



[itobiad], 2019, 8 (1): 64/90

**TIMSS 2011 Fen Bilimleri Testinin 2016 İzmir ve 2011 Türkiye
Verilerinin Karşılaştırılması**

Comparison of 2016 İzmir and 2011 Turkey Data by TIMSS 2011
Science Test

Suat TÜRKOĞUZ

Doç.Dr., Dokuz Eylül Üni., Buca Eğitim Fakültesi
Assoc.Prof.Dr., Dokuz Eylül Uni., Buca Faculty of Education
suat.turkoguz@gmail.com

Orcid ID: 0000-0002-7850-2305

Ali Günay BALIM

Prof.Dr., Dokuz Eylül Üni., Buca Eğitim Fakültesi
Prof.Dr., Dokuz Eylül Uni., Buca Faculty of Education
agbalim@gmail.com

Orcid ID: 0000-0003-2010-1696

Vefa BARDAKÇI

Eğitim Uzmanı, Milli Eğitim Bakanlığı
M.Sci. of Education, Ministry of National Education

Orcid ID: 0000-0003-2707-8599

Makale Bilgisi / Article Information

Makale Türü / Article Types : Araştırma Makalesi / Research Article
Geliş Tarihi / Received : 12.10.2018
Kabul Tarihi / Accepted : 05.02.2019
Yayın Tarihi / Published : 03.03.2019
Yayın Sezonu : Ocak-Şubat-Mart
Pub Date Season : January-February-March

Atıf/Cite as: TÜRKOĞUZ, S , BALIM, A , BARDAKÇI, V . (2019). TIMSS 2011 Fen Bilimleri Testinin 2016 İzmir ve 2011 Türkiye Verilerinin Karşılaştırılması. İnsan ve Toplum Bilimleri Araştırmaları Dergisi, 8 (1), 64-90. Retrieved from <http://www.itobiad.com/issue/43055/469696>

İntihal /Plagiarism: Bu makale, en az iki hakem tarafından incelenmiş ve intihal içermediği teyit edilmiştir. / This article has been reviewed by at least two referees and scanned via a plagiarism software. <http://www.itobiad.com/>

Copyright © Published by Mustafa YİĞİTOĞLU- Karabük University, Faculty of Theology, Karabük, 78050 Turkey. All rights reserved.

TIMSS 2011 Fen Bilimleri Testinin 2016 İzmir ve 2011 Türkiye Verilerinin Karşılaştırılması

Öz

Bu çalışmada, TIMSS (Trends in International Science and Mathematics Study) 2011 sekizinci sınıf Fen bilimleri testi (TIMSS₁₁FBT₈) aracılığıyla 2016 yılında İzmir’de tekrar toplanan verilerle 2011 yılında Türkiye’de toplanan verilerinin karşılaştırılması amaçlanmıştır. Bu çalışmada betimsel tarama modeli kullanılmıştır. Bu amaçla TIMSS₁₁FBT₈’nin seçenekli 80 sorusu 2016 yılında İzmir’de 775 ortaokul sekizinci sınıf öğrencisine uygulanmıştır. Veriler, betimsel istatistik teknikleriyle analiz edilmiş; güvenilirlik ve korelasyonel değerlerine bakılmıştır. Çalışma sonucunda, TIMSS₁₁FBT₈ 2011 yılı Türkiye verilerine göre güvenirligi 2016 yılı İzmir verilerine göre biraz daha düşük bulunmuştur. Ayrıca TIMSS₁₁FBT₈ 2011 yılı Türkiye verilerine göre öğrenme alanı ortalaması 2016 yılı İzmir verilerine göre tutarlı olduğu söylenebilir. Fizik öğrenme alanı diğer öğrenme alanlarına göre düşük düzeyde kalmaktadır. Buna ek olarak; TIMSS₁₁FBT₈’nin bilişsel, uygulama ve sorgulama düzeylerine ilişkin 2016 yılı Türkiye ve 2011 yılı İzmir verileri karşılaştırıldığında sonuçların benzer olduğu görülmektedir. Sorgulama düzeyi diğer düzeylere göre düşük kalmaktadır. Bu kapsamda öğrenme alanlarının sorgulama etkinlikleri ile desteklenmesi gerekmektedir.

Anahtar Kelimeler: TIMSS, Fen Bilimleri, Eğitim, Ölçme, Türkiye.

A Comparison of 2016 İzmir and 2011 Turkey Data by TIMSS 2011 Science Test

Abstract

This study aims to compare 2016 İzmir data and 2011 Turkey data collected through the 8th grade Science Test (TIMSS₁₁FBT₈). This study is a descriptive research model. Data were gathered by 80 questions of TIMSS₁₁FBT₈. The participants were 775 students at elementary schools in İzmir, 2016. Data were analyzed by descriptive statistical techniques. Additionally; reliability and correlational values were examined. According to the results of the study, the reliability of 2011 Turkey data in TIMSS₁₁FBT₈ is lower by 2016 İzmir data. Moreover, it shows that 2016 Turkey data are consistent with 2016 İzmir data by the cognitive domains of TIMSS₁₁FBT₈. Compared to the others, the physic learning area is low. Additionally; through a comparison of 2016 İzmir data and 2016 Turkey data with regard to the cognitive levels of TIMSS₁₁FBT₈, it may be seen that the results bear similarity. The reasoning area of TIMSS₁₁FBT₈ is low compared to other areas. In this context, learning activities should be supported by reasoning activities.

Keywords: TIMSS, Science, Education, Measurement, Turkey.



Giriş

2005 yılında Fen bilgisi dersi öğretim programı, fen ve teknoloji dersi öğretim programı olarak reform niteliğinde değiştirilmiş, 2013 yılında Fen bilimleri dersi öğretim programı şeklinde yenilenmiştir. 2017 yılında ise programın içeriğinde yeniden güncelleme çalışmaları başlatılmıştır. Programdaki bu yeniliklerle içerik kapsamının uluslararası Fen bilimleri program içeriklerine yaklaştırılması, öğrenci merkezli etkinliklerle sorgulayan, tasarlayan, girişimci ruhlu 21. yüzyıl becerileriyle donanımlı bireylerin yetiştirilmesi amaçlanmıştır. Program içeriklerinde yapılan değişikliklerde ulusal ve uluslararası toplumsal, ekonomik ve siyasi gereksinimler, akademik araştırmalar, ulusal merkezi ve uluslararası testlerden alınan genel veriler etkilidir. Ülkeler kendi eğitim sistemlerini kendi içlerinde ve uluslararası düzeyde karşılaştırabilecekleri ve uzun vadede veri tabanı oluşturabilecekleri çalışmaları 1990'lı yılların başından itibaren uluslararası sınavlara katılmakla gerçekleştirmektedir (Abazaoğlu, Yıldızhan ve Yıldırım, 2014). Uluslararası bu sınavlara TIMSS (Third International Mathematics and Science Study-TIMSS), PISA ve PIRLS gibi sınavlar örnek verilebilir. Uluslararası testlerden elde edilen bulgular, Türkiye'nin öğretim programlarını geliştirme ve uygulama konusunda beklenen düzeyde olmadığı kanısını oluşturmaktadır.

TIMSS, uluslararası kapsamda yapılan Matematik ve Fen alanlarında sorunları ortaya koymayı amaçlayan çok boyutlu ve farklı bileşenleri de kullanarak inceleyen bir değerlendirme sınavıdır (Ersoy, 2016). TIMSS gibi sınavlarla ulusal ve uluslararası boyutta Fen ve Matematik eğitimindeki sorunlar belirlenebilir. Bunun yanı sıra, öğretim programları elde edilen bulgulara göre güncellenebilir ve öğrencilerin Fen ile Matematik başarılarına etki eden faktörler ortaya çıkarılabilir (Ceylan ve Berberoğlu, 2007). TIMSS, okullarda verilen Fen ve Matematikle ilgili sorunlara yanıt olması için bu alanlara yönelik güncel ihtiyaçlara göre dönüt olan betimsel bir çalışma yürütür. TIMSS'in bulguları hiyerarşik bir model içinde genel ülke puanları boyutuyla sunulduğundan derinlemesine bilgilere ulaşmak zordur. Ancak bu zorluk bazı araştırmacıların ilgisi ve merakı ile derinlemesine yapılan analizler sonucunda elde edilen araştırmalarla aşılabilmektedir. TIMSS, Türkiye hakkında genel bilgiler verirken, bölgesel sonuçları ise bilgi olarak sunmamaktadır.

Doğrudan gözlenemeyen davranışlar; başarı testi, anket, ilgi, algı, tutum ölçeği gibi değişik araçlarla dolaylı yollardan ölçülmeye çalışılır. Dolaylı yollardan yapılan tüm ölçmelerde hata kaynakları mevcuttur (Erkuş, 2010). Hata kaynaklarından arındırılmış gerçek veri puanlarının hesaplanmasında çeşitli ölçme kuramlarından faydalanılmaktadır. Ulusal merkezî sistem ve uluslararası ölçmelerle yapılan sınavlarda öğrencilerin davranışları bilişsel ve duyuşsal test maddeleri ile dolaylı olarak ölçülmektedir. Bu nedenle dolaylı yollarla toplanan veriler genellikle birbirleriyle ilişkili ve iç içe



geçmiş durumdadır. Sözgelimi TIMSS, PISA ve PIRLS gibi uluslararası sınavlarda iç içe geçmiş verileri değerlendirme işlemlerinde madde tepki kuramı (MTK) modelleri, öğrencilerin gizil yeteneklerini ölçmede geliştirilen testleri hiyerarşik lineer modeliyle analiz etmede ve puanlamada kullanılır. (Anıl, 2009; Atar, 2010; Uzun, Gelbal ve Öğretmen, 2010). Günümüzde HLM'nin (Hiyerarşik Lineer Model) kullanıldığı araştırmalar yaygınlaşmakta ve YEM (Yapısal Eşitlik Modeli) ile birleştirilerek daha derin değerlendirmelere gidilebilmektedir. Başarı üzerine yapılan istatistiksel modellemeler öğretim programlarını değerlendirme açısından önemli bulgular sunabilir. Uluslararası sınavlarda HLM uygulandığından akademik başarı puanlarının güvenilirlik hesaplarından kısmi korelasyon ve korelasyon katsayılarından faydalanılmaktadır. Hata kaynaklarının araştırılmasında güvenilirlik katsayısı önemli yer tutmaktadır. HLM modeline duyulan güvenden dolayı bu tür araştırmalarda Alpha Chronbach ve KR 20-21 gibi güvenilirlik katsayısı verilerine ise çok rastlanılmamaktadır.

