



A Test Development Study Based On TIMSS Knowledge Application Reasoning Cognitive Domains in Force and Energy Unit

Gonca Gül Domaç^{1*} and Salih Ateş²

¹Science Teacher, Ministry of National Education, Ankara, Türkiye

²Gazi University, Gazi Faculty of Education, Ankara, Türkiye

ABSTRACT

In this study, the Force and Energy Unit Test (KET) was developed in alignment with the cognitive domains defined by TIMSS (2019), encompassing the learning objectives of the Force and Energy unit in the 7th grade Science curriculum. The test was constructed to assess students' individual performance across the cognitive domains of knowing, applying, and reasoning. To ensure the content and construct validity of the test items, expert consultation was employed. A pilot implementation of the initial 35-item version of the test was conducted with a sample of 371 eighth-grade students from the central districts of Ankara. Confirmatory Factor Analysis (CFA) was utilized to evaluate the structure of the test, and item discrimination and difficulty indices were calculated. Based on these analyses, 8 items were excluded, resulting in a final test comprising 27 items equally distributed among the three cognitive domains (9 items per domain). Item discrimination index, calculated using the point-biserial correlation method, yielded an average discrimination value of 0.38, while the average item difficulty was found to be 0.42. In the pilot study of the test, the KR-20 reliability coefficient was computed as 0.86, and the Cronbach alpha reliability coefficient obtained as a result of the CFA was calculated as 0.85. The test was applied to a sample of 619 7th graders and the KR-20 reliability coefficient was computed as 0.86. As a result, a valid and reliable criterion-referenced assessment tool was developed to measure students' knowledge, application, and reasoning skills within the context of the Force and Energy unit.

Keywords: Force and Energy, test development, TIMSS, knowledge application and reasoning cognitive domains

ARTICLE INFO

Article History:

Received:27.05.2025

Received in revised form:21.06.2025

Accepted:25.06.2025

Available online:28.06.2025

Article Type: Research Article

To Cite This Article: Domaç, G. G., & Ateş, S. (2025). A test development study based on TIMSS knowledge application reasoning cognitive domains in force and energy unit. *Journal of Individual Differences in Education*, 7(1), 54-74, DOI: 10.47156/jide.1707289

1. Extended Summary

1.1 Introduction

Science education aims to enable students to acquire scientific knowledge—such as theories, laws, principles, and concepts proposed by scientists—as well as an understanding of how scientists explore the natural order of the world (NRC, 2015). Since the implementation of the 2005 curriculum reform in Turkey, educational programs have increasingly emphasized the development of students' reasoning, argumentation, and scientific process skills, as opposed to the mere transmission of factual knowledge (MEB, 2006; 2013; 2018; 2024). Consequently, assessments of scientific literacy should not only

*Corresponding author's address: Ministry of National Education, Ankara, Türkiye
e-mail: goncaguldmac@gmail.com

measure memorized knowledge but also evaluate students' ability to apply their knowledge and reason within new contexts.

International assessments such as PISA and TIMSS categorize test items according to specific cognitive domains. This study utilized the taxonomy provided by TIMSS (2019) to evaluate the cognitive processes targeted by the learning outcomes of the Force and Energy unit within the 7th grade Science curriculum. Accordingly, a three-dimensional criterion-referenced test was developed, addressing the cognitive domains of *knowing*, *applying*, and *reasoning*. A review of the existing literature revealed a lack of assessments specifically designed to evaluate students' performance across these cognitive domains in the context of the Force and Energy unit, particularly those aligned with the TIMSS (2019) framework. Therefore, this study aimed to develop a criterion-referenced test to fill this gap.

1.2 Method

A survey model was employed for the research design. The study population consisted of middle schools located in the central districts of Ankara, Turkey. A random sampling method was used to select the sample, which included 371 eighth-grade students who had previously completed the Force and Energy unit.

The *Force and Energy Unit Test* (KET), developed as a criterion-referenced test, was structured in accordance with the TIMSS (2019) cognitive domains: *knowing*, *applying*, and *reasoning*. During the test development process, a specification table was created and revised based on expert feedback, and the learning outcomes from the curriculum were refined or supplemented as necessary. The initial draft of the test comprised 35 multiple-choice items, which were reviewed by field experts for alignment with the learning outcomes, cognitive domains, and multiple-choice test design principles.

Confirmatory Factor Analysis (CFA) was conducted to validate the theoretical three-dimensional structure of the test. In addition to CFA, item discrimination indices were calculated using the point-biserial correlation method.

1.3 Findings

The inter-rater agreement rate for the classification of items into cognitive domains was 63%, with a chance agreement rate of 34.4%. The Fleiss' Kappa coefficient was calculated as 0.44, indicating a moderate to good level of agreement among experts. In the pilot study of the test, the KR-20 reliability coefficient was computed as 0.86, and the Cronbach alpha reliability coefficient obtained as a result of the CFA was calculated as 0.85. The test was applied to a sample of 619 7th graders and the KR-20 reliability coefficient was computed as 0.86. Model fit indices from the CFA (RMSEA, CFI, SRMR, chi-square, and chi-square/df) demonstrated a good fit to the data. The variance explained by the three factors was 22.6% for *knowing*, 22.3% for *applying*, and 18.0% for *reasoning*. Covariance among the factors ranged between 0.76 and 0.83.

Based on CFA results, six items (Items 5, 6, 9, 14, 16, and 35) were excluded due to low factor loadings. An additional three items (Items 11, 22, and 23) were removed to achieve balance across the cognitive domains and to reduce the total item count. The final version of the KET consisted of 27 items, with an equal distribution of 9 items for each cognitive domain.

The item discrimination indices ranged from 0.23 to 0.59, while the item difficulty indices ranged from 0.21 to 0.78. After item elimination, the mean item discrimination index was 0.38, and the mean item difficulty index was 0.42.

1.4 Conclusion and Discussion

The mean test score was calculated as 14.65, with a standard deviation of 6.79. The average item discrimination index of 0.38 is consistent with previous studies on similar topics (İdin & Aydoğdu, 2016; Güngörmez & Akgün, 2018; Yüzbaşıoğlu & Kurnaz, 2023). Similarly, the average item difficulty index of 0.42 aligns with findings from related research (Aydoğdu et al., 2012; Açıkgöz & Karşı, 2015; İdin & Aydoğdu, 2016; Güngörmez & Akgün, 2018; Yüzbaşıoğlu & Kurnaz, 2023).

The test's reliability coefficient ($KR-20 = 0.86$) exceeds the generally accepted threshold of 0.70, confirming its reliability and comparability to similar assessments in the literature (Akınoğlu & Tandoğan, 2007; Günaydın, 2010; Aydoğdu et al., 2012; Akbulut & Çepni, 2013; Karatay & Doğan, 2016; Özden & Yenice, 2017; Güngörmez & Akgün, 2018; Yüzbaşıoğlu & Kurnaz, 2023).

This study successfully developed a criterion-referenced test capable of assessing students' scientific literacy in terms of their cognitive processes of knowing, applying, and reasoning. The three-dimensional structure of the test distinguishes it from other assessments in the field and contributes to a more comprehensive evaluation of students' conceptual understanding and scientific thinking skills.

Kuvvet ve Enerji Ünitesi Bilme Uygulama Akıl Yürütme TIMSS Bilişsel Alanları Temelinde Bir Test Geliştirme Çalışması

Gonca Gül Domaç^{1*}  ve Salih Ateş² 

¹Fen Bilgisi Öğretmeni, Millî Eğitim Bakanlığı, Ankara, Türkiye

²Gazi Üniversitesi, Gazi Eğitim Fakültesi, Ankara, Türkiye

ÖZ

Bu araştırmada 7. sınıf Fen Bilimleri öğretim programı ünitelerinden biri olan Kuvvet ve Enerji Ünitesi kazanımlarını içeren TIMSS bilişsel alanları temelinde Kuvvet Enerji Ünitesi Testi (KET) geliştirilmiştir. Testte öğrencilerin bireysel başarılarını belirlemek için kriter olarak TIMSS (2019) tarafından belirlenen “bilme”, “uygulama” ve akıl yürütme” bilişsel alanlarında hazırlanan maddeler kullanılmıştır. Test maddelerinin kapsam geçerliği ve yapı geçerliğinin sağlanmasında uzman görüşüne başvurulmuştur. 35 maddeden oluşan testin pilot uygulaması Ankara merkez ilçelerindeki 371 tane 8.sınıf öğrencisi ile gerçekleştirilmiştir. Test maddelerine DFA uygulanmış, madde ayırt ediciliği ile güçlük indeksleri hesaplanmış ve analizler sonucunda 8 madde testten çıkarılarak 9 bilme, 9 uygulama, 9 akıl yürütme bilişsel alanına ait 27 maddeden oluşan bir test elde edilmiştir. Test maddelerinin madde ayırt edicilik indeksi nokta çift serili korelasyon yöntemi hesaplanmış ve maddelerin ortalama ayırt edicilik değeri 0.38; güçlük düzeyi 0.42 olarak belirlenmiştir. Testin pilot uygulamada KR – 20 güvenirlik katsayısı 0.86 olarak, DFA sonucunda elde edilen Cronbach alfa güvenirlik katsayısı ise 0.85 olarak hesaplanmıştır. Test 7.sınıflardan oluşan 619 kişilik örnekleme grubuna uygulanmıştır ve KR- 20 güvenirlik katsayısı 0.86 olarak hesaplanmıştır. Çalışmada öğrencilerin Kuvvet ve Enerji Ünitesi’nde bilme, uygulama ve akıl yürütme düzeylerini belirlemede kullanılabilecek geçerli ve güvenilir kriter referanslı bir test elde edilmiştir.

Anahtar Kelimeler: Kuvvet ve enerji, test geliştirme, TIMSS, bilme uygulama ve akıl yürütme bilişsel alanları

MAKALE BİLGİ

Makale Tarihi:

Alındı: 27.05.2025

Düzeltilmiş hali alındı: 21.06.2025

Kabul edildi: 25.06.2025

Çevrimiçi yayınlandı: 28.06.2025

Makale Türü: Araştırma Makalesi

Bu Makaleye Atıfta Bulunmak İçin: Domaç, G. G., & Ateş, S. (2025). Kuvvet ve enerji ünitesi bilme, uygulama, akıl yürütme TIMSS bilişsel alanları temelinde bir test geliştirme çalışması. *Journal of Individual Differences in Education*, 7(1), 54–74., DOI: 10.47156/jide.1707289

1. Giriş

Bilim evreni ve evrendeki olguları ve varlıkları bilimsel yöntem kullanılarak sistemli olarak inceleyen bulguların modellerle ifade edildiği ilgi, istek ve ihtiyaç temelli bir insan uğraşdır (Ateş, 2019). Ortaokul yıllarında öğrenciler bilim ile ilgili algılarının gelişmesine bağlı olarak kendileri ve günlük yaşamları ile ilgili bilinçli olarak kararlar almaya özendirilir. Bu sayede yetişkin olduklarında bilimsel olanı kurgudan ayıran, toplumsal, çevre, sağlık veya ekonomik temelli sorunların altında yatan bilimsel nedenleri anlayabilen bireyler olarak yetiştirilmeleri amaçlanır (Mullis ve Martin, 2017). 21. Yüzyılda fen okuryazarlığı tüm öğrenciler için vazgeçilmez bir hedef halindedir ve bu eğitim sayesinde öğrenciler içinde bulundukları dünyayı anlamanın yanı sıra bilimin prensiplerini yaşamlarında uygulamayı da öğrenirler (NRC, 2015).

