

BİLGİSAYAR DESTEKLİ ÇOKTAN SEÇMELİ VE İKİ AŞAMALI KİMYA TESTİNE KATILAN ÖĞRENCİLERİN YANITLAMA SÜRELERİ VE DİKKAT DÜZEYLERİNİN İNCELENMESİ

INVESTIGATION OF THE RESPONSE TIMES AND ATTENTION LEVELS OF STUDENTS TAKING THE COMPUTER AIDED MULTIPLE-CHOICE AND TWO-TIER CHEMISTRY TESTS

Recep AKYEL

Dokuz Eylül Üniversitesi, Buca-İzmir, Türkiye

ORCID: <https://orcid.org/0000-0003-1843-0992>

akyelrecep@hotmail.com

Suat TÜRKOGUZ

Prof.Dr., Dokuz Eylül Üniversitesi, Buca-İzmir, Türkiye

ORCID: <https://orcid.org/0000-0002-7850-2305>

suat.turkoguz@gmail.com

Ali KAYALAR

Dokuz Eylül Üniversitesi, Buca-İzmir, Türkiye

ORCID: <https://orcid.org/0000-0003-3324-7950>

alikayalar1907@gmail.com

Received: January 07, 2023

Accepted: April 14, 2023

Published: April 30, 2023

Suggested Citation:

Akyel, R., Türkoguz, S., & Kayalar, A. (2023). Bilgisayar destekli çoktan seçmeli ve iki aşamalı kimya testine katılan öğrencilerin yanıtlama süreleri ve dikkat düzeylerinin incelenmesi. *International Journal of New Trends in Arts, Sports & Science Education (IJTASE)*, 12(2), 122-142.



This is an open access article under the [CC BY 4.0 license](https://creativecommons.org/licenses/by/4.0/).

Öz

Türkiye’de farklı amaçlarla kullanılan birçok ölçme ve değerlendirme araçları yer almaktadır. En yaygın olarak kullanılanlardan birisi de çoktan seçmeli testlerdir. 1980’li yıllarda çoktan seçmeli testlerin, olumlu yönlerini kullanıp olumsuz yönlerini en aza indirmek için iki aşamalı testler geliştirilmiştir. Bu aşamalı testlerin amacı öğrencinin ne kadar öğrendiğini tespit etmektir. Öğrenme içsel bir süreç olup duyuşsal farkındalık, dikkat, tanıma, dönüştürme, bilginin alınmasını ve işlenmesini içermektedir. Bilginin işlenebilmesi için bilginin alınma süreci dikkat ile başlar. Dikkat, bir nesneye veya düşünce dizisine yönelik olarak zihnin net ve canlı bir pozisyon almasıdır. Dikkatin temelinde odaklanma ve konsantrasyon vardır. Testlerde dikkatin yanı sıra yanıtlama süresi de önemlidir. İlkel çağlardan günümüze hızlı olmak, hayatta kalma mücadelesinin ilk adımı olmuştur. Hızlı olmak kadar etkili olmak da gereklidir. Bu yüzden dikkat düzeyi kadar yapılan işin süresi de önemlidir. Bu çalışmanın amacı, bilgisayar ortamında yapılan çoktan seçmeli kimya testi ile iki aşamalı kimya tanı testinde öğrencilerin dikkat düzeyleri ile yanıtlama sürelerinin incelenerek analiz edilmesidir. Bu kapsamda çoktan seçmeli ve aşamalı testler süresince katılımcıların NeuroSky beyin sensörüyle ölçülen dikkat verilerinin ve yanıtlama sürelerinin kullanılması düşünülmüştür. Bu çalışmada veriler korelasyonel ve nedensel araştırma yöntemine dayalı olarak toplanmıştır. Çalışmanın katılımcıları 2021-2022 eğitim öğretim yılının güz yarıyılında bir devlet üniversitesinde öğrenim gören Fen Bilgisi Öğretmen adaylarıdır. Çalışmada çoktan seçmeli kimya testini alan öğrencilerle İAKTT (İki Aşamalı Kimya Tanı Testi)’yi alan öğrencilerin NeuroSky Beyin sensörüne dayalı olarak ürettiği dikkat sinyallerine ilişkin tepe değerleri ile test yanıtlama süreleri incelenmiştir. Öğrencilerin yanıtlama süresi ve Neurosky dikkat düzeyi arasındaki karşılaştırmalarda parametrik olmayan analiz yöntemlerinden Mann-Whitney U ve Wilcoxon Testinden yararlanılmıştır. ÇSKTT (Çoktan Seçmeli Kimya Tanı Testi)’ye katılan öğrencilerin yanıtlama süreleri ile İAKTT’ye katılan öğrencilerin yanıtlama süreleri arasında anlamlı bir farklılık görülmemiştir. İAKTT’ye katılan öğrencilerin birinci ve ikinci aşamalarının yanıtlama süreleri karşılaştırıldığında anlamlı bir farklılık belirlenmiş; öğrencilerin ikinci aşamayı daha kısa sürede tamamladığı anlaşılmıştır. İAKTT’nin ikinci aşamasında yanıtlama süreleri ile dikkat frekansları arasında orta düzeyde bir ilişkinin olduğu bulunmuştur.

Anahtar Terimler: Bilgisayar destekli sınav, dikkat, yanıtlama süresi, çoktan seçmeli test, iki aşamalı test.

Abstract

In Turkey, there are many measurement and evaluation tools used for different purposes. One of the most commonly used is multiple-choice tests. In the 1980s, two-tier tests were developed to use the positive aspects of multiple-choice tests and to

minimize the negative aspects. The purpose of the tests is to determine how much the student has learned. Learning is an internal process that includes sensory awareness, attention, recognition, transformation, acquisition, and processing of information. In order for the information to be processed, the process of obtaining the information starts with attention. Attention is a clear and vivid position of the mind toward an object or set of thoughts. Focus and concentration are the foundation of attention. In tests, besides attention, response time is also important. Being fast has been the first step in the struggle for survival since primitive times. It is necessary to be effective as well as being fast. Therefore, the duration of the work is as important as the level of attention. The aim of this study was to analyze the students' attention levels and response times in the multiple-choice chemistry test and the two-tier chemistry diagnostic test (TTCDDT) performed in a computer environment. In this context, it was considered to use the participants' response times and attention data measured with the NeuroSky brain sensor during the multiple-choice and tiered diagnostic tests. In this study, data were collected together based on correlational and causal research methods. The participants were the pre-service science teachers studying at a state university in the fall semester of the 2021-2022 academic year. In this context, the peaks related to the attention signals produced by the NeuroSky Brain sensor and the test response times of the students who took the multiple-choice chemistry test and the students who took the TTCDDT (Two-Tier Chemistry Diagnostic Test) were examined. Mann-Whitney U and Wilcoxon Tests, which are non-parametric analysis methods, were used in comparisons between students' response time and Neurosky attention level. There was no significant difference between the response times of the students who participated in the Multiple-Choice Chemistry Diagnostic Test and the response times of the students who participated in the TTCDDT. When the response times of the students participating in TTCDDT were compared for the first and second tiers, a significant difference was determined; It was understood that the students completed the second tier in less time. It was found that there was a moderate relationship between response times and attention frequencies in the second tier of TTCDDT.

Keywords: Computer-aided exam, attention, response time, multiple-choice test, a two-tier test.

GİRİŞ

Geçmişten günümüze insanlar; joystickler, fareler, klavyeler, mikrofonlar ve dokunmatik yüzeyler aracılığıyla bilgisayarlarla iletişim kurmanın yollarını aramışlardır. Fakat tüm bu arayüzler, kas sisteminin hareketlerine dayalı olarak çalışmaktadır. Oysaki günümüzde bu durum biraz değişime uğramıştır. Günümüz teknolojisiyle geliştirilen arayüzler ve insan vücudundaki çok düşük enerjili sinyaller artık ölçülebilmekte ve makinalar kontrol edilebilmektedir. Örneğin, çeşitli sensörler kullanılarak beyinden ölçülen sinyaller ve arayüzler vasıtasıyla dönüştürülen dijital veriler bilgisayarlara gönderilebilmekte ve belirli algoritmalarla nesnelere kontrol edilebilmektedir. Beyin dalgaları gibi elde edilen bazı veriler, EEG (Elektroensefalografi) cihazlarının gelişmesiyle nesnelere kontrolünden oyun sektörüne, birçok makinenin kontrolü için kaynak oluşturmaktadır (Demirci, 2011; Soraghan, Matthews, Kelly, Ward, Markham, Pearlmuter & O'Neill, 2006). Beyin, barındırdığı yaklaşık 10 trilyon sinir hücresi (nöron) ile tüm vücut işlevlerini yöneten ve aralarında işbirliği sağlayan kontrol merkezi olup, sinir hücreleri arası iletişimi temel olarak elektrik sinyalleri ve sinir hücreleri arasındaki bağlantıların (synapse) salgıladığı bazı kimyasal maddelerle (neurotransmitter) sağlanmaktadır. Beynin nasıl çalıştığını ortaya çıkarmak amacıyla geçmişten günümüze çeşitli çalışmalar yapılmıştır. Nöroloji ve psikoloji alanlarındaki gelişmelere paralel olarak teknolojiye de kaydedilen gelişmeler beyin işlevlerini ve çeşitli beyin hastalıklarını daha iyi anlayabilmemize olanak tanımıştır. Aynı zamanda, beyin dalgalarını ölçmek için Elektro Beyin Grafisi (EEG) ve beyin dalgalarının yapısını görmek için ise Tomografi ve Manyetik Rezonans (MRI) gibi yeni görüntüleme teknikleri ve cihazları geliştirilmiştir (İşcan, 2009; Sevinç, 2006).

Eğitim-öğretim süreci zamanın değişen şartlarına göre kendini yenilemektedir. Öğretmen ve öğrenciye biçilen roller, yüklenen görevler ve beklentiler değişmektedir. Öğretmenlerin kullandığı ölçme araçları da günümüz şartlarına ve eğitim anlayışına göre değişiklik göstermektedir. Bilgisayarın eğitim alanında yaygın olarak kullanılmasıyla sınavlar online olarak yapılabilen, kısa zamanda sonuçlar açıklanabilmekte ve birçok konuda analiz yapılarak öğrenci ve öğretmene dönüt olarak veriler sunulabilmektedir. Dikkat, bütün sınavlarda sonuca tesir etmekte ve kişisel özelliklere bağlı olarak değişiklikler gösterebilmektedir. Öğrencinin dikkati ve yanıtlama süresi, ölçme araçlarından alınan sonuçları etkileyen önemli etmenlerdir. Bu etmenlerden aldığımız veriler büyük önem arz etmektedir. Vücudumuzda gerçekleşen reaksiyonlardan elde edeceğimiz veriler bize önemli bir kaynak teşkil edecektir. NeroSky Beyin Sensörü cihazıyla yapılan ölçümlerin bize daha gerçekçi ve daha az manipüle edilmiş veri sunacağı düşünülmektedir.

Ülkemizde farklı amaçlarla kullanılan birçok test yer almaktadır. En yaygın olarak kullanılanlardan birisi de çoktan seçmeli testlerdir. Eğitim sistemimizde sıklıkla kullanılan çoktan seçmeli testlerin özelliği, sorulan bir sorunun cevabının sunulan seçenekler arasından bulunması istenir. Sorunun

zorluk düzeyi seçeneklerdeki çeldiricilerin güçlü olması ile paralellik göstermektedir. Hazırlanması zor fakat değerlendirme açısından kolaylık sağlayan bu testlerin yetersiz kaldığı bazı kısımlar da vardır. Özellikle kavramlarla ilgili yapılan araştırmalarda bu dezavantaj kendisini göstermektedir. Mülakatlar, açık uçlu sorular, kompozisyonlar neticesinde ortaya çıkarılan öğrencilerin sahip oldukları kavram yanlışları, çeldiriciler olarak seçeneklere yerleştirilmektedir (Palmer, 1998; Tamir, 1971). Tüm bu çeldiricilere rağmen öğrenciler, konuyu bilmedikleri halde şans başarısıyla (faktörüyle) doğru seçeneği işaretleme olasılıkları bulunmaktadır. Bu durumda bile öğrenciler, sınavda doğru cevabı bilmedikleri halde, tesadüfen doğru seçeneği işaretleyebilirler. Bu tür testlerde, öğrencilerin işaretledikleri seçeneği seçme gerekçeleri ve bunun altında yatan sebeplerle ilgili araştırmacının yorum yapabilmesi veya yeterli bilgiye sahip olması çok da mümkün değildir. Verilen belirli kalıplar içerisinde, öğrencinin fikirlerini belirlemede bu testler yetersiz kalmaktadır (Mintzes, Wandersee & Novak, 2001). 1980'li yıllarda çoktan seçmeli testlerin, olumlu yönlerini kullanıp olumsuz yönlerini en aza indirmek için iki aşamalı testler geliştirilmiştir. Süreç içerisinde birçok araştırmacı tarafından fen bilimlerinin farklı alanlarında yaygın olarak kullanılmaktadır (Tan, Goh, Chia & Treagust, 2002). İki aşamalı testler iki kısımdan oluşmaktadır. İlk kısım çoktan seçmeli testlerle aynıdır. İkinci kısımda ise öğrencinin ilk aşamada işaretlediği seçeneğin gerekçesini belirtmesi gerekmektedir. Testin ikinci aşamasında, literatür incelemesi veya mülakatlardan elde edilen bulgulara dayalı olarak belirlenen kavram yanlışları, çoktan seçmeli ya da açık uçlu bir format ile sunulabilir. Açık uçlu bir yapıda hazırlanan testler, öğrencilerin muhakeme yeteneğini ölçmelerine ve daha önce belirlenen kavram yanlışlarından farklı alternatif kavram yanlışlarının olup olmadığını tespit etmemize olanak sağlar (Mann & Treagust, 1998).

