



A Comparison of Psychometric Properties of a General Ability Test Which Administered In Paper-Pencil and Computer Based Form¹

Eren Can AYBEK²

R. Nükhet DEMİRTAŞLI³

ABSTRACT. The aim of this research is to compare the students' performance according to their attitudes towards computer, computer familiarity and gender; and to compare the psychometric properties of a general ability test which is administered in paper-pencil and computer based form. The study group includes a total of 136 primary school students who live in central district of Eskisehir province. A 60-item general ability test, which includes verbal and computational reasoning sub-tests, the attitude towards computer scale, and the computer familiarity scale were administered to the study group. The findings showed that paper-pencil and computer based forms were highly equal and there was no significant difference between students' computer based test performances in terms of computer familiarity and gender. On the other hand, there was a significant difference between students' computer based test performances in their attitudes towards computers.

Keywords: computer based test, general ability test, attitude towards computer, computer familiarity

SUMMARY

Purpose and significance: This study is aimed at comparing psychometric properties of a general ability test (GAT) which is administered in paper-pencil and computer based form. Besides, students' performances on computer based form were investigated by their attitudes towards computer, computer familiarity and gender.

Methods: The study group includes a total of 136 primary school students, 73 male and 63 female, who live in central district of Eskişehir. In this study, several instruments were used to collect data. First, a general ability test with 60 multiple choice items, which includes verbal and numerical reasoning sub-tests, was used. Other instruments were a 17-item scale, which aimed at measuring students' attitude towards computer and the computer familiarity scale were administered to the study group. The mean scores of GAT, validity and reliability coefficients and item statistics were computed. The means of dependent samples were compared by *t*-test's. Feldt's test was performed for comparison of reliability coefficients and z-test was used to compare validity coefficients and item statistics of two forms of general ability test. And also independent samples that obtained from two forms of were compared by *t*-test in terms of attitudes towards computer, computer familiarity and gender.

Results: KR-20 reliability coefficients of computer based and paper-pencil forms were compared and there was no significant difference between coefficients for whole test and sub-tests. Students' performances on different forms of general ability test were compared and students' performances were significantly higher on paper-pencil form of verbal reasoning sub-test. Yet, there was no significant difference between forms of computational reasoning sub-test. The results showed that students' performances on computer based form of general ability test were not significantly differentiated by gender and computer familiarity. However, the opposite was true when it came to the attitudes. Low and high attitude score groups' performances on computer based form of general ability test were significantly different.

Discussion and Conclusions: Paper-pencil and computer based forms were highly equal and there was no significant difference between students' computer based test performances according to computer familiarity and gender. Apart from that there was significant difference between students' computer based test performances according to attitude towards computers.

¹ This study is a summary of the thesis submitted to the Ankara University for the Master degree in Education

² Res. Asst. Eskişehir Osmangazi University, Faculty of Education, Department of Educational Sciences. – e-mail: erencan@aybek.net

³ Prof. Dr., Ankara University, Faculty of Educational Sciences, Department of Measurement and Evaluation, Ankara, Turkey. e-mail: rnukhet@yahoo.com

Bilgisayar Ortamında Ve Kağıt-Kalem Formunda Uygulanan Genel Yetenek Testinin Psikometrik Özelliklerinin Karşılaştırılması¹

Eren Can AYBEK²

R. Nükhet DEMİRTAŞLI³

ÖZ. Bu çalışmada, bir genel yetenek testi (GYT) kağıt-kalem ve bilgisayar ortamında uygulanarak, testin psikometrik özelliklerinin ve bilgisayara yönelik tutum, bilgisayar aşinalığı ve cinsiyete göre öğrenci performanslarının karşılaştırılması amaçlanmıştır. İlişkisel türde betimsel bir araştırma olan bu çalışmanın araştırma grubunu Eskişehir Büyükşehir Belediyesi sınırları içerisinde kalan ilköğretim okullarından seçilen; 73 erkek, 63 kız toplam 136 6. sınıf öğrencisi oluşturmaktadır. Araştırma grubuna sözel ve sayısal akıl yürütme alt testlerinden oluşan, toplam 60 maddelik genel yetenek testi kağıt-kalem olarak ve bilgisayar ortamında uygulanmıştır. Araştırma grubuna aynı zamanda bilgisayara yönelik tutum ölçeği ve bilgisayar aşinalığı ölçekleri de verilmiştir. Araştırmadan elde edilen bulgular; testin kağıt-kalem ve bilgisayar ortamında uygulanan formlarının ölçtükleri özellik bakımından denk olduğunu; bilgisayar aşinalığı ve cinsiyete göre, testin bilgisayar ortamında uygulanan formundan alınan puanların farklılaşmadığını; buna karşın, bilgisayara yönelik tutum bakımından bilgisayar ortamında uygulanan formdan alınan GYT puanlarının bilgisayara yönelik olumlu tutuma sahip olanlar lehinde farklılaştığını göstermektedir.

Anahtar sözcükler: bilgisayar ortamında uygulanan test, genel yetenek testi, bilgisayara yönelik tutum, bilgisayar aşinalığı

GİRİŞ

Gelişen bilgisayar ve internet teknolojisi eğitim-öğretim ve ölçme işlemlerinin bilgisayar ortamına aktarılmasını giderek artırmıştır. Bu eğilim beraberinde teknolojinin öğrenme, öğretme ve değerlendirme süreçleri üzerindeki etkilerinin araştırılmasını önemli hale getirmiştir.

Amerika Birleşik Devletleri'nde eğitim ortamı dışındaki amaçlar için; örneğin, sürücü belgesi testleri, askeri eğitim testleri, özel sektörde işe alım ve üniversiteye giriş sürecinde uygulanan testler bilgisayar ortamında uygulanmaktadır. Avrupa'da ise Danimarka ve İzlanda, ulusal çapta uyguladığı kağıt-kalem testlerini bilgisayar ortamına taşımıştır. Bununla birlikte Norveç ve Singapur da ulusal çapta uygulanan testleri bilgisayar ortamına geçirme hazırlıklarını sürdürmektedir (Programme for International Student Assessment [PISA], 2010; Russel, Goldberg ve O'Connor, 2003; Clariana, Wallace, 2002).

Bilgisayar ortamında uygulanan testler konusunda yürütülen çalışmaların ardından Amerikan Psikologlar Derneği (APA), 1986 yılında *Bilgisayar Ortamında Uygulanan Testler ve Yorumlama Rehberi (Guidelines for computer-based psychological tests and interpretations)* adında bir rehber yayınlamıştır. Bu rehberde, bilgisayar ortamında uygulanan test geliştiricilerine ve uygulayıcılarına belirli önerilerde bulunulmuştur. Bunun yanında Uluslararası Test Komisyonu (International Test Commission, ITC), 2005 yılında *Bilgisayar ve İnternet Ortamında Uygulanan Testler İçin Uluslararası Kılavuz (International Guidelines on Computer-Based and Internet Delivered Testing)* adı altında bir kılavuz hazırlamıştır. Bu kılavuzda, bilgisayar ya da internet ortamında uygulanan testlerin; geliştiricileri, yayıncıları ve

¹ Bu çalışma "Kağıt - Kalem Formu Ve Bilgisayar Ortamında Uygulanan Genel Yetenek Testinin Psikometrik Özelliklerinin Karşılaştırılması" isimli yüksek lisans tez çalışmasından özetlenerek hazırlanmıştır.