Fen okuryazarı olan bir birey doğal dünyayı yakından tanımakla birlikte hareket, canlılardaki hücrelerin görevleri ve maddenin özellikleri gibi temel konularda ilke, yasa, teori ve kavramları anlayabilir (NRC, 2015). Fen eğitiminde yeni nesil fen bilimleri standartlarına (NGSS) göre fen okuryazarlığı; bilimin içeriği (disiplinler arası temel fikirler), bilimsel bilgiye ulaşma ve anlama yolları (bilim ve mühendislik uygulamaları) ve bilimlerin disiplinler arası evrensel kavramlar aracılığıyla birbirine bağlanması olmak üzere üç boyutta tanımlanmıştır (NSTA, 2024). Burada da görüldüğü gibi bilimsel bilgi birçok bilimsel okuryazarlık tanımında olduğu gibi burada da temel bileşen olarak

² Sorumlu yazar adresi: Millî Eğitim Bakanlığı, Ankara, Türkiye
e-posta: goncaguldomaç@gmail.com

karşımıza çıkmaktadır. Bu da bilimde içerik bilgisinin önemini ve öğrencilerin fen okuryazarlık düzeylerinin belirlenmesinde mutlaka dikkate alınması gerektiğini ortaya koymaktadır.

Fen okuryazarlığının başlıca bileşenleri arasında bilimsel bilgi, bilimin yöntem ve süreç becerileri, bilim ve araştırma sorgulamanın doğası gibi boyutlar bulunmaktadır (O' Hearn 1976; Miller, 1983; Fives, Huebner, Birnbaum ve Nicolich 2014). Türkiye'de öğretim programlarında fen okuryazarlığı ilk kez 2005 yılında hazırlanan öğretim programında yer almıştır (MEB, 2006). Bu programda fen okuryazarlığının bilgi boyutu "anahtar fen kavramları" şeklinde tanımlanmıştır. Daha sonra uygulanan 2013 ve 2018 öğretim programlarında da fen okuryazarlığının geliştirilmesi temel hedeflerden biri olarak ele alınmıştır. Fen okuryazarlığının uluslararası düzeyde ölçüldüğü sınavlarda da bilgi ve beceri boyutlarının her zaman yer aldığı görülmektedir. Ekonomik İşbirliği ve Kalkınma Örgütü (OECD) tarafından hazırlanan Türkiye'de de uygulanan Uluslararası Öğrenci Değerlendirme Programı'nda (PISA) fen okuryazarlığı; yeterlikler, bilimsel bilgi ve gerçek yaşam bağamları üzerinden tanımlanmıştır (OECD, 2019). Uluslararası Eğitim Başarılarını Değerlendirme Kuruluşu (IEA) tarafından yapılan Türkiye'de de uygulanan Uluslararası Matematik ve Fen Eğilimleri Araştırması (TIMSS) sınavında öğrenciler, öğrenme boyutu ve düşünme süreçlerini içeren bilişsel boyut olmak üzere iki boyutta değerlendirilmektedir (Mullis ve Martin, 2017). Bilimsel bilgi içeriği bu uygulamada da ön planda tutulmaktadır.

Öğrencilerin aldıkları fen bilimleri eğitimi ve yaşamlarında edindikleri deneyimin sonucunda; bilim insanları tarafından ortaya konmuş kuram, yasa, ilke ve kavramlar gibi bilimsel bilgileri ve bilim insanlarının dünyanın doğal düzenini keşfetmeleri sürecinin gerçekleşme şeklini de öğrenmeleri beklenmektedir (NRC, 2015). Bu amaçlarla fen bilimleri öğrenimi sürecinde öğretim yöntem ve tekniklerinin geliştirilmesi ve öğretim programlarının yenilenmesi gibi çalışmalar yapılmaktadır. Bütün bunlar öğretim sonuçlarını değerlendirme aşamasında gerçekleştirilen ölçme faaliyetlerini de etkilemektedir. Türkiye'de 2005 yılında uygulanmaya başlanan Fen Bilimleri Dersi Öğretim Programı ve sonrasında uygulanan tüm öğretim programlarında öğrencilerin bilgiyi doğrudan almaları yerine öğrenmenin merkezinde olmaları, bilgiyi yapılandırmaları, akıl yürütme, argümantasyon, bilimsel süreç ve bilimsel yöntemler konularındaki becerilerinin geliştirilmesi beklenmektedir (MEB, 2006; 2013; 2018; 2024). Dolayısıyla fen okuryazarlığının bilgi boyutunun ölçülmesinde yalnızca ezberlenen bilgi değil, öğrencilerin bilgiyi yeni durumlara uygulama yollarının ve yeni bağlamlarda akıl yürütme yollarının da dikkate alınması gerekmektedir. Bu durumda öğretim programındaki kazanımların ve bu kazanımların içerdiği becerilerin sınıflandırılmasında bir araca ihtiyaç duyulmaktadır. Kazanımlar doğrultusunda ölçülmesi beklenen bilgi ve becerilerin sınıflandırılmasında TIMSS (2019) tarafından belirlenen "bilme", uygulama" ve "akıl yürütme" şeklindeki bilişsel alanları sınıflandırılma taksonomisi en yaygın kullanılan sınıflandırma olarak göze çarpmaktadır. Bu çalışmada Fen Bilimleri dersi öğretim programındaki kazanımların bilişsel süreçlerinin ölçülmesinde TIMSS'in bilme, uygulama ve akıl yürütme şeklindeki taksonomisi kullanılmıştır.

TIMSS uygulamasında fen bilimleri alanında bilimsel bilginin ölçülmesinde öğrencilerin "bilme", "uygulama" ve "akıl yürütme" bilişsel süreçlerindeki düzeyleri belirlenir (Suna, Şensoy, Parlak ve Özdemir 2020). Bu bilişsel süreçlerden "bilme" boyutunun içeriğinde; öğrencilerden süreç, olgu, kavramlar, ilişkiler ve araçlar konularındaki bilgilerini kullanarak tanımlama, hatırlama ve örnek verme gibi zihinsel işlemleri gerçekleştirmeleri beklenir. "Uygulama" bilişsel sürecinde aynı bilgileri farklı bir problem durumunda kullanma, karşılaştırma, sınıflandırma, ilişkilendirme, açıklama, modeller kullanma veya bilgiyi yorumlama gibi zihinsel işlemleri gerçekleştirmeleri beklenir. "Akıl yürütme" bilişsel alanında ise bu bilgileri kullanarak çeşitli problem durumlarını analiz etmeleri, bunlardan sonuç çıkarmaları, farklı durumlarla sentez yapabilmeleri, bir araştırmayı planlayabilmeleri, genelleme ve değerlendirme yapabilmeleri gibi zihinsel işlemleri gerçekleştirmeleri beklenir (Mullis ve Martin, 2017).

Alan yazın incelendiğinde fen bilimleri dersi kapsamında Kuvvet ve Enerji Ünitesiyle ilgili geliştirilen bazı testlerin bulunduğu görülmektedir. Bunların; bilimsel süreç becerilerini ölçmeye yönelik (Aydoğdu, Tatar, Yıldız ve Buldur, 2012; Karatay ve Doğan, 2016), bağlam temelli olarak geliştirilen

(Yüzbaşıoğlu ve Kurnaz, 2023), alternatif ölçme-değerlendirme tekniklerinin kullanıldığı (Açıkgöz ve Karşı, 2015), kavramsal anlama düzeylerini ve kavram yanlışlarını tespit etmeye yönelik testler (Özden ve Yenice, 2017) olduğu görülmektedir. Bunların dışında öğrencilerin öğretim programı kazanımlarına ulaşma düzeylerinin ölçüldüğü testler de geliştirilmiştir (Akınoğlu ve Tandoğan, 2007; Akbulut ve Çepni, 2013; Güngörmez ve Akgün, 2018).

Bu testlerin öğretim programı kazanımlarında hedeflenen kavramsal ve yöntemsel bilgilerin ölçülmesine yönelik olduğu görülmektedir. Fakat bu testlerin öğrencilerin hedeflenen kavramsal ve yöntemsel bilgilerin hangi bilişsel süreçteki (bilme, anlama, uygulama, sentez ve değerlendirme vb.) düzeylerinin ölçülmesine yönelik olduğu belirli değildir. Uluslararası değerlendirmelerin yapıldığı PISA veya TIMSS gibi sınavların hazırlanma süreçlerinde de test maddelerini belirli alanlara göre sınıflandırma işlemlerinin yapıldığı ve öğrencilerin düzeylerinin bu sınıflandırmalara göre hazırlanan maddelere verdikleri yanıtlar üzerinden belirlendiği görülmektedir. Alan yazınındaki çalışmalarda öğrencilerin Kuvvet ve Enerji Ünitesi bağlamında “bilme”, “uygulama” ve “akıl yürütme” bilişsel süreçlerindeki düzeylerini tespit etmekte kullanılabilecek bir teste rastlanmamıştır. Ayrıca öğrencilerin bireysel başarılarının değerlendirilmesi amacıyla hazırlanan maddelerde TIMSS (2019) bilişsel alanları ve alt kategorileri kullanılarak sınıflandırma yapılan herhangi bir çalışmaya da rastlanmamıştır. Belirtilen nedenlerle öğrencilerin başarılarını değerlendirmede “bilme”, “uygulama” ve “akıl yürütme” bilişsel süreçlerinin kriter olarak belirlendiği bu test geliştirme çalışmasının alan yazınına katkı sağlayacağı düşünülmektedir.

Bir testin öğretim programlarındaki kazanımlar temel alınarak ve TIMSS (2019) tarafından belirlenen “bilme”, “uygulama” ve “akıl yürütme” bilişsel süreçleri içerecek şekilde hazırlanması bu testin kriter referanslı bir test olduğu anlamına gelmektedir. Kriter referanslı testler öğrencilerin belirli bir konu alanındaki bireysel ilerleme düzeylerini belirlenmede, öğrenme eksiklerini tespit etmede, öğretim programlarının etkinlik düzeylerini belirlemede, eğitim ve sosyal faaliyet gibi programlarda, çeşitli sertifika programlarında ve lisans verme amaçları için kullanılabilmektedir (Hambleton ve Eignor, 1979).

1.1 Amaç ve Önem

Bu araştırmanın amacı Fen Bilimleri Dersi 7. sınıf Kuvvet ve Enerji Ünitesi’nin kazanımları bağlamında öğrencilerin “bilme”, “uygulama” ve “akıl yürütme” bilişsel süreçlerindeki düzeylerini ölçmek için kriter referanslı bir test geliştirmektir.