Ulusal ya da uluslararası sınavlarda elde edilen puanların yorumlanmasında grup karşılaştırmaları yapılırken grupların nitelikleri göz önünde bulundurulması ve puanlamaya katkıları hesaplanması gerekliliği konusunda vurgu yapan çalışmalar da vardır (Kelly, 2002; Ceylan ve Berberoğlu, 2007; Öztürk, ve Uçar, 2010; Uzun, Gelbal ve Öğretmen, 2010; Ersoy, 2018). TIMSS maddeleri hazırlanırken tüm katılımcı ülkelerin öğrencilerinin düzeyleri düşünülerek hazırlanmaktadır. TIMSS test maddeleri hazırlanırken sınava dâhil olan ülkelerdeki uzmanların görüşlerine de başvurulmaktadır. Bu bakımdan test maddelerinin hatalardan arınık bir şekilde hazırlandığı düşünülebilir. Her ne kadar bu ölçütlere dikkat edilse de soru maddelerinin anlaşılmasında dilsel ve kültürel farklılıklar test puanlarına olumsuz bir şekilde yansiyabilmektedir (Deniz, 2007). Çünkü grup nitelikleri gruptan gruba kültürden kültüre farklılık göstermekte ve dolaylı olarak puan ortalamalarında muhtelif sapmalar meydana getirebilmektedir. Bunun en önemli göstergesi ülkeler arası puan farklılıklarıdır. PISA, TIMSS ve PIRLS gibi uluslararası sınavlardan elde edilen veriler her ülkenin kendi bünyesindeki okulları karşılaştırma olanakları sunabilmektedir. PISA 2006 raporu, öğrenci başarıları arasındaki farklılığın üçte birinin okullar arasındaki farklılıklardan kaynaklandığını açıklamaktadır (Ceylan, 2009). Uluslararası sınavların verileri okullar arasındaki farklılık gösterebildiği gibi yıllar arasında ve ülkenin bünyesinde yer alan farklı coğrafi yerlere göre de değişim gösterebilir. Farklı coğrafyalardan sınava katılanların dil, kültür ve okul farklılıkları uluslararası sınavlardan elde edilen puanlar arasında farklılıklar oluşturabilir.



TIMSS Fen Bilimleri Testi Hakkında Genel Bir Bilgi

TIMSS sınavı, ilk olarak 1995 yılında 41 ülkenin katılımıyla uygulanmıştır. Devamında dört yıllık döngülerle 1999, 2003, 2007, 2011 ve 2015 yıllarında gerçekleştirilmiştir (Ceylan ve Berberoğlu, 2007). Bu sınav uluslararası bir ölçme değerlendirme kuruluşu olan IEA (International Association for the Evaluation of Educational Achievement), tarafından gerçekleştirilmektedir. TIMSS uygulamaları, katılan ülkelerin 4. ve 8. sınıflarında yer alan öğrencilerin Fen ve Matematik alanlarındaki başarılarını değerlendirmektedir (Abazaoğlu, Yıldızhan ve Yıldırım, 2014; Bayraktar, 2010). Bu çalışmada, 2011 yılında yapılan TIMSS₁₁FBT₈'ye ilişkin dokümanlar ve sayısal veriler kullanılmakta ve bunlara ilişkin istatistik veriler değerlendirilmektedir.

TIMSS₁₁FBT₈'nin İçeriği

2011 yılında yapılan TIMSS araştırmasına 67 ildeki 259 ilköğretim okulunun 8. sınıfında öğrenim gören 7286 öğrenci katılmıştır. 2011 TIMSS₁₁FBT₈ biyoloji (%35), fizik (%25), kimya (%20) ve yer bilimleri (%20) disiplinlerinden dört alanı kapsamaktadır. Bu dört konu alanındaki sorular bilişsel (%35), uygulama (%35) ve sorgulama (%30) düzeyinde düzenlenmiştir. Bilişsel düzeydeki soruların tanım yapmaya, örneklendirmeye, hatırlamaya, bilimsel süreçleri ve yöntemleri kullanmaya yönelik kazanımları karşıladığı; uygulama düzeyindeki soruların karşılaştırma yapma, çözüm üretme, açıklamada bulunma, yorumlama, modellendirme ve tasnifleme gibi kazanımları karşıladığı; sorgulama düzeyindeki soruların tasarım yapma, çıkarımda bulunma, genelleme, ispatlama, tahminde bulunma, analiz ve sentez yapma gibi kazanımları karşıladığı görülmüştür (Abazaoğlu, Yıldızhan ve Yıldırım, 2014; Büyükoztürk, Çakan, Tan ve Atar, 2014).

TIMSS₁₁FBT₈'nin Uygulanması, Puanlanması ve Analizi

TIMSS₁₁FBT₈'nin içerisinde yer alan soru maddeleri, uygulandığı 2011 yılındaki Fen ve Teknoloji dersi öğretim program kazanımları ve 2013 yılındaki Fen bilimleri dersi öğretim programı kazanımlarıyla karşılaştırmalı olarak uzmanlar tarafından incelenmiş ve onların görüşleri doğrultusunda testin kapsam geçerliliğinin yüksek olduğu anlaşılmıştır. Testte 172 adet soru kullanılmış, testin yarısı çoktan seçmeli testten, diğer yarısını açık uçlu yoruma dayalı sorulardan oluşmaktadır. Ayrıca araştırmaya katılan öğrenciler testle beraber kendi yaşadıkları ortamlarla eğitim gördükleri ortamlar hakkında bilgiler verdikleri maddeleri de cevaplamaktadırlar. Bununla birlikte öğretmenler, okul müdürleri, öğretimsel uygulamalar, sınıf içindeki deneyimler, okul olanakları ve okuldaki öğrenme iklimi hakkında bilgiler vermektedir. Öğrencilere 172 adet sorunun olduğu testin tamamı değil, Hiyerarşik Çoklu Regresyon Analiz Yaklaşımına dayalı olarak soruların bir kısmı verilerek test yapılmaktadır. Bu nedenle TIMSS₁₁FBT₈ 14



kitapçık halinde öğrencilere dağıtılmıştır. Bu nedenle 172 adet test maddesinden oluşan TIMSS₁₁FBT₈, 25-30 maddelik testler halinde 14 kitapçığa bölünerek katılımcılara uygulanmıştır. 14 test kitapçığı katılımcılara aynı gün ve saatte dağıtılarak gerçekleştirilmiştir. Öğrencilere 45 dakikada yanıtlayabilecekleri sorular verilmiştir ve her kitapçıkta sorular farklıdır. Test kitapçıkları “Hiyerarşik Çoklu Regresyon Analiz Yapısı”na uygun olacak şekilde düzenlenmiştir. Her bir kitapçık için sınav süresi 45 dakika olarak belirlenmiştir. Başarı testinden sonra öğrencilerin dinlenmesi için 30 dakika süre verilmektedir. Ardından anket uygulanmaktadır. Başarı testleri arasında 5 dakika, anket uygulaması öncesinde ise 15 dakika ara verilmiştir. Çoktan seçmeli sorular dört seçeneqli olarak hazırlanmakta ve her bir soru maddesinin 1 puan değerinde tek doğru cevabı vardır. Yanlış verilen cevaplar doğru verilen cevapları ya da puanları etkilememektedir. Açık uçlu sorularda ise öğrenciler kendi yanıtlarını yazılı olarak yorum, çizim, grafik ve açıklama katarak vermektedir. Açık uçlu sorular puanlama anahtarıyla değerlendirilmekte ve her bir sorunun puanlama anahtarı farklılık göstermektedir. TIMSS₁₁FBT₈ başarı puanları ve yeterlik düzeyleri dört gruba bölünerek düzenlenmiştir. Alt düzeydekiler için 400/475; orta düzeydekiler için 475/550; yüksek düzeydekiler için 550/625; ileri düzeydekiler için ise 625 ve üstü başarı puanı olarak tanımlanmıştır. 400 puan altı alanlarda alt düzey altı olarak tanımlanmıştır. TIMSS₁₁FBT₈’den beklenen ortalama başarı puanı 500; yani en az orta düzeyde olması beklenmektedir (Abazaoğlu, Yıldızhan ve Yıldırım, 2014; Büyüköztürk, Çakan, Tan ve Atar, 2014).

Bu çalışmada, TIMSS₁₁FBT₈’nin 2011 Türkiye verileriyle 2016 İzmir verilerinin betimsel ve ilişkisel karşılaştırmasının yapılması amaçlanmıştır. Bununla birlikte korelasyonel ve Path analiziyle TIMSS₁₁FBT₈’de yer alan öğrenme alanlarının ve ayrıca bilgi, sorgulama ve uygulama düzeylerinin TIMSS₁₁FBT₈ testine yönelik katkı düzeylerinin ortaya konulması hedeflenmiş; 2011 Türkiye ve 2016 İzmir TIMSS₁₁FBT₈ test verilerinden elde edilen test puanlarına olan katkılarının karşılaştırılması düşünülmüştür. Bu kapsamda;

1. TIMSS₁₁FBT₈’de yer alan biyoloji, kimya, fizik ve yer bilimleri öğrenme alanlarının ortalama ve standart sapma değerlerine bakılarak 2011 Türkiye ve 2016 İzmir verileriyle karşılaştırılması;
2. TIMSS₁₁FBT₈’de yer alan bilgi, sorgulama ve uygulama düzeylerinin ortalama ve standart sapma değerlerine bakılarak 2011 Türkiye ve 2016 İzmir verileriyle karşılaştırılması;
3. TIMSS₁₁FBT₈’de yer alan biyoloji, kimya, fizik ve yer bilimleri öğrenme alanlarına ilişkin puanların birbirleri ve TIMSS₁₁FBT₈ toplam puanı arasındaki korelasyonel ve Path analizi değerlerine bakılarak 2011 Türkiye ve 2016 İzmir verileriyle ilişkisel karşılaştırılması;
4. TIMSS₁₁FBT₈’de yer alan bilgi, sorgulama ve uygulama düzeylerine ilişkin puanların birbirleri ve TIMSS₁₁FBT₈ toplam puanı arasındaki korelasyonel ve



Path analizi değerlerine bakılarak 2011 Türkiye ve 2016 İzmir verileriyle ilişkisel karşılaştırılması;

5. TIMSS₁₁FBT₈'de yer alan biyoloji, kimya, fizik ve yer bilimleri öğrenme alanlarına ilişkin puanların birbirleri ve TIMSS₁₁FBT₈ toplam puanı arasındaki 2016 korelasyonel ve Path analizi değerlerine bakılarak 2016 İzmir verileriyle ve 2016 TEOG Fen bilimleri testi puanıyla ilişkisel karşılaştırılması;

6. TIMSS₁₁FBT₈'de yer alan bilgi, sorgulama ve uygulama düzeylerine ilişkin puanların birbirleri ve TIMSS₁₁FBT₈ toplam puanı arasındaki 2016 korelasyonel ve Path analizi değerlerine bakılarak 2016 İzmir verileriyle ve 2016 TEOG Fen bilimleri testi puanıyla ilişkisel karşılaştırılması amaçlanmıştır.

