

The Comparison Of Academic Success Of Students With Low And High Anxiety Levels In Tests Varying In Item Difficulty*

Çilem Doğan Gül

Hakkari University, Faculty of Education

Ömay Çokluk Bökeoğlu

Ankara University, Faculty of Educational Sciences

Abstract

The aim of this research is to study the effects of varying item order in multiple choice tests on the test scores of students with low and high anxiety levels and on item and test statistics according to their difficulty indice. The study is in accordance with casual comparative research since the comparison between varying item order and test scores of students with low and high anxiety levels and item and test statistics have been made. Data of the study were gathered from seven schools in four central districts of Ankara province. "Success Test on the Structure of Atom" developed by Tağ (2012) and "Test Anxiety Inventory" adapted into Turkish Culture by Öner (1990) were used in data gathering. As a result, it was observed that some item difficulty indices and item discrimination indices varied after the test items were randomly ordered from easy to difficult and from difficult to easy. In comparison between reliability coefficients among three forms, significant differences were found between the form ordered from easy to difficult and the one ordered from difficult to easy, and between the form ordered from easy to difficult and randomly ordered one.

Keywords: : Item order, multiple choice tests, test anxiety, academic achievement



Inönü University
Journal of the Faculty of Education
Vol 19, No 3, 2018
pp. 252-265
DOI: 10.17679/inuefd.341477

Received : 03.10.2017
Accepted : 19.07.2018

Suggested Citation

Doğan Gül, Ç. & Çokluk Bökeoğlu, Ö. (2018). The Comparison of Academic Success Of Students With Low And High Anxiety Levels In Tests Varying In Item Difficulty, *Inonu University Journal of the Faculty of Education*, 19(3), 252-265. DOI: 10.17679/inuefd.341477

*This study is a part of Çilem DOĞAN GÜL'S master dissertation under the supervision of Ömay ÇOKLUK BÖKEOĞLU

EXTENDED ABSTRACT

Introduction

Success tests are quite important in evaluating academic success and making decisions on students. It is essential to use tests with high validity, reliability and practicality in assessing academic success in education. In accordance with the aim of the test, these tests may be prepared by using different item formats. In all entrance exams centrally conducted in Turkey (high school entrance exam, university entrance exam), only multiple choice item format has been used for many years and it is still being used.

When the item order in a test is studied, there might be some problems such as the most successful candidates having difficulties answering the items, spending too much time on them and consequently the lack of reflection of the real performance on assessment process if one starts the test from the difficult items. Exposing an individual who starts answering a success test to difficult questions immediately may unavoidably cause the anxiety level to increase even more and the test performance to be negatively affected. Therefore, this situation will result in significant differences between the test performances of students with different anxiety levels (low and high test anxiety, etc.) (Ankara University, 2011).

Purpose

The aim of this research is to study the effects of varying item order in multiple choice tests on the test scores of students with low and high anxiety levels and on item and test statistics according to their difficulty indice.

Method

This research is in accordance with casual comparative model. Data of the research were applied to 578 seventh graders in science and technology lesson's success test related to "The Structure of the Atom" in three different forms in which the items were placed from easy to difficult (FormKZ), from difficult to easy (FormZK) and random (FormS). In addition to the success test, test anxiety inventory was also applied to the students.

In this study, "Success Test on the Structure of Atom" developed by Tağ (2012) and "Test Anxiety Inventory" adapted into Turkish Culture by Öner (1990) were used in data gathering.

Assumptions necessary for the analysis were realized before data analysis. Then, the students were divided as low success and high success by using "median" value and grouped as "high test anxiety" and "low test anxiety" according to median of the total score they got from Test Anxiety Inventory. Success test scores of students with low and high test anxiety levels were calculated in equalized groups (in low and high success groups). Success score means and item and test statistics of the students with low and high test anxiety levels were comparatively analysed.

Findings

There are not significant differences between the scores of students with low and high test anxiety levels in different forms of the success test in low and high success group. When the findings related to the item statistics obtained from the different forms of the success test are studied:

Findings related to item difficulty indice: it is seen that the difficulty levels of items 2 and 18 significantly differ in paired comparison of FormKZ and Form ZK. It is seen that the item difficulty indice of item number 2 in FormKZ is lower in FormZK. It is also seen that the item difficulty indice of item number 18 in FormKZ is higher in FormZK. It is noticed that the items 1, 2, 6 and 9 significantly differ in the comparison between FormKZ and FormS. It was found that the item difficulty indices of items 1, 2, and 9 were higher in FormKZ than in FormS, and that the item difficulty indice of item number 6 was higher in FormS. It is observed that items 1, 2, 6 and 9 are items that are placed in the beginning or in the middle of the test form. In the comparison between FormZK and FormS, a significant difference is seen in item 9. It was found out that the difficulty indice of this item had been lower in FormS and higher in FormKZ.

Findings related to item difficulty indice: it is seen that the item discrimination indices of FormKZ in items 9 and 13 significantly differ in comparison of FormKZ and Form ZK. In comparison of FormKZ and FormS, it is also seen that the item discrimination indices of FormS in items 6, 11 and 18 significantly differ from FormKZ. There was not found any significant differences in terms of item discrimination indice in comparison between FormZK and FormS.

Findings related to reliability coefficients: a significant difference was found between the reliability coefficients of FormKZ-FormZK and FormKZ-FormS. No significant difference was found between the reliability coefficients of FormZK-FormS.

Discussion & Conclusion

Changing the order of items in the test according to the findings obtained from the research does not cause a significant difference between the different forms of success test scores of students grouped as low and high test level anxiety in low success group. In high success group, however, there is significant difference between the different forms of success test scores of students grouped as low and high test level anxiety whereas there is no significant difference between the different forms of success test scores of students when there is not a group division. This finding is similar to some researches in the literature (Barcikovski and Olsen, 1975; Newman, Kundert, Lane, Bull, 1988; Smouse and Munz, 1968) while it does not match the findings of the research by Hambleton and Traub (1974).

It was seen in the order where the items were placed from easy to difficult and randomly that some item difficulty indices had increased in both forms whereas the number of these items had been more in easy to difficult form. It was also seen in the order where the items were placed from difficult to easy and randomly that some item difficulty indices had changed in both forms. One of the reasons there were different findings in the research is that the items were randomly ordered in the test. Item difficulty and discrimination indices might vary in every random order to be made.

It was noticed that easy items that are placed in the beginning of the test caused a decrease in some item discrimination indices and difficult items that are placed in the beginning of the test caused an increase in some item discrimination indices. It was observed in the form where the items were placed from easy to difficult that items which have lower item discrimination indices were more in number compared to randomly ordered form. Furthermore, it was observed that some item discrimination indices had changed when randomly ordered from difficult to easy. This finding is partially parallel to the findings conducted by Brenner (1964). When the orders of items from difficult to easy and from easy to difficult were compared, it was noticed that items which have higher item discrimination indices were more in number in difficult to easy order. When the order of items from difficult to easy and random order were compared, it was noticed that items which have higher item discrimination indices were more in number in random order. While the item discrimination index of 13 items in form whose items are randomly ordered are higher than the form whose items are from easy to difficult, the item discrimination index of 12 items in form whose items are randomly ordered are higher than the form whose items are from difficult to easy. The item discrimination index of 14 items in form whose items are ordered from difficult to easy are higher than the form whose items are from easy to difficult.