Kriter referanslı testlerden elde edilen puanlar testin uygulandığı öğrencilerin neler yapabileceği konusunda betimleme yapabilmeyi sağlarken öğrenme etkinliklerinin sonuçları ile ilgili kararlar verebilmeyi, öğretim programlarının etkinliğini değerlendirmeyi de sağlar. Ancak bu testler geliştirilirken amaç öncelikle öğrencilerin kendi bireysel başarılarını değerlendirebilmektir (Hambleton ve Eignor, 1979). Buradan hareketle öğrencilerin belirli konu alanlarındaki bireysel başarılarının değerlendirilmesinde kriter referanslı testlerin kullanılmasının uygun olacağı düşünülmektedir. Öğrencilerin bireysel başarılarının değerlendirilmesini amaçlayan uluslararası sınavların yapısına bakıldığında da çeşitli sınıflandırmaların yapıldığı (bilme, uygulama, akıl yürütme bilişsel alanları gibi) ve öğrenci başarısının bu kriterlere göre belirlendiği görülmektedir. Bununla birlikte öğrenci başarısı ölçülürken belirli konu alanlarının temel alındığı da görülmektedir. Bu test Kuvvet ve Enerji Ünitesinin bilişsel olan tüm kazanımları dikkate alınarak hazırlanmış ve kapsam geçerliği sağlanmaya çalışılmıştır. Bu testin geliştirilmesi; testin bilme, uygulama ve akıl yürütme olmak üzere üç boyuttan oluşması ve önceki çalışmalarda çok sık rastlanmayan öğrencilerin değişik bilişsel süreçlerdeki düzeylerini ölçebilmeye imkân verebilecek olması açısından da önem taşımaktadır.

2. Yöntem

2.1. Araştırma Deseni

Bu araştırmada tarama modeli kullanılmıştır. Tarama modeli belirlenmiş bir evrenin özelliklerini, mevcut durumunu ve özelliklerini açıklamaya ve genelleme yapmaya imkân tanıyan bir yöntemdir (Kish, 1965). Tarama modelinin iki türü vardır ve bunlardan biri olan kesitsel tarama yöntemi bu araştırmada kullanılmıştır. Evreni temsil etmek üzere belirli bir zaman aralığında evrenden seçilen örnekleme başarı testi uygulanmıştır.

2.2. Evren ve Örneklem

Araştırmanın evrenini Ankara ili merkez ilçelerdeki ortaokullar oluşturmaktadır. Bu evreni temsil etmek üzere belirlenen örneklem ise seçkisiz örnekleme yöntemiyle oluşturulmuştur. Seçkisiz örnekleme yönteminde her bir kişinin örnekleme dâhil olma ihtimali birbirine eşittir (Fraenkel, Wallen ve Hyun, 2012). Örneklemen evreni temsil etme gücü diğer yöntemlerden daha yüksek olduğu için genelleme yapmaya da imkân sağlar (Büyüköztürk, 2010). Ankara ili merkezinde bulunan sekiz ilçe belirlendikten sonra bu ilçelerdeki ortaokullar rastgele olarak belirlenmiş ve her okuldan ikişer tane 8. sınıf rastgele olacak şekilde belirlenmiştir. Pilot uygulama testin kapsamını oluşturan üniteyi daha önceden öğrenmiş olma koşulunu sağlayan 8.sınıf öğrencileri ile yürütülmüştür. Test maddelerinin uygulanacağı örneklem büyüklüğünün belirlenmesinde alan yazındaki çeşitli öneriler değerlendirilmiştir. Bryman ve Cramer (2001), örneklem büyüklüğünün test madde sayısının beş katı; Nunnally (1978), test maddelerinin on katı kadar olmasının gerektiğini belirtirken; Cohen ve Cohen (1983), her test maddesinin en az on katılımcıya uygulanması gerektiğini (Akt. Uyumaz ve Sırgancı, 2020), Stevens (2002), testin her maddesinin 5 ile 10 arasında katılımcıya uygulanmasını önermektedir. Bu öneriler doğrultusunda araştırmada her bir test maddesinin yaklaşık 10 katı kadar öğrenciye uygulanmasına karar verilmiş ve her okuldan yaklaşık 50 öğrenci alınarak araştırma örneklemi 400 kişi olacak şekilde tasarlanmıştır. Bu öğrencilerden bazıları teste katılmadığı için testin pilot uygulaması 371 kişilik bir örnekleme uygulanmıştır. Pilot uygulama sonucunda geliştirilen testin uygulamasının gerçekleştirildiği araştırmanın evrenini Ankara ili Çankaya ilçesindeki Milli Eğitim Bakanlığına bağlı ortaokullarda okuyan 7. sınıf öğrencileri oluşturmaktadır. Bu evreni temsil etmek üzere seçilen örneklem seçkisiz tabakalı örnekleme yöntemi ile belirlenmiştir. Bu işlemde farklı alt grupların yeterince temsil edilmesi için evren belirli gruplara ayrılır daha sonra bu grupların her birine bireyler seçilir (Gravetter ve Forzano, 2018). Testin uygulanacağı okullar Çankaya İlçe Milli Eğitim Müdürlüğünden alınan beyana dayalı olarak bulundukları bölgede birbirine yakınlık esasına göre yedi farklı tabaka oluşturacak şekilde düzenlendikten sonra test Ankara ili Çankaya ilçesindeki ortaokullarda öğrenim gören 619 7. sınıf öğrencisine uygulanmıştır.

2.3. Veri Toplama Aracı

Bu araştırmada veri toplamak amacı ile kriter referanslı bir test olarak geliştirilen Kuvvet ve Enerji Ünitesi Testi (KET) kullanılmıştır. Kriter referanslı testlerin ilk tanımları Glaser (1963) ve Popham (1978) tarafından yapılmıştır (Hambleton ve Eignor (1979). Glaser'a (1963) göre, norm referanslı testlerde bireylerin testteki başarı düzeyleri teste katılan grubun diğer bireyleriyle karşılaştırılarak yorumlanır, kriter referanslı testlerde ise bireyin başarısı gruptaki teste katılan diğer bireylerden bağımsız bir şekilde belirli davranışsal ifadeler veya bunların tamamına dayalı olmaktadır (Jackson, 1970). Popham (1978), kriter referanslı testler bireylerin önceden belirlenmiş bir alandaki düzeylerini belirlemek için kullanıldığını ifade eder (Akt. Hambleton ve Eignor, 1979).

KET araştırmacılar tarafından TIMSS (2019) yapısında tanımlanan “bilme”, “uygulama” ve “akıl yürütme” bilişsel alanlarına uygun olacak şekilde hazırlanmıştır. Kriter referanslı bir test olan KET geliştirilirken 2018 Fen Bilimleri Dersi Öğretim Programındaki 7.sınıf kazanımları incelenmiştir. Bu kazanımlar öncelikle TIMSS bilişsel alanlarına göre sınıflandırılmış ve ilgili kazanım ve maddenin yazılması düşünülen bilişsel alanlar belirlenmiştir. Bu amaçla TIMSS sınavlarında yer alan ve yayımlanan sorular her birinin ait olduğu bilişsel alanın özellikleri ve alt kategorileri de dikkate

alınarak incelenmiştir. Madde havuzu Kuvvet ve Enerji Ünitesi kapsamındaki kazanımlar dâhilinde hazırlanan 35 maddeden oluşmaktadır. Pilot uygulama sonrasında yapılan analizlerin sonucunda sekiz madde test kapsamından çıkarılmıştır. Araştırmanın sonucunda öğretim programı kazanımları dâhilinde bilme, uygulama ve akıl yürütme bilişsel alanlarına yönelik her alandan 9 sorunun yer aldığı toplam 27 sorudan oluşan çoktan seçmeli bir test geliştirilmiştir.

2.4. Verilerin Analizi

Test maddelerinin pilot uygulaması sonucunda elde edilen veriler üzerinde doğrulayıcı faktör analizi yapılmıştır. Hambleton ve Eignor (1979), kriter referanslı testlerin madde istatistiklerinde norm referanslı testlerde olduğu gibi madde ayırt edicilik indekslerinin kullanılmasının (nokta-çift serili korelasyon katsayısı gibi) çeşitli sorunlara yol açtığını ifade ederek kriter referanslı testlerde bir test maddesinin birden fazla hedefi ölçüp ölçmediğini belirlemek amacıyla faktör analizi yapılmasını önermektedirler. Bu nedenle test maddelerinde yapı geçerliğinin sağlanması için test maddelerinde doğrulayıcı faktör analizi (DFA) yapılmıştır. DFA bir madde grubunun yapıyı tanımlayıp tanımlamadığını ortaya koymak için kullanılır ve amacı testlerin teorik yapılarının doğrulanmasıdır (Schumacker ve Lomax, 2004 s.16).

Crocker ve Algina (1986), bir testte ölçülen faktörlerin genel yapısı ve faktörlerin sayısı öncesinde belirlenmemiş olsa da eğer araştırmacının bir hipotezi bulunmuyorsa burada yapılan faktör analizinin açıklayıcı olduğunu, ancak ölçülen faktörlerin yapısı ile sayısı biliniyorsa ve bu tür hipotezler bulunuyorsa yapılması gereken faktör analizinin DFA olması gerektiğini belirtmişlerdir. Bu çalışmada geliştirilen testin bilme, uygulama ve akıl yürütme boyutlarını oluşturacak şekilde hazırlanması ve bu alanların belirlenmesinde uzman görüşlerinin de alınmış olması nedeniyle KET'nin üç faktörlü yapısının doğrulanması ve yapı geçerliğinin sağlanması için DFA yapılmıştır.

Kriter referanslı testlerde bir maddenin biçimi ve kapsamı, geliştirilen test için belirlenmiş ve özelleştirilmiş alanın hedeflenen boyutunu ölçmek için uygun olup olmadığının belirlenmesinde uzman görüşlerinin alınması önerilmektedir (Davis ve Diamond, 1974; Hambleton ve Eignor, 1979). Geliştirilen testin maddeleri önceden belirlenen “bilme”, “uygulama” ve “akıl yürütme” boyutlarından oluşmaktadır. KET'nin yapısının planlandığı şekilde üç faktörlü olup olmadığını belirlemek ve kapsam geçerliğine dair kanıt sağlamak için uzman görüşü alınmıştır. Maddeler alanında uzman üç akademisyene gönderilmiş uzmanlardan maddeleri bilme, uygulama ve akıl yürütme kategorilerinden hangisine dâhil olmaları gerektiğini belirtmeleri istenmiştir. Uzmanlardan elde edilen görüşlerden maddelerin bilişsel alanlara göre sınıflandırılması üzerinde uzmanların uyuşma oranı Fleiss Kappa katsayısı ile hesaplanmıştır.

Daha sonra bu maddeler üzerinde var olan yapıyı doğrulamak için DFA yapılmıştır. DFA yapılırken “R” programında yer alan (R Core Team, 2023) (Tierney ve Cook, 2023) ve MVN (Korkmaz, Göksülük ve Zararsız, 2014) paketlerinden; Aybek (2021) tarafından oluşturulan ve R tabanlı analiz işlemi yapılmasına olanak sağlayan web sitesinden ve JASP programından (2021) yararlanılmıştır.