Yöntem

Bu çalışmada betimsel tarama modeli kullanılmıştır. TIMSS uluslararası araştırmalarından sonra topladığı verileri, verilerin analizlerini, test içeriklerini ve kullandıkları yöntemleri genel rapor halinde web (<http://timssandpirls.bc.edu/>) sayfasından paylaşmaktadır. Bu kapsamda çalışmada TIMSS₁₁FBT₈'nin içeriği, hedefleri ve madde yapıları ilgili web sayfasından indirilerek incelenmiştir. TIMSS₁₁FBT₈'de yer alan soru maddelerinin serbest bırakılan maddeleri "<https://timss.bc.edu>" linkinden indirilerek ve TIMSS₁₁FBT₈ ile ilgili soru maddeleri literatürden bulunarak oluşturulan soru havuzu, 2013 Fen bilimleri ders içeriği kazanımlarına göre karşılaştırılmış ve 2016 yılında İzmir ilinde yeniden uygulanması amacıyla 80 maddelik tek kitapçıktan oluşan TIMSS₁₁FBT₈ hazırlanmıştır. 80 soru maddesi 2013 Fen bilimleri dersinin öğretim programında yer alan kazanımlara göre belirlenmiştir. Hazırlanan bu 80 maddelik kitapçık, 2011 yılında hazırlanan kitapçıktan biçimsel olarak farklı bir biçimde düzenlenmiş olup ve MEB tarafından Fen bilimleri dersinin kazanımlarının güncellenmesinden dolayı yeniden uzman görüşüne başvurularak öneriler alınmıştır. Ayrıca TIMSS'in web sayfasında paylaştığı veriler indirilerek Türkiye ile ilgili TIMSS₁₁FBT₈'ye ilişkin güvenilirlik, puanlama biçimleri incelenmiştir. 80 maddelik tek kitapçık formundaki TIMSS₁₁FBT₈ İzmir ilindeki 8. sınıflara 2016 yılında uygulanmıştır. Elde edilen veriler TIMSS₁₁FBT₈ puanlama ölçütlerine göre puanlanarak veriler excel programında girilmiş ve istatistik programlarında betimsel analizleri yapılmıştır.

Evren ve Örneklem

TIMSS₁₁FBT₈ 2011 Türkiye katılımcıları YEĞİTEK ulusal merkezi ve IEA'nın WinW3S veri tabanı kullanılarak tesadüfi örnekleme yöntemiyle belirlenmektedir. Bunun öncesinde Milli Eğitim Bakanlığı nezdinde IEA araştırma şirketinin örnekleme planlaması kapsamında okullara belirlenen



kriterler ışığında kodlar atanarak IEA araştırma şirketinin veri tabanına yüklenmektedir. Bu aşamadan sonra belirlenen örnekleme programları sayesinde tesadüfi örnekleme yapılmaktadır. 2011 yılında Türkiye’de TIMSS₁₁FBT₈’ye katılan öğrencilerin yaş ortalaması 14’tür. Ayrıca katılımcıların %49’u kız, %51’i erkektir. Verilerin analizi sırasında eksik veri silme işlemleri gerçekleştirildiğinden bu oran kızlar için %51 ve erkekler için %49’a doğru bir değişim göstermiştir. 2011 yılında yapılan TIMSS araştırmasına 67 ilden 259 ilköğretim okulu ve sekizinci sınıftan 7286 öğrenci katılmıştır. 2011 yılı için TIMSS₁₁FBT₈ için Türkiye evreninden seçilen okulların örnekleme yüzdesi % 0.8’dir. 2011 yılında TIMSS₁₁FBT₈’ye katılan 5441 sekizinci sınıf öğrencisi 2016 yılında İzmir’de araştırma kapsamında yinelenen TIMSS₁₁FBT₈’ne katılan 775 öğrencilerden elde edilen verilerle karşılaştırılmak üzere seçilmiştir. 7286 öğrenciden 5441 öğrenciye indirgenmesinde 2016 yılında İzmir’de uygulanan TIMSS₁₁FBT₈ için seçilen 80 soru maddesinin etken olmasıdır. 80 soru, 2013 Fen bilimleri dersi öğretim programında yer alan kazanımlara göre seçilmiştir. 2011 yılında yapılan TIMSS sınavının soru maddeleri, 2004 Fen ve Teknoloji dersi öğretim programı kazanımlarına göre yapılandırılmıştır. Bu nedenle kapsam farklılığı görülebilir. Karşılaştırmalarda kapsam farklılığından oluşacak hataları azaltmak için ortak kazanımlardan yararlanılarak sorular belirlenmiş ve bu sorulara göre analizler yapılmıştır. Bu nedenle 2011 yılındaki TIMSS sınavına katılan öğrencilerden (n=7286) ortak belirlenen sorulara yanıt veren öğrenciler (n=5441) analizlere dâhil edilmiştir. Çünkü TIMSS₁₁FBT₈’de 6 kısımda 172 adet test maddesi yer almakta ve öğrencilere birbirinden bağımsız olarak HLM’ye dayalı 14 kitapçık formunda dağıtılmaktadır. Bu nedenle 2016 yılında yeniden yapılan uygulamada karşılaştırmak amacıyla TIMSS₁₁FBT₈’ni yanıtlayanların 5441’i katılımcı olarak kabul edilmiştir. 2016 yılında İzmir ilinde ortaokullar küme örneklemeyle gruplandırılarak rastgele olarak gruplar içinden seçilmiştir. Okullardaki mevcut sekizinci sınıflar teste tabi tutulmuştur. 2016 yılında İzmir ilinde mevcut 512 ortaokuldan küme örneklemeyle rastgele seçilen 20 ortaokuldan 775 ilköğretim 8. sınıf öğrencisi çalışmanın örnekleme dâhil edilmiştir. Örnekleme dâhil edilen okul sayısı yaklaşık %2’dir. Bu oran 2011 yılındaki verilere göre İzmir’deki katılımı temsil etme açısından yeterli bir orandır. 2016 yılında İzmir’de TIMSS₁₁FBT₈’ye katılan öğrencilerin yaş ortalaması 14’tür. Ayrıca 2016 yılında İzmir’de TIMSS₁₁FBT₈’ye katılan kız öğrencilerin oranı %56, erkek öğrencilerin oranı %44’tür. Bu değerlerin 2011 yılında Türkiye genelinde katılım gösteren orana yakın bir değerdir. İzmir ilinde yapılan çalışmada mağduriyet yaşamamaları bakımından okulların, öğrencilerin, öğretmenlerin ve okul idarecilerinin isimleri gizli tutulmuştur. Ayrıca bu çalışma, 2016 yılında İzmir Milli Eğitim Müdürlüğü’nün desteği ve yasal izinleri kapsamında denetlenerek yapılmıştır. TIMSS Türkiye temsilciliği ile ilgili mail yoluyla yapılan yazışmalar ile testin uygulanması için gerekli izinler alınmıştır. 2016 yılında TIMSS₁₁FBT₈’ye yönelik örnekleme seçiminde, IEA araştırma şirketinin yapmış olduğu örnekleme



programlarına dayalı tesadüfi örnekleme gidilmemiştir. Her iki araştırma örnekleme seçiminde farklılık bu noktada görülmektedir.

Veri Toplama Araçları ve Analizleri

Bu çalışmada veri toplama aracı olarak, 14 farklı kitapçık halinde uygulanan TIMSS₁₁FBT₈'nin 2011 yılı Türkiye genelindeki test formunda yer alan çoktan seçmeli soru maddelerinden 80 tanesi kullanılmıştır. Açık uçlu sorular çalışma kapsamına dâhil edilmemiştir. 2011 yılı Türkiye genelinde uygulanan TIMSS₁₁FBT₈ biyoloji, kimya, fizik, yer bilimleri alanlarındaki kavramlarına yönelik 4 seçenekli çoktan seçmeli soruları ve açık uçlu kavramsal anlamayı ölçen toplam 172 adet test maddesinden oluşmaktadır. Ayrıca 2011 yılı Türkiye genelinde uygulanan TIMSS₁₁FBT₈ bilişsel düzeyi, sorgulama düzeyi ve uygulama düzeyi sorularını da kapsamaktadır. Bu çalışmada TIMSS₁₁FBT₈'nin 2011 yılı Türkiye genelindeki test formunda yer alan soru maddelerine TIMSS'in resmi web sayfasından ve ilgili alan yazını taramasından elde edilen makalelerden Türkçe ve İngilizce formlarına ulaşılmıştır. 2004 Fen ve Teknoloji dersi öğretim programı kazanımlarıyla 2013 Fen bilimleri dersi öğretim programı kazanımlarının uyumuna bakılarak ortak kazanımlar belirlenmiş ve soru maddeleri ortak kazanıma göre oluşturulmuştur. Yeniden yapılandırılan 80 maddelik testin biçimsel ve kazanımlara göre kapsam değerlendirmesi için yeniden uzman görüşlerine başvurulmuştur. Çoktan seçmeli sorular, çalışmanın amacına uygun olması ve değerlendirme kolaylığı açısından seçilmiştir. Çünkü TIMSS₁₁FBT₈'nin 2011 yılı Türkiye genelindeki test formunda yer alan soru maddelerinin puanlanmasında farklılıklar görülmektedir. Bu test formunda çoktan seçmeli soru maddeleri 1 puan ile değerlendirilirken, açık uçlu soru maddeleri oluşturulan cevap anahtarına göre belirlenmektedir. TIMSS₁₁FBT₈'nin 2016 yılı İzmir'de uygulanan test formunda yer alan çoktan seçmeli sorular TIMSS₁₁FBT₈'nin 2011 yılı Türkiye'de uygulanan test formunda olduğu gibi dört seçeneklidir ve bir doğru cevabı vardır. Aynı şekilde her bir çoktan seçmeli sorunun doğru cevabı 1 puandır ve yanlış cevaplar doğru cevapları etkilememektedir.