As a result of the comparison of the KR-20 reliability coefficients that were calculated for the easy to difficult, difficult to easy and random forms of the success test, the coefficients were found high ($>.70$) or closer to each other in both forms. It was also found that the coefficient of the form whose items had been placed from easy to difficult (Form KZ; 0.69) was medium level, but lower than others. Klossner and Gellman (1973) have determined the reliability approximately 0.68 in the form whose items were ordered according to subject; approximately 0.67 in the form ordered from easy to difficult within the topic and approximately 0.59 in the form ordered according to item difficulty, regardless of the subjects. Carlson and Ostrosky (1992), on the other hand, have determined a quite close reliability of the two forms they presented in four different tests prepared in order of subjects of the lessons in random order.

Madde Güçlüklerine Göre Farklı Sıralanan Testlerde Düşük ve Yüksek Kaygılı Öğrencilerin Akademik Başarılarının Karşılaştırılması¹

Çilem Doğan Gül

Hakkari University, Eğitim Fakültesi

Ömay Çokluk Bökeoğlu

Ankara Üniversitesi, Eğitim Bilimleri Fakültesi

Öz

Bu araştırmanın amacı, çoktan seçmeli testlerde yer alan maddelerin güçlük düzeylerine göre farklı şekillerde sıralanmasının, düşük ve yüksek sınav kaygılı öğrencilerin test puanına ve madde ve test istatistiklerine etkisini incelemektir. Farklı madde sıralamasının düşük ve yüksek sınav kaygılı öğrencilerin test puanları ve madde ve test istatistikleri açısından karşılaştırılması yapıldığından, çalışma nedensel karşılaştırma modeli ile uyumludur. Çalışmanın verileri Ankara ilinin dört merkez ilçesinden seçilen toplam yedi okuldan toplanmıştır. Veriler toplanırken Tağ (2012) tarafından geliştirilmiş olan "Atomun Yapısı Konusuna İlişkin Başarı Testi" ve öğrencilerin kaygı düzeylerini ölçmek için Öner (1990) tarafından Türk Kültürü'ne uyarlanmış olan "Sınav Kaygısı Envanteri" kullanılmıştır. Sonuç olarak, testin maddeleri kolaydan zora, zordan kolayla ve seçkisiz sıralandığında, bazı madde güçlük indeksleri ve madde ayırıcılık gücü indekslerinin farklılaştığı gözlemlenmiştir. Üç form arasında yapılan güvenilirlik katsayıları karşılaştırmasında kolaydan zora sıralanmış form ile zordan kolayla sıralanmış form ve kolaydan zora sıralanmış form ile seçkisiz sıralanmış form arasında manidar farklılık bulunmuştur

Anahtar Kelimeler: Madde sıralaması, çoktan seçmeli testler, sınav kaygısı, akademik başarı..



Inönü Üniversitesi
Eğitim Fakültesi Dergisi
Cilt 19, Sayı 3, 2018
ss. 252-265
DOI: 10.17679/inuefd.341477

Gönderim Tarihi : 03.10.2017
Kabul Tarihi : 19.07.2018

Önerilen Atıf

Doğan Gül, Ç. & Çokluk Bökeoğlu, Ö. (2018). Madde Güçlüklerine Göre Farklı Sıralanan Testlerde Düşük ve Yüksek Kaygılı Öğrencilerin Akademik Başarılarının Karşılaştırılması. *Inönü Üniversitesi Eğitim Fakültesi Dergisi*, 19(3), 252-265. DOI: 10.17679/inuefd.341477

¹ Bu çalışma Ömay ÇOKLUK BÖKEOĞLU danışmanlığı altında yürütülen Çilem DOĞAN GÜL'ün yüksek lisans tezinin bir bölümüdür.

GİRİŞ

Ölçme ve değerlendirme, eğitim sürecinde öğrenme ortamları, eğitim programlarının etkililiği ve öğrencilerin öğrenme düzeyleri gibi bazı temel bileşenler ile ilgilenmektedir. Öğrenme düzeylerinin belirlenmesi, öğrencilerin başarı durumlarının düzenli aralıklarla ve yansız bir şekilde ölçülmesi ve değerlendirilmesi ile mümkündür. Öğrenci başarısı farklı biçimlerde algılanan ve tanımlanan bir kavramdır. Genel anlamda başarı, istenilen bir sonuca ulaşma yönünde bir ilerlemedir. Okul ortamında ise başarı, belirli bir ders ya da akademik programdan bireyin ne derece yararlandığının bir ölçüsü ya da göstergesidir (Good, 1973). Başarı kavramıyla; akademik programda yer alan derslerde gösterilen gelişim, öğretmen notları, test puanları vb. ile belirlenen bilgi, beceri gibi kazanımları ifade eden "akademik başarı" kastedilmektedir (Good, 1973; Özgüven, 2002).

Akademik başarının değerlendirilmesi ve öğrenciler hakkında kararlar verilmesinde başarı testleri çok önemlidir. Başarı testleri farklı sınıf seviyelerinde, öğrencilerin öğrenme seviyelerini belirleme, olası öğrenme güçlüklerini saptama gibi amaçlarla öğrencilerin öğrenmelerine ilişkin somut kanıtlar elde etmek için sıklıkla kullanılmaktadır. Ancak, ölçme araçları kullanarak doğru kararlar verebilmek için, öncelikle bu araçların geliştirildikleri amaca hizmet etmeleri gerekir. Ölçme aracının amaca hizmet etmesi, ilgili özelliği hatalardan olabildiğince arınık olarak ölçmesiyle yakından ilişkilidir. Bunun sağlanması ise, ölçme araçlarının güvenilirlik, geçerlik ve kullanılabilirlik gibi psikometrik niteliklere sahip olması ile mümkündür.

Eğitimde akademik başarının ölçülmesinde geçerliği, güvenilirliği ve kullanılabilirliği yüksek testlerden yararlanmak gerekmektedir. Bu testler öğretmenler veya uzmanlar tarafından geliştirilebilir ve testin amacına uygun olarak da farklı madde formatları kullanılarak hazırlanabilirler. Madde formatlarına ilişkin en yaygın sınıflamanın "öğrencinin yanıt verme biçimine" göre yapılan sınıflama olduğu ifade edilebilir. Öğrencinin yanıt verme biçimi; "yanıtı seçerek verdiği maddeler (doğru-yanlış, eşleştirme, çoktan seçmeli)" ve "yanıt kendisinin oluşturduğu / yapılandırdığı (açık uçlu)" maddelerdir (Crocker ve Algina, 2006). Öğrencinin yanıtı seçerek verdiği madde türleri içerisinde ise çoktan seçmeli maddelerin en yaygın kullanıma sahip madde türü olduğu belirtilebilir. Türkiye'de merkezi olarak uygulanan tüm giriş sınavlarında (liselere giriş sınavı, üniversiteye giriş sınavı vb.) uzun yıllar boyunca yalnızca çoktan seçmeli madde formatı kullanılmış ve halen de kullanılmaya devam etmektedir. Çoktan seçmeli maddelerin (ya da bu maddelerden oluşan testlerin) akademik başarıyı belirlemede yaygın olarak kullanılması nedenleri arasında; puanlanmasının objektif ve kolay olması, kısa sürede puanlanabilmesi, çok sayıda soru sorulmasına olanak sağladığı için kapsam geçerliğinin yüksek olması, farklı öğrenim düzeylerinde kullanılabilirliği vb. sayılabilir.

Öğrenci başarısının ölçülmesinde önemli bir role sahip olan testlerin hazırlanması önemli bir süreçtir. Test geliştirme, özellikleri önceden belli bir yaklaşımla kestirilebilen bir test hazırlama işidir ve bazı işlemlerin gerçekleştirilmesini gerektirir (Baykul, 2010). Test geliştirilirken test düzeninin sağlanmasında uyulması gereken pek çok ilke vardır. Bu ilkeler; maddelerin test içerisinde sıralanışına karar verme, yanıt ve puanlama anahtarlarını düzenleme, öğrencilerin yanıt verme biçimine karar verme gibi birçok noktada yol gösterici olmaktadır.

