks	Uyum Düzeyi												
	Orijinal	Olumlu	Olumsuz	Sözel	Sayısal	Beşli	Dörtlü	Beşli	Sekizli	Beşli	Dokuzlu	Dörtlü	Sekizli
χ^2/sd	M	M	M	OD	Kötü	M	M	OD	OD	M	OD	M	OD
RMSEA	K	K	K	K	K	K	K	K	K	K	K	K	K
GFI	K	K	K	K	K	K	K	K	K	K	K	K	K
AGFI	K	K	K	K	K	K	K	K	K	K	K	K	K
SRMR	V	V	V	V	V	V	V	V	V	V	V	V	V
NNFI	İ	İ	İ	İ	KE	İ	İ	KE	KE	İ	KE	İ	KE
CFI	İ	İ	İ	İ	İ	İ	İ	KE	İ	İ	KE	İ	İ

M: Mükemmel, İ: İyi, OD: Orta Düzeyde, KE: Kabul Edilebilir, V: Vasat, K:Kötü

Standardize edilmiş RMR değerleri tüm formlarda vasat uyuma karşılık gelmektedir (Kline, 2005). NNFI değerleri; orijinal, olumlu ve olumsuz formlarda, sözel formda, beşli ve dörtlü formlarda, beşli ve dokuzlu formların karşılaştırıldığı grupta beşli formda, dörtlü ve sekizli formların karşılaştırıldığı grupta dörtlü formda iyi uyuma karşılık gelmektedir (Tabachnick & Fidell, 2001). Sayısal formda, beşli ve sekizli formların karşılaştırıldığı grupta iki formda da, beşli ve dokuzlu formların karşılaştırıldığı grupta dokuzlu formda, dörtlü ve sekizli formların karşılaştırıldığı grupta sekizli formda kabul edilebilir uyuma karşılık gelir (Yılmaz & Çelik, 2009).

CFI değerleri; orijinal, olumlu ve olumsuz formlarda, sözel ve sayısal formda, beşli ve dörtlü formların karşılaştırıldığı grupta ikisinde de, dörtlü ve sekizli formların karşılaştırıldığı grupta ikisinde de, beşli ve sekizli formların karşılaştırıldığı grupta sekizli formda, beşli ve dokuzlu formların karşılaştırıldığı grupta beşli formda iyi uyuma karşılık gelir (Tabachnick & Fidell, 2001). Beşli ve sekizli formların karşılaştırıldığı grupta beşli formda, beşli ve dokuzlu formların karşılaştırıldığı grupta dokuzlu formda kabul edilebilir uyuma karşılık gelir (Yılmaz & Çelik, 2009).

Analizler sonucu elde edilen uyum indeksi değerleri, standartlaştırılmış yükler, açıklanan varyans, hata varyansı ve t değerleri birlikte ele alındığında; orijinal, olumlu ve olumsuz formların karşılaştırıldığı grupta her üç form, beşli ve dörtlü Likert tipi formların karşılaştırıldığı grupta her iki form için de uyum indeksi değerlerinin aynı uyum düzeylerine karşılık geldiği, sözel formun uygulanmasıyla elde edilen puanlara ilişkin uyum indeksi değerlerinin sayısal forma göre daha yüksek uyum düzeyine karşılık geldiği, beşli ve sekizli Likert tipi formlardan sekizli Likert tipi formda, beşli ve

dokuzlu Likert tipi formlardan beşli Likert tipi formda, dördlü ve sekizli Likert tipi formlardan dördlü Likert tipi formda uyum indeksi değerlerinin daha yüksek düzeyde olduğu saptanmıştır.

Formlardan alınan toplam puanlar arasındaki ilişki.

OYTÖ'nün farklı şekillerde düzenlenen formlarında bireylerin uygulamalardan aldıkları toplam puanlar arasındaki ilişkiler Pearson korelasyon katsayısı hesaplanarak incelenmiştir. Katsayılar hesaplanmadan önce derecelendirme farklılıklarından kaynaklanan puan ranjını standartlaştırmak için puanlar yüzdelik puanlara çevrilmiştir. Formlardan alınan toplam puanlar arasındaki korelasyonlar Tablo 6'da verilmiştir.