TIMSS₁₁FBT₈ 2011 yılında Türkiye genelinde 7286 kişiye 14 kitapçık halinde 6 kısma ayrılarak uygulanmıştır. Bu nedenle 7286 öğrencinin tamamı 172 adet soruyu yanıtlamamıştır. Kendilerine dağıtılan kitapçıkta yer alan 25 ile 30 maddeden oluşan testten sorumlu olmuştur. HLM'ye göre göre test kitapçıkları birbirleriyle ilişkilendirilerek Madde Tepki Kuramını esas alarak analiz edilmiştir. Bu çalışmada, 2016 yılında İzmir'de 775 öğrenciye farklı günlerde üç ders saati boyunca 80 maddelik testten oluşan TIMSS₁₁FBT₈ uygulanmıştır. Aynı okulda yer alan öğrencilere testin uygulama saati ve gününe dikkat edilerek farklılık giderilmeye çalışılmıştır. Çalışmaya seçilen okulların ulaşım açısından uzak noktalarda olmasından dolayı sınavın sonuçlarına bozacak olumsuz durumların oluşmadığı varsayılmıştır. Bu nedenle testin uygulama zorluğu açısından farklı noktalarda yer alan



okulların sınavları farklı günlerde yapılmıştır. Buradaki amaç testin bütününden elde edilen bulguların TIMSS₁₁FBT₈ 2011 yılında Türkiye genelinden HLM ile elde edilen bulguların karşılaştırılmasını sağlamaktır. Ayrıca 2016 yılında İzmir örnekleminde elde edilen verilerin değerlendirilmesinde Klasik Test Kuramı esas alınmış ve 2011 yılında Türkiye genelinden toplanan verilerin de Klasik Test Kuramına esasla analizleri yapılarak karşılaştırılmaları yapılmaya çalışılmıştır. Bu kapsamda bu araştırmada incelemeye alınan çoktan seçmeli sorulardan dolayı 5441 katılımcıdan elde edilen veriler analize dâhil edilmiştir. Türkiye genelindeki 7286 örneklem sayısı ile bu çalışmaya dâhil edilen 5441 örneklem sayısının farklılığı seçilen 80 soru maddesine katılan öğrenci sayısından kaynaklanmaktadır.

Verilerin analizinde Klasik Test Kuramı esas alınarak betimsel analiz teknikleri kullanılmıştır. Güvenirlik katsayısı hesabında Chronbach Alpha katsayısı ele alınmış ve betimsel verilerin analizinde ortalama değerlere dikkat edilmiştir. TIMSS₁₁FBT₈'nin biyoloji, fizik, kimya ve yer bilimleri öğrenme alanlarıyla bilgi, sorgulama ve uygulama düzeyleri arasındaki ilişkilendirmelerde regresyona dayalı ilişkisel path analizleri yapılmış tablo ve grafiklerle sunulmaya çalışılmıştır. IEA araştırma şirketi verilerin analizinde HLM kullanmakta ve eksik verileri analizden çıkarmaktadır. Bu çalışmada ise HLM kullanılmamış ve var olan veriler üzerinden betimsel analizlerle karşılaştırma yoluna gidilmiştir.

Tablo 1'de TIMSS₁₁FBT₈'nin 2011 yılının Türkiye verileriyle 2016 yılının İzmir verilerinin güvenirlik karşılaştırılması verilmiştir. Tablo 1'de TIMSS₁₁FBT₈'nin 2011 Türkiye verilerinde hesaplanan güvenirlik oranları IEA araştırma şirketi tarafından hesaplanan değerler değildir. IEA araştırma şirketinin "https://timss.bc.edu" adresinden paylaştığı verilerden yeniden hesaplanarak 2016 İzmir verileriyle karşılaştırılmaya gidilmiştir.

Tablo 1. TIMSS₁₁FBT₈'nin 2011 Türkiye Verileriyle 2016 İzmir Verilerine Göre Güvenirlik Açısından Karşılaştırılması

| Bölüm | Madde Sayısı | TIMSS ₁₁ FBT ₈ | | | |
|---------|--------------|--------------------------------------|-----|------------|-----|
| | | 2011 Türkiye | | 2016 İzmir | |
| | | r_{xy} | n | r_{xy} | n |
| 1.kısım | 13 | 0.74 | 998 | 0.68 | 775 |
| 2.kısım | 12 | 0.74 | 503 | 0.76 | 775 |
| 3.kısım | 12 | 0.71 | 985 | 0.77 | 775 |
| 4.kısım | 14 | 0.64 | 984 | 0.70 | 775 |
| 5.kısım | 16 | 0.67 | 983 | 0.74 | 775 |
| 6.kısım | 13 | 0.67 | 988 | 0.75 | 775 |

2011 yılında TIMSS₁₁FBT₈'de yer alan 172 adet test maddesi her kısımda 25-30 madde yer alacak şekilde 6 kısma bölünerek 14 kitapçık formatında katılımcılara dağıtılmaktadır. Bu çalışmada 172 adet test maddesinden 80



soru maddesi seçilmiş ve soru maddeleri Tablo 1 de gösterildiği gibi 6 kısım halinde harmanlanmıştır. 6 kısma bölünmüş 80 maddelik TIMSS₁₁FBT₈'ye ilişkin 2011 Türkiye ve 2016 İzmir verilerine ilişkin güvenilirlik bulguları Tablo 1'de sunulmuştur. Tablo 1'deki değerlere göre TIMSS₁₁FBT₈'nin 2011 Türkiye verilerinin ortalama güvenilirliği 0.70 olarak bulunmuştur. Bununla birlikte TIMSS₁₁FBT₈'nin 2016 İzmir verilerinin 6 kısımdan elde edilen ortalama güvenilirliği 0.73; toplam 80 maddeye göre hesaplanan güvenilirliği 0.93 olarak saptanmıştır. İki örneklem grubundaki güvenilirlik katsayısının farklı olması 2011 Türkiye verilerinin HLM'ye göre toplanması ve toplam test güvenilirliğinin hesaplanamaması neden olarak gösterilebilir. 2016 İzmir verilerinin güvenilirlik ortalamasının yüksek çıkması verilerin toplam 80 maddelik teste göre testin uygulanması ve hesaplamaların buna göre yapılmasına işaret edebilir. Madde sayısı azaldıkça güvenilirlik değerleri kullanılan formül gereği düşebilmektedir.

Bulgular

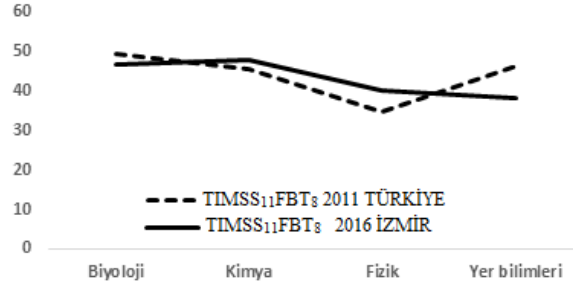
Çalışmanın bu kısmında TIMSS₁₁FBT₈'nin 2011 Türkiye verileriyle 2016 İzmir verilerinin betimsel ve ilişkisel bulguları verilmiş; Korelasyonel ve Path analiziyle TIMSS₁₁FBT₈'ye yönelik öğrenme alanlarının katkı düzeyleri ortaya konulmaya çalışılmıştır.

Çalışmanın Birinci Amacı Olarak; Tablo 2'de TIMSS₁₁FBT₈'nin 2011 Türkiye verileriyle 2016 İzmir verilerinin biyoloji, kimya, fizik ve yer bilimleri öğrenme alanlarına göre ortalama ve standart sapma gibi betimsel değerler açısından karşılaştırılması yapılmıştır.

Tablo 2. TIMSS₁₁FBT₈'nin 2011 Türkiye Verileriyle 2016 İzmir Verilerinin Öğrenme Alanlarına Göre Betimsel Değerler Açısından Karşılaştırılması

| | TIMSS ₁₁ FBT ₈ 2011 Türkiye | | TIMSS ₁₁ FBT ₈ 2016 İzmir | |
|----------------------|--|-------|--|-------|
| | \bar{x} | ss | \bar{x} | ss |
| Biyoloji | 49.42 | 30.83 | 46.50 | 22.30 |
| Kimya | 45.65 | 31.24 | 47.91 | 22.59 |
| Fizik | 34.70 | 26.42 | 39.94 | 21.35 |
| Yer Bilimleri | 46.18 | 39.48 | 38.20 | 20.55 |





Grafik 1. TIMSS₁₁FBT₈'nin 2011 Türkiye Verileriyle 2016 İzmir Verilerinin Öğrenme Alanlarına Göre Betimsel Değerler Açısından Karşılaştırıldığı Grafik

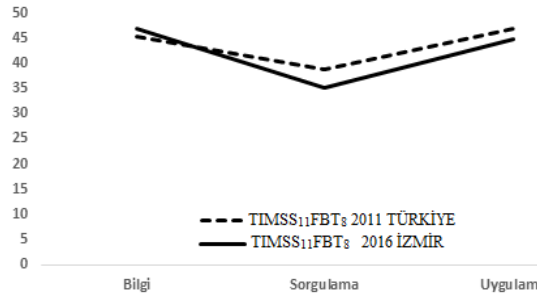
TIMSS₁₁FBT₈'nin 2016 İzmir verilerine göre biyoloji konularına ilişkin puan ortalaması 2016 Türkiye verilerine göre daha düşük bulunmuştur. Aynı şekilde buna benzer bulgu TIMSS₁₁FBT₈'nin Yer Bilimleri alanındaki veriler içinde söylenebilir. Fakat TIMSS₁₁FBT₈'nin 2016 İzmir verilerine göre fizik konularına ilişkin puan ortalaması 2011 Türkiye verilerine göre biraz yüksek bulunmuştur. Aynı şekilde buna benzer bulgu TIMSS₁₁FBT₈'nin kimya alanındaki veriler içinde geçerlidir.

