T.C. KOCAELİ ÜNİVERSİTESİ

SOSYAL BİLİMLER ENSTİTÜSÜ

EĞİTİM BİLİMLERİ ANABİLİM DALI

EĞİTİMDE ÖLÇME VE DEĞERLENDİRME BİLİM DALI

TEST VE MADDE İSTATİSTİKLERİNİN HESAPLANMASINA YÖNELİK BİR ÇEVRİMİÇİ ORTAM GELİŞTİRİLMESİ

YÜKSEK LİSANS TEZİ

Gözde TEKALMAZ

KOCAELİ, 2019

T.C. KOCAELİ ÜNİVERSİTESİ

SOSYAL BİLİMLER ENSTİTÜSÜ

EĞİTİM BİLİMLERİ ANABİLİM DALI

EĞİTİMDE ÖLÇME VE DEĞERLENDİRME BİLİM DALI

TEST VE MADDE İSTATİSTİKLERİNİN HESAPLANMASINA YÖNELİK BİR ÇEVRİMİÇİ ORTAM GELİŞTİRİLMESİ

YÜKSEK LİSANS TEZİ

Gözde TEKALMAZ

Dr. Öğretim Üyesi Fatih KEZER

KOCAELİ, 2019

T.C. KOCAELİ ÜNİVERSİTESİ

SOSYAL BİLİMLER ENSTİTÜSÜ

EĞİTİM BİLİMLERİ ANABİLİM DALI

EĞİTİMDE ÖLÇME VE DEĞERLENDİRME BİLİM DALI

TEST VE MADDE İSTATİSTİKLERİNİN HESAPLANMASINA YÖNELİK BİR ÇEVRİMİÇİ ORTAM GELİŞTİRİLMESİ

YÜKSEK LİSANS TEZİ

Tezi Hazırlayan: Gözde TEKALMAZ
Tezin Kabul Edildiği Enstitü Yönetim Kurulu Karar ve No: 20/11/2019 - 29

Prof. Dr. Satılmış TEKİNDAL (imza)

Doç. Dr. Turgay ALAKURT (imza)

Dr. Öğretim Üyesi Fatih KEZER(İmza)

KOCAELİ, 2019

TEZ BİLDİRİMİ

Tez çalışmam içerisinde yer alan bütün bilgilerin etik davranışlar ve akademik kurallar çerçevesinde elde edildiğini, ayrıca tez yazım kurallarına uygun olarak hazırlanan bu çalışmada bana ait olmayan her bilginin kaynağının belirtilerek atıf yapıldığını bildiririm.

Gözde TEKALMAZ

TEŞEKKÜR

Bu çalışmanın gerçekleştirilmesinde, bilgisini daima benimle paylaşan, tez konusunu seçerken isteklerimi göz önünde bulundurup bana yardımcı olan, yüksek lisans eğitimim ve tez sürecim boyunca kendisine danıştığım her anda, elinden gelen yardımı esirgemeyen, herhangi bir sorun yaşadığımda yanımda olup destek veren, çalışkanlığı, titizliği, kibarlığı ve vizyonuyla bana daima örnek olan ve örnek olmaya devam edecek olan, tez sürecimin tamamlanmasından sonra da kendisinden fikir alıp faydalanabileceğimi düşündüğüm değerli tez danışmanım Dr. Öğretim Üyesi Fatih KEZER' e sonsuz teşekkürlerimi sunarım.

Bütün eğitim hayatım, tez çalışma sürecim ve tüm kendimi geliştirme çabalarımda her zaman yanımda olup beni destekleyen ve güç veren, tez yazma sürecinin bütün zorluklarını benimle birlikte göğüsleyen, tez yazma süreciyle birlikte gelen kaygı, stres gibi durumlarımı sağlıklı bir şekilde yönetmemi sağlayıp beni cesaretlendiren, hayatım boyunca da her durumda bana destek ve ilham olan değerli eşim Utku TEKALMAZ'a sonsuz teşekkürlerimi sunarım.

Bununla birlikte tez yazım aşamasında ve yazılım geliştirme aşamalarında fikrine danıştığım, görüşlerini benden esirgemeyen, arkadaşlığıyla da her zaman yanımda olup bu süreci benim için kolaylaştırmaya çalışan değerli arkadaşım Özhan YILDIZ' a teşekkür ederim.

Aslında hiçbir şey için yüksek derecede stres yaşanmaması gerektiğini kanıtlamak istercesine sakinliği ve miskinliğiyle beni daima mutlu eden ve rahatlatan kedimiz Şeftali'ye ve süreç boyunca az ya da çok bana destek olan hayatımdaki herkese teşekkür ederim.

İÇİNDEKİLER

TEŞEKKÜR i
İÇİNDEKİLER ii
ÖZET iii
ABSTRACT iv
ŞEKİLLER LİSTESİ v
TABLO LİSTESİ vi
GİRİŞ 1
BİRİNCİ BÖLÜM 3
1. PROBLEM 3
1.1. PROBLEM DURUMU 3
1.2. AMAÇ 10
1.3. ÖNEM 10
1.4. SAYILTILAR 12
1.5. SINIRLILIKLAR 13
İKİNCİ BÖLÜM 14
2. İLGİLİ ARAŞTIRMALAR 14
ÜÇÜNCÜ BÖLÜM 17
3. YÖNTEM 17
3.1. ARAŞTIRMANIN MODELİ 17
3.2. VERİLER 18
3.3. VERİLER VE TOPLANMASI 18
3.4. VERİLERİN ANALİZİ VE YORUMLANMASI 21
DÖRDÜNCÜ BÖLÜM 37
4. ARACIN GELİŞTİRİLMESİ 37
4.1. Uygulamanın Tasarım Aşaması 37
4.2. Uygulamanın Geliştirilmesi 39
BEŞİNCİ BÖLÜM 56
5. BULGULAR VE YORUMLAR 56
5.1. Geliştirilen uygulamanın SPSS ve TAP paket programı ile karşılaştırılması sonucu elde edilen bulgular 56
5.2. Geliştirilen uygulamanın kullanılabilirlik anketine bağlı olarak elde edilen bulgular. 77
5.3. Uygulamayla ile ilgili genel görüşler 80
SONUÇ VE ÖNERİLER 82
KAYNAKÇA 85

ÖZET

TEST VE MADDE İSTATİSTİKLERİNİN HESAPLANMASINA YÖNELİK BİR ÇEVRİMİÇİ ORTAM GELİŞTİRİLMESİ

Tekalmaz, Gözde

Yüksek Lisans, Eğitimde Ölçme ve Değerlendirme Bilim Dalı

Tez Danışmanı: Dr. Öğretim Üyesi Fatih KEZER

Kasım, 2019

Eğitimin, davranış değişikliği geliştiren bir süreç olduğu düşünüldüğünde, istenilen davranışın bireye kazandırılıp kazandırılmadığı ya da ne ölçüde kazandırıldığını anlamak büyük önem taşımakta, bunu anlamak da ölçme ve değerlendirmeyle mümkün olmaktadır. Okullar göz önüne alındığında, özellikle kalabalık gruplardaki işlevselliği sayesinde çoktan seçmeli maddeler oldukça fazla kullanılmaktadır. Fakat öğrencinin adımlarını, izlediği yolu, testin güvenirliğini görmek sadece maddelere bakarak mümkün değildir.

Bu araştırmada, bahsedilen ihtiyaca yönelik, test ve madde istatistiklerinin, pratik, detaylı ve anlamlı bir analizi için web ortamında bir uygulama geliştirilmiş ve bu uygulamanın geçerliği ve güvenirliği simülasyon veriler kullanılarak belirlenmiştir. Bu çalışmada amaç, ölçme işlemlerinin istatistiksel sonuçlarına ulaşmak isteyip de karmaşıklığından korkarak kullanamayan bireylere ve aynı zamanda tüm istatistiksel verilerini diğer paket programlar yoluyla hesaplayan bireylere de pratik bir istatistiksel web tabanlı uygulama hizmeti sunmaktır.

Bu amaçla hazırlanan uygulama, simülasyon veri elde edilerek Spss ve TAP gibi muadil uygulamalarla karşılaştırıldığında anlamlı derecede geçerli ve güvenilir olduğu görülmüştür. Geriye dönük hizmet verebilmesi ve detaylı rapor çıktısı gibi biricik özellikleriyle de öğretmenler tarafından daha sık kullanılması ve daha kaliteli ölçme ve değerlendirme işlemlerine yol gösterici olması beklenmektedir.

Anahtar kelimeler: Ölçme ve değerlendirme, madde analizi, istatistik.

ABSTRACT

DEVELOPMENT OF A WEB-BASED ENVIRONMENT FOR TEST AND ITEM STATISTICS

Tekalmaz, Gözde

Master of Education, Department of Assessment and Evaluation

Supervisor: Asst. Dr. Fatih KEZER

November, 2019

Considering that education is a process that develops behavior change, it is very important to understand whether or not the desired behavior is gained to the individual, and to understand this through measurement and evaluation.

When schools are considered, multiple-choice items are widely used due to their functionality especially in crowded groups. However, it is not possible to see the steps of the student, the way he / she follows and the reliability of the test, only by looking at the items.

In this study, a web application has been developed for a practical, detailed and meaningful analysis of test and item statistics for the mentioned need and the validity and reliability of this application has been determined by using simulation data. The aim of this study is to provide a practical statistical web-based application service to individuals who want to reach the statistical results of the measurement processes and cannot use them fearing the complexity, as well as individuals who calculate all statistical data through other packet programs.

The application prepared for this purpose was found to be significantly valid and reliable when compared to the equivalent applications such as Spss and TAP by obtaining simulation data. expected.

Key words: Measurement and evaluation, item analysis, statistics.

ŞEKİLLER LİSTESİ

ŞEKİL 1: Örnek CSV dosyası 43
ŞEKİL 2: Kullanıcı Karşılama Ekranı 50
ŞEKİL 3: Javascript Kod Ekranı 51
ŞEKİL 4: Hesap Oluşturma Ekranı 52
ŞEKİL 5: Kullanıcı Oturum Açma Ekranı 52
ŞEKİL 6: Test İstatistiklerini Elde Edebilmek İçin Gerekli Olan Verilerin Giriş Ekranı 53
ŞEKİL 7: Uygulamanın Verdiği Raporun Bir Kısmının Görüntüsü 54
ŞEKİL 8: Eski Sınav İstatistiklerine Ulaşma Ekranı 54

TABLO LİSTESİ

TABLO 1: Kullanılabilirlik anketi 19
TABLO 2: Kullanılabilirlik Anketinde Yer Alan Maddelerin İşlevleri 20
TABLO 3: 30 Kişilik Öğrenci Grubuna Ait 1. Teste Dair İstatistikler 57
TABLO 4: 30 Kişilik Öğrenci Grubuna Ait 2. Teste Dair İstatistikler 58
TABLO 5: 30 Kişilik Öğrenci Grubuna Ait 3. Teste Dair İstatistikler 59
TABLO 6: 100 Kişilik Öğrenci Grubuna Ait 1. Teste Dair İstatistikler 60
TABLO 7: 100 Kişilik Öğrenci Grubuna Ait 2. Teste Dair İstatistikler 61
TABLO 8: 100 Kişilik Öğrenci Grubuna Ait 3. Teste Dair İstatistikler 62
TABLO 9: 500 Kişilik Öğrenci Grubuna Ait 1. Teste Dair İstatistikler 63
TABLO 10: 500 Kişilik Öğrenci Grubuna Ait 2. Teste Dair İstatistikler. 64
TABLO 11: 500 Kişilik Öğrenci Grubuna Ait 3. Teste Dair İstatistikler 65
TABLO 12: 30 Kişilik Öğrenci Grubuna Uygulanan 1. Sınavın Madde İstatistikleri 66
TABLO 13: 30 Kişilik Öğrenci Grubuna Uygulanan 2. Sınavın Madde İstatistikleri 68
TABLO 14: 30 Kişilik Öğrenci Grubuna Uygulanan 3. Sınavın Madde İstatistikleri 69
TABLO 15: 100 Kişilik Öğrenci Grubuna Uygulanan 1. Sınavın Madde İstatistikleri 70
TABLO 16: 100 Kişilik Öğrenci Grubuna Uygulanan 2. Sınavın Madde İstatistikleri 71
TABLO 17: 100 Kişilik Öğrenci Grubuna Uygulanan 3. Sınavın Madde İstatistikleri 72
TABLO 18: 500 Kişilik Öğrenci Grubuna Uygulanan 1. Sınavın Madde İstatistikleri 73
TABLO 19: 500 Kişilik Öğrenci Grubuna Uygulanan 2. Sınavın Madde İstatistikleri 74
TABLO 20: 500 Kişilik Öğrenci Grubuna Uygulanan 3. Sınavın Madde İstatistikleri 75
TABLO 21: 30 Kişilik Öğrenci Grubuna Uygulanan 1. Sınavın Z Ve T Puan Tablosu 76
TABLO 22: Kullanıcıların Tıklanma Sayısı Ve Erişim Süreleri 78
TABLO 23: Kullanıcıların Kullanılabilirlik Anketi Aracılığıyla Uygulamaya Verdikleri Puanlar 79

GİRİŞ

Eğitim, belli bir amaç doğrultusunda bireye istendik davranışlar kazandırma sürecidir (Özçelik, 2014). İstendik davranışların kazandırılması bir sürece ihtiyaç duyar. Bu süreç boyunca, bireye bu davranışların kazandırılma derecesini anlamak ise ölçme ve değerlendirmenin konusudur.

Ölçme değerlendirme işlemleri, eğitimin en önemli aşamalarındandır. Bu yüzden anlamlı bir ölçme değerlendirme süreci var olan ve gelecek eğitim durumlarını da olumlu yönde etkileyebileceği için doğru şekilde yapılması önemlidir. Ölçme ve değerlendirme sürecinde kullanılan ve bu amaca hizmet eden birden fazla araç vardır. Çoktan seçmeli testler, açık uçlu sorular, doğru yanlış testleri bunlardan bazılarıdır. Seçme ve yerleştirme sınavları, sınıf içi ölçümler de göz önüne alındığında, ülkemizde en sık tercih edilen ölçme araçlarından biri çoktan seçmeli testlerdir. Çoktan seçmeli testler, kalabalık gruplarda kullanım kolaylığı ve daha fazla kazanımın kazandırılma derecesini ölçebilme özelliği nedeniyle tercih edilmektedir. Çoktan seçmeli testler, bahsedilen konularda avantajlı olsa da, öğrencinin cevap kağıdına bakarak izlediği adımları anlayabilme gibi konularda da dezavantajlıdır. Bu noktada test ve madde istatistiklerine bakarak teste dair birçok konuda fikir sahibi olmak mümkündür. Aritmetik ortalama, standart sapma, testin güvenirliği, maddenin güvenirliği, maddenin güçlüğü ve ayırt ediciliği bunlardan bir kaçıdır.

Bu araştırmada, öğretmene test ve madde istatistiklerini pratik bir yolla hesaplayabilen, detaylı bir rapor sunan ve bilgilerini geriye dönük depolayabileceği bir çevrim içi ortam geliştirmek amaçlanmıştır. Bu amaç doğrultusunda alan yazın taraması yapılıp, uzman görüşü de alınarak çevrimiçi ortamın geliştirilme süreci başlatılmıştır. Geliştirme ve tasarım süreci tamamlandıktan sonra, çevrim içi ortamın güvenirliği hakkında fikir sahibi olabilmek için TAP ve SPSS programlarının verdiği çıktılarla karşılaştırma yapılmıştır. Karşılaştırma yapılırken kullanılan veriler TAP programı yardımıyla simülatif olarak elde edilmiştir. Elde edilen veriler sonucunda geliştirilen ortamın güvenilir olabildiği görülmüştür. Bununla beraber, uygulamanın kullanılabilirliği hakkında fikir sahibi olabilmek adına öğretmenlere kullanılabilirlik anketi uygulanmıştır. İlk uygulamadan sonra öğretmenlerden görüşler alınıp, bu doğrultuda düzenlemeler yapıldıktan sonra tekrar anket uygulanmıştır.

Düzeltmelerden sonra öğretmenlerin görüşlerinde olumlu anlamda düzelmeler görülmüştür.

BİRİNCİ BÖLÜM

1. PROBLEM

Bu bölümde, araştırmaya ilişkin problem durumu, araştırmanın amacı, önemi, sayıltılar ve sınırlılıklar yer almaktadır.

1.1. PROBLEM DURUMU

Değişen ve gelişen dünyada, keşfedilen ve dünyaya sunulan her yenilik, hem bireyleri hem de toplumu etkileyerek onları değişime mecbur bırakmaktadır. Bu değişime, hızlı bir şekilde adapte olup, elde ettiği bilgileri de anlamlı bir şekilde hayata geçirebilen bireylerden oluşan toplumların daha başarılı ve öncü toplumlar olabileceği ifade edilebilir. Değişime direnmeyi bir meziyet olarak görmekten ziyade, kendini doğru bir şekilde yenileyebilen toplumlar, geleceklerini kendi elleriyle şekillendirecekleri gibi, ona daha emin adımlarla yaklaşmış olup öncü de olurlar. Böyle bir toplumu oluşturan bireyleri yetiştirebilmenin en net ve doğru yolu, eğitimdir. Eğitim, genel anlamıyla insan davranışını istendik yönde değiştirme veya insana yeni davranışlar kazandırma işinin en verimlice nasıl yapılabileceği ile ilgili tüm bilgi, beceri ve işlemleri kapsayan bir disiplindir (Ertürk, 1972).

En genel tanımıyla eğitim; bireyin davranışında kendi yaşantısı yoluyla ve kasıtlı olarak istendik değişme meydana getirme sürecidir (Ertürk,1991, s.12). Tanımdan da görüldüğü üzere eğitimin bir süreç olması durumu, davranış değişikliklerinin hızlı bir şekilde oluşturulamayacağını göstermektedir. Eğitim sistemleri uzun, planlı ve programlı bir şekilde bireyde istenilen davranış değişikliklerini yaratmayı amaçlamaktadır. Bir toplumun bireylerine kazandırmak istenilen davranışlar, o ülkenin eğitim programının ve sisteminin temel taşlarını oluşturur.

Eğitim sisteminin başlıca dört aşamadan meydana geldiği söylenebilir: Bunlar; girdi, süreç, çıktı ve değerlendirmedir (Demirel, 2005). Bu dört öğe şu şekildedir. Girdi; eğitimin var olması için gerekli kaynaklardır denilebilir. Kullanılacak araç ve gereçler, öğretim programları, öğretim ilke ve yöntemleri, para bunlardan bazılarıdır. Süreç, eğitim çalışmalarının yapıldığı zaman dilimidir. Çıktı, bu süreç boyunca hedefler doğrultusunda elde edilen ürünler, bilgi ve beceriler, tutum ve alışkanlıklardır

denilebilir. Değerlendirme ise, elde edilen ürünlerin kalite ve kontrolünün yapıldığı aşamadır.

Eğitim sistemini oluşturan bu dört öğe, birbiriyle etkileşim içerisindedir. Bütün sistemin birbirini etkilediği göz önüne alındığında, bu öğelerden birinde ortaya çıkabilecek bir aksaklık, tüm sistemi etkileyebilir. Eğitim sisteminde var olan ya da olabilecek bir aksaklığın durumu, bireye kazandırılmak istenen davranışların kazanılıp kazanılamadığı ya da ne ölçüde kazanıldığını görebilmek bir kontrol mekanizmasını zorunlu kılmaktadır. Eğitimde ölçme ve değerlendirmenin önemi de tam bu noktada ortaya çıkmaktadır. Eğitim boyunca, istendik davranışların kazandırılıp kazandırılmadığı ya da ne ölçüde kazandırıldığını öğrenmek ileride yapılacak eğitim ve öğretim faaliyetlerinin de niteliğini etkileyeceğinden dolayı oldukça önemlidir.

Bir eğitim sisteminin niteliği hakkında bilgi verebilecek en nitelikli kaynak öğrencinin davranışlarıdır denilebilir. Eğitim boyunca yapılacak ölçme ve değerlendirme işlemleriyle birlikte, öğrenci davranışlarının hangi düzeyde olduğu, yetersizlik varsa hangi noktalarda olduğu ve hatta istenmeyen davranışların oluşup oluşmadığı belirlenebilir (Kutlu, 2003). Başarı ya da başarısızlık durumunun ve derecesinin bilinmesi, daha sonra gerçekleşebilecek eğitim etkinliklerinin daha gerçekli ve esaslı planlanmasına yardım eder (Turgut ve Baykul, 2010).

Ölçme işlemi, aslında hayatın birçok alanında kullanılan, hatta insanlık tarihi kadar eski bir kavramdır denilebilir. Ölçme, gözlenen niteliklerin bir takım sayı ve sembollerle sayısallaştırılarak ifade edilmesidir. Değerlendirme ise, ölçme işleminden elde edilen verilerin belirli ölçütlerle karşılaştırılarak, ölçülen o niteliğe dair bir karar verme sürecidir (Tekin, 2012). Ölçme ve değerlendirme birbirine çok karıştırılan kavramlar olsa da, aslında birbirinden oldukça farklı kavramlardır. Fakat aynı zamanda birbirlerini de tamamlamaktadırlar. Çünkü eğer sonunda bir değerlendirme işlemi yapılmayacaksa, bir ölçme işlemi de yapmanın pek bir anlamı yoktur ya da bir amaca hizmet etmeyeceği söylenebilir.

Ölçme ve değerlendirme, muhakkak eğitimin vazgeçilemez bir öğesidir. Yenilenen eğitim programlarıyla birlikte benimsenen yapılandırmacı yaklaşımın, bireyin öğrenme süresi boyunca, edindiği yeni bilgiler ve eskiden sahip olduğu bilgiler arasında bir bağ kurarak kendine özgü bir dünya görüşü ve bakış açısı oluşturduğu kabul edilir (Brooks, 2002). Davranışçı yaklaşımın benimsendiği klasik sınıflarda ise

ölçme değerlendirme genellikle açık uçlu sorular, doğru-yanlış testleri ve çoktan seçmeli testlerle yapılmaktadır (Turan-Oluk ve Ekmekçi, 2017). Seçme ve yerleştirme sınavlarında ve sınıf içi ölçmelerde de sıklıkla tercih edilen ölçme aracı çoktan seçmeli testlerdir. Çoktan seçmeli testlerde ise sadece öğrencinin cevabına bakarak detaylı bir bilgi sahibi olabilmek zordur. Bu durumda, sınıfta öğretmen tarafından verilen bilgilerin ne ölçüde kazanıldığını nasıl öğrenilebilir? Bu sorunun cevabı, geçerliği ve güvenirliği sağlanmış ölçme araçlarının kullanıldığı ölçme ve değerlendirme işlemleridir.

Eğitim ve öğretimin var olduğu bir yerde, öğretimin kontrol mekanizması olan ölçme ve değerlendirme mutlaka sürecin bir parçası olmalıdır. Eğitimin ve öğretimin sağlıklı bir şekilde devam edebilmesi için, öğrencilerin hazırbulunuşluk seviyelerini bilmeye, başarısız öğrencilerin öğrenme eksikliklerini gidermede ve amaçların gerçekleşme düzeyi hakkında bir yargıya varma konusunda ölçme ve değerlendirme büyük bir öneme sahiptir.

Derslerde edinilmesi beklenen kazanımların gerçekleşme düzeyini saptamak için nitelikli ölçme araçlarına ihtiyaç vardır. Bu ölçme araçlarıyla elde edilen ölçümlerin geçerlik ve güvenirliklerinin de yüksek olması beklenmekte ve gerekmektedir. Geçerlik, ölçme sonuçlarının ölçülen şeyi yansıtma derecesidir (Tekindal, 2014). Cronbach (1990) ise geçerliği, test puanlarının ne anlama geldiği hakkındaki düşünceleri desteklemek için, test geliştiricilerin kanıt toplama süreci olarak tanımlar. Gronlund ve Linn (1995)' e göre, geçerlik, bir varlıktan ya da yokluktan ziyade, bir dereceyi ifade eder. Buna göre ölçümlere bakarak "geçerlidir" ya da "geçerli değildir" şeklinde bir karara varmaktan ziyade, yüksek, orta ya da düşük derecede geçerlikten söz edilebilir. Aslında geçerlik kavramıyla, bir ölçümden elde edilen ölçme sonucunun, ölçülmek istenen değişkeni, temsil edip etmediği; temsil ediyorsa da ne derece temsil edebildiği ifade edilir. Bir aracın da geçerli olabilmesinin ön şartı güvenilir olabilmesidir.

Geçerlik ile birlikte, iyi bir ölçme aracında bulunması gereken özelliklerden biri de güvenirliktir. Genel olarak güvenirlik, bir testin ölçme sonuçlarının (test puanlarının) tesadüfi hatalardan arınık olma derecesi olarak tanımlanmaktadır (Turgut, 1992). Diğer bir görüşe göre ise güvenirlik, ölçme sonuçlarının tutarlılığı ya da tekrarlanabilirliğidir (Crocker ve Algina, 1986). Testten elde edilen cevaplara göre,

farklı yöntemler kullanılarak testin güvenirliği hakkında fikir sahibi olunabilir. İyi bir ölçme aracında geçerlik ve güvenirliğin yüksek olması gerekir.

Ölçme aracı olarak farklı tiplerde testler kullanılabilir. Hazırlanan bu testler, öğrenciler için uygun fiziksel ve psikolojik şartların oluşturulmasıyla uygulanır ve elde edilen verilere göre iyileştirilme yapılır. Her test geliştirme işleminin sonunda, elde edilen veriler ışığında aslında daha geçerli ve güvenilir bir ölçme aracının oluşturulması için de çalışılır. Test geliştirme işleminin, birçok aşamadan oluşan bir sistematiği olduğu söylenebilir. Bu aşamalar, uygulanacak olan testin türünün ve düzeyinin öğrencilere önceden bildirilmesi, soru havuzunun oluşturulması, testte öğrencilere sorulacak soruların belirtke tablosuna göre seçilmesi, testin düzenlenmesi, öğrencilere uygulanması ve madde analizi yapılması şeklinde sıralanabilir (Özçelik, 1992).

Uygulanan testler hakkında genel bilgi sahibi olmak için, madde analizinin yanı sıra, teste dair de bazı istatistiksel bilgilere ihtiyaç vardır. Güvenirlik katsayısı, standart sapma, aritmetik ortalama, mod, medyan gibi değerler kişiye testin geneli ya da öğrencilerin başarı düzeyi hakkında bilgi verirken, madde güçlük indeksi, madde ayırt edicilik indeksi gibi değerler teker teker o testte var olan her bir soru hakkında bilgi verir. Tüm bu basamaklar titizlik ve kararlılıkla uygulandığında eğitimde ölçme ve değerlendirme işlemlerinin kalitesini önemli derecede iyileştirebilir. Bu noktada, her uygulanan test sonrasında, birden fazla sınıfta görev yapan bir öğretmen için, her seferinde uzun matematiksel işlemler yapmak ve daha sonraki ölçme işlemleri için de yine her sınıf adına bütün sonuçları takip edebilmek gerçekten karmaşık olabilir. Bu noktada, teknolojiden destek alabilmek, bir paket programdan ya da web tabanlı bir uygulamadan destek alabilmek büyük önem taşır.