Özellikle araştırmacıların ilgilendikleri ölçme aracı bir başarı testi olduğundan, maddelerin test içerisinde sıralanmasında "konu alanına", "madde türüne" ve "madde güçlüğüne" göre sıralama yapılmasının önemli olduğu vurgulanmakta ya da bu tarz sıralamalar önerilmektedir. Bu tarz madde sıralamalarının, öğrenci açısından yanıtlamayı, uygulayıcılar açısından da puanlamayı kolaylaştırdığı ifade edilmektedir (Aiken, 2000; Kline, 1986). Genel olarak Türkiye'de yapılan merkezi sınavlarda ve öğretmen yapımı başarı testlerinde madde sıralanışına dikkat edilmediği, hatta güvenlik gerekçesi ile maddelerin rastgele karıştırılması yoluyla farklı test formlarının oluşturulduğu ve uygulamalarda bu şekilde kullanıldığı görülmektedir.

Öğrencilerin akademik başarılarının belirlenmesi sürecinde farklı uygulamaların benimsendiği birçok sınava girmeleri ve bu sınavlarda başarılı olmaları gerekmektedir. Öğrencinin "başarılı" olarak değerlendirmesinde en önemli ölçütü oluşturan sınavlar, sınav kaygısının olumsuz etkilerini de beraberinde getirmektedir. Sınava yönelik tutum, sınavın güçlüğü, sınav ortamı gibi faktörlerin ise sınav kaygısının nesnel belirleyicileri olduğu belirtilmektedir (Zeidner, 1998). Alanyazında yer alan çalışmalarda, sınav hakkındaki tutum ve sınavın güçlüğüne ilişkin algının, test maddelerinin test içerisindeki düzeni veya sıralanışının (Covington, 1992; Plake, Anson, Parker, Lkowsky, 1982: 55; Vispoel, 1998; Wise, Roos, Plake, Nebelsick-Gullet, 1994) sınav kaygısı ile ilişkili olduğu vurgulanmaktadır.

Bir test içerisinde maddelerin sıralanması konusu incelendiğinde, teste kolay ya da çok kolay maddelerle başlanması ve test ilerledikçe maddelerin güçlük düzeyinin artmasının, öğrencilerin bu maddeleri yanıtlama güdülerini ve kendilerine güvenlerini açısından çok önemli olduğu, aynı zamanda sınav kaygısının belirli bir düzeyde tutulabilmesine katkı sağladığı gibi görüş ve ilkelere rastlandığından söz edilebilir. Teste zor maddelerle başlandığında, en başarılı adayların bile bu maddelere yanıt vermekte zorlanabildiği, bu maddelerde fazla zaman haralayabildiği ve sonuçta da gerçek performansının ölçme sürecine yansıtılamadığı gibi sorunlarla karşılaşılabilir. Daha testin ilk maddelerini cevaplamada güçlük çeken bir öğrencinin doğal

olarak sonraki maddeleri doğru cevaplamaya ilişkin özgüveni azalabilmekte, motivasyonu düşmekte, buna karşılık kaygı düzeyi ise yükselbilmektedir. Bir bireyin başarısının sınındığı herhangi bir durum ile karşı karşıya getirilmesi başlı başına kaygı yaratıcı iken, bir başarı testini yanıtlamaya başlayan bireyin, teste başlar başlamaz zor sorulara maruz bırakılması, ister istemez kaygı düzeyinin daha da yükselmesine ve test performansının olumsuz yönde etkilenmesine neden olabilecektir. Bu durum ise farklı kaygı düzeylerine sahip öğrencilerin (düşük ve yüksek sınav kaygılı vb.) test performansları arasında önemli farklılıklar oluşması ile sonuçlanacaktır (Ankara Üniversitesi, 2011).

Yukarıdaki açıklamalardan da anlaşılacağı gibi, uygulayıcılar genellikle madde sıralamasının test performansı (test puanı) üzerindeki etkisini göz ardı eden uygulamalar yapmakta, bu durumu dikkate almadıkları anlaşılmaktadır. Oysaki test düzeni oluşturma ile ilgili en temel ölçme ilkelerinden biri, maddelerin test içindeki düzeninin doğru bir mantıkla oluşturulmasıdır. Maddelerin yerlerinin değiştirilmesi ile elde edilen formlar bu temel ilke ile ters düşmekte, farklı test formları almak, adil bir ölçme yapılmasını engellemektedir. Maddelerin farklı test formlarında farklı yerlerde olması, bazı istatistiksel parametreleri de etkilediğinden, test formlarının eşitlikleri de bozulmaktadır. Bu durum maddelerin bazı öğrencilerce daha zor (ya da daha kolay) algılanabilmesine neden olmakta ve dolayısıyla da bir yanlılık kaynağı oluşturabilmektedir. Tüm bu gerekçeler, çoktan seçmeli testlerde madde sıralamasının düşük ve yüksek sınav kaygılı öğrencilerin test performansını nasıl etkilediğinin incelenmesini önemli kılmakta ve bu araştırmanın da problemini oluşturmaktadır.

Amaç

Bu araştırmanın genel amacını, çoktan seçmeli testlerde yer alan maddelerin güçlük düzeylerine göre farklı şekillerde sıralanmasının, düşük ve yüksek sınav kaygılı öğrencilerin test puanına ve madde ve test istatistiklerine etkisini incelemek oluşturmaktadır.

Bu genel amaç doğrultusunda yedinci sınıf fen ve teknoloji dersi Atomun Yapısı konulu başarı testi üç farklı form şeklinde düzenlenmiştir. Birinci formda maddeler kolaydan zora (Form KZ), ikinci formda zordan kolayla (Form ZK) ve üçüncü formda ise seçkisiz (random) olarak (Form S) sıralanmış ve aşağıdaki sorulara yanıt aranmıştır:

1. Düşük ve yüksek başarılı öğrenci gruplarında ayrı ayrı olmak üzere:
 - a. Öğrencilerin düşük ve yüksek sınav kaygılı grupta olduğuna bakılmaksızın, farklı formlardaki başarı puan ortalamaları arasında manidar bir farklılık var mıdır? (Form temel etkisi)
 - b. Öğrencilerin hangi test formunu aldıklarına bakılmaksızın, düşük ve yüksek sınav kaygılı gruplardaki öğrencilerin başarı puan ortalamaları arasında manidar bir farklılık var mıdır? (Grup temel etkisi)
 - c. Farklı test formları alan düşük ve yüksek sınav kaygılı öğrenci gruplarının başarı puan ortalamaları arasında manidar bir farklılık var mıdır? (Form X Grup ortak etkisi / etkileşim etkisi)
2. Başarı testinin üç formunun (Form KZ, Form ZK, Form S):
 - a. Madde güçlük indeksleri arasında manidar bir farklılık var mıdır?
 - b. Madde ayırt edicilik indeksleri arasında manidar bir farklılık var mıdır?
3. Başarı testinin üç formuna (Form KZ, Form ZK, Form S) ilişkin güvenilirlik katsayıları arasında manidar bir farklılık var mıdır?

YÖNTEM

Bu bölümde araştırmanın modeli, araştırma grupları, veri toplama araçları ve verilerin analizi hakkında bilgi verilmiştir.

Araştırma Modeli

Bu çalışmada, farklı madde sıralamasının düşük ve yüksek sınav kaygılı öğrencilerin test puanları ve madde ve test istatistikleri açısından karşılaştırılması yapıldığından, çalışma nedensel karşılaştırma modeli ile uyumludur. Nedensel karşılaştırma araştırmaları gruplar arasındaki farklılıkları, koşullar ve katılımcılar üzerinde herhangi bir müdahale olmaksızın belirlemeyi amaçlayan çalışmalardır (Büyüköztürk, Çakmak, Akgün, Karadeniz, Demirel, 2012).