² Arş. Gör., Eskişehir Osmangazi Üniversitesi, Eğitim Fakültesi, Eğitim Bilimleri Bölümü – e-posta: erencan@aybek.net

³ Prof. Dr., Ankara Üniversitesi, Eğitim Bilimleri Fakültesi, Ölçme ve Değerlendirme Bölümü – e-posta: mukhet@yahoo.com

uygulayıcıları için standartlar tanımlanmıştır. ITC (2005) yayınladığı bu kılavuzda; bir testin kağıt-kalem formundan bilgisayar ortamında uygulanacak bir formunun geliştirilmesi halinde aşağıdaki eşdeğerlik kanıtlarının sağlanması gerektiğini belirtmiştir:

1. Kağıt-kalem ve bilgisayar formlarının güvenilirlikleri benzer olmalıdır.
2. Güvenilirlik tahminleri manidar bir ilişki göstermelidir.
3. Bir dış ölçüte göre geçerlilikleri benzer olmalıdır.
4. Ortalama ve standart sapmaları benzer olmalıdır.

Son yıllarda yapılan bazı araştırmalara bakıldığında, öğrencilerin bilgisayar ortamında uygulanan testler ile kağıt-kalem testlerinde gösterdikleri performanslar arasında manidar bir fark olup olmadığının araştırıldığı, ancak bu araştırmalarda geliştirilen testlerden alınan puanlara ait psikometrik kanıtların (geçerlik, güvenilirlik) eşdeğerliğine ilişkin bulgulara yer verilmediği ya da sınırlı olarak yer verildiği görülmüştür (Anakwe, 2008; Akdemir, Oğuz, 2007; Bayazıt, 2007; Ergün, 2002; İlci, 2004; Maguire, Smith, Brallier ve Palm, 2009).

Kağıt-kalem testlerinin geribildirim özellikleri kapsamında çalışmalar yapan Hyland (2000) ve O'Donovan, Price ve Rust (2001), geri bildirim testlere dayalı durum belirleme sürecinin en önemli bölümünü oluşturduğunu belirtmiştir. Bu öneme rağmen kağıt-kalem testlerinde geri bildirim test uygulamasından belirli bir süre sonra verilmektedir ve bu genellikle sadece tüm-test / alt-test puanlarından oluşmaktadır. Ülkemizde ulusal çapta uygulanan testlerde ise verilen geri bildirim yalnızca test puanı, doğru yanıt sayısı gibi nicel sonuçlardan oluşmaktadır.

Kağıt-kalem testlerinde, test gizliliğinin sağlanması konusunda da çeşitli sorunlarla karşılaşabilmektedir. Kullanılan test materyallerinin genel merkezlerden sınav merkezlerine ve sınav merkezlerinden genel merkezlere ulaştırılmasında, test kitapçıklarının basımı ya da cevap kağıtlarının puanlanması aşamalarında güvenliğin ve gizliliğin sağlanması oldukça önemlidir. Geçmiş yıllarda Ölçme, Seçme ve Yerleştirme Merkezi (ÖSYM)'nin ulusal çapta uyguladığı bazı testlerde, test kitapçıklarının çalınması (Hürriyet, 2013; ÖSYM, 2013) sınavın iptal edilmesine sebep olmuştur. Daha sonra testler yeniden hazırlanmış ve bu durum kuruma ve ülkeye fazladan maddi külfet getirmiştir. Kağıt-kalem testlerinde geri bildirim sınırlılığı ve test gizliliğini sağlamanın güç ve masraflı oluşu, bilgisayar ortamında uygulanan testlerin avantajlarıyla aşılabılır.

Bilgisayar ortamında uygulanan testlerle kağıt-kalem olarak uygulanan testlerdeki performanslar çeşitli psikolojik değişkenler bakımından karşılaştırılmıştır. Ergün (2002), yaptığı araştırmada bilgisayar kaygısının bilgisayar ortamında uygulanan testteki performansı olumsuz yönde farklılaştırdığını ifade etmiştir. Bununla birlikte yaş, cinsiyet, geçmiş bilgisayar ve içerik deneyimleri, test konu alanı gibi değişkenlerin, bilgisayar ortamında uygulanan testlerde gösterilen performansı etkileyip etkilemediği farklı araştırmacılar tarafından araştırılmış ancak araştırma bulgularının tutarlı olmadığı görülmüştür. Bazı araştırmalarda testlerin kağıt-kalem ve bilgisayar formlarında gösterilen performanslar arasında manidar bir fark bulunmazken (Anakwe, 2008; Akdemir, Oğuz, 2007, Bayazıt, 2007; PISA, 2010), Clariana ve Wallace (2002) tam tersi bir bulgu elde etmiştir.

Bilgisayar ortamında uygulanan testlerde, bilgisayara yönelik tutum ile bilgisayar aşinalığı arasındaki ilişkilerin araştırıldığı çalışmalar da bulunmaktadır. Smith, Caputi ve Rawstorne (2000) bilgisayara yönelik tutumu; bir bireyin bilgisayara ya da bilgisayar etkinliklerine karşı genel algısı, olumlu ya da olumsuz duygu durumu olarak tanımlamış ve bilgisayar aşinalığı ile bilgisayar tutumları arasında pozitif ve manidar bir ilişki olduğu sonucuna ulaşmışlardır. Taylor, Kirsch, Eignor ve Jamieson (1999) ise geniş bir yanıtlayıcı grubu üzerinde yaptıkları araştırmada, bilgisayar aşinalığının TOEFL'in bilgisayar ortamında uygulanan formunda gösterilen performansla manidar bir ilişki göstermediğini belirtmiştir.

Kağıt-kalem testlerinin yukarıda belirtilen sınırlılıkları sebebiyle bilgisayar ortamında uygulanan testlere geçiş yapmak istenildiğinde öğrenci performansının ve testlerin psikometrik özelliklerinin bilgisayara yönelik tutum, bilgisayar aşinalığı ve cinsiyet gibi değişkenlere göre değişip değişmediğinin belirlenmesi, bilgisayarlı test uygulamalarının pratikte yaygınlaşması bakımından önemlidir. Diğer taraftan genel yetenek testleri çeşitli alt boyutlarda (sözel akıl

yürütme, sayılarla düşünme) genel öğrenme gücünün ölçüsünü veren ölçme araçlarıdır. Genel yetenek testleri, gelecekte bireylerin yöneleceği okul programlarını tercih ederken ve işe eleman alma sürecinde işte başarılı olacakları tahmin etmek üzere uygulanan ölçme araçlarıdır. Geçmişte bu konu üzerine yapılan çalışmaların farklı bulgulara sahip oluşu ve genel yetenek testleri için bu türden karşılaştırmaların yapıldığı bir araştırmaya rastlanmamış olması böyle bir araştırmanın yapılmasına olan ihtiyacı artırmaktadır.