Teknolojik gelişmeler toplumsal yaşamın her alanında değişikliklere neden olmaktadır. Bu değişimler, eğitim kurumlarının da işleyişlerine etki etmektedir. Gelişen ve değişen dünyada bütün sistemler, teknolojiyi kullanabilen bireylerin yetişmesini beklemekte ve önceliğini teknolojiye ayak uydurabilmiş ve içselleştirebilmiş bireylerden yana kullanmaktadır. Teknolojiyi kullanabilen ve içselleştirebilmiş bireylerin varlığı, teknolojiyi kullanabilen ve içselleştirebilmiş öğretmenlerin yetiştirilmesine ihtiyaç duyar. Dolayısıyla eğitim sistemi de öğretmenlerden diğer her sistem gibi bu işlevi beklemektedir. Üstelik bu beklenti

sadece teknoloji kullanımını öğretmeyi değil, aynı zamanda öğretim tekniklerinde bu işlevi kullanmayı da kapsamaktadır. Bu durumda okullarda hali hazırda görev yapan öğretmenlerin ve fakültelerde yetiştirilmekte olan öğretmen adaylarının teknoloji konusunda eğitimli olmaları gerekmektedir. Bu durumda, Türkiye'de öğretmenlerin teknolojiye dayalı yeterlikleri üzerine alan yazın taraması yapıldığında, Özdemir'in (2007) İngilizce öğretmenleri ile gerçekleştirdiği çalışmasından elde edilen sonuçlara göre; öğretmenler, teknoloji kullanma açısından, kendilerini diğer alanlara göre daha az yeterli görmektedir. Fendi'nin (2007) ve Ertürk'ün (2008) gerçekleştirmiş oldukları tez çalışmalarında da benzer sonuçlara ulaşılmıştır.

Diğer yandan, öğretmenlerin mesleki yeterliği de eğitim kalitesini doğrudan etkileyen önemli bir konudur. MEB tarafından yayınlanan Öğretmenlik Mesleği Genel Yeterlikleri Kılavuzu'na göre, öğretmenlerin sahip olması gereken en önemli yeterliklerden birisi de ölçme ve değerlendirme konusundadır. Yapılan çalışmaların önemli bir bölümü öğretmenlerin sınıf içi ölçme ve değerlendirme bilgi ve becerilerinin alternatif ölçme ve değerlendirme yöntemleri konusunda bilgilerinin yeterli olmadığını istenen ve olması gereken düzeyin altında olduğunu göstermiştir (Bıçak ve Çakan, 2004; Daniel ve King, 1998; Güven, 2001). Kubiszyn ve Borich (1996) öğretmenlerin bilgi ve beceri sahibi olması gereken ölçme ve değerlendirme konularını şu şekilde sıralamıştır:

- Farklı amaçlarla yapılması planlanan ölçmeler için hangi test türlerinin uygun olduğu,
- Dersin yoklanması gereken hedeflerinin neler olduğunu doğru tespit edilebilmesi,
- Yazılı yoklama türü testlerin nasıl geliştirilmesi gerektiği bilgi ve becerisi,
- Testlerin güvenirlik ve geçerliğinin nasıl sağlanacağı, temel test istatistikleri,
- Test puanlarının nasıl kullanılması gerektiği
- Öğrenci başarılarının ya da notlarının etkili bir iletişim tekniği kullanarak ailelere etkili ve yararlı olacak şekilde nasıl iletilmesi gerektiği.

İlk ve ortaöğrenimde görevli öğretmenlerle yapılan bir araştırmada (Daniel ve King, 1998), öğretmenlerin ölçme ve değerlendirme konularında yeterli seviyede kapsamlı bir bilgiye sahip olmadıkları araştırma sonucu olarak ortaya konmuştur. Aynı

zamanda öğretmenlerin, ölçme ve değerlendirme çalışmalarında kullanılan basit istatistiki bilgilere de sahip olmadıkları gözlemlenmiştir. Bununla birlikte, öğretmenlerin programda yer alan ölçme ve değerlendirme uygulamaları konusunda problemler yaşadığı; bu konuda kendilerini diğer alanlara göre daha yetersiz gördükleri; ölçme ve değerlendirme konusunda eğitim ihtiyacı içinde olduklarını ifade ettikleri, çeşitli araştırmalarla belirlenmiştir (Yapıcı ve Demirdelen 2007; Gözütok Akgün ve Karacaoğlu 2005; Yaşar vd, 2005). Öğretmenlerin ölçme ve değerlendirme yeterlikliklerine ilişkin yapılan araştırmalar (Birgin ve Gürbüz, 2008; Gelbal ve Kelecioğlu, 2007) öğretmenlerin daha çok geleneksel ölçme yöntemlerini tercih ettiklerini, ölçme tekniklerinin kullanılması ve hazırlanması konusunda öğretmenlerin eğitime ihtiyaç duyduklarını ortaya koymuştur.

Eğitim fakültelerinde yetiştirilmekte olan öğretmen adaylarının eğitimi konusunda ölçme ve değerlendirmeye yeterince yer verilmemesi ve öğretmenlerin ve öğretmen adaylarının ölçme ve değerlendirme konusunda büyük eksikliklerinin olduğu kabul edilen bir durumdur. Halbuki öğretim sürecini uygun bir şekilde ölçüp değerlendirmek kadar, sonrasında elde edilen verilerin istatistiksel analizini yapıp, bir sonraki ölçmeleri şekillendirmek de bir o kadar önemlidir. Yapılan alan yazın taramalarında, önceden de bahsedildiği üzere öğretmenlerin ölçme değerlendirme ve istatistiksel analizler kısmındaki yeterliklerinin seviyesi ve teknoloji kullanımındaki gelişmeye açık durumları ortadadır. (Bıçak ve Çakan, 2004; Güven, 2001) Bundan dolayı, öğretmen adaylarının lisans eğitimi boyunca aldıkları ölçme ve değerlendirme ders içerikleri ve öğretmenlerin teknoloji yeterlik algıları göz önüne alındığında, tüm bu istatistiksel hesaplamaları bilgisayar ortamında gerçekleştiren SPSS, Jmetrik, TAP ve ITEMAN gibi programların kullanımı, gerekli fakat pek de pratik görünmemektedir.

Kısaca bahsetmek gerekirse, SPSS, Jmetrik ve Iteman gibi paket programlar, kullanıcının bilgisayarına yükleyerek kullanabildiği, daha önceden bahsi geçen istatistiksel hesaplamaları ve dahasını yapabilen ve kullanıcıya rapor olarak bildiren paket programlardır. Jmetrik ve TAP tamamen ücretsiz olmakla birlikte, Iteman programının belli temel işlevlerini ücretsiz kullanabilip, daha detaylı analizler için ise, yıllık abonelik fiyatı ödenmesi gerekmektedir. SPSS ise ücretli bir paket programdır. Bahsi geçen paket programların, Türkçe dilinde kullanımının olmaması da, kullanıcı

açısından negatif bir durumdur. Bununla birlikte analizden önce verilerin (öğrenci cevapları, cevap anahtarı gibi) kullanıcı tarafından sisteme girilmesinde takip edilmesi gereken süreç, oldukça karmaşık olabilmektedir. Tüm bu analizlerin sonucuna bağlı olarak verilen raporlar incelendiğinde ise, hangi değerin ne gibi bir duruma karşılık geldiğini bilmek için, ciddi derecede ölçme ve değerlendirme bilgisine ihtiyaç duyulduğu görülmektedir. Öğretmenlerin ölçme ve değerlendirme ve teknoloji konusundaki yeterliklerini içeren çalışmalara bakıldığında, tüm bu aşamaların öğretmenlerin çok küçük bir kısmı tarafından sağlıklı bir şekilde yürütülmesi belki mümkün olabilir.(Ertürk, 2008; Fendi, 2007; Özdemir, 2007)

Ölçme işlemlerinin bu kadar önemli bir aşamasının nitelikli yapılamaması daha kaliteli bir eğitimin önünde engel olduğu düşünülebilir. Testin geneline ve tüm sorulara yönelik tek tek analiz yapabilen, kullanımı kolay, detaylı ve anlaşılır bir rapor veren, bununla birlikte öğretmenin bütün sınıflara ait verilerinin takibini rahatça yapabilmesi için, kendi sisteminde bu bilgileri depolayabilen ve gerektiğinde öğretmene geri sunabilecek bir uygulamanın boşlukları dolduracak şekilde ihtiyaçlara cevap vereceği düşünülebilir. Bu noktada, araştırmacının geliştirdiği web tabanlı uygulama daha önce bahsedilen tüm bu ihtiyaçları karşılamakta, üstelik zaman içerisinde güncellenerek daha geniş işlevlerinin de olabileceğinin vaadini vermektedir.

Sınıf içi ölçmeler sonucunda elde edilen puanlar üzerinden test ve madde analizleri doğru bir şekilde analiz edilip incelendiğinde, o test hakkında ciddi anlamda bilgiler verebilir. Bu bilgiler ışığında da kademeli olarak bir dahaki ölçme işlemlerinin kalitesi artırılabilir. Test ve madde istatistikleri incelendiğinde, birçoğunun öğretmen tarafından elle hesaplanmasının pratik olmadığı düşünülebilir. Üstelik bütün istatistiksel hesapların bir şekilde yapıldığı düşünüldüğünde de (paket programlardan yararlanarak ya da hesap makinesi kullanarak) açık ve anlaşılır bir yorumuna ulaşabilmek, ancak kişinin kendisinde bu bilgiler mevcutsa mümkündür. Bunun dışında, herhangi bir hesaplama sonrasında kişinin yorumlanmış bir rapora ve sınıf gelişimini takip edebileceği anlaşılır bir düzene ulaşması mümkün değildir.

Bu doğrultuda araştırma problemi, öğretmenlerin uyguladıkları testlerden sonra test ve madde istatistiklerine ve buna dair anlaşılır bir rapora ulaşabilecekleri, raporlarını kendilerine ait bir çevrimiçi ortamda saklayıp takibini yapabilecekleri, kişisel bir web uygulaması geliştirmek ve kullanılabilirliği hakkında fikir edinmektir.

1.2. AMAÇ

Bu araştırmanın amacı, test ve madde istatistiklerine ilişkin hesaplamaları yapabilen, kullanıcıya geçmişe yönelik çalışmalarını takip etme imkanı sunan, pratik ve kullanıcı dostu olarak detaylı ve anlaşılır bir rapor sunan, öğretmenlerin kolayca kullanabileceği geçerli ve güvenilir sonuçlar üreten bir çevrimiçi ortam oluşturmaktadır. Bu çevrimiçi ortam oluşturulurken aşağıdaki sorulara yanıt aranmıştır:

1.3. ÖNEM

Ölçme ve değerlendirme, eğitim en önemli ve vazgeçilemez parçalarından biridir. Bir eğitim ya da öğretim süreci boyunca, istenilen davranışların ne derece kazandırıldığı ya da kazandırılıp kazandırılamadığını görmek sürecin en mühim parçasıdır. Ölçme ve değerlendirme işlemlerinin, doğru ve nitelikli bir şekilde yapılması doğrudan eğitim kalitesini etkiler. Bir kazanımın ne derece kazanıldığını doğru sorularla doğru şekilde ölçebilmek, doğru öğrenci grubunu doğru sorularla karşılaştırabilmek, sınavın güvenirliğinden emin olabilmek, öğrencilerin bireysel ve genel başarısını tespit edip, takip edebilmek bu sürecin ülkemizde çoğunlukla ihmal edilen, fakat en önemli parçalarıdır. Burada bahsedilen tüm aşamalar, sistemli ve kararlı bir şekilde uygulanarak, önlem alınması gereken yerlerde doğru adımlar atılarak ilerlendiğinde, tüm ölçme faaliyetlerinin ve dolayısıyla da eğitim kalitesinin artacağı düşünülebilir.

Bahsi geçen adımların ihmal edilmesinin birden fazla sebebi olabilir. Kısaca bahsedilecek olursa;

- Tüm bu hesapları, hesap makinesi kullanarak ya da elle hesaplamak hiç pratik bir yol değildir.
- Bu hesaplamaları yapabilen Jmetrik, SPSS, TAP, Iteman gibi paket programlar bulunmaktadır. Bu programlardan, Jmetrik ve TAP tamamen ücretsiz bir program olup, Iteman'i tam anlamıyla kullanmak isteyen bir kullanıcı yıllık abonelik satın almak zorundadır. SPSS ise ücretli bir paket programdır.
- Bahsi geçen programların, ara yüzlerinin, veri girişlerinin ve raporlarının dili anadilde olmayıp, İngilizce'dir.

- Öğretmenler bahsedilen programları kullanabilecek yeterliğe sahip olabilecekleri bir eğitim almamaktadır.
- Bahsi geçen programlar incelendiğinde, kullanıcıyı kolay bir şekilde kullanıma yönlendirememektedir, yani kullanıcı dostu değildir denilebilir. Bununla birlikte, programları doğru bir şekilde kullanabilmeyi anlatan rehberler, İngilizce'dir ve anlatımlarda görsel ve duyuşsal olarak sınırlı kalıp, düz bir anlatım yolunu seçmişlerdir. Bu noktada, anlatım konusunda açık olmadıkları söylenebilir.
- Aynı zamanda, bu programların kullanıcıya verdikleri raporlar, tamamen sayısal verilerden oluşmakta, tüm verilerin kullanıcı tarafından anlaşılıp, yorumlanması gerektiği beklenmektedir.
- Iteman, SPSS ve TAP programlarında kişiye özel geriye dönük veri depolanamazken, Jmetrik paket programında belli bir kotaya kadar veriler kullanıcı için programda depolanmaktadır.
- Bahsi geçen programlar paket program olduğu için, sadece indirilen bilgisayarda kullanılabilir. Örneğin, evindeki bilgisayarında işlem yapan bir öğretmen, ertesi gün okulunda bilgilerine ulaşmak istese, programın en baştan kurulumunu yapıp, aynı datayı tekrardan analiz etmesi gerekmektedir. Cep telefonu üzerinden ulaşabilmek ya da işlem yapabilmek mümkün bile değildir. Çağımızda mobilitenin ne kadar önemli olduğu göz önüne alındığında, bu oldukça önemli bir problem sayılabilir.

Bu noktada, tüm burada bahsi geçen açıkları kapatmasıyla öne çıkan uygulamanın avantajları şu şekilde sıralanabilir:

- Veri girişinden rapor çıktısına kadar tüm aşamalarda kullanıcı dostudur.
- Kullanım dili, ara yüz, veri girişi, rapor çıktısı, kullanıcı rehberi dahil olmak üzere dili Türkçe'dir.
- Programın çıktı olarak verdiği rapor, salt sayılardan oluşmamakta, aynı zamanda yorum da içermektedir. Örneğin; madde ayırt edicilik indeksi 0.15 olan bir soru için sadece bu değeri yazmaktan ziyade, raporda geliştirilip değiştirilemezse mutlaka testten çıkarılması gerektiği bilgisi de yer alır.
- Her kullanıcı kendi e-mail adresiyle profil oluşturulabilir.

- Her kullanıcının geriye dönük uygulamaları kendi kullanıcı profilinde depolanır.
- Google Firebase sunucu olanaklarıyla yapıldığından dolayı, hem güvenlidir, hem de oldukça yüksek ve artırılabilir bir depolama alanı mevcuttur.
- Bu noktada uygulamada kişinin e mail adresini kullanarak sisteme kayıt olması ve giriş yapması sağlanırken, diğer yandan hiç bir kişisel bilgisine erişim izni istenmez.
- Uygulamanın kullanımı tamamen ücretsizdir, kullanım içi satın almalar gerektirmez.
- Kullanıcı için yol gösterimi niteliğinde oluşturulan rehber, ekran görüntüleri kullanılarak pratikleştirilmiş olup, dili de tamamen Türkçe'dir.
- Rapordaki yorumlardan daha fazla bilgiye ulaşmak isteyen kullanıcılar için, test ve madde istatistiklerinin örneklerle anlatılıp yorumlandığı bir bölüm mevcuttur.
- Tamamen online bir platform olduğu için, her hangi bir zaman diliminde istenilen bilgisayardan ya da telefondan ulaşılıp işlem yapılabilir. Geriye dönük bilgilerin hepsine de kullanıcı girişi yapılarak ulaşılabilir.

Bu araştırma ve geliştirme süreci sonunda, bahsi geçen açıkların kapatıldığı düşünülmektedir. Öğretmenin genel olarak atladığı test ve madde istatistikleri hesaplamaları oldukça pratik bir hale getirildiği için, kullanılıp bu aşamaların atlanılmasının önüne geçilmesi beklenmektedir. Uzun vadede, uygulamanın fazla ve istikrarlı kullancıya ulaşması hedeflendiğinden, bu hedef doğrultusunda ölçme faaliyetlerinin dolayısıyla da eğitimin niteliğinin sistemli bir artışı umulmaktadır.

1.4. SAYILTILAR

Bu araştırmada, simulatif olarak elde edilen verinin, gerçeğe yakın ve anlamlı bir veri ürettiği var sayılmıştır. Aynı zamanda araştırmada, kullanılabilirlik anketinin uygulandığı 15 kişilik öğretmen grubunun ankete samimi bir şekilde cevap verdiği varsayılmıştır.

1.5. SINIRLILIKLAR

Bu araştırma;

1. Simülatif olarak elde edilen verilere dayalı sonuçlarla sınırlıdır.
2. Kullanılabilirlik anketinin uygulandığı 15 kişilik öğretmen grubuyla sınırlıdır.

İKİNCİ BÖLÜM

2. İLGİLİ ARAŞTIRMALAR

Alanyazın taraması yapıldığında, eğitimde ölçme değerlendirme etkinliklerine ilişkin uygulama geliştirilmesi ya da web tasarımına dayalı çalışmalar ve bazı web sayfalarının kullanılabilirliğinin test edildiği çalışmalar bulunmaktadır. Bilgisayar ortamında ölçme değerlendirmeye dayalı yazılım geliştirilmesi çalışmasına örnek olarak Bahçeşehir Üniversitesi'nde, 2008 yılında bir yüksek lisans tez çalışması yapılmıştır. Bu çalışmada, bireye uyarlanmış bilgisayar destekli ölçme ve değerlendirme yazılımı geliştirilmiştir. Geliştirilen bu yazılımla öğrencilerin Microsoft Windows ve Office işletim sistemini kullanmaya yönelik becerileri ölçülmeye çalışılmıştır. Bilgisayar ortamında bireye uyarlanmış test kapsamında, öğrencilerin geleneksel sınav sistemlerine göre daha az soruyla seviyelerinin belirlenmesinden dolayı zaman anlamında tasarruf edilmiş ve öğrencilerin seviyelerinin belirlenmesine bir etkisi olmayacak sorularla uğraşması engellenmiştir. Bireye uyarlanmış testlerde, öğrencinin başarı seviyesi arttıkça soru sayısı azalmaktadır, bu bağlamda avantaj olarak da puan sapmalarının diğer tekniklere göre daha az olduğu saptanmıştır (Cömert, 2008).

Bununla birlikte; Tekedere ve Mahiroğlu'nun (2012) ilkyardım dersi için web tabanlı probleme dayalı öğrenme yazılımının geliştirilmesi çalışması da bulunmaktadır. Bu çalışmada, Sağlık Hizmetleri Meslek Yüksekokulu'nun öğretim programında yer alana ilkyardım dersinin probleme dayalı öğrenme stratejisi ile web üzerinden yürütülebilmesi için web tabanlı bir eğitim yazılımı geliştirilmiştir. Web tabanlı bu yazılım probleme dayalı öğrenme stratejisinin öğretim süreçleri de dikkate alındığından dolayı bu yazılımın, öğrencilerin işbirliği yapmalarına imkân verecek şekilde farklı disiplinlerde de uygulanabilirliği kolay olan bir model olma özelliği taşıdığı da görülmüştür.

Aynı zamanda Aruk'un (2008) , bilişim teknolojilerinin zihinsel engellilerin e-eğitiminde kullanılması ve örnek bir uygulama geliştirilmesi çalışması bulunmaktadır. Bilişim Teknolojilerinin "E-Eğitim" ortamı her yerde her insana ulaşan yapısı sayesinde, zihinsel engelli öğrencilerin eğitiminde var olan problemlerin en aza indirgenmesi, fazladan harcanan maliyet ve zamanın geri kazanılması, eğitimin

mekândan bağımsız bir hale getirilmesiyle engelli bireyin eğitiminde daha hızlı bir gelişim sürecinin yaşanmasına fırsat tanınmaya çalışılmıştır. Bu çalışmada aynı zamanda, Zihinsel Engelliler İçin E-Eğitim Uygulaması yaparak eğitimin daha kolay, daha az masraflı, daha kısa bir sürede, daha zevkli ve daha verimli bir hale gelmesini sağlamak amaçlanmıştır. Tüm bu yönleriyle birlikte de, zihinsel engelli öğrencilerin eğitiminin teknoloji ile birebir ilerlemesine yardımcı olarak, eğitimi daha kalıcı ve daha ilgi çekici bir hale getirdiği gözlemlenmiştir.

Tercan'ın (2010) da bilgisayar uyarlamalı test uygulayabilen bir araç geliştirdiği bir çalışması bulunmaktadır. Çalışmada aynı zamanda klasik test teorisine dayalı testlerin ve bilgisayar uyarlamalı testlerin sınıf içi ve uzaktan eğitimde kullanılmasına yönelik de öğretmen görüşleri yer almaktadır. Araştırmada Sakarya Üniversitesi Eğitim Fakültesi Bilgisayar ve Öğretim Teknolojileri Eğitimi Bölümü'nde öğretim gören 5 öğretmen adayıyla görüşme yapılmıştır. Yarı yapılandırılmış görüşme tekniğinin kullanıldığı bu görüşmelerde, öğretmen adaylarının klasik test teorisi ve bilgisayar test uygulama süreçleri kullanımına dair görüşleri alınmıştır. Araştırmanın sonucunda, öğretmen adaylarının, bilgisayar uyarlamalı test kavramı ile daha önceden karşılamadıkları, eğer önceden hazırlanmış soru havuzları olursa, bilgisayar uyarlamalı testlerden yararlanmaya istekli oldukları, uzaktan eğitimde de bilgisayar uyarlamalı test kullanımının etkili bir ölçme süreci sağlayacağını düşündükleri ortaya çıkmıştır.

Kezer (2013) doktora çalışmasında bilgisayar ortamında bireye uyarlanmış test imkanı sunan bir çevrimiçi ortam geliştirmiştir. Çalışmada, geleneksel kağıt kalem test yöntemi ile bilgisayar ortamında bireye uyarlanmış test yöntemi karşılaştırılmış ve aynı zamanda da bilgisayar ortamında bireye uyarlanmış test yöntemine ilişkin farklı stratejilerin karşılaştırılması amaçlanmıştır. Araştırmanın verilerini, 2012-2013 eğitim öğretim yılında Ankara Üniversitesi Yabancı Diller Yüksekokulu bünyesinde hazırlık sınıfında öğrenim görmekte olan 1166 öğrenciden toplayan araştırmacı, aynı zamanda bireye uyarlanmış test uygulamalarının yapılabilmesi için bir çevrimiçi ortam geliştirmiştir. Araştırma sonucunda, bireye uyarlanmış test uygulamasında, kağıt kalem testine göre madde sayılarında büyük oranda tasarruf sağlandığı görülmüştür.

Bununla birlikte web sayfalarının kullanılabilirliği üzerine de çalışmalar bulunmaktadır. Bu çalışmaların bir kısmı bilgi merkezleriyle ya da eğitim

kurumlarıyla ilgilidir. Kullanılabilirlik çalışmalarının yapıldığı çalışmalara örnek olarak, 2003 yılında Niagara Üniversitesi Kütüphanesi tarafından uygulanan kütüphane web sayfasının kullanılabilirlik çalışması verilebilir. İncelenen bu çalışmada, 14 lisans öğrencisine web sitesinin dolaşım araçları ile ilgili 12 sorunun yöneltildiği ve bu sorulara öğrencilerin verdiği cevaplar doğrultusunda web sitesinin kullanım alanlarının belirlendiği görülmüştür (Niagara, 2003).

Alanyazın incelendiğinde, kütüphanelerin web sayfaları için yapılan kullanılabilirlik çalışmalarından biri de Maryland Üniversitesi'ne aittir. Bu testte farklı bölümlerden dört öğrenci belirlenerek kütüphane web sitesi üzerinde tanımlanmış görevlere erişmeleri istenmiştir. Test sonucunda kullanıcıların kütüphane web sitesini kullanırken hangi alanlarda sıkıntı yaşadıkları saptanmaya çalışılmıştır (University of Maryland, 2007).

Bunlar dışında veri tabanı servis sağlayıcılarının da kullanıcılarına daha iyi hizmetler sunabilmek için çeşitli kullanılabilirlik testlerini kendi ara yüzlerine uyguladığı ve bu doğrultuda tasarımlarını geliştirdikleri görülmektedir. Örnek olarak EBSCOhost veri tabanının 2006 yılında arama ekranları üzerinde yaptığı kullanılabilirlik testi gösterilebilir. Bu çalışmada James Madison Üniversitesi Kütüphanesi kullanılabilirlik laboratuvarından yararlanılmış ve iki grup halindeki 8 öğrencinin, belirlenen arama görevlerini tamamlamaları istenmiştir ve böylelikle bir kullanılabilirlik çalışması gerçekleştirilmiştir (Fagan, 2006). EBSCOhost web sitesinin kullanılabilirliği test edilirken, kullanılabilirlik testinin yanı sıra EBSCOhost web sitesinin çeşitli ara yüzleri, Kent State Üniversitesi Kütüphanecilik ve Bilgi Bilim Okulu kullanılabilirlik laboratuvarlarında göz izleme cihazları ile sayfanın hangi alanlarına kullanıcıların daha çok odaklandığı da belirlenmeye çalışılmış ve elde edilen bu sonuçlar, tekrardan tasarımda değerlendirilmeye alınmıştır (EBSCO, 2009).

Alanyazında yer alan çalışmalara genel olarak bakıldığında, kullanılabilirlik araştırmalarının büyük bir çoğunluğunun araştırma kütüphaneleri üzerine, yazılım geliştirme çalışmalarına bakıldığında ise bireye uyarlanmış test geliştirme, e-öğrenme ortamları üzerine olduğu görülmektedir.