Çok fazla uyarının bulunduğu ortamlarda duyu sistemleri, bilgiyi işleyebildiği ölçüde etkin çalışır. Bu kapasitenin üzerinde bilgi yüklendiği zamanlarda ise başarısız olmaktadır (Solso, Maclin & Maclin, 2009). Öğrenme içsel bir süreç olup duysal farkındalık, dikkat, tanıma, dönüştürme, bilginin alınmasını ve işlenmesini içermektedir. Bilginin işlenebilmesi için bilginin alınma süreci dikkatle başlar (Atkinson & Shiffrin, 1968). Dikkat, bir nesneye veya düşünce dizisine yönelik olarak zihnin net ve canlı bir pozisyon almasıdır. Dikkatin temelinde odaklanma ve konsantrasyon vardır. Deneysel psikoloji tarihinde 1890 yılında "Principles of Psychology" adlı kitabında ilk defa William James tarafından ele alınan dikkat, "herkesin sezgisel olarak ne olduğunu bildiği bir olgu. Aynı anda birkaç olası nesne veya düşünce dizisinden birinin zihnin açık ve canlı bir biçimde ele geçirilmesidir." şeklinde tanımlamıştır. Günümüze gelene kadar birçok araştırmacının konusu olmuş ve çok farklı şekillerde tanımlama yoluna gidilmiş ve çeşitli metaforlarla da tanımlanmaya çalışılmıştır. Seçici filtre (selective filter) metaforuyla tanımlayan Broadbent (1958), seçme işlevine; darboğaz (bottleneck) metaforuyla Pashler (1984), sınırlı kapasiteye; spot ışığı (spot light) metaforuyla Posner ve ark., (1984) seçilecek bilgilere tutulan bir ışığa; geçiş kapısı (gate) metaforuyla Hunt ve Ellis (2004), bilginin bilince ulaşmak için aldığı yoldaki bir kapıya; modülatör (modulator) metaforuyla Andrewes (2009), algı ve bellek gibi bilişsel süreçler üzerindeki düzenleyici etkisine vurgu yaparak tanımlanmaya çalışmışlardır (Broadbent, 1958; Pashler, 1984; Posner, Walker, Friedrich & Rafal, 1984; Hunt & Ellis, 2004; Andrewes, 2009).

İlkel çağlardan günümüze hızlı olmak, hayatta kalma mücadelesinin ilk adımı olmuştur. Av olmaktan kurtulmak için hızlı olmak zorunda olan insanlığı, günümüzde makinalara karşı yarışta hızlı olma yarışının içerisine girmiştir. Zamanı verimli bir şekilde yönetmek önem kazanmıştır. Hızlı olmak mı güçlü olmak mı? Hızlı olmak kadar etkili olmak da önemlidir. Hız konusunda günümüzün makinalarıyla yarışmak çok da mümkün değildir. İnsan, kapasitesinin sınırlarını zorlayarak etkili sonuçlar elde edebilir. Zeki olmanın bir göstergesi olarak kabul edilen pratiklik, motor becerilerde değil bilişsel becerilerde kendini göstermektedir. Bunun yanında, çok duyduğumuz bir hikâye olan kaplumbağa ile tavşanın yarışında, kaplumbağanın istikrarı ve çabası tavşanın hızını yenmiştir. Toplumda hızlı olanın daha zeki olduğu algısı yaygın olarak kabul edilmektedir. Öğretmenler ve aileler, çabuk öğrenen çocukları daha zeki olarak görmektedirler. Öğrenmede beklenen hız sınavlarda da beklenmektedir. Bazı araştırmalar zekâyı açıklamak için işleyen bellek ve hızın tek başına yeterli olabileceğini öne sürmektedir. Bazı araştırmalar ise tek bir yetenek puanıyla beyin fonksiyonlarının açıklanamayacak kadar karmaşık bir yapıda olduğunu öne sürmektedirler (Stankov & Roberts, 1997). Hız ile zekâ arasında bağ kurulmuş ve bunun sonucunda bazı görevlerin belli sürelerde yapılabilmesi

öngörülmüştür. Öğrenciler girdikleri bütün sınavlarda, tanımlanan süre içerisinde süreci tamamlamak zorundadırlar. Testte doğru cevapların bulunmasından ziyade, öğrencinin verilen süre içinde doğru cevabı verebilmesi beklenmektedir. Uygulanan testin türüne, soruların zorluk düzeyine, kişilik özelliklerine gibi birçok hususta öğrencilerin yanıtlama süreleri değişebilmektedir. Sınavlarda soruların algılanması, başarıma düzeyi ve yanıtlama süresinin etkilerinin belirlenmesi gerekmektedir.

Araştırmanın Amacı

Günümüzde bilgisayar teknolojileriyle eğitimde, tıpta, mühendislikte ve hatta günlük yaşamda hayatın her anında karşılaşmak mümkündür. 70'li yıllarda Amerikan askeri çalışanları ve klinik psikologları bilgisayar destekli testleri geliştirerek eğitim alanında bu teknolojileri kullanmaya başlamıştır (Russell, 2003). Iowa State Üniversitesinde Test Merkezinde 2003 güz yılından 2015 güz yılına kadar olan 12 yıllık süreçte çevrimiçi olarak yapılan sınavların sayısı 38,6 kat artış göstermiştir (Prisacari & Danielson, 2017). İnternet üzerinden ya da merkezi ağlar üzerinden yapılan çevrimiçi sınavlara gün geçtikçe açıkça talebin arttığı söylenebilir. Bilgisayar destekli testlere geçişte kullanıcıların bilgisayarlara alışık olması, ekran karşısında fazla zaman geçirmesi, günlük yaşamla ilgili ihtiyaçları ekranlar üzerinden çözüm üretme çabaları en önemli sebeplerdir. Öğrenciler bilgisayar destekli testlere günün herhangi bir zamanında katılabilir ve sonuçlarını çok kısa sürede öğrenebilir. Ayrıca bireysel yetenekleriyle ilgili kişisel raporlara kısa sürede ulaşabilir. Bilgisayar destekli testler eğitim sürecinde tekrar ve esnek kullanım kolaylığı sağlayabilir. Buna ek olarak eğitim ve öğretim planlayıcılarına kısa sürede geri dönütler sunabilir (Luecht, 2005). Bilgisayar destekli testlerin böylesi yararlarının yanında bazı sınırlılıklarının da bulunması olasıdır. Test katılımcıları, bilgisayar başında bazı notları almada zorlanabilecekleri gibi uzun okuma metinleriyle karşılaştıklarında göz yorgunluğu yaşayabilirler. Bilgisayar destekli testler öğrencilere teknoloji okuryazarlığı bilme zorunluluğundan kaynaklı ek bilişsel yükler getirebilir (Ommerborn, 2001). Test Katılımcılarının bilgisayar destekli testlerdeki test maddelerine doğru yanıtlama performanslarıyla birlikte yanıtlama süresi gibi otomatik kaydedilen kişisel performanslarla birlikte bilgisayarla bağlantılı dikkat vb. duyuşsal parametreler eğitim içeriklerini ve süreçlerini değerlendirme amacıyla kullanılabilir. Kavramsal anlama ve diğer tanısal ölçme araçları geliştirilirken test katılımcılarının yanıt süreleri madde zorluğunun analiz edilmesinde ilginç bir alternatif veya ek bir yaklaşım olabilir (Direnga, Timmermann, Presentati, & Brose, 2015). Bu kompleks yapılar, öğrencilerin test maddelerine doğru yanıt vermenin yanında hızlı ya da yavaş yanıt verme, dikkatli ya da stres altında yanıt verme gibi parametrelerle birlikte araştırmacılara tahmin edilebilen tümleşik ortak kişisel parametreleri açıklamada fırsatlar sunabilir. Örneğin yanıtlama süresiyle motivasyon, dikkat ve akademik başarı gibi bilişsel ve duyuşsal yetiler arasındaki doğrusal ilişkiler bu kompleks yapılardan toplanan verilerle günümüzde kolaylıkla açıklanabilmektedir (Wainer, Dorans, Flaugher, Green & Mitlevy, 2000). Yanıtlama süresi, test madde zorluğu ve test maddesine ilişkin kişisel yetenekler kestirilebilir (Yang, O'neill & Kramer, 2002; Mayerl, 2013). Bilgisayarlı testlerle test katılımcılarının dikkatleri, tepki süreleri gibi performansları ölçülebilmekte ve testte sergiledikleri performans ile matematiksel ilişkilendirmeler yapılarak örtük olarak katılımcının yetenek ve başarı kestirimine gidilmektedir. Ayrıca farklı matematiksel fonksiyonlar içinde test yanıt süresi, madde yanıt süresi, madde doğruluğu, madde yanlışlığı, madde güçlüğü vb. parametreler model kurma süreçlerinde kullanılmaktadır (Chan, Lu & Tsai, 2014). Öğrencilerin test yanıtlama performanslarını etkileyen stres, kaygı, motivasyon, dikkat gibi birçok duyuşsal durum söz konusudur. Stres ve kaygının öğrencilerin test performansına ve yanıtlama sürelerine olumsuz etkileri bilinmektedir. Test katılımcılarının düşük yeteneğe, düşük motivasyona, dikkat eksikliğine sahip olması test performanslarını olumsuz etkileyebilir ve testi kısa sürede tamamlamasına sebep olabilir (Weeks, Von Davier & Yamamoto, 2016). Düşük motivasyona sahip öğrenciler stres ve kaygı altında sınavlarda tahmin davranışında bulunabilmekte ve testi daha kısa sürede tamamlayabilmektedir (Gass & Curiel, 2011; Trifoni & Shahini, 2011; Wise & Kong, 2005). Yüksek motivasyona sahip olan öğrencilerin yanıtlama performanslarının iyi olduğu ve testi ortalamanın üstünde makul sürede tamamladığı görülmektedir (Fernández-Castillo & Caurcel, 2015; Putwain & Remedios, 2014, Singh & Bhargava, 1985). Dikkat düzeyi yüksek olan öğrencilerin akademik performanslarının da iyi olduğu düşünülmektedir (Pacheco-Unguetti, Acosta, Lupianez, Roman & Derakshan, 2012; Shah & Saleem, 2015; Hotulainen, Thuneberg, Hautamäki, & Vainikainen, 2014). Dikkatin ders dinleme ve derse

katılımla ilgili olumlu katkıları gözlenirken test performanslarına ve yanıtlama sürelerine ilişkin katkıları bilinmemektedir (Yang vd., 2002; Mayerl, 2013). Ayrıca bilgisayar destekli testlere katılan öğrencilerin dikkatini ve akademik başarısını etkileyen birçok durumdan bahsedilebilir. Örneğin bilgisayar ekranının boyutu, ekran ışığı, ortamın havası ve sesi, ekranda yer alan test biçimleri öğrencilerin dikkatini ve akademik başarısını etkilemesi söz konusudur. Öğrencilerin bilgisayar kullanım becerileri, kullanılan cihazların uyumsuzluğu, internet hızı, laboratuvar koşulları bu sonuçları etkileyebileceği öngörüsü ileri sürülmektedir. Bilgisayar testlerinde kullanılan test türleri elbetteki kâğıt testlerine yönelik zayıf ya da güçlü yanlar bakımından farklı sonuçları ortaya çıkarabilmektedir (Bodman & Robinson, 2004). Öğrencilerin klavye ve mouse kullanma becerileri, test maddelerinin ekrana farklı biçimlerde yansması, ekrana farklı sürelerde gelmesi, öğrencilerin test maddelerine dönememesi ya da öğrencilerin test maddelerini tekrar gözden geçirme süreçlerinin farklılaşması bilgisayar testlerine dezavantaj durumlar oluşturmaktadır (Lee, Moreno, & Sympson, 1986; Russell, 1999; Clariana & Wallace, 2002; Spray, Ackerman, Reckase, & Carlson, 1989; Mason, Patry & Bernstein, 2001; Bodmann & Robinson, 2004). Test maddesinin zorluğu ve uzunluğu, testin bilişsel yapısı ve karmaşık içeriğe sahip olması öğrencinin test performansını etkileyerek onları tahmin yapmasına zorlayabilir. Ayrıca testin başında gösterilen performans ile sonunda gösterilen performans farklılık göstererek test sonunda başarı azalır ve öğrenciler test sonunda tahminde bulunabilir. Testin başında aşına olunan test maddeleri ve kelimeler test sonunda yer alan test maddelerini olumsuz etkileyebilir (Schnipke & Scrams, 1999; Setzer, Wise, Heuvel & Ling, 2013). Soru kökenlerinin yapısı da test performansını ve tahmini de artırabilir. Wise ve Ma (2012), ikinci ve üçüncü yanıt seçenek pozisyonlarında doğru cevapları olan orantısız olarak daha fazla sayıda madde bulunması nedeniyle, hızlı tahminlerin, şans eseri beklenenin biraz üzerinde bir oranda doğru olduğunu bulmuşlardır. DeMars (2000), yapılandırılmış cevap maddelerinin düşük bahisli testler için çoktan seçmeli maddelere göre daha kötü olduğunu tespit etmiştir. Streiner'a (2003) göre test maddesinin performansı ve güvenilirliğinin test maddesinin cevabının başka bir soru maddesinin cevabına bağlı olarak ilişkilendirildiğinde ve test maddesinin kapsamı farklı boyutlarda genişletildiği durumlarda düştüğünü ifade etmiştir. Bayazıt (2013) yaptığı çalışmada çevrimiçi sınavda katılımcıların göz izleme programıyla göz hareketlerini izlemiş ve soru tarzı, biçim ve niteliklerinin cevaplama sürelerini ve yanıtlamada doğruluk oranını etkilediği kanıtlamıştır. Bu durumda iki aşamalı ve üç aşamalı testlerde bu durumların bulunması söz konusudur. İki ya da üç aşamalı kavramsal tanı testleri ve güç testleri birbirini izleyen cevaplar olması ve aşama içerisinde bir aşamanın cevabı diğer aşamayı ilgilendirmesi bu testlerin performansının ve güvenilirliğinin düşmesine neden olabilir. Bireysel farklılıklar ve özellikler madde yanıt süresi ve doğruluğunu etkileyebilir (Semmes, Davison, & Close, 2011). Yanıtlama süresi, test katılımcısının yararlandığı bilgisayar ve iletişim teknolojilerinden, ekran rengi, büyüklüğü, çözünürlüğü, testin hazırlandığı bilgisayara yazılımı teste verilen tepki süresini ayrıca test katılımcısının teknoloji okuryazarlığı ve bilgisi teste yönelik tutumunu ve dolaylı olarak da test süresini de etkileyebilir ve değiştirebilir. (Nikou & Economides, 2013). Testlerdeki yanıtlama süresine cinsiyet faktörünün etkili olmadığı ortaya çıkmıştır (Truell, 2005; Truell, Zhao, & Alexander, 2005). Yanıtlama süresi test katılımcısının katıldığı test etkilenebilir. Testin bilgisayar ortamında ya da kâğıt kalem testinde olması gibi durumlarda teste verilen tepki süresini etkileyebilir. Örneğin Bernt ve Bugbee (1988), Bayazıt (2007) ve Delen (2015) yaptığı çalışmada öğrencilerin bilgisayar ortamında oldukları sınavı kâğıt kalem testlerine göre daha uzun sürede tamamladıkları ortaya çıkmıştır. Bayazıt (2007) her iki test türünde de test katılımcılarının yaklaşık olarak aynı başarıyı gösterdiklerini ifade etmiştir. Test maddelerine verilen doğru yanıtlar (zorluk indeksi) ile test maddesine harcanan süre arasında korelasyon kurulduğunda ters bir ilişki görülmüştür. Buna benzer şekilde Swanson ve diğ. (2001) Amerika Birleşik Devletleri Tıbbi Lisanslama Sınavında “madde zorluğu ve ortalama yanıt süresi arasında nispeten güçlü bir ilişki” bulmuştur. Beklendiği gibi, tüm korelasyon katsayıları ve eğimleri negatiftir, bu da daha zor maddelerin, yani daha küçük zorluk indeksi olan maddelerin daha uzun yanıt süreleri gerektirdiğini göstermektedir. Maddenin zorluğu ve cevap süreleri gerçekten birbiriyle ilişkili olduğundan, yanıt sürelerinden ötürü madde zorluğunu tahmin etmek için ters model incelenebilir (Direnga, Timmermann, Presentati, & Brose, 2015). İlgili literatür ışığında testlerin yanıtlanma süresiyle ve performanslarıyla ilgili bulgular bilgisayar ortamında elde edilmiş, diğer dikkat, motivasyon gibi duyuşsal bulgular kâğıt testler üzerinden katılımcıların katılım düzeyi beyanlarına göre ulaşılmıştır. Bu çalışmada ise yanıtlama performansları