Araştırmanın genel amacı; bir genel yetenek testinin kağıt-kalem formunda ve bilgisayar ortamındaki uygulanmasından elde edilen puanlara ait madde ve test istatistiklerinin anlamlı biçimde farklı olup olmadığı ve genel yetenek testinin her iki uygulama biçiminden alınan puanların bilgisayara yönelik tutum, bilgisayar aşinalığı ve cinsiyetleri farklı olan öğrenciler arasında manidar bir farklılık gösterip göstermediğini ortaya koymaktır. Bu genel amaç doğrultusunda aşağıdaki sorulara yanıt aranmıştır:

1. Genel yetenek testinin ve alt testlerin kağıt-kalem ve bilgisayar formlarından elde edilen puanlara dayalı test istatistikleri (\bar{X} , Ss, KR-20) hangi düzeydedir? İki formdan elde edilen istatistikler arasında manidar bir fark var mıdır?
2. Bir geçerlilik kanıtı olarak genel yetenek testinin kağıt-kalem ve bilgisayar formlarından alınan sözel ve sayısal akıl yürütme alt test puanlarının, sırasıyla Türkçe ve matematik dersi başarı notuyla gösterdikleri ilişkinin düzeyi ve yönü nedir?
3. Bilgisayara yönelik tutum ve bilgisayar aşinalığı puanları bakımından düşük ve yüksek düzeyde yer alan öğrencilerin, genel yetenek testinin bilgisayar formundan alınan puanları arasında manidar bir fark var mıdır?
4. Genel yetenek testinin bilgisayar formundan alınan puan ortalamaları cinsiyete göre manidar bir farklılık göstermekte midir?

YÖNTEM

Araştırma modeli

Bu araştırma, bilgisayar ortamı ve kağıt-kalem formunda uygulanacak genel yetenek testinin psikometrik özelliklerinin karşılaştırılması ve bazı öğrenci özelliklerinin bu formlardan elde edilen test puanlarında farklılığa yol açıp açmadığını sınımayı amaçlayan ilişkisel türde betimsel bir araştırmadır.

Çalışma grubu

Araştırmada, ulaşılabilirlik bakımından uygun olan Eskişehir Büyükşehir Belediyesi sınırları içerisinde kalan ve bilgisayar donanımı yeterli görülen ilköğretim okulları belirlenmiştir. Bu aşamada, teknik özellikleri *yeterli* sayılan (en az 800MHz işlemci, 256MB RAM, 17" monitör, bellek girişi ya da cd sürücü, standart klavye ve fareye sahip) bilgisayarlar çalışmada kullanılmıştır. Belirlenen okullar arasından bilgisayar sayısı, donanım ve yazılım özellikleri dikkate alınarak araştırmanın yürütülmesine uygun olan altı okul seçilmiştir. Seçilen okulların ilköğretim 6. Sınıf düzeyindeki öğrencilerinden oluşan sekiz sınıfa çalışmaya dahil edilmiş ve toplam 200 öğrenci ile araştırma yürütülmüştür. Ancak veri toplanma sürecinde her iki uygulamada da (kağıt-kalem, bilgisayarlı uygulama) bulunamayan öğrenciler olması, eksik yanıtlamalar vb. nedenlerden dolayı analizler 73 erkek, 63 kız; toplam 136 öğrenciden elde edilen veriler üzerinde yapılmıştır. Araştırmanın yürütülebilmesi için gereken izinler Milli Eğitim Bakanlığı'ndan alınmıştır.

Veri toplama araçları

Genel Yetenek Testi (GYT). Araştırmada, Ankara Üniversitesi Eğitim Bilimleri Enstitüsü Ölçme ve Değerlendirme Anabilim Dalı'nda lisansüstü düzeyde açılan *OD506 Genel Yetenek Testlerinin Geliştirilmesi* adlı ders kapsamında 2000 – 2010 yılları arasında geliştirilmiş ve deneme uygulamaları yapılarak nitelikli ölçme yaptığı saptanmış sözel ve sayısal yeteneği ölçen testlerdeki maddeler kullanılmıştır. Bu bağlamda bu maddelerden kapsamı temsil eden ve

psikometrik nitelikleri iyi olduğu bilinen 60 madde seçilmiştir. Madde sayısının belirlenmesinde, GYT'nin yaklaşık bir ders saati içerisinde cevaplanabilecek özellikte olmasına dikkat edilmiştir. Testi oluşturan 60 maddenin, 30'u sözel akıl yürütme ve 30'u sayısal akıl yürütme alt testlerini oluşturmuştur. Sözel akıl yürütme alt testi, 15 madde *sözcükler arasındaki ilişkiler* ve 15 madde *sözcük dağarcığı* bölümlerinden; sayısal akıl yürütme alt testi ise 25 madde içeren *sayı dizileri* ve 5 madde içeren *sayılarla işlem* bölümlerinden oluşmaktadır. Teste alınan maddelerin güçlükleri .26 ile .80 arasında, ayırt edicilik indeksleri ise .31 ile .67 arasında değişmektedir.

Oluşturulan genel yetenek testinin, kağıt-kalem ve bilgisayar formları hazırlanmış ve öğrencilere bu iki form farklı zamanlarda uygulanmıştır. Testin bilgisayar formu, araştırmacı tarafından Microsoft Visual Basic® 2005 programlama dili ve derleyicisi ile hazırlanmış olup, kağıt-kalem formundaki maddeler aynı sıra ve yapı ile bilgisayar ortamında yer almış ve bu şekilde uygulanmıştır. Bilgisayar formu oluşturulurken testin bilgisayar formunun yanıtlanmasında, öğrencilerin aynı kağıt-kalem formunda olduğu gibi bir sonraki, bir önceki maddeye veya diğer alt teste geçebilmesine, ya da maddeyi yanıtızsız bırakabilmesine fırsat tanınmıştır. Öğrenciler maddeleri yalnızca fare yardımıyla yanıtlayabilmektedir. Uygulama sonrasında öğrenciler doğru, yanlış ve boş yanıt sayılarını ekranda görebilmişlerdir.

Bilgisayara Yönelik Tutum Ölçeği. Araştırmacı tarafından tutum nesnesi ile ilgili bilişsel, duyuşsal ve davranışsal öğeleri dengeli olarak temsil eden, yaklaşık olarak yarısı olumlu ve yarısı olumsuz tutum bildiren madde; tutum maddesi yazma ölçütleri (Tavşancıl, 2006) gözetilerek hazırlanmıştır. Bununla birlikte, Bindak ve Çelik (2006) tarafından geliştirilen Bilgisayar Tutum Ölçeği'nin bazı maddeleri de (5, 6, 7, 9, 10, 14, 16, 19, 20, 21) izin alınarak ölçekte kullanılmıştır. Ölçeğin 24 maddelik deneme formu ilköğretim ikinci kademe öğrencilerinden oluşan 200 kişilik bir gruba uygulanmış ve yapılan temel bileşenler faktör analizi sonucunda ölçeğin toplam varyansın %33.10'unu açıklayan tek bir boyuta sahip olduğu görülmüştür (Çizelge 1). Madde seçiminde faktör yük değeri olarak 0.30 ölçüt alınmış ve 0.30'un altında faktör yük değeri veren bir madde ölçekten çıkarılmıştır.

Çizelge 1. Bilgisayara Yönelik Tutum Ölçeği Temel Bileşenler Analizi Sonuçları

Madde No	Bileşen Ortak Varyansı	Boyut 1	Madde No	Bileşen Ortak Varyansı	Boyut 1
M9	.476	.690	M18	.341	.584
M15	.456	.675	M4	.312	.559
M16	.439	.663	M12	.294	.542
M17	.431	.657	M22	.271	.521
M20	.411	.639	M13	.262	.512
M10	.409	.638	M1	.261	.511
M19	.407	.638	M3	.258	.507
M7	.407	.638	M5	.256	.506
M8	.388	.623	M11	.210	.458
M14	.368	.607	M23	.136	.369
M2	.363	.603	M21	.103	.320
M6	.355	.595			

Açıklanan Toplam Varyans: %33.10

Kalan maddeler üzerinden ölçeğe ait Cronbach Alfa iç tutarlık katsayısı hesaplanmış ve .90 değeri elde edilmiştir. Bu değer, ölçeğin iç tutarlılık anlamındaki güvenilirliğinin oldukça yüksek olduğunu göstermektedir.