ÜÇÜNCÜ BÖLÜM

3. YÖNTEM

Bu bölümde araştırmanın modeli, veriler, verilerin toplanması ve verilerin analizi bahsedilecektir.

3.1. ARAŞTIRMANIN MODELİ

Ölçme faaliyetleri sonrasında yapılması gereken test ve madde istatistikleri daha pratik bir şekilde hesaplanıp, daha anlamlı bir rapor sunması hedeflenen bu araç geliştirme sürecinde nicel araştırma yöntemi kullanılmıştır. Bu araştırma üç aşamada geliştirilmiştir:

1. Araştırmanın tüm yönleriyle ortaya konulabilmesi için öncelikle uzman görüşü alınıp, alan yazın taraması yapılmıştır.
2. Ürün geliştirme süreci tamamlanmış ve daha sonrasında ortaya çıkan ürün kullanılarak simülatif verilerle birlikte uygulama yapılmıştır.
3. Son olarak ürünün genel etkisi, getirdiği katkı, güçlü yönleri ve geliştirilmesi gereken yönleri belirlenerek ortaya koyulmuştur.

Bu araştırma tasarım ve geliştirme araştırmasıdır. Tasarım ve geliştirme araştırmaları gerçek hayatta karşılaşılabilen problemlere yönelik, somut ve uygulanabilir çözümler ya da ürünler ortaya koymaya yönelik bir yöntemdir. Richey ve Klein (2008, 2014) tasarım ve geliştirme araştırmasını, araç, öğretim materyali, teknoloji gibi ürünlerin geliştirilme süreci ya da bu ürünlerin geliştirilmesinde kullanılan modellerin geliştirilme süreci olarak tanımlamaktadır. Bu iki hedefe yönelik olarak tasarım ve geliştirme araştırması iki kategori altında toplanmaktadır: Tip 1 ve Tip 2. Tip 1 tasarım ve geliştirme aşamasında ürünlerin geliştirilmesi yani ana ürünün ortaya çıkarılması araştırma sürecinin hedefiyken, Tip 2'de bu ürünlerin geliştirilme aşamalarını, süreçlerini içeren modellerin geliştirilmesi araştırmanın temel hedefini oluşturmaktadır. Tasarım ve geliştirme araştırmaları geliştirme odaklı araştırmalar olmasından dolayı, özellikle proje çalışmalarında bu araştırma yönteminin kullanılması önerilmektedir. Tasarım ve geliştirme araştırmaları, doğası gereği yoğun emek harcanması beklenen ve bunun sonucunda da patent alma potansiyeli olabilecek ya da olabilen bir ürünün ortaya çıkabileceği bir araştırma modelidir.

3.2. VERİLER

Araştırmanın veri grubunu; araştırma kapsamında simülatif veriler ve kullanılabilirlik aşamasında, kullanılabilirlik anketi uygulanan 15 öğretmenden toplanan görüşler oluşturmaktadır. Araştırmaya katkıda bulunan bu 15 öğretmenden altı tanesi ortaokul, dokuz tanesi lise düzeyinde öğretmenlik yapmaktadır. Ortaokul düzeyinde öğretmenlik yapanlar öğretmenlerin branşları iki matematik öğretmeni, üç ingilizce öğretmeni, bir görsel sanatlar öğretmeni şeklinde dağılım gösterirken; lise düzeyinde görev yapan öğretmenlerin branşları üç matematik öğretmeni, bir kimya öğretmeni, bir fizik öğretmeni, bir tarih öğretmeni, bir Türk dili ve edebiyatı öğretmeni, iki ingilizce öğretmeni şeklinde dağılım göstermektedir. Öğretmenlerden gelen geri bildirimlere göre, uygulama revize edilmiş olup ve son haliyle de yeniden aynı öğretmen grubuna uygulanarak son görüşleri alınmıştır. Kullanılabilirlik anketinin uygulandığı 15 kişilik öğretmen grubu uygun örnekleme yöntemi ile belirlenmiştir. Uygun örnekleme yöntemi, araştırmacının ihtiyaç duyduğu bir gruba ulaşana kadar en ulaşılabilir yanıtlayıcılardan başlayarak elde ettiği örnekleme yöntemine denir (Büyüköztürk vd., 2015).

Uygulamanın güvenirliğinin test edilmesi için, muadil program çıktılarıyla karşılaştırıldığı aşamada, ortaya konan web tabanlı uygulamanın da çıktılarını oluşturabilmek için TAP programı yoluyla dokuz farklı simülatif veri grubu oluşturulmuştur. Buradaki veriler oluşturulurken 30, 100 ve 500 kişilik veri gruplarının, her birine 0.40, 0.60 ve 0.90 güçlük seviyesinde testlerin uygulandığı varsayılmıştır. Böylece her bir veri grubu için toplamda üç farklı güçlük seviyesinde veri elde edilerek, dokuz farklı durum incelenmesine olanak sağlanmıştır.

3.3. VERİLER VE TOPLANMASI

Uygulamayı, SPSS ve TAP programlarıyla karşılaştırmak için ihtiyaç duyulan veriler daha önce de bahsedildiği üzere, TAP programı yoluyla simülatif bir şekilde elde edilmiştir. Uygulamanın kullanılabilirliğine dair veri toplamak adına da 15 kişilik bir öğretmen grubuna bir anket uygulanmıştır. Bu noktada, geliştirilen web tabanlı uygulamanın kullanılabilirliğini ölçmek adına, Kocaeli'de çalışan 15 öğretmene

Oğuzhan Erdinç tarafından Türkçe'ye uyarlanan kullanılabilirlik anketi uygulanmıştır. Anket ve anket uygulamasına dair özellikler aşağıdaki gibidir.

Tablo 1. Kullanılabilirlik Anketi

1.Genel olarak, yazılımın kullanım kolaylığından memnunum.	1 2 3 4 5 6 7
2.Yazılımı kullanmak basittir.	1 2 3 4 5 6 7
3.Yazılımı kullanarak işlerimi etkin bir şekilde yapabiliyorum.	1 2 3 4 5 6 7
4.Yazılımı rahatlıkla kullanabiliyorum.	1 2 3 4 5 6 7
5.Yazılımı kullanmayı öğrenmem kolay oldu.	1 2 3 4 5 6 7
6.Yazılımı kullanarak kısa zamanda üretken hale geldiğime inanıyorum.	1 2 3 4 5 6 7
7.Yazılımın verdiği hata mesajları, problemleri nasıl gidereceğimi açıkça anlatmaktadır.	1 2 3 4 5 6 7
8.Yazılımın verdiği bilgiler (çevrim-içi yardım, ekran mesajları, diğer bilgiler, vb.) açık ve nettir.	1 2 3 4 5 6 7
9.Yazılımın verdiği bilgiler kolayca anlaşılmaktadır.	1 2 3 4 5 6 7
10.Yazılımın arayüzünü beğendim.	1 2 3 4 5 6 7
11.Yazılımın arayüzünü kullanmak hoşuma gidiyor.	1 2 3 4 5 6 7
12.Yazılım, beklediğim bütün işlevlere sahiptir ve yeterlidir.	1 2 3 4 5 6 7
13.Genel olarak yazılım tatmin edicidir.	1 2 3 4 5 6 7

*Kesinlikle katılıyorum =1, Kesinlikle katılmıyorum =7

*Ölçek ters puanlanmaktadır, düşük puan = iyi kullanılabilirlik

Tablo 1'de maddeleri yer alan anket, IBM'de geliştirilen Computer System Usability Questionnaire (CSUQ), kullanılabilirlik alanında yaygın anketlerden birinin Türkçe'ye uyarlanmış halidir. Ankette bulunan maddeler incelendiğinde; kullanıcıya yazılımın, işlevi, yeterliği, kullanıcı dostu olması, anlaşılır ve bilgilendirici olması gibi konularda sorular yöneltilmiştir. Kullanıcılardan tüm bunları 1'den 7'ye kadar puanlaması istenmiştir. Bu noktada anketin ters kurulduğuna tekrardan dikkat çekmek gerekebilir. 1 puan "kesinlikle katılıyorum" anlamını taşırken, 7 puan "kesinlikle katılmıyorum" anlamını taşımaktadır. Kullanıcının yaşadığı deneyimi bu aralıkta puanlaması beklenmiştir. Bu noktada, anketteki her maddenin ölçmek istediği işlevleri teker teker belirlemekte fayda vardır. Bu sebeple, her maddeye karşılık gelen işlevler Tablo 2'de verilmiştir.

Tablo 2. Kullanılabilirlik Anketinde Yer Alan Maddelerin İşlevleri

Madde No	Maddelerin İşlevi
1	Genel olarak kullanıcı memnuniyeti
2	Yazılımın kullanım kolaylığı
3	Yazılımın etkin kullanımı
4	Yazılımın kullanım kolaylığı kolaylığı
5	Yazılımın öğrenme kolaylığı
6	Yazılımın üretkenliğe etkisi
7	Yazılımın hata mesajlarının anlaşılırlığı
8	Yazılım bilgilerinin (ekran mesajı- çevrimiçi yardım) anlaşılırlığı
9	Yazılılım verdiği bilgilerin anlaşılırlığı
10	Yazılımın ara yüzünün kullanıcı dostu olması
11	Yazılımın ara yüzünün kullanıcıyı kullanıma teşvik etmesi
12	Yazılımın yeterliği
13	Yazılımın genel tatmin ediciliği

Yukarıdaki tablo incelendiğinde; her bir maddenin anketteki işlevine dair bilgilendirme bulunmaktadır. Daha önce de bahsedildiği üzere, kullanıcının yazılımı kullandığı süredeki deneyimini ölçmek ve bu deneyimin etkinliğinin, etkililiğinin, kullanıcı dostu olmasının ölçülmesi üzerine yazılmış olan bu maddeler, uygulanan ankette bulunmaktadır.

Maddelerin işlevlerine ek olarak, anketteki sorular 4 farklı alt boyut oluşturacak şekilde gruplanmıştır ve bu alt boyutlar aşağıdaki gibidir:

Sistem yararlılığı: 1-6. Maddeler

Bilgi kalitesi: 7-9. Maddeler

Arayüz kalitesi: 10-12. Maddeler

Genel memnuniyet: 13. Madde

Aynı zamanda Tablo 22'de belirtilen tıklanma sayıları ve erişim süreleri verilerini elde edebilmek için sayaç ve kronometre kullanılmıştır. Sayaç ve kronometre başlangıç ve bitiş zamanlamalarının kontrolü için gönüllü öğretmenlerden destek alınmıştır. Araştırmacının ve gönüllü öğretmenlerin dahil olduğu bu grup, kullanıcıların kronometre başlangıç ve bitiş sürelerinin kontrolünü ve sayaç işlemlerinin kontrolünü sağlamışlardır.

3.4. VERİLERİN ANALİZİ VE YORUMLANMASI

Toplanan veriler için test ve madde istatistikleri hesaplanmış ve karşılaştırılmıştır. Bu hesaplamalar yapılırken eş zamanlı olarak hem SPSS paket programı, hem de TAP paket programı kullanılarak analiz edilmiştir ve geliştirilen uygulama test edilmiştir. Bu noktada amaç, uygulamanın güvenirliği hakkında fikir sahibi olmaktır. Elde edilen verilerle SPSS paket programının kullanılma sebebi, test istatistiklerine dair, geliştirilen uygulamanın sonuçlarıyla karşılaştırılacak istatistikler elde etmektir. SPSS programıyla madde analizinin yapılamamasından dolayı da, daha önce simülatif veri elde etme konusunda kullanılan TAP programı kullanılmıştır. Simülatif yolla elde edilen veriler, sadece sayısal olarak aynı doğrulukta değeri verip vermemek üzerine incelenmiştir.

Geliştirilen bu web tabanlı uygulamada, tüm bu analizler yapılırken SPSS programının kullandığı formülleri açıkladığı el kitabından ve alan yazındaki kaynaklarda yer alan istatistik formüllerinden yararlanılmıştır. Bundan sonraki kısımda bu hesaplamalardan bahsedilmiştir.

Bu bağlamda test ve madde analizlerine ilişkin bazı önemli bilgiler aşağıdaki gibidir:

Aritmetik ortalama, verilerin toplamının veri sayısına bölünmesiyle elde edilen bir değerdir (Büyüköztürk vd., 2015). Medyan ve ranjla kıyaslandığında ve sadece uç değerlerle ya da sadece yığılma noktalarıyla ilgilenmediği göz önüne alındığında, daha anlamlı ve ileri düzeyde bir fikir verdiği söylenebilir. İleri derecede yapılacak analizlerde kullanılmak için daha elverişlidir. Ortalama hesaplanırken, veri grubundaki tüm verilerin kapsanması, bu tekniğin oldukça önemli bir avantajıdır. Ancak puanların dağılımının çarpıklık göstermesi ya da aşırı derecede uç değerlerin olması gibi bir durumda, sadece aritmetik ortalamaya bakılarak yorum yapılması yanıltıcı olabilir. Bu nedenle aritmetik ortalamanın, dağılımın normale yakın dağılım gösterdiği ya da uç değer ayıklamasının yapıldığı durumlarda kullanılması önerilmektedir. Bu durum;

$$\mu = \frac{\sum X}{N} \quad (1)$$

$\mu =$ Aritmetik ortalama;

$\sum X$ = Verilerin toplamı ;

N= Veri sayısı;

şeklinde formülize edilebilir.

Mod (Tepe Değeri); Crocker ve Algina'ya (2008) göre; bir veri grubunda en çok tekrar eden değerdir. Başka bir deyişle, bir sınıfta en çok alınan puan mod değerini verir. Mod, ortalamanın ve medyanın hesaplanamadığı durumlarda kullanılabilen yüzeysel bir betimleme olanağı veren bir merkezi eğilim ölçüdür (Büyüköztürk vd., 2015). Bir dağılım için mod, genelde bir tanedir, fakat dağılımın normal olmadığı durumlarda birden fazla mod değeri olabilir.

Medyan (Ortanca) ise; bir veri grubundaki değerler küçükten büyüğe doğru sıralandığında en ortada kalan değerdir. Buna göre ortanca, küçükten büyüğe doğru sıralanmış bir ölçüm grubunun orta puanını gösterir (Howitt ve Cramer, 1997). Medyan, bir veri grubunun en ortasındaki değerle ilgilenir. Bu noktada medyan için, uç değerlere yönelik bir çıkarım yapılamayacağı (en düşük ve en yüksek değerler gibi), daha çok yığılmanın olduğu kısımlar hakkında bilgi sahibi olunabileceği söylenebilir. Bu yüzden, daha çok, dağılımın normalden uzak olması, sağı ya da solu çarpık olması durumunda kullanılır.

Ranj; bir veri grubundaki en büyük ve en küçük verinin arasındaki farktır (Tekindal, 2014). Ranj değeri, uç değerlere bakılarak elde edilen bir değer olduğu için, bir test hakkında ranja bakarak karar vermek araştırmacı için tehlikeli olabilir. Ranj değerine bakılarak, grubun homojenliği ya da heterojenliği hakkında fikir sahibi olunabilir. Bu durum;

$$R = D_{max} - D_{min} \quad (2)$$

R = Ranj;

D_{max} = Maksimum değer;

D_{min} = Minimum değer;

şeklinde formulize edilebilir.

Standart sapma; bir veri grubundaki verilerin aritmetik ortalamadan olan farklarının kareleri ortalamasının karekökü̈dür (Büyüköztürk vd., 2015). Aynı zamanda, bir dağılımdaki verilerin tümünü işleme katan ve ileri derecede matematiksel hesaplar için uygun olan standart sapma değeri, en güvenilir ve kararlı değişim ölçüsüdür. Standart sapma, veri grubundaki tüm verilerin ortalamadan farklarının karelerinin ortalamasının karekökü̈dür denilebilir. Evrenleri temsilen seçilen örneklemlerin standart sapma değerlerinin gerçek evren standart sapmasından küçük olma eğilimi vardır. Bu nedenle, örneklem için, standart sapma hesaplanırken formülde payda değerinden 1 eksiltilerek düzeltme işlemi uygulanır. Standart sapma, mod, medyan, aritmetik ortalama gibi değerlere göre daha geniş anlamda bilgi verir ve üzerine yorum yapılabilir. Standart sapma ne kadar büyükse, o veri grubundaki veriler arasında o kadar farklılaşma vardır denilebilir. Bu durum veri grubunu daha heterojen yaparken, aynı zamanda bu veri grubu eğer bir sınıfsa, testin ayırt ediciliğini daha yüksek kılıp, güvenilir bir test olduğunu gösterebilir. Standart sapmanın daha küçük olduğu durumlarda ise, veri grubundaki veriler arası daha az farklılaşma olduğu, yani eğer bu veri grubu bir sınıf ise, herkesin birbirine yakın puanlar aldığı söylenebilir. Bu duruma bakılarak da, ayırt edicilik daha düşük ve sınav daha az güvenilirdir yorumu yapılabilir. Formülü ise aşağıdaki gibidir:

$$S.S = \sqrt{\frac{\sum(X-\mu)^2}{N-1}} \quad (3)$$

S.S = Standart sapma;

X= Her bir öğrencinin puanı;

$\mu =$ Aritmetik Ortalama;

N= Veri (öğrenci) sayısı;

$\sum(X-\mu)^2 =$ Her bir öğrencinin puanından ortalamanın çıkarılarak karesinin alınmasıyla elde edilen toplam;

şeklinde formülize edilebilir.

Varyans; standart sapmanın karesine eşittir. Varyans için de, standart sapmaya ilişkin yapılan yorumlar yapılabilir. Bu nedenle burada sadece kısaca bahsedilmiştir. Formülü ise aşağıdaki gibidir:

$$V = \frac{\sum(X-\mu)^2}{N-1} \quad (4)$$

V = Varyans;

X= Her bir öğrencinin puanı;

$\mu =$ Aritmetik Ortalama;

N= Veri (öğrenci) sayısı;

$\sum(X-\mu)^2$ = Her bir öğrencinin puanından ortalamanın çıkarılarak karesinin alınmasıyla elde edilen toplam;

şeklinde formülize edilebilir

Testin ortalama güçlük katsayısı; bir testin uygulandığı gruba ne kadar zor ya da kolay geldiği hakkında bilgi veren bir katsayıdır. İki farklı yolla bulunabilmektedir. Bunlardan birisi, testin aritmetik ortalamasının (doğru cevaplar 1 yanlış cevaplar için 0 şeklinde puanlanmalıdır), testteki soru sayısına bölünmesiyle elde edilir. Diğer bir yol da, testteki bütün soruların güçlük indeksinin, toplam soru sayısına bölünmesiyle elde edilir. Her ikisi de testin ortalama güçlüğü hakkında yorum yapılmasına olanak sağlar. 0.50 güçlük katsayısına sahip bir test için, ortalama güçlüğe sahiptir denilebilir. Bununla birlikte, bu değer 1 e yaklaştıkça sınavın kolay olduğu, 0 yaklaştıkça sınavın zor olduğu şeklinde yorum yapılabilir. Ortalama güçlük katsayısını veren her iki formül aşağıdaki gibidir:

$$P = \frac{\mu}{K} \quad (5)$$

P = Testin ortalama güçlük katsayısı;

μ = Testin aritmetik ortalaması;

K= Testteki soru sayısı;

şeklinde ya da;

$$P=\frac{\sum P_j}{K} \quad (6)$$

P_j = madde güçlük katsayısı;

$\sum P_j$= madde güçlük katsayılarının toplamı;

K= Testteki soru sayısı;

şeklinde formülize edilebilir.

Testin güvenirlik katsayısı; testin maddelerinin birbiriyle ne derece tutarlı olduğunu gösteren bir katsayıdır (Alıcı vd., 2014). Bu katsayı 1 e yaklaştıkça testin güvenilirliği artar, 0 a yaklaştıkça da testin güvenirliği azalır. Bu çalışmada, çoktan seçmeli testlerde (doğru cevapların 1 yanlış cevapların 0 olarak puanlanabildiği testlerde) güvenirliği belirleme yöntemlerinden olan KR-20 ve KR-21 olmak üzere iki farklı formül iç tutarlılık anlamında kullanılmıştır. KR-20 formülü madde istatistiklerinden yararlanarak bir sonuç elde ederken, KR-21 formülü test istatistiklerine dayalı olarak bir katsayı verir. Her ikisine de ait formüller aşağıdaki gibidir:

$$\text{KR-20} =\frac{K}{K-1}\left(1-\frac{\sum p.q}{V}\right) \quad (7)$$

K= Testteki soru sayısı;

p= Bir soruyu doğru yapanların yüzdesi;

q= Bir soruyu yanlış yapanların yüzdesi;

$\sum p.q$ = Tüm soruların doğru yapılma ve yanlış yapılma yüzdelerinin çarpımının toplamı;

şeklinde ya da;

$$\text{KR-21} =\frac{K}{K-1}\left(1-\frac{K.\mu-\mu^2}{K.V}\right) \quad (8)$$

K= Testteki soru sayısı;

V = Varyans;

$\mu =$ Aritmetik Ortalama;

şeklinde formülize edilebilir.

Testin standart hata katsayısı; test sonuçlarına karışan hata miktarı hakkında oransal olarak değil, test puanı cinsinden bilgi veren bir katsayıdır.(Alıcı vd., 2014) Standart hata katsayını veren formül, testin standart sapma değerine ve güvenirlik katsayısına bağlı olarak elde edilir. Bu durumda, testin standart hatasını veren formül aşağıdaki gibidir:

$$S_H = S.S -\sqrt{(1 - r_{xx})} \quad (9)$$

S_H = Standart hata;

S.S = Standart sapma;

r_{xx}= Herhangi bir yolla hesaplanan testin güvenirlik katsayısı;

şeklinde hesaplanabilir.

Bağıl değişkenlik katsayısı; standart sapmanın, aritmetik ortalamaya bölünerek 100 ile çarpılması sonucu elde edilen bir değerdir(Alıcı vd., 2014). İçerisinde hem standart sapma, hem de ortalamayı barındırdığı için, daha çok bilgi verici bir değerdir denilebilir. Bağıl değişkenlik katsayısının 20'den küçük olması, standart sapmanın aritmetik ortalamaya göre küçük olduğunu, dolayısıyla da sınıftaki puanların birbirine yakın olduğu bilgisini verir. Bu katsayının, 25'den büyük olması ise standart sapmanın aritmetik ortalamaya göre daha büyük olduğu bilgisini verir. Dolayısıyla sınıfta alınan puanlar birbirinden daha uzaktır denilebilir. Standart sapmada da bahsedildiği üzere, puanların yakın olması grubun homojenliğini, birbirine yakın puanlar aldığını, dolayısıyla da sınavın daha az ayırt edici ve güvenilir olduğu bilgisini verebilir. Puanların birbirinden uzak olması durumu ise, grubun heterojenliğini ve birbirinden daha farklı puanlar aldığını gösterir. Bu durumda sınav daha ayırt edici ve güvenilirdir yorumu yapılabilir. Formülü ise aşağıdaki gibidir:

$$B.D.K = \frac{S.S}{\mu} \times 100 \quad (10)$$

B.D.K= Bağıl değişkenlik katsayısı;

$\mu =$ Aritmetik Ortalama;

S.S = Standart sapma;

şeklinde hesaplanabilir.

Çarpıklık katsayısı; pozitif ve negatif değerler alabildiği gibi 0 değerini de alabilir. Katsayının pozitif olması, *sağı çarpık (pozitif kayışlı)* bir dağılım gösterdiğini, negatif olması da *solu çarpık (negatif kayışlı)* bir dağılım gösterdiği bilgisini verir (Arıcı, 2001; Baykul, 2000). Bir veri grubunun sağı çarpık dağılım göstermesi, veri grubunda 'mod < medyan < aritmetik ortalama' şeklinde bir ilişki olduğu bilgisini verir. Bu ilişki incelendiğinde, sınavda en çok tekrar eden notun ve en orta değerdeki notun bile genel ortalamanın altında kaldığı söylenebilir. Bu da sınavın çok zor olduğu veya grubun çok başarısız olduğu şeklinde yorumlanabilir. Bir veri grubunun solu çarpık dağılım göstermesi ise, veri grubunda 'aritmetik ortalama < medyan < mod' şeklinde bir ilişki olduğu bilgisini verir. Bu ilişki incelendiğinde, sınavda en çok tekrar eden notun ve en orta değerdeki notun bile genel ortalamanın üstünde olduğu söylenebilir. Bu da sınavın çok kolay olduğu veya grubun çok başarılı olduğu şeklinde yorumlanabilir. Çarpıklık katsayısının 0 olması ise, grubun normal bir dağılım gösterdiği bilgisini verir. Sınav orta güçlükte olduğu ya da sınav başarısının orta düzeyde olduğu söylenebilir. Çarpıklık katsayısı formülü ise aşağıdaki gibidir:

$$Ç.K = \frac{3.(\mu - Ortanca)}{S.S} \quad (11)$$

Ç.K= Çarpıklık katsayısı;

$\mu =$ Aritmetik Ortalama;

S.S = Standart sapma;

Ortanca = Test puanlarının ortanca değeri;

şeklinde formülize edilir.

Basıklık katsayısı; çarpıklık katsayısı gibi dağılımın simetrikliği hakkında bilgi verir. Çarpıklık katsayısı gibi, pozitif ve negatif değerler alabildiği gibi, 0 değerini de alır. Basıklık katsayısının pozitif olması, dağılımın normalden sivri olduğu bilgisini verir. Bu da grupta alınan puanların birbirine yakın olduğu ve grubun homojen olduğu bilgisini verir. Basıklık katsayısının negatif olması, dağılımın normalden daha basık olduğu bilgisini verir. (Arıcı, 2001; Baykul, 2000) Bu da grupta alınan puanların birbirinden uzak olduğu ve grubun heterojen olduğu bilgisini verir.

$$B.K = \frac{\sum(X-\mu)^4 / n}{SS^4} - 3 \quad (12)$$

B.K = Basıklık katsayısı;

$\mu =$ Aritmetik Ortalama;

S.S = Standart sapma;

n = Öğrenci sayısı;

$(X - \mu)$ = Her bir öğrenci puanının aritmetik ortalamadan farkı

Standart puan (Z ve T puanları); Z puanı, bir testten elde edilen puanların aritmetik ortalaması 0 ve standart sapması 1 olan ve normal dağılım gösteren standart bir puana dönüştürür (Büyüköztürk vd., 2015). Z puanı, verilen bir puanın ortalamanın ne kadar altında ya da ne kadar üstünde olduğunun anlaşılmasına yardımcı olur. Z değeri, verilen ortalamadan uzaklıklarının standart sapmaya oranını gösterir.