Tablo 6

Pearson Korelasyon Katsayıları

Formlar	R	p
Orijinal - Olumlu	0.815	0.000
Orijinal - Olumsuz	0.807	0.000
Olumlu - Olumsuz	0.854	0.000
Sözel – Sayısal	0.806	0.000
Beşli Likert -Dördlü Likert	0.603	0.000
Beşli Likert - Sekizli Likert	0.652	0.000
Beşli Likert - Dokuzlu Likert	0.831	0.000
Dördlü Likert - Sekizli Likert	0.848	0.000

Tablo 6'da yer alan korelasyonlar incelendiğinde bireylerin; orijinal, olumlu ve olumsuz formlardan aldıkları toplam puanlar arasında yüksek düzeyde, pozitif ve manidar bir ilişki olduğu ($r=0.815$, $p<0.01$), ($r=0.807$, $p<0.01$), ($r=0.854$, $p<0.01$); sözel ve sayısal formlardan aldıkları toplam puanlar arasında yüksek düzeyde, pozitif ve manidar bir ilişki olduğu ($r=0.806$, $p<0.01$); beşli ve dördlü ile beşli ve sekizli Likert tipi derecelendirilmiş formlardan aldıkları toplam puanlar arasında orta düzeyde, pozitif ve manidar bir ilişki olduğu ($r=0.603$, $p<0.01$), ($r=0.652$, $p<0.01$); beşli ve dokuzlu ile dördlü ve sekizli Likert tipi derecelendirilmiş formlardan aldıkları toplam puanlar arasında ise yüksek düzeyde, pozitif ve manidar bir ilişki olduğu ($r=0.831$, $p<0.01$), ($r=0.848$, $p<0.01$) görülmektedir. Buna göre bireylerin tekrarlı uygulama yapılan formlardan aldıkları toplam puanların birbirine yakın olduğu ifade edilebilir.

Araştırmanın ikinci alt amacına ilişkin bulgular

Çalışmanın ikinci alt amacı Likert tipi bir ölçeğin;

- Orijinal, olumlu ve olumsuz formlarının,
- Sözel ve sayısal formlarının karşılaştırıldığı gruplarda yanıtlayıcı tutumları farklılaşmakta mıdır? şeklindedir.

OYTÖ'nün olumlu ve olumsuz maddelerin seçkisiz dağıldığı (orijinal form), olumlu maddelerle başlayan (olumlu form) ve olumsuz maddelerle başlayan (olumsuz form) formlarında yanıtlayıcı tutumlarının farklılaşp farklılaşmadığı incelenmiştir. Ölçeğin orijinal formu ile farklı düzenlenen formları arasında yanıtlayıcı tutumları bakımından manidar fark olup olmadığı belirlemek amacıyla formlardan elde edilen verilere ilişkili örneklem için tek faktörlü varyans analizi yapılmıştır. OYTÖ'nün orijinal, olumlu ve olumsuz formlarının ortalama ve standart sapma değerleri Tablo 7'de verilmiştir.

Tablo 7

Okula Yönelik Tutum Ölçeği'nin Orijinal, Olumlu ve Olumsuz Formlarının Ortalama ve Standart Sapma Değerleri

Tutum	<i>n</i>	\bar{X}	S
Orijinal Form	74	128.68	21.82
Olumlu Form	74	132.00	20.49
Olumsuz Form	74	129.67	19.82

OYTÖ'nün orijinal, olumlu ve olumsuz formlarına ait ANOVA sonuçları Tablo 8'de verilmiştir.