Çalışmanın İkinci Amacı Olarak; Tablo 3'te TIMSS₁₁FBT₈'nin 2011 Türkiye verileriyle 2016 İzmir verilerinin bilişsel, sorgulama ve uygulama düzeylerine göre ortalama ve standart sapma gibi betimsel değerler açısından karşılaştırılması yapılmıştır.

Tablo 3. TIMSS₁₁FBT₈'nin 2011 Türkiye Verileriyle 2016 İzmir Verilerinin Bilişsel, Uygulama ve Sorgulama Düzeylerine Göre Betimsel Değerler Açısından Karşılaştırılması

| | TIMSS ₁₁ FBT ₈ 2011 Türkiye | | TIMSS ₁₁ FBT ₈ 2016 İzmir | |
|------------------|--|-------|--|-------|
| | \bar{x} | SS | \bar{x} | SS |
| Bilgi | 45.38 | 22.10 | 46.85 | 20.07 |
| Sorgulama | 38.80 | 32.97 | 35.14 | 19.57 |
| Uygulama | 46.93 | 31.98 | 44.77 | 24.17 |





Grafik 2. TIMSS₁₁FBT₈'nin 2011 Türkiye Verileriyle 2016 İzmir Verilerinin Bilişsel, Uygulama ve Sorgulama Düzeylerine Göre Betimsel Değerler Açısından Karşılaştırıldığı Grafik

TIMSS₁₁FBT₈'nin 2016 İzmir verilerine göre bilgi alanına ilişkin puan ortalaması 2011 Türkiye verilerine göre birbirine yakın bulunmuştur. TIMSS₁₁FBT₈'nin 2016 İzmir verilerine göre uygulama alanına ilişkin puan ortalaması 2011 Türkiye verilerine göre daha düşük olduğu görülmüştür. Aynı şekilde buna benzer bulgu TIMSS₁₁FBT₈'nin sorgulama alanındaki veriler içinde söylenebilir.

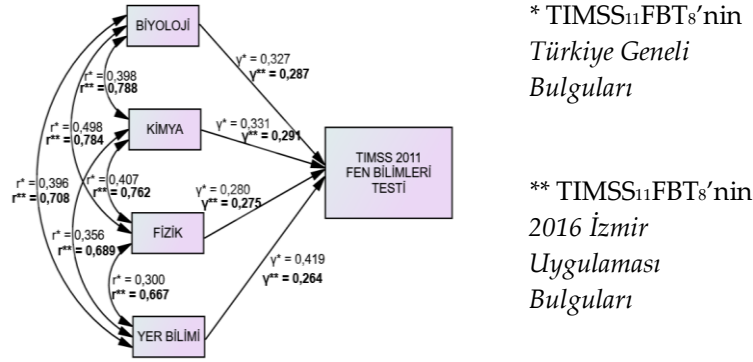
Çalışmanın Üçüncü Amacı Olarak; Tablo 4'te TIMSS₁₁FBT₈'nin biyoloji, kimya, fizik ve yer bilimleri öğrenme alanlarına göre 2011 Türkiye verileriyle 2016 İzmir verilerinin ilişkisel karşılaştırılması yapılmış ve TIMSS₁₁FBT₈'ye yönelik ağırlıklarının belirlendiği Path analiziyle katkı değerleri belirlenmiştir.

Tablo 4. TIMSS₁₁FBT₈'nin 2011 Türkiye Verileriyle 2016 İzmir Verilerinin Öğrenme Alanlarına Göre İlişkisel Açıdan Path Analiziyle Karşılaştırılması

| | | TIMSS ₁₁ FBT ₈ 2011 Türkiye | | TIMSS ₁₁ FBT ₈ 2016 İzmir | |
|------------|----------|--|-------|--|-------|
| Korelasyon | | r_x | Y | r_x | Y |
| Biyoloji | FBT | 0.764 | 0.327 | 0.919 | 0.287 |
| Kimya | FBT | 0.724 | 0.331 | 0.909 | 0.291 |
| Fizik | FBT | 0.703 | 0.280 | 0.898 | 0.275 |
| Yer Bil. | FBT | 0.750 | 0.419 | 0.851 | 0.264 |
| Biyoloji | Fizik | 0.498 | - | 0.784 | - |
| Biyoloji | Kimya | 0.398 | - | 0.788 | - |
| Biyoloji | Yer Bil. | 0.396 | - | 0.708 | - |
| Kimya | Yer Bil. | 0.356 | - | 0.689 | - |
| Fizik | Kimya | 0.407 | - | 0.762 | - |
| Fizik | Yer Bil. | 0.300 | - | 0.667 | - |

TIMSS₁₁FBT₈'nin 2016 İzmir verilerine göre biyoloji, fizik, kimya, yer bilimlerine ilişkin korelasyonları, 2011 Türkiye verilerine göre yüksek korelasyonların elde edildiği görülmüştür.





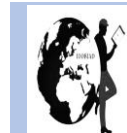
Şekil 1. TIMSS₁₁FBT₈'nin 2011 Türkiye Verileriyle 2016 İzmir Verilerinin Öğrenme Alanlarına Göre İlişkisel Açından Path Analiz Diyagramı

Şekil 1'de TIMSS₁₁FBT₈'nin 2011 Türkiye verileriyle 2016 İzmir verilerinin Biyoloji, Fizik, Kimya ve Yer Bilimleri ile olan doğrusal ilişkileri verilmiştir. Diyagramdaki dışsal (TIMSS₁₁FBT₈) ve içsel (TIMSS₁₁FBT₈'nin Biyoloji, Fizik, Kimya ve Yer Bilimleri) değişkenler arasındaki ilişkiler yol analiziyle incelenmiştir. Analiz sonucunda, TIMSS₁₁FBT₈'nin 2011 Türkiye verilerine göre öğrenme alanlarının TIMSS₁₁FBT₈ test puanlarına katkı değerleri incelendiğinde; Biyolojinin katkı değeri $\gamma = 0.327$, Fiziğin katkı değeri $\gamma = 0.280$, Kimyanın katkı değeri $\gamma = 0.331$ ve Yer Bilimlerinin katkı değeri $\gamma = 0.419$ olarak hesaplanmıştır. TIMSS₁₁FBT₈ puanlarına en fazla katkının yer bilimleri öğrenme alanından en az katkının fizik öğrenme alanından geldiğini söyleyebiliriz. Aynı şekilde TIMSS₁₁FBT₈'nin 2016 İzmir verilerine göre öğrenme alanlarının TIMSS₁₁FBT₈ test puanlarına katkı değerleri incelendiğinde; Biyolojinin katkı değeri $\gamma = 0.287$, Fiziğin katkı değeri $\gamma = 0.275$, Kimyanın katkı değeri $\gamma = 0.291$ ve Yer Bilimlerinin katkı değeri $\gamma = 0.264$ olarak hesaplanmıştır. TIMSS₁₁FBT₈'nin 2016 İzmir verilerine göre en fazla katkının kimya öğrenme alanından en az etkinin yer bilimleri öğrenme alanından geldiğini söyleyebiliriz.

Çalışmanın Dördüncü Amacı Olarak; Tablo 5'te TIMSS₁₁FBT₈'nin bilgi, sorgulama ve uygulama düzeylerine göre 2011 Türkiye verileriyle 2016 İzmir verilerinin ilişkisel karşılaştırılması yapılmış ve TIMSS₁₁FBT₈'ye yönelik ağırlıklarının belirlendiği Path analiziyle katkı değerleri belirlenmiştir.

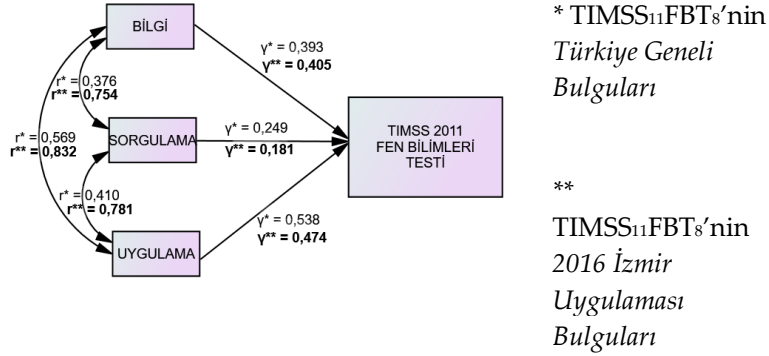
Tablo 5. TIMSS₁₁FBT₈'nin 2011 Türkiye Verileriyle 2016 İzmir Verilerinin Bilişsel, Uygulama ve Sorgulama Düzeylerine Göre İlişkisel Açından Path Analiziyle Karşılaştırılması

| TIMSS ₁₁ FBT ₈ 2011 Türkiye | TIMSS ₁₁ FBT ₈ 2016 İzmir |
|--|--|
|--|--|



| Korelasyon | | r_x | Y | r_x | Y |
|------------|-----------|-------|-------|-------|-------|
| Bilgi | FBT | 0.793 | 0.393 | 0.937 | 0.405 |
| Sorgulama | FBT | 0.617 | 0.249 | 0.857 | 0.181 |
| Uygulama | FBT | 0.864 | 0.538 | 0.953 | 0.474 |
| Bilgi | Sorgulama | 0.376 | - | 0.754 | - |
| Bilgi | Uygulama | 0.569 | - | 0.832 | - |
| Sorgulama | Uygulama | 0.410 | - | 0.781 | - |

TIMSS₁₁FBT₈'nin 2016 İzmir verilerine göre bilgi, sorgulama ve uygulama alanına ilişkin korelasyonları, 2011 Türkiye verilerine göre yüksek korelasyonların elde edildiği görülmüştür.