T puanı ise, Z puanı ile ilgili bazı problemlerin (negatif puanlar, kesirli puanlar) üstesinden gelmek için elde edilen bir puan türüdür. Bu noktada, Z puanını kullanarak T puanını elde etmek için, Z puanı ilk önce 10 ile çarpılıp, daha sonra 50 eklenir.

Z ve T puanlarının hesabına yönelik puanlar aşağıdaki gibidir:

$$Z = \frac{X-\mu}{SS} \quad (13)$$

Z= Z puanı;

μ = Aritmetik Ortalama;

S.S = Standart sapma;

X=Z puanı hesaplanmak istenen öğrencinin puanı;

şeklinde hesaplanabilir. T puanı ise;

$$T = 10Z + 50 \quad (14)$$

T= T puanı;

Z= Z puanı;

şeklinde hesaplanabilir.

Madde ayırt edicilik indeksi; her bir test maddesinin testle ölçülmek istenen özelliği yansıttığı mantığına dayalıdır (Alıcı vd., 2014). Madde ayırt edicilik indeksi, en basit haliyle, bir sorunun istenilen amaca hizmet edip etmediğini ölçer. Başka bir deyişle, madde ayırt edicilik indeksi, bilenle bilmeyeni ayırt etmeye yarar. Örneğin bir sorunun, ortalamanın üstündeki öğrenciler tarafından doğru cevaplanıp, ortalamanın altındaki öğrenciler tarafından yanlış cevaplanması, o sorunun bilenle bilmeyeni ayırt etme noktasında başarılı olduğu hakkında fikir verir. Ayırt ediciliği yüksek sorulardan oluşan bir sınav, güvenirliğinin de yüksek olduğu konusunda fikir verebilir. Ayırt edicilik indeksi, -1 ve +1 arasında değer alır ve aldığı bu değerlere göre bazı yorumlar yapılabilir. 0,40 ve üstü değer alan maddelerin ayırt ediciliği en yüksek olanlardır. 0.30-0.39 arası için ayırt ediciliği oldukça iyidir denilebilir. 0.20-0.29 arası değer alan maddeler düzeltilmeli ve geliştirilmelidir. 0.00-0.19 arası maddeler ayırt edicilik noktasında çok zayıftır. Eğer düzeltilemiyorsa, mutlaka testten çıkarılmalıdır. 0 ve altında değer alanlarsa, mutlaka çıkarılmalıdır. Madde ayırt edicilik gücünü hesaplamanın birçok yolu vardır. Bunlar korelasyona dayalı olanlar ve grup farklılıklarına dayalı olanlar olarak ikiye ayrılabilir. Korelasyona dayalı olanlar aşağıdaki gibidir (Alıcı vd., 2014):

- Pearson momentler çarpımı korelasyon katsayısı
- Phi katsayısı
- Tetrakorik korelasyon kat sayısı

- Çift serili korelasyon kat sayısı
- Nokta çift serili korelasyon katsayısı

Bunlar arasında en sık kullanılanlar çift serili korelasyon katsayısı ve nokta çift serili korelasyon katsayısıdır. Bu iki katsayıya ait formül ise aşağıdaki gibidir: Çift serili korelasyon katsayısı;

$$R_{çift} = r_{jx} = \frac{\mu_{JD} - \mu}{SS} \cdot \frac{P_J}{Y_J} \quad (15)$$

$R_{çift}$ = Çift serili korelasyon katsayısı (madde ayırt edicilik gücü indeksi);

μ_{JD}= Maddeyi doğru cevaplayanların test puanları ortalaması;

μ = Test puanlarının aritmetik ortalaması;

SS = Test puanlarının standart sapması;

P_J= Maddenin güçlük indeksi;

Y_j =Standart normal dağılımda, maddenin p_j sini q_j den ayıran ordinat değeri;

şeklinde hesaplanır.

Nokta çift serili korelasyon katsayısı ise;

$$Rn\text{-}_{çift} = r_{jx} = \frac{\mu_{JD} - \mu}{SS} \cdot \sqrt{\frac{P_J}{q_J}} \quad (16)$$

$Rn\text{-}_{çift}$ = nokta çift serili korelasyon katsayısı (madde ayırt edicilik gücü indeksi);

μ_{JD}= Maddeyi doğru cevaplayanların test puanları ortalaması;

μ = Test puanlarının aritmetik ortalaması;

SS =Test puanlarının standart sapması;

P_J= Maddenin güçlük indeksi;

q_j = Maddeyi yanlış yapanların oranı;

şeklinde hesaplanabilir.

Bunun dışında diğer yaygın kullanımlardan biri olan grup farklılıklarına dayalı olan yöntem, alt grup ve üst grup analiziyle madde ayırt edicilik indeksini elde eder. Kısaca anlatmak gerekirse, alt grup-üst grup analizine dayalı madde ayırt edicilik indeksi bulunurken şu yollar izlenir.

- Testten elde edilen sonuçlar büyükten küçüğe sıralanır.
- Kişi sayısı 0.27 ile çarpılarak %27 lik dilimin kaç kişiye denk geldiği hesaplanır.
- 2. Adımda elde edilen sayıya göre, en yüksek puan alan ve en düşük puan alan kısımdan o kadar kişi alt grup ve üst grup olarak ayrılır.
- Bu noktada, her gruptaki kişilerin hangi sorulara ne cevap verdiğine bakılır.
- Daha sonra aşağıdaki formül kullanılarak madde ayırt edicilik indeksi hesaplanır.

$$R_{jx} = \frac{n_{dü} - n_{da}}{n} \quad (17)$$

R_{jx} = Ayırt edicilik indeksi;

$n_{dü}$ = Maddeyi üst grupta doğru cevaplayan bireylerin sayısı;

n_{da} = Maddeyi alt grupta doğru cevaplayan bireylerin sayısı;

n= Alt ya da üst grupta yer alan birey sayısı

şeklinde formülize edilir.

Madde güçlük indeksi; testi cevaplayan öğrencilerin doğru ya da yanlış cevap vermelerine göre değer kazanan bir indekstir. Tıpkı ayırt edicilik indeksindeki gibi; alt grup-üst grup farklılıklarına göre değer kazanan bir indekstir. Formülü incelenerek ifade edilmek istenen daha iyi anlaşılabilir:

$$P_j = \frac{n_{dü} + n_{da}}{N} \quad (18)$$

PR_j = Güçlük indeksi;

$n_{dü}$ = Maddeyi üst grupta doğru cevaplayan bireylerin sayısı;

n_{da} = Maddeyi alt grupta doğru cevaplayan bireylerin sayısı;

N= Alt ve üst grupta yer alan toplam birey sayısı;

şeklinde hesaplanabilir.

Bu indeks, o sorunun sınıfa ya da o gruba ne kadar kolay ya da zor geldiğinin bilgisini verir. Madde güçlük indeksi, 0 ile 1 arasında değer alır. 1 değerini alması, sınıftaki herkesin o soruyu doğru cevapladığını gösterir, 0 değer alması da kimsenin doğru cevaplayamadığını gösterir. Her iki durum da, sorunun bireyler arası farklılıklara ilişkin pek bir bilgi vermediği ortadadır. Bu yüzden hazırlanan bir sınavı oluşturan tüm soruların bu değerlerden oluşması, sınavın amacına hizmet etmesini engeller. Bu noktada, madde güçlük indeksi 1 e yaklaştıkça, sorunun sınıfa kolay geldiği, 0 yaklaştıkça da zor geldiği yorumu yapılabilir. 0,50 değeri tam olarak orta güçlükte bir sorunun sahip olacağı madde güçlük indeksi değeridir.

Madde varyansı; her bir doğru cevaba 1, her yanlış cevaba da 0 puan verilerek puanlanan bir test maddesine ait varyans aşağıdaki gibi hesaplanır (Alıcı vd., 2014):

$$V_{madde} = p_j.q \quad (19)$$

V_{madde} = Madde varyansı;

p_j = Madde güçlük indeksi;

q = Maddenin yanlış cevaplanma oranı;

şeklinde hesaplanabilir.

Madde standart sapması da; madde varyansının karekökünü alarak aşağıdaki gibi hesaplanabilir.

$$SS_{madde} = \sqrt{pj.q} \quad (20)$$

SS_{madde} = Madde standart sapması;

p_j = Madde güçlük indeksi;

q = Maddenin yanlış cevaplanma oranı;

şeklinde hesaplanabilir.

Madde varyansı ya da standart sapması maddelerin ölçülen özellik açısından bireyler arası farklılıkları ne derece ortaya koyduğunun bilgisini verir. Bir maddenin

varyansının maksimum olması istenen bir durumdur. Maddenin varyansı, maksimum değere ne kadar yakınsa, bireyler arası farklılık hakkında o kadar iyi bilgi verir.

Madde güvenirlik katsayısı; testteki her bir maddenin güvenirliği hakkında bilgi veren katsayıdır. Madde ayırt edicilik indeksiyle, madde standart sapmasının çarpılmasıyla elde edilir.

$$R_j = r_{jx}.ss_{madde} \quad (21)$$

R_j = Madde güvenirlik katsayısı;

r_{jx} = Madde ayırt edicilik gücü

ss_{madde} = Madde standart sapması

şeklinde hesaplanabilir.

Eşitlikten de anlaşılacağı üzere, madde ayırt edicilik ve madde varyansına bağlı olan güvenirlik katsayısı, bu iki değer ne kadar yüksekse o kadar iyi bilgi verir ya da o madde o kadar güvenilirdir denilebilir.

Tüm bunların dışında aynı zamanda, 15 kişilik bir öğretmen grubuna, uygulamanın kullanılabilirliğine dair veri elde edebilmek adına uygulamanın demosu kullandırılmış ve bununla birlikte de bir kullanılabilirlik anketi uygulanmıştır. Geliştirilen herhangi bir yazılım aracının ya da uygulamanın kullanılabilirliğini anlayabilmek, bu araç hakkında fikir sahibi olmak adına oldukça büyük önem arz etmektedir. Bu bağlamda kullanılabilirlik kavramından ve bu kavramın öneminden bahsetmek gerekmektedir.

Kullanılabilirlik, kişilerin çevresi ile olan etkileşimini, özelliklerini, sınırlılıklarını inceleyen ve elde edilen verilere bakılarak, sistem tasarımına yönelik kullanımı ifade eden bir kavramdır (Alexander, 2003). Coşkunserçe ve Dursun (2008) kullanılabilirlik için " bir ürünü kullanan insanların işlerini hızlı ve kolay bir biçimde yerine getirmesi" olarak ifade etmişlerdir. Kullanılabilirlik Profesyonelleri Derneği (Usability Professionals' Association, UPA) ise kullanılabilirlik için, "yazılım, donanım ya da herhangi bir ürünün, o ürünü kullanan kişiler için uygunluğu ve kolay kullanımını sağlayan bir ölçüt" olarak tanımlamaktadır (UPA, 2008). Jacob Nielsen

(2003) bu tanımlardan farklı olarak kullanılabilirliği "basit bir kullanıcı arayüzünün nasıl oluşturulacağını değerlendiren bir kalite niteliği" olarak tanımlamıştır.

Bu bağlamda kullanılabilirlik, sadece genel olarak ürün ya da insan-bilgisayar iletişimi açısından değil, herhangi bir hizmetin tasarlanmasında da önemli bir durum olarak karşımıza çıkmaktadır. Kullanılabilirliğe dair yapılan tanımları incelediğimizde de, kullanıcı merkezli bir kalite ölçütü olması ve hizmetlerin basit ve çabukluk gibi özelliklerini gösteren bir kavram olduğundan bahsedilebilir. Ürünlerin ve hizmetlerin tasarımında kullanılabilirlik kavramı ne kadar göz önüne alınırsa, kullanıcıların gereksinimlerinin en iyi şekilde karşılanmasına yardımcı olur ve aynı zamanda benzer hizmeti sunan diğer kurum ya da ürünlere karşı üstünlük sağlar.

Uluslararası Standartlar Organizasyonu (International Organization for Standardization, ISO) tarafından kullanılabilirlikle ilgili olarak ISO 9241 standardının oluşturulduğu görülmektedir (ISO, 2008). ISO 9241 standardından yola çıkılarak kullanılabilirlik ile ilgili farklı araştırmalar yapılmıştır. Bu araştırmalara göre, kullanılabilirliği belirleyen beş ölçüt bulunmaktadır ve alan yazında bu ölçütler "5E" olarak nitelendirilmektedir. Bu ölçütlerden birincisi, etkililiktir (effectiveness). Etkililik ölçütü, uygulamanın ya da çevrimiçi bir ortamın kullanıcılara sunması gereken bilgiyi tam olarak verip vermediğini ölçer. İkinci ve diğer bir ölçüt ise, etkinlik (efficient) ölçütüdür. Kullanıcıların bir çevrimiçi ortamda bir görevi tamamlarken harcadıkları bilişsel çabayı ele alır. Üçüncü olarak karşımıza caziplik (engaging) kavramı çıkmaktadır. Bu ölçüt de, kullanıcıların memnuniyetlerini ve tatmin olma durumlarını ele alır. Literatürde bahsedilen bir diğer ölçüt de hatalara karşı olumlu yaklaşımdır (error tolerant). Bu ölçüt, kullanıcıların hata yapmalarını engelleyecek bir arayüzün olması ve kullanıcılara yardımcı unsurların çevrimiçi ortamda yer almasını ele alır. Sonuncusu ise, kullanıcının uygulamayı ya da çevrimiçi ortamı kullanırken yapacağı işlemleri nasıl gerçekleştireceğini kolay bir şekilde tahmin etmesini ve uygulamadaki bölümlerin bir bütün olarak tutarlı olmasını ele alan kolay öğrenilebilirlik (easy to learn) ölçütüdür (Alexander, 2003).

Bu ölçütlerden yola çıkılarak çevrimiçi ortamların ya da uygulamaların kullanılabilirliklerinin belirlenmesine yönelik yapılan çalışmalara kullanılabilirlik testi adı verilmektedir. Bu konu ile ilgili yapılan tanımlar incelendiğinde, kullanılabilirlik testlerinin, tasarımların ya da arayüzlerin kullanıcı ile karşılaşmadan önce yapılan ve

tasarımın kullanılabilirliğini ölçen testler olduğu söylenebilir (Dalcı, 2008). Kullanılabilirlik testleri bir ürünün, sistemin ya da web sayfasının daha önceden belirlenen amaçlar ve ölçütler doğrultusunda ne derece etkin, verimli ve tatmin edici bir kullanım sunduğunu araştıran testlerdir (ÜTEST, 2008).

Bir ürünün kullanılabilirliğinin ölçülebilmesi için bazı unsurların göz önüne alınması ve bu unsurların araştırma öncesinde belirlenmesi gerekir. Bir ürünün kullanılabilirlik çalışmasını yaparken öncelikle;

• Kullanılabilirlik çalışması yapılacak konu

• Kullanılabilirlik çalışması uygulanacak arayüz ve bu çalışma sırasında kullanılacak araçlar

• Hedef kullanıcı grubu

• Verilerin hangi tekniklerle toplanacağının belirlenmesi gerekir (Alaçam ve Çağıltay, 2007).

Bu ölçümler yapılırken farklı yöntemler kullanılmakla birlikte, genelde ISO 9241 standardında yer alan etkinlik, etkililik ve tatmin edicilik fonksiyonları hakkında veriler elde edilmeye çalışılır. Bu nedenle kullanılabilirlik testlerinde hazırlanan soruların bu üç unsuru belirlemeye yönelik ortak özellikler taşıdığı söylenebilir. Kullanılabilirlik testlerinde yer alan etkililik unsuru, sistemin bu amaca yönelik kullanımının nicel olarak ölçülmesini ifade etmektedir (Bağış, 2003). Etkililik ile ilgili sorular ise, genelde kullanıcıların yaptıkları hata sayısı, görevleri başarıyla tamamlayan kullanıcıların yüzdesi gibi sayısal ve istatistiksel verilerin elde edilmesine yöneliktir. Etkinlik unsurunun saptanmaya çalışılması yoluyla da uygulamanın ya da ürünün oluşturulma amacına ne denli ulaşıldığının belirlenmesi ölçülmek istenmektedir. Etkililik için verilen bir görevi gerçekleştirmek için harcanan süre ve çaba, yardım kullanmada harcanan zaman gibi veriler elde edilmeye çalışılmaktadır. Son olarak ise bireylerin web uygulamasını ya da sitesini kullanımları doğrultusunda şikâyet sıklıkları, olumlu ya da olumsuz görüşleri ve bu görüşlerin oranı gibi hem nitel hem de nicel verilerden yararlanılarak kullanılabilirliğin memnun olma unsuru hakkında bilgi elde edilmektedir (Bağış, 2003). Buraya kadar bahsi geçen ve bir ürünü etkileyen unsurlar ölçülürken, kullanılabilirlik testlerinde dikkat edilmesi gereken iki kavram vardır. Bu iki önemli etken “güvenirlik” ve “geçerlik” kavramlarıdır (Elaina,

2002). Bu kavramlardan güvenirlik, Elaina'nin (2002); Hernon'un 1994 yılında yaptığı bir çalışmaya göre ürünün hedef kullanıcı grubu arasından seçilen farklı örneklemlerden benzer sonuçların elde edilmesidir. Bu bağlamda kullanılabilirlik testleri gerçekleştirilirken seçilen örneklemin güvenilir bir özellik taşıması gerekmektedir. Geçerlikte ise, hem araştırmacıların ölçmek istedikleri unsurların tam olarak ölçülüp ölçülmediği ifade edilmeye çalışılırken, hem de kullanılabilirlik testi sonrasında elde edilen bulguların genellenebilme durumu ifade edilmektedir. Bu bağlamda kullanılabilirlik testi uygulanırken kullanılan yöntem ve tekniklerin bu ölçütleri sağlayacak şekilde yapılandırılması gerekir. Bu durum, araştırmanın tutarlılığı açısından büyük bir öneme sahiptir. Kullanılabilirlik testinin uygulanması gereken çalışmalar için beş ile on beş kişi arasında değişen gruplar yeterli olmakta, az sayıda kişi üzerinde sınama yapılmasına rağmen yüksek tutarlılık sağlanabilmektedir (Nielsen, 2000). Bu araştırmalarda yer alacak kullanıcıları belirlerken öncelikle web sitesi ile nelerin hedeflendiği ve kimlere yönelik olarak bu tasarımın yapıldığının belirlenmesi gerekir (Lazar, 2001).

DÖRDÜNCÜ BÖLÜM

4. ARACIN GELİŞTİRİLMESİ

Bu bölüm, uygulamanın tasarım aşaması ve uygulamanın geliştirilmesi olmak üzere 2 bölümde incelenmiştir.

4.1. Uygulamanın Tasarım Aşaması

Uygulamanın fikir olarak ortaya çıkış noktası, araştırmacının kişisel deneyimlerine dayanmaktadır. Öğretmenlerin sınıf için uygulamaları dikkat alınarak yapılan gözlemler sonucunda, alan yazında da belirtildiği gibi öğretmenlerin madde ve test analizi yaparken yaşadığı zorluklar gözlemlenmiş ve bu gözlemler çerçevesinde böyle bir uygulamaya ihtiyaç duyulduğu düşünülmüştür.

İlk olarak detaylı bir alan yazın araştırması yapılarak, fikir aşamasında olan uygulamanın henüz tam anlamıyla bir muadilinin olmadığı görülmüştür. Bu noktada alan yazında yapılan okumalar doğrultusunda da, istatistiksel hesaplamaların pratikleştirebileceği ve hem ölçme işlemlerinin kalitesinin artmasına dolaylı olarak da eğitimin kalitesinin artışına katkı sağlanabileceği düşünülmüştür.

Alanyazında araştırması yapıldıktan ve alanyazında okumalar yapıldıktan sonra, bir uzman görüşü alınmıştır. Uzman görüşü alındıktan sonra madde ve test analizi yapılacak bu uygulamada mutlaka bulunması gereken istatistiksel hesaplar aşağıdaki gibi belirlenmiştir:

- Aritmetik Ortalama
- Mod
- Medyan
- Ranj
- Standart Sapma
- Varyans
- Test Ortalama Güçlüğü
- Test Güvenirlik Katsayısı
- Testin Standart Hatası

- Bağıl Değişkenlik Katsayısı
- Çarpıklık Katsayısı
- Basıklık Katsayısı
- Z Puanları
- T Puanları
- Madde Güçlük İndeksi
- Madde Ayırt Edicilik İndeksi
- Madde Varyansı
- Madde Standart Sapması
- Madde Güvenirlik Katsayısı

Bununla birlikte, uygulamada verilecek raporun tasarımı, içeriği, bilgiyi sunuş şekli konusunda da süreç boyunca uzman görüşü alınarak yapılandırılmıştır. Uzman görüşüyle birlikte, programın hali hazırda olan muadillerinin de verdiği rapor çıktıları incelenmiştir. Bu raporlardaki, eksik yanlar da göz önünde bulundurularak raporun yapısı belirlenmiştir.

Aynı zamanda, öğretmenin sınıflarının gelişimini takip edebilmesi için, elde ettiği verilerin geriye dönük olarak saklanmasının önemli olduğu ortak görüşüne varılmıştır. Bu bağlamda, her kullanıcının kendi mail adresiyle giriş yapabileceği bir sistemin geliştirilmesi planlanmış, kendi profilinde sınavlarını geriye dönük olarak saklayabilmesi ve istenildiği zaman bu bilgilere ulaşılabilmesi mümkün olacak şekilde tasarlanmıştır.

Tüm bunların dışında, uygulamanın nasıl kullanılacağını detaylı bir şekilde anlatan bir kullanıcı rehberinin de oluşturulması gerektiği düşünülmüştür. Kullanıcı dostu olması planlanarak geliştirilen uygulamanın, aynı açıklıkta kullanıcı dostu bir rehberini de bulunmaktadır. Bununla birlikte, rapor çıktısındaki analizlerin nasıl yapıldığını ve daha derin olarak ne anlama geldiğini merak eden kullanıcılar için de ayrı bir bölüm oluşturulması planlanmaktadır.

4.2. Uygulamanın Geliştirilmesi

Uygulamanın geliştirilmesinde HTML, CSS, Javascript, Bootstrap, Firebase gibi teknolojilerden yararlanılmıştır. Uygulamanın geliştirilme aşamasına geçmeden önce bu teknolojilerden bahsetmek gereklidir.

HTML, İsviçre'deki CERN araştırma enstitüsünde bir fizikçi olan Tim Berners-Lee tarafından yaratılmıştır. Hyper Text Markup Language kelimelerinin kısaltması olan Html, kullanıcılara web sayfaları ve uygulamalar için paragraflar, başlıklar ve yapılar oluşturma seçeneği sunar. Zannedilenin aksine bir programlama dili olmayan Html, dinamik bir işleve sahip değildir. Aslında, bir nevi Microsoft Word gibi belgeleri biçimlendirmeyi sağlar. Görsel, yazı ve video gibi içeriklerin web sayfasında konumlandırılmasını ve oluşturulan bu içeriklerin de doğru şekilde internet sitelerinde görüntülenmesini sağlar. Arama motorlarına web siteleri hakkında haber verir.

Css (Cascading Style Sheets) ise, web sayfasında görsel olarak bulunan her yapıyı düzenlemeyi sağlayan bir markup dilidir. Sayfaya ait tüm ögelerin (buton, arka plan, form gibi) renk, boyut, font konum gibi özelliklerinin düzenlenmesini sağlar.

JavaScript, yaygın olarak web tarayıcılarında kullanılmakta olan bir programlama dilidir. JavaScript sayesinde tarayıcının kullanıcıyla etkileşimde bulunması, tarayıcının kontrol edilmesi, asenkron bir şekilde sunucu ile iletişime geçilmesi ve web sayfası içeriğinin değiştirilmesi gibi işlevler sağlanır. JavaScript, Node.js gibi platformlar sayesinde sunucu tarafında da yaygın olarak kullanılmaktadır.

İlk olarak bir Netscape çalışanı olan Brendan Eich tarafından geliştirilen ve 1997 yılından beri bir ECMA standardı olan JavaScript, günümüzde Mozilla Vakfı öncülüğünde özgür yazılım topluluğu tarafından geliştirilmekte ve bakımı yapılmaktadır. Web sayfalarının görünümünün mükemmelleştirilmesi, işlevselliğinin artırılması ve dinamik web sayfaları tasarlanması amacıyla JavaScript, HTML kodu içinde gömülü olarak kullanılabilir. JavaScript, bir programlama dili disiplini ve özelliklerine sahiptir.

Ücretsiz bir CSS framework'u olan *Bootstrap*, açık kaynak kodlu bir tasarım aracıdır. Web sayfanızın, telefon, tablet ve masaüstü bilgisayarlar için farklı ve

cihazınızın büyüklüğüyle orantılı şekilde görünmesini sağlayan tema ve tasarımların kolaylıkla yapılmasını sağlar. Bir site için gerekli olan bütün unsurları içerisinde barındıran Bootstrap ile tasarım yaparken bu hazır unsurları kullanarak tüm cihazlara uygun tasarımlar yapmak mümkündür. Bu da demek oluyor ki, her şeyi hazır olan kodlarla yeni bir tasarım yaratmak oldukça kolay ve pratiktir. Stiller, imajlar ve JavaScript'ler daha öncesinden Bootstrap'in içine monte edilmiştir. Yapmanız gereken tek şey bunları çağırmanızdır. Bootstrap'in son sürümü olan Bootstrap 4, Chrome, Safari, Firefox, Internet Explorer, Opera gibi neredeyse her internet tarayıcısının son sürümlerinden sorunsuz şekilde çalışmaktadır.

Google Firebase; web ve mobil uygulamalarının sunucu tarafıyla geliştiricinin uğraşmasına gerek kalmadan kullanıcı giriş yetkilendirmeli ve verilerini gerçek zamanlı ve senkron bir şekilde tutulmasını sağlayan bir platformdur. Günümüzdeki projeler tüm markete hitap etmesi açısından iOS , Android ve web platformlarında geliştirilir fakat her platformun kendine ait yazılım dili ve bağlantı şekilleri vardır. Sunucu tarafı adı verilen arka plandaki verilerin tutulması ve gerektiği zaman kullanılması her platformun ortak sorunudur ve Google Firebase bu konuda geliştirilmiş ortak bir çözümdür.

Google Firebase'in başlıca özellikleri aşağıdaki gibidir:

- Gerçek zamanlı veritabanı
- Kullanıcı giriş yetkilendirmesi
- Depolama
- Makine Öğrenmesi Kitleri
- Performans ve hata test ortamları
- Platformlar arası ortak uygulama analizi
- Ortak fonksiyon
- Toplu bildirim
- Reklam araçları

Bu özelliklerin birçoğu her platformda tek tek uğraşılması gereken işler olması gerekirken geliştiriciler Google Firebase ile tek seferde ve kolay bir çözüm olanağı bulabilir. Kullanıcı girişlerinin olduğu ve belirli verilerin saklandığı birden fazla

platform için (iOS, Android vb.) geliştirilecek bir proje geliştirilmekteyse Google Firebase bu konuda oldukça fayda sağlayabilir.