Tablo 8

Okula Yönelik Tutum Ölçeği'nin Orijinal, Olumlu ve Olumsuz Formlarına Ait Puanların Karşılaştırılması

Varyansın Kaynağı	Kareler Toplamı	sd	Kareler Ortalaması	F	<i>p</i>
Deneklerarası	82119.449	73	1124.924		
Ölçüm	429.780	2	214.890	2.628	0.076
Hata	11937.791	146	81.766		
Toplam	94487.020	221			

Tablo 8 incelendiğinde, bireylerin OYTÖ'nün orijinal, olumlu ve olumsuz formlarından aldıkları toplam puanlar arasında manidar bir fark olmadığı görülmektedir, $F_{(2,146)} = 2.628$, $p > 0.05$. Alanyazında olumlu ve olumsuz anlam yüküne sahip maddelerin ölçek içinde seçkisiz olarak dağıtılması, düzenli bir sıra izlememesi ve gruplandırılmaması gerektiği belirtilmektedir (Tezbaşaran, 1997). Maddelerin ölçek içindeki düzeni bireylerin tutumlarını etkileyen bir durumdur. Ancak bu çalışmadan elde edilen bulgular, maddelerin seçkisiz dağıtılması ile olumlu maddelerin yada olumsuz maddelerin bir arada sunulmasının yanıtlayıcı tutumlarını farklılaştırmadığını göstermektedir.

OYTÖ'nün her bir derecenin sözel ifadelendirildiği (orijinal form) ve her bir derecenin sayısal ifadelendirildiği (sayısal form) formlarında yanıtlayıcı tutumlarının farklılaşp farklılaşmadığı incelenmiştir. Ölçeğin orijinal formu ile düzenlenen formu arasında yanıtlayıcı tutumları bakımından manidar fark olup olmadığı belirlemek amacıyla; her bir derecenin sözel ifadelendirildiği (orijinal form) ve her bir derecenin sayısal ifadelendirildiği (sayısal form) formlardan elde edilen verilere ilişkili örneklem için t-testi yapılmıştır. OYTÖ'nün sözel ve sayısal formlarına ait t testi sonuçları Tablo 9'da verilmiştir.

Tablo 9

Okula Yönelik Tutum Ölçeği'nin Sözel ve Sayısal Formlarına İlişkin Puanların Karşılaştırılması

Ölçüm	n	\bar{X}	S	sd	t	p
Sözel	139	115.77	25.77	138	3.146	0.002
Sayısal	139	111.42	26.60			

Tablo 9 incelendiğinde, bireylerin OYTÖ'nün sözel ve sayısal formlarından aldığı toplam puanlar arasında manidar bir fark olduğu görülmektedir, $t_{(138)} = 3.146$, $p < 0.01$. Sözel form puanlarının ortalaması $\bar{X} = 115.77$ iken sayısal form puanlarının ortalaması $\bar{X} = 111.42$ 'dir. Literatürde her bir derecenin sözel ifadelendirildiği ve her bir derecenin sayısal ifadelendirildiği formlarda yanıtlayıcı tutumlarının farklılık göstereceği vurgulanmaktadır. Bunun nedeni bireylerin sayılara yüklemiş olduğu anlamların karıştırıcı etki yapmasıdır (Tezbaşaran, 1997). Bu çalışmada fark bulunmasına karşın Garland (1990)'ın yaptığı çalışmada sözel ve sayısal ifadelendirilmiş formlar arasında manidar fark bulunmamıştır.

Araştırmanın üçüncü alt amacına ilişkin bulgular.

Çalışmanın son üçüncü alt amacı, Likert tipi bir ölçeğin orijinal, olumlu ve olumsuz formlarında yanıtlayıcılarda tepki kurulumu oluşmakta mıdır? şeklindedir. OYTÖ'nün olumlu ve olumsuz maddelerin seçkisiz olarak dağıtıldığı (orijinal form), olumlu maddelerle başlayan (olumlu form) ve olumsuz maddelerle başlayan (olumsuz form) formlarında yanıtlayıcılarda tepki kurulumu olup olmadığı incelenmiştir. Yanıtlayıcılarda tüm olumlu maddeleri ya da tüm olumsuz maddeleri üst üste yanıtlamanın tepki kurulumuna neden olup olmadığını belirlemek amacıyla, bireylerin oluşturulan formlara ait olumlu ve olumsuz maddelerden aldıkları toplam puanlar ayrı ayrı hesaplanarak ilişkili örneklem için t-testi yapılmıştır. OYTÖ'nün olumlu ve olumsuz formlarından, bireylerin olumlu maddelerden aldıkları toplam puanlar ve olumsuz maddelerden aldıkları toplam puanları arasındaki farkın manidarlığı için yapılan t testlerine ilişkin sonuçlar Tablo 10'da verilmiştir.