Araştırma Grupları

Çalışmanın verileri Ankara ilinin dört merkez (Mamak, Çankaya, Altındağ, Sincan) ilçesinden seçilen toplam yedi okuldan toplanmıştır. Okullar seçilirken araştırmacının koşulları ve okulların kolay ulaşılabilirliği göz önünde bulundurulmuştur. Seçilen okullardan 578 yedinci sınıf öğrencisine fen ve teknoloji dersine ilişkin "Atomun Yapısı" konulu test, maddelerinin kolaydan zora, zordan kolayla ve seçkisiz yerleştirildiği üç farklı form halinde uygulanmıştır. Öğrencilere başarı testi ile birlikte sınav kaygısı envanteri de uygulanmıştır.

Veri Toplama Araçları

Bu çalışmada veriler toplanırken Tağ (2012) tarafından geliştirilmiş olan "Atomun Yapısı Konusuna İlişkin Başarı Testi" ve öğrencilerin kaygı düzeylerini ölçmek için Spielberger (1980) ve bir grup doktora öğrencisi tarafından geliştirilmiş, Öner (1990) tarafından Türk Kültürü'ne uyarlanmış olan "Sınav Kaygısı Envanteri" kullanılmıştır.

Atomun Yapısı Konusuna İlişkin Başarı Testi

Bu çalışmada veri toplamak amacıyla Tağ (2012) tarafından geliştirilmiş olan fen ve teknoloji dersi "Atomun Yapısı Konusuna İlişkin Başarı Testi" kullanılmıştır. Bu test öğrencilerin atomun yapısı konusundaki başarılarını belirlemeyi amaçlamaktadır. Test, dört seçenekli 20 çoktan seçmeli maddeden oluşmaktadır. Testin KR-20 güvenirlik katsayısı .73'tür ve Tağ (2012) tarafından geliştirilen testin orijinaline ait maddelere ilişkin ayırt edicilik ve madde güçlük indeksleri Tablo 1'de sunulmaktadır.

Tablo 1

Fen ve Teknoloji Dersi Atomun Yapısı Konusu Testindeki Maddelerin Güçlük ve Ayırt Edicilik İndeksleri

Madde	Düzyey	Madde Güçlük İndeksi(p_j)	Madde Ayırt Edicilik İndeksi(r_{ix})
1. Madde	Kolay	0.60	0.30
2. Madde	Zor	0.39	0.48
3. Madde	Kolay	0.64	0.20
4. Madde	Zor	0.33	0.51
5. Madde	Orta	0.45	0.27
6. Madde	Orta	0.56	0.46
7. Madde	Kolay	0.80	0.57
8. Madde	Kolay	0.64	0.50
9. Madde	Orta	0.54	0.34
10. Madde	Kolay	0.89	0.41
11. Madde	Zor	0.15	0.30
12. Madde	Kolay	0.73	0.43
13. Madde	Orta	0.46	0.69
14. Madde	Zor	0.31	0.39
15. Madde	Zor	0.26	0.32
16. Madde	Orta	0.53	0.43
17. Madde	Zor	0.36	0.31
18. Madde	Zor	0.31	0.51
19. Madde	Kolay	0.69	0.31
20. Madde	Orta	0.40	0.45

Madde güçlük indeksi maddenin uygulandığı grupta doğru yanıtlanma oranıdır ve 0 ile 1 arasında değerler alır. Madde güçlük indeksinin 0'a yakın değerler alması maddenin zor olduğuna, 1'e yakın değerler alması ise maddenin kolay olduğuna işaret eder. Çizelge 4'te yer alan güçlük düzeylerinin gruplanmasında 0-.39 (zor), .40 - .59 (orta) ve .60 - 1.00 (kolay) aralıkları referans alınmıştır (Kubiszyn ve Borich, 2013).

Sınav Kaygısı Envanteri

Öğrencilerin kaygı düzeylerini ölçmek için Spielberger (1980) ve bir grup doktora öğrencisi tarafından Güney Florida Üniversitesi'nde, 1974-1979 yılları arasında geliştirilmiştir. El kitabı ile birlikte ilk kez 1980 yılında yayınlanmıştır. Türk Kültürü'ne Öner (1990) tarafından uyarlanmıştır. Sınav Kaygısı Envanteri dörtlü Likert tipinde 20 maddelik bir ölçektir. Yanıtlayıcılar, sınavdan önce, sınavda ya da sınavdan sonra yaşadıkları olumsuz duyguların sıklık derecesini bu seçeneklerden birini işaretlemek suretiyle belirtmektedirler. Envanterden elde edilen puanların yüksekliği, sınav kaygısının yüksek olduğu anlamına gelmektedir. Test-tekrar test güvenirliği iki hafta ara ile yapılan uygulamada .80 ve 6 ay ara ile uygulamada ise .60 olarak hesaplanmıştır. Envanterin Türkçe formunun farklı çalışmalardan elde edilen Cronbach alfa güvenirlik katsayısı .70 ile .90 arasında değişmektedir. Test-tekrar test güvenirliği ise üç haftalık ara ile yapılan uygulamada .70 ile .90 arasında bulunmuştur (Öner, 1990).

Verilerin Analizi

Verilerin analizine geçilmeden önce verilerin analize hazırlanması sürecinde bazı işlemler yapılmıştır. Öncelikle verilerdeki uç değerler incelenmiştir. Bu amaçla tüm madde puanları Z puanlarına dönüştürülmüş ve ∓ 3 standart sapma aralığı dışındaki puanlar uç değer olarak kabul edilmiştir. Bu incelemede veri setinde hiçbir uç değere rastlanmamıştır.

Ardından gruplar önce yüksek ve düşük başarılı olarak denkleştirilmiş, sonrasında ise her grup kendi içinde düşük ve yüksek sınav kaygılı olarak sınıflandırılmıştır. Aynı formun aynı gruba üç kez uygulanması, taşıma etkisi (carry-over effects) (Köse, 2012; Kubiszyn ve Borich, 2013) olarak adlandırılan temel bir sorunu da beraberinde getirmektedir. Bir diğer deyişle ilk uygulamanın ardından öğrenciler verdikleri cevapları

hatırlayabilir ya da öğrenebilirler. Bu sorunu ortadan kaldırabilmek için gruplar okul yönetiminden alınan 100'lük sisteme göre verilmiş olan birinci dönem fen ve teknoloji dersi akademik not ortalamaları kullanılarak denkleştirilmiş ve her form farklı bir gruba uygulanmış ancak grupların denk olmasından dolayı, her form bireylere üç kez uygulanmış gibi kabul edilmiştir. Tablo 2'de farklı formları alan grupların denk olduğuna ilişkin betimsel istatistikler yer almaktadır. Tablo 3'te ise başarı testi puanlarının formlara göre karşılaştırıldığı ANOVA sonuçları yer almaktadır.

Tablo 2

Başarı Testinin Formlarına İlişkin Betimsel İstatistikler

Form	N	X	SS
FormKZ	195	48.87	18
FormZK	186	47.66	21
FormS	197	46.60	21

Tablo 3

Başarı Testi Puanlarının Formlara Göre ANOVA Sonuçları

Varyansın Kaynağı	Kareler Toplamı	sd	Kareler Ortalaması	F	p
Gruplararası	506.811	2	253.405	.627	.535
Gruplarıçi	232530.776	575	404.401		
Toplam	233037.587	577			

Tablo 2 ve 3'te görüldüğü gibi aritmetik ortalama ve standart sapma değerleri birbirine yakın ve bu değerler arasında yapılan karşılaştırma sonuçları ortalamalar arasında anlamlı bir farklılık olmadığını göstermektedir, [F(2, 575)= .627, p>.05]. Analiz sonuçları farklı formları alan uygulama gruplarının birbirine denk olduğunu göstermektedir.