Bilgisayar Aşinalığı Ölçeği: Bilgisayar aşinalığı; bilgisayar deneyimi, bilgisayar kullanım sıklığı, devam edilen bilgisayar kurs sayısı, bilgisayara sahip olma, bilgisayara erişme ve bilgisayara yönelik tutumlarla ilişkilendirilmiştir (Taylor ve diğerleri, 1999). Araştırmada da bu ilişkilendirmeye uygun maddeler barındıran ve PISA'nın 2009 yılı uygulamasında kullanılan Bilgi İletişim Teknolojileri Aşinalığı Anketi (*Information Communication Technology [ICT]*)

Familiarity Component For The Student Questionnaire)’nden faydalanılmıştır. Anket, 10 başlık altında yer alan maddelerden oluşmakta ve öğrencilerin bilgi iletişim teknolojilerine dair aşinalıklarını belirlemeyi amaçlamaktadır (PISA, 2009). Maddelerin araştırma amacına uygunluğu gözetildiğinde, 4. başlık altında yer alan maddelerin bilgisayar kullanımı ve kullanım sıklığı bakımından bilgisayar aşinalığını yansıttığı görülmüş ve araştırma kapsamına bu maddeler alınmıştır. Maddeler tutum ifadesi içermeyip, yalnızca bilgisayarda gerçekleştirilen etkinlikleri kapsadığından, Türkçe’ye araştırmacı tarafından çevrilmiş ve daha sonra iki dil uzmanının görüşlerine başvurulmuştur. Yapılan temel bileşenler analizi sonucunda 4. başlık altında yer alan maddelerin toplam varyansın %45.88’ini açıklayan tek bir boyut altında toplandığı görülmüş ve analiz sonuçları Çizelge-2’de verilmiştir.

Çizelge 2. BİT Aşinalığı Anketi 4. Madde Temel Bileşenler Analizi Sonuçları

Madde No	Bileşen Ortak Varyansı	Boyut 1
4f	.636	.798
4g	.613	.783
4e	.559	.748
4i	.553	.744
4d	.507	.712
4c	.452	.672
4b	.330	.575
4h	.240	.489
4a	.238	.488
Açıklanan toplam varyans: %45.88		

Buna göre, bu başlık altında yer alan maddelerden bilgisayar aşinalığı ölçeği oluşturulmuş ve ölçeğe verilen yanıtlar toplanarak, öğrencilerin bilgisayar aşinalığı puanları elde edilmiştir. Ölçekten alınan puanın yükselmesi, yüksek bilgisayar aşinalığının; puanın düşmesi ise, bilgisayar kullanımının az olduğunun göstergesidir.

Verilerin toplanması

Veri toplama sürecinden önce, bilgisayar laboratuvarlarında yer alan bilgisayarlar; ekran çözünürlüğü, ekran parlaklığı gibi bilgisayara bağlı değişkenler açısından standart hale getirilmiş ve öğrencilerin testi alması bu şekilde sağlanmıştır. Standartlara uymayan bilgisayarlar araştırmada kullanılmamıştır. Uygulama sırasında öğrencilere sayısal akıl yürütme alt testinde kullanılan işlemleri yapabilmeleri için kağıt ve kalem de dağıtılmıştır.

Veriler, hazırlanan ölçme araçları yardımıyla ilköğretim 6. sınıf öğrencilerinden toplanmıştır. Araştırma grubuna dahil edilen şubeler öncelikle seçkisiz olarak iki gruba ayrılmış; ilk grupta yer alan şubelerde okuyan öğrenciler önce testin kağıt-kalem formunu alırken, diğer grup öncelikle bilgisayar formunu almıştır. Farklı formların uygulanması arasında geçecek süre ise iki hafta olarak belirlenmiştir. Bu sürenin belirlenmesinde, öğrencilerin bir önceki forma verdikleri yanıtları hatırlamamaları ve ölçülmek istenen özellikte değişiklik meydana gelmemesi gerektiği dikkate alınmıştır.

Bilgisayar formu uygulanırken öğrencilere örnek maddelerle birlikte testin yapısı açıklanmıştır. Açıklamaların standart olabilmesi için uygulayıcıya bir yönerge hazırlanmış ve açıklamalar bu yönerge doğrultusunda yapılmıştır. Test uygulamaları boyunca araştırmacı ve ders öğretmeni gözetmenlik yaparak, öğrencilerin özellikle bilgisayar uygulamasında yöneltebileceği sorulara veya karşılaşılabileceği sorunlara yardımcı olmuştur.

Bilgisayara Yönelik Tutum Ölçeği ve Bilgisayar Aşinalığı Ölçeği, GYT uygulama süresi dışında başka bir derste öğrencilere uygulanmıştır.

Verilerin analizi

Verilerin analizinde kullanılan parametrik testlerin en önemli varsayımı, puanların normal dağılım göstermesidir. İlgili analizler için bu varsayımın sağlanıp sağlanmadığı araştırılmıştır. Dağılımın normal dağılımdan aşırı sapma gösterdiği ya da göstermediği kararı içinse; normallik testi, histogram grafiği, çarpıklık ve basıklık değerleri ile ortalama, tepedeğer ve ortanca değerleri incelenmiştir. Çarpıklık ve basıklık değerleri için kabul edilebilir sınırlar ± 1 aralığı olarak kabul edilmiştir (Büyüköztürk, 2009). Normallik varsayımının sağlanmadığı durumlarda ise, parametrik olmayan testler kullanılmıştır.

Testin kağıt-kalem ve bilgisayar formlarına ait ortalama, standart sapma değerleri ile madde gücü ve madde ayırt edicilikleri gibi test ve madde istatistikleri hesaplanmıştır. Madde gücü ve madde ayırt edicilik indekslerinin, test formuna göre farklılık gösterip göstermediği Fisher'in z testi ile araştırılmıştır. Puan ortalamaları arasındaki farkın manidarlığı bağımlı ve bağımsız gruplar için t-testi ile güvenilirlik katsayıları arasındaki farkın manidarlığı ise Feldt testi (Alsawalmeh ve Feldt, 1994) ile sınanmıştır.

İki ölçme aracının güvenilirliği karşılaştırılırken, güvenilirlik katsayılarının Pearson momentler çarpım korelasyon katsayısı ya da Spearman sıra farkları korelasyon katsayısı cinsinden elde edildiği durumlarda Akhun (1984)'un belirttiği korelasyon karşılaştırma teknikleri kullanılabilir. Ancak Cronbach Alfa ve KR-20 gibi, madde kovaryanslarına dayalı olarak hesaplanan güvenilirlik katsayısı karşılaştırmalarında Alsawalmeh ve Feldt (1994), aşağıdaki bağıntıyı önermiştir:

Bağıntı 1. Feldt Testi

$$W = \frac{(1 - \alpha_2)}{(1 - \alpha_1)}$$

α_1 : 1. teste ait güvenilirlik katsayısı

α_2 : 2. teste ait güvenilirlik katsayısı

Bu bağıntıdan elde edilen W değerinin, $F_{(N-1, N-1)}$ serbestlik derecesi değeri ile karşılaştırılması gerektiği ifade edilmiştir. Yukarıdaki bağıntı bağımsız gruplar için güvenilirlik katsayısı karşılaştırmalarında kullanılmakla birlikte; k, tüm ölçme araçlarındaki madde sayısını göstermek üzere $(k - 1) \cdot (N - 1)$ değeri 1000'in üzerine çıktığı durumlarda bağımlı gruplar için de bu formülün kullanılabilmesi belirtilmiştir. Bu bağlamda gerek alt test gerekse tüm teste yönelik güvenilirlik karşılaştırmalarında $(k - 1) \cdot (N - 1)$ değeri 1000'in üzerinde olduğundan, bu bağıntının kullanılması uygun görülmüştür.