Bu teknolojiler kullanılarak yazılan bu çevrimiçi web ortamının nasıl bir alt yapı ve mantıkla çalıştığını daha iyi anlayabilmek adına, kodun belli parçalarının incelenmesi anlamlı olabilir. Aşağıda kodun temel fonksiyonlarının nasıl çalıştığına dair açıklamalar bulunmaktadır.

İhtiyaç duyulan şey, bir sınıftaki öğrencilerin cevaplarının belirli bir cevap anahtarı ile işleme sokularak doğru cevapların ve bazı matematiksel ifadelerin hafızada saklanması olduğu için; sınav ve öğrenciler bir obje olarak modellenirken, matematiksel ifadeler ise sınav objesinin taşıdığı özellikler('attribute') olarak kodlanmıştır. Ayrıca öğrenci objelerini bir arada tutabilmek, yani bir sınıf modellemesi yapabilmek adına; oluşturulan her öğrenci bir diziye eleman olarak eklenmiştir. Bu dizi(ogrenciListesi), uygulama sırasında bir sınıf özelliği gösterecek ve öğrencilere erişimi sağlamaya yarayacaktır. Ayrıca öğrencilerin ve sınavın, programlama dilinde bir 'tür'(class) olarak oluşturulması, sonradan geliştirme yapılmasına açık bir yapıda olmasını sağlamaktadır. Programın geliştirilebilir olması, ilerde yapılabilecek ekstra çalışmaları da kolaylaştırması açısından; fonksiyonlarını yerine getiriyor olması kadar önemlidir. Öğrenci türü tanımı kod içerisinde 'OgrenciClass' ve sinav türü de 'SinavClass' olarak bulunmaktadır.

Öğrenci türünde bir obje;

- Öğrenci ad-soyadı (ogrenci_adi)
- Öğrencinin cevapları (ogrenci_cevap)

özelliklerini göstermektedir.

Sınav türünde bir obje ise özellik olarak şunları içermektedir:

- Sınav adı (sinavAdi)
- Sınavın yapıldığı şube(sinavSube)
- Sınavın yapıldığı okul (sinavOkul)
- Sınav tarihi (sinavTarihi)
- Sınav kimlik bilgisi(sinavId)
- Medyan (median)
- Ranj (ranj)
- Sınıf ortalaması (sinifOrtalama)
- Mod (max_tekrar)

- Alınan en yüksek not (max_puan)
- Alınan en düşük not(min_puan)
- Varyans (varyans)
- Standart sapma (standart_sapma)
- Bağıl değişkenlik (bagil_degiskenlik)
- Çarpıklık katsayısı (pearson_carpiklik_katsayisi)
- Testin güçlük katsayısı (test_gucluk)
- Basıklık katsayısı (basiklik_katsayisi)
- Standart hata (standart_hata)
- Kr 20 – güvenirlik katsayısı (kr_20)
- Maddelerin ayırt edicilik indeksi (madde_ayirt_edicilik)
- Maddelerin güçlük indeksi (maddelerin_gucluk_katsayisi)
- Maddelerin varyansı (maddelerin_varyansi)
- Maddelerin standart sapması (maddelerin_standart_sapmasi)
- Z puanları (z_puanlari)
- T puanları (t_puanlari)
- Madde seçenek analizi (maddelerin_secenek_analizi)

Bu özellikler kod içerisinde parantez içinde verildiği şekliyle bulunmaktadır. İstenildiğinde çok rahat bir şekilde yeni özellikler sisteme eklenebilip, mevcut olanlarda düzenlemeye gidilebilir olarak bulunmaktadır. Sistem rapor sayfasına bu özellikleri barındıran objeyi(mevcutSinav objesi) tek bir paket olarak aktarmakta, ayrıca Google Firebase veritabanına da bu objeyi göndermektedir.

Alınan verilerin virgülle ayrılmış değerler yani CSV(comma seperated values) formatında olması, kullanıcıya aynı veri listesinde tekrar kullanabileceği bir sınıf listesi oluşturmasını sağlarken sistem içinde alınan verinin özel belirlenmiş bir formatta olmasını sağlamaktadır. Kullanıcı verileri tek tek de girebilirken, bir sınıf için her seferinde aynı CSV dosyasında sadece cevapları değiştirerek sisteme girdi olarak sokabilir ve çok daha rahat bir kullanım sağlayabilir. Sistemde gelen dosyanın da her seferinde aynı olmasının, verilerin okunması açısından önemi çok kritiktir. Örnek vermek gerekecek olursa; A öğretmeni B sınıfı için yaptığı ilk matematik sınavını bir CSV dosyası olarak Microsoft Excel veya Microsoft Not Defteri uygulamasında

hazırlayıp kaydettikten sonra sisteme girdi olarak vermektedir. Daha sonra yapacağı ikinci bir sınav için (hatta başka branşlarda da aynı sınıf olduğu sürece) tekrar bu CSV dosyasında öğrencilerin cevaplarını ve sisteme gireceği cevap anahtarını değiştirerek aynı dosyayı ve öğrencileri tekrar tekrar giriş yapmadan kullanabilir. Aynı sınıf için başka bir dersin öğretmeni de bu dosyayı kullanabilir. CSV dosyasında dikkat edilmesi gereken kurallar kullanma kılavuzunda da belirtildiği gibi; ilk değer alanında öğrencinin adı soyadı birlikte yer alacak şekilde girilirken, kalan bütün değer sütunlarına tek bir cevap gelecek şekilde giriş yapılmalıdır. Aşağıda sistemin kabul edeceği bir CSV dosyasından birkaç satırlık bir örnek verilmiştir. Programın arka planında çalışan kodda yer alan her "not" ibaresi puan anlamında kullanılmıştır.

	A	B	C	D	E	F	G	H	I	J	K	L
1	Sabahattin Ali	C	A	D	D	A	A	B	C	D	C	A
2	Ahmet Ümit	C	D	D	D	A	A	B	C	B	C	A
3	Yaşar Kemal	C	A	D	D	A	A	B	C	B	C	A
4	Cahit Arf	C	A	D	D	A	A	B	C	B	C	A
5	Sait Faik Abasıyanık	C	A	D	D	A	A	B	C	B	C	A

Şekil 1. Örnek CSV dosyası

'Sınav Verilerini Hesapla' butonu ile çalışan verileriHesapla fonksiyonun çalışma şekli aşağıda belirtilmiştir.

```
" sorularin_dogru_yapilma_sayisi = new Array(cevapAnahtariDizisi.length).fill(0);
  ogrencilerin_dogru_sayisi = new Array(ogrenciListesi.length).fill(0);
  mevcutSinav = new SinavClass();

  mevcutSinav.setSinavAdi(_sinavAdi.value);
  mevcutSinav.setSinavSube(_sinifSubesi.value);
  mevcutSinav.setSinavTarihi(_sinavTarihi.value);
  mevcutSinav.setSinavId();
  mevcutSinav.setcevapAnahtariDizisi(cevapAnahtariDizisi); "
```

İlk olarak soruların doğru yapılma sayılarını ve her bir öğrencinin kaç doğru yaptığını kaydetmemizi yarayacak olan diziler, soru sayısı ve öğrenci listesinin uzunluğu kadar bir uzunlukta olacak şekilde oluşturulur. Başlangıçta dizilerin bütün elemanları sıfır olacak şekilde ayarlanacağı için fill(0) komutu kullanılmıştır.

Sınav türünün(SinavClass) özelliklerini edinip kullanabilmek ve kaydedebilmek adına bu türde bir obje türetmek için ‘mevcutSinav’ objesi de yeni bir obje olarak oluşturulmaktadır.

Daha sonra mevcutSinav objesine, site üzerinde giriş yapılan sınav adı, sınav şubesi ve sınav tarihi değerleri aktarılmaktadır. Burada kilit nokta, sınavın kimlik numarası niteliğindeki ‘sinavId’ özelliğidir. Bu özellik girilen sınavın adını, tarihini ve şubesini birleştirerek bir kimlik kodu ortaya çıkartmakta, bu kod ise daha sonradan veri tabanından bu sınavı belirleyip bilgileri tekrar çekmek için kullanılmaktadır. Yani 3A sınıfında 12 Ocak 2019 tarihinde yapılan bir Matematik sınavının kimlik kodu “3A-Matematik-12-01-2019” olarak belirlenecektir. Bu kimlik kodu bir nevi sınava eşsizlik sağlamaktadır. Sonraki adımda girilen cevap anahtarı, mevcutSinav objesinin ‘cevapAnahtariDizisi’ özelliğine girilmektedir.

```
“for (var i = 0; i < (ogrenciListesi.length); i++) {
    for (var j = 0; j < (cevapAnahtariDizisi.length); j++) {
      if ((ogrenciListesi[i].getAnswers[j]) == cevapAnahtariDizisi[j]) {
        sorularin_dogru_yapilma_sayisi[j] += 1;
        ogrencilerin_dogru_sayisi[i] += 1;
        ogrenciListesi[i].ogrencininDogruları[j] = 1;
        ogrenciListesi[i].ogrencininDogruSayisi += 1;
      } else {
        ogrenciListesi[i].ogrencininDogruları[j] = 0;
      }
    }
  } “
```

Bu döngüler içerisinde öğrenci cevaplarının cevap anahtarı ile kıyaslaması yapılmakta ve kalan tüm algoritmaya şekil verecek olan diziler oluşturulmaktadır. İlk for döngüsü içerisinde, ‘ogrenciListesi’ dizisinde bulunan her bir öğrenci tutularak, ikinci for döngüsünde ise bu tutulan öğrencinin cevap anahtarına ‘.getAnswers’ metodu ile erişilmektedir. ‘.getAnswers’ metodu, öğrenci objesi içerisinde tanımlı olan bir fonksiyon olup, öğrencinin cevaplarını bir dizi olarak geri döndürmektedir. Yani “AABDBCDBADBA” şeklinde bir dizi geri döndürerek, bu dizi üzerinde

indeksleme(yani ‘j’) yaparak öğrencinin her bir cevabına sıra ile ulaşılmaktadır. Burada ‘j’ indeksi, öğrencinin ilk cevabından sonra cevabına kadar tüm karakterleri sırasıyla ve tek tek tutmaktadır. Aynı şekilde cevapAnahtariDizisi üzerinde de aynı indeksli cevaba erişmek için yine aynı numaralı ‘j’ göstergeci kullanılarak kıyaslama sağlanılmaktadır. Bir örnekleme yapmak açısından; ‘i’ indeksinde bulunan öğrencinin, ‘j’ indeksinde bulunan cevabı ile cevap anahtarının ‘j’ indeksinde bulunan karakter kıyaslanarak (eğer sorgusu ile eşitlik kontrolü şeklinde) işlem yapılmaktadır. Yani ‘i’ indeksi 3 olduğunda sinifListesi[3] numaralı öğrenci çağırılacak, ‘j’ indeksi ise 0,1,2,3,4… değerlerini alarak öğrencinin cevabına ve cevap anahtarındaki cevaba erişecektir. Eğer kıyaslanan iki karakter eşitse, cevap doğru olacağı için öğrencinin doğru sayısını tutan(ogrencininDogruSayisi) özellik(attribute), öğrencilerin doğru sayısını tutan dizinin(ogrencilerin_dogru_sayisi) ilgili indeksteki(i) elemanı ve hangi sorunun kaç defa doğru yapıldığını tutan (sorularin_dogru_yapilma_sayisi) dizinin ilgili indeksteki(j) elemanı birer artırılır. Ayrıca burada daha sonra ‘alt grup’ ve ‘üst grup’ kümelerinde yer alan öğrencilerin hangi soruyu doğru yaptığını kullanmak üzere mevcut öğrencinin hangi soruları doğru yaptığını ‘1’ ve ‘0’ olarak tutan bir dizi olan ‘ogrencininDogruları’ dizisindeki ilgili indeks(j) elemanı 1 yapılmaktadır. Eğer karakterler kıyaslandığında aynı çıkmıyorsa, bir başka deyişle cevap yanlışsa; mevcut öğrencinin doğrularını tutan ‘ogrencininDogruları’ elemanında ilgili indeks(j) elemanı 0 yapılmaktadır. Bu kısım bittiğinde elimizde; hangi öğrenci hangi soruyu doğru yapmış, kaç doğru yapmış ve hangi soru kaç defa doğru yapılmış, hangi öğrenciler tarafından doğru yapılmış gibi sonuçların tamamı geçmektedir.

```
“ maddelerin_secenek_analizi = new Array();
   for (var i = 0; i < cevapAnahtariDizisi.length; i++)
  {
    maddelerin_secenek_analizi.push(new Array(6).fill(0));
  } ”
```

Bu kısımda ise, hangi seçenek, doğru yanlış bakılmaksızın öğrenciler tarafından kaç defa işaretlenmiş analizini yapmak adına ‘maddelerin_secenek_analizi’ oluşturulmaktadır. A,B,C,D,E ve Boş seçeneklerini tutmak için 6 elemanlı bir dizi, cevap anahtarı uzunluğunca eklenmektedir.

```
“ for (var k = 0; k < ogrenciListesi.length; k++)
  {
    for (var m = 0; m < cevapAnahtariDizisi.length; m++)
    {
     switch ( (ogrenciListesi[k].getAnswers[m]) )
{
       case "A" : maddelerin_secenek_analizi[m][0] +=1; break;
       case "B" : maddelerin_secenek_analizi[m][1] +=1; break;
       case "C" : maddelerin_secenek_analizi[m][2] +=1; break;
       case "D" : maddelerin_secenek_analizi[m][3] +=1; break;
       case "E" : maddelerin_secenek_analizi[m][4] +=1; break;
       default: maddelerin_secenek_analizi[m][5] +=1;
        }
     }
  }
mevcutSinav.setMaddelerinSecenekAnalizi(maddelerin_secenek_analizi); ”
```

Bu döngülerle ‘ogrenciListesi’ içerisinde her ‘k’ indeksli öğrencinin ‘m’ indeksli cevabı için; maddelerin seçenek analizini tuttuğumuz dizide(maddelerin_secenek_analizi) ‘A’ şıkkı için m = 0 indeksli, ‘B’ şıkkı için m = 1 indeksli, ‘C’ şıkkı için m = 2 indeksli, ‘D’ şıkkı için m = 3 indeksli, ‘E’ şıkkı için m = 4 indeksli ve boş olan her şık için m = 6 indeksli elemanlar 1 artırılır. Yani öğrenci listesinde 4. sırada bulunan bir öğrenci için ‘k’ indeksi, dizilerde indeksleme 0’dan başladığı için 3 olup; bu öğrencinin 7. soruya verdiği cevap bulunurken m=6 indeksi kullanılıyor olur. Aynı şekilde ‘maddelerin_secenek_analizi’ dizisinde de m=6 indeksli elemana ulaşılarak 7. soruya verilen şık ne ise o seçeneğin karşılık geldiği indeks bir artırılır. Dizinin tüm değerleri yazılıp döngü tamamlandığında; ‘maddelerin_secenek_analizi’ dizisi; ‘mevcutSinav’ objesinin ilgili özelliğine atama yapılır.

```
“ for (var m = 0; m < ogrenciListesi.length - 1; m++) {
    for (var n = 0; n < ogrenciListesi.length - 1; n++) {
```

```
    if (ogrenciListesi[n].ogrencininDogruSayisi > ogrenciListesi[n + 1].ogrencininDogruSayisi) {
      var geciciDegisken = ogrenciListesi[n + 1];
      ogrenciListesi[n + 1] = ogrenciListesi[n];
      ogrenciListesi[n] = geciciDegisken;
    }
   }
  } ”
```

Bu kısımda artık kaç doğru yaptığı ve hangi soruları doğru yaptığı bilinen öğrencilerin olduğu ‘ogrenciListesi’ dizisinde, for döngüsü ile indeksleme yapılarak, dizi içerisindeki öğrencilerin doğru sayılarına göre sıralanmaları sağlanmıştır. Burada sıralama daha sonra alt grup ve üst gruptaki öğrencilere erişimi sağlayabilmek adına yapılmıştır.

```
“ siraliNotlar = ogrencilerin_dogru_sayisi.slice().sort(); ”
```

Öğrencilerin doğru sayılarının sıralandığı dizi(‘ogrencilerin_dogru_sayisi’) üzerinde Javascript’in küçükten büyüğe sıralayıp atama yapmayı sağlayan ‘.slice()’ ve ‘.sort()’ metodları kullanılarak yeni bir dizi oluşturulmuş ve ‘siraliNotlar’ dizisinde saklanmıştır.

```
“ max_puan = siraliNotlar[siraliNotlar.length - 1];
  min_puan = siraliNotlar[0]; ”
```

‘siraliNotlar’ dizisi küçükten büyüğe sıralı bir dizi olduğu için ilk elemanı en küçük puanı vermektedir. Diziler ‘0’ indeksinden başladığı için bu işlem ‘siraliNotlar[0]’ şeklinde yapılmaktadır. Sonrasında kullanılmak üzere bu değer ‘min_puan’ değişkenine atanmıştır. Tam tersi olarak en büyük puanı elde etmek için dizinin son terimine gidilmelidir. Bunu yapmak için ise diziler ‘0’ indeksinden başladığı için ‘eleman sayısı-1’ indeksine gidildiğinde son terime ulaşılabilir. Örneğin 8 elemanlı bir dizinin son terimine ulaşmak için o dizinin ‘7’ numaralı indeksine; dizi[7] şeklinde erişilebilir.

“ altSinir = Math.floor((siraliNotlar.length - 1) / 2);
ustSinir = Math.ceil((siraliNotlar.length - 1) / 2);
median = (siraliNotlar[altSinir] + siraliNotlar[ustSinir]) / 2;
mevcutSinav.setMedian(Number.parseFloat(parseFloat(median).toFixed(3))); ”

Medyan hesabında bilindiği üzere sıralanmış puanlar arasından tam ortadaki değer seçilmesi gerektiği için, ‘siraliNotlar’ dizisi kullanılmaktadır. Ancak bu bir takım problemleri beraberinde getirmektedir. Eğer dizi 3-5 gibi tek elemanlı olursa rahatlıkla bulunabilen ortanca eleman, çift sayılı dizilerde, buçuklu sayının altı ve üstünün ortalaması şeklinde hesaplanır. Verilen kod bloğunda da, ortadaki buçuklu eleman bir üste, bir alta yuvarlanılarak toplanmış ve daha sonra ikiye bölünerek ortalaması alınmıştır. Bu çift elemanlı dizi problemini çözdüğü gibi, tek elemanlı dizilerde de herhangi bir soruna sebep olmamaktadır. Son satırda, bulunan medyan değeri ‘mevcutSinav’ objesine özellik olarak eklenilir.

“ ranj = siraliNotlar[siraliNotlar.length - 1] - siraliNotlar[0]; //ranj hesabı
mevcutSinav.setRanj(Number.parseFloat(parseFloat(ranj).toFixed(3))); ”

Ranj hesabının gerçeklemesi, daha önceden sıralanarak oluşturulmuş ‘siraliNotlar’ dizisinin ilk ve son elemanları birbirinden çıkarılarak yapılır. Daha sonra yapılan bu çıkarma işleminin sonucu ‘mevcutSinav’ objesine özellik olarak eklenmiştir. Ekleme yapılırken kod boyu virgülden sonra 3 basamak almaya özen gösterildiği için Javascript’in ‘toFixed(3)’ metodu kullanılmıştır.

“ sinifToplami = 0;
for (var i = 0; i < siraliNotlar.length; i++) {
sinifToplami += siraliNotlar[i];
} ”

Sınıfın puan toplamını bulmak için, ‘siraliNotlar’ dizisinin tüm elemanları kümülatif olarak ‘sinifToplami’ değişkenine eklenmiştir. Böylelikle bulunan toplam not değeri ileriki aşamalarda kullanılmak üzere saklanmıştır.

“ sinifOrtalama = sinifToplami / ogrenciListesi.length;

mevcutSinav.setSinifOrtalama(Number.parseFloat(parseFloat(sinifOrtalama).toFixed(3))); ”

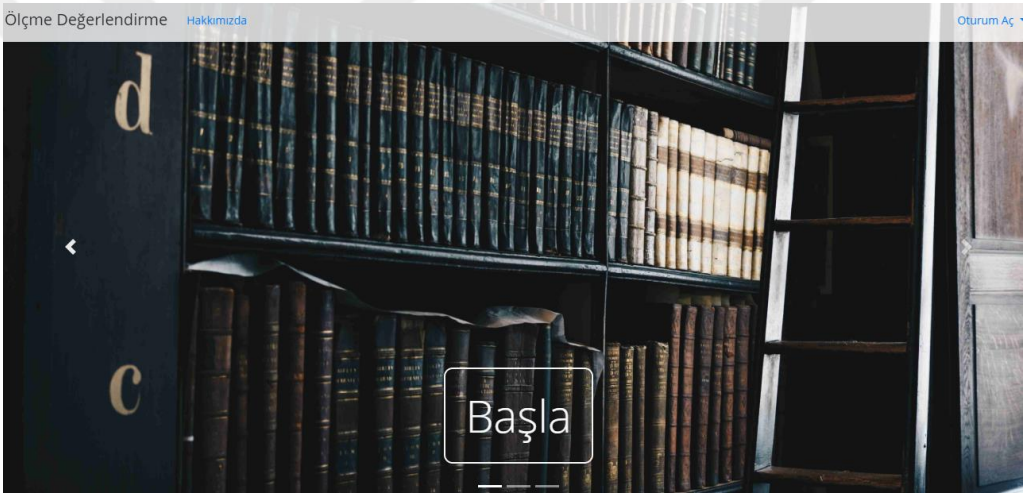
Bulunan sınıf toplamı, sınıf mevcuduna bölünerek ortalama bulunabilir. Bu sebeple ‘sinifToplami’ değerini, ‘ogrenciListesi’ dizisinin uzunluğuna bölünmüş ve bulunan değer ‘sinifOrtalama’ değerinde saklanılmıştır. ‘mevcutSinav’ objesine bu değer eklenilirken, önceden de yapıldığı gibi virgülden sonra 3 basamak almak için ilgili ‘toFixed()’ metodu kullanılmıştır.

```
“ var enCokSayilan = 0;
  for (var i = 0; i < siraliNotlar.length; ++i)
 {
  var geciciEnCok = 0;
  for (var j = 0; j < siraliNotlar.length; ++j)
   {
    if (siraliNotlar[j] == siraliNotlar[i]) ++geciciEnCok;
   }
  if (geciciEnCok > enCokSayilan)
  {
   enCokSayilan = geciciEnCok;
   max_tekrar = siraliNotlar[i];
  }
 }

mevcutSinav.setMax_tekrar(Number.parseFloat(parseFloat(max_tekrar).toFixed(3)));
 mevcutSinav.setMax_puan(Number.parseFloat(parseFloat(max_puan).toFixed(3)));
mevcutSinav.setMin_puan(Number.parseFloat(parseFloat(min_puan).toFixed(3))); ”
```

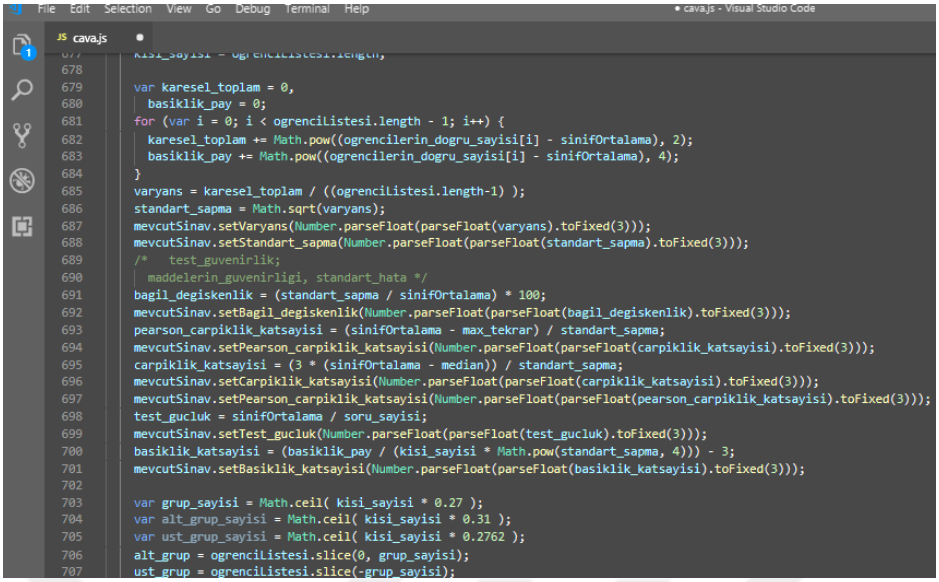
Araştırmanın bu bölümünde, buraya kadar bahsi geçen geliştirme sürecinin sonucunda araştırmacı tarafından geliştirilen çevrimiçi uygulamadan bahsedilmiştir. Araştırmanın amaçları doğrultusunda, uygulamanın yapılması için geliştirilen çevrim içi ortam ve simülasyon çalışmaları hakkında bilgi verilmiştir.

Bilgisayar ortamında, test ve madde istatistiklerinin uygulanabilmesi için araştırmacı tarafından bir çevrimiçi ortam oluşturulmuştur. Uygulamanın farklı platformlar için kurulum gerektirmemesi, cep telefonu, tablet ve bilgisayar gibi cihazlarda rahatça kullanılabilmesi, erişimin kolay olması gibi nedenlerden dolayı çevrimiçi ortam tercih edilmiştir. Araştırmacı tarafından www.testanalizi.net adresi alınmıştır.



ŞEKİL 2. Kullanıcı karşılama ekranı

Bu çevrim içi ortam, Javascript programlama dili ile yazılmış ve veri tabanı yönetimi ve sunucu işlemleri için Google Firebase kullanılmıştır.



ŞEKİL 3. Javascript kod ekranı

Aynı zamanda uygulamanın güvenirliğini test etmek amacıyla 9 farklı veri grubuyla çalışılmıştır. Bu 9 farklı veri grubu, 0.40 , 0.60 ve 0.90 test güçlüğüne sahip 30-100 ve 500 kişilik öğrenci gruplarından oluşmaktadır. Simülasyon yoluyla elde edilen veriler TAP (Test Analysis Program) yardımıyla oluşturulmuştur. Aynı zamanda uygulamanın analiz ederek verdiği değerlerin güvenirliğinin test edilmesinde de değerler, SPSS programının verdiği çıktılarla karşılaştırılmıştır.