Kriter referanslı testlerden elde edilen puanların sonuçları, norm referanslı testlerde yapıldığı gibi, birbiriyle veya teste katılan gruplar arasında karşılaştırma yapılmasını sağlar. Ancak bu durum kriter referanslı testlerin geliştirilmesinde öncelikli bir amaç değildir (Hambleton ve Eignor, 1979). Buradan yola çıkılarak kriter referanslı testlerin ve norm referanslı testlerin geliştirilme amacının farklı olduğunu görülmektedir. Geliştirilen kriter referanslı bir testin puanlarının eğer amaçlanıyorsa teste katılan öğrencilerin başarı düzeylerini birbiriyle karşılaştırılarak sıralamada da kullanılabileceği görülmektedir. Bu nedenle KET'de faktör analizine ek olarak nokta çift serili korelasyon yöntemi kullanılarak madde analizi de yapılmış ve madde ayırt edicilik indeksleri de hesaplanmıştır. Nokta çift serili korelasyon yöntemi kategorik ve sürekli değişkenlerin arasındaki ilişkinin ölçüldüğü zamanlarda kullanılan bir korelasyon yöntemidir ve testlerde madde analizinde ve madde ayırt ediciliğinin hesaplanmasında sıklıkla kullanılır. Burada elde edilen korelasyon katsayısı herhangi bir maddenin toplam test puanı ile arasındaki ilişkinin belirlenmesiyle bulunur (Crocker ve Algina, 1986). Test maddelerinin ayrıca güçlük indeksleri de hesaplanmıştır. Testin öğrencilerin başarı düzeylerini

sıralamada kullanılması durumunda teste dâhil edilen tüm maddelerin ayırt edicilik değerlerinin alan yazında önerilen referans aralığı içinde olduğu görülmüştür.

2.5. Etik İlkelerle Uygunluk

Bu araştırma Gazi Üniversitesi Eğitim Bilimler Enstitüsü'nde yürütülen bir doktora tezi verilerinden üretilmiştir. Araştırma Gazi Üniversitesi Rektörlüğü Etik Komisyonu'nun 16.08.2024 tarihli ve E.1018581sayılı 2024 - 1314 kod numaralı etik kurul izniyle gerçekleştirilmiştir. Araştırma için geliştirilen testlerin Milli Eğitim Bakanlığı'na bağlı resmi ortaokullarda yürütülme süreci Milli Eğitim Bakanlığı Telim ve Terbiye Kurulu Başkanlığı tarafından verilen 23.08.2024 tarihli MEB.TT.2024.001005 sayılı izin ile gerçekleştirilmiştir.

3. Bulgular

3.1. Testin Amacı ve Kapsamının Belirlenmesi

Bu araştırma için geliştirilen test 7. sınıf Fen Bilimleri Dersi Öğretim Programı Kuvvet ve Enerji Ünitesi kazanımları kullanılarak hazırlanmıştır. Bu testin amaçlarından biri Kuvvet ve Enerji konusunda öğrencilerin bilme, uygulama ve akıl yürütme bilişsel düzeyindeki sorulara cevap verme durumunun ölçülmesi olduğu için kapsam geçerliği oldukça önemli bir kriterdir. Öğretim programı incelendiğinde bazı kazanım açıklamalarının konu alanı ve gerektirdiği beceri açısından ana kazanımın yanı sıra yeni içerik alanlarına sahip olduğu görülmüştür ve bunlar alınan uzman görüşleri doğrultusunda yeni bir kazanım olarak belirtke tablosuna eklenmiştir. Bazı kazanımlarda ise öğrencilerden beklenen becerilerin eksiksiz ölçülebilmesi için belirsiz ifadelerde uzman görüşleri temelinde düzenleme yapılmıştır. Yeni kazanımlar ve düzenlemeler ile birlikte test kapsamını oluşturan diğer kazanımlar da Tablo 1'de belirtke tablosu içinde gösterilmiştir. Bu kazanımların ilgili olabileceği bilişsel alan belirlendikten sonra her bir kazanım için en az iki soru olacak şekilde maddeler yazılmıştır. Bir kazanım için uygulama bilişsel alanından hazırlanan bir madde olabileceği gibi kazanım incelendiğinde beklenen beceriler açısından kapsam geçerliğinin sağlanması açısından aynı kazanım için bilme bilişsel alanında da madde yazılmıştır. Benzer şekilde bir kazanım için akıl yürütme maddesi yazılırken TIMSS (2019) tarafından belirtilen akıl yürütme bilişsel alanı alt kategorileri de dikkate alınarak madde sayısı artırılmış ve bu kategoriler açısından kapsam geçerliğinin sağlanması amaçlanmıştır.

MEB (2018) Fen Bilimleri Öğretim Programındaki kazanımlar ve her biri için hazırlanan test maddesi sayılarının yer aldığı belirtke tablosu Tablo 1'de gösterilmiştir. Fen Bilimleri Öğretim Programındaki kazanımlardan biri olan "F.7.3.3.3 Hava veya su direncinin etkisini azaltmaya yönelik bir araç tasarlar." kazanımı tasarım, bilimsel süreç ve diğer psikomotor becerileri de gerektirdiği için testin kapsamı içine dâhil edilmemiştir. Test maddelerinin hazırlanması aşamasında bu belirtke tablosuna göre kazanımların bilme, uygulama ve akıl yürütme bilişsel alanlarında ölçülebileceği maddelerin yazım süreci gerçekleştirilmiştir.

3.2. Madde Havuzunun Hazırlanması ve Uzman Görüşünün Alınması

KET, belirtke tablosu hazırlandıktan sonra TIMSS(2019) bilişsel alanlarının ve alt kategorilerinin özellikleri de dikkate alınarak toplam 35 maddelik taslak test hazırlanmıştır. Bu maddelerin ait olduğu kazanımlar ile ilgili yapılan düzenlemeler için de uzman görüşleri alınmıştır. Test maddeleri hazırlanırken bilme, uygulama ve akıl yürütme bilişsel alanlarının alt kategorilerindeki kapsam geçerliğini sağlamak için her düzeyden maddenin yer almasına dikkat edilmiştir. Testin kriterlerini oluşturan "bilme", "uygulama" ve "akıl yürütme" bilişsel alanlarına ait madde örneklerine sırasıyla Resim1, Resim 2 ve Resim 3'te yer verilmiştir.

Tablo 1. Kuvvet ve Enerji Testi Belirtke Tablosu

Kazanımlar	Madde Sayısı
F.7.3.1.1. Kütleyle etki eden yer çekimi kuvvetini ağırlık olarak adlandırır. a. Ağırlığın bir kuvvet olduğu vurgulanır. b. Dinamometre kullanılarak ağırlık ölçümü yaptırılır. <i>Yeni kazanım: Dinamometre kullanarak çeşitli cisimlerin ağırlığını ölçer.</i>	2
F.7.3.1.2. Kütle ve ağırlık kavramlarını karşılaştırır.	6
F.7.3.1.3. Yer çekimini kütle çekimi olarak gök cisimleri temelinde açıklar. <i>Matematiksel bağıntılara girilmez.</i>	2
F.7.3.2.1. Fiziksel anlamda yapılan işin, uygulanan kuvvet ve alınan yolla ilişkili olduğunu açıklar. <i>Düzenleme: Bir cisim üzerine yapılan işin nelere bağlı olduğunu açıklar.</i> a. İşin birimi joule olarak verilir. b. Matematiksel bağıntılara girilmez.	3
F.7.3.2.2. Enerjiyi iş kavramı ile ilişkilendirerek, kinetik ve potansiyel enerji olarak sınıflandırır. <i>Düzenleme: Enerjiyi iş kavramı ile ilişkilendirerek, kinetik, kütle çekim potansiyel enerjisi ve esneklik potansiyel enerjisi olarak sınıflandırır.</i> a. Potansiyel enerji, çekim potansiyel enerjisi ve esneklik potansiyel enerjisi şeklinde sınıflandırılır. b. Potansiyel enerjinin kütle ve yüksekliğe, kinetik enerjinin kütle ve sürata bağlı olduğu belirtilir. <i>Yeni kazanım: Kütle çekim potansiyel enerjisinin ve kinetik enerjinin nelere bağlı olduğunu açıklar.</i> c. Matematiksel bağıntılara girilmez.	2
F.7.3.3.1. Kinetik ve potansiyel enerji türlerinin birbirine dönüşümünden hareketle enerjinin korunduğu sonucunu çıkarır.	11
F.7.3.3.2. Sürtünme kuvvetinin kinetik enerji üzerindeki etkisini örneklerle açıklar. <i>Düzenleme: Hareket eden cisme sürtünme kuvvetinin, hava direncinin ve su direncinin cismin kinetik enerjisi üzerindeki etkisini örneklerle açıklar.</i> a. Sürtünme kuvvetinin kinetik enerji üzerindeki etkisinin örneklendirilmesinde sürtünmeli yüzeyler, hava direnci ve su direnci dikkate alınır. b. Sürtünen yüzeylerin ısındığı, basit bir deneyle gösterilerek kinetik enerji kaybının ısı enerjisine dönüştüğü vurgulanır. <i>Yeni Kazanım: Sürtünen yüzeylerde kinetik enerjinin ısıya dönüştüğünü basit bir deneyle gösterir.</i>	2
F.7.3.3.3. Hava veya su direncinin etkisini azaltmaya yönelik bir araç tasarlar. a. Hava veya su direncinin farklı taşıtların tasarımındaki etkisine değinilir. b. Tasarımlar çizimle ortaya konulur, üç boyutlu bir ürüne dönüştürülmez.	0

Kazanım: F.7.3.1.2. Kütle ve ağırlık kavramlarını karşılaştırır.

Soru No: 7. Uygun ölçme araçlarıyla 2kg ve 5N şeklinde ölçülen iki değer sırasıyla hangi kavramlara aittir?

- A) Ağırlık- kütle
- B) Kuvvet – kütle
- C) Kütle – enerji
- D) Kütle – ağırlık

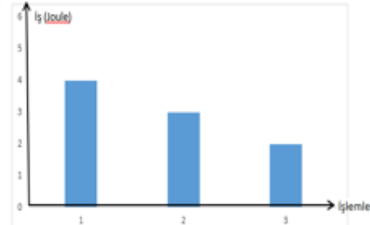
TIMSS (2019) Bilişsel Alanı: Bilme

Alt Kategori - Hatırlama/ Tanıma: Olguları, ilişkileri ve kavramları ayırt eder veya ifade eder. Belirli canlıların, materyallerin ve süreçlerin özelliklerini tanımlar; bilimsel araçların uygun kullanımlarını belirler; bilimsel kelime, sembol, kısaltma, birim ve ölçekleri doğru şekilde kullanır.

Resim 1. Bilme bilişsel alanına ait madde örneği

Kazanım: Bir cisim üzerine yapılan işin nelere bağlı olduğunu açıklar.

Soru No: 15. Bir cisim üzerinde üç farklı işlem yapıldığında, cisim üzerinde fiziksel olarak yapılan işlerin miktarı grafikteki gibi oluyor.



Buna göre, aşağıdakilerden hangisi yapılan bu işlemlerden biri olamaz?

- A) Cisim yerden alınarak aynı kuvvetle farklı yüksekliklere kaldırılmıştır.
- B) Cisim yerden alınarak aynı kuvvetle aynı yüksekliklere kaldırılmıştır.
- C) Cisim aynı ortamda farklı büyüklükte kuvvetler uygulanarak aynı mesafeler boyunca itilmiştir.
- D) Cisim aynı ortamda aynı büyüklükte kuvvetler uygulanarak farklı mesafeler boyunca itilmiştir.

TIMSS (2019) Bilişsel Alan: Uygulama

Alt Kategoriler:

Modeller Kullanma: Bir süreci, döngüyü, ilişkiyi veya sistemin gösterimini yapmak veya fen problemlerine çözümler bulmak için diyagram veya farklı bir model kullanır.