Öğrenciler "ortanca" değeri kullanılarak düşük ve yüksek başarılı olarak ayrılmış ve ardından Sınav Kaygısı Envanterinden elde ettikleri toplam puanın ortancasına göre "yüksek sınav kaygılı" ve "düşük sınav kaygılı" olarak gruplandırılmıştır. Ortanca kullanılmasının nedeni uç değerlerden etkilenmemesi ve mod ve aritmetik ortalamaya göre daha etkili olmasıdır (Baykul, 2010). Tablo 4'te başarı durumları, kaygı durumları ve formların uygulandığı sıralama özetlenmektedir.

Tablo 4

Düşük ve Yüksek Başarılı Grupta Düşük ve Yüksek Sınav Kaygılı Öğrencilerin Aldıkları Formlar

Başarı	Kaygı Düzeyi	Alınan Formlar		
Düşük	Düşük	FormKZ	FormZK	FormS
	Yüksek	FormKZ	FormZK	FormS
Yüksek	Düşük	FormKZ	FormZK	FormS
	Yüksek	FormKZ	FormZK	FormS

Araştırmanın birinci alt amacı, öğrencilerin düşük ve yüksek kaygılı grupta olduğuna bakılmaksızın, farklı formlardaki başarı puan ortalamaları arasında manidar farklılık olup olmadığı (form temel etkisi), öğrencilerin hangi test formunu aldıklarına bakılmaksızın düşük ve yüksek kaygılı gruplardaki öğrencilerin başarı puan ortalamaları arasında manidar farklılık olup olmadığı (grup temel etkisi) ve farklı test formlarını alan düşük ve yüksek sınav kaygılı öğrencilerin başarı puan ortalamaları arasında manidar farklılık olup olmadığı (form ve grup ortak etkisi) karışık ölçümler için iki faktörlü ANOVA kullanılarak incelenmiştir.

Araştırmanın ikinci alt amacının a maddesi için madde güçlük indeksleri hesaplanmış ve testin üç formundan elde edilen madde güçlük indeksleri arasındaki fark, her madde için ikişerli olarak incelenmiştir. İki yüzde arasındaki fark (1) no'lu eşitlikten yararlanılarak hesaplanmıştır (Akhun, 1991):

$$z = \frac{P_1 - P_2}{\sqrt{PQ\left(\frac{1}{N_1} + \frac{1}{N_2}\right)}} \quad (1)$$

- P1: Birinci örneklemin yüzdesi
- P2: İkinci örneklemin yüzdesi
- P: İki örneklemin yüzdesinin ağırlıklı ortalaması
- Q: 100-P
- N1: Birinci örneklemin büyüklüğü
- N2: İkinci örneklemin büyüklüğü

Araştırmanın ikinci alt amacının b maddesi için araştırmada kullanılan üç test formunda yer alan maddelerin ayırt edicilik indeksleri için nokta çift serili korelasyon katsayısı hesaplanmış ve hesaplanan madde ayırt edicilik indeksleri arasındaki fark, her madde için ikiyeşerli olarak incelenmiştir. Bu incelemede, önce korelasyon katsayıları Fisher'in Zr'sine dönüştürülmüştür. Bu dönüştürme için (2) no'lu eşitlikten yararlanılmıştır (Akhun, 1991).

$$Z_r = \frac{1}{2} \log_e \frac{1+r}{1-r} \quad (2)$$

Zr: Korelasyon katsayısının Fisher'in z katsayısına dönüşümü

R: Korelasyon katsayısının değeri

Fisher'in Zr'sine dönüştürülen korelasyon katsayıları arasındaki farkın manidar olup olmadığı (3) no'lu eşitlikten yararlanılarak hesaplanmıştır (Akhun, 1991).

$$Z = \frac{Z_{r1} - Z_{r2}}{\sqrt{\frac{1}{n_1 - 3} + \frac{1}{n_2 - 3}}} \quad (3)$$

Z: İki korelasyon arasındaki farka ilişkin test istatistiği

Zr1: Birinci korelasyon katsayısı

Zr2: İkinci korelasyon katsayısı

n1: Birinci örneklem büyüklüğü

n2: İkinci örneklem büyüklüğü

Araştırmanın üçüncü alt amacı için test formlarının güvenilirlik katsayıları KR-20 güvenilirlik katsayısı ile hesaplanmıştır. İki güvenilirlik katsayısı arasındaki farkın manidarlığının test edilmesi için iki aritmetik ortalama, iki yüzde ya da iki diğer istatistik arasındaki manidarlığın test edilmesinde izlenen yaklaşımla aynıdır; önce iki katsayı arasındaki fark bulunur, sonra farkın standart hatası hesaplanır ve farka bölünür. Elde edilen kritik oran ya da Z değeri bu iki korelasyon katsayısı arasındaki farkın istatistiksel olarak manidar olup olmadığına karar verilmesinde yardımcı olur (Akhun, 1984).

BULGULAR

Bu bölümde araştırmanın amaçları doğrultusunda elde edilen bulgulara ve yorumlara yer verilmiştir. Araştırma sorularına yanıt aramaya çalışırken öncelikle denkleştirilmiş gruplarda düşük ve yüksek sınav kaygılı öğrencilerin başarı testinden elde ettikleri test puanları hesaplanmıştır. Düşük ve yüksek sınav kaygılı öğrenci gruplarının başarı puan ortalamalarının karşılaştırılmasına ilişkin bulgular denkleştirilmiş gruplarda (düşük ve yüksek başarılı gruplarda) ayrı ayrı incelenmiştir.

Düşük ve yüksek Sınav Kaygılı Öğrenci Gruplarının Başarı Puan Ortalamalarının Karşılaştırılmasına İlişkin Bulgular

A. Düşük Başarılı Grup

Düşük başarılı grubun madde sıralanışı olan farklı test formlarından aldıkları ortalama puanların manidar bir farklılık gösterip göstermediğine ilişkin grup ve form temel etki ve grup*form ortak etki testlerine ilişkin iki faktörlü ANOVA sonuçları Tablo 5'te sunulmaktadır.

Tablo 5

Düşük Başarılı Grupta Form ve Kaygı Değişkenlerine Göre Başarı Testi Puanlarının Karşılaştırılması

Varyansın Kaynağı	KT	sd	KO	F	p
Deneklerarası	26634.249	90			
Grup	2.381	1	2.381	.008	.929
Hata	26631.868	89	299.234		
Denekleriçi	47406.943	182			
Form	936.796	2	468.398	1.799	.168
Grup*Form	128.005	2	64.002	.246	.782
Hata	46342.142	178	260.349		
Toplam	74041.192	272			

Tablo 5'e göre, düşük ve yüksek sınav kaygılı öğrencilerin başarı testinin farklı formlarından (Form KZ, Form ZK, Form S) elde ettikleri toplam puan ortalamaları arasında manidar bir fark bulunmamıştır [F(1, 89)= .008, p>.05]. Form temel etkisi ile ilgili olarak da benzer bir sonuç elde edilmiştir. Grup ayrımı (düşük ve yüksek kaygılı grup) yapmaksızın araştırmada yer alan öğrencilerin farklı formlardan elde ettikleri test puan ortalamaları arasında manidar bir fark bulunmamaktadır [F(1, 89)= 468.398, p>.05].