ve süreleri ilgili verilere benzer biçimde bilgisayar ortamında ulaşılmaya çalışılırken literatürden farklı olarak katılımcıların duyuşsal verilerden dikkat verileri bilgisayar ortamında NeuroSky beyin sensörü verisiyle eş zamanlı ulaşılmaya çalışılmıştır. Bu kapsamda günümüzde bilgisayar destekli testlerin tercih edildiđi sınav ortamlarında öğrencilerin dikkatlerinin, yanıtlama performans ve yanıtlama sürelerinin eş zamanlı ölçülmesi ve değerlendirilmesi eğitimciler ve araştırmacılara önemli bilgiler verebilir. Bu bilgiler ışığında çalışmada bilgisayar destekli yapılan İAKTT ve ÇSKTT’de öğrencilerin test maddelerini yanıtlama sürelerinin ve test esnasındaki dikkat düzeylerinin NeuroSky cihazıyla birlikte incelenmesi amaçlanmıştır.

Problem Cümlesi

Bu çalışmanın esas problemi “NeroSky Beyin Sensörü cihazıyla yapılan ölçümlere göre Diagnostik fen tanı testi ve çoktan seçmeli testini alan öğrencilerin dikkat düzeyi ve yanıtlama süreleri nasıl değişmektedir?” biçiminde ifade edilmiştir. Bu ana problem aşğıdaki alt problemlere ayrılarak çözülmeye çalışılmıştır. Bu durumda alt problemler;

- a. NeroSky Beyin Sensörü cihazıyla yapılan ölçümlere göre Diagnostik fen tanı testini alan öğrencilerin yanıtlama süreleri ile çoktan seçmeli teste alan öğrencilerin yanıtlama süreleri arasında nasıl bir değişim vardır?
- b. NeroSky Beyin Sensörü cihazıyla yapılan ölçümlere göre Diagnostik fen tanı testini alan öğrencilerin dikkat düzeyleri ile çoktan seçmeli teste alan öğrencilerin dikkat düzeyleri arasında nasıl bir değişim vardır?
- c. NeroSky Beyin Sensörü cihazıyla yapılan ölçümlere göre Diagnostik fen tanı testini alan öğrencilerin dikkat düzeyleri ile yanıtlama süreleri arasında nasıl bir ilişki vardır?

YÖNTEM

Araştırmanın Modeli

Fen bilgisi öğretmen adaylarının genel kimya dersi kapsamında dikkat düzeyleri ile yanıtlama süreleri arasındaki birbirine etki düzeylerini ortaya koymak için nicel araştırma yöntemlerinden korelasyonel araştırma yöntemi kullanılmıştır. Korelasyonel araştırma yönteminde iki veya daha fazla değişken arasındaki ilişki incelenmektedir (Fraenkel & Wallen, 2000). Nedensel araştırma yöntemlerinde, ortaya çıkmış var olan bir durum ya da olayın nedenlerini, bu nedenleri etkileyen değişkenleri veya bu etkinin sonuçlarını belirlemeye yönelik bir araştırma türüdür. Bu tür araştırmalarda bir durumun oluşmasında nelerin etkili olduđu bulunmaya çalışılır. Araştırma deseninde sadece bir tek grup varsa nedensel araştırma yöntemi kullanılır (Büyüköztürk, Çakmak, Akgün, Karadeniz, & Demirel, 2008).

Bu çalışmada korelasyonel ve nedensel araştırma yöntemi birlikte kullanılmıştır. Bu kapsamda çoktan seçmeli kimya testini alan öğrencilerle İAKTT’yi alan öğrencilerin NeuroSky Beyin sensörüne dayalı olarak ürettiđi dikkat sinyalleri ile test yanıtlama performans verilerinin ilişkileri incelenmiştir.

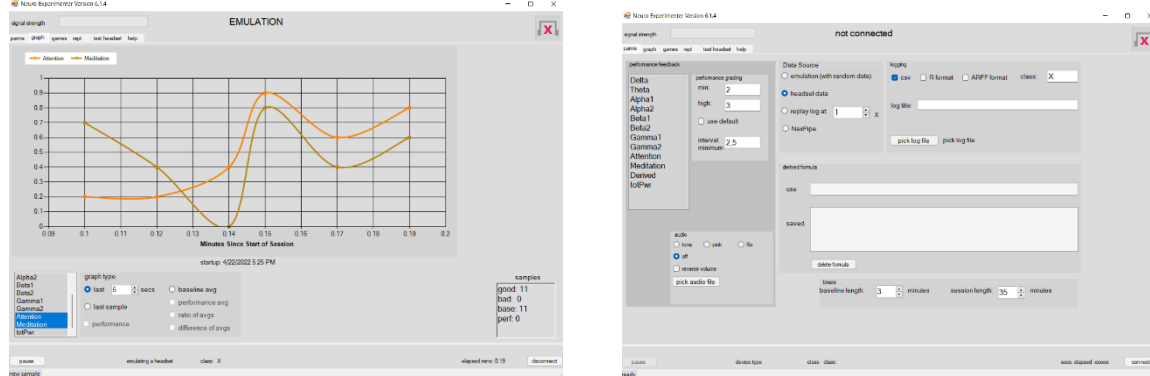
Araştırmanın Evreni ve Örneklemi

Bu çalışma 2021-2022 eğitim öğretim yılında Dokuz Eylül Üniversitesi’nde öğrenim gören 1. sınıf Fen Bilgisi Öğretmen adaylarıyla çalışılmıştır. Teste katılan 66 öğrenciden 47 kişi kız, 19 kişi ise erkek öğrencidir. 35 kişi A şubesinde, 31 kişi B şubesinde yer almaktadır. Katılımcılar Akdeniz bölgesinden (n:10), Dođu Anadolu (n:1), Güneydođu Anadolu (n:3), İç Anadolu (n:8), Karadeniz (n:6), Marmara (n:6) ve Ege (n:32) bölgelerinde ikamet etmektedir. Katılımcıların 33’ü İAKTT’ye, diđer 33’ü ÇSKTT’ye katılmış ancak ÇSKTT’ye katılan 5 katılımcının verisi cihazdan kaynaklı kayıplardan dolayı analizlerin dışında bırakılmıştır. İki grupta toplam 37 kişi, Kimya dersinde 100’lük puan sisteminde 45 ortalamının altında kalırken, 29 kişi ise 45 ortalamının üstünde başarılı olmuştur.

Veri Toplama Aracı

Araştırmada kullanılan iki aşamalı kimya testi, Mutlu ve Şeşen (2016) tarafından geliştirilen ve Şeker (2017) tarafından yürütölen yüksek lisans tezinde konu kapsamı genişletilerek kullanılan bir ölçme aracıdır. Test konularının kapsamı termodinamik, kimyasal kinetik, kimyasal denge, asit-baz,

elektrokimya konularına ilişkin kavramlardan oluşmaktadır. Testin ilk aşaması, çoktan seçmeli test olarak çalışmada kullanılmıştır. Test 44 sorudan oluşan bir kimya testidir. Öğrencilerin dikkat düzeylerinin tespit edilmesi amacıyla NeuroSky cihazı kullanılmıştır. Bu kapsamda cihazın resmi internet adresinde yer alan ve ücretsiz olarak hizmete sunulmuş olan Neuroexperimenter programını indirilerek cihazlarla bağlantısı sağlanmıştır.

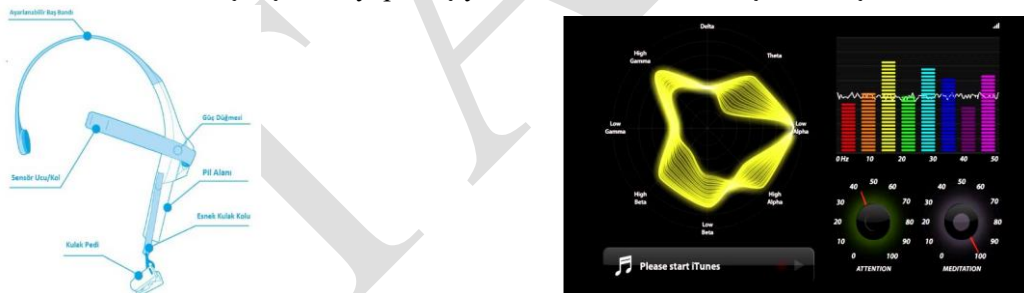


Şekil 1. Neuroexperimenter Programı Arayüzü

Bilgisayarlara indirilen programlarla cihaz bağlantısı kurulduktan sonra karışıklığa neden olmaması için etiketleme yapılarak sınıflama yoluna gidilmiştir.

NeuroSky

NeuroSky MindWave jel kullanmaya ihtiyaç duyulmadan elektrotları yerleştirmeyi sağlayan kuru aktif teknolojisine sahip, genel kullanıma uygun düşük maliyetli kablosuz EEG kulaklığıdır. İlk zamanlarda fiziksel açıdan yetersiz kalan kişilere hizmet amacıyla geliştirilen cihazın kullanım alanı son yıllarda büyük artış göstermiştir. Beyin dalgaları üzerine çalışma yapan araştırmacıların dikkatini çeken cihaz ile farklı alanlarda çalışmalar yapılmış yeni analiz metotları oluşturulmuştur.



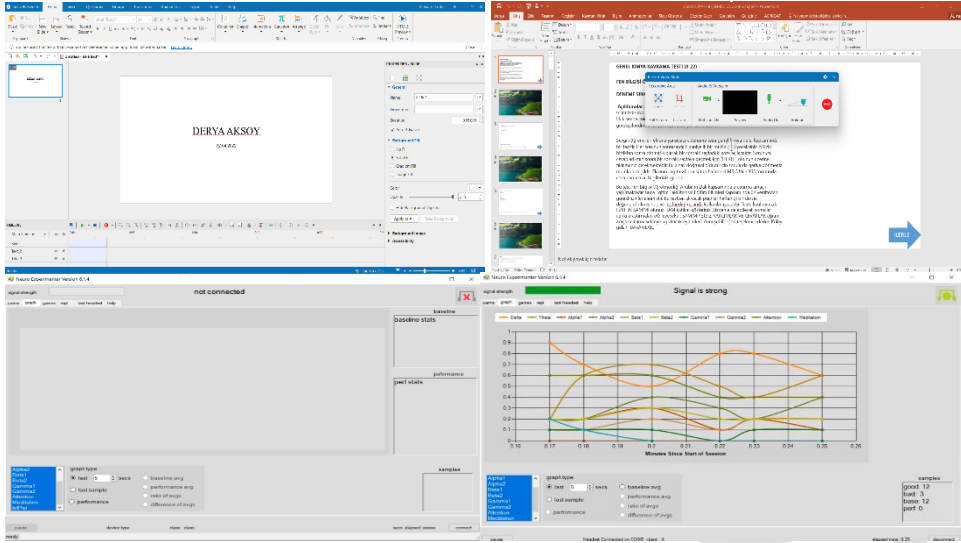
Şekil 2. EEG Başlığı Neurosky Arayüzü

İnel (2014) "Sosyal Bilgiler Öğretiminde Kullanılan Bilgisayar Temelli Materyallerin 6. Sınıf Öğrencilerinin Dikkat ve Motivasyon Düzeylerine Etkisi" adlı doktora tezinde NeuroSky cihazını kullanmıştır. Özkan (2017) "EMDR Cihazının Tasarımı ve Optimum Çalışma Parametrelerinin Sinyal İşleme Teknikleri ile Belirlenmesi" adlı yüksek lisans çalışmasında NeuroSky cihazını kullanmışlardır. Demirel, Kandemir ve Köse (2018) "Ekstraoküler Kasları Kullanarak EEG Cihazı ile Robot Kontrolü" adlı yaptıkları çalışmada NeuroSky beyin sensörü ile robot kontrol etmişlerdir. Kullanılan NeuroSky cihazından elde edilen veriler çalışmalarda farklı şekillerde analize tabi tutulmuştur.