Bilgiyi Yorumlama: Metin, resim, tablo veya grafik formundaki bilgileri yorumlamak için fen kavramlarını kullanır.

Resim 2. Uygulama bilişsel alanına ait madde örneği

Kazanım: F.7.3.3.1. Kinetik ve potansiyel enerji türlerinin birbirine dönüşümünden hareketle enerjinin korunduğu sonucunu çıkarır.

Soru No: 30. Aşağıda sürtünmelerin ihmal edildiği bir ortamda bir yay ve bir tane kutu kullanılarak yapılan deneyin aşamaları verilmiştir.

- I. Yere dik şekilde sabitlenen bir yayın ucuna bir kutu konuluyor.
- II. Kutuya elle bastırılarak yayın bir miktar sıkışması sağlanıyor. Kutu yayın üstünde bir süre bu şekilde tutuluyor.
- III. Kutu serbest bırakılıyor ve düşey olarak yukarı doğru hareket ederek yaydan ayrılıyor.
- IV. Bir süre sonra kutu yayın üzerine düşüyor yay yeniden bir miktar sıkışıyor.



Bu deneyden yayın veya kutunun sahip olduğu kinetik ve potansiyel enerjisiyle ilgili hangi sonuç çıkarılabilir?

- A) Yayın üzerinde kutu tutulurken yay kinetik enerjiye sahip olur.
- B) Yayın üzerindeki kutu serbest bırakıldığında kutunun yere göre kütle çekim potansiyel enerjisi azalır.
- C) Kutu yaya doğru düşerken esneklik potansiyel enerjisi yere göre kütle çekim potansiyel enerjisine dönüşür.
- D) Kutu yayın üzerine ilk konduğunda ve son aşamada yayın üzerine düştüğünde yayda esneklik potansiyel enerjisi depolanır.

TIMSS (2019) Bilişsel Alan: Akıl yürütme

Alt Kategoriler:

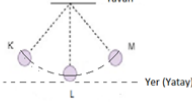
Sentez: Birçok farklı bileşenin veya ilişkili kavramların birlikte düşünülmesini gerektiren soruları cevaplar.

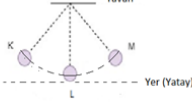
Sonuç çıkarma: Gözlemlere, kanıta ve/veya diğer bilgi kaynaklarına dayanarak çıkarımlar yapar. Soruları veya hipotezleri ele alan uygun sonuçlar üretir, sebep-sonuç ilişkilerini dikkate alır.

Resim 3. Akıl yürütme bilişsel alanına ait madde örneği

Test maddeleri hazırlandıktan sonra fizik eğitimi alan uzmanlarına gönderilerek uzman görüşüne başvurulmuştur. Uzman görüşü formunda her bir madde için kazanıma uygun olup olmadığı, TIMSS (2019) bilişsel alanlarından hangisine ait olduğu ve çoktan seçmeli test maddesi yapısına uygun olup olmadığı sorulmuştur. Ayrıca uzmanlara TIMSS açıklanan soru örneklerinden bilme, uygulama ve akıl yürütme bilişsel alanına ait soru örnekleri ve bilişsel alanların alt kategorileri ile ilgili açıklamalar da gönderilmiştir. Uzmanların bilişsel alanlar için yaptığı sınıflandırmanın uyuşma oranının % 63 ve şans oranının % 34.4 olduğu görülmüştür. Fleiss Kappa katsayısı 0.44 olarak hesaplanmıştır. Fleiss vd. (2003) 0.40 ve 0.75 arasındaki değerleri iyi olarak değerlendirmiştir ve bu değer uzmanlar arasında iyi düzeyde uyum olduğunu göstermektedir. Aşağıda görülen Resim 4'te uzmanlara gönderilen görüş formundaki örnek bir madde ve uzman görüşü yer almaktadır.

F.7.3.3.1. Kinetik ve potansiyel enerji türlerinin birbirine dönüşümünden hareketle enerjinin korunduğu sonucunu çıkarır.

27. Sürtünmesiz bir ortamda bir ipin ucuna bağlanan küçük bir cisim K noktasındaki konumdan serbest bırakıldığında sırasıyla  şekilde görülen yer seviyesinde kabul edilebilecek L noktası ve L den daha yüksekteki M noktalarına doğru hareket etmiştir. Top M noktasına ulaştığı anda geldiği yolu izleyerek tekrar K noktasına geri gelmiştir. Daha sonra aynı hareketi tekrarlamıştır.



Topun hareketi sırasında çeşitli aşamalarda sahip olduğu kinetik ve kütle çekim potansiyel enerjisiyle ilgili açıklamalardan hangisi doğrudur?

- A) K'den L'ye doğru hareket ederken yere göre kütle çekim potansiyel enerjisi azalır.
- B) L'den M noktasına ulaşmaya kadar her an kinetik enerjiye sahiptir.
- C) L'de yere göre hem kütle çekim potansiyel enerjiye hem de kinetik enerjiye sahiptir.
- D) M'de yere göre sadece kinetik enerjiye sahiptir.

Bu soru verilen kazanıma uygun mudur?	Bu sorunun TIMSS bilişsel alanlarından hangisinin kapsamında olduğunu düşünüyorsunuz?	Bu sorunun çoktan seçmeli test yapısına uygun olduğunu düşünüyor musunuz?
Evet (X) Hayır ()	Bilme () Uygulama (X) Akıl Yürütme ()	Evet (X) Hayır () Cevabınız hayır ise gerekçenizi yazabilir misiniz?

Microsoft Office User
Bu seçenek yerine daha uygun bir ifade yazılabilir. M noktasına ulaşmaya kadar ifadesinden "M noktası" dahil anlamı çıkmayabilir. Yani tartışmaya açık bir ifade gibi geldi..

Resim 4. Uzman görüşü örneği

3.3. Güvenirlik Analizi

KET'nin pilot uygulamasından elde edilen verilere Kuder-Richardson KR-20 ve KR - 21 formülü uygulanmıştır. Büyüköztürk, Akgün, Karadeniz, Demirel ve Kılıç (2013)'a göre KR-20 ve KR-21 formülü, bir test maddesine verilen cevaplar doğru/ yanlış ya da 1/0 olarak puanlanabilir olduğu durumlarda kullanılabilir. Ayrıca maddelerin güçlük indeksi incelenmiş ve bu değerlerin kolay, orta ve zor olarak sınıflandırıldığı görülmüştür. Büyüköztürk vd. (2013) tarafından da önerildiği gibi madde güçlük indeksinin ortalama güçlükte olmaması durumunda KR-20 formülünün uygulanmasına karar verilmiştir. Bu değer KR-20 değeri 0.70 ve üzerinde olan testlerin güvenilir olduğu söylenebilir (Büyüköztürk, 2010). KET 35 maddeye sahiptir ve testin KR-20 değeri 0.86 olarak hesaplanmıştır. Testin DFA sonucunda elde edilen Cronbach alfa güvenirlilik katsayısı ise 0.85 olarak hesaplanmıştır. Pilot uygulama sonunda 27 madde olarak geliştirilen testin 619 7.sınıf öğrencisine uygulanması sonucunda elde edilen KR-20 güvenirlilik katsayısı 0.86 olarak hesaplanmıştır.

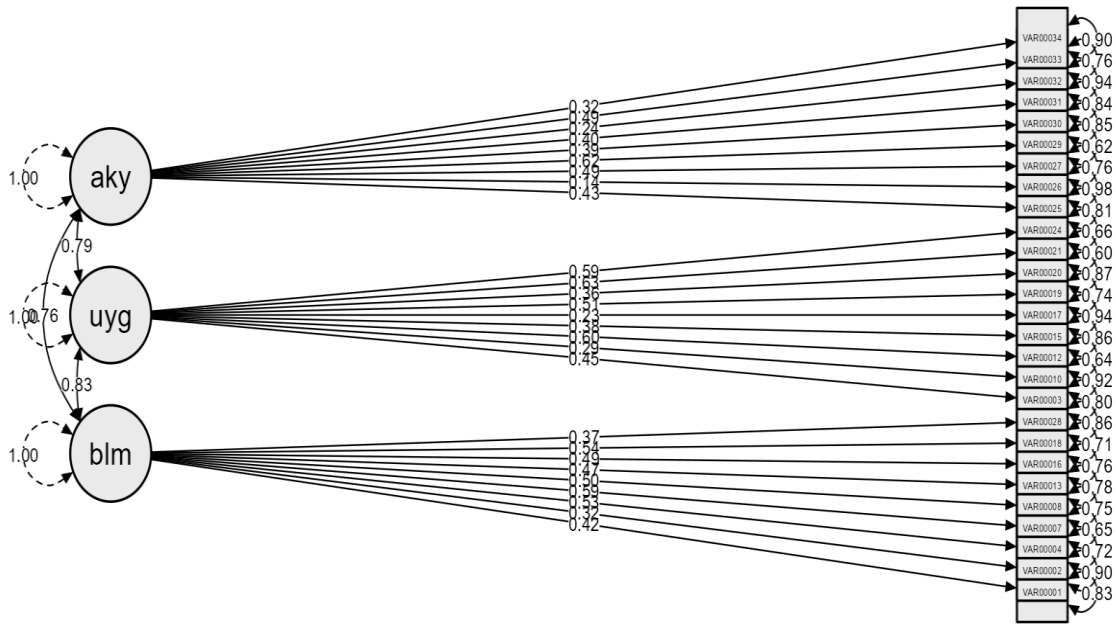
3.4. Doğrulayıcı Faktör Analizine ait Bulgular

Testin DFA sürecine başlamadan önce incelenen varsayımların ilkinde veri setinde kayıp değer ya da uç değer olmadığı görülmüştür. Öğrencilerin boş bıraktığı maddeler kayıp değer olarak işaretlenmiş ve analiz sürecinde kayıp değerlerle başa çıkma yöntemi olarak çift yönlü (pairwise) yöntem seçilmiştir. Çok değişkenli normallik varsayımını test etmek için kullanılan Mardia testi sonucunda verilerin çok değişkenli normallik varsayımını sağlamamasından dolayı Koğar ve Yılmaz Koğar'ın (2016) kategorik verilerde DFA uygulanırken Diyagonal Ağırlıklandırılmış En Küçük Kareler (DWLS) yöntemi kullanılması halinde daha düşük hatayla kestirim yapılabileceği önerilerinden yola çıkılarak analize DWLS ile devam edilmiştir. Çoklu bağlantı varsayımının kontrolü için maddeler arası korelasyonlar incelenerek bu problemin olmadığı belirlenmiştir.