Grup*form ortak etkisine ilişkin test sonucu incelendiğinde, düşük başarılı grupta, düşük ve yüksek sınav kaygılı öğrencilerin başarı testinin farklı formlarından aldıkları puanlar arasında manidar farklılık bulunmamaktadır [F(1, 89)=1.799, p>.05].

B. Yüksek Başarılı Grup

Yüksek başarılı grubun madde sıralanışı farklı olan test formlarından aldıkları puanların manidar bir farklılık gösterip göstermediğine ilişkin iki faktörlü ANOVA sonuçları Tablo 6'da sunulmaktadır.

Tablo 6

Yüksek Başarılı Grupta Form ve Kaygı Değişkenlerine Göre Başarı Testi Puanlarının Karşılaştırılması

Varyansın Kaynağı	KT	sd	KO	F	p
Deneklerarası Grup	26316.868	82			
Hata	1231.417	1	1231.417	3.976	.045
Denekleriçi Form	25085.451	81	309.697		
Hata	57870.923	166			
Grup*Form	253.453	2	126.727	.360	.698
Hata	568.313	2	284.156	.807	.448
Toplam	57049.157	162	352.155		
	84187.791	248			

Tablo 6'ya göre, düşük ve yüksek sınav kaygılı öğrencilerin başarı testinin farklı formlarından (Form KZ, Form ZK, Form S) elde ettikleri puanların ortalamaları arasında manidar fark vardır [$F(1, 81)=3.976, p<.05$]. Form temel etkisi incelendiğinde, grup ayrımı yapılmaksızın araştırmada yer alan öğrencilerin aldıkları farklı formlardan elde ettikleri test puan ortalamaları arasında manidar farkın olmadığı görülmektedir [$F(1, 81)=.360, p>.05$].

Grup*Form ortak etkisine ilişkin test sonucu incelendiğinde, yüksek başarılı grupta, düşük ve yüksek sınav kaygılı öğrencilerin testin farklı formlarından aldıkları puanların manidar farklılık göstermediği saptanmıştır [$F(1, 81)=.807, p>.05$].

Başarı Testinin Farklı Formlarından Elde Edilen Madde İstatistiklerine İlişkin Bulgular Madde Güçlük İndekslerine İlişkin Bulgular

Araştırmada yer alan başarı testinin üç farklı formundan elde edilen madde güçlük indeksleri arasında manidar bir fark olup olmadığını belirlemek için öncelikle formlara ait madde güçlük indeksleri hesaplanmış ve ardından karşılaştırmalar yapılmıştır. Sonuçlar Tablo 7'de sunulmaktadır.

Tablo 7

Başarı Testinin Farklı Formlarından Elde Edilen Madde Güçlük İndeksleri ve Karşılaştırılması

Madde No	Madde Güçlük İndeksi			Form KZ – Form ZK $P_{j1}-P_{j2}$	Form KZ- Form S $P_{j1}-P_{j3}$	Form ZK- Form S $P_{j2}-P_{j3}$
	Form KZ	Form ZK	Form S			
1	0.82	0.79	0.71	0.74	2.59**	1.81
2	0.78	0.66	0.66	2.63**	2.66**	0.00
3	0.53	0.58	0.60	0.98	1.41	0.40
4	0.54	0.49	0.48	0.97	1.20	0.19
5	0.59	0.60	0.55	0.20	0.80	0.99
6	0.48	0.50	0.59	0.39	2.20**	1.77
7	0.36	0.42	0.39	1.21	0.62	0.60
8	0.73	0.71	0.69	0.44	0.88	0.43
9	0.59	0.55	0.45	0.79	2.79**	1.97*
10	0.56	0.56	0.53	0.00	0.58	0.59
11	0.52	0.46	0.45	1.18	1.39	0.20
12	0.27	0.25	0.29	0.39	0.44	0.88
13	0.34	0.30	0.34	0.84	0.00	0.84
14	0.40	0.37	0.38	0.60	0.41	0.25
15	0.46	0.45	0.40	0.20	1.21	0.99
16	0.54	0.54	0.55	0.00	0.20	0.20
17	0.40	0.45	0.41	1.10	0.20	0.79
18	0.38	0.49	0.45	2.16**	1.61	0.78
19	0.15	0.10	0.13	1.48	0.57	0.93
20	0.23	0.27	0.26	0.91	0.69	0.22

* $p<.05$ ** $p<.01$

Tablo 7'ye göre, Form KZ ile Form ZK arasında yapılan ikili karşılaştırmalara göre 2. ve 18. maddelerin güçlük düzeylerinin manidar bir şekilde farklılaştığı görülmektedir. Form KZ'de 2. sırada yer alan maddenin güçlük indeksinin Form ZK'de daha düşük olduğu görülmektedir. Form KZ'de 18. sırada yer alan maddenin ise güçlük indeksinin Form ZK'de daha yüksek olduğu görülmektedir. 2. maddenin Form KZ'de testin başında, 18. maddenin ise Form ZK'de testin sonunda yer aldığı gözlenmektedir. Test formu maddeleri kolaydan zora sıralandığında başta yer alan bu maddelerin, zordan kolayla sıralandığında ise bazı zor maddelerin madde güçlük indekslerinin yüksek çıktığı görülmektedir. Buna göre maddelerin kolaydan zora ve zordan kolayla sıralandığı formlarda, ilgili maddelerin test formunun başında yer almasının, bu maddelerin test formunun sonunda yer almasına göre madde güçlük indekslerinde yükselmeye neden olduğu düşünülebilir.

Form KZ ve Form S arasında yapılan karşılaştırmada 1., 2., 6. ve 9. maddelerin manidar bir şekilde farklılaştığı görülmektedir. 1., 2. ve 9. maddelerin güçlük indekslerinin Form KZ'de Form S'den daha yüksek olduğu, 6. maddenin ise güçlük indeksinin Form S'de daha yüksek olduğu bulunmuştur. 1., 2., 6. ve 9. maddelerin test formunun başında veya ortasında yer alan maddeler olduğu gözlenmektedir. Bu durum yapılacak her seçkisiz atamada değişeceğinden, madde güçlük indeksini etkilediği ifade edilebilir.

Form ZK ve Form S arasında yapılan karşılaştırmada ise 9. maddede manidar bir farklılık görülmektedir. Bu maddenin güçlük indeksinin Form S'de daha düşük, Form KZ'de ise daha yüksek olduğu belirlenmiştir.

Madde Ayırt Edicilik İndekslerine İlişkin Bulgular

Başarı testinin üç farklı formundan elde edilen madde ayırt edicilik indeksleri arasında manidar bir fark olup olmadığını belirlemek amacıyla öncelikle formlara ait madde ayırt edicilik indeksleri hesaplanmış daha sonra karşılaştırmalar yapılmış ve Tablo 8'de sunulmuştur.

Tablo 8

Başarı Testinin Farklı Düzenlenmiş Formlarından Elde Edilen Madde Ayırt Edicilik İndekslerinin Karşılaştırması

Madde No	FormKZ-ZK	FormKZ-S	FormZK-S	Madde No	FormKZ-ZK	FormKZ-S	FormZK-S
1	-1.08	-1.18	0.00	11	-0.68	-2.17*	-1.47
2	-1.18	-1.09	0.09	12	-1.47	-0.29	1.18
3	-0.49	-0.49	0.00	13	-2,15*	0,09	2.25
4	-0.39	-0.69	-0.29	14	0.49	-0.39	-0.88
5	-0.49	0.09	0.59	15	-0.88	0.49	1.37
6	-1.18	-2.77**	-1.56	16	-1.47	-0.89	0.59
7	-0.78	0.30	1.08	17	-0.78	-1.48	-0.68
8	-0.98	-0.99	0.00	18	-1.47	3.16**	-1.66
9	-2.15*	-1.38	0.78	19	-1.47	-0.40	1.08
10	1.76	1.28	-0.49	20	-0.88	-0.99	-0.09

Tablo 8 incelendiğinde, Form KZ ve Form ZK arasında yapılan karşılaştırmada 9. ve 13. maddelerde, Form ZK'nın madde ayırt edicilik gücü indekslerinin, Form KZ'den manidar bir şekilde farklılaştığı görülmektedir. Kolay maddelerin test formunun başında yer alması, bu maddelerin formun sonunda yer almasına göre bazı ayırt edicilik indekslerinde düşüşe yol açtığı, aynı şekilde zor maddelerin test formunun sonunda yer almasının ise formun başında yer almasına göre, bazı ayırt edicilik indekslerinde yükselmeye neden olduğu ifade edilebilir.