Uygulamanın kullanılabilirliğinin test edilebilmesi için farklı branşlar ve yaş gruplarından 10 öğretmen uygulamada sunulan rehber yardımıyla, uygulamayı kullanmış ve Oğuzhan Erdinç tarafından Türkçe'ye uyarlanan, kullanılabilirlik anketini doldurmuşlardır. Bu anket yoluyla elde edilen verilere göre de uygulamanın kullanılabilirliği belirlenmiştir.

Geliştirilen uygulamaya ilişkin bazı ekran görüntüleri aşağıda verilmiştir. Her şeyden önce analiz yapabilmek, uygulamayı kullanabilmek ve çıktı olarak bir rapora ulaşabilmek için kullanıcıların sisteme üye olması gerekmektedir.



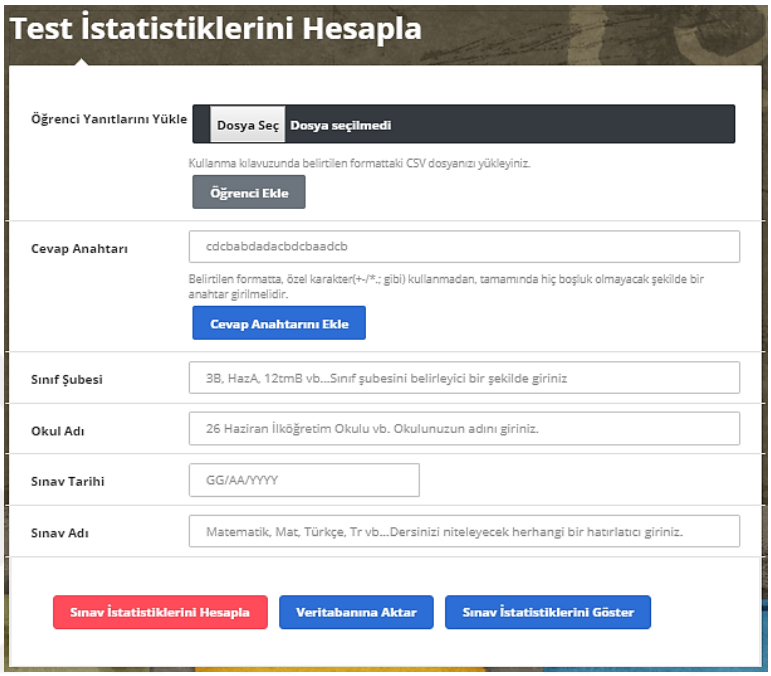
ŞEKİL 4. Hesap oluşturma ekranı

Üye olduktan sonra kullanıcıların e-posta adreslerine bir doğrulama maili gönderilip, doğrulanma işlemi sağlandıktan sonra sitede aynı zamanda oturum açmış olacaklardır. Kullanıcılar bir kez üyelik aldıktan sonra, bundan sonraki işlemlerinde üyelikleriyle sadece oturum açarak analizlerini yine kolaylıkla yapabileceklerdir.



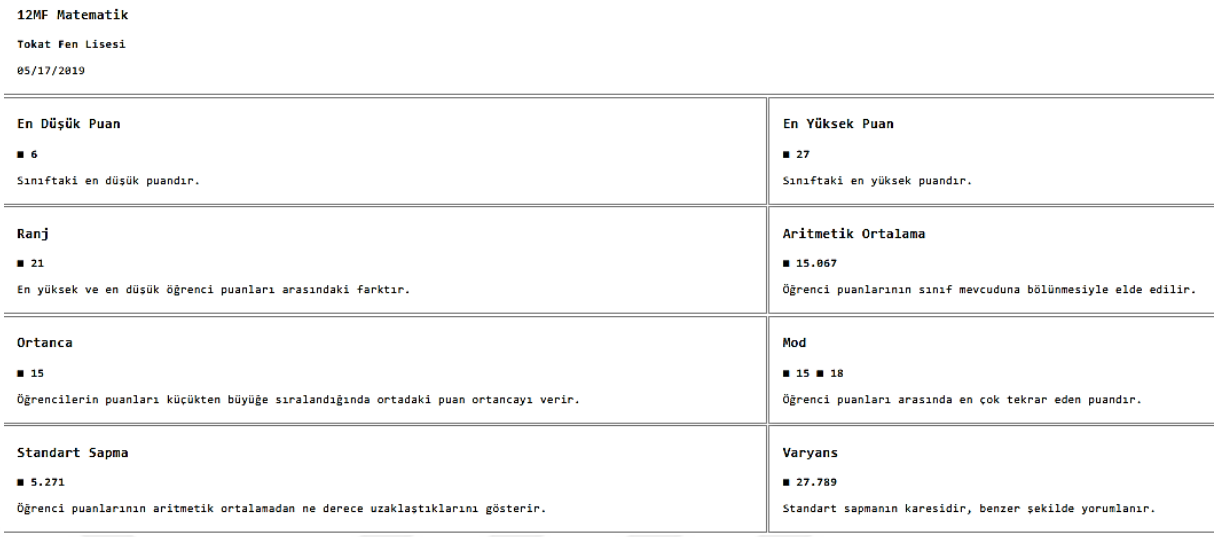
ŞEKİL 5. Kullanıcı oturum açma ekranı

Oturum açtıktan sonra, verilerini kolaylıkla uygulamaya yükleyebilen kullanıcılar, cevap anahtarı, okul adı, sınıf adı, ders adı ve tarih gibi bilgilerini de girerek, başka hiç bir işlem yapmadan analiz sonucu rapor olarak elde edebilmektedirler.



ŞEKİL 6. Test istatistiklerini elde edebilmek içiniçin gerekli olan verilerin giriş ekranı

Kullanıcıya verilen rapor, uygulamayı muadillerinden ayıran en önemli noktalardan biridir. Test ve madde istatistiklerini detaylı bir şekilde kullanıcıya sunan bu uygulama, raporda da yorumu kullanıcıya bırakmayıp net ifadelerle verilerin analizinin sonucu hakkında fikir vermektedir.



12MF Matematik

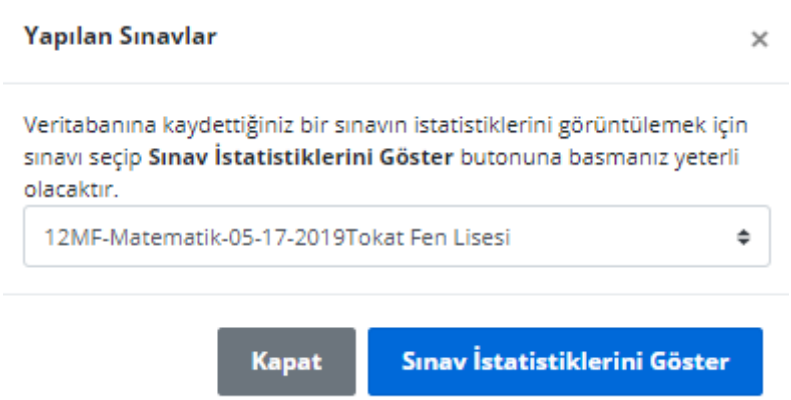
Tokat Fen Lisesi

05/17/2019

En Düşük Puan ■ 6 Sınıftaki en düşük puandır.	**En Yüksek Puan** ■ 27 Sınıftaki en yüksek puandır.
Ranj ■ 21 En yüksek ve en düşük öğrenci puanları arasındaki farktır.	**Aritmetik Ortalama** ■ 15.067 Öğrenci puanlarının sınıf mevcuduna bölünmesiyle elde edilir.
Ortanca ■ 15 Öğrencilerin puanları küçükten büyüğe sıralandığında ortadaki puan ortancayı verir.	**Mod** ■ 15 ■ 18 Öğrenci puanları arasında en çok tekrar eden puandır.
Standart Sapma ■ 5.271 Öğrenci puanlarının aritmetik ortalamadan ne derece uzaklaştıklarını gösterir.	**Varyans** ■ 27.789 Standart sapmanın karesidir, benzer şekilde yorumlanır.

ŞEKİL 7. Uygulamanın verdiği raporun bir kısmının görüntüsü

Aynı zamanda, sınav verilerini sisteme aktaran kullanıcı, daha sonra geriye dönüp ulaşmak istediğinde rahatlıkla sistemde kayıtlı olan raporlarını yeniden inceleyebilir.



ŞEKİL 8. Eski sınav istatistiklerine ulaşma ekranı

Bununla birlikte; kullanıcıya uygulamayı rahatça kullanabilmesi için bir rehber oluşturulmuştur. Böylelikle kullanıcının, uygulamayı kullanma noktasında rahat bir erişimi olması sağlanmıştır.

Aynı zamanda, test ve madde analizleri konusunda rapordakinden daha fazla bilgiye erişmek isteyen ve istatistiksel hesaplamaların nasıl yapıldığını merak eden

kullanıcılar için, ayrı bir bölüm hazırlanmıştır. Bu bölümde, bütün istatistiksel hesaplamaların formülleri ve daha geniş bir anlatımla ne anlama geldikleri anlatılmıştır.

BEŞİNCİ BÖLÜM

5. BULGULAR VE YORUMLAR

Bu bölümde, verilerin analizi sonucunda elde edilen bulgulara ve yorumlara yer verilmiştir. Elde edilen bulgular, öncelikle uygulamanın güvenirliği daha sonra, kullanılabilirliği üzerinedir.

5.1. Geliştirilen uygulamanın SPSS ve TAP paket programı ile karşılaştırılması sonucu elde edilen bulgular

Bu aşamada, daha önce de bahsedildiği üzere, uygulamanın güvenirliğine ait bulgular paylaşılmıştır. 30, 100, 500 kişilik öğrenci gruplarını temsil edecek şekilde TAP programı yoluyla simülatif bir şekilde veriler üretilmiştir. Veriler üretilirken her grup için, 0.40, 0.60 ve 0.90 ortalama güçlükler seçilmiştir. Bu gruplar üzerinden elde edilen sonuçlar, SPSS ve TAP paket programlarıyla elde edilen sonuçlar ile karşılaştırılmış ve tablolaştırılmıştır. Bu bağlamda her iki uygulamadan da elde edilen veriler test istatistikleri ve madde istatistikleri olmak üzere gruplanarak farklı tablolar üzerinde karşılaştırılmıştır. Aşağıda test ve madde istatistiklerine ait oluşturulmuş tablolar ve tablolara ait yorumlar bulunmaktadır. Tablolarda daha pratik bir adlandırma yapabilmek adına aşağıdaki gibi bir yol izlenmiştir.

TEST 1: 30 kişilik gruba ait 0.40 güçlüklü test ve madde analizinden elde edilen istatistikler;

TEST 2: 30 kişilik gruba ait 0.60 güçlüklü test ve madde analizinden elde edilen istatistikler;

TEST 3:30 kişilik gruba ait 0.90 güçlüklü test ve madde analizinden elde edilen istatistikler;

TEST 4:100 kişilik gruba ait 0.40 güçlüklü test ve madde analizinden elde edilen istatistikler;

TEST 5:100 kişilik gruba ait 0.60 güçlüklü test ve madde analizinden elde edilen istatistikler;

TEST 6:100 kişilik gruba ait 0.90 güçlüklü test ve madde analizinden elde edilen istatistikler;

TEST 7: 500 kişilik gruba ait 0.40 güçlüklü test ve madde analizinden elde edilen istatistikler;

Test 8: 500 kişilik gruba ait 0.60 güçlüklü test ve madde analizinden elde edilen istatistikler;

TEST 9:500 kişilik gruba ait 0.90 güçlüklü test ve madde analizinden elde edilen istatistikler olarak adlandırılmıştır.

Araştırma sonuçlarına ait sunulan bulgularda araştırmacı tarafından geliştirilen uygulama "GU" olarak kullanılmıştır.

Tablo 3. 30 Kişilik Öğrenci Grubuna Ait 1. Teste Dair İstatistikler

TEST 1 (N=30)	SPSS	G.U
En Düşük Puan	0.000	0.000
En Yüksek Puan	22.000	22.000
Mod	9.000	9.000
Medyan	9.000	9.000
Ranj	22.000	22.000
Aritmetik Ortalama	9.400	9.400
Standart Sapma	5.834	5.834
Varyans	34.041	34.041
Bağıl değişkenlik katsayısı	-	62.069
Çarpıklık katsayısı	0.749	0.749
Basıklık Katsayısı	-0.182	-0.182
Testin Ortalama Güçlüğü	0.313	0.313
KR-20 Güvenirlik Katsayısı	0.863	0.863
KR-21 Güvenirlik Katsayısı	0.838	0.838

Yukarıdaki tabloda da görüldüğü üzere, geliştirilen uygulamanın test ve madde analiz sonuçlarına göre elde edilen test istatistikleri, SPSS programı ile birebir aynıdır. SPSS uygulamasından farklı olarak, geliştirilen yazılımın bağıl değişkenlik katsayısını sunduğu da görülebilir.

Tablo 4. 30 Kişilik Öğrenci Grubuna Ait 2. Teste Dair İstatistikler

TEST 2 (N=30)	SPSS	G.U
En Düşük Puan	6.000	6.000
En Yüksek Puan	27.000	27.000
Mod	15.000-18.000	15.000-18.000
Medyan	15.000	15.000
Ranj	21.000	21.000
Aritmetik Ortalama	15.067	15.067
Standart Sapma	5.271	5.271
Varyans	27.789	27.789
Bağıl değişkenlik katsayısı	-	34.988
Çarpıklık katsayısı	0.147	0.147
Basıklık Katsayısı	-0.375	-0.375
Testin Ortalama Güçlüğü	0.502	0.502
KR-20 Güvenirlik Katsayısı	0.809	0.809
KR-21 Güvenirlik Katsayısı	0.755	0.755

Tablo 4 incelendiğinde, geliştirilen uygulamadan elde edilen test istatistikleri ile SPSS programından elde edilen test istatistiklerinin aynı olduğu görülmektedir.

Tablo 5. 30 Kişilik Öğrenci Grubuna Ait 3. Teste Dair İstatistikler

TEST 3 (N=30)	SPSS	G.U
En Düşük Puan	16.000	16.000
En Yüksek Puan	30.000	30.000
Mod	25.000	25.000
Medyan	25.000	25.000
Ranj	14.000	14.000
Aritmetik Ortalama	24.367	24.367
Standart Sapma	3.755	3.755
Varyans	14.102	14.102
Bağıl değişkenlik katsayısı	-	15.412
Çarpıklık katsayısı	-0.512	-0.512
Basıklık Katsayısı	-0.584	-0.584
Testin Ortalama Güçlüğü	0.812	0.812
KR-20 Güvenirlik Katsayısı	0.749	0.749
KR-21 Güvenirlik Katsayısı	0.699	0.699

Tablo 5 incelendiğinde, yine tüm istatistikler için geliştirilen uygulamadan elde edilen test istatistikleri ile SPSS programından elde edilen test istatistiklerinin aynı olduğu görülmektedir.

Tablo 6. 100 Kişilik Öğrenci Grubuna Ait 1. Teste Dair İstatistikler

TEST 4 (N=100)	SPSS	G.U
En Düşük Puan	1.000	1.000
En Yüksek Puan	25.000	25.000
Mod	25.000	25.000
Medyan	5.000-6.000	5.000-6.000
Ranj	24.000	24.000
Aritmetik Ortalama	9.860	9.860
Standart Sapma	5.356	5.356
Varyans	28.687	28.687
Bağıl değişkenlik katsayısı	-	54.321
Çarpıklık katsayısı	0.592	0.592
Basıklık Katsayısı	-0.009	-0.009
Testin Ortalama Güçlüğü	0.329	0.329
KR-20 Güvenirlik Katsayısı	0.815	0.815
KR-21 Güvenirlik Katsayısı	0.796	0.796

Tablo 6 incelendiğinde, geliştirilen uygulamadan elde edilen test istatistikleri ile SPSS programından elde edilen test istatistiklerinin aynı olduğu görülmektedir.

Tablo 7. 100 Kişilik Öğrenci Grubuna Ait 2. Teste Dair İstatistikler

TEST 5 (N=100)	SPSS	G.U
En Düşük Puan	2.000	2.000
En Yüksek Puan	28.000	28.000
Mod	15.000	15.000
Medyan	15.000	15.000
Ranj	24.000	24.000
Aritmetik Ortalama	15.400	15.400
Standart Sapma	5.963	5.963
Varyans	35.556	35.556
Bağıl değişkenlik katsayısı	-	38.720
Çarpıklık katsayısı	-0.056	-0.056
Basıklık Katsayısı	-0.707	-0.707
Testin Ortalama Güçlüğü	0.513	0.513
KR-20 Güvenirlik Katsayısı	0.835	0.835
KR-21 Güvenirlik Katsayısı	0.816	0.816

Tablo 7'de de görüldüğü gibi, geliştirilen uygulamadan elde edilen test istatistikleri ile SPSS programından elde edilen test istatistiklerinin aynı olduğu görülmektedir.

Tablo 8. 100 Kişilik Öğrenci Grubuna Ait 3. Teste Dair İstatistikler

TEST 6 (N=100)	SPSS	G.U
En Düşük Puan	15.000	15.000
En Yüksek Puan	30.000	30.000
Mod	28.000	28.000
Medyan	26.500	26.500
Ranj	15.000	15.000
Aritmetik Ortalama	25.520	25.520
Standart Sapma	3.356	3.356
Varyans	11.262	11.262
Bağıl değişkenlik katsayısı	-	13.150
Çarpıklık katsayısı	-0.978	-0.978
Basıklık Katsayısı	0.450	0.450
Testin Ortalama Güçlüğü	0.851	0.851
KR-20 Güvenirlik Katsayısı	0.701	0.701
KR-21 Güvenirlik Katsayısı	0.684	0.684

Tablo 8'de de görüleceği gibi, test istatistikleri hem geliştirilen uygulamada hem de SPSS programında aynı değerlere sahiptir.

Tablo 9. 500 Kişilik Öğrenci Grubuna Ait 1. Teste Dair İstatistikler

TEST 7 (N=500)	SPSS	G.U
En Düşük Puan	1.000	1.000
En Yüksek Puan	28.000	28.000
Mod	11.000	11.000
Medyan	13.000	13.000
Ranj	27.000	27.000
Aritmetik Ortalama	13.132	13.132
Standart Sapma	5.313	5.313
Varyans	28.227	28.227
Bağıl değişkenlik katsayısı	-	40.458
Çarpıklık katsayısı	0.186	0.186
Basıklık Katsayısı	-0.530	-0.530
Testin Ortalama Güçlüğü	0.438	0.438
KR-20 Güvenirlik Katsayısı	0.791	0.791
KR-21 Güvenirlik Katsayısı	0.764	0.764

Tablo 9 incelendiğinde, geliştirilen uygulamadan elde edilen test istatistikleri ile SPSS programından elde edilen test istatistiklerinin aynı olduğu görülmektedir.

Tablo 10. 500 Kişilik Öğrenci Grubuna Ait 2. Teste Dair İstatistikler

TEST 8 (N=500)	SPSS	G.U
En Düşük Puan	4.000	4.000
En Yüksek Puan	29.000	29.000
Mod	20.000	20.000
Medyan	19.500	19.500
Ranj	25.000	25.000
Aritmetik Ortalama	19.008	19.008
Standart Sapma	5.296	5.296
Varyans	28.044	28.044
Bağıl değişkenlik katsayısı	-	-
Çarpıklık katsayısı	-0.355	-0.355
Basıklık Katsayısı	-0.458	-0.458
Testin Ortalama Güçlüğü	0.634	0.634
KR-20 Güvenirlik Katsayısı	0.802	0.802
KR-21 Güvenirlik Katsayısı	0.778	0.778

Tablo 10 incelendiğinde, geliştirilen uygulamadan elde edilen test istatistikleri ile SPSS programından elde edilen test istatistiklerinin aynı olduğu görülmektedir.

Tablo 11. 500 Kişilik Öğrenci Grubuna Ait 3. Teste Ait İstatistikler

TEST 9 (N=500)	SPSS	G.U
En Düşük Puan	16.000	16.000
En Yüksek Puan	30.000	30.000
Mod	29.000	29.000
Medyan	28.000	28.000
Ranj	14.000	14.000
Aritmetik Ortalama	27.228	27.228
Standart Sapma	2.408	2.408
Varyans	5.800	5.800
Bağıl değişkenlik katsayısı	-	8.845
Çarpıklık katsayısı	-1.128	-1.128
Basıklık Katsayısı	1.324	1.324
Testin Ortalama Güçlüğü	0.908	0.908
KR-20 Güvenirlik Katsayısı	0.603	0.603
KR-21 Güvenirlik Katsayısı	0.586	0.586

Tablo 11 incelendiğinde, geliştirilen uygulamadan elde edilen test istatistikleri ile SPSS programından elde edilen test istatistiklerinin aynı olduğu görülmektedir.

Yukarıda 9 farklı veri setine ait istatistikler incelendiğinde, geliştirilen uygulamanın, SPSS paket programıyla çoğunlukla aynı sonuçlar verdiği görülmüştür. Ondalık kısımda oluşan farklılıklar ise, iki yazılımın da arka planda virgülden sonra kaç basamağı yuvarlayarak değerleri rapora yansıttığına bağlı olarak değişkenlik göstermektedir. Bu noktada, yazılımın çıktı olarak verdiği rapor değerlerinin, en bilinen muadiliyle aynı ya da daha yakın değerler vermesi, yazılımın güvenirliği açısından oldukça önemlidir. Aynı veri setleri kullanılarak, sadece test istatistikleri değil, aynı zamanda madde istatistikleri de hesaplanmıştır. Hesaplanan madde istatistikleri TAP programıyla karşılaştırılarak tablolaştırılmıştır. Bu noktada şunu belirtmekte fayda vardır. Araştırmacı tarafından geliştirilen uygulamada, madde ayırt ediciliği için üç farklı değer hesaplanmaktadır. Bunlar alt grup- üst gruba dayalı madde ayırt ediciliği, çift serili korelasyon ve nokta çift serili korelasyon değerleridir. Fakat, TAP programında nokta çift serili korelasyona dayalı bir istatistik raporu

sunulduğunda dolayı, bundan sonraki verilecek tabloların tümünde karşılaştırılan madde ayırt ediciliği değeri, nokta çift serili korelasyon değeridir.

Tablo 12. 30 Kişilik Öğrenci Grubuna Uygulanan 1. Sınavın Madde İstatistikleri

Test1	Madde Güçlük İndeksi		Madde Ayırt Edicilik İndeksi		Madde Güvenirliği		Madde Varyansı	
	TAP	G.U	TAP	G.U	TAP	G.U	TAP	G.U
M1	0.330	0.333	0.440	0.440	-	0.471	-	0.222
M2	0.070	0.067	0.240	0.230	-	0.249	-	0.062
M3	0.300	0.300	0.640	0.630	-	0.458	-	0.210
M4	0.100	0.100	0.270	0.260	-	0.300	-	0.09
M5	0.430	0.433	0.630	0.620	-	0.496	-	0.246
M6	0.200	0.200	0.420	0.410	-	0.400	-	0.160
M7	0.470	0.467	0.520	0.510	-	0.499	-	0.249
M8	0.130	0.133	0.400	0.390	-	0.341	-	0.116
M9	0.130	0.133	0.690	0.680	-	0.341	-	0.116
M10	0.530	0.533	0.580	0.570	-	0.499	-	0.249
M11	0.230	0.233	0.510	0.500	-	0.423	-	0.179
M12	0.170	0.167	0.510	0.510	-	0.373	-	0.139
M13	0.470	0.467	0.670	0.660	-	0.499	-	0.249
M14	0.470	0.467	0.470	0.460	-	0.499	-	0.249
M15	0.430	0.433	0.490	0.480	-	0.496	-	0.246
M16	0.200	0.200	0.370	0.370	-	0.400	-	0.160
M17	0.170	0.167	0.530	0.520	-	0.373	-	0.139
M18	0.670	0.667	0.410	0.400	-	0.471	-	0.222
M19	0.330	0.333	0.170	0.170	-	0.471	-	0.222
M20	0.070	0.067	0.030	0.030	-	0.249	-	0.062
M21	0.270	0.267	0.630	0.620	-	0.443	-	0.196
M22	0.630	0.633	0.080	0.080	-	0.482	-	0.232
M23	0.370	0.367	0.380	0.370	-	0.482	-	0.232
M24	0.300	0.300	0.470	0.470	-	0.458	-	0.210
M25	0.130	0.133	0.230	0.230	-	0.341	-	0.116
M26	0.470	0.467	0.270	0.270	-	0.499	-	0.249
M27	0.270	0.267	0.430	0.420	-	0.443	-	0.196
M28	0.530	0.533	0.390	0.380	-	0.499	-	0.249
M29	0.300	0.300	0.630	0.620	-	0.458	-	0.210
M30	0.230	0.233	0.590	0.580	-	0.423	-	0.179

Tablo 12 incelendiğinde, madde güçlük ve madde ayırt edicilik indekslerinin TAP programıyla karşılaştırabildiği görülmektedir. Madde güvenirliği ve madde varyansına ait istatistikler TAP programı tarafından kullanıcıya sunulmamaktadır. Tablo incelendiğinde, madde güçlük ve ayırt edicilik indekslerine ait değerlerin bazı maddelerde aynı iken, bazı değerlerde farklılık gösterdiği görülmektedir. Madde

güçlük indeksi ele alındığında, 6 maddenin güçlük indeksinin tamamen aynı olduğu, 24 maddenin de birbirine oldukça yakın olduğu görülmektedir. Bu farklılaşma 0.000 - 0.010 şeklinde bir aralıkta değişkenlik göstermektedir. Bu farklılıklar da uygulamaların, arka planda hesaplamaları yaparken, virgülden kaç basamak sonrasını kullanmalarına bağlı olarak değişkenlik göstermektedir. Bu noktada aynı formüller kullanılmasına rağmen, matematiksel basamak yuvarlama işlemlerinin farklılaşmasından ötürü, elde edilen istatistiklerde de farklılaşma gözlenebilmektedir.