Model veri uyumunu değerlendirmek için Yaklaşık Hataların Ortalama Karekökü (RMSEA), Karşılaştırmalı Uyum İndeksi (CFI), Standardize Edilmiş Artık Ortalamaların Karekökü (SRMR), ki-kare ve ki-kare/serbestlik derecesi kullanılmıştır. DFA sonucunda p değeri 0.01 olarak bulunmuştur. CFI, modeli değişkenler ve bağımsızlık modeli arasında ilişkinin olmadığı bir model ile karşılaştırır. Bu değer 0.90 ve üstü olduğunda model uyumunun iyi olduğunu gösterir (Tabachnick & Fidell, 2012). KET için yapılan DFA sonucunda CFI değerinin 0.96 olarak elde edilmesi model uyumunun iyi olduğunu göstermektedir. RMSEA değeri 0'a yaklaştıkça mükemmel uyumu, 0.05'ten küçük değerler mükemmel uyumu gösterir. Bu testte elde edilen RMSEA değerinin 0.40 olarak belirlenmesi modelin uyumunun iyi olduğunu göstermektedir. SRMR değerleri 0 ile 1 arasında değerler alır ve 0.08'den küçük değerler iyi uyumu 0.05'ten küçük değerler ise çok iyi uyumu gösterir (Brown, 2015). Bu

çalışmada yapılan DFA sonucunda elde edilen 0.53 değeri modelin uyumunun iyi olduğunu göstermektedir. Ki-Kare “kötü uyum indeksi” olarak da bilinmesi sebebiyle bu faktör analizinin sonucunda elde edilen 514.197 (321) değerinin küçük ve anlamlı olmaması model uyumun iyi olduğunu göstermektedir. Ancak büyük örneklemelerde ki-kare derecesinin anlamlı çıkma ihtimali vardır. Bu nedenle ki-kare/serbestlik derecesi (sd) alan yazında kullanılan uyum indekslerinden biridir. Bu değerin 2 veya daha küçük olması modelin uyumunun iyi olduğunu gösterir (Tabachnick ve Fidell, 2012). Kline (2019) ki-kare/sd değerinin küçük örneklemeler söz konusu olduğunda 2,5’in altında olmasının modelin iyi uyumlu olduğunu büyük örneklemeler söz konusu olduğunda ise 3’ün altında olmasının iyi uyumlu olduğunu belirtmiştir.

DFA sonucunda elde edilen modelin hata varyanslarının ve parametre kestirimlerinin yer aldığı yol diyagramı Şekil 1’de verilmiştir.



Şekil 1. DFA Modeline Ait Yol Diyagramı

Şekil 1’e bakıldığında faktör yüklerinin 0.14 ile 0.62 arasında ve hata varyanslarının 0.61 ile 0.94 arasında değiştiği görülmektedir. DFA sonucunda kestirim yapılan tüm iç tutarlılık katsayıları ile parametreler ve faktörlere ait ortalama açıklanan varyans Tablo 2’de verilmiştir.

Tablo 3’teki bulgulara göre tüm kovaryansların ve faktör yüklerinin istatistiksel olarak anlamlı olduğu görülmüştür ($p < 0.05$). Bu bulguya göre faktörlerin kendi altında yer alan maddelerin anlamlı bir yordayıcısı olduğu söylenebilir. Standartlaştırılmamış faktör yüklerinin 0.06 ile 0.31 arasında olacak şekilde değiştiği standart hatanın tüm maddeler için .01 olduğu belirlenmiştir. AVE değerleri incelendiğinde bilme faktöründe varyansın %22,6’sının; uygulama faktöründe varyansın %22,3’ünün; akıl yürütme faktöründe varyansın %18’inin açıklandığı görülmüştür. Güvenirlik katsayılarına bakıldığında Cronbach Alfa katsayıları incelendiğinde faktörler için 0.63 ile 0.71 arasında değiştiği ve tüm test için 0.85 olduğu belirlenmiştir. Cronbach Alpha değer aralığı; 0.84 ve üzerinde olması iç tutarlılık için kabul edilebilir bir değerdir (Sarmiento ve Costa, 2017). Diğer bir güvenilirlik katsayısı olan Omega katsayılarına bakıldığında faktörler için 0.62 ile 0.72 arasında değiştiği ve tüm test için 0.86 olarak hesaplanmıştır. Tüm bu bulgular değerlendirildiğinde modelin veri ile iyi uyum gösterdiği ve ölçülmek istenen yapının geçerlik ve güvenilirlik kanıtlarının sağlandığı sonucuna ulaşılmaktadır. Faktörlerin birbirleriyle gösterdiği kovaryansın 0.76 ile 0.83 arasında değiştiği ve bu değer istatistiksel olarak anlamlı olduğu görülmektedir.

Tablo 2. Faktör Yükleri, Faktörler Arası Kovaryanslar ile Geçerlik ve Güvenirlik Katsayıları

	B	SH_B	z	p	β	ε	AVE	α	ω
Bilme							.226	.712	.723
M01	.209	.012	17.346	< .001	.418	.825			
M02	.160	.012	13.485	< .001	.320	.897			
M04	.261	.012	21.502	< .001	.526	.723			
M07	.297	.012	23.887	< .001	.593	.649			
M08	.250	.012	20.712	< .001	.499	.751			
M13	.230	.011	20.275	< .001	.471	.778			
M16	.207	.010	19.979	< .001	.486	.764			
M18	.263	.012	22.062	< .001	.537	.711			
M28	.181	.011	15.819	< .001	.369	.864			
Uygulama							.223	.687	.707
M03	.179	.010	18.325	< .001	.445	.802			
M10	.142	.011	13.051	< .001	.288	.917			
M12	.298	.012	24.523	< .001	.601	.639			
M15	.188	.011	16.858	< .001	.376	.858			
M17	.111	.011	10.510	< .001	.235	.945			
M19	.253	.011	22.121	< .001	.508	.742			
M20	.163	.011	15.422	< .001	.359	.871			
M21	.310	.012	25.349	< .001	.634	.598			
M24	.293	.012	24.367	< .001	.587	.656			
Akıl Yürütme							.180	.632	.622
M25	.211	.012	17.867	< .001	.434	.812			
M26	.061	.010	6.187	< .001	.143	.980			
M27	.244	.012	20.066	< .001	.491	.759			
M29	.306	.013	23.450	< .001	.620	.616			
M30	.173	.010	16.519	< .001	.388	.849			
M31	.193	.011	17.090	< .001	.402	.838			
M32	.106	.010	10.461	< .001	.242	.941			
M33	.239	.012	20.138	< .001	.491	.759			
M34	.155	.011	13.656	< .001	.316	.900			
Tüm Test								.841	.852
bilme↔ uygulama	.830	.031	26.423	< .001					
bilme↔ akıl_yürütme	.756	.033	22.790	< .001					
uygulama↔akıl_yürütme	.790	.034	23.314	< .001					

B: Standartlaştırılmamış Faktör Yüğü; SH_B : Standartlaştırılmamış Faktör Yüğünün Standart Hatası; β : Standartlaştırılmış Faktör Yüğü; ε : Standartlaştırılmış Hata Varyansı; AVE: Ortalama Çıkarılan Varyans; α : Cronbach Alfa Katsayısı; ω : Omega Katsayısı

Yapılan DFA sonucunda KET'in 5., 6., 9., 14., 16. ve 35. maddeleri, faktör yükleri düşük olduğu için testten çıkarılmıştır. Bunların dışındaki 11., 22. ve 23. maddeler ise farklı nedenlerden dolayı çıkarılmıştır. Bu çalışma öğrencilerin farklı bilişsel süreçlerdeki soruları cevaplama düzeylerini karşılaştırmayı amaçladığı için bilme uygulama ve akıl yürütme bilişsel alanlarındaki soru sayılarının birbirine eşit olması hedeflenmektedir. Ayrıca kapsam geçerliği açısından yeterli madde bulunması durumunda fazla olan maddeler çıkarılabilmektedir çünkü bu test ortaokul 7. sınıf öğrencilerine uygulanacaktır ve onların dikkat süreleri de göz önünde bulundurularak testin madde sayısının çok yüksek olmamasına dikkat edilmiştir. KET'in 11., 22., ve 23. maddeleri uygulama bilişsel sürecindeki soru sayısının azaltılması için aynı kazanıma ait birden fazla madde yer alması sebeplerinden dolayı testten çıkarılmıştır. Bu maddeler çıkarılırken ait oldukları bilişsel alanın alt kategorileri dikkate alınmıştır. Örneğin uygulama alanında karşılaştırma sınıflama alt kategorisinde yeterli madde varsa buradan azaltma yapılmıştır. Aksi durum tercih edilmemiş ve bilişsel alanların ait olduğu alt

kategorilerin de testte yeterince kapsama dâhil edilmesi sağlanmıştır. Son durumda KET 9 bilme, 9 uygulama, 9 akıl yürütme alanından olmak üzere toplam 27 maddeden oluşmaktadır.

3.5. Madde İndeksleri

KET'in öğrenci başarılarını karşılaştırmada veya sıralamada kullanılmak istenmesi halinde test maddelerinin ayırt edicilik ve güçlük indeksleri de hesaplanmıştır. KET'in pilot uygulaması sonucunda elde edilen veriler nokta çift serili korelasyon yöntemi ile örneklemin tamamı üzerinden analiz edilmiş ve elde edilen madde ayırt edicilik indeksleri ve madde güçlüğü indeksleri Tablo 2'de gösterilmiştir. Öğrencilerin boş bıraktığı maddeler kayıp değer olarak işaretlenmiş ve analiz sürecine dâhil edilmemiştir.

Tablo 2. Madde ayırt edicilik indeksleri ve madde güçlük düzeyleri

Madde Numarası	Madde Ayırt Ediciliği (rpb)		Madde Güçlüğü (p)	
1	0.43	Çok iyi	0.48	Orta
2	0.31	İyi	0.46	Orta
3	0.43	Çok iyi	0.78	Kolay
4	0.46	Çok iyi	0.42	Orta
5	0.24	Düzenlenmeli	0.37	Orta
6	0.18	Düzenlenmeli	0.27	Zor
7	0.52	Çok iyi	0.49	Orta
8	0.46	Çok iyi	0.46	Orta
9	0.13	Çok zayıf	0.29	Zor
10	0.27	Düzenlenmeli	0.34	Orta
11	0.45	Çok iyi	0.73	Kolay
12	0.54	Çok iyi	0.54	Orta
13	0.46	Çok iyi	0.54	Orta
14	0.28	Düzenlenmeli	0.31	Orta
15	0.35	İyi	0.39	Orta
16	0.50	Çok iyi	0.72	Kolay
17	0.27	Düzenlenmeli	0.29	Zor
18	0.48	Çok iyi	0.55	Orta
19	0.48	Çok iyi	0.45	Orta
20	0.52	Çok iyi	0.65	Orta
21	0.59	Çok iyi	0.57	Orta
22	0.30	İyi	0.36	Orta
23	0.52	Çok iyi	0.53	Orta
24	0.56	Çok iyi	0.47	Orta
25	0.41	Çok iyi	0.33	Orta
26	0.15	Düzenlenmeli	0.20	Zor
27	0.43	Çok iyi	0.36	Orta
28	0.36	İyi	0.36	Orta
29	0.54	Çok iyi	0.37	Orta
30	0.35	İyi	0.21	Zor
31	0.34	İyi	0.28	Zor
32	0.23	Düzenlenmeli	0.20	Zor
33	0.43	Çok iyi	0.31	Orta
34	0.33	İyi	0.32	Orta
35	0.13	Çok zayıf	0.25	Zor
Ortalama	0.38		0.43	

35 maddeden oluşan testin madde ayırt ediciliği ve madde güçlüğü değerleri nokta çift serili korelasyon yöntemi belirlendikten sonra öncelikle madde ayırt ediciliği değerleri incelenmiştir.