Şekil 2. TIMSS₁₁FBT₈'nin 2011 Türkiye Verileriyle 2016 İzmir Verilerinin Bilişsel, Uygulama ve Sorgulama Düzeylerine Göre İlişkisel Açından Path Analiz Diyagramı

Şekil 2'de TIMSS₁₁FBT₈'nin 2011 Türkiye'de verileriyle 2016 İzmir verilerinin bilgi, sorgulama ve uygulama alanlarıyla olan doğrusal ilişkileri verilmiştir. Diyagramdaki dışsal (TIMSS₁₁FBT₈) ve içsel (TIMSS₁₁FBT₈'nin bilgi, sorgulama ve uygulama alanları) değişkenler arasındaki ilişkiler yol analiziyle incelenmiştir. Analiz sonucunda, TIMSS₁₁FBT₈'nin Türkiye verilerine göre bilişsel, sorgulama ve uygulama düzeylerinin TIMSS₁₁FBT₈ test puanlarına katkı değerleri incelendiğinde; bilgi düzeyinin katkı değeri $Y = 0.393$, sorgulama düzeyinin katkı değeri $Y = 0.249$ ve uygulama düzeyinin katkı değeri $Y = 0.538$ olarak hesaplanmıştır. TIMSS₁₁FBT₈'nin Türkiye verilerine göre en fazla katkının uygulama düzeyinden en az etkinin sorgulama düzeyinden geldiği söylenebilir. Aynı şekilde TIMSS₁₁FBT₈'nin 2016 İzmir verilerine göre bilgi, sorgulama ve uygulama düzeylerinin TIMSS₁₁FBT₈ test puanlarına katkı değerleri incelendiğinde; bilgi düzeyinin katkı değeri $Y = 0.405$, Sorgulama düzeyinin katkı değeri $Y = 0.181$ ve Uygulama düzeyinin katkı değeri $Y = 0.474$ olarak hesaplanmıştır. TIMSS₁₁FBT₈'nin 2016 İzmir verilerine göre en fazla katkının uygulama düzeyinden en az katkının sorgulama düzeyinden geldiği söylenebilir.

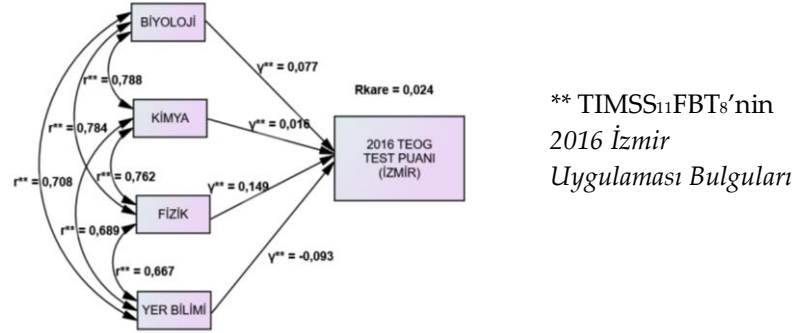


Çalışmanın Beşinci Amacı Olarak; Tablo 6’da TIMSS₁₁FBT₈’nin biyoloji, kimya, fizik ve yer bilimleri öğrenme alanlarına göre 2016 İzmir verileriyle aynı öğrencilerin 2016 TEOG puanlarının ilişkisel karşılaştırılması yapılmış ve 2016 TEOG puan ağırlıklarının belirlendiği Path analiziyle katkı değerleri belirlenmiştir.

Tablo 6. TIMSS₁₁FBT₈’nin 2016 İzmir Verilerine Göre Öğrenme Alanlarından Elde Edilen Puanların TEOG Sınavı Puanları Açısından Path Analiziyle Karşılaştırılması

| MEB TEOG TESTİ | TIMSS ₁₁ FBT ₈ 2016 İzmir (Öğrenme Alanlarına Göre) | r_x | Y |
|------------------------------|---|-------|--------|
| TEOG 2016 Fen Bilimleri Test | Biyoloji | 0.142 | 0.077 |
| | Kimya | 0.159 | 0.149 |
| | Fizik | 0.129 | 0.016 |
| Toplam Puanı | Yer Bilimleri | 0.075 | -0.093 |

TIMSS₁₁FBT₈’nin 2016 İzmir verilerine göre biyoloji, fizik, kimya konularına ilişkin puanları, Kasım 2016 TEOG puanlarıyla karşılaştırıldığında çok az ilişkili korelasyonların elde edildiği görülmüştür.



Şekil 3. TIMSS₁₁FBT₈’nin 2016 İzmir Verilerine Göre Öğrenme Alanlarından Elde Edilen Puanların TEOG sınavı Puanları Açısından Path Analiz Modeli

Şekil 3’te 2016 yılında TEOG Fen bilimleri test puanlarıyla TIMSS₁₁FBT₈’nin 2016 İzmir verilerine göre Biyoloji, Fizik, Kimya ve Yer Bilimleri öğrenme alanlarıyla olan doğrusal ilişkileri verilmiştir. Diyagramdaki dışsal (TEOG 2016 Fen bilimleri testi) ve içsel (TIMSS₁₁FBT₈’nin Biyoloji, Fizik, Kimya ve Yer Bilimleri) değişkenler arasındaki ilişkiler yol analiziyle incelenmiştir. Analiz sonucunda, TIMSS₁₁FBT₈’nin 2016 İzmir verilerine göre 2016 TEOG Fen bilimleri test puanlarına öğrenme alanlarının katkı değerleri incelendiğinde; Biyoloji öğrenme alanının katkı değeri $Y = 0.077$, Fizik öğrenme alanının katkı değeri $Y = 0.016$, Kimya öğrenme alanının katkı değeri $Y = 0.149$ ve Yer Bilimleri öğrenme alanının katkı değeri $Y = -0.093$ olarak hesaplanmıştır. TIMSS₁₁FBT₈’nin 2016 İzmir verilerine göre 2016 yılı TEOG Fen bilimleri testi puanlarına öğrenme alanlarının katkısının olmadığı söylenebilir.

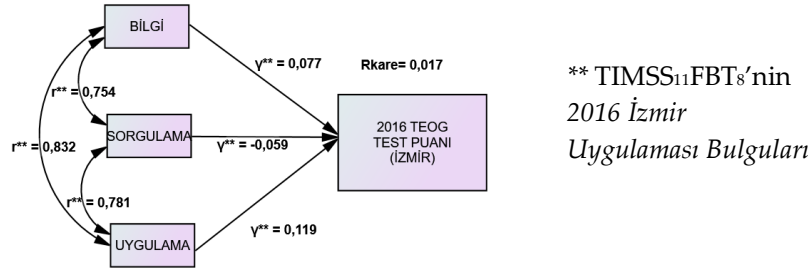


Çalışmanın Altıncı Amacı Olarak; Tablo 7'de TIMSS₁₁FBT's'nin bilgi, sorgulama ve uygulama düzeylerine göre 2016 İzmir verileriyle aynı öğrencilerin 2016 TEOG Fen bilimleri testi puanlarının ilişkisel karşılaştırılması yapılmış ve 2016 TEOG puan ağırlıklarının belirlendiği Path analiziyle katkı değerleri belirlenmiştir.

Tablo 7. TIMSS₁₁FBT's'nin 2016 İzmir Verilerine Göre Bilişsel, Uygulama ve Sorgulama Düzeylerinden Elde Edilen Puanların TEOG Sınavı Puanları Path Analiziyle Karşılaştırılması

| | | TIMSS ₁₁ FBT's 2016 İZMİR | |
|------------------------------|-----------|---|--------|
| Korelasyon | | r_x | Y |
| TEOG 2016 Fen Bilimleri Test | Bilgi | 0.133 | 0.077 |
| | Sorgulama | 0.093 | -0.059 |
| | Uygulama | 0.138 | 0.119 |

TIMSS₁₁FBT's'nin 2016 İzmir verilerine göre bilgi ve uygulama alanına ilişkin puanları, Kasım 2016 TEOG sınavının puanları karşılaştırıldığında düşük düzeyde korelasyonların elde edildiği görülmüştür.



Şekil 4. TIMSS₁₁FBT's'nin 2016 İzmir Verilerine Göre Bilişsel, Uygulama ve Sorgulama Düzeylerinden Elde Edilen Puanların TEOG sınavı puanları Path Analiz Modeli

Şekil 4'te 2016 yılında TEOG Fen bilimleri test puanlarıyla TIMSS₁₁FBT's'nin 2016 İzmir verilerine göre bilgi, sorgulama ve uygulama alanlarıyla olan doğrusal ilişkileri verilmiştir. Diyagramdaki dışsal (TEOG 2016 Fen bilimleri testi) ve içsel (TIMSS₁₁FBT's'nin bilgi, sorgulama ve uygulama) değişkenler arasındaki ilişkiler yol analizi ile incelenmiştir. Analiz sonucunda, TIMSS₁₁FBT's'nin 2016 İzmir verilerine göre 2016 TEOG Fen bilimleri test puanlarına katkı değerleri incelendiğinde; bilgi düzeyinin katkı değeri $Y = 0.077$, sorgulama düzeyinin katkı değeri $Y = -0.059$ ve uygulama düzeyinin katkı değeri $Y = 0.119$ olarak hesaplanmıştır. TIMSS₁₁FBT's'nin 2016 İzmir verilerine göre 2016 yılı TEOG Fen bilimleri testi puanlarına bilgi, sorgulama ve uygulama düzeylerinin katkısının olmadığı söylenebilir.