Tablo 13. 30 Kişilik Öğrenci Grubuna Uygulanan 2. Sınavın Madde İstatistikleri

Test 2	Madde Güçlük İndeksi		Madde Ayırt Edicilik İndeksi		Madde Güvenirliği		Madde Varyansı	
	TAP	G.U	TAP	G.U	TAP	G.U	TAP	G.U
M1	0.670	0.667	0.230	0.220	-	0.471	-	0.222
M2	0.130	0.133	0.370	0.370	-	0.341	-	0.116
M3	0.500	0.500	0.330	0.330	-	0.500	-	0.250
M4	0.370	0.367	0.260	0.250	-	0.482	-	0.232
M5	0.830	0.833	0.450	0.450	-	0.373	-	0.139
M6	0.570	0.567	0.370	0.670	-	0.496	-	0.246
M7	0.630	0.633	0.570	0.560	-	0.482	-	0.232
M8	0.200	0.200	0.270	0.260	-	0.400	-	0.160
M9	0.370	0.367	0.100	0.100	-	0.482	-	0.232
M10	0.700	0.700	0.530	0.520	-	0.458	-	0.210
M11	0.470	0.467	0.620	0.610	-	0.499	-	0.249
M12	0.130	0.133	0.410	0.400	-	0.341	-	0.116
M13	0.700	0.700	0.610	0.600	-	0.458	-	0.210
M14	0.600	0.600	0.500	0.490	-	0.490	-	0.240
M15	0.770	0.767	0.330	0.320	-	0.423	-	0.179
M16	0.330	0.333	0.350	0.340	-	0.471	-	0.222
M17	0.330	0.333	0.280	0.270	-	0.471	-	0.222
M18	0.800	0.800	0.220	0.210	-	0.400	-	0.160
M19	0.400	0.400	0.420	0.420	-	0.490	-	0.240
M20	0.130	0.133	0.240	0.240	-	0.341	-	0.116
M21	0.570	0.567	0.600	0.590	-	0.496	-	0.246
M22	0.730	0.733	0.440	0.440	-	0.443	-	0.196
M23	0.630	0.633	0.160	0.150	-	0.482	-	0.232
M24	0.530	0.533	0.090	0.090	-	0.499	-	0.249
M25	0.130	0.133	0.620	0.610	-	0.341	-	0.116
M26	0.830	0.833	0.200	0.190	-	0.373	-	0.139
M27	0.370	0.367	0.460	0.450	-	0.482	-	0.232
M28	0.730	0.733	0.460	0.450	-	0.443	-	0.196
M29	0.570	0.567	0.660	0.650	-	0.496	-	0.246
M30	0.330	0.333	0.440	0.430	-	0.471	-	0.222

Tablo 13 incelendiğinde, istatistikler arası farklılaşmanın en çok 0.010 kadar olduğu görülmektedir. Bu farklılaşmanın sebebi ise, her iki uygulamanın arka planda matematiksel basamak yuvarlama işlemlerini, aynı ondalık basamağa kadar yapmıyor oluşudur.

Tablo 14. 30 Kişilik Öğrenci Grubuna Uygulanan 3. Sınavın Madde İstatistikleri

Test 3	Madde Güçlük İndeksi		Madde Ayırt Edicilik İndeksi		Madde Güvenirliği		Madde Varyansı	
	TAP	G.U	TAP	G.U	TAP	G.U	TAP	G.U
M1	0.970	0.967	-0.080	-0.080	-	0.179	-	0.032
M2	0.400	0.400	0.530	0.520	-	0.490	-	0.240
M3	0.900	0.900	0.270	0.270	-	0.300	-	0.090
M4	0.700	0.700	0.660	0.650	-	0.458	-	0.210
M5	0.970	0.967	-0.180	-0.180	-	0.179	-	0.032
M6	0.830	0.833	0.480	0.470	-	0.373	-	0.139
M7	0.930	0.933	-0.120	-0.120	-	0.249	-	0.062
M8	0.530	0.533	0.440	0.430	-	0.499	-	0.249
M9	0.770	0.767	0.550	0.540	-	0.423	-	0.179
M10	0.830	0.833	0.000	0.000	-	0.373	-	0.139
M11	0.770	0.767	0.630	0.620	-	0.423	-	0.179
M12	0.730	0.733	0.550	0.540	-	0.443	-	0.196
M13	0.870	0.867	0.620	0.610	-	0.341	-	0.116
M14	0.830	0.833	0.340	0.330	-	0.373	-	0.139
M15	0.930	0.933	0.240	0.240	-	0.249	-	0.062
M16	0.700	0.700	0.640	0.630	-	0.458	-	0.210
M17	0.770	0.767	0.480	0.470	-	0.423	-	0.179
M18	1.000	1.000	****	Null	-	0.000	-	0.000
M19	0.770	0.767	0.550	0.550	-	0.423	-	0.179
M20	0.500	0.500	0.060	0.060	-	0.500	-	0.250
M21	0.930	0.933	0.350	0.350	-	0.249	-	0.062
M22	0.930	0.933	0.280	0.280	-	0.249	-	0.062
M23	0.930	0.933	0.030	0.030	-	0.249	-	0.062
M24	0.930	0.933	0.280	0.280	-	0.249	-	0.062
M25	0.600	0.600	0.450	0.440	-	0.490	-	0.240
M26	0.900	0.900	0.480	0.480	-	0.300	-	0.090
M27	0.800	0.800	0.320	0.320	-	0.400	-	0.160
M28	0.970	0.967	0.070	0.070	-	0.179	-	0.032
M29	0.970	0.967	-0.030	-0.030	-	0.179	-	0.032
M30	0.700	0.700	0.220	0.220	-	0.458	-	0.210

Tablo 14 incelendiğinde, istatistikler arası farklılaşmanın en çok 0.010 kadar olduğu görülmekle beraber, bu farklılaşmanın sebebi de, her iki uygulamanın arka planda matematiksel basamak yuvarlama işlemlerini, aynı ondalık basamağa kadar

yapmıyor oluşundan kaynaklanmaktadır. Madde 18 incelendiğinde de, ayırt edicilik indeksi ihmal edilebilecek kadar küçük bir değer verdiğinde, TAP uygulamasında "****" , geliştirilen uygulamada ise "null" şeklinde geri dönüş olmaktadır.

Tablo 15. 100 Kişilik Öğrenci Grubuna Uygulanan 1. Sınavın Madde İstatistikleri

Test 4	Madde Güçlük İndeksi		Madde Ayırt Edicilik İndeksi		Madde Güvenirliği		Madde Varyansı	
	TAP	G.U	TAP	G.U	TAP	G.U	TAP	G.U
M1	0.210	0.210	0.310	0.310	-	-0.407	-	0.166
M2	0.380	0.380	0.530	0.520	-	0.486	-	0.236
M3	0.350	0.350	0.220	0.220	-	0.476	-	0.227
M4	0.150	0.150	0.240	0.240	-	0.358	-	0.128
M5	0.190	0.190	0.450	0.450	-	0.392	-	0.154
M6	0.270	0.270	0.400	0.390	-	0.444	-	0.197
M7	0.450	0.450	0.390	0.380	-	0.498	-	0.248
M8	0.590	0.590	0.460	0.460	-	0.492	-	0.242
M9	0.270	0.270	0.520	0.520	-	0.444	-	0.197
M10	0.510	0.510	0.440	0.440	-	0.500	-	0.250
M11	0.450	0.450	0.400	0.400	-	0.498	-	0.248
M12	0.250	0.250	0.410	0.400	-	0.434	-	0.188
M13	0.200	0.200	0.360	0.360	-	0.400	-	0.160
M14	0.440	0.440	0.450	0.440	-	0.496	-	0.246
M15	0.410	0.410	0.520	0.520	-	0.492	-	0.242
M16	0.230	0.230	0.530	0.530	-	0.421	-	0.177
M17	0.470	0.470	0.310	0.310	-	0.499	-	0.249
M18	0.470	0.470	0.430	0.430	-	0.499	-	0.249
M19	0.190	0.190	0.350	0.350	-	0.392	-	0.154
M20	0.370	0.370	0.350	0.350	-	0.483	-	0.233
M21	0.490	0.490	0.530	0.530	-	0.500	-	0.250
M22	0.140	0.140	0.310	0.310	-	0.346	-	0.120
M23	0.240	0.240	0.220	0.220	-	0.427	-	0.182
M24	0.500	0.500	0.310	0.310	-	0.500	-	0.250
M25	0.490	0.490	0.510	0.500	-	0.500	-	0.250
M26	0.160	0.160	0.430	0.420	-	0.366	-	0.134
M27	0.320	0.320	0.270	0.270	-	0.467	-	0.218
M28	0.120	0.120	0.310	0.310	-	0.326	-	0.106
M29	0.370	0.370	0.420	0.410	-	0.483	-	0.233
M30	0.180	0.180	0.470	0.460	-	0.385	-	0.148

Tablo 15'te de görüldüğü gibi, istatistikler arası farklılaşma en çok 0.010 kadardır. Bu farklılaşmanın sebebi de, her iki uygulamanın arka planda matematiksel basamak yuvarlama işlemlerini, aynı ondalık basamağa kadar yapmıyor oluşundan kaynaklanmaktadır.

Tablo 16. 100 Kişilik Öğrenci Grubuna Uygulanan 2. Sınavın Madde İstatistikleri

Test 5	Madde Güçlük İndeksi		Madde Ayırt Edicilik İndeksi		Madde Güvenirliği		Madde Varyansı	
	TAP	G.U	TAP	G.U	TAP	G.U	TAP	G.U
M1	0.500	0.500	0.400	0.400	-	0.500	-	0.250
M2	0.490	0.490	0.490	0.480	-	0.500	-	0.250
M3	0.600	0.600	0.360	0.360	-	0.490	-	0.240
M4	0.290	0.290	0.370	0.370	-	0.454	-	0.206
M5	0.410	0.410	0.500	0.500	-	0.492	-	0.242
M6	0.410	0.410	0.390	0.380	-	0.492	-	0.242
M7	0.680	0.680	0.340	0.340	-	0.467	-	0.218
M8	0.600	0.600	0.440	0.440	-	0.490	-	0.240
M9	0.470	0.470	0.450	0.440	-	0.499	-	0.249
M10	0.740	0.740	0.290	0.290	-	0.438	-	0.192
M11	0.530	0.530	0.390	0.390	-	0.499	-	0.249
M12	0.360	0.360	0.370	0.370	-	0.480	-	0.230
M13	0.320	0.320	0.430	0.420	-	0.467	-	0.218
M14	0.570	0.570	0.330	0.330	-	0.495	-	0.245
M15	0.610	0.610	0.420	0.410	-	0.488	-	0.238
M16	0.410	0.410	0.330	0.330	-	0.492	-	0.242
M17	0.650	0.650	0.420	0.420	-	0.476	-	0.227
M18	0.720	0.720	0.360	0.360	-	0.449	-	0.202
M19	0.330	0.330	0.310	0.310	-	0.470	-	0.221
M20	0.680	0.680	0.490	0.480	-	0.467	-	0.218
M21	0.790	0.790	0.250	0.240	-	0.407	-	0.166
M22	0.280	0.280	0.350	0.350	-	0.449	-	0.202
M23	0.420	0.420	0.510	0.500	-	0.494	-	0.244
M24	0.660	0.660	0.400	0.400	-	0.473	-	0.224
M25	0.660	0.660	0.360	0.360	-	0.473	-	0.224
M26	0.390	0.390	0.570	0.570	-	0.488	-	0.238
M27	0.550	0.550	0.480	0.480	-	0.497	-	0.247
M28	0.320	0.320	0.400	0.400	-	0.467	-	0.218
M29	0.550	0.550	0.540	0.540	-	0.497	-	0.247
M30	0.410	0.410	0.600	0.600	-	0.492	-	0.242

Tablo 16 incelenirse, istatistikler arası farklılaşmalar bulunmakta ve bu farklılaşmalar 0.000-0.010 arası değişiklik göstermektedir. Bu durumun sebebi ise, her iki uygulamanın arka planda matematiksel basamak yuvarlama işlemlerini, aynı ondalık basamağa kadar yapmıyor olmasıdır.

Tablo 17. 100 Kişilik Öğrenci Grubuna Uygulanan 3. Sınavın Madde İstatistikleri

Test 6	Madde Güçlük İndeksi		Madde Ayırt Edicilik İndeksi		Madde Güvenirliği		Madde Varyansı	
	TAP	G.U	TAP	G.U	TAP	G.U	TAP	G.U
M1	0.740	0.740	0.460	0.460	-	0.438	-	0.192
M2	0.890	0.890	0.310	0.310	-	0.313	-	0.098
M3	0.900	0.900	0.340	0.340	-	0.300	-	0.090
M4	0.730	0.730	0.310	0.310	-	0.444	-	0.197
M5	0.780	0.780	0.460	0.460	-	0.415	-	0.172
M6	0.840	0.840	0.290	0.290	-	0.366	-	0.134
M7	0.920	0.920	0.170	0.170	-	0.272	-	0.074
M8	0.940	0.940	0.330	0.330	-	0.237	-	0.056
M9	0.810	0.810	0.400	0.390	-	0.392	-	0.154
M10	0.930	0.930	0.380	0.380	-	0.255	-	0.065
M11	0.900	0.900	0.250	0.250	-	0.300	-	0.090
M12	0.700	0.700	0.250	0.240	-	0.458	-	0.210
M13	0.750	0.750	0.440	0.440	-	0.434	-	0.188
M14	0.910	0.910	0.350	0.350	-	0.286	-	0.082
M15	0.880	0.880	0.520	0.520	-	0.326	-	0.106
M16	0.820	0.820	0.260	0.260	-	0.385	-	0.148
M17	0.900	0.900	0.370	0.370	-	0.300	-	0.009
M18	0.940	0.940	0.270	0.260	-	0.237	-	0.056
M19	0.790	0.790	0.460	0.460	-	0.407	-	0.166
M20	0.930	0.930	0.250	0.250	-	0.255	-	0.065
M21	0.920	0.920	0.230	0.230	-	0.272	-	0.074
M22	0.760	0.760	0.250	0.250	-	0.427	-	0.182
M23	0.760	0.760	0.490	0.480	-	0.427	-	0.182
M24	0.940	0.940	0.060	0.060	-	0.237	-	0.056
M25	0.970	0.970	0.080	0.080	-	0.170	-	0.029
M26	0.830	0.830	0.270	0.270	-	0.375	-	0.141
M27	0.920	0.920	0.290	0.290	-	0.272	-	0.074
M28	0.740	0.740	0.260	0.250	-	0.438	-	0.192
M29	0.850	0.850	0.380	0.370	-	0.358	-	0.128
M30	0.830	0.830	0.310	0.310	-	0.375	-	0.141

Tablo 17 incelendiğinde, istatistikler arası farklılaşmaların olduğu görülmekte ve bu durum 0.000-0.010 arası değişiklik göstermektedir. Bu durumun sebebi ise, her iki uygulamanın arka planda matematiksel basamak yuvarlama işlemlerini, aynı ondalık basamağa kadar yapmıyor olmasıdır.

Tablo 18. 500 Kişilik Öğrenci Grubuna Uygulanan 1. Sınavın Madde İstatistikleri

Test 7	Madde Güçlük İndeksi		Madde Ayırt Edicilik İndeksi		Madde Güvenirliği		Madde Varyansı	
	TAP	G.U	TAP	G.U	TAP	G.U	TAP	G.U
M1	0.230	0.230	0.360	0.350	-	0.421	-	0.177
M2	0.470	0.474	0.340	0.340	-	0.499	-	0.249
M3	0.360	0.360	0.390	0.390	-	0.480	-	0.230
M4	0.230	0.230	0.360	0.350	-	0.421	-	0.177
M5	0.490	0.488	0.330	0.330	-	0.500	-	0.250
M6	0.300	0.298	0.370	0.370	-	0.457	-	0.209
M7	0.220	0.220	0.330	0.330	-	0.415	-	0.172
M8	0.390	0.390	0.420	0.420	-	0.488	-	0.238
M9	0.490	0.490	0.410	0.410	-	0.500	-	0.250
M10	0.360	0.360	0.330	0.330	-	0.480	-	0.230
M11	0.570	0.572	0.370	0.370	-	0.495	-	0.245
M12	0.300	0.302	0.400	0.400	-	0.459	-	0.211
M13	0.690	0.686	0.370	0.370	-	0.464	-	0.215
M14	0.460	0.460	0.430	0.430	-	0.498	-	0.248
M15	0.220	0.222	0.380	0.380	-	0.416	-	0.173
M16	0.310	0.308	0.360	0.360	-	0.462	-	0.213
M17	0.480	0.476	0.470	0.470	-	0.499	-	0.249
M18	0.420	0.416	0.380	0.380	-	0.493	-	0.243
M19	0.640	0.644	0.360	0.360	-	0.479	-	0.229
M20	0.450	0.446	0.390	0.390	-	0.497	-	0.247
M21	0.370	0.374	0.380	0.380	-	0.484	-	0.234
M22	0.770	0.768	0.360	0.360	-	0.422	-	0.178
M23	0.520	0.518	0.390	0.390	-	0.500	-	0.250
M24	0.690	0.694	0.370	0.370	-	0.460	-	0.212
M25	0.660	0.658	0.380	0.380	-	0.474	-	0.225
M26	0.300	0.304	0.430	0.430	-	0.460	-	0.212
M27	0.250	0.248	0.340	0.340	-	0.431	-	0.186
M28	0.610	0.612	0.340	0.340	-	0.487	-	0.237
M29	0.600	0.598	0.390	0.390	-	0.490	-	0.240
M30	0.290	0.286	0.350	0.350	-	0.452	-	0.204

Tablo 18'de görüldüğü gibi, değerler karşılaştırıldığında istatistikler arası farklılaşmalar bulunmakta ve bu farklılaşmalar 0.000-0.010 arası değişiklik göstermektedir. Bu durum, her iki uygulamanın arka planda matematiksel basamak yuvarlama işlemlerini, aynı ondalık basamağa kadar yapmıyor olmasından kaynaklanmaktadır.

Tablo 19. 500 Kişilik Öğrenci Grubuna Uygulanan 2. Sınavın Madde İstatistikleri

Test 8	Madde Güçlük İndeksi		Madde Ayırt Edicilik İndeksi		Madde Güvenirliği		Madde Varyansı	
	TAP	G.U	TAP	G.U	TAP	G.U	TAP	G.U
M1	0.440	0.440	0.430	0.430	-	0.496	-	0.246
M2	0.730	0.726	0.390	0.390	-	0.446	-	0.199
M3	0.540	0.544	0.450	0.440	-	0.498	-	0.248
M4	0.410	0.412	0.380	0.380	-	0.492	-	0.242
M5	0.700	0.700	0.380	0.380	-	0.458	-	0.210
M6	0.490	0.492	0.390	0.390	-	0.500	-	0.250
M7	0.420	0.422	0.450	0.450	-	0.494	-	0.244
M8	0.610	0.614	0.350	0.350	-	0.487	-	0.237
M9	0.700	0.698	0.410	0.410	-	0.459	-	0.211
M10	0.590	0.588	0.450	0.450	-	0.492	-	0.242
M11	0.760	0.756	0.380	0.380	-	0.429	-	0.184
M12	0.520	0.524	0.440	0.440	-	0.499	-	0.249
M13	0.850	0.850	0.270	0.270	-	0.358	-	0.128
M14	0.680	0.676	0.390	0.390	-	0.468	-	0.219
M15	0.400	0.400	0.350	0.350	-	0.490	-	0.240
M16	0.490	0.492	0.460	0.460	-	0.500	-	0.250
M17	0.680	0.678	0.390	0.390	-	0.467	-	0.218
M18	0.640	0.636	0.420	0.420	-	0.482	-	0.232
M19	0.840	0.836	0.350	0.350	-	0.370	-	0.137
M20	0.660	0.658	0.370	0.370	-	0.474	-	0.225
M21	0.580	0.584	0.430	0.430	-	0.493	-	0.243
M22	0.910	0.910	0.210	0.210	-	0.286	-	0.082
M23	0.760	0.756	0.360	0.360	-	0.429	-	0.184
M24	0.830	0.828	0.300	0.300	-	0.377	-	0.142
M25	0.830	0.828	0.320	0.320	-	0.377	-	0.142
M26	0.470	0.466	0.450	0.450	-	0.499	-	0.249
M27	0.430	0.432	0.420	0.420	-	0.495	-	0.245
M28	0.780	0.784	0.340	0.340	-	0.411	-	0.169
M29	0.770	0.774	0.350	0.350	-	0.418	-	0.175
M30	0.500	0.504	0.360	0.360	-	0.500	-	0.250

Tablo 19'a göre, istatistikler arası farklılaşmalar bulunmakta ve bu durumun 0.000-0.010 arası değiştiği görülmektedir. Bu durumun sebebi ise, her iki uygulamanın arka planda matematiksel basamak yuvarlama işlemlerini, aynı ondalık basamağa kadar yapmıyor olmasıdır.

Tablo 20. 500 Kişilik Öğrenci Grubuna Uygulanan 3. Sınavın Madde İstatistikleri

Test 9	Madde Güçlük İndeksi		Madde Ayırt Edicilik İndeksi		Madde Güvenirliği		Madde Varyansı	
	TAP	G.U	TAP	G.U	TAP	G.U	TAP	G.U
M1	0.830	0.828	0.320	0.320	-	0.377	-	0.142
M2	0.960	0.956	0.200	0.200	-	0.205	-	0.042
M3	0.870	0.870	0.270	0.270	-	0.336	-	0.113
M4	0.810	0.808	0.370	0.370	-	0.394	-	0.155
M5	0.950	0.950	0.260	0.260	-	0.219	-	0.048
M6	0.860	0.856	0.360	0.360	-	0.351	-	0.123
M7	0.810	0.810	0.370	0.370	-	0.392	-	0.154
M8	0.890	0.886	0.390	0.390	-	0.318	-	0.101
M9	0.920	0.920	0.230	0.230	-	0.272	-	0.074
M10	0.900	0.898	0.370	0.370	-	0.303	-	0.092
M11	0.950	0.950	0.210	0.210	-	0.219	-	0.048
M12	0.850	0.854	0.380	0.380	-	0.354	-	0.125
M13	0.980	0.982	0.060	0.060	-	0.134	-	0.018
M14	0.920	0.924	0.270	0.270	-	0.265	-	0.070
M15	0.790	0.790	0.440	0.440	-	0.407	-	0.166
M16	0.870	0.872	0.280	0.280	-	0.335	-	0.112
M17	0.930	0.930	0.240	0.240	-	0.255	-	0.065
M18	0.940	0.942	0.250	0.250	-	0.235	-	0.055
M19	0.980	0.976	0.170	0.170	-	0.152	-	0.023
M20	0.940	0.936	0.180	0.180	-	0.245	-	0.060
M21	0.900	0.900	0.290	0.290	-	0.300	-	0.090
M22	0.970	0.974	0.200	0.200	-	0.158	-	0.025
M23	0.960	0.958	0.210	0.210	-	0.200	-	0.040
M24	0.980	0.982	0.150	0.150	-	0.134	-	0.018
M25	0.970	0.968	0.210	0.210	-	0.176	-	0.031
M26	0.890	0.888	0.310	0.310	-	0.315	-	0.099
M27	0.850	0.852	0.340	0.340	-	0.355	-	0.126
M28	0.950	0.948	0.230	0.230	-	0.221	-	0.049
M29	0.980	0.978	0.150	0.150	-	0.148	-	0.022
M30	0.840	0.842	0.380	0.380	-	0.365	-	0.133

Daha önce de belirtildiği üzere, bütün veri gruplarına dair madde istatistikleri teker teker incelendiğinde, TAP programı ile tamamen aynı değerlerin ya da oldukça ihmal edilebilir bir farklılaşmanın olduğu görülmektedir. Bunun da madde ayırt edicilik, madde varyans, madde güçlük ve madde güvenirlik değerleri hesaplanırken kullanılan diğer değişkenlerin ondalık kısmının kaç basamak olarak dahil edilip, yuvarlanma işleminin yapılıp yapılmadığına bağlı olarak değişkenlik gösterdiği bilinmektedir. Bununla birlikte, araştırmacı tarafından geliştirilen uygulama,

kullanıcıya Z ve T puanlarına dair de bir sonuç vermektedir. Z ve T puanlarına dair bir karşılaştırma tablosu aşağıda okuyucuya sunulacaktır. Fakat kullanışlılık açısından, tablo olarak sadece 30 kişilik bir Z ve T puanı tablosu verilecektir. 100 kişilik ve 500 kişilik gruba dair Z ve T değerleri ise ekte sunulacaktır.

Tablo 21. 30 Kişilik Öğrenci Grubuna Uygulanan 1. Sınavın Z Ve T Puan Tablosu

Test 1	Z Puanları		T Puanları	
	TAP	G.U	TAP	G.U
Ö1	-1.640	-1.611	-	33.890
Ö2	-1.120	-1.097	-	39.030
Ö3	-1.120	-1.097	-	39.030
Ö4	-1.120	-1.097	-	39.030
Ö5	-1.120	-1.097	-	39.030
Ö6	-0.940	-0.926	-	40.740
Ö7	-0.940	-0.926	-	40.740
Ö8	-0.770	-0.754	-	42.460
Ö9	-0.770	-0.754	-	42.460
Ö10	-0.420	-0.411	-	45.890
Ö11	-0.420	-0.411	-	45.890
Ö12	-0.240	-0.240	-	47.600
Ö13	-0.240	-0.240	-	47.600
Ö14	-0.240	-0.240	-	47.600
Ö15	-0.070	-0.069	-	49.310
Ö16	-0.070	-0.069	-	49.310
Ö17	-0.070	-0.069	-	49.310
Ö18	-0.070	-0.069	-	49.310
Ö19	-0.070	-0.069	-	49.310
Ö20	0.100	0.103	-	51.030
Ö21	0.100	0.103	-	51.030
Ö22	0.100	0.103	-	51.030
Ö23	0.280	0.274	-	52.740
Ö24	0.450	0.446	-	54.460
Ö25	1.320	1.303	-	63.030
Ö26	1.500	1.474	-	64.740
Ö27	1.670	1.645	-	66.450
Ö28	1.670	1.645	-	66.450
Ö29	2.020	1.988	-	69.880
Ö30	2.200	2.160	-	71.600

Tablo 21 incelendiğinde, araştırmacı tarafından geliştirilen yazılımın verdiği Z puanlarıyla TAP programının verdiği değerler arasında 0.001 ile 0.040 arasında değişebilen farklılıklar vardır. Bu farklılıklar, Z puanı hesaplanırken kullanılan

ortalama ve standart sapma değerlerinden kaynaklanmaktadır. SPSS programında, arka planda hesaplama yapılırken ondalık kısımlarının kaçıncı basamaktan itibaren matematiksel yuvarlama yapılmasına bağlı olmasından kaynaklanmaktadır. T puanları incelendiğinde ise, TAP programının buna dair bir çıktısı olmadığı görülmektedir. T puanının pratiklik ve yorumlama açısından daha kullanışlı olduğu genel görüşü dikkate alınırsa, bu durumun kullanıcı açısından geliştirilen yazılımı öne çıkaran ve avantaj sayılabilecek bir durum olduğu söylenebilir.