Madde ayırt edicilik değerleri 0.20'nin altındaysa ayırt edicilik düşüktür ve maddenin testten çıkarılması önerilir. Ayırt edicilik düzeyi 0.20 ve 0.30 arasındaki değerlerde ise maddenin yeniden düzenlenerek teste alınması önerilir. 0.30 ve 0.40 arasındaki maddeler iyi düzeyde ayırt edici olarak değerlendirilir değişiklik yapılması zorunlu değildir, 0.40 ve üzerindeki maddeler ise çok iyi düzeyde ayırt edici madde olarak değerlendirilir (Crocker ve Algina, 1986). Madde gücü 0.20'nin altında bir değer olduğunda çok zor bir madde olur 0.50 olduğunda orta güçlükte bir madde olarak sınıflandırılır. Bu değer 0.80 ve üzeri olduğunda ise çok kolay bir maddedir (Crocker ve Algina, 1986). KET'de pilot uygulama sonrasında 5., 6., 9., 11., 14., 16., 22., 23. ve 35. maddeleri testten çıkarılmıştır. Pilot uygulama sonucunda madde ayırt edicilik indeksi 0.23 ve 0.59 arasında; madde güçlük indeksi 0.21 ve 0.78 arasında olan 27 maddelik bir test elde edilmiştir. Belirtilen maddeler çıkarıldıktan sonra ortalama ayırt edicilik değeri 0,38; güçlük düzeyi 0.42 olan bir test geliştirilmiştir.

4. Sonuç Tartışma ve Öneriler

4.1. Sonuç ve Tartışma

Bu çalışmada 7. sınıf öğrencilerinde fen okuryazarlığının bilgi ve beceri boyutlarının ölçülmesi amacıyla Kuvvet ve Enerji Ünitesi kazanımlarının temel alındığı öğrencilerin bilme, uygulama ve akıl yürütme bilişsel süreçlerine dayalı soruları çözebilme düzeylerini belirlemeye yönelik kriter referanslı bir test geliştirilmiştir. Kuvvet ve Enerji Ünitesi konularının ölçülmesine yönelik testler alan yazında yer almaktadır ancak bu testlerin kazanımların doğrudan ölçülmesine yönelik olduğu görülmektedir (Akinoğlu ve Tandoğan, 2007; Kıray, 2010; Öztürk, 2014; Akbulut ve Çepni, 2013; Güngörmez ve Akgün, 2018). Alan yazındaki bazı çalışmalarda da bilimsel süreç becerilerinin ölçmesi için geliştirilen testlerde Kuvvet ve Enerji konuları bağlam olarak kullanılmıştır (Aydoğdu vd., 2012; Karatay ve Doğan, 2016). Ancak Kuvvet ve Enerji konularının TIMSS (2019) tarafından tanımlanan “bilme”, “uygulama” ve “akıl yürütme” bilişsel süreçlerine dayalı soruları çözebilme durumlarının ölçülmesine yönelik bir araştırma bulunmamaktadır. Dolayısıyla araştırmada geliştirilen testin kapsamını oluşturan konu bu araştırmanın diğer test geliştirme araştırmalarından farklı olan yönlerinden biridir.

Test ile ölçülmek istenen öğrenci başarıları “bilme”, “uygulama” ve “akıl yürütme” bilişsel alanlarındaki başarı düzeyleri olarak tanımlandığı için kriter referanslı bir test geliştirilmiş olup pilot uygulama sonuçları da bu bağlamda değerlendirildiğinde testin bu yapısının doğrulanmasında Crocker ve Algina (1986) tarafından da önerildiği gibi DFA sonuçları dikkate alınmış ve modelin veri ile iyi uyum gösterdiği belirlenmiştir. Çalışmada geliştirilen testin pilot uygulamasında ortalaması 14,65 standart sapması 6,79 olarak hesaplanmıştır. Standart sapmanın yorumlanması testin koşullarına göre değişmekle birlikte başarı testlerinde genelde standart sapmanın büyük olması beklenir (Atılğan, Kan ve Aydın, 2018 s. 274). Testin pilot uygulaması sonucunda elde edilen ortalama madde ayırt edicilik değeri 0.38 olarak hesaplanmıştır. İdin ve Aydoğdu (2016) araştırmalarında geliştirdikleri “Kuvvet ve Enerji Ünitesi Başarı Testi”nin ortalama ayırt ediciliğini 0.31 olarak hesaplamışlar, Yüzbaşıoğlu ve Kurnaz (2023), 0.60 olarak; Güngörmez ve Akgün (2018), 0.520 olarak hesaplamıştır. Bu sonuçlar testin alan yazındaki diğer çalışmalar ile benzer şekilde kabul edilebilir bir ayırt edicilik düzeyinde olduğunu göstermektedir.

Testin ortalama madde güçlük indeksi ise 0.42 olarak hesaplanmıştır. Özçelik (1992), başarı testlerinde madde güçlük indeksinin 0.20 ile 0.80 arasındaki değerlerde olabileceğini belirtmiştir. Bu araştırmada geliştirilen testte de kolay orta ve zor olarak sınıflandırılan maddeler yer almaktadır. Alan yazındaki benzer çalışmalarda da testlerin ortalama güçlük indeksleri Aydoğdu vd. (2012) tarafından 0.54 olarak; Açıkgöz ve Karşı (2015) tarafından 0.44; İdin ve Aydoğdu (2016) tarafından 0.367; Güngörmez ve Akgün, (2018) tarafından 0.550; Yüzbaşıoğlu ve Kurnaz (2023) tarafından 0.49 olarak raporlanmıştır. Bu araştırmalarda da testlerin ortalama güçlüğü benzer düzeyde olduğu görülmektedir.

Kuvvet ve enerji konuları kullanılarak geliştirilen çeşitli başarı testlerinde güvenilirlik değerlerinin 0.70'in üzerinde olduğu görülmektedir (Akinoğlu ve Tandoğan, 2007; Günaydın, 2010; Aydoğdu vd.

2012; Akbulut ve Çepni, 2013; İdin ve Aydoğdu, 2016; Karatay ve Doğan, 2016; Özden ve Yenice, 2017; Güngörmez ve Akgün, 2018; Yüzbaşıoğlu ve Kurnaz, 2023). Testin 619 7. sınıf öğrencisine uygulanması sonucunda güvenilirlik değeri KR-20 formülü uygulanarak 0.86 olarak belirlenmiştir ve bu değerin test geliştirilen diğer araştırmalarda da olduğu gibi 0.70'in üzerinde bir değerde olduğu görülmektedir.

Çalışma Kuvvet ve Enerji Ünitesinde test geliştirmekten ziyade bu ünitenin kazanımlarının öğrenciler tarafından bilme, uygulama ve akıl yürütme bilişsel alanlarını gerçekleştirme düzeylerinin belirlenmesi amacına yönelik bir çalışmadır. Dolayısıyla maddelerin hazırlık sürecinde de doğrudan odaklanılan kavramlar, kazanımları Anderson ve Krathwohl (2001) tarafından revize edilen Bloom taksonomisine göre sınıflandırmak değil öğrenciler bu kazanımı gerçekleştirirken onlardan hangi bilişsel alanlardaki becerileri gerçekleştirmesinin beklendiğidir. Bu nedenle kazanımlar tek bir bilişsel sürece birebir yerleştirilmeyip her birinde öğrenciden beklenen beceri/beceriler belirlenip belirtke tablosu buna göre oluşturulmuştur. Örneğin, Fen Bilimleri Dersi Öğretim Programı'nda yer alan "F.7.3.1.2. Kütle ve ağırlık kavramlarını karşılaştırır." kazanımının ölçülmesinde sadece karşılaştırma yaparak "uygulama" bilişsel alanının ölçülmesi değil, kütle ve ağırlık kavramlarının temel düzeyde ölçümüyle ilgili "bilme" becerileriyle ilgili maddeler de hazırlanmıştır. Bu ve benzeri işlemlerin sonucunda öğrencinin ünite sonunda gerçekleştirmesi beklenen beceriler daha kapsamlı ve öğretim programlarıyla uluslararası kurumların raporlarında belirtildiği gibi fen okuryazarı bireylerde bulunması gerektiği vurgulanan üst düzey düşünme becerileri de ölçülecek şekilde bir test geliştirilmiştir. Bilişsel alan sınıflandırmasının ve alt kategorilerin de dikkate alınarak ölçümün genişletilmesi bu testi alan yazındaki diğer testlerden farklı kılan diğer bir özelliğidir.

Bu çalışmada geliştirilen KET'in aynı konuda geliştirilen diğer testlerden farklı olan yönlerinden bir diğeri de üç boyutlu olarak tasarlanmış olmasıdır. Testi geliştirmenin amacı 7.sınıf öğrencilerinin Kuvvet ve Enerji konusundaki bilme, uygulama ve akıl yürütme düzeylerinin ölçülmesinde kullanılacak bir ölçme aracının bulunmaması olduğu için hazırlık aşamasında madde yazımı ve bu bilişsel alanlar ve alt kategorileri temelinde hazırlanmıştır. Testin pilot uygulamasının tamamlanmasından sonra test yapısını doğrulamak için yapılan DFA sonuçları testin üç boyutlu yapısını oluşturan modelinin uyumunun iyi olduğunu göstermektedir. DFA sonuçları ile testin kullanılabilirliğini sağlamak üzere yapılan analizler sonucunda 35 maddelik havuzdan sekiz madde çıkarılmış ve sonuçta dokuz tane bilme, dokuz tane uygulama, dokuz tane akıl yürütme bilişsel alanının ölçülmesine yönelik 27 maddelik bir test geliştirilmiştir. Araştırma ile kapsam ve yapı geçerliği sağlanmış, öğrenci başarısını değerlendirmede uluslararası kriterleri içeren, öğrencilerin bireysel başarılarını çeşitli boyutlar açısından belirleyebilmede ve aynı zamanda başarılarını sıralamada da kullanılabilecek bir ölçme aracı elde edilmiştir.

4.2. Öneriler

Bu çalışmada geliştirilen test çoktan seçmeli test maddeleri kullanılarak hazırlanmıştır. Araştırmacılar aynı konuda aynı bilişsel alanların farklı ölçme formatları kullanılarak ölçülmesini sağlayan bir ölçme aracı geliştirebilir. Çalışmanın örneklemini Ankara merkez ilçeleri ile sınırlıdır. Bu nedenle farklı örneklemelerde de testin uygulanması sağlanabilir. Testin pilot uygulamasına katılan öğrencilerin farklı okullarda olmaları nedeniyle test kapsamını oluşturan konuları öğrenme sürecinin birbirinden farklı olması bu çalışmanın sınırlılıklarından biridir. Araştırmacılar Kuvvet ve Enerji konusunun farklı öğrenme yöntemleri uygulanarak öğrenildiği sınıflarda uygulanan öğrenme yönteminin etkililiğini belirlemek veya öğrenme yöntemlerinin öğrencilerin bilme, uygulama ve akıl yürütme düzeylerine etkilerini belirlemek için çalışmalar yapabilirler.

Etik Beyannamesi

Araştırmada Yükseköğretim Kurumları ve Milli Eğitim Bakanlığı Bilimsel Araştırma ve Yayın Etiği Yönergesinde yer alan maddeler çerçevesinde gerçekleştirilmiştir. Araştırmanın etik kurul bilgileri ve araştırma izinleri yöntem bölümünde belirtilmiştir. Herhangi bir etik ihlal durumunun olmadığını beyan ederiz.