BULGULAR

Araştırmanın birinci sorusuna yönelik olarak, genel yetenek testinin kağıt-kalem ve bilgisayar formlarından elde edilen test puanlarına dayalı test istatistikleri (\bar{X} , KR-20, S) belirlenmiş ve bunlar arasındaki farkın manidarlığı sınanmıştır.

Testin kağıt-kalem ve bilgisayar ortamında uygulanan formlarına yönelik olarak hesaplanan test istatistikleri Çizelge 3'te verilmiştir.

Çizelge 3. Kağıt-Kalem ve Bilgisayar Ortamında Uygulanan Formlara Ait Test İstatistikleri

Form	N	\bar{X}	S	$S_{\bar{X}}$	Min.	Maks.	KR-20
Kağıt-Kalem	136	33.56	11.46	.98	11	56	.92
Sözel Akıl Yürütme		16.91	5.93	.51	4	30	.84
Sayısal Akıl Yürütme		16.65	6.50	.56	3	28	.87
Bilg. Ort. Uygulanan	136	31.55	10.77	.92	11	56	.90
Sözel Akıl Yürütme		15.12	5.48	.47	4	26	.80
Sayısal Akıl Yürütme		16.43	6.39	.55	4	28	.87

Çizelge 3'te yer alan ortalama (\bar{X}) ve güvenilirlik (KR-20) değerleri arasındaki farkların manidarlığı araştırılmış ve ortalamaların karşılaştırılmasında bağımlı gruplar için t-testi kullanılmıştır. KR-20 güvenilirlik katsayılarının karşılaştırılmasında ise Felt testinden faydalanılmıştır. Buna göre, genel yetenek testinin kağıt-kalem ve bilgisayar ortamında uygulanan formlarına ait puan ortalamaları karşılaştırılması sonucu elde edilen bulgular Çizelge 4'te verilmiştir.

Çizelge 4. Kağıt-Kalem ve Bilgisayar Ortamında Uygulanan Formlara Ait Puan Ortalamalarına ait t-testi Sonuçları

Alt Test / Form	N	\bar{X}	S	sd	t	p
Sözel Akıl Yürütme						
Kağıt-Kalem	136	16.91	5.93	135	4.99*	.00
Bilg. Ort. Uyg.	136	15.12	5.48			
Sayısal Akıl Yürütme						
Kağıt-Kalem	136	16.65	6.50	135	.54	.59
Bilg. Ort. Uyg.	136	16.43	6.39			
Tüm Test						
Kağıt-Kalem	136	33.56	11.46	135	3.52*	.00
Bilg. Ort. Uyg.	136	31.55	10.77			

*p<.05

Çizelge 4'te görüldüğü üzere GYT'nin kağıt-kalem ve bilgisayar ortamında uygulanan formlarına ait sözel akıl yürütme alt testi puan ortalamaları arasında manidar bir fark bulunmuştur ($t = 4.99$; $p < .05$). Öğrenciler genel yetenek testinin kağıt-kalem formuna ait sözel akıl yürütme alt boyutunda, bilgisayar ortamında uygulanan forma göre manidar olarak daha yüksek puan almışlardır. Aynı zamanda etki büyüklüğü, eta-kare korelasyon katsayısından faydalanılarak hesaplanmış ve $\eta^2 = .16$ olarak bulunmuştur. Bu değer geniş bir etki büyüklüğünü temsil etmekte olup, sözel akıl yürütme testi puanlarında gözlenen varyansın %16'sının testi farklı formda almaya bağlı olduğu ifade edilebilir (Büyüköztürk, 2009).

Genel yetenek testinin kağıt-kalem ve bilgisayar ortamında uygulanan formlarına ait sayısal yetenek puan ortalamaları arasında manidar bir fark bulunmamıştır ($t = .54$; $p > .05$). Bu durum, genel yetenek testini farklı formlarda almanın bu grupta yer alan öğrencilerin sayısal akıl yürütme puan ortalamalarını farklılaştırmadığını göstermektedir.

Genel yetenek testinin tümünden elde edilen puan ortalamaları, testin uygulandığı forma göre manidar bir farklılık göstermektedir ($t = 3.52$, $p < .05$). Öğrenciler, genel yetenek testinin bütünü dikkate alındığında kağıt-kalem formunda, bilgisayar ortamında uygulanan forma göre daha yüksek puan almışlardır. Etki büyüklüğü katsayısı, eta-kare değeri hesaplandığında, $\eta^2 = .08$ olarak bulunmuştur. Bu değer orta düzeyde bir etki büyüklüğünü temsil etmektedir ve buna göre genel yetenek testinin tümünden elde edilen puanlarda gözlenen varyansın %8'inin testi farklı formlarda almaya bağlı olduğu ifade edilebilir (Büyüköztürk, 2009).

Testten elde edilen puan ortalamalarının karşılaştırılmasının ardından, her iki forma ait tüm test ve alt test güvenilirlik katsayılarının karşılaştırılması yapılmıştır. Yapılan karşılaştırmada Feldt testi uygulanmıştır.

Genel yetenek testinin kağıt-kalem ve bilgisayar ortamında uygulanan formlarına ait sözel akıl yürütme alt testinin güvenilirlik karşılaştırması Çizelge 5'te verilmiştir.

Çizelge 5'te yer alan bilgilerden görülebileceği gibi, genel yetenek testinin sözel akıl yürütme ve sayısal akıl yürütme alt testlerinin güvenilirlikleri ve tüm test için elde edilen güvenilirlikler için kağıt-kalem ve bilgisayar ortamında uygulanan formlar arasında manidar bir fark bulunmamıştır ($W < F_{(135,135)}$).

Çizelge 5. Genel Yetenek Testi Kağıt-Kalem ve Bilgisayar Ortamında Uygulanan Formlarına Ait Güvenilirlik Karşılaştırmaları Sonuçları

Form	N	k	KR-20	W
Sözel Akıl Yürütme				
Kağıt-Kalem	136	30	.84	1.25
Bilg. Ort. Uyg.	136	30	.80	
Sayısal Akıl Yürütme				
Kağıt-Kalem	136	30	.87	1.00
Bilg. Ort. Uyg.	136	30	.87	
Tüm Test				
Kağıt-Kalem	136	60	.92	1.25
Bilg. Ort. Uyg.	136	60	.90	

*p<.05

Güvenilirlik değerlerine ilişkin tüm bu bulgular doğrultusunda, genel yetenek testinin kağıt-kalem ve bilgisayar ortamında uygulanan formlarının güvenilirlik değerleri bakımından denk olduğu söylenebilir.

Ölçüt Geçerliliği Kanıtları

Araştırmanın ikinci sorusuna yönelik olarak, öğrencilerin genel yetenek testinin her iki formuna ait sözel ve sayısal akıl yürütme alt test puanları ile okulda göstermiş oldukları sırasıyla Türkçe ve matematik dersi başarıları arasındaki ilişki Spearman sıra farkları korelasyon katsayısı olarak hesaplanmış ve sonuçlar Çizelge 6'da verilmiştir.