5.2. Geliştirilen uygulamanın kullanılabilirlik anketine bağlı olarak elde edilen bulgular

Geliştirilen çevrimiçi uygulama ile elde edilen istatistiklerin hali hazırda kullanılan ve geçerli güvenilir sonuç üreten farklı programlar ile karşılaştırılması dışında aynı zamanda sistemin kullanılabilirliği de öğretmenlere uygulanan kullanılabilirlik anketi ile belirlenmiştir. Tablo 2'de gösterilen kullanılabilirlik anketi ve alt boyutları incelendiğinde, kullanılabilir bir yazılımda var olması gerekenlerin belirleyicisi olan 5 özelliğe yönelik olduğu görülmektedir. Bu 5 özellik etkililik, etkinlik, caziplik, hatalara karşı olumlu yaklaşım ve kolay öğrenilebilirliktir (Alexander, 2003). Anketin alt boyutlarına bakıldığında da bilgi kalitesi boyutu etkililik özelliğini; arayüz kalitesi ve sistem yararlılığı alt boyutları etkinlik, hatalara karşı olumlu yaklaşım, kolay öğrenilebilirlik ve caziplik özelliklerini; genel memnuniyet alt boyutu ise kullanıcın sisteme karşı olumlu yaklaşımlarının derecesini ölçmektedir.

Geliştirilen çevrimiçi uygulama kullanılabilirliğini ölçmek adına, Kocaeli ilinde çalışmakta olan 15 farklı branştan ve yaş grubundan öğretmene Oğuzhan Erdinç tarafından Türkçe'ye uyarlanan, kullanılabilirlik anketi uygulanmıştır. Aynı zamanda, uygulamayı kullanma aşamasında soruları yanıtlarken sesli düşünmeleri istenmiştir. Kullanıcıların her birinin uygulamayı kullanım sürelerince (kendilerine verilen temsili test verileriyle) tıklama sayılarının ve erişim sürelerinin kayıtları tutulmuştur, bu süreç kullanıcının en baştan kendisine verilen datayı analiz etmesi, rapora ulaşması, ulaştığı raporu kaydetmesi ve kaydettiği raporu yeniden görüntülemesi süreçlerini kapsamaktadır. Tıklama sayıları ve erişim sürelerinin elde edilmesi uygulamanın

etkinlik, kolay öğrenebilirlik yönlerini ortaya koyma amacıyla yapılmıştır. Anket uygulamasına dair bulgular aşağıdaki gibidir.

Tablo 22. Kullanıcıların tıklanma sayısı ve erişim süreleri

Kullanıcı No	Tıklama Sayısı	Erişim Süresi (dk:sn)
K_1	22	01:22
K_2	27	01:37
K_3	19	01:20
K_4	25	01:35
K_5	24	01:32
K_6	32	01:40
K_7	35	01:45
K_8	28	01:40
K_9	26	01:35
K_{10}	31	01:32
K_{11}	27	01:35
K_{12}	27	01:37
K_{13}	29	01:40
K_{14}	28	01:32
K_{15}	37	01:45

Tablo 22 incelendiğinde, her kullanıcının kendi deneyimleri boyunca kaç kez tıklama yaptıkları ve bu deneyim boyunca toplam ne kadar zaman harcadıkları belirtilmiştir. Veriler incelendiğinde, kullanıcıların ortalama tıklama sayısının 27.800 ve uygulamada geçen toplam sürenin 1:35 olduğu görülmektedir. Harcanan zaman, sarf edilen efor, analiz sonucunda ulaşılan raporun detayı ve işlevi düşünüldüğünde, yazılımın oldukça işlevsel olduğu söylenebilir.

Bununla birlikte, anket uygulanan öğretmenlerin, her bir maddeye verdikleri puan aşağıda Tablo 23'te verilmiştir. Bu noktada öğretmenlerin, genel olarak ara yüzle alakalı memnuniyetsizliklerinin olduğu görülmüştür. Daha sonra web sitesi ile ilgili görüşler bölümünde de belirtileceği üzere, kullanıcılar genellikle ara yüzün rengi, bir

işlem yapmak için tıklandığında, o işlemin gerçekleşip gerçekleşmediğinin belli olmaması gibi durumlardan dolayı ara yüz hakkında olumsuz görüş bildirmişlerdir. Bu geri bildirimler doğrultusunda uygulamada değişiklikler yapılmış ve bu 15 öğretmenin yazılımı tekrardan kullanması sağlanmıştır. Değişikliklerden sonra öğretmen görüşlerinde olumlu yönde değişimler olduğu görülmüştür. Öğretmenlerin anket puanlarına ilişkin sonuçlar Tablo 23'te sunulmuştur.

Tablo 23. Kullanıcıların kullanılabilirlik anketi aracılığıyla uygulamaya verdikleri puanlar

Kullanıcı No	İlk Anket Puanı	İkinci Anket Puanı
1	12	12
2	14	12
3	18	13
4	15	13
5	18	13
6	16	14
7	16	12
8	15	13
9	15	13
10	12	12
11	15	12
12	14	12
13	16	13
14	15	13
15	18	13

Tablo 23 incelendiğinde, öğretmenlerden gelen geri bildirimlere göre uygulamada değişiklikler yapıldığında, öğretmenlerin %13.33'lük bir diliminin fikrinde değişiklik gözlemlenmediği, % 86.67'lik diliminde ise olumlu anlamda fikir değişikliği gözlemlenmiştir.

5.3. Uygulamayla ile ilgili genel görüşler

Daha önce de belirtildiği üzere, kullanıcılar uygulamayı kullanılırken yüksek sesle verdikleri geri bildirimler kayıt altına alınmıştır. Kayıt altına alınan bu düşünceler doğrultusunda da, uygulamada ve uygulamanın verdiği rapor çıktısı konusunda düzeltmeler yapılmıştır. Bu bağlamda, kullanıcıların uygulamayı kullanırken verdiği, geri bildirimlere bu bölümde yer verilmiştir ve aşağıdaki gibidir.

K_1 uygulamanın verdiği rapor hakkındaki olumlu görüşünü,

K_1-*"...Raporda her değerin altında, değer aralıklarının olması, hepsinin ne anlama geldiğinin bize açıklanması gerçekten güzel olmuş..."* şeklinde belirtmiştir.

K_7, uygulamayı kullanırken bazı aşamalarda, uygulamanın kullanıcı tarafından yapılan işlemin gerçekleştirilip gerçekleştirilmedi konusunda yetersizlik olduğunu vurgulamış ve bu geri bildirim sonrasında uygulama düzeltilmiştir. Yazılımın ara yüzüne dair geliştirilen bu eleştiri şu şekilde bildirilmiştir.

K_7-*"...Bu cevap anahtarını ekledikten sonra tıklıyorum, bir şey olmuyor, ekledi mi, eklemedi mi ne olduğunu anlayamadım..."*

Yazılıma dair gelen eleştirilerden biri de, çıktı olarak verilen raporun sadece test istatistiklerine dair bir sonuç sunuyor olmasıdır. Bu noktada, raporda öğrencilere dair bilgi görmek istediğini belirten K_5;

K_5- *"...Aslında sadece sınav istatistikleri değil de, öğrenci doğru yanlış sayılarını da görsek mesela, daha iyi olabilir..."* şeklinde görüş bildirmiştir.

Ara yüze dair gelen eleştirilerden biri de, K_7 'nin belirttiği problemle paralel göstermekle birlikte, veri tabanına aktarma sisteminin gerçekleşip gerçekleşmemesi ile ilgilidir. K_3 bunu şu sözlerle dile getirmiştir.

K_3- *"...Sınav istatistiklerini hesapladıktan sonra veri tabanına aktarmak istedim, tıkladım, bir daha tıkladım, ama hiçbir şey olmadı. Ama sonradan fark ettim ki; zaten tıkladığımda veri tabanına kaydetmiş, ama bana bunu hiçbir şekilde bildirmediği için, anlayamadım maalesef..."*

K_{12} uygulamada toplu olarak öğrenci bilgisi yüklemenin CSV dosya yoluyla olmasına dair görüşlerini aşağıdaki gibi belirtmiştir.

K_{12}- *"...Öğrencileri toplu olarak CSV dosyasıyla yükleme durumu beni ilk başta biraz korkuttu, nasıl yaparım diye ama sitede güzel bir şekilde anlatıldığından yapabildiğimi düşünüyorum..."*

K_9 ise uygulamada üyelik aldıktan sonra kişilere gönderilen onay mailinin zamanlamasına dair görüş bildirmiştir.

K_9- *"...Ben üyeliğimi alırken onay maili biraz geç geldi, orda bir süre bekledim, yani gecikmesi bu tarz şeylerin pek hoş olmuyor açıkçası, ama internetten mi kaynaklıydı belki de, onu da bilemiyorum. Ama eğer uygulamayla ilgiliyse bence düzeltilmeli ..."*

SONUÇ VE ÖNERİLER

Bu bölümde, araştırmanın bulgularına ilişkin elde edilen sonuçlarına yer verilmiş ve bu sonuçlar eşliğinde bazı önerilerde bulunulmuştur.

Bu çalışmada, öğretmenler için test ve madde istatistiklerini hesaplayıp, detaylı bir rapor elde ederek bunun kaydını tutabilecekleri bir platform geliştirmek, aynı zamanda, geliştirilen bu platformun güvenirliği ve kullanılabilirliğini belirleyebilmek, öğretmenlerin kolayca kullanabileceği geçerli ve güvenilir sonuçlar üreten bir analiz platformu oluşturmak amaçlanmıştır. Bu amaçla geliştirilen bu web tabanlı uygulama simülatif yolla elde edilen verilerle test edilmiş olup, SPSS ve TAP gibi uygulamalardan elde edilen analizlerle karşılaştırılmıştır. Test edilme aşamasında araştırmacı tarafından geliştirilen yazılımın, test ve madde istatistikleri değerlerinin SPSS ve TAP paket programlarıyla karşılaştırılması sağlanmıştır. Tap programından simülatif yolla dokuz farklı veri grubu elde edilirken, aynı zamanda madde istatistiklerinin karşılaştırılmasında da yararlanılmıştır. SPSS programından ise test istatistiklerinin karşılaştırılması yönünde yararlanılmıştır. Bununla birlikte 15 kişilik bir öğretmen grubuna farklı zaman aralıklarında iki kez yazılımı kullanma deneyimi sunulmuştur ve bu deneyimlerini kendilerine sunulan anket doğrultusunda puanlamaları istenmiştir. Bu aşamalardan sonra aşağıdaki sonuçlara ulaşılmıştır.

SONUÇLAR

Uygulamada kullanıcıya sunulan rapordaki tüm istatistikler, sadece dosyayı web ortamına yükleyip, cevap anahtarı girişi yapma işlemlerine bağlı olmasından dolayı, kullanıcılar oldukça hızlı bir şekilde analiz sonuçlarının yer aldığı rapora ulaşmışlardır. Bu noktada rapora ulaşmak, kaydetmek ve yeniden görüntülemek için geçen süre bilgilerini içeren Tablo 22, araştırma içerisinde sunulmuştur. Bu tablo incelendiğinde uygulamanın bu bağlamda, oldukça pratik olduğu ve zaman tasarrufu sağladığı görülmüştür. Bununla birlikte, kullanım şeklinin sadece küçük bir kaç aşamadan oluşmasından dolayı, daha aşamalı işlemlerden geçerek bir rapora ulaşılmayı sağlayan muadil programlara göre, kolay ve neredeyse hatasız bir yol izleme seçeneği sunduğu görülmüştür. Bu noktada detaylı bir anlatım sunan kullanıcı rehberinin de önemli katkısının olduğu fark edilmiştir.

Aynı zamanda, kullanım dili, rehberi, veri girişi, raporu dahil tüm bölümlerinde Türkçe dilinin kullanılmasının, iyi seviyede İngilizce kullanamayan kullanıcılar için pozitif bir durum olduğu görülmektedir.

Muadilleriyle kıyaslandığında, programdan elde edilen rapor çıktısının daha anlaşılır, net ve kullanıcıyı yönlendiren taraflarıyla öne çıktığı görülmüştür. Ayırt edici olmadığı düşünülen sorunun sınavdan çıkarılması gerektiğinin kullanıcıya bildirilmesi bunlardan biridir. Raporun, kullanıcının test ve madde istatistiklerini yorumlama gücü ve bilgisine terk edilmemesi, uygulamanın öne çıkan durumlarından biri olmuştur.

Her kullanıcının kendi e posta adresiyle oluşturduğu profilinde, sınavlarını sınıf sınıf ve sınav sınav geriye dönük olarak depolayabilmesinin uzun vadede başarı takibinde kolaylaştırıcı bir etkisinin olabileceği düşünülmektedir.

Öğretmenlere uygulanan kullanılabilirlik anketi doğrultusunda da, uygulamanın kullanılabilirlik değerinin oldukça yüksek olduğu söylenebilir. 15 kişilik öğretmen grubuna yazılımın ilk kez deneyimletilmesinden sonar elde edilen geri bildirimlere göre yeniden birşekillendirme sağlanmış ve ikinci kullanım aşamasında, kullanıcıların deneyimlerini ilk deneyime göre olumlu anlamda puanladıkları görülmüştür.

ÖNERİLER

Uygulamanın geliştirilmesi gereken yönleri de muhakkak ki bulunmaktadır. Bu bağlamda, uygulamanın daha fazla geliştirilmesi konusunda araştırmacılara bir kaç öneri verilebilir.

- Programın uluslararası bir kullanımının da olabilmesi için İngilizce dil seçeneğinin olması düşünülebilir.
- Verilerin çalışmalarda direkt olarak kullanılabilmesi için histogram, çizgi, daire grafiklerini ve normal dağılım eğrilerini de çıktı olarak vermesi seçeneklerinin üzerinde durulabilir.
- Bu program, daha kaliteli ölçme ve değerlendirme yapmak isteyen her öğretmen tarafından, bütün sınıfların, öğrencilerin ve o sınıflara ait test ve madde analiz sonuçların kaydını tutmak amacıyla kullanılabilir.

- Aynı zamanda bu uygulama, SPSS ve TAP gibi programları kullanmayı bilmeyen ya da madde ve test analiz sonuçlarını daha pratik bir şekilde elde etmek isteyen her araştırmacı tarafından kullanılabilir.
- Geliştirilen web tabanlı uygulama mobil kullanım desteklidir. Ancak Google Play Store ya da Apple Store'da bir mobil uygulama olarak bulunmamaktadır. Bu durumda diğer araştırmacılara, bu uygulamanın daha kapsamlı bir mobil deneyim sunabilen bir versiyonunun geliştirilmesi önerilebilir.

KAYNAKÇA

KİTAPLAR

Alıcı, D., Demir, S.B., Çakan, M., Güler, N., Kan, A., Karaca, E., Özbek, Ö.Y., Yaşar, M. (2008). S. Tekindal (Ed.), Eğitimde Ölçme ve Değerlendirme. Ankara: Pegem Akademi.

Arıcı, H. (2001). İstatistik: Yöntemler ve Uygulamalar. Ankara: Meteksan Basımevi.

Baykul, Y. (2000). Eğitimde ve Psikolojide Ölçme. Ankara: ÖSYM.

Büyüköztürk, Ş. (2015). Sosyal bilimler için veri analizi el kitabı. Ankara, Pegem Akademi.

Büyüköztürk,Ş., Çokluk, Ö. ve Köklü, N. (2015). Sosyal Bilimler İçin İstatistik. Ankara, Pegem Akademi.

Crocker, L., and Algina, J. (1986). Introduction classical and modern test theory. New York: Harcourt Brace Javonovich Colege Publishers.

Cronbach, J. L. (1990). Essentials of psychological testing. New York: Harper ve Row Publishers.

Demircioğlu, G. ve Gündoğdu, K. (2008). Ölçme ve Değerlendirme. Ankara: Pagem Akademi.

Demirel, Ö. (2005). Öğretme Sanatı: Öğretimde planlama ve değerlendirme. Ankara: Pegem A Yayıncılık.

Ertürk, S. (1972). Eğitimde program geliştirme. Ankara: Hacettepe Üniversitesi Basımevi.

Hernon, P. (1994). Statistics: A Component of the research process. Norwood, N.J.:Ablex.

Howitt, D. ve Cramer, D. (1997). An Introduction to Statistics in Psychology: a Complete Guide for Students. London: Prentice Hall.

Gronlund, N.E., ve Linn, R.L. (1995). Measurement and Assessment in Teaching (7th ed.) New Jersey: Prentice-Hall.

Kubiszyn, T. ve Borich, G. (2003). Education Testing and Measurement (7th ed.), Hoboken: John Wiley.

Lazar, J. (2001). User-centered Web Development. Boston: Jones and Bartlett Publishers.

Özçelik, Ali Durmuş (2014). Eğitim Programları ve Öğretim. (3. Baskı). Ankara: Pegem A Yayıncılık.

Tekin, H. (2012). Eğitimde Ölçme ve Değerlendirme. Ankara: Yargı Matbaası.

Tekindal, S. (2014). Okullarda Ölçme ve Değerlendirme Yöntemleri. 4. Baskı Ankara: Nobel Akademik Yayıncılık.

Turgut, F., ve Baykul, Y.(2010). Eğitimde Ölçme ve Değerlendirme. Ankara :Pegem Akademi Yayıncılık.

Turgut, F. (1992). Eğitimde ölçme ve değerlendirme metotları. Ankara :Yargı Matbaası.

MAKALELER, BİLDİRİLER, DİĞER BASILI YAYINLAR

Alaçam, Ö. ve Çağıltay, K. (2007). ODTÜ BİDB İnsan Bilgisayar Etkilesimi Laboratuvarı ve Yürütülen Çalışmalar, Akademik Bilişim Konferansı Bildirileri, Kütahya Dumlupınar Üniversitesi.

Alexander, D. (2003). Usability and human factors. Proceedings of Web Workshop series at Monash University. 12 Mart 2019 tarihinde http://www.its.monash. edu.au/staff/web/slideshows/usability-humanfactors/ adresinden erişildi.

Aruk, İ. (2008). Bilişim teknolojilerinin zihinsel engellilerin e-eğitiminde kullanılması ve örnek bir uygulama geliştirilmesi, Yüksek Lisans Tezi, Trakya Üniversitesi, Fen Bilimleri Enstitüsü, Edirne.

Bağış, A. (2003). Arayüz tasarımlarının karşılaştırmalı değerlendirilmesinde kullanılabilirlik yaklaşımı, Makine ve Mühendis Dergisi, 522, 25-31.

Bıçak, B. ve Çakan, M. (2004). Lise öğretmenlerinin sınıf içi ölçme ve değerlendirme uygulamalarına dönük görüşleri. MEB, Orta Öğretimde Yeniden Yapılanma Sempozyumu, Ankara.

Birgin, O. Ve Gürbüz, R. (2008). Sınıf Öğretmeni Adaylarının Ölçme ve Değerlendirme Konusundaki Bilgi Düzeylerinin İncelenmesi. Selçuk Üniversitesi Sosyal Bilimler Enstitüsü Dergisi, 20, 163-179.

Brooks, V. (2002). Assesment in Secondary Schools. Buckingham: Open University Pres. Bloom, B. S. (1956). Taxonomy of Educational Objectives, The Classification of Educational Goals. Handbook I. Cognative Domain. New York, David McKay Company Inc.

Coşkunserçe, O. ve Dursun, Ö. Ö. (2008). Web sitelerinin değerlendirilmesinde kullanılan yöntemler. 8th International Educational Technology Conference Proceedings, 12 Mart 2019 tarihinde http://ietc2008.home.anadolu.edu.tr/ ietc2008/172.doc adresinden erişildi.

Cömert, M. (2008). Bireye uyarlanmış bilgisayar destekli ölçme ve değerlendirme yazılımı geliştirilmesi. Yayınlanmamış Yüksek Lisans Tezi. Bahçeşehir Üniversitesi Fen Bilimleri Enstitüsü, İstanbul.

Daniel, L. G. ve King, D. (1998). A knowledge and use of testing and measurement literacy of elementary and secondary teachers. Journal of Educational Research, 91(6), 331-344.

Elaina, N. (2002). Usability testing for library websites: a hands-on guide. Chicago: American Library Association.

Erdinç O., Lewis J. R., (2013) Psychometric Evaluation of the T-CSUQ: The Turkish Version of the Computer System Usability Questionnaire, International Journal of Human-Computer Interaction, 29, 319-326.

Ertürk, H. (2008). Matematik Öğretmenlerinin Teknoloji Kullanma Yeterliliklerinin Verimliliğe Etkisi. Yüksek Lisans Tezi, Yeditepe Üniversitesi, Sosyal Bilimler Enstitüsü, İstanbul.

Fagan, J.C. (2006). Basic search and visual search: usability tests with EBSCOhost. Conference on Electronic Resources & Libraries.

Fendi, F. (2007) İlköğretim Öğretmenlerinin Teknoloji Kullanım Yeterliliği, Yüksek Lisans Tezi. Yeditepe Üniversitesi, Sosyal Bilimler Enstitüsü, İstanbul.

Gelbal, S., Kelecioğlu, H. (2007). Öğretmenlerin Ölçme Ve Değerlendirme Yöntemleri Hakkındaki Yeterlik Algıları Ve Karşılaştıkları Sorunlar. Hacettepe Üniversitesi Eğitim Fakültesi Dergisi, 33 (33) , 135-145.

Gözütok, F. D., Akgün, Ö. E. ve Karacaoğlu, Ö. C. (2005). İlköğretim Programlarının Öğretmen Yeterlikleri Açısından Değerlendirilmesi. Yeni İlköğretim Programlarını Değerlendirme Sempozyumu (Kayseri). 17-40. Ankara: Sim Matbaası.

Güven, S. (2001). Sınıf Öğretmenlerinin Ölçme ve Değerlendirmede Kullandıkları Yöntem ve Tekniklerin Belirlenmesi, 10.Ulusal Eğitim Bilimleri Kongresi, Abant İzzet Baysal Üniversitesi, Bolu.

Kezer, F. (2013). Bilgisayar Ortamında Bireye Uyarlanmış Test Stratejilerinin Karşılaştırılması. Doktora Tezi, Ankara Üniversitesi, Eğitim Bilimleri Enstitüsü, Ankara.

Kutlu, Ö. (2003). Cumhuriyetin 80. yılında: Ölçme ve değerlendirme, Milli Eğitim Dergisi,160.

Özdemir, Ö., (2007), İlköğretim Birinci Kademe İngilizce Öğretmenlerinin Eğitim Durumunda Yöntem-Teknik ve Araç-Gereç (Teknoloji) Kullanma Yeterlilikleri. Yüksek lisans tezi, Çanakkale Onsekiz Mart Üniversitesi, Sosyal Bilimler Enstitüsü, Çanakkale.

Richey, R. C., ve Klein, J. D. (2008). Research on design and development. In J. M. Spector, M. D. Merrill, J. van Merrienboer, M. P. Driscoll (Eds.), Handbook of Research for Educational Communications and Technology (3rd Ed., pp. 748-757). Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.

Richey, R. C., ve Klein, J. D. (2014). Research on design and development. In J. M. Spector, M. D. Merrill, J. van Merrienboer, M. P. Driscoll (Eds.), Handbook of Research for Educational Communications and Technology (pp. 141-150). NewYork :Springer.

Tekedere, H., ve Mahiroğlu, A. (2012). İlkyardım dersi için web tabanlı probleme dayalı öğrenme yazılımının geliştirilmesi. Ahi Evran Üniversitesi Kırşehir Eğitim Fakültesi Dergisi, 13(2), 95-114.

Tercan, S.S. (2010). Çevrimiçi uyarlamalı bir test aracı geliştirilmesi, uygulanması ve öğrenci görüşlerinin belirlenmesi. Yüksek Lisans Tezi, Sakarya Üniversitesi, Sosyal Bilimler Enstitüsü, Sakarya.

Turan-Oluk, N. ve Ekmekci, G. (2017). Alternatif değerlendirme teknikleri ile geleneksel değerlendirme tekniklerinnin öğrenci başarısını ölçme açısından karşılaştırılması. JRES, 4(2), 172-199.

Yapıcı, M. and Demirdelen C., (2007). Teachers' views with regard to the primary 4th grade social sciences curriculum. Elementary Education Online, 6(2), 204-212.

Yaşar, Ş., Gültekin, M., Türkan , B., Yıldız, N., Girmen, P. (2005). Yeni İlköğretim Programlarının Uygulamasına İlişkin Sınıf Öğretmenlerinin Hazır bulunuşluk Düzeylerinin ve Eğitim Gereksinimlerinin Belirlenmesi (Eskişehir İli Örneği). Yeni İlköğretim Programlarını Değerlendirme Sempozyumu, Kayseri.

ELEKTRONİK KAYNAKLAR

Mediaclick, Bootstrap Nedir. (t.y.). Erişim: 28.04.2019, https://www.mediaclick.com.tr/blog/html-nedir

Mediaclick, Css Nedir. Erişim: 28.04.2019, https://www.mediaclick.com.tr/blog/css-nedir

Mediaclick, Bootstrap Nedir. (t.y.). Erişim: 28.04.2019, https://www.mediaclick.com.tr/blog/bootstrap-nedir

Medium, Javascript. (t.y.). Erişim: 28.04.2019, https://medium.com/@learningzone/asp-net-javascript-jquery-a387ae856be9

Medium, Google Firebase. (t.y.). Erişim: 28.04.2019, https://medium.com/furkanpacikgoz/google-firebase-nedir-ae013e495a74

Niagara. (2003). Niagara University Library web site usability study. 19 Mart 2019 tarihinde http://www.niagara.edu/library/webstudy.pdf adresinden erişildi.

Nielsen, J. (2000). Why you only need to test with 5 users. Jakob Nielsen's Alertbox web sitesinden 9 Mart 2019 tarihinde http://www.useit.com/ alertbox/20000319.html adresinden erişildi.

Nielsen, J. (2003). Usability 101: introduction to usability. Jakob Nielsen's Alertbox web sitesinden 9 Mart 2019 tarihinde http://www.useit.com/ alertbox/20030825.html adresinden erişildi.

University of Maryland. (2007). UM Library's Web Site Usability Testing. 17 Mayıs 2019 tarihinde http://www.lib.umd.edu/itd/ux/site07/index.html adresinden erişildi.

UPA. (2008). UPA, 9 Mart 2019 tarihinde http://www.upassoc.org/usability_resources/about_usability/index.html adresinden erişildi.

ÜTEST. (2008). Orta Doğu Teknik Üniversitesi BİLTİR / ÜTEST Ürün Kullanılabilirlik Test Birimi. 9 Mart 2019 tarihinde http://www.utest.metu. edu.tr/kullanilabilirlik.html adresinden erişildi.

Video Eğitim Portalı, Javascript Akademi. Erişim: 28.04.2019, http://firat.videoegitim.com/kurs/132