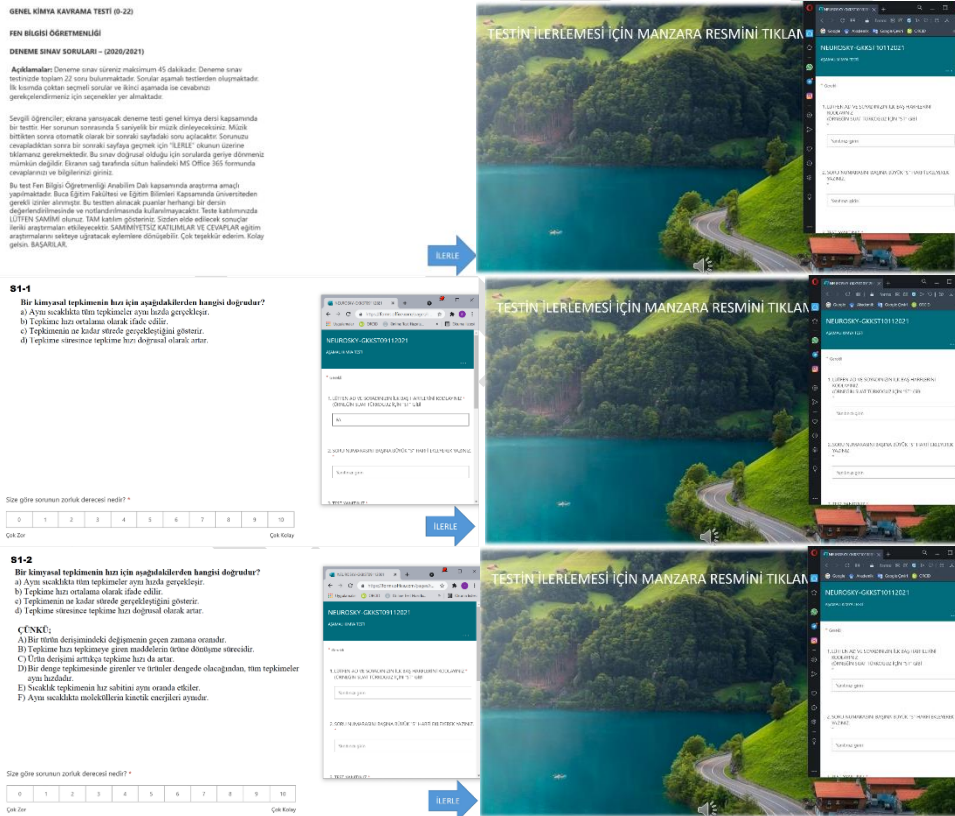
Verilerin Toplanması ve Analizi

Bu çalışmada 7 adet laptop ve 15 adet NeuroSky cihazı kullanılarak veri toplanmıştır. İki aşamalı kimya testi, powerpoint programı üzerinden uygulanmıştır. Her bir soru için bir slayt sayfası açılmış ve sorularda lineer bir ilerleme sağlanmıştır. Sorularda geriye dönme olmadığı gibi sorular arası geçişte 10 saniyelik manzara resmi eşliğinde müzik dinletisi yapılmıştır. Soruların cevaplarının kaydedilmesi için Office-365'te cevap anahtarı formu hazırlanmış ve her bir sorunun yanıtlanma süresi ve yanıtlama performansı bu cevap anahtarı aracılığıyla kaydedilmiştir. Powerpoint'teki sorular ve Office-365'teki cevap anahtarının ekranda aynı anda bulunması için DeskPins programı ile

yanıtlama formu bilgisayar ekranının sağ tarafında sabitlenmiştir. Veri kaybının önüne geçmek, katılımcıları teste ve test ortamına alıştırmak için kimya testini uygulanmadan önce 3 sorudan oluşan küçük bir deneme testi gerçekleştirilmiştir. Bu süreçte katılımcılara testin ilerleyişi ve yanıtlama süreci ile ilgili deneyim kazanmaları sağlanmıştır. Katılımcılar kendilerini hazır hissettiklerinde ve merakları giderildikten sonra kimya testine geçilmiştir.



Şekil 3. Ekran Kaydı Başlangıç Aşaması



Şekil 4. Aşamalı Kimya Testinin Ekran Görüntüleri

Veriler toplandıktan sonra tarihe göre, yapılan test türüne göre sınıflandırılarak harddisklere yedeklemesi yapılmıştır ve analiz aşamasına geçilmiştir. Araştırma verileri iki biçimde analiz edilmiştir.

Öğrencilerin yanıtlama süresi ve Neurosky dikkat düzeyinin maksimum düzeydeki frekansları bağımsız değişkenler (IAKTT'ye katılım-çoktan seçmeli teste katılım) arasında karşılaştırmalı ve ilişkisel analiz edilmiştir. Beyin dalgalarının sinyali, kişinin belirli özelliklerine ve özelliklerine sahiptir. Beyin dalgaları kişiler tarafından taklit edilemez veya yorumlanamaz, dolayısıyla benzerlik şansı yoktur. Kimlik tanıma, kişinin özelliklerini belirlemek için gereklidir. Bireyin beyninin alınan uyarılara verdiği sinyal örüntüsü ve tepki hızı farklıdır. Verilen uyarana karşı farklı insan beyin dalgası aktivitesi nedeniyle her yanıtlayan için farklı bir değer gösterecektir. Doğru uyarın kullanılarak doğruluk sonuçlarındaki farklılıklar, elde edilen doğruluğu önemli ölçüde etkiler. Cihazdan elde edilen verilerin tepe değerleri, sonuçları yorumlama açısından önemlidir (Yudhana, Muslim, Wati, Puspitasari, Azhari & Mardhia, 2020).

Browarska ve diğ. (2021) yaptığı çalışmada, uygulanan yöntem, analiz edilen EEG sinyallerinde bulunan karakteristik "zirvelere" dayalı derecelendirme kriterlerine dayalı olarak filtrelerin verimlilik tahminini yumuşatmak amacıyla uyguladılar. Görsel tahminin sinyal filtreleme kalitesi deneyiminin genellikle yetersiz olması nedeniyle, filtreleme verimliliği değerlendirme kriterlerini EEG sinyalinin işaretleme karakteristik "tepe noktalarına" dayalı olarak uygulamaya karar vermişlerdir. Öğrencilerin yanıtlama süresi ve Neurosky dikkat düzeyinin maksimum düzeydeki frekans verileri iki aşamalı kimya test grubuna katılanlar ile çoktan seçmeli test grubuna katılanlar arasındaki karşılaştırmalarda parametrik olmayan analiz yöntemlerinden Mann-Whitney U testinden yararlanılmıştır. Sebep olarak verilerin normal dağılım göstermemesinden ve yeterli örneklem sayısına ulaşamamasından kaynaklanmaktadır. Anlamlılık düzeyi 0,05 olarak belirlenmiştir. Benzer şekilde sadece İAKTT'ye katılan öğrencilerin yanıtlama süresi ve Neurosky dikkat düzeyinin maksimum düzeydeki frekans verileri İAKTT'nin I. aşaması ile II. aşaması arasındaki karşılaştırmalarda parametrik olmayan analiz yöntemlerinden Wilcoxon testinden yararlanılmıştır. Yapılan analizde normal dağılım göstermediği sonucuna ulaşılmıştır. Anlamlılık düzeyi 0,05 olarak belirlenmiştir. Öğrencilerin yanıtlama süresi ve Neurosky dikkat düzeyinin maksimum düzeydeki frekans verileri arasındaki korelasyon değerleri çoktan seçmeli testin, İAKTT'nin I. ve II. aşamasına göre ayrı ayrı incelenmiştir. Aşağıdaki düzeylere göre yorum yapılmıştır.

Tablo 1. Korelasyonel ilişki düzeyleri

İlişki Düzeyi	Puan
İlişki yok	0,00-0,200
Az düzeyde ilişki	0,201-0,400
Orta düzeyde ilişki	0,401-0,600
İyi ilişki	0,601-0,800
Çok iyi ilişki	0,801-1,000

BULGULAR

Bu bölümde bulgular, araştırmanın problem ve alt problemlerine ilişkin sorulara yanıt olacak şekilde sırasıyla verilmiştir. İlk olarak NeroSky Beyin Sensörü cihazıyla yapılan ölçümlere göre Diagnostik fen tanı testini alan öğrencilerin yanıtlama süreleri ile çoktan seçmeli teste katılan öğrencilerin yanıtlama süreleri arasındaki değişimler verilmiştir. İkinci olarak NeroSky Beyin Sensörü cihazıyla yapılan ölçümlere göre Diagnostik fen tanı testini alan öğrencilerin dikkat düzeyleri ile çoktan seçmeli teste alan öğrencilerin dikkat düzeyleri arasındaki değişimler verilmiştir. Son olarak NeroSky Beyin Sensörü cihazıyla yapılan ölçümlere göre Diagnostik fen tanı testini alan öğrencilerin dikkat düzeyleri ile yanıtlama süreleri arasındaki ilişkiler verilmiştir.

Çalışmada ilk olarak diagnostik fen tanı testini alan öğrencilerin yanıtlama süreleri ile çoktan seçmeli teste katılan öğrencilerin yanıtlama süreleri bağımsız örneklem için (ilişkisiz) t testi ve nonparametrik bir istatistik yöntem olan Mann-Whitney U testi ile karşılaştırılmıştır. Mann-Whitney U testi, iki ilişkisiz örneklemden elde edilen puanların birbirlerinden anlamlı bir şekilde farklılık gösterip göstermediğini test eder. Başka bir anlatımla, bu test iki ilişkisiz grubun ilgilenilen değişken bakımından evrende benzer dağılımlara sahip olup olmadığını test eder (Büyüköztürk, 2002). Tablo 2, 3 ve 4 bu karşılaştırmalara ilişkin bulguları göstermektedir.

Tablo 2. Yanıtlama sürelerinin ortalaması

Grup	ORTALAMA (sn)
İAKTT'nin I. aşamasının yanıtlama süresi (N:33)	54,02
ÇSKT'nin yanıtlama süresi (N:28)	52,43
İAKTT'nin II. aşamasının yanıtlama süresi (N:33)	49,90

Tablo 2’de, İki Aşamalı Kimya Tanı Testinin (İAKTT) I. ve II. aşamasının yanıtlama süreleri ile Çoktan Seçmeli Kimya Testinin (ÇSKTT) yanıtlama sürelerinin karşılaştırıldığı betimsel ortalama değerleri bulunmaktadır. Tablo 2’de bu betimsel ortalama değerlere göre İAKTT’nin I. aşamasının yanıtlama süresine ilişkin ortalaması 54,02 sn, II. aşamasının yanıtlama süresine ilişkin ortalaması 52,43 sn ve ÇSKTT’nin yanıtlama süresine ilişkin ortalaması 49,90 sn olarak hesaplanmıştır. Bu sonuçlara göre İAKTT’nin I ve II. aşamasının ortalama yanıtlama süreleri ile ÇSKTT’nin ortalama yanıtlama sürelerinin birbirine çok yakın olduğu söylenebilir.

İstatistiksel olarak anlamlı farklılıkları gözetlemek için ilk olarak İAKTT’nin I. aşamasının yanıtlama süresi ile ÇSKTT’nin yanıtlama süresi ikili bağımsız grup karşılaştırmalarında kullanılan Mann-Whitney U (MWU) testi ile karşılaştırılmıştır. İkinci olarak İAKTT’nin I. aşamasının yanıtlama süresi ile II. aşamasının yanıtlama süresi ikili bağımlı grup karşılaştırmalarında kullanılan Wilcoxon Testi (WT) ile analiz edilmiştir. Tablo 3’te, İki Aşamalı Kimya Tanı Testinin (İAKTT) I. aşamasının yanıtlama süreleri ile Çoktan Seçmeli Kimya Testinin (ÇSKTT) yanıtlama sürelerinin karşılaştırıldığı Mann-Whitney U testi analiz sonuçları bulunmaktadır.

Tablo 3. Yanıtlama sürelerinin karşılaştırılmasına ilişkin MWU test sonuçları

Grup	N	Ortalama Sırası	Sıra Toplamı	MWU	p
İAKTT'nin I. aşamasının yanıtlama süresi	33	7,13	57,00	21,000	0,699
ÇSKTT'nin yanıtlama süresi	28	8,00	48,00		

Tablo 3’te MWU analiz sonuçlarına göre İAKTT’nin I. aşamasının yanıtlama süresine ilişkin ortalama sırası 7,13 (Sıra Toplamı: 57,00) ve ÇSKTT’nin yanıtlama süresine ilişkin ortalama sırası 8,00 (Sıra Toplamı 48,00) olarak hesaplanmıştır. Bu sonuçlara göre İAKTT’nin I. aşamasının yanıtlama süresinin ortalama sırası ile ÇSKTT’nin yanıtlama süresinin ortalama sırası arasında anlamlı bir farklılığın olmadığı belirlenmiştir ($U=21,000$; $p>0,05$). Bu bulguya göre İAKTT’nin I. aşamasının yanıtlama süresi ile ÇSKTT’nin yanıtlama süresi arasında benzer olduğu sonucu çıkarılabilir. Tablo 4’te, İAKTT’nin I. aşamasının yanıtlama süreleri ile II. aşamasının yanıtlama sürelerinin karşılaştırıldığı Wilcoxon testi analiz sonuçları bulunmaktadır.

Tablo 4. İAKTT’nin I ve II. aşamasının yanıtlama sürelerinin Wilcoxon Test ile karşılaştırılması

	N	Ortalama Sırası	Sıra Toplamı	Wilcoxon Test	p
Negatif Sıralar	8 ^a	4,50	36,00	-2,521	0,012 Anlamlı
Pozitif Sıralar	0 ^b	0,00	0,00		
Eşlenik Sıralar	0 ^c				

^a. İAKTT’nin II. aşamasının yanıtlama süresi < İAKTT’nin I. aşamasının yanıtlama süresi;

^b. İAKTT’nin II. aşamasının yanıtlama süresi > İAKTT’nin I. aşamasının yanıtlama süresi;

^c. İAKTT’nin II. aşamasının yanıtlama süresi = İAKTT’nin I. aşamasının yanıtlama süresi

İAKTT’nin I. ve II. Aşamasının yanıtlama süreleri arasındaki korelasyon değeri $r=0,743$; $p=0,035$; Yüksek İlişki

Tablo 4’te Wilcoxon testi analiz sonuçlarına göre, negatif sıraların ortalama sırası 4,50 ve pozitif sıraların ortalama sırası 0,00 olarak hesaplanmıştır. Diğer bir ifadeyle İAKTT’nin I. aşamasının yanıtlama süresi İAKTT’nin II. aşamasının yanıtlama süresinden biraz daha fazla olduğudur. Bu sonuçlara göre İAKTT’nin I. aşamasının yanıtlama süresinin ortalama sırası ile II. aşamasının yanıtlama süresinin ortalama sırası arasında anlamlı bir farklılığın olduğu saptanmıştır ($WT=-2,521$; $p<0,05$). Bu bulguya göre İAKTT’nin I. aşamasının yanıtlama süresinin İAKTT’nin II. aşamasının yanıtlama süresinden biraz daha fazla olduğu söylenebilir.

Yanıtlama süresine ilişkin genel olarak elde edilen bulgularda çoktan seçmeli teste katılan öğrencilerin yanıtlama süreleri ile İAKTT’ye katılan öğrencilerin yanıtlama süreleri arasında anlamlı bir farklılık görülmemiştir. İAKTT’ye katılan öğrencilerin birinci ve ikinci aşamaları yanıtlama süreleri karşılaştırıldığında anlamlı bir farklılık belirlenmiş; öğrencilerin ikinci aşamayı daha kısa

sürede tamamladığı anlaşılmıştır. İAKTT'nin II. aşamasında yanıtlanma süreleri ile dikkat frekansları arasında orta düzeyde bir ilişkinin olduğu bulunmuştur.

Çalışmada ikinci olarak diagnostik fen tanı testini alan öğrencilerin Neurosky dikkat dalga tepe değer sayıları ile çoktan seçmeli teste katılan öğrencilerin Neurosky dikkat dalga tepe değer sayıları bağımsız örneklem için (ilişkisiz) t testi ve nonparametrik bir istatistik yöntem olan Mann-Whitney U testi ile karşılaştırılmıştır. Tablo 5, 6 ve 7 bu karşılaştırmalara ilişkin bulguları göstermektedir.

Tablo 5. NeuroSky dikkat dalga tepe değer sayılarının ortalamaları

Grup	ORTALAMA (Tepe değer sayısı)
İAKTT'nin I. aşamasının NeuroSky dikkat dalga tepe değer sayısı (N:33)	10,30
ÇSKTT'nin Neurosky dikkat dalga tepe değer sayısı (N:28)	10,19
İAKTT'nin II. aşamasının Neurosky dikkat dalga tepe değer sayısı (N:33)	9,87

Tablo 5'te, İki Aşamalı Kimya Tanı Testinin (İAKTT) I. ve II. aşamasının Neurosky dikkat dalga tepe değer sayısı ile Çoktan Seçmeli Kimya Testinin (ÇSKTT) Neurosky dikkat dalga tepe değer sayısının karşılaştırıldığı betimsel ortalama değerleri bulunmaktadır. Tablo 5'te bu betimsel ortalama değerlere göre İAKTT'nin I. aşamasının Neurosky dikkat dalga tepe değer sayısına ilişkin ortalaması 10,30; II. aşamasının Neurosky dikkat dalga tepe değer sayısına ilişkin ortalaması 10,19 ve ÇSKTT'nin Neurosky dikkat dalga tepe değer sayısına ilişkin ortalaması 9,87 olarak hesaplanmıştır. Bu sonuçlara göre İAKTT'nin I ve II. aşamasının ortalama Neurosky dikkat dalga tepe değer sayısı ile ÇSKTT'nin ortalama Neurosky dikkat dalga tepe değer sayısının birbirine çok yakın olduğu söylenebilir.

İstatistiksel olarak anlamlı farklılıkları gözetlemek için ilk olarak İAKTT'nin I. aşamasının Neurosky dikkat dalga tepe değer sayısı ile ÇSKTT'nin Neurosky dikkat dalga tepe değer sayısı ikili bağımsız grup karşılaştırmalarında kullanılan Mann-Whitney U (MWU) testi ile karşılaştırılmıştır. İkinci olarak İAKTT'nin I. aşamasının Neurosky dikkat dalga tepe değer sayısı ile II. aşamasının Neurosky dikkat dalga tepe değer sayısı ikili bağımlı grup karşılaştırmalarında kullanılan Wilcoxon Testi (WT) ile analiz edilmiştir. Tablo 6'da, İki Aşamalı Kimya Tanı Testinin (İAKTT) I. aşamasının Neurosky dikkat dalga tepe değer sayısı ile Çoktan Seçmeli Kimya Testinin (ÇSKTT) Neurosky dikkat dalga tepe değer sayısının karşılaştırıldığı Mann-Whitney U testi analiz sonuçları bulunmaktadır.

Tablo 6. Neurosky dikkat dalga tepe değer sayısının karşılaştırılmasına ilişkin MWU test sonuçları

Grup	N	Ortalama Sırası	Sıra Toplamı	MWU	p
İAKTT'nin I. aşamasının Neurosky dikkat dalga tepe değer sayısı	33	31,70	1046,00	439,000	0,739
ÇSKTT'nin Neurosky dikkat dalga tepe değer sayısı	28	30,18	845,00		

Tablo 6'da MWU analiz sonuçlarına göre İAKTT'nin I. aşamasının Neurosky dikkat dalga tepe değer sayısına ilişkin ortalama sırası 31,70 (Sıra Toplamı: 1046,00) ve ÇSKTT'nin Neurosky dikkat dalga tepe değer sayısına ilişkin ortalama sırası 30,18 (Sıra Toplamı 845,00) olarak hesaplanmıştır. Bu sonuçlara göre İAKTT'nin I. aşamasının Neurosky dikkat dalga tepe değer sayısının ortalama sırası ile ÇSKTT'nin Neurosky dikkat dalga tepe değer sayısının ortalama sırası arasında anlamlı bir farklılığın olmadığı belirlenmiştir ($U=439,000$; $p>0,05$). Bu bulguya göre İAKTT'nin I. aşamasının Neurosky dikkat dalga tepe değer sayısı ile ÇSKTT'nin Neurosky dikkat dalga tepe değer sayısı arasında benzer olduğu sonucu çıkarılabilir. Tablo 7'de İAKTT'nin I. aşamasının Neurosky dikkat dalga tepe değer sayısı ile II. aşamasının Neurosky dikkat dalga tepe değer sayısı karşılaştırıldığı Wilcoxon testi analiz sonuçları bulunmaktadır.

Tablo 7. İAKTT'nin I ve II. aşamasının NeuroSky tepe değer sayısının Wilcoxon Test ile karşılaştırılması

	N	Ortalama Sırası	Sıra Toplamı	Wilcoxon Test	p
Negatif Sıralar	23 ^a	17,83	410,00		
Pozitif Sıralar	10 ^b	15,10	151,00	-2,314	0,021
Eşlenik Sıralar	0 ^c				Anlamlı

^a. İAKTT'nin II. aşamasının Neurosky dikkat dalga tepe değer sayısı < İAKTT'nin I. aşamasının Neurosky dikkat dalga tepe değer sayısı;
^b. İAKTT'nin II. aşamasının Neurosky dikkat dalga tepe değer sayısı > İAKTT'nin I. aşamasının Neurosky dikkat dalga tepe değer sayısı;
^c. İAKTT'nin II. aşamasının Neurosky dikkat dalga tepe değer sayısı = İAKTT'nin I. aşamasının Neurosky dikkat dalga tepe değer sayısı

İAKTT'nin I. ve II. Aşamasındaki Neurosky dikkat dalga tepe değer sayısı arasındaki korelasyon değeri r=0,885; p=0,000; Çok Yüksek Düzeyde İlişki

Tablo 7'de Wilcoxon testi analiz sonuçlarına göre, negatif sıraların ortalama sırası 17,83 ve pozitif sıraların ortalama sırası 15,10 olarak hesaplanmıştır. Diğer bir ifadeyle İAKTT'nin I. aşamasının Neurosky dikkat dalga tepe değer sayısının İAKTT'nin II. aşamasının Neurosky dikkat dalga tepe değer sayısından biraz daha fazla olduğudur. Bu sonuçlara göre İAKTT'nin I. aşamasının Neurosky dikkat dalga tepe değer sayısının ortalama sırası ile II. aşamasının Neurosky dikkat dalga tepe değer sayısının ortalama sırası arasında anlamlı bir farklılığın olduğu saptanmıştır (WT=-2,314; p<0,05). Bu bulguya göre İAKTT'nin I. aşamasının Neurosky dikkat dalga tepe değer sayısının İAKTT'nin II. aşamasının Neurosky dikkat dalga tepe değer sayısından biraz daha fazla olduğu söylenebilir.

Neurosky dikkat dalga tepe değer sayısına ilişkin genel olarak elde edilen bulgularda çoktan seçmeli teste katılan öğrencilerin Neurosky dikkat dalga tepe değer sayıları ile İAKTT'ye katılan öğrencilerin Neurosky dikkat dalga tepe değer sayıları arasında anlamlı bir farklılık görülmemiştir. İAKTT'ye katılan öğrencilerin birinci ve ikinci aşamalardaki Neurosky dikkat dalga tepe değer sayıları karşılaştırıldığında anlamlı bir farklılık belirlenmiş; öğrencilerin ikinci aşamada daha az dikkat performansı gösterdiği görülmüştür. İAKTT'nin I. ve II. aşamasının Neurosky dikkat dalga tepe değer sayıları arasında çok yüksek düzeyde bir ilişkinin olduğu bulunmuştur.

Çalışmada üçüncü olarak NeroSky Beyin Sensörüyle yapılan ölçümlere göre İAKTT ve ÇSKTT'yi alan öğrencilerin dikkat düzeyleri ile yanıtlama süreleri ilişkisel olarak karşılaştırılmış ve bu analize yönelik korelasyonel bulgular verilmiştir. Tablo 8, bu analizlere ilişkin bulguları göstermektedir.

Tablo 8. İAKTT'nin I ve II. aşamasının NeuroSky dikkat dalga tepe değer sayısı ile yanıtlama sürelerinin korelasyon tablosu

	I. aşama Neurosky dikkat dalga tepe değer sayısı	II. aşama Neurosky dikkat dalga tepe değer sayısı
I. aşama yanıtlama süresi	0,918 (P=0,000)	-0,074 (P=0,634)
II. aşama yanıtlama süresi	0,069 (P=0,656)	0,960 (P=0,000)
I. aşama Neurosky dikkat dalga tepe değer sayısı	1	-0,041 (P=0,791)

Tablo 8'de, İki Aşamalı Kimya Tanı Testinin (İAKTT) I. ve II. aşamasının yanıtlama süreleri ile İki Aşamalı Kimya Tanı Testinin (İAKTT) I. ve II. aşamasının NeuroSky dikkat dalga tepe değer sayısının korelasyonel değerleri bulunmaktadır. Tablo 8'deki bu korelasyonel değerlere göre İAKTT'nin I. aşamasının yanıtlama süresi ile İAKTT'nin I. aşamasının NeuroSky dikkat dalga tepe değer sayısı arasındaki korelasyonel değer 0,918'dir. İAKTT'nin II. aşamasının yanıtlama süresi ile İAKTT'nin I. aşamasının NeuroSky dikkat dalga tepe değer sayısı arasındaki korelasyonel değer 0,069'dur. İAKTT'nin I. aşamasının yanıtlama süresi ile İAKTT'nin II. aşamasının NeuroSky dikkat dalga tepe değer sayısı arasındaki korelasyonel değer -0,074'tür. İAKTT'nin II. aşamasının yanıtlama

süresi ile İAKTT'nin II. aşamasının NeuroSky dikkat dalga tepe değer sayısı arasındaki korelasyonel değer 0,960'tır. İAKTT'nin I. aşamasının NeuroSky dikkat dalga tepe değer sayısı ile İAKTT'nin II. aşamasının NeuroSky dikkat dalga tepe değer sayısı arasındaki korelasyonel değer -0,041'dir. I. aşamanın yanıtlama süresi ile I. aşamanın dikkat dalga tepe değer sayısı arasında çok yüksek düzeyde bir ilişki tespit edilmiştir. II. aşamanın yanıtlama süresi ile II. aşamanın dikkat dalga tepe değer sayısı arasında çok yüksek düzeyde bir ilişki tespit edilmiştir. Yapılan diğer karşılaştırmalarda ise çok düşük düzeyde bir ilişki tespit edilmiştir.

TARTIŞMA

Bu çalışma genel kimya dersi kapsamındaki konuları içeren, aşamalı ve çoktan seçmeli testlerin yanıtlama süreleri ve dikkat düzeyi puanlarının ortalamalarını incelemektedir. Bu bölümde tartışma, çalışmanın alt problemlerin çözümüne yönelik sırasıyla yapılmıştır. Her alt problemin bulgu sunumu yapıldıktan sonra ilgili alandaki çalışmalarla karşılaştırılmış ve bir sonuca ulaşılmıştır.

Bilgisayar tabanlı testler, test maddelerinin yanıtlanması süresini ölçerek test katılımcıları ya da testin yapısı ile ilgili çok daha yararlı bilgiler sunabilir (Wirth, 2008). Bilgisayarlı testler, test yanıtlama performanslarının ve yanıt sürelerinin toplanmasını kolaylaştırarak sınava girenlerin değerlendirmeler sırasında sorulara yanıt verme davranışlarının yönlerini ortaya çıkarabilir. Sorulara verilen tepki süresi, sınava girme çabasının geçerli bir davranışsal göstergesi olabilir (Wise & Kong, 2005). Bu kapsamda çalışmanın ilk alt probleminde, NeroSky Beyin Sensörü cihazıyla yapılan ölçümlere göre Diagnostik fen tanı testini alan öğrencilerin yanıtlama süreleri ile çoktan seçmeli teste alan öğrencilerin yanıtlama süreleri arasındaki fark incelenmiştir. İAKTT ve ÇSKTT'ye katılan öğrencilerin ortalama yanıtlama süreleri her iki test grubu alanlar arasında karşılaştırıldığında ortalama yanıtlama sürelerinde fark görülmemiş ve yanıtlama sürelerinin çok yakın olduğu saptanmıştır. Aşamalı testin ilk aşaması ile çoktan seçmeli testin yanıtlama sürelerinin birbirine eşdeğer olduğu anlaşılmıştır. Aşamalı testin ikinci aşamasının, birinci aşamaya göre daha kısa sürede yanıtlandığı görülmüştür. Bu farklılığa öğrencilerin çözüm davranışında bulunması, tahminde bulunması, test maddesinin soru kökünün biçimsel ve metinsel yapılarının farklı olması, test maddelerinin zorluk ve ayırt ediciliklerinin farklı olması, kişisel özellikler ve öğrencinin test maddesine verdiği yanıtta güven neden olabilir. Sınava giren öğrenciler, test sorularını yanıtlarken iki tür davranış sergileyebilirler: çözüm davranışı (doğru yanıtı bulmaya çalışmak) ve hızlı tahmin davranışı (soruyu dikkate almadan hızla yanıt vermek) (Wise & Kong, 2005). Bir test maddesine yanıt verirken çözüm çabasında istekli olarak sınava giren kişiler, doğru yanıtı aramak için zaman harcayarak çözüm davranışında bulunurken, ilgisiz sınav katılımcıları, hızlı bir yanıt vererek veya maddeyi hızlı bir şekilde atlayarak hızlı tahmin etme veya hızlı atlama davranışı sergileyebilir (Wise & Kong, 2005; Wise & Gao, 2017). Sınava giren bir kişinin belirli bir soruyu cevaplamak için harcadığı süre, sınavı başarıya çabası için nesnel bir gösterge olarak kabul edilir; çok hızlı yanıtların bir test maddesiyle etkileşim eksikliğinin göstergesi olduğunu varsaymak daha mantıklıdır (Wise, 2017). Hızlı tahmin etme, yani bir maddeden beklenenleri okumak ve anlamak için yeterli süre geçmeden çok hızlı yanıt vermek, teste karşı ilgisiz bir tutum sergileme davranışıyla ilişkilendirilmiştir (Schnipke, 1995; Wise, 2017). Ayrıca çaba göstererek veya ilgisiz bir şekilde sınava girmenin, test süresi boyunca tutarlı olmadığı, ancak testin maddeleri arasında kendine özgü görüldüğü iddia edilmiştir (Wise & Kingsbury, 2016). Örneğin, daha uzun maddeler, testte daha sonradan görünen maddeler ve yardımcı okuma materyali içeren maddeler gibi belirli madde özellikleri, yanıt süresiyle ölçüldüğü üzere öğrenciler tarafından daha az çaba sarf etme eğilimindedir (Setzer vd., 2013). Halkitis ve diğ. (1996), madde yanıt süresinin kelime sayısından, maddenin zorluk derecesinden ve ayırt edicilik düzeyinden etkilendiği tespit edilmiştir. Farklı yetenek düzeylerine (düşük, orta, yüksek) göre bir İngilizce testinde madde istatistikleri ile madde yanıt süresi arasındaki ilişkinin incelendiği bir çalışmada, yanıt süresinin maddelerin zorluk düzeyine göre değiştiği görülmektedir. Üst düzey öğrenciler kolay soruları daha kısa sürede yanıtlar ve zor konulara daha fazla zaman harcarlar. Madde ayırt edicilik indeksinin öğrencilerin yetenek düzeylerine göre cevaplama sürelerine etkisinin olmadığı sonucuna ulaşılmış; bir test maddesinin yanıtlanma süresinin öğrencinin yetenek düzeyine ve maddelerin güçlük düzeyine göre farklılık gösterdiği görülmüştür (Altuner, 2019; Goldhammer, Naumann, Stelter, Tóth, Rölke & Klieme, 2014; İlgün-Dibek, 2020;

Yavuz, 2019). Ayrıca başka bir çalışmada ise test maddesinin yanıtlanmasına ilişkin öğrencilerin güven düzeyinin düşük olması test maddelerini yanıtlamak için daha fazla zaman harcamalarına neden olduğu görülmüştür (Lasry, Watkins, Mazur & Ibrahim, 2013). Türkoğuz (2020a) yaptığı çalışmada aşamalı tanı testinin ikinci aşamadaki test performanslarının ve yanıtlama sürelerinin farklılığını test maddesinin zorluğuna ve test maddesi yapılarına bağlamıştır. Ayrıca aşamalı testlerde öğrencilerin I. aşamaya aşına olmalarının II. aşamadaki performanslarını etkilediğini, yanıtlamayı daha kısa sürede tamamladığını ve daha düşük yanıtlama performansı gösterdiğini ifade etmiştir. Ayrıca öğrencilerin ikinci aşamada daha hızlı tahmin davranışı gösterdiğini bulmuştur. Buna benzer olarak Türkoğuz (2020b), başka bir çalışmada aşamalı testlerde öğrencilerin birinci aşamada test başarısının daha yüksek olduğunu gözlemlemiştir. Buna gerekçe olarak birinci aşamada kullanılan seçenek sayılarının daha az olması, ikinci aşamada seçenek sayılarının daha fazla olmasını göstermiştir. Bu çalışmada ilk aşamada 3 seçenek ikinci aşamada 4 seçenek kullanılmıştır. İkinci aşamada 4 seçenek olmasına rağmen ve öğrencilerin kelime sayısına göre daha uzun sürede yanıt vermesi gerekirken kısa sürede yanıtlaması öğrencilerin ikinci aşamada tahmin etme davranışında olduğu sonucu çıkarılabilir. Bu nedenle birinci ve ikinci aşamanın yanıtlama sürelerini ayrı ayrı değerlendirilmesi tahmin davranışını ortaya çıkarmada yardımcı olabileceği gibi testin daha gerçekçi sonuçlarla değerlendirilmesine yardımcı olabilir. Böylelikle birinci ve ikinci aşamanın birlikteliği ile şans faktörü olasılığı azaltılmış olur.

İkinci olarak, NeroSky Beyin Sensörü cihazıyla yapılan ölçümlere göre Diagnostik fen tanı testini alan öğrencilerin dikkat düzeyleri ile çoktan seçmeli testi alan öğrencilerin dikkat düzeyleri arasındaki fark incelenmiştir. İAKTT ve ÇSKTT'ye katılan öğrencilerin ortalama Neurosky dikkat dalga tepe değer sayısı her iki test grubu alanlar arasında karşılaştırıldığında anlamlı fark görülmemiş ve NeuroSky dikkat dalga tepe değer sayısının çok yakın olduğu saptanmıştır. İAKTT'ye katılan öğrencilerin I. ve II. aşamasındaki ortalama Neurosky dikkat dalga tepe değer sayısı aşamalar arasında karşılaştırıldığında I. aşama lehine anlamlı fark görülmüştür. Cheng ve diğ. (2015) yaptıkları çalışmada okuma ile dikkat arasında ilişki olduğunu tespit etmişlerdir. Bu çalışmada İAKTT'nin I. ve II. aşamaları arasında dikkat düzeyinde farkların olması metinlerin uzunluğundan kaynaklanabilir. Metinlerin uzunluğu öğrencilerin yanıtlama sürelerini etkilemiş olabilir. Shadiev ve Huang (2020) İngilizce çeviri eğitimiyle ilgili yaptıkları bir çalışmada bilişsel yükün dikkat düzeyine etkisinin olduğunu ifade etmiştir. Bu çalışmada İAKTT'nin ikinci aşaması öğrencilerde ek bilişsel yük ve yorgunluk getirmesi dikkati ve yanıtlama süresini etkileyebilir. Ayrıca bu çalışmada öğrencilerin ikinci aşamada hızlı tahmin davranışında bulunmaları soruyu yanıtlamada gereken dikkati vermediklerini de gösterebilir. Böylelikle öğrencilerin ilk aşamaya aşına olmaları II. aşamadaki Neurosky dikkat dalga tepe değer sayısı da etkileyebilir.

Üçüncü olarak, “NeroSky Beyin Sensörü cihazıyla yapılan ölçümlere göre Diagnostik fen tanı testini alan öğrencilerin dikkat düzeyleri ile yanıtlama süreleri arasındaki ilişki incelenmiştir. Sonuçlara göre İAKTT'nin I. ve II. aşamasında yanıtlama süreleri ile NeuroSky dikkat dalga tepe değer sayısı arasında yüksek düzeyde bir ilişkinin olduğu belirlenmiştir. Kokubo ve Shoji (2018) yaptıkları çalışmada odaklanma süresiyle dikkat düzeyleri arasında yüksek ilişki bulmuşlar; daha uzun süre odaklanan öğrencilerin daha başarılı olduğunu gözlemlemiştir. Yang ve diğ. (2002) çözüme yönelik davranış sergileyen kişilerin, hızlı yanıt veren kişilere göre zor maddelere verdikleri yanıt süresinin fazla olduğunu bulmuşlardır. Soland ve diğ. (2019) yetenek düzeyi ile hızlı tahmin davranışı arasındaki ilişkiyi incelediği çalışmada, çok kısa sürede soruya yanıt veren kişilerin yetenek düzeyinin düşük olduğunu bulmuştur. Ponce ve diğ. (2020), öğrencilerin bilgisayar ortamında kişilerin gerçekleştirdikleri eylemlerin yanıt sürelerini ve yanıt doğruluklarını nasıl etkilediğini araştırmışlardır. Boşluk doldurma testi üzerinden yapılan çalışmada, farklı arayüzlere dayalı yanıtlama seçeneklerinin öğrencilerin yanıtlama sürelerinde farklılaşmaya sebep olduğunu tespit etmişlerdir. Yavuz (2019) çalışmada maddeleri yanıtlama süreleri ve maddeyi yanıtlarken gerçekleştirdikleri eylem sayıları arasındaki ilişkiyi incelemiştir. Maddeleri yanıtlamak için gerçekleştirilen eylem sayısı ve maddeleri yanıtlama süreleri arttıkça maddeleri doğru yanıtlama olasılıklarının arttığını bulmuşlardır.

Test sırasında madde yanıt süresi, öğrencilerin tutumlarını ve öğelerinin kalitesini anlamak için önemli bir özelliktir. Ancak, test yapanlar genellikle tepki süresine dikkat etmemektedir (Schatz & Brownkye, 2002). Öğrencilerin fare hareketleri ve madde seçim taktikleri, test puanlarına göre izlenebilir ve yorumlanabilir. Klavye ya da bilgisayar faresi gibi giriş cihazlarının kullanımının toplanan verilerin geçerliliğini etkileyebileceğine dikkat edilmelidir (Wirth, 2008). Daha önce yapılan araştırmalar, rastgele tahminleri içerebilen tipik bir cevap modelini saptamak için kişilere uygun istatistikleri uygulasa da eğitimsel testlerde bilgisayar tabanlı değerlendirmelerin kullanılması araştırmacıların, rastgele tahmin etmede yeni yöntemler destekleyen yanıt süresi bilgilerini toplamasına izin vermektedir. Bu yöntemlerde, yetenekleri ne olursa olsun, motive olmamış test katılımcılarının, madde kökünü ve tüm cevap seçeneklerini dikkatli bir şekilde okumak için gerekli zamanı almaması varsayımdır (Swerdzewski, Harmes & Finney, 2011). Lasry ve diğ. (2013), Hestenes ve diğ. (1992) tarafından geliştirilen Force Concept Envanteri (FCI) öncesi ve sonrası eğitim uygulamalarına yanıt süresini analiz etmiştir. Ayrıca, yanıt süresine ilişkin doğruluğun etkisini de araştırmışlar ve “yanıt sürelerinin yanlış cevaplardan doğru olanlardan daha uzun olduğunu” bulmuşlardır.

Çıkarımlar

TED konuşmacılarının süreleri 18 dakika ile, Twitter’da atılan tweetlerdeki karakterin kelime sayısı 280 kelime ile, okullarda yapılan dersler 40 dakika ile sınırlandırılmıştır. Bunun gibi örnekler çoğaltılabilir. Bu sürelerin belirlenmesinde ve sınırlamaların yapılmasında birçok araştırma yapılarak ortalama bir değer belirlenmeye çalışılmıştır. Bireylerin bir işe odaklanma süreleri yapılan işe göre, zorluğuna göre, kişiye göre, şartlara göre ve bunun benzeri birçok etkiye göre değişebilmektedir. Yapılan sınavlarda, öğrencinin dikkat süresi birçok etkene göre değişim gösterebilmektedir. Ortalama bir sürenin tespit edilebilmesi için sürecin çok boyutlu olarak değerlendirilmesi gerekmektedir. Gök (2010), okul öncesi öğretmenlerinin bilişim teknolojileri kullanımında öğrenci ilgi ve dikkat düzeyine etkisine ilişkin algı ve görüşlerinin incelenmesine yönelik çalışmada okul öncesi öğretmenlerinin bilişim teknoloji araçlarını yeterli sürede ve kontrollü bir şekilde kullandıkları zaman, ilgi ve dikkat düzeylerine olumlu yönde etki ettiğine ilişkin bulgular elde edilmiştir. Aşamalı testin sorularının çok ve uzun olması öğrencilerin sıkılmasına ve dikkat sürelerinin kısalmasına sebep olduğu gözlemlenmiştir. Tabi bunun yanında dikkat düzeyi ile yanıtlama süresi arasında güçlü bir bağın olduğu ve birbirini dolaylı yoldan etkilediği söylenebilir. Kişiden kişiye bunun çok fazla farklılık gösterdiği de iddia edilebilir. Bunların göz ardı edilecek tarafları olsa da çok boyutlu yönleri bulunduğu aklın bir kenarında tutularak değerlendirmeye tabi tutulmalıdır. Bunun yanı sıra kişilerin odaklanmalarına bağlı olarak uzun süreli dikkat sağlamaları da mümkün olabilmektedir. Yapılan çalışmanın kişinin ilgisini çekmesi, zorluk düzeyi veya kişisel özellikler öğrencinin dikkat düzeyini ve süresini etkileyebilmektedir.

Araştırmanın Sınırlılıkları

Bu araştırmanın bazı sınırlılıkları vardı. İlk olarak, öğrencilere uygulanan aşamalı testin 44 sorudan oluşması, öğrencilerin sıkılmasına sebep olduğu anlaşılmıştır. Aynı zamanda sürenin uzun olmasından kaynaklı olarak dikkat verilerinin kopması sonucu bazı kişilerde veri kayıplarının olduğu görülmüştür. Analiz aşamasında bu veri kayıpları değerlendirme dışında tutulmuştur. İkinci olarak, testin PowerPoint programı ile uygulanması veri toplam sürecini ve analizleri zorlaştırmıştır. Üçüncü olarak, öğrencilerin çok hoşuna gitse de sorular arasında geçişlerde manzara resmi ve yağmur sesi kullanmanın çok da gerekli olmadığı anlaşılmıştır. Aslında manzara resmi ve yağmur sesini kullanılmasının sebebi sorular arasındaki geçişlerde Neurosky beyin sensörü cihazında toplanan dikkat düzeyleri verilerini soru bazlı olarak net olarak görebilmektedir. Test ekranı yeni soruya geçiş yaptığında, cihazda müzik ve görsel destekli rahatlama seanslarını görerek sorulardaki kayıt başlangıç anını belirleyebilmek mümkün hale gelmektedir. Analizler sürecinde bu bilgi destek olsa da testin biraz daha uzamasına sebep olmuştur. Son olarak, öğrencilerin makyaj yapması, küpe takması ve kullandıkları takıların NeuroSky cihazında veri elde ederken kopmalara sebep olduğu gözlemlenmiş ve kolonyalı mendille temas noktaları temizlenerek bu sorun çözülmeye çalışılmıştır.

Sonuçlar

Testlerin yanıtlanma süreleri arasında anlamlı bir farklılık görülmemiş, aşamalı testin aşamaları arasında öğrencilerin II. aşamayı daha kısa sürede tamamladığı tespit edilmiştir. Testlerde, NeuroSky dikkat dalga tepe değer sayısının birbirine çok yakın olduğu bulunmuştur. Aşamalı testte, II. aşamanın NeuroSky dikkat dalga tepe değer sayısının daha fazla olduğu görülmüştür. Aşamalı testte, yanıtlanma süresi ile NeuroSky dikkat dalga tepe değer sayısı arasında yüksek düzeyde bir ilişki tespit edilmiştir. Bilgisayar destekli sınavların yaygınlaşması ile dijital araçlarla ölçüm yapabilme imkânı testlerin değerlendirilmesinde farklı bir boyut kazandırmıştır. Gelişen teknolojiler sayesinde, öğrencilerin dikkat düzeylerini NeuroSky cihazı gibi bazı cihazlarla ölçmek mümkün hale geldiği anlaşılmış ve farklı öğrenme süreçlerinde kullanımının uygunluğu bu çalışmayla test edilmiştir. Bununla birlikte öğrencilerin her bir soruya ne kadar vakit ayırdığı, bu süre içerisinde dikkat düzeylerinin nasıl değiştiğinin ölçülebilmesi süreci değerlendirme açısından önemli ipuçları verebilir. Pandemi, yaşam şartlarının değişmesi gibi durumlardan kaynaklı bireysel öğrenmenin yaygınlaştığı eğitim süreçleri düşünüldüğünde, öğrencilere öğrenme süreci ile ilgili olarak test performanslarının yanında yanıtlanma, dikkat, kaygı vb. duyuşsal verileri içeren geri dönütlerin verilmesi eğitimi katkı kapsamında daha faydalı olacağı düşünülmektedir. Eğitimciler, bu imkânlardan faydalanabildiği ölçüde elinde daha fazla veri olacak ve öğrenme sürecini yapılandırmasında faydalı geri dönütleri elde etmiş olacaktır.

Öneriler

Bu çalışmada veriler koronavirüs pandemisinin hemen sonrasında toplanmıştır. Bu nedenle öğrencilerin bilgi ve hazırbulunuşluk düzeylerinin pandemi sürecinden etkilenmesi söz konusudur. Belirli bir zaman sonra çalışmanın tekrarından çıkacak sonuçların pozitif ya da negatif yönleri merak edilmektedir. Böylelikle pandemi sürecinin bu çalışmaya olan etkilerinin pozitif ya da negatif yönleri öğrenilebilir. Bu çalışmada öğrencilerin dikkat ölçümleri NeuroSky MindWave cihazıyla gerçekleştirilmiştir. Dikkat ölçümleri yapan başka cihazlarla karşılaştırması yapılabilir. Örneğin dikkat ölçümü yapan cihazlarla toplanan dikkat verileri göz takip cihazı gibi göz odaklanmasının belirlendiği beyin dalgası dışındaki farklı verilerle ilişkilendirilmesi gerçekleştirilebilir. Katılımcıların test yanıtlanma performanslarıyla yanıtlanma sürelerinin birlikte kullanıldığı matematiksel modellemelerle ilgili çalışmalara rastlanılmıştır. Test yanıtlanma performansı ve yanıtlanma süresine ek olarak dikkat verilerinin olduğu değişkenlerle üçlü bir matematiksel model geliştirilerek kişi ve test yeteneklerinin değerlendirilmesinde farklı boyuta gidilebilir. Bu çalışma çevrimiçi ortamda gerçekleştirilerek dikkat ölçümleri gerçekleştirilmiştir. Benzer çalışma kâğıt testleri üzerinde de denenebilir. Hatta sınavlarda salon etkisi bu yöntemle de incelenebilir. Buna benzer olarak öğrencilerin dikkat düzeyleri ders ortamlarında değerlendirilebilir.

Etik ve Çıkar Çatışması Beyanı

Bu çalışma, Dokuz Eylül Üniversitesi Bilimsel Araştırma Projeleri Koordinasyon Birimi tarafından Doktora Tez Projesi kapsamında desteklenen 2021.KB.EGT.001 nolu ve "Diagnostik Fen Testlerinin Dikkat, Yanıtlanma Süresi ve Başarı Düzeylerinin Beyin Dalgalarına Göre İncelenmesi" başlıklı projeden üretilmiştir. Ayrıca bu çalışma 23-24 Nisan 2022 tarihinde 2. Uluslararası Bursa Bilim Kongresinde sözlü bildiri olarak sunulmuş ve kongrenin en iyi sözlü bildirisi olarak kongre bilim kurulu tarafından ödüllendirilmiştir. Yazarlar etik ilkelere uygun olarak davrandıklarını, ve yazarlar arasında çıkar çatışmasının olmadığını beyan etmişlerdir.

KAYNAKÇA

- Altuner, F. (2019). *Investigation of the relationship between item statistics and item response time* (Unpublished Master's Thesis). Mersin University, Institute of Education Sciences, Mersin.
- Andrewes, D. (2009). *Neuropsychology: from theory to practice*. Psychology Press.
- Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In K. W. Spence, & J. T. Spence(Eds.), *The psychology of learning and motivation*. New York: Academic Press. Doi: 10.1016/s0079-7421(08)60422-3.
- Balcı, A. (2005). *Sosyal bilimlerde araştırma: Yöntem, teknik ve ilkeler (5.Baskı)*. Ankara: Pegem Publishing.
- Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: Freeman and Company.

- Bayazit, A. (2013). *Investigating the effects of different question modalities on eye movements, performance and response time*. (Unpublished Doctorate Dissertation). Hacettepe University Graduate School of Science and Engineering, Ankara.
- Bayazit, Ö. (2007). An examination of current collaborative supply chain practices. *International Journal of Business Innovation and Research*, 1(3), pp. 253-266. Doi:10.1504/IJBIR.2007.012110
- Bernt, F. M., & Bugbee, A. C. (1988). *Your time is up! An assessment of time limits for American college students*. Examination Research Report No. 88-1. Bryn Mawr, PA: The American College.
- Bodmann, S. M. & Robinson, D. H. (2004). Speed and performance differences among computer-based and paper pencil tests. *Journal of Educational Computing Research*, 31(1), pp. 51-60. Doi: 10.2190/grqq-yt0f-7lkb-f033
- Broadbent, D. E. (1958). *Perception and communication*. Pergamon Press, Fitzroy Square, London, UK.
- Browarska, N., Zygarlicki, J., Pelc, M., Niemczynowicz, M., Zygarlicka, M., & Kawala-Sterniuk, A. (2021, August). Pilot study on using innovative counting peaks method for assessment purposes of the EEG data recorded from a single-channel non-invasive brain-computer interface. In *2021 25th International Conference on Methods and Models in Automation and Robotics (MMAR)* (pp. 68-72). IEEE. Doi: 10.1109/mmar49549.2021.9528447
- Buchberger, F., Campos, B. P., Kallos, D., & Stephenson, J. (1500). *Green paper on teacher education in Europe*. Umeå, Sweden: Thematic Network on Teacher Education in Europe. 30 March 2016 retrieved from <http://www.cep.edu.rs/sites/default/files/greenpaper.pdf>
- Büyüköztürk, Ş. (2002). *Veri analizi el kitabı (1. Baskı)*. Ankara: Pegem A Yayıncılık.
- Büyüköztürk, Ş., Çakmak, E. K., Akgün, Ö. E., Karadeniz, Ş., & Demirel, F. (2008). *Bilimsel araştırma yöntemleri*. Ankara: Pegem Akademi Yayıncılık.
- Chan S.-C., Lu T.-S., & Tsai R.-C. (2014). Incorporating response time to analyze test data with mixture structural equation modeling. *Psychological Testing*, 61, pp. 463-488.
- Chen, C. M., & Wu, C. H. (2015). Effects of different video lecture types on sustained attention, emotion, cognitive load, and learning performance. *Computers & Education*, 80, pp. 108-121. Doi: 10.1016/j.compedu.2014.08.015
- Clariana R., Wallace P. (2002). Paper-based versus computer-based assessment: Key factors associated with the test mode effect. *British Journal of Educational Technology*, 33, pp. 593-602. Doi: 10.1111/1467-8535.00294
- Delen, E. (2015). Enhancing a computer-based testing environment with optimum item response time. *Eurasia Journal of Mathematics, Science and Technology Education*, 11(6), pp.1457-1472. Doi: 10.1111/1467-8535.00294
- DeMars, C.E: (2000). Test stakes and item format interactions. *Applied Measurement in Education*, 13(1), pp. 55-77. Doi: 10.1207/s15324818ame1301_3. Doi: 10.1207/s15324818ame1301_3
- Demirci, E. (2011), Beyin Dalgalarıyla Oyun Oynamak, *TÜBİTAK Bilim Teknik Dergisi*, 44 (520), ss. 18-24.
- Demirel, Ç., Kandemir, H., & Köse, H. (2018). *Controlling a robot with extraocular muscles using EEG device*. 26th Signal Processing and Communications Applications Conference (SIU). Doi: 10.1109/SIU.2018.8404157
- Direnga, J, Timmermann, D., Presentati, B., Brose, A., & Kautz, C. (2015, July). Do students spend more time on difficult questions? Analysis of item response time versus correctness in the SCI/CATS. Research in Engineering Education Symposium (REES 2015).
- Direnga, J., Presentati, B., Timmermann, D., Brose, A., & Kautz, C.H. (2015, June). *Does it stick? – Investigating long-term retention of conceptual knowledge in mechanics instruction*. In presented at 2015 ASEE Annual Conference & Exposition, Seattle, Washington. 10.18260/p.23897
- Eells, R. J. (2011). *Meta-analysis of the relationship between collective teacher efficacy and student achievement*, Unpublished Doctorate Thesis, Loyola University Chicago, Chicago, IL.
- Fernández-Castillo, A., & Caurcel, M. J. (2015). State test-anxiety, selective attention and concentration in university students. *International Journal of Psychology*, 50(4), pp. 265-271. Doi: 10.1002/ijop.12092
- Fraenkel, J. R., & Wallen, N. (2000). *How to design and evaluate research in education* (4th ed.). NY: McGraw-Hill.
- Gass, C S., & Curiel, R.E. (2011). Test anxiety in relation to measures of cognitive and intellectual functioning. *Archives of Clinical Neuropsychology*, 26(5), pp. 396-404. Doi:10.1093/arclin/acr034
- Goddard, R. G., Hoy, W. K., & Hoy, A. W. (1504). Collective efficacy: Theoretical development, empirical evidence, and future directions. *Educational Researchers* 33(3), pp. 3-13.
- Goldhammer, F., Naumann, J., Stelter, A., Tóth, K., Rölke, H., & Klieme, E. (2014). The time on task effect in reading and problem solving is moderated by task difficulty and skill: Insights from a computer-based large-scale assessment. *Journal of Educational Psychology*, 106(3), pp. 608-626. Doi:10.1037/a0034716

- Halkitis, P. N., Jones, J. P., & Pradhan, J. (1996, April 8-12). Estimating testing time: The effects of item characteristics on response latency. In *presented annual meeting of the American Educational Research Association*, New York.
- Hestenes, D., Wells, M. ve Swackhamer, G. (1992). Force concept inventory. *Phys. Teach.* 30(3), pp. 141-158. Doi: 10.1119/1.2343497
- Hunt, R. R., & Ellis, H. C. (2004). *Fundamentals of cognitive psychology (Edition 7)*. McGraw-Hill.
- İlgün-Dibek, M. (2020). Silent predictors of test disengagement in PIAAC 2012. *Journal of Measurement and Evaluation in Education and Psychology*, 11(4), pp. 430-450. Doi:10.21031/epod.796626
- İnel, Y. (2014). *The effects of computer based instructional materials used in social study of sixth grade students' attention and motivation levels* (Unpublished Doctorate Dissertation), Gazi University, Institute of Education Sciences, Ankara.
- Kokubo, Y., & Shoji, Y. (2018). Relationship between Brain Waves and Examination Achievements. *Information Engineering Express*, 4(1), pp. 53-62. Doi: 10.52731/iee.v4.i1.254
- Lasry, N., Watkins, J., Mazur, E., & Ibrahim, A. (2013). Response times to conceptual questions. *American Journal of Physics*, 81(9), pp. 703-706. Doi: 10.1119/1.4812583
- Lee J., Moreno K. E., Sympson J. B. (1986). The effects of mode of test administration on test performance. *Educational and Psychological Measurement*, 46, pp. 467-473. Doi: 10.1177/001316448604600224
- Luecht, M.(2005). Some useful costbenefit criteria for evaluating computerbased test delivery models and systems. *Journal of Applied Testing Technology*, 7(2), pp. 1-31.
- Mann, M., & Treagust, D. F. (1998). *A pencil and paper instrument to diagnose students' conception of breathing, gas exchange and respiration*. *Australian Science Teachers Journal*, 44(2), pp. 55-59.
- Mason B. J., Patry M., & Bernstein D.J. (2001). An examination of the equivalence between non-adaptive computer-based and traditional testing. *Journal of Educational Computing Research*, 24, pp. 29-39. Doi: 10.2190/9epm-b14r-xqwt-wvnl
- Mayerl, J. (2013). Response latency measurement in surveys. detecting strong attitudes and response effects. *Survey Methods: Insights from the Field*. Retrieved from <https://surveyinsights.org/?p=1063>. Doi: 10.13094/SMIF-2013-00005
- Mintzes, J. J., Wandersee, J. H., & Novak, J. D. (2001). *Assessing understanding in biology*. *Journal of Biological Education*, 35(3), pp. 118-125. Doi: 10.1080/00219266.2001.9655759
- National Association for the Education of Young Children [NAEYC] (1509). *NAEYC standards for early childhood professional preparation programs*. 28 July 2015 retrieved from <https://www.naeyc.org/files/naeyc/file/positions/ProfPrepStandards09.pdf>
- Nikou, S., & Economides, A. A. (2013). Student achievement in paper, computer/web and mobile-based assessment. In *presented Balkan Conference in Informatics, BCI '13*, Thessaloniki, Greece.
- Ommerborn, R., & Schuemer, R. (2001). Using computers in distance study: Results of a survey amongst disabled distance students. FernUniversität-Gesamthochschule in Hagen.
- Özkan, N. (2017). *EMDR device design and determination of optimum operation parameters with signal processing techniques* (Unpublished Master's Thesis). Afyon Kocatepe University, Institute of Education Sciences, Afyonkarahisar.
- Pacheco-Unguetti, A. P., Acosta, A., Lupianez, J., Roman, N., & Derakshan, N. (2012). Response inhibition and attentional control in anxiety. *Quarterly Journal of Experimental Psychology*, 65, pp. 646-660. Doi:10.1080/17470218.2011.637114.
- Palmer, D. H. (1998). *Measuring contextual error in the diagnosis of alternative conceptions in science*. *Issues in Educational Research*, 8(1), pp. 65-76.
- Pashler, H.E. (1984). Processing stages in overlapping tasks: Evidence for a central bottleneck. *Journal of Experimental Psychology: Human Perception and Performance*, 10, 358-377. Doi: 10.1037/0096-1523.10.3.358
- Ponce, H. R., Mayer, R. E., Sithiworachart, J., & López, M. J. (2020). Effects on response time and accuracy of technology-enhanced cloze tests: An eye-tracking study. *Educational Technology Research and Development*, 68(5), pp. 2033-2053. Doi: 10.1007/s11423-020-09740-1
- Posner, M. I., Walker, J. A., Friedrich, F. J., & Rafal, R. D. (1984). *Effects of parietal injury on covert orienting of attention*. *The Journal of Neuroscience*, 4(7), pp. 1863-1864. Doi: 10.1523/jneurosci.04-07-01863.1984
- Prisacari, A.A., & Danielson, J. (2017). Computer-based versus paper-based testing: Investigating testing mode with cognitive load and scratch paper use. *Computers in Human Behavior*, 77, pp. 1-10. Doi: 10.1016/j.chb.2017.07.044

- Russell M. (1999). *Testing on computers: A follow-up study comparing performance on computer and on paper*. (Unpublished doctoral dissertation). Boston College, The Graduate School of Education, USA.
- Russell, M., & O'Connor, K. (2003). Computer-based testing and validity: a look back and into the future. Lynch School of Education in TASC Publications, Boston College.
- Schatz, P., & Browndyke, J. (2002). Applications of computer-based neuropsychological assessment. *Journal of Head Trauma Rehabilitation*, 17(5), pp. 395-410. Doi: 10.1097/00001199-200210000-00003
- Schnipke, D. L. (1995). *Assessing speededness in computer-based tests using item response times* (Unpublished doctoral dissertation). Johns Hopkins University, Baltimore, MD.
- Schnipke, D.L., & Scrams, D.J. (1999). *Representing response time information in item banks* (LSAC Computerized Testing Report No. 97-09). Newtown, PA: Law School Admission Council.
- Semmes, R., Davison, M. L., & Close, C. (2011). Modeling individual differences in numerical reasoning speed as a random effect of response time limits. *Applied Psychological Measurement*, 35(6), pp. 433-446. Doi: 10.1177/0146621611407305
- Setzer, J.C., Wise, S.L., Heuvel, J.R., & Ling, G. (2013) An Investigation of examinee test-taking effort on a large-scale assessment. *Applied Measurement in Education*, 26(1), pp. 34-49. Doi: 10.1080/08957347.2013.739453
- Sevinç E., (2006). Beyin Bilgisayar Arayüzleri, http://www.rehabilitasyon.com/action/makale/1/Beyin_Bilgisayar_Arayuzleri-2299 (15 aralık 2022).
- Shadiev, R., & Huang, Y. M. (2020). Investigating student attention, meditation, cognitive load, and satisfaction during lectures in a foreign language supported by speech-enabled language translation. *Computer Assisted Language Learning*, 33(3), pp. 301-326. Doi: 10.1080/09588221.2018.1559863
- Soland J., Wise S.L., & Gao, L. (2019). Identifying disengaged survey responses: New evidence using response time metadata. *Applied Measurement in Education*, 32(2), pp. 151-165. Doi: 10.1080/08957347.2019.1577244
- Solso, R., Maclin, K. M., & Maclin, O. H. (2009). *Cognitive psychology [Bilişsel psikoloji] (in English)* (4. Baskı, A. Ayçiçeği-Dinn, Trans.). İstanbul: Kitabevi.
- Soraghan, C., Matthews, F., Kelly, D., Ward, T., Markham, C., Pearlmutter, B.A., & O'Neill, R., (2006, November). A dual-channel optical brain-computer interface in a gaming environment. In presented *CGAMES 2006 - 9th International Conference on Computer, UK*.
- Spray J. A., Ackerman T.A., Reckase M.D., & Carlson J.E. (1989). Effect of the medium of item presentation on examinee performance and item characteristics. *Journal of Educational Measurement*, 26, pp. 261-271. Doi: 10.1111/j.1745-3984.1989.tb00332.x
- Stankov, L. and Roberts, R.D. (1997). Mental speed is not the 'basic' process of intelligence. *Personality and Individual Differences*, 22(1), pp. 69-84. Doi: 10.1016/s0191-8869(96)00163-8
- Streiner, D.L. (2003) Being Inconsistent About Consistency: When Coefficient Alpha Does and Doesn't Matter. *Journal of Personality Assessment*, 80(3), pp. 217-222. Doi: 10.1207/S15327752JPA8003_01
- Swanson, D.B., Case, S.E., Ripkey, D.R., Clauser, B.E., & Holtman, M.C. (2001). Relationships among Item Characteristics, Examinee Characteristics, and Response Times on USMLE Step 1. *Academic Medicine*, 76, pp. 114-116. Doi: 10.1097/00001888-200110001-00038
- Swerdzewski, P. J., Harmes, J. C., & Finney, S. J. (2011). Two approaches for identifying low-motivated students in a low-stakes assessment context. *Applied Measurement in Education*, 24, pp. 162-188. Doi: 10.1080/08957347.2011.555217
- Tamir, P. (1971). *An Alternative Approach to The Construction of Multiple Choice Test Items*. *Journal of Biological Education*, 5, pp. 305-307. Doi: 10.1080/00219266.1971.9653728
- Tan, K. C. D., Goh, K. N., Chia, S. L., & Treagust, D. F. (2002). *Development and application of a two-tier multiple choice diagnostic instrument to assess high school students' understanding of inorganic chemistry qualitative analysis*. *Journal of Research in Science Teaching*, 39(4), pp. 283-301. Doi: 10.1002/tea.10023
- Truell, A. D. (2005). Comparing student performance on two computer-based user interfaces and paper-and pencil-test formats. *NABTE Review*, 32, pp. 29-35.
- Truell, A. D., Zhao, J. J., & Alexander, M. W. (2005). The impact of settable test item exposure control interface format on postsecondary business student test performance. *Journal of Career and Technical Education*, 22(1), pp. 31-41. Doi: 10.21061/jcte.v22i1.668
- Türkoguz, S. (2020a). Investigation of three-tier diagnostic and multiple choice tests on chemistry concepts with response change behaviour. *International Education Studies*, 13(9), pp. 10-22. Doi: 10.5539/ies.v13n9p10

- Türkoguz, S. (2020b). Comparison of threshold values of three-tier diagnostic and multiple-choice tests based on response time. *Anatolian Journal of Education*, 5(2), pp. 19-36. Doi: 10.29333/aje.2020.522a
- Wainer, H, Dorans, N. J., Flaugher, R., Green B. F., & Mislevy R. J. (2000). Computerized adaptive testing: A primer. Routledge.
- Weeks, J. P. , Von Davier, M.& Yamamoto, K. (2016). Using response time data to inform the coding of omitted responses. *Psychological Test and Assessment Modeling*, 58 (4), pp. 671-701
- Wirth, J. (2008). *Computer-based tests: alternatives for test and item design*. In J. Hartig, E. Klieme, & D. Leutner (Eds.), *Assessment of competencies in educational contexts* (pp. 235–252). Göttingen: Hogrefe.
- Wise, S. & Kong, X. (2005). Response time effort: A new measure of examinee motivation in computer-based tests. *Applied Measurement in Education*, 18(2), pp. 163–183. Doi: 10.1207/s15324818ame1802_2
- Wise, S. L. (2017). Rapid-guessing behavior: Its identification, interpretation, and implications. *Educational Measurement: Issues and Practice*, 36(4), pp. 52–61. Doi: 10.1111/emip.12165
- Wise, S. L., & Gao, L. (2017). A general approach to measuring test-taking effort on computer-based tests. *Applied Measurement in Education*, 30(4), pp. 343-354. Doi: 10.1080/08957347.2017.1353992
- Wise, S. L., & Kingsbury, G. G. (2016). Modeling student test-taking motivation in the context of an adaptive achievement test. *Journal of Educational Measurement*, 53(1), pp. 86-105. Doi:10.1111/jedm.12102
- Wise, S. L., & Ma, L. (2012, January). *Setting response time thresholds for a CAT item pool: The normative threshold method*. In presented at the annual meeting of the National Council on Measurement in Education, Vancouver, Canada.
- Yang, C. L., O'neill, T. R., & Kramer, G. A. (2002). Examining item difficulty and response time on perceptual ability test items. *J. Appl. Measure.* 3, pp. 282–299.
- Yavuz, H. Ç. (2019). The effects of log data on students' performance. *Journal of Measurement and Evaluation in Education and Psychology*, 10(4), pp. 378-390. Doi: 10.21031/epod.564232
- Yudhana, A., Muslim, A., Wati, D. E., Puspitasari, I., Azhari, A., & Mardhia, M. M. (2020). Human emotion recognition based on EEG signal using fast fourier transform and K-Nearest neighbor. *Adv. Sci. Technol. Eng. Syst. J.*, 5, pp. 1082-1088. Doi: 10.25046/aj0506131

EXTENDED ABSTRACT

People from past to present; they have sought ways to communicate with computers through joysticks, mice, keyboards, microphones, and touchpad, all these interfaces work based on the movements of the muscular system. Today, however, this situation has changed somewhat. Interfaces developed with today's technology and very low energy signals in the human body can now be measured and machines can be controlled. For example, digital data converted by using of signals measured from the brain and interfaces can be sent to computers, and objects can be controlled with certain algorithms. With the development of EEG (Electroencephalography) devices, some data obtained, such as brain waves, is a source for the control of many machines, from object control to the game industry. Additionally, the education-teaching process renews itself according to the changing conditions of the time. The roles assigned to the teacher and student, the assigned tasks, and expectations are changing. The measurement tools used by teachers also vary according to today's conditions and education understanding. With the widespread use of computers in the field of education, exams can be made online, results can be announced in a short time, and data can be presented as feedback to students and teachers by analyzing many subjects. Attention affects the results in all exams and may vary depending on personal characteristics. The student's attention and response time are important factors affecting the results obtained from the measurement tools. The data we received from these factors is of great importance. The data we will obtain from the reactions occurring in our body will be an important source for us. It is thought that measurements made with the NeroSky Brain Sensor device will provide us with more realistic and less manipulated data. NeuroSky MindWave is a general-use, low-cost wireless EEG headset with dry active technology that allows electrodes to be placed without the need for gel. The usage area of the device, which was developed to serve people who were physically incapable in the early days, has increased greatly recently. With the device, which attracted the attention of researchers working on brain waves, studies in different fields were made and new analysis methods were created. There are many measurement and evaluation tools used for different purposes in our country. One of the most widely used is multiple choice tests. In our country, there are many measurement and evaluation tools used for different purposes. One of the most commonly used is multiple choice tests. In the 1980s, two-tier tests were developed to use the positive aspects of multiple choice tests and to minimize the negative aspects. The purpose of the tests is to determine how much the student has learned. Learning is an internal process that includes sensory awareness, attention, recognition, transformation, acquisition, and processing of information. In order for the information to be processed, the process of obtaining

the information starts with attention. Attention is a clear and vivid position of the mind toward an object or set of thoughts. Focus and concentration are the foundation of attention. In tests, besides attention, response time is also important. Being fast has been the first step in the struggle for survival since primitive times. It is necessary to be effective as well as being fast. Therefore, the duration of the work is as important as the level of attention. It is impossible to compete with today's machines in terms of speed. A person can achieve effective results by pushing the limits of his capacity. Practicality, which is accepted as an indicator of being intelligent, is to show itself in cognitive skills, not motor skills. It is widely accepted in society that the faster person is the smarter. Teachers and parents see children who learn quickly as more intelligent. The speed expected in learning is also expected in exams. A link was established between speed and intelligence, and as a result, it was foreseen that some tasks could be done in certain times. Students must complete the process within the defined time in all exams they take. Apart from finding the correct answers, the student is expected to give the correct answer within the given time. Students' response times may vary in many aspects, such as the type of test administered, the difficulty level of the questions, and their personality traits. It is necessary to determine the effects of perception of questions, level of achievement, and response time in exams. The aim of this study was to analyze the students' attention levels and response times in the multiple choice chemistry test and the two-tier chemistry diagnostic test (TTCDDT) performed in a computer environment. In this context, NeuroSky brain sensor and attention data and response time data were used during the tests applied to pre-service science teachers studying at a state university in the fall semester of the 2021-2022 academic year. In this study, correlational and causal research methods were used together. In this context, the peaks related to the attention signals produced by the NeuroSky Brain sensor and the test response times of the students who took the multiple-choice chemistry test and the students who took the TTCDDT (Two-Tier Chemistry Diagnostic Test) were examined. Mann-Whitney U and Wilcoxon Tests, which are non-parametric analysis methods, were used in comparisons between students' response time and NeuroSky attention level. There was no significant difference between the response times of the students who participated in the Multiple Choice Chemistry Diagnostic Test and the response times of the students who participated in the TTCDDT. When the response times of the students participating in TTCDDT were compared for the first and second tiers, a significant difference was determined; it was understood that the students completed the second tier in less time. It was found that there was a moderate relationship between response times and attention frequencies in the second tier of TTCDDT.