Form KZ ve Form S arasında yapılan karşılaştırmada 6., 11. ve 18. maddelerde, Form S'nin madde ayırt edicilik gücü indekslerinin Form KZ'den manidar bir şekilde farklılaştığı görülmektedir. Form ZK ve Form S arasında yapılan karşılaştırmada ise madde ayırt edicilik gücü indeksleri açısından manidar bir fark bulunmamıştır.

Başarı Testinin Üç Farklı Formundan Elde Edilen Güvenirlilik Katsayılarına İlişkin Bulgular

Güvenirlilik kestirimlerine ilişkin bulgular incelendiğinde, Form KZ'in güvenirliliği 0,69; Form ZK'nin güvenirliliği 0,79; Form S'nin güvenirliliği ise 0,78 olarak bulunmuştur. Her üç form için elde edilen güvenirlilik katsayılarının yüksek olduğu; Form KZ için elde edilen güvenirliliğin diğer formların güvenirliliklerine göre görece daha düşük olduğu ifade edilebilir.

Başarı testinin farklı formlarına ilişkin elde edilen KR-20 ve güvenirlilik katsayılarının manidar farklılık gösterip göstermediğini belirlemek amacıyla uygulanan Fisher'in Z Testi sonuçları Tablo 9'da sunulmuştur.

Tablo 9

Başarı Testinin Farklı Düzenlenmiş Formlarına İlişkin Güvenirlik Katsayılarının Karşılaştırılması

Formlar	KR-20	Z _r	z
KZ	0.69	0.84	
ZK	0.79	1.07	-2.25*
KZ	0.69	0.84	
S	0.78	1.04	-1.98*
ZK	0.79	1.07	
S	0.78	1.04	0.29

Tablo 9 incelendiğinde Form KZ-Form ZK ve Form KZ-Form S'nin güvenilirlik katsayıları arasında manidar fark olduğu görülmektedir. Form ZK-Form S'nin güvenilirlik katsayıları arasında ise manidar bir farklılık bulunamamıştır.

TARTIŞMA, SONUÇ VE ÖNERİLER

Bu araştırmada maddeleri farklı sıralanan çoktan seçmeli testlerde düşük ve yüksek sınav kaygılı öğrencilerin başarıları, düşük ve yüksek başarılı gruplarda ayrı ayrı karşılaştırılmıştır. Ayrıca madde sıralanmasının bazı madde istatistiklerinde farka yol açıp açmadığı incelenmiştir.

Araştırmadan elde edilen sonuçlara göre başarı testinde maddelerin test içindeki sıralamasının değiştirilmesi, düşük başarılı grup içerisinde; düşük ve yüksek sınav kaygılı olarak ayrılan öğrencilerin başarı testinin farklı formlarından elde ettikleri test puanları arasında manidar bir fark yaratmamaktadır. Yüksek başarılı grup içerisinde ise; düşük ve yüksek sınav kaygılı olarak ayrılan öğrencilerin başarı testinin farklı formlarından elde ettikleri test puanları arasında manidar fark bulunmakta, grup ayırımı yapılmadığında ise öğrencilerin farklı formlardan elde ettikleri test puanları arasında manidar farka rastlanmamaktadır. Bu bulgu, alanyazındaki bazı araştırmalar ile benzerlik gösterirken (Barcikovski ve Olsen, 1975; Nevman, Kundert, Lane, Bull, 1988; Smouse ve Munz, 1968); Hambleton ve Traub'un (1974) yaptıkları araştırmadan elde edilen bulgularla da uyumlanmaktadır.

Başarı testinde maddeler, madde güçlük indekslerine göre kolaydan zora, zordan kolaya ve seçkisiz sıralanarak üç form halinde düzenlendiğinde, bazı maddelerin güçlük düzeylerinde farklılaşma olduğu belirlenmiştir. Maddeleri kolaydan zora ve zordan kolaya sıralanmış formlarda; maddelerin test formunun başında yer almasının, bu maddelerin test formunun sonunda yer almasına göre bazı madde güçlük indekslerinde yükselmeye neden olduğu saptanmıştır. Maddelerin kolaydan zora ve seçkisiz olarak sıralanmasında, bazı madde güçlük indekslerinin her iki formda da yükseldiği ancak bu maddelerin sayısının kolaydan zora formda daha fazla olduğu görülmüştür. Maddelerin zordan kolaya ve seçkisiz olarak sıralanmasında ise, bazı madde güçlük indekslerinin her iki formda da değiştiği gözlenmiştir. Araştırmada formlar arasındaki karşılaştırmalara göre farklı sonuçlar elde edilmesinin bir nedeni, maddelerin test içinde seçkisiz olarak sıralanmasıdır. Yapılacak her farklı seçkisiz sıralamada madde güçlük ve ayırt edicilik gücü indeksleri farklılaşabilir.

Maddeler, madde güçlük düzeylerine göre kolaydan zora, zordan kolaya ve seçkisiz formlar halinde düzenlendiğinde, bazı madde ayırt edicilik gücü indekslerinde farklılaşma olduğu gözlenmiştir. Kolay ve zor maddelerin test formunun başında ya da sonunda yer almasına göre bazı madde ayırt edicilik gücü indeksleri farklılaşmaktadır. Kolay maddelerin test formunun başında yer almasının, bazı madde ayırt edicilik gücü indekslerinin düşmesine; zor maddelerin test formunun başında yer almasının ise bazı madde ayırt edicilik gücü indekslerinin yükselmesine neden olduğu görülmüştür. Maddeleri kolaydan zora sıralanan test formunda, seçkisiz sıralanan forma göre madde ayırt edicilik gücü indeksleri daha düşük olan maddelerin sayıca daha fazla olduğu gözlenmiştir. Ayrıca, test maddeleri zordan kolaya ve seçkisiz olarak sıralandığında bazı madde ayırt edicilik indekslerinin değiştiği görülmüştür. Bu bulgu, Brenner (1964) tarafından yapılan araştırmanın bulguları ile kısmen paralellik göstermektedir. Carlson ve Ostrosky (1992) ise, yaptıkları dört testte, madde ayırt ediciliklerinin değişim yüzdesini hesaplamışlardır. Bu yüzdelerin, iki testte maddeleri konulara göre sıralanan formda artığı, iki testte ise karışık sıralanan formda arttığını ortaya koymuşlardır. Bu bulgu, araştırmanın bulguları ile benzemektedir. Maddelerin kolaydan zora sıralanması seçkisiz sıralanması ile karşılaştırıldığında, madde ayırt edicilik indekslerinin kolaydan zora sıralamada daha düşük olduğu belirlenmiştir. Maddelerin zordan kolaya sıralanması ve kolaydan zora sıralanması karşılaştırıldığında, zordan kolaya sıralamada madde ayırt edicilik indeksi yüksek olan madde sayısının daha fazla olduğu saptanmıştır. Maddelerin zordan kolaya sıralanması ve seçkisiz sıralanması karşılaştırıldığında ise, seçkisiz sıralamada madde ayırt edicilik indeksi yüksek olan madde sayısının daha fazla olduğu belirlenmiştir. Maddeleri seçkisiz sıralanan formun 13 maddesinin ayırt edicilik gücü indeksi, maddeleri kolaydan zora sıralanan formdan daha yüksekken; maddeleri seçkisiz sıralanan formun 12 maddesinin ayırt edicilik gücü indeksi, maddeleri zordan kolaya

sıralanan formdan daha yüksektir. Maddeleri zordan kolayca sıralanan formun 14 maddesinin ayırt edicilik gücü indeksi, maddeleri kolaydan zora sıralanan formdan daha yüksektir.