Çizelge 6. Ölçüt Geçerliliği Kanıtları

Ders / Form	Kağıt-Kalem	Bilg. Ort. Uygulanan
Türkçe	.38*	.51*
Matematik	.54*	.49*

*p<.05

Çizelge 6'da görüldüğü gibi, öğrencilerin Türkçe dersi başarı puanları ile genel yetenek testinin sözel akıl yürütme alt testinden elde edilen puanlar arasında her iki form için de orta düzeyde ancak manidar ilişkiler bulunmuştur ($p < .05$). Elde edilen korelasyon katsayıları arasında istatistiksel olarak bir fark olup olmadığı araştırılmış ve Bilgisayarlı GYT uygulaması Türkçe dersi başarı notlarıyla anlamlı düzeyde daha yüksek ilişki gösterdiği görülmüştür ($t_{135} = -2.39$; $p < .05$). Bulunan ilişkiler öğrencilerin Türkçe dersi başarı puanlarının elde edildiği sınıf ölçme işlemlerinin ne denli geçerli ve güvenilir olduğuna bağlıdır ve öğrenci ders başarı notları, dersi yürüten öğretmenin kanaatinden etkilenebilmektedir.

Matematik dersi başarıları ile sayısal akıl yürütme testi puanları arasında her iki form için de orta düzeyde manidar ilişkiler bulunmuştur. Türkçe ders başarı puanları gibi, matematik ders başarı puanlarının da elde edilmesinde kullanılan ölçme araçlarının geçerliği ve güvenilirliği ile dersi yürüten öğretmenin kişisel görüşleri elde edilen korelasyon katsayılarını etkileyebilmektedir.

Bulunan korelasyon katsayıları arasındaki fark araştırılmış ve iki korelasyon katsayısı arasında manidar bir fark bulunamamıştır ($t_{135} = 0.96$, $p > .05$). Matematik ders başarıları, sayısal akıl yürütme için bir ölçüt olarak kabul edildiğinde; sayısal akıl yürütme testinin iki formu ölçüt geçerlikleri bakımından eşdeğer kabul edilebilir.

Formlar Arası

Genel yetenek testinin kağıt-kalem ve bilgisayar ortamında uygulanan formlarından elde edilen puanlar arasındaki ilişki araştırılmış ve sonuçlar Çizelge 7’de verilmiştir.

Çizelge 7. Kağıt-Kalem ve Bilgisayar Ortamında Uygulanan Formlardan Elde Edilen Puanlar Arasındaki İlişki

Alt Test	r
Sözel Akıl Yürütme	.74*
Sayısal Akıl Yürütme	.74*

*p<.05

Çizelge 7’den görülebileceği üzere, sözel ve sayısal akıl yürütme alt testine ait kağıt-kalem ve bilgisayar ortamında uygulanan formlardan elde edilen puanlar arasında .74 düzeyinde pozitif yönde ve yüksek derecede; manidar bir ilişki bulunmaktadır. Bu sonuçlar, öğrencilerin genel yetenek testinin kağıt-kalem ve bilgisayar ortamında uygulanan formlarında benzer performans sergilediğini göstermektedir.

Bilgisayara Yönelik Tutum, Bilgisayar Aşinalığı ve Cinsiyet Bakımından Öğrenci Performans Farklılıkları

Araştırmanın üçüncü sorusu için Bilgisayara Yönelik Tutum Ölçeği ve Bilgisayar Aşinalığı Ölçeğinden elde edilen puanlar doğrultusunda, ortanca puan dikkate alınarak, tutum ve aşinalık puanları düşük ve yüksek gruplar belirlenmiştir. Buna göre testin bilgisayar ortamında uygulanan formundan alınan puanların bilgisayara yönelik tutuma ve bilgisayar aşinalığına göre farklılık gösterip göstermediği bağımsız gruplar için t-testi ile incelenmiştir.

Bilgisayara Yönelik Tutum

Sözel ve sayısal akıl yürütme alt testine ilişkin sonuçlar Çizelge 8’de verilmiştir.

Çizelge 8. Bilgisayara Yönelik Tutuma Göre Genel Yetenek Testinin Bilgisayar Ortamında Uygulanan Formundan Elde Edilen Puanlara Ait t-testi Sonuçları

Bilgisayara Yönelik Tutum	N	\bar{X}	S	sd	t	p
Sözel Akıl Yürütme						
Düşük tutum	58	14.10	5.71	130	-1.977*	.00
Yüksek tutum	74	15.93	4.91			
Sayısal Akıl Yürütme						
Düşük tutum	58	14.51	6.31	130	-3.275*	.00
Yüksek tutum	74	17.99	5.78			

*p<.05

Çizelge 8’de görüldüğü üzere, bilgisayara yönelik tutumları düşük ve yüksek grupların, sözel akıl yürütme alt testinin bilgisayar ortamında uygulanan formundan aldıkları puanlar arasında manidar fark olduğu görülmektedir. Bilgisayara yönelik tutumları yüksek olan öğrencilerin puan ortalamaları ($\bar{X} = 15.93$), bilgisayara yönelik tutumları düşük olan öğrencilerin puan ortalamalarından ($\bar{X} = 14.10$) daha yüksek bulunmuştur. Bilgisayara yönelik tutumları düşük ve yüksek grupların, sayısal akıl yürütme alt testinin bilgisayar ortamında uygulanan formundan aldıkları puan ortalamaları arasında da manidar bir fark vardır. Bilgisayara yönelik tutumları yüksek olan öğrencilerin puan ortalamaları ($\bar{X} = 17.99$), bilgisayara yönelik tutumları düşük olan öğrencilerin puan ortalamalarından ($\bar{X} = 14.51$) daha yüksek bulunmuştur.

Bilgisayar Aşinalığı

Sözel ve sayısal akıl yürütme alt testine ilişkin t-testi sonuçları Çizelge 9’da verilmiştir.

Çizelge 9. Bilgisayara Aşinalık Düzeylerine Göre Genel Yetenek Testinin Bilgisayar Ortamında Uygulanan Formundan Elde Edilen Puanlara Ait t-testi Sonuçları

GYT Bilgisayarlı Form	N	\bar{X}	S	sd	t	p
Sözel Akıl Yürütme						
Bilgisayar aşinalığı düşük	59	14.54	5.80	117	-1.347	.18
Bilgisayar aşinalığı yüksek	60	15.92	5.32			
Sayısal Akıl Yürütme						
Bilgisayar aşinalığı düşük	59	15.31	6.89	117	-1.781	.08
Bilgisayar aşinalığı yüksek	60	17.42	6.02			

*p<.05

Çizelge 9’da görüldüğü gibi, sözel ve sayısal akıl yürütme alt testlerinin bilgisayar ortamında uygulanan formundan elde edilen puanlar, bilgisayar aşinalığı düşük ve yüksek gruplar arasında manidar bir fark göstermemektedir.

Cinsiyet

Araştırmanın beşinci sorusu için genel yetenek testinin bilgisayar ortamında uygulanan formundan elde edilen puanlar arasında cinsiyete göre manidar bir fark araştırılmış ve yapılan bağımsız gruplar için t-testi sonuçları Çizelge 10’da verilmiştir.