Tablo 10

Okula Yönelik Tutum Ölçeği'nin Orijinal, Olumlu ve Olumsuz Formlarına Ait Puanların Karşılaştırılması

Toplam Puan	Ölçüm	n	\bar{X}	S	sd	t	p
Olumlu Maddeler	Olumlu Form	74	64.84	12.30	73	1.543	0.127
	Olumsuz Form	74	63.45	10.54			
Olumsuz Maddeler	Olumlu Form	74	67.16	10.04	73	1.325	0.189
	Olumsuz Form	74	66.22	10.33			

Tablo 10 incelendiğinde, bireylerin OYTÖ'nün hem olumlu formuna ait olumlu maddelerden aldıkları toplam puanlar ile olumsuz maddelerden aldığı toplam puanlar arasında ($t_{(73)} = 1.543$; $p > 0.05$) hem de olumsuz formuna ait olumlu maddelerden aldıkları toplam puanlar ile olumsuz maddelerden aldıkları toplam puanlar arasında manidar bir fark olmadığı sonucuna ulaşılmıştır, ($t_{(73)} = 1.325$; $p > 0.05$). Alanyazın incelendiğinde Likert tipi ölçeklerde en büyük sorunlardan birinin, kategori etiketlerinin aynı yönde ve altalta gelmesinden kaynaklanan tepki kurulumu (response setting) olduğu görülmektedir. Boş bırakma tepkisinden veya kalıp yargılara dayalı tepkilerden kaçınmak için ölçekle ölçülen yapının her iki yanını da temsil edecek (olumlu ve olumsuz) madde sayısı eşit olacak şekilde düzenlenmelidir (Erkuş, 2012; Tezbaşaran, 1997; Thorndike, 1997; Turgut & Baykul, 1992). Araştırmanın sınırlılıkları çerçevesinde, bu çalışma grubundan elde edilen verilerde olumlu ya da olumsuz maddelerin ölçek içerisinde biraraya toplanmasının ve yanıtlayıcının aynı yöndeki tüm maddeleri üst üste almasının tepki kurulumuna neden olmadığı görülmektedir. Bu durumun nedeni çalışma grubundaki birey sayısının az olması veya ölçekteki madde sayısının az olması nedeniyle 17 olumlu maddeyi üst üste yanıtlamanın bireylerin tutumunu olumlu ya da 16 olumsuz maddeyi üst üste yanıtlamanın bireylerin tutumunu olumsuzla çevirmek için yeterli olmaması olabilir.

Sonuç ve Tartışma

Bu çalışmada tutum ölçeklerinde maddelerin ölçek içerisindeki düzeninde ve derecelendirmede yapılan değişikliklerin ölçeğin psikometrik özelliklerine ve yanıtlayıcı tutumlarına etkisi incelenmiştir.

Bu doğrultuda Okula Yönelik Tutum Ölçeği'nin orijinal, olumlu ve olumsuz formları; sözel ve sayısal formları; dördü, beşli, sekizli ve dokuzlu derecelendirilmiş formları oluşturulmuş ve çalışmanın alt amaçları doğrultusunda gerekli incelemeler yapılmıştır.

Araştırma sonuçları değerlendirilirken; araştırmanın gerçek verilerle yapılmış olması, bu nedenle çalışma gruplarındaki birey sayılarının az olması nedeniyle homojen bir grup oluşmuş olması ve tek bir tutum ölçeğinin farklı formları kullanılarak karşılaştırmalar yapıldığı dikkate alınmalıdır. Bu sınırlılıklar değerlendirilerek,

bulguların tüm ölçeklere ve tüm araştırma gruplarına genellenmesinin doğru olmayacağı göz önünde bulundurulmalıdır. Aşağıda çalışmadan elde edilen sonuçlar maddeler halinde özetlenmiştir.