Tartışma ve Sonuç

Bir ölçme aracının tek boyutta ölçme yaptığının ve ölçeğin iç tutarlılığının yüksek olduğunun kanıtı olarak ölçme aracının test maddeleriyle kendi içinde ve toplam test puanıyla yüksek korelasyona sahip olması gösterilebilir (Kan, 2008). Ölçme aracı geliştirilirken ölçeğin tekrar uygulanma durumu olmadığı durumlarda ya da ona benzer testlerin uygulanmadığı durumlarda Cronbach Alfa güvenilirlik ve Spearman-Brown güvenilirlik analizleri yapılabilmektedir (Tavşancıl, 2002; Tekin, 1991). Ölçme araçlarının iç tutarlılık göstergesi olan güvenilirlik katsayılarının 0.70 ve üstü değerlerde olması güvenilirlik açısından kritik değer olarak kabul görmektedir (Tavşancıl, 2002). TIMSS₁₁FBT's'nin 2011 Türkiye verileriyle 2016 İzmir verilerin güvenilirlik katsayıları 0.70'in üstünde çıkması yapılan çalışmada toplanan verilerin güvenilir olduğunu göstermektedir. TIMSS₁₁FBT's'nin 2016 İzmir verilerin güvenilirlik katsayısının daha yüksek çıkmasının nedeni olarak testin tek boyut olarak bir kerede bir örnekleme uygulanması gösterilebilir. Güvenirlik katsayıları formülleri ölçekteki ya da testteki madde sayısı ile ilişkilendirildiğinden madde sayısı arttıkça güvenilirlik katsayısında da bir artış olabilmektedir. TIMSS₁₁FBT's altı boyuta bölünerek farklı örneklemeler üzerinden 2011 yılında Türkiye'de uygulandığından ve elde edilen güvenilirlik katsayısı değeri daha az madde içeriğiyle hesaplandığından değeri daha düşük olması doğaldır. Çünkü 80 maddelik test altı boyutta ve altı örneklemeden elde edilmiştir. Her boyut için hesaplanan güvenilirlik katsayısı, kullanılan formülünde bir hata değeri olduğundan testin tümü için 6 defa hata oranı miktarı karışmaktadır. TIMSS₁₁FBT's'nin 2016 yılında İzmir ilinde tekrarlanması ve güvenilirlik katsayısı değerinin yüksek çıkması karışan hatanın büyüklüğü hakkında yaklaşık olarak bilgi vermektedir.

TIMSS₁₁FBT's'nin 2016 İzmir verilerine göre biyoloji ve yer bilimleri konularındaki ortalama puanları TIMSS₁₁FBT's'nin 2011 Türkiye verilerine göre daha düşüktür. Fakat TIMSS₁₁FBT's'nin 2016 İzmir verilerine göre kimya ve fizik konularındaki ortalama puanları TIMSS₁₁FBT's'nin 2011 Türkiye verilerine göre daha ileri düzeydedir. Karamustafaoğlu ve Sontay'ın (2012) TIMSS 2011 sınavına katılan öğrencilerle ve öğretmenlerle yaptığı görüşmeler doğrultusunda öğrencilerin ve öğretmenlerin bu sınava yönelik bir hazırlık yaptıklarını belirtmişlerdir. Ayrıca öğretmenler bu görüşlerinde sınavın cevaplama süresinin yetersiz kaldığını, öğrencilerin soruları anlamada zorluk çektiklerini, sorgulama türü soruları cevaplama yetersiz kaldıklarını ve TIMSS 2011 sınavı öncesi hazırlık süresinin yetersiz olduğunu ifade etmişlerdir. TIMSS₁₁FBT's'nin İzmir ilinde 2016 yılında uygulanan öğrenci örnekleme önceden TIMSS sınavına ilişkin hiçbir eğitim almamıştır. Eğitim almamasına rağmen 2011 verilerine yakın ortalama puanlar elde edilmiştir. TIMSS sınavları öncesi öğrencilerin eğitim aldığı düşünüldüğünde İzmir ilinin bu sınava yönelik az da olsa artılarının olduğu söylenebilir. Ayrıca öğretmen görüşlerine göre öğrencilerin sorgulama



sorularını yapmada yetersiz kalması bu çalışmanın bulgularıyla örtüşmektedir. Bunun kaynağı olarak dil yetersizlikleri gösterilebilir. Delil ve Tetik (2015) çalışmasında Türkiye’de yapılan ulusal merkezi sınavları günümüze kadar yapılan TIMSS sınavlarıyla bilişsel düzeyde karşılaştırmış ve sonuç olarak dramatik düzeyde farklılıklar olduğunu görmüştür. Bu sonuca gerekçe olarak ulusal, merkezi olarak yapılan sınavlarda soru hazırlamayla ilgili standart bir sınav çerçevesinin olmadığını belirtmiştir. Öğrencilerin ve öğretmenlerin belirli standart bir sınav ölçme ve değerlendirmelerine alışık olmaması 2011 yılında ülke genelinde yapılan ve 2016 yılında İzmir ilinde tekrarlanan TIMSS₁₁FBT’s’nin sonuçlarına yansdığı göstermektedir.

TIMSS₁₁FBT’s’nin 2016 İzmir verilerine göre uygulama ve sorgulama alanlarındaki ortalama puanları TIMSS₁₁FBT’s’nin 2011 Türkiye verilerine göre daha düşüktür. TIMSS₁₁FBT’s’nin 2016 İzmir verilerine göre bilgi alanındaki ortalama puanı TIMSS₁₁FBT’s’nin 2011 Türkiye verilerine göre benzerlik göstermektedir. TIMSS₁₁FBT’s’nin 2011 Türkiye verilerine göre sorgulama alanına yönelik ortalama puanı TIMSS₁₁FBT’s’nin 2016 İzmir verilerine göre daha düşüktür. Bilgi alanındaki zihinsel faaliyetler uygulama ve sorgulama alanlarına kıyasla daha temel ve kolay faaliyetlerdir (Yıldırım, Yıldırım, Ceylan ve Yetişir, 2013). Yıldırım vd. (2013) yaptığı araştırmalarının bulgularında farklılıklar görülmektedir. Bu çalışmada incelemeye alınan sorular seçenekli sorular olup karşılaştırmalar buna göre yapılmıştır. Bu bulgulara göre hem ülke genelindeki hem de İzmir ilindeki öğrencilerin bilgi düzeyinde daha başarılı oldukları, sorgulama ve uygulama düzeyinde bu düzeyin düştüğü görülmektedir. Yıldırım vd. (2013) çalışmasında ise bu başarı düzeyleri terslik göstermektedir. Çünkü Yıldırım vd. (2013) yaptığı incelemelerde Fen bilimleri testi puanı hem diğer bağıl değişkenlere göre hesaplanmış, hem de açık uçlu değerlendirmelerde hesaba dâhil edilmiştir. Bu bulgular doğrultusunda açık uçlu sınav sisteminin öğrencilerin uygulama ve sorgulama düzeyindeki düzeyleri ölçmede oldukça etkili olduğu söylenebilir. MEB’in ve ÖSYM’nin açık uçlu soru sistemine geçmesinin uluslararası sınavlarda öğrencilerin uygulama ve sorgulama düzeylerini geliştirmede alışkanlık kazandırması açısından önemli olduğu söylenebilir.

Abazaoğlu vd. (2014) TIMSS 2011 verilerini okul yerinin Fen bilgisi başarısına etkisine yönelik etkilerini araştırdığında; nüfusu 100.000 üzeri olan yerleşim yerlerinde elde edilen Fen bilimleri başarı puanının nüfus yoğunluğuna paralel olarak azaldığını belirlemiştir. TIMSS₁₁FBT’s’nin 2016 İzmir verilerine göre ortalama puanların TIMSS₁₁FBT’s’nin 2011 Türkiye verilerine göre puanların altında olması çalışma sırasında 2016 yılında İzmir’den seçilen örneklemin şehir merkezinden yapılması gösterilebilir. Çalışma sırasında köy, kasaba ve kırsal kesimlerden örnekleme dair bir uygulama yapılmamıştır. Bu çalışmadan elde edilen verilerin ülke geneline



göre düşük olmasında nüfus yoğunluğunun etkisi olduğu söylenebilir. Ancak öğrenci başarılarını ve davranışlarını tek bir nedene bağlamak doğru olmaz. Eğitim ve öğretim sürecine dayalı davranış değiştirme eylemleri olduğundan ve bu süreçte öğrencilerin birden fazla duruma maruz kaldığından nüfus yoğunluğundan başka etkenler de incelenebilir. Abazaoğlu vd. (2014) TIMSS 2011 verilerine göre Fen bilimlerinde Türkiye için en sorunlu alanların kimya ile dünya ve yer bilimleri olduğu ifade edilmiştir. Abazaoğlu vd. çalışmalarında bu yorumu yaparken TIMSS 2011 testinin tümünü ve açık uçlu soruları da kapsayan puanlama sistemine göre değerlendirmişlerdir. Bu çalışmada ise sadece seçenekli sorular dikkate alınmıştır. Seçenekli sonuçlara göre ise Türkiye için en sorunlu alan Fizik, İzmir ili için en sorunlu alan Fizik ve Yer Bilimleri olduğu anlaşılmaktadır. Öğrencilerin açık uçlu sorularda Fizik ve Yer Bilimleri hakkındaki düşüncelerini rahatlıkla açıklayabildikleri, testlerde ise bunu yapamadıkları söylenebilir. Fizik kavramları, anlaşılması zor ve soyut kavramlardır. TIMSS çalışmalarının Fen ve Matematik eğitiminin gelişmesine birçok katkılarının yanında bu çalışmalara eleştirel yaklaşan çalışmalarda bulunmaktadır. Bu çalışmalarda çoktan seçmeli testlerle ve kısa cevaplı sorularla beceri ve uygulama düzeyindeki kavramları ölçmede yetersiz olduğu ifade edilmiştir (Öztürk ve Uçar, 2010). TIMSS’de başarısız olan öğrencilerle yapılan görüşmeler sonucunda öğrencilerin aslında daha başarılı olduğu, TIMSS testlerini dilsel olarak anlamadıklarını ve dolayısıyla cevaplayamadıkları ortaya çıkmıştır. Testlerde öğrencilerin yaşadıkları ortamın diline uygun bilimsel dil kullanılmadığı için TIMSS’ de başarı sağlanamadığı anlaşılmıştır (Harlow ve Jones, 2004; Kelly, 2002; Olkun ve Aydoğdu, 2003; Öztürk ve Uçar, 2010).