Çizelge 10. Cinsiyete Göre Genel Yetenek Testi Puanlarına Ait t-testi Sonuçları

Cinsiyet	N	\bar{X}	S	sd	t	p
Sözel Akıl Yürütme						
Erkek	73	14.93	5.20	134	-.425	.67
Kız	63	15.33	5.81			
Sayısal Akıl Yürütme						
Erkek	73	17.30	6.35	134	1.689	.09
Kız	63	15.44	6.34			
Tüm Test						
Erkek	73	32.22	10.62	134	.777	.44
Kız	63	30.77	10.96			

*p < .05

Çizelge 10’da görülebileceği gibi, genel yetenek testinin bilgisayar ortamında uygulanan formuna ait sözel ve sayısal akıl yürütme alt testleri ve tüm testten elde edilen puanlar cinsiyete göre manidar bir farklılık göstermemektedir.

TARTIŞMA ve SONUÇ

Bu çalışmada, bir genel yetenek testinin kağıt-kalem ve bilgisayar ortamında uygulanan formlarına ait psikometrik özellikleri ve öğrencilerin iki uygulamadaki performansları karşılaştırılmış; ayrıca bilgisayara yönelik tutum, bilgisayar aşinalığı ve cinsiyet değişkenlerinin

bilgisayar ortamında uygulanan testte gösterilen performansı farklılaştırıp farklılaştrmadığı incelenmiştir.

Bilgisayar ortamında uygulanan testlerin, kağıt-kalem formunda uygulanan testlerin yerine kullanılabilmesi için öncelikle formların psikometrik özellikler bakımından eşdeğer olması gerekmektedir (ITC, 2005). Bu sebeple genel yetenek testinin kağıt-kalem ve bilgisayar ortamında uygulanan formlarına ait KR-20 güvenilirlik katsayıları karşılaştırılmıştır. Karşılaştırma sözel akıl yürütme, sayısal akıl yürütme ve tüm teste ait güvenilirlik katsayıları üzerinden yapılmış ve bu katsayılar arasında manidar bir fark bulunamamıştır.

Testin ölçüt geçerliliği kanıtı için, sözel akıl yürütme ve sayısal akıl yürütme alt testlerinden elde edilen puanlarla, sırasıyla Türkçe ve matematik dersi başarıları arasındaki ilişkiler araştırılmıştır. Daha sonra kağıt-kalem ve bilgisayar ortamında uygulanan formlardan elde edilen geçerlilik katsayılar karşılaştırılmış ve sayısal akıl yürütme alt testi için ölçüt geçerliliklerinin testin uygulandığı forma göre manidar olarak farklılaşmadığı, buna karşın sözel akıl yürütme alt testine ait ölçüt geçerliliklerinin testin uygulandığı forma göre manidar olarak farklılaştığı görülmüştür. Kullanılan ölçütün sınıf içi ölçmelere dayalı öğretmen tarafından belirlenen öğrenci başarı puanları olduğu göz önüne alındığında, bu farklılığın ölçüt nedeniyle oluşmuş olabileceği de göz ardı edilmemelidir.

Bu sonuçlardan yola çıkarak, genel yetenek testinin *sayısal akıl yürütme* alt testine ait kağıt-kalem ve bilgisayar ortamında uygulanan formlarının psikometrik özellikler bakımından eşdeğer kabul edilebileceği ve birbirinin yerine kullanılabilmesi sonucuna ulaşılabilir. *Sözel akıl yürütme* alt testi içinse formların ölçüt geçerlilikleri ve öğrenci performansları farklılık göstermesine rağmen güvenilirlik katsayılarının formlara göre farklılık göstermemesi ve formlar arası korelasyon katsayılarının yüksek olmasından dolayı, sözel akıl yürütme alt testi için formların büyük ölçüde eşdeğer olduğunu söylemek mümkündür. Formlar arasındaki yüksek korelasyon, öğrencilerin farklı formlardan aldıkları puanlara göre sıralandığında, sıralamalarının büyük ölçüde aynı kaldığı anlamına gelmektedir.

Öğrencilerin genel yetenek testinde gösterdikleri performansların karşılaştırılması sonucunda öğrencilerin sözel akıl yürütme testinin kağıt-kalem formunda manidar olarak daha yüksek puan aldığı görülmüştür. Sayısal akıl yürütme alt testinin iki formundan elde edilen puanlar arasında ise manidar bir farklılık bulunamamıştır. Bu sonuç, sözel testlerin kağıt-kalem ve bilgisayar ortamında uygulanan formlardan elde edilen puanlar arasında manidar farklılığa rastlamayan ancak sayısal testin iki formundan elde edilen puanlar arasında manidar bir fark bulan Kingston (2009)'un elde ettiği bulgular ile çelişmektedir. Sözel akıl yürütme alt testinde öğrencilerin kağıt-kalem formunda daha yüksek puan ortalamasına sahip olması, yeni test uygulamaları ve nitel çalışmalar ile detaylı olarak incelenebilir. Sayısal akıl yürütme alt testinde ise öğrencilerin aldıkları puan ortalamaları arasında fark olmayışı test uygulaması için beklenen ve istenilen bir durumdur. Bununla birlikte sayısal akıl yürütme alt testinde öğrencilerin işlem yapması gereken madde sayısının az olması ve işlemlerini yapabilmeleri için öğrencilere kağıt ve kalem dağıtılmasının böyle bir sonuca neden olmuş olabileceği unutulmamalıdır. Tüm test dikkate alınarak yapılan karşılaştırmada, öğrencilerin kağıt-kalem testinden manidar olarak daha yüksek puanlar aldığı sonucuna ulaşılmıştır. Bu sonuçlar Anakwe (2008), Akdemir ve Oğuz (2007), Bayazıt (2007) ve PISA (2010)'da verilen bulgular ile çelişki, Clariana ve Wallace (2002)'in elde ettiği bulgular ile tutarlılık göstermektedir.

Cinsiyete göre yapılan karşılaştırmada, öğrencilerin sözel ve sayısal akıl yürütme alt testlerine ait bilgisayar ortamında uygulanan formundan elde ettikleri puanlar arasında manidar bir farklılık bulunamamıştır. Bu sonuç, Akdemir ve Oğuz (2008), Anakwe (2008), Clariana ve Wallace (2002)'in bulguları ile örtüşmekte; PISA (2010)'da verilen bulgular ile çelişmektedir.

Bilgisayar aşinalığına göre testin bilgisayar ortamında uygulanan formundan elde edilen puanlar arasındaki farklar araştırılmış ve bilgisayar aşinalığı düşük ve yüksek grupların, testin bilgisayar ortamında uygulanan formundan elde ettikleri puan ortalamaları arasında manidar bir farkın bulunmadığı görülmüştür. Bu sonuç Taylor, Kirsch, Eignor ve Jamieson (1999)'in elde ettiği bulgularla kısmen örtüşmektedir. Buna karşın bilgisayara yönelik tutumu düşük ve yüksek grupların; sözel ve sayısal akıl yürütme alt testlerinin bilgisayar ortamında uygulanan

formlarından elde ettikleri puan ortalamaları arasında manidar farklılıklar bulunmuştur. Başka bir deyişle bilgisayara yönelik tutumları yüksek olan öğrenciler, bilgisayar ortamında uygulanan formdan daha yüksek puan almışlardır.

Araştırma sonuçlarına göre testin kağıt-kalem ve bilgisayar ortamında uygulanan formlarının büyük ölçüde eşdeğer olduğu görülmektedir. Bu alandaki araştırma sayısının artırılarak, merkezi olarak uygulanan bazı testlerin bilgisayar ortamına aktarılması sağlanabilir.