1. Likert tipi ölçeklerde maddelerin ölçek içindeki sıralamasının değiştirilmesi, derecelendirmenin sözel ya da sayısal şekilde ifadelendirilmesi ve derecelendirmedeki kategori sayısının değiştirilmesinin, istatistiksel olarak manidar bir fark yaratmamakla birlikte güvenilirlik katsayılarını farklılaştırdığı bulunmuştur. Benzer farklılıklar aynı zamanda ölçeğin geçerliği için de söz konusudur.
2. Olumlu ve olumsuz maddelerin ölçek içerisinde seçkisiz dağıtılması durumunda güvenilirlik katsayısı, olumlu ve olumsuz maddelerin biraraya toplandığı duruma göre daha yüksek değer almıştır. Ölçek geliştirme çalışmalarında maddelerin ölçek içerisindeki dağılımının seçkisiz yapılması, olumlu ya da olumsuz maddelerin biraraya toplanmamasına özen gösterilmesi gerekmektedir. Bu çalışmada maddelerin ölçek içerisinde seçkisiz dağıtılmasının ölçeğin geçerliğini etkilemediği sonucuna ulaşılmıştır.
3. Her bir derecenin sözel olarak ifadelendirildiği durumda güvenilirlik katsayısı, derecelerin sayısal ifadelendirildiği duruma göre daha yüksek değer almıştır. Benzer şekilde derecelendirmenin sözel olarak ifadelendirildiği forma ait geçerlik kanıtları, bu formun uyum değerlerinin daha yüksek uyum düzeyine karşılık geldiğini göstermektedir. Araştırmacıların ölçekteki dereceleri sözel olarak açık bir şekilde ifade etmesi, yanıtlayıcıları sayılara yüklenen anlamın karıştırıcı etkisinden kurtaracak ve ölçeğin psikometrik niteliklerini arttıracaktır.
4. İncelenen formlarla yapılan karşılaştırmalarda ölçekteki kategori sayısı arttıkça, ölçeğin güvenilirliğinin arttığı sonucuna ulaşılmıştır.
5. Ölçeğin orijinal formu ile olumlu ve olumsuz formları arasında yanıtlayıcı tutumlarının farklılaşp farklılaşmadığı incelendiğinde; maddelerin ölçek içerisindeki düzenin değiştirildiği formlarla orijinal form arasında, bireylerin aldıkları toplam puanlar açısından manidar bir fark bulunmamıştır. Bir başka deyişle, olumlu ve olumsuz maddelerin ölçek içerisinde seçkisiz dağıtılması ya da olumlu ve olumsuz maddelerin biraraya toplanması, bu çalışmada yanıtlayıcıların söz konusu tutum objesine yönelik tutumlarını etkilememektedir.
6. Ölçeğin sözel formu ile sayısal formu arasında yanıtlayıcı tutumlarının farklılaşp farklılaşmadığı incelendiğinde; derecelendirmenin sözel ve sayısal olarak yapılandırıldığı formlardan alınan toplam puanlar arasında manidar fark bulunmuştur. Derecelendirmenin sözel veya sayısal olarak yapılandırılmasının bireylerin sayılara yüklemiş olduğu anlamların karıştırıcı etkisi nedeniyle, yanıtlayıcı tutumlarını farklılaştırdığı görülmüştür. Bu farklılaşma sahip olunan tutum düzeyiyle ilgili değil, bireyin tutumunu kendisine verilen yanıt seçeneklerinden hangisiyle ifadelendirdiği ile ilgilidir. Bir başka deyişle, farklılaşan bireyin tutumu değil, tutumunu ifade ederken dereceleri seçme biçimidir. Derecelendirme sözel olarak ifade edildiğinde birey tutumunun hangi

aralığa geldiğini kolayca seçebilir, ancak sayısal ifadelerin tam karşılıklarını belirlemek birey için daha zordur.