Uzun vd. (2010) TIMSS 1999 ve 2007 verilerine göre Türk öğrencilerin çoğunluğunun Fen bilimleri dersine yönelik olumlu bir tutuma sahip olduklarını ifade etmiş ve bu öğrencilerin bilişsel başarı sıralamasında ilk beşe giren ülke öğrencilerini Fen dersine yönelik tutum açısından geride bıraktıklarını belirtmişlerdir. Aynı çalışmada bilişsel başarı sıralamasında ilk beşe giren ülke öğrencilerinin haftalık ödevlere zaman ayırma sürelerine göre incelendiğinde Türk öğrencilerin ödevlerine daha fazla zaman ayırdığı belirlenmiştir. Türkiye’de öğrencilerin ödevlere fazla zaman ayırmasının Fen bilimleri başarısında çok etkili olmadığı görülmüştür. Türkiye’nin TIMSS 1999 ve 2007 Fen bilimleri test verilerine göre genel ortalamanın altında kalmıştır. Bu durumda Türkiye için Fen bilimleri dersi tutumuyla Fen bilimleri başarısı arasındaki ters korelasyon tartışılır olmuştur. Çeken (2011) çalışmasında 2005-2010 yıllarında TÜBİTAK tarafından düzenlenen Benim Eserim Projelerinin Türkiye değerlendirmesini yapmıştır. En fazla proje kazanan iller olarak birinci sırada İstanbul ve ikinci sırada İzmir yer almıştır. Ayrıca ödül kazanılmasa dahi bu iki il, proje başvuru ve değerlendirme sıralamasında başı çekmektedir. TIMSS₁₁FBT’s’nin 2016 İzmir verilerine göre akademik başarının TIMSS₁₁FBT’s’nin 2011 Türkiye verilerine göre daha



düşük olması, ayrıca uygulama ve sorgulama alanında ortalama puanların düşük çıkması öğrencilerin proje ödevleriyle meşgul edilmesi gösterilebilir. Çünkü öğrencilerin tutumları ve ödevlere zaman ayırmayla tutumları arasındaki ters orantı bu sonucu doğurabilir. Türkiye’de ödev ve proje ödevlerine ilişkin olarak ailelerin yetersiz kalmasından dolayı öğrencilerin daha fazla zaman ayırmak zorunda kalmaları gösterilebilir (Aksu ve Karaçöp, 2015; Duru ve Çoğmen, 2017). Çünkü veliler Fen bilimleri dersi konusunda yeterli deneyimleri olmadığı için konunun anlaşılması, malzemelerin tedarik edilmesi ve öğrencilere yardım konusunda gerekli desteği sağlayamamaktadır. Bu nedenle öğrenciler gereğinden fazla zaman ayırmakta, öğrenme sürecinde projede verilen kavramın odağından uzaklaşmakta ve ödev/proje bir hobi deneyimine dönüşmektedir.

2013 yılında Fen bilimleri programıyla araştırma-sorgulama yaklaşımını temel alan bir programa geçilmiştir. MEB, bu konudaki eksikliği görüp Fen bilimleri eğitiminde güncellemeye gitmiştir. MEB, 2017 ve 2018 yıllarında tekrar Fen bilimleri dersi öğretim programını FeTeMM, Kodlama gibi yeni içeriklerle güncelleyerek araştırma-sorgulama yaklaşımını güçlendiren uygulamaları programa dâhil etmiştir (MEB, 2017; MEB, 2018). Bu çalışmadan elde edilen bulgular, MEB’in araştırma-sorgulama yaklaşımıyla ilgili güncellemeye yönelik tutumları desteklemektedir. Bu konudaki anlamlı değişimi anlayabilmek için zamana ihtiyaç vardır. Ayrıca TIMSS₁₁FBT₈ verilerine göre hem Türkiye’de hem de İzmir’de fizik, kimya, biyoloji konuları arasında yüksek korelasyon görülürken, yer bilimlerinin bu üç konu alanındaki düşük korelasyona sahip olması ilginç bir bulgudur. Düşük korelasyonun nedenleri; 1. Öğretmenlerin yer bilimleri konusunda yeterlilikleri, 2. Yer bilimleri ünitesinin okulun son günlerine denk gelmesi, 3. Yer bilimleri konu içeriğinin bu üç konu alanıyla yeterli ilişkilendirilememesi olabilir. MEB Fen bilimleri dersi öğretim programının içeriği 2017 yılında yeniden güncellenmiştir. Bu güncellemede bu çalışmayla ilişkili olduğu düşünülen güncelleme ise konuların ve ünitelerin yerlerinin değiştirilmesidir. Eğitim-öğretim yılının sonuna denk gelen bazı üniteler, uluslararası sınavlardaki başarısızlıklar ve yıl sonunda yapılan öğrencilerin devamsızlıkları gerekçe gösterilerek eğitim-öğretim yılının başlarına çekilmiştir (MEB, 2017; MEB, 2018). Milli Eğitim Bakanlığı’nın bu güncellemesi bu araştırmanın bulgularını destekler nitelik taşımaktadır. Ayrıca hem Türkiye’de hem de İzmir’de uygulama ve bilgi alanında yüksek korelasyon görülürken, sorgulama alanının bu iki konu alanındaki düşük korelasyona sahip olması ilginç bir bulgudur. Sorgulama yaklaşımını etkili kullanabilmek için ders içerikleri uygulama alanı içinde sorgulama ile düzenlenmesi gerekmektedir.

TIMSS₁₁FBT₈’nin 2016 İzmir verilerine göre biyoloji, fizik, kimya konularına ilişkin puanları, Kasım 2016 TEOG puanlarıyla karşılaştırıldığında çok az ilişkili korelasyonların elde edildiği görülmüştür. Yer Bilimleri alanının



neredeyse TEOG Fen bilimleri testi puanıyla ilişkili olmadığı söylenebilir. TEOG sınavı TIMSS ile ilişkili olmadığı ortaya çıkmaktadır. Eğer biz TIMSS gibi sınavlarda etkin bir başarı yakalamak istiyorsak, TEOG gibi ulusal sınavları bu tür sınavların içeriğine yaklaştırmalıyız. Bu konuda MEB, uzmanlarından destek almalı ve Ankara merkezinin dışında kurulacak bölgesel merkezli birimler açarak araştırmalarını derinleştirmelidir.

TIMSS₁₁FBT's'nin 2016 İzmir verilerine göre bilgi ve uygulama alanına ilişkin puanları, Kasım 2016 TEOG sınavının puanları karşılaştırıldığında düşük düzeyde korelasyonların elde edildiği görülmüştür. Sorgulama alanı ile Kasım 2016 TEOG puanları arasında korelasyonel bir ilişki bulunamamıştır. Bu bulgu ile bir önceki bulguda görüldüğü gibi TEOG sınavı TIMSS ile ilişkili olmadığı ortaya çıkmaktadır. MEB; TEOG sınavında açık-uçlu sınav sorularını sormak istemesi TIMSS sınavlarına ilişkin değerlendirmelerini takip ettiği söylenebilir. TIMSS sınavlarında açık-uçlu sorular ve farklı değerlendirmeleri bulunmaktadır. TEOG sınavları sadece dört seçenekli sorulardan oluşmaktadır. Bu farklılık karşılaştırmaları etkilemektedir.

Öneriler

Bu araştırmada, 2016 yılında İzmir merkezinde yer alan ortaokul sekizinci sınıf öğrencilerine TIMSS₁₁FBT's'nin uygulanması sonucu veriler toplanmış ve 2011 yılının TIMSS ortaokul sekizinci sınıf Fen bilimleri testinin verileriyle karşılaştırılmıştır. Bu nedenle farklı illerden seçilen örneklerle karşılaştırmalara gidilmesi önerilir. Ayrıca karşılaştırma sadece Fen bilimleri testiyle sınırlıdır. TIMSS kapsamında kullanılan farklı testlerle ilgili olarak bölgesel karşılaştırmaların yapılması anlamlı sonuçlar üretebilir.

Bu araştırmada TIMSS verileriyle TEOG sınavı verilerinin ilişkili olmadığı ve sınav türlerinin birbirine benzemediği bulgularla ortaya konmuştur. Ulusal sınavlarda kullanılan soru madde ve içeriklerinin uluslararası sınav içeriklerine uyarlanacak çalışmalara ağırlık verilmelidir. Öğrencilerin uluslararası sınavlarda yer alan test maddelerine uyum sağlaması için öğretmenlerin bu sınavlarda yer alan soru maddeleriyle ilgili ölçme ve değerlendirme becerilerinin geliştirilmesi ve konuya ilişkin yeterliliklerin kazandırılması önerilir.

Bu araştırmanın bulgularında, öğrencilerin sorgulama düzeylerinin her iki örneklem grubunda da yeterli düzeyde olmadığı görülmektedir. Öğrencilerin sorgulama düzeylerinin iyileştirilmesi için sınıf içi etkinliklerde sorgulamaya yönelik alternatif değerlendirme yaklaşımları benimsenmelidir. Akademik başarı değerlendirmeleri sorgulamaya dayalı olarak planlanmalıdır.

Bu araştırmanın 2016 yılı İzmir örneğinde veri toplama biçimiyle TIMSS 2011 Fen bilimleri testine dayalı olarak veri toplama biçimi farklılıklar göstermektedir. 2016 yılı İzmir örneğinde veriler, kısıtlı imkanlar ve olanaklar dâhilinde toplandığından böyle bir biçim farklılığı oluşmuştur.



Eğitim bilimlerinden elde edilen istatistiki değerlendirmelerle ilgili genellemeler, pozitif bilimlerdeki gibi kesin olarak ya da büyük olasılık oranlarında çok küçük hata oranlarıyla yapılamamaktadır. Çok küçük hata oranlarıyla sonuç üretebilmek için testlerin aynı anda, aynı süre aralığında ve büyük destek hizmetleriyle yapılması gerekmektedir. Daha kesin sonuçlara ulaşabilmek ve gerekli önlemleri alabilmek için uluslararası sınavlardan sonra ara dönemler halinde Milli Eğitim Bakanlığı desteğiyle böylesi sınavlar yapılmalıdır.

Kaynaklar

Abazaoğlu, İ., Yıldızhan, Y. ve Yıldırım, O. (2014). TIMSS 2011 Türkiye 8. Sınıf Fen Bilimleri Sonuçlarının Değerlendirilmesi. *Eğitim ve Öğretim Araştırmaları Dergisi*, 3(1), 278-288.

Aksu, F.F. ve Karaçöp, A. (2015). Öğrencilerinin Ev Temelli Fen Öğrenme Etkinliklerine Aile Katılımı. *Erzincan Ü. Eğitim Fakültesi Dergisi*, 17(2), 456-476. doi:10.17556/jef.52040

Anıl, D. (2009). Uluslararası Öğrenci Başarılarını Değerlendirme Programı (PISA)'nda Türkiye'deki Öğrencilerin Fen