Her araştırma gibi bu araştırmanın da bazı sınırlılıkları bulunmaktadır. Bu nedenle gelecek araştırmalarda bu sınırlılıkları giderici aşağıdaki önerilerde bulunulması mümkündür:

1. Ölçüt geçerliği karşılaştırmalarında öğrencilerin Türkçe ve matematik ders notlarından faydalanılmıştır. Bunun yerine geçerliği ve güvenilirliği önceden belirlenmiş bir ölçüt test kullanılması faydalı olabilir. Bu çalışmada kullanılan genel yetenek testinde kullanılan maddelerin psikometrik nitelikleri gözden geçirilerek, nitelikleri daha iyi ve daha geniş bir madde bankasına dayalı bir test kullanılabilir.
2. Benzer araştırmalarda sözel akıl yürütme puanlarının bilgisayar ve kağıt-kalem formlarında nasıl değiştiği hem nicel hem de nitel veri toplama ve çözümlene teknikleriyle incelenebilir.
3. Öğrencilerin bilgisayar formundaki puanlarının bilgisayara yönelik tutumlarına göre farklılaştığı dikkate alınarak, bilgisayarlı uygulamalara geçilmeden önce tutum değişkeni kontrol edilebilir ya da bilgisayara yönelik tutumları düşük olan bireylere yönelik etkinlikler düzenlenebilir.
4. Yapılan araştırmada genel yetenek testinin kağıt-kalem formu aynı şekilde bilgisayar ortamına aktarılmıştır. Bilgisayar teknolojisinden yararlanılarak farklı türde, etkileşimli maddeler kullanmanın öğrenci performansını nasıl değiştirdiği üzerine çalışmalar yürütülebilir.
5. Araştırmada kullanılan genel yetenek testi sözel ve sayısal akıl yürütme üzerine kurgulanmıştır. Sadece genel yetenek değil, çeşitli ders konu alanlarındaki öğrenci başarısını ölçmeyi amaçlayan testler üzerinde de benzer araştırmalar yürütülebilir.
6. Araştırma çalışma grubundaki bireylerden toplanan veriler ile sınırlıdır. Farklı çalışma grupları ve farklı yaş gruplarında benzer araştırmalar yürütülebilir.
7. Testi alanların her iki uygulamada gösterdikleri test alma davranışlarına ilişkin gözlemleri ve görüşleri araştırılabilir.

KAYNAKÇA

- Akdemir, O. & Oguz, A. (2008). Computer based testing: An alternative for assessment of Turkish undergraduate students. *Computers & Education*. 51, 1198-1204.
- Akhun, İ. (1984). İki Korelasyon Katsayısı Arasındaki Farkın Manidarlığının Test Edilmesi. *Ankara Üniversitesi Eğitim Bilimleri Fakültesi Dergisi*. Cilt:17. Sayı:1.
- Alsawalmeth, Y.M., Feldt, L.S. (1994). A Modification of Feldt's Test of the Equality of Two Dependent Alpha Coefficients. *Psychometrika*. Vol. 59. No:1. 49-57.
- Anakwe, B. (2008). Comparison of Student Performance in Paper-Based Versus Computer-Based Testing. *Journal of Education for Business*. September-October, 13-17.
- Bayazıt, A. (2007). *Çevirim İçi Sınavlar ve Kağıt-Kalem Sınavları Arasındaki Sınav Süresi ve Öğrenci Başarım Farklılıkları*. Yayımlanmamış yüksek lisans tezi. Hacettepe Üniversitesi Fen Bilimleri Enstitüsü, Ankara.
- Bindak, R. & Çelik, H.C. (2006). Öğretmenler İçin Bilgisayar Tutum Ölçeğinin Güvenirlik Geçerlik Çalışması, *Eurasian Journal of Educational Research-EJER*. 22: 38 - 47.
- Büyüköztürk, Ş. (2009). *Sosyal Bilimler İçin Veri Analizi El Kitabı*. PegemA Yayıncılık: Ankara.
- Clariana, R., Wallace, P. (2002). Paper-based versus computer-based assessment: key factors associated with the test mode effect. *British Journal of Educational Technology*. 33, 593-602.

- Ergün, E. (2002). *Üniversite Öğrencilerinin Bilgisayar Destekli Ölçmeden Elde Ettikleri Başarının Kalem-Kağıt Testi Başarısı, Bilgisayar Kaygısı ve Bilgisayar Tecrübeleri Açısından İncelenmesi*. Yayınlanmamış yüksek lisans tezi. Anadolu Üniversitesi Eğitim Bilimleri Enstitüsü, Eskişehir.
- Hürriyet. (2013). *1.5 Milyon Öğrenci Yıkıldı*. Web: <http://webarsiv.hurriyet.com.tr/1999/05/02/113426.asp> 04 Ocak 2014 tarihinde alınmıştır.
- Hyland, F. (2000). ELS writers and feedback: giving more autonomy to students. *Language Teaching Research*.4:33.
- İlci, B. (2004). *Geleneksel Kağıt-Kalem Yöntemi ile ve Bilgisayarda Online Uygulanan Çoktan Seçmeli Sayısal Yetenek ve Sözel Yetenek Testlerine Ait Madde ve Test İstatistiklerinin Karşılaştırılması*. Yüksek lisans tezi.Hacettepe Üniversitesi Sosyal Bilimler Enstitüsü, Ankara.
- International Test Commission. (2005). *International Guidelines on Computer-Based and Internet Delivered Testing*. Web: <http://www.intestcom.org/guidelines> 16 Kasım 2010 tarihinde alınmıştır.
- Kingston, N.M. (2009). Comparability of Computer- and Paper-Administered Multiple Choice Tests for K-12 Populations: A Synthesis. *Applied Measurement in Education*. 22: 22-37
- Maguire, K.A., Smith, D.A., Brallier, S.A., Palm, L.J. (2009). Computer-Based Testing: A Comparison of Computer-Based and Paper-and-Pencil Assessment. *Allied Academies International Internet Conference*. Vol.11.
- O'Donovan, B., Price, M., Rust, C. (2000). The Student Experience of Criterion-Referenced Assessment. *Innovations in Education and Teaching International*. 38,1.
- Öğrenci Seçme ve Yerleştirme Merkezi (ÖSYM). (2013). 2009-PMYO: Sınav İptali. Web: <http://www.osym.gov.tr/belge/1-11560/2009-pmyo-sinav-iptali-01102009.html> 4 Ocak 2014 tarihinde alınmıştır.
- Programme for International Student Assessment. (2010). *PISA Computer-Based Assessment of Student Skills in Science*. OECD Publications.
- Programme for International Student Assessment. (2009). ICT Familiarity Component for the Student Questionnaire. Web: <http://pisa2009.acer.edu.au/downloads.php> 10 Şubat 2010 tarihinde alınmıştır.
- Russel, M., Goldberg, A., O'Connor, K. (2003). Computer Based Test and Validity: a look back into the future. *Assessment in Education*. 10, 279-293.
- Smith, B., Caputi, P., Rawstorne, P. (2000). Differentiating computer experience and attitudes toward computers: an empirical investigation. *Computers in Human Behavior*. 16, 59-81.
- Tavşancıl, E. (2006). *Tutumların Ölçülmesi ve SPSS ile Veri Analizi*. Nobel Yayın: Ankara.
- Taylor, C., Kirsh, I., Eignor, D., Jamieson, J. (1999). Examining the Relationship Between Computer Familiarity and Performance on Computer-Based Language Tasks. *Language Learning*. 49:2, 219-274.
- Tezbaşaran, A. (1996). *Likert Tipi Ölçek Geliştirme Kılavuzu*. Psikologlar Derneği Yayınları: Ankara.