7. Yanıtlayıcılarda tepki kurulumu oluşumu incelendiğinde; ölçeğin olumlu ve olumsuz formlarının, hem olumlu maddelerinden hem de olumsuz maddelerinden alınan toplam puanlar arasında manidar bir fark olmadığı, bir başka deyişle tüm olumlu maddeleri ya da tüm olumsuz maddeleri üst üste yanıtlamanın yanıtlayıcılarda tepki kurulumuna neden olmadığını sonucuna ulaşılmıştır.

Öneriler

1. Bu çalışmada bir tutum ölçeğinin dördü, beşli, sekizli ve dokuzlu Likert tipi derecelendirilmiş formlarına ait psikometrik özellikler incelenmiş ve karşılaştırmalar yapılmıştır. Benzer şekilde farklı sayıda derecelendirilmiş formların da psikometrik özellikleri incelenebilir.
2. Çalışmada yalnız derecelendirme sayısı değiştirilmiştir. Derecelendirme sayısı ile birlikte madde sayısı da değiştirilerek inceleme ve karşılaştırmalar yapılabilir.
3. Araştırma Likert tipi ölçeklerde geçerlik ve güvenilirliği etkileyen etmenlerden bazılarının (kategori sayısı, derecelendirmenin sayı ya da ifade oluşu ve maddelerin ölçek içerisindeki dağılımı) incelenmesiyle sınırlandırılmıştır. Bundan sonra yapılacak çalışmalarda diğer etmenlerin (derecelendirmenin aşırı uçluluğu, sosyal beğenirlik etkisi vb.) de ölçeğin psikometrik özelliklerini nasıl etkilediği araştırılabilir.
4. Çalışma, Likert tipi ölçekleme tekniği ile geliştirilmiş bir ölçek kullanılarak gerçekleştirilmiştir. Farklı ölçek tipleri için de derecelendirme ve madde düzeni farklılıklarının ölçeğin psikometrik özelliklerine etkisi incelenebilir.
5. Çalışmanın sınırlılıkları çerçevesinde, olumlu ya da olumsuz maddelerin ölçek içerisinde bir araya toplanmasının ve yanıtlayıcının aynı yöndeki tüm maddeleri birarada almasının tepki kurulumuna neden olmadığı görülmektedir. Aynı durum madde sayısının daha fazla olduğu ölçeklerde de incelenebilir. Ayrıca tutum objesinin daha uç tutumları ortaya çıkartacak bir obje olması, farkın belirginleşmesine neden olabilir.
6. Olumlu ve olumsuz maddelerin ölçek içerisinde seçkisiz dağıtılması durumunda güvenilirlik katsayısı daha yüksek değer aldığından, ölçek geliştirecek olan araştırmacıların maddelerin ölçek içerisinde seçkisiz dağıtılmasına dikkat etmeleri gerekir.
7. Her bir derecenin sözel olarak açıkça ifadelendirildiği durumda güvenilirlik katsayısı daha yüksek değer aldığından, ölçek geliştirecek olan araştırmacılar derecelendirmenin sözel olarak ifade edilmesine dikkat etmeleri yararlı olacaktır.
8. Çalışma sınırlı sayıda bireyden oluşan bir araştırma grubu ile yapılmıştır. Bu nedenle elde edilen sonuçların genellenebilirliği mümkün değildir. Genelleme yapılabilmesi için sistematik simülasyon çalışmaları ve meta analiz çalışmaları yapılması yararlı olacaktır.