Başarı testinin maddelerin kolaydan zora, zordan kolayca ve seçkisiz olarak sıralandığı formlar için hesaplanan KR-20 güvenilirlik katsayılarının karşılaştırılması sonucunda, maddeleri zordan kolayca ve seçkisiz sıralanan formlarda yüksek (>.70) ve birbirine daha yakın bulunmuştur. Maddeleri kolaydan zora sıralanan formun katsayısının ise (Form KZ:, 0.69) orta düzeyde, ancak diğerlerinden daha düşük olduğu bulunmuştur. Klosner ve Gellman (1973), maddeleri konuya göre sıralanmış formda güvenilirliği yaklaşık 0,68; konu başlığı içinde kolaydan zora sıralanan formda yaklaşık 0,67 ve konulara bakılmaksızın madde güçlüğüne göre sıralanan formda ise yaklaşık 0,59 olarak belirlemişlerdir. Carlson ve Ostrosky (1992) ise, dört ayrı sınavda konuları derste işlenen sırada ve karışık olarak sundukları iki formun güvenilirliklerin birbirine oldukça yakın olduğunu saptamışlardır.

Bu çalışmada bir başarı testinin farklı formlarına ait test istatistikleri ve öğrencilerin bu testten aldıkları puanlar karşılaştırılmıştır ve taşıma etkisinin meydana getireceği problemten dolayı formların uygulandığı öğrenci grupları akademik not ortalamalarına göre denkleştirilmiştir ve bu gruplar üzerinde analiz yapılmıştır. Benzer karşılaştırmalar bir testin paralel formları hazırlanarak incelenebilir. Çalışmada 20 maddelik bir başarı testi kullanılmıştır. Madde sayısı daha fazla olan bir test ile araştırma tekrarlanabilir. Test hazırlama çalışmalarında test formunda yer alan maddelerin, madde güçlük düzeylerine göre sıralanmasında cinsiyet, okul türü, yaş, test türü gibi değişkenler de göz önüne alınarak benzer çalışmalar yapılabilir. Bu çalışmada uygulanan test ile öğrencilere not verilmemiştir. Öğretmenlerin not vereceği gerçek bir sınav ortamında araştırma tekrarlanabilir.

KAYNAKÇA/REFERENCES

- Aiken, L. R. (2000). Psychological testing and assessment. USA: Allyn and Bacon.
- Ankara Üniversitesi, (2011). Ankara Üniversitesi Eğitim Bilimleri Fakültesi'nin YGS Hakkında Görüşü. Web: <http://rigel2.cc.ankara.edu.tr/dyr.php?id=1170> 01 Nisan 2014 tarihinde alınmıştır.
- Akhun, İ. (1991). İstatistiklerin manidarlığı ve örneklem. Ankara: Kendi Yayını.
- Barcikovski, R. S. ve Olsen, H. (1975). Test item arrangement and adaptation level. The Journal of Psychology, 90, 87-93.
- Baykul, Y. (2010). Eğitimde ve psikolojide ölçme klasik test teorisi ve uygulaması. Ankara: ÖSYM Yayınları.
- Brenner, M. H. (1964). "Test difficulty, reliability, and discriminations of item difficulty order". Journal of Applied Psychology, 48(2), 98-100.
- Büyüköztürk, Ş., Kılıç Çakmak, E., Akgün, Ö. E., Karadeniz, Ş., ve Demirel, F. (2012). Bilimsel araştırma yöntemleri. Ankara: PegemA.
- Carlson, J. L. ve Ostrosky, A. L. (1992). Item sequence and student performance on multiple-choice exams: further evidence. Journal of Economic Education, 2(3), 232-235.
- Covington, M. V. (1992). Making the Grade: A self worth perspective on motivation and school reform. New York: Cambridge University Press.
- Crocker, L. ve Algina J. (2006). Introduction to classical and modern test theory. N.Y.: Nelson Education.
- Good, V. C. (1973). Dictionary of education. N.Y.: McGraw Hill Book Company.
- Hambleton, R. K. ve Traub, R. E. (1974). The effects of item order on test performance and stres. Journal of Experimental Education, 43(1), 40-46.
- Kline, P. (1986). A handbook of test construction: introduction to psychometric desing. London: Methuen&Co.
- Klosner, N.C.ve Gellman, E.K. (1973). "The effect of item arrangement on classroom test performance: implications for content validity". Educational and Psychological Measurement, 33, 413-418.
- Köse, İ.A. (2012). Ölçmede güvenilirlik. N. Çikrikçı-Demirtaşlı, (Ed.), Eğitimde ölçme ve değerlendirme (93-115). Ankara Elhan Yayınları.
- Kubiszyen, T. ve Borich, G. (2013). Educational testing and measurement: classroom application and practice. (10th ed.) John Wiley ve Sons, INC.
- Nevman, D.L, Kundert, D.K, Lane, D.S ve Bull K. S. (1988). Effect of varying item order on multiple-choice test scores: importance of statistical and cognitive difficulty. Applied Measurement in Education, 1 (1), 89-97.
- Öner, N. (1990). Sınav kaygısı envanteri el kitabı. İstanbul: Yöret Vakfı.
- Özguven, G. E. (2002). Psikolojik testler. Ankara: Nobel Yayın Dağıtım.
- Plake, B. S., Ansorge, C. J., Parker, C. S., ve Lkowsky, S. R. (1982). Source effects of item arrangement, knowledge of arrangement test anxiety and sex on test performance. Journal of Educaitional Measurement, 19 (1), 49-57.

- Smouse, Albert D. ve Munz, David C. (1968). The effects of anxiety and item difficulty sequence on achievement testing scores. *The Journal of Psychology*, 68, 181-184.
- Spielberger, C. D. (1980). Preliminary professional manual for the test anxiety inventory. California, USA: Consulting Psychologist Press.
- Tağ, S. M. (2012). Atomun yapısı konusunu öğrenmede klasik yöntemler ile bilgisayar destekli öğretimin öğrenci başarısına etkileri (Yüksek Lisans Tezi, Fırat Üniversitesi, 2012). YÖK Ulusal Tez Merkezi.
- Vispoel, W. P. (1998). Psychometric characteristics of computer-adaptive and self-adaptive vocabulary tests: The role of answer feedback and test anxiety. *Journal of Educational Measurement*, 35 (2), 155-167.
- Wise, S. L., Roos, L. L., Plake, B. S. ve Nebelsick-Gullet, L. J. (1994). The relationship between examinee anxiety and preference for self-adapted testing. *Applied Measurement in Education*, 7 (1), 81-91.
- Zeidner, M. (1998). Test anxiety-the state of art. New York, USA: Plenum Press.

İletişim/Correspondence

Arş. Gör. Çilem DOĞAN GÜL
cilemdogangul@hakkari.edu.tr

Ömay ÇOKLUK BÖKEOĞLU
cokluk@education.ankara.edu.tr