Çıkar Çatışması Beyanı

Bu çalışmanın yayınlanmasında akademik ya da finansal bir çıkar çatışması olmadığını beyan ederiz. Çalışma birinci yazarın ikinci yazarın danışmanlığında yürütmekte olduğu doktora tezi verilerinden üretilmiştir.

Bilgilendirilmiş Onam

Katılımcılar araştırmaya gönüllü olarak katılmışlardır. Katılımcıların ortaokul 8.sınıf öğrencisi olmaları sebebiyle bilgilendirilmiş onam formu olarak veli onam formları alındığını beyan ederiz.

Kaynakça

- Açıkgöz, M., & Karslı, F. (2015). Alternatif ölçme-değerlendirme yaklaşımları kullanılarak iş ve enerji konusunda geliştirilen başarı testinin geçerlilik ve güvenirlik analizi. *Amasya Üniversitesi Eğitim Fakültesi Dergisi*, 4(1), 1-25.
- Akbulut, H. İ., & Çepni, S. (2013). Bir üniteye yönelik başarı testi nasıl geliştirilir?: İlköğretim 7. sınıf kuvvet ve hareket ünitesine yönelik bir çalışma. *Amasya Üniversitesi Eğitim Fakültesi Dergisi*, 2(1), 18-44.
- Akınoğlu, O., & Tandoğan, Ö. R. (2007). The effects of problem-based active learning in science education on students' academic achievement, attitude and concept learning. *Eurasia Journal of Mathematics, Science & Technology Education*, 3(1), 71-81.
- Anderson, L.W. (Ed.), Krathwohl, D.R. (Ed.), Airasian, P.W., Cruikshank, K.A., Mayer, R.E., Pintrich, P.R., Raths, J., & Wittrock, M.C. (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's Taxonomy of Educational Objectives* (Kısaltılmış baskı). New York: Longman.
- Ateş, S. (2019). *Bilimsel Muhakeme (Akıl Yürütme)*. Ankara, Palme Yayınevi.
- Atılğan, H., Kan, A., & Aydın, B. (2017). *Eğitimde ölçme ve değerlendirme*. (11. Baskı). Ankara, Anı Yayıncılık.
- Aybek, E. C. (2021). "Data preparation for factor analysis." from <https://shiny.eptlab.com/dp2fa/>
- Aydoğdu, B., Tatar, N., Yıldız, E., & Buldur, S. (2012). İlköğretim öğrencilerine yönelik bilimsel süreç becerileri ölçeğinin geliştirilmesi. *Journal of Theoretical Educational Science*, 5(3), 292-311.
- Brown, T. A. (2015). *Confirmatory factor analysis for applied research* (2nd ed.). The Guilford Press.
- Bryman, A., & Cramer, D. (2001). *Quantitative data analysis with SPSS release 10 for Windows: A guide for social scientists*. Routledge.
- Büyüköztürk, Ş. (2010). *Veri Analizi El Kitabı* (12. bs). Ankara, PegemA Yayıncılık.
- Büyüköztürk, Ş., Akgün, Ö. E., Karadeniz, Ş., Demirel, F. ve Kılıç, E. (2013). *Bilimsel araştırma yöntemleri*. Ankara: Pegem Akademi
- Crocker, L., & Algina, J. (1986). *Introduction to classical and modern test theory*. Holt, Rinehart and Winston, 6277 Sea Harbor Drive, Orlando, FL 32887
- Davis, F. B., & Diamond, J. J.(1974). *The preparation of criterion-referenced tests for the CSE monograph series in evaluation*. Los Angeles:CenterStudy of Evaluation, University of California
- Fives, H., Huebner, W., Birnbaum, A. S., & Nicolich, M. (2014). Developing a measure of scientific literacy for middle school students. *Science Education*, 98(4), 549-580.
- Fleiss, J. L., Levin, B., & Paik, M. C. (2003). Statistical methods for rates and proportions.
- Fraenkel, J. R., Wallen, N. E., & Hyun, H. H. (2012). *How to design and evaluate research in education*.(8th ed.). New York: McGraw-Hill.

- Gravetter, F. J., & Forzano, L. A. B. (2018). *Research methods for the behavioral sciences* (6th ed.). Boston: Cengage.
- Güngörmez, H. G., & Akgün, A. (2018). Ortaokul Öğrencilerinin Fen Bilimleri Dersindeki Kuvvet ve Enerji Ünitesine Yönelik Akademik Başarı Testi Geliştirme Çalışması. *Diyalektolog-Uluslararası Sosyal Bilimler Dergisi*, (18).
- Hambleton, R. K., & Eignor, D. R. (1979). *A practitioner's guide to criterion-referenced test development, validation, and test score usage* (2. Baskı). University of Massachusetts, Amherst.
- İdin, Ş., & Aydoğdu, C. (2016). Kuvvet ve hareket ünitesi başarı testi geçerlik ve güvenirlik araştırması. *Eğitim Bilim ve Teknoloji Araştırmaları Dergisi*, 1(1), 14-33.
- Jackson, R. (1970). *Developing criterion-referenced tests*. ERIC Clearinghouse on Tests, Measurement, and Evaluation.
- Karatay, R., & Doğan, F. (2016). Ortaokul 7. Sınıf Fen ve Teknoloji Dersi Bilimsel Süreç Becerileri Ölçeğinin Geliştirilmesi. *Dicle Üniversitesi Ziya Gökalp Eğitim Fakültesi Dergisi*, (27), 1-8.
- Kıray, S. A. (2010). *İlköğretim ikinci kademedeki uygulanan fen ve matematik entegrasyonunun etkililiği*. (Doktora Tezi). Yüksek Öğretim Merkezi Ulusal Tez Merkezinden alınmıştır. (Tez No: 381067).
- Kline, R. B. (2019). *Yapısal eşitlik modellemesinin ilke ve uygulaması*. Ankara: Nobel Akademi.
- Koğar, H., & Yılmaz Koğar, E. (2016). Comparison of Different Estimation Methods for Categorical and Ordinal Data in Confirmatory Factor Analysis. *Journal of Measurement and Evaluation in Education and Psychology*, 6(2). <https://doi.org/10.21031/epod.94857>
- Korkmaz, S., Göksülük, D., & Zararsız, G. (2014). MVN: An R package for assessing multivariate normality. *The R Journal* 6(2): 151-162.
- MEB (2018). Milli Eğitim Bakanlığı Fen Bilimleri Dersi Öğretim Programı (İlkokul ve Ortaokul3,4,5,6,7,8. Sınıflar). <https://mufredat.meb.gov.tr/Dosyalar/201812312311937-FEN%20B%C4%B0L%C4%B0MLER%C4%B0%20C3%96%C4%9ERET%C4%B0M%20PROGRAMI2018.pdf> adresinden 06.01.2021 tarihinde edinilmiştir.
- MEB (2006). Talim ve terbiye kurulu başkanlığı, ilköğretim fen ve teknoloji dersi (6, 7 ve 8. Sınıflar) öğretim programı. Ankara: Devlet Kitapları Müdürlüğü.
- MEB (2013). Talim ve terbiye kurulu başkanlığı, ilköğretim kurumları fen bilimleri dersi (3, 4, 5, 6, 7 ve 8. Sınıflar) öğretim programı. Ankara: Devlet Kitapları Müdürlüğü.
- MEB (2024), Türkiye Yüzyılı Maarif Modeli. Fen Bilimleri Öğretim Programı. <https://tymm.meb.gov.tr/ogretim-programlari/fen-bilimleri-dersi/8> adresinden 29.10.2024 tarihinde edinilmiştir.
- Miller, J. D. (1983). Scientific literacy: A conceptual and empirical review. *Daedalus*, 29-48.
- Mullis, I. V.S., & Martin, M. O. (2017). *TIMSS 2019 assessment frameworks*. TIMSS and PIRLS International Study Center. Chestnut Hill, MA: Lynch School of Education, Boston College.
- NRC (2015) <https://www.nagb.gov/content/dam/nagb/en/documents/publications/frameworks/science/2015-science-framework.pdf> adresinden 30.10.2024 tarihinde edinilmiştir.
- NSTA (2024) <https://www.nsta.org/applying-knowledge-context> adresinden 4.11.2024 tarihinde edinilmiştir.
- Nunnally, J. C. (1978). *Psychometric theory* (2. Bölüm). McGraw-Hill.
- OECD (2019). *PISA 2018 Assessment and Analytical Framework*. Paris: OECD Publishing. doi:<https://doi.org/10.1787/b25efab8-en>.
- O'Hearn, G. T. (1976). Science literacy and alternative futures. *Science Education*, 60(1),

103-114. DOI: 10.1002/sce.3730600114.

R Core Team. (2023). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. <https://www.R-project.org/>

Sarmiento, R. P., & Costa, V. (2019). Confirmatory factor analysis--a case study. arXiv preprint arXiv:1905.05598.

Schumacker, R. E., & Lomax, R. G. (2004). *A beginner's guide to structural equation modeling*. (2. Baskı). Lawrence Erlbaum Associates, Publishers Mahwah, New Jersey London.

Stevens, J. (2002). *Applied multivariate statistics for the social sciences*. (5. Baskı). Mahwah, NJ: Lawrence Erlbaum Associates.

Suna, H. E., Şensoy, S., Parlak, B. & Özdemir, E. (2020). TIMMS 2019 Türkiye Ön Raporu. Milli Eğitim Bakanlığı Eğitim Analiz ve Değerlendirme Raporları Serisi. No:15. Aralık.

Özçelik, D. A. (2010). *Test Hazırlama Kılavuzu* (4. baskı). Ankara, PegemA Yayıncılık.

Öztürk, M. (2014). 8. Sınıf kuvvet ve hareket ünitesine yönelik bilgisayar destekli öğretim uygulamalarının etkililiğinin araştırılması. Yüksek Lisans Tezi. Karadeniz Teknik Üniversitesi.

Özden, B., & Yenice, N. (2017). "Kuvvet ve Enerji" Ünitesine Yönelik Üç Aşamalı Kavramsal Anlama Testi Geliştirme Çalışması. *Necatibey Eğitim Fakültesi Elektronik Fen ve Matematik Eğitimi Dergisi*, 11(2), 432-463.

Tabachnick, B. G., & Fidell, L. S. (2012). *Using Multivariate Statistics* (Sixth ed.): Pearson.

Tierney, N. and D. Cook (2023). "Expanding Tidy Data Principles to Facilitate Missing Data Exploration, Visualization and Assessment of Imputations." *Journal of Statistical Software* 105(7): 1-31.

Uyumaz, G., & Sırgancı, G. (2020). Doğrulamalı faktör analizi için gerekli örneklem büyüklüğü kaç kişidir?: Bayes Yaklaşımı ve maksimum olabilirlik kestirimi. *OPUS International Journal of Society Researches*, 16(32), 5302-5340.

Yüzbaşıoğlu, M. K., & Kurnaz, M. A. (2023). Fen Bilimleri Dersi" Kuvvetin Ölçülmesi ve Sürtünme" Ünitesine Yönelik Bağlam Temelli Öğrenme Durumları Testi Geliştirme Çalışması. *Adıyaman Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, (45), 32-65.