Kaynakça

- Ajzen, I. (2005). *Attitudes, personality and behavior* (2th ed.). New York: Open University.
- Anderson, L. W. (1988a). *Attitudes and their measurement; educational research, methodology, and measurement: An nternational handbook*. Pergamon. Edited by: Keeves, J. P. (Ed), 421-426
- Anderson, L. W. (1988b). *Likert scales; educational research, methodology, and measurement: An international handbook*. Pergamon. Edited by: Keeves, J. P. (Ed), 227-228
- Akhun, İ. (1984). İki korelasyon katsayısı arasındaki farkın manidarlığının test edilmesi. *Ankara Üniversitesi Eğitim Bilimleri Fakültesi Dergisi*, 17(1), 1-7.
- Atılğan, H. & Saçkes, M. (2004). Ölçeklerin ikili ve çok kategorili puanlanmasının psikometrik özelliklerinin karşılaştırılması. *İnönü Üniversitesi Eğitim Fakültesi Dergisi*, 5(7).
- Crocker, L. & Algina, J. (2008). *Introduction to classical & modern test theory*. USA: Cengage Learning.
- Erkuş, A. (2003). *Psikometri üzerine yazılar*. Ankara: Türk Psikologlar Derneği.
- Erkuş, A. (2012). *Psikolojide ölçme ve ölçek geliştirme-1, temel kavramlar ve işlemler*. Ankara: PegemA.
- Garland, R. (1990). A comparison of three forms of the semantic differential. *Marketing Bulletin*, 1, 19-24.
- Gülleroğlu, H. D. (2008). Okula ilişkin tutum ölçeği. *Eğitim Bilimleri ve Uygulama Dergisi*, 7(13), 77-86.
- Kan, A. (2009). Effect of scale response format on psychometric properties in teaching self-efficacy. *Eurasian Journal of Educational Research*, 34, 215-228.
- Kline, R.B. (2005). *Principles and practice of structural equation modelling*. New York, NY: Guilford.
- Köklü, N. (1997). Tutumların ölçülmesi ve Likert tipi ölçeklerde kullanılan seçenekler. *Ankara Üniversitesi Eğitim Bilimleri Fakültesi Dergisi*, 28(2), 81-93.
- Leung, S-O. (2011). A comparison of psychometric properties and normality in 4-, 5-, 6-, and 11-point likert scales. *Journal of Social Service Research*, 37(4), 412-421.
- Mertler, C. A., & Vannatta, R. A. (2005). *Advanced and multivariate statistical methods*. California: Pyrczak.
- Oskamp, S. (1977). *Attitudes and opinions*. New Jersey: Prentice-Hall.
- Özdamar, K. (2004). *Paket programlar ile istatistiksel veri analizi (çok değişkenli analizler)*. Eskişehir: Kaan.
- Özgüven, İ. E. (2011). *Psikolojik testler*. Ankara: PDREM Yayınları.

- Sümer, N., (2000). Yapısal eşitlik modelleri: Temel kavramlar ve örnek uygulamalar. *Türk Psikoloji Yazıları*, 3(6), 49-74.
- Tate, R.L., Simpson, G.K., Soo, C.A., & Lane-Brown, A.T. (2011). Participation after acquired brain injury: Clinical and psychometric considerations of the sydney psychosocial reintegration scale (SPRS). *Journal of Rehabilitation Medicine*, 43(7), 609–618.
- Tavşancıl, E. (2010). *Tutumların ölçülmesi ve SPSS ile veri analizi* (4. baskı). Ankara: Nobel.
- Tabachnick, B. G. & Fidell, L. S. (2001). *Using multivariate statistics* (5th Ed.) USA: Pearson.
- Tekindal, S. (2009). *Duyuşsal özelliklerin ölçülmesi için araç oluşturma* (2.Basım). Ankara: PegemA.
- Tezbaşaran, A. (1997). *Likert tipi ölçek geliştirme kılavuzu*. Ankara: Türk Psikologlar Derneği.
- Thorndike, R. (1997). *Measurement and evaluation in psychology and education*. New Jersey: Prentice-Hall.
- Turgut, F. M., & Baykul, Y. (1992). *Ölçekleme teknikleri*. Ankara: ÖSYM.
- Weng, L-J. (2004). Impact of the number of response categories and anchor labels on coefficient alpha and test-retest reliability. *Educational and Psychological Measurement*, 64(6), 956-972.
- Yılmaz, V. & Çelik, H. E. (2009). *LISREL ile yapısal eşitlik modellemesi-1*. Ankara: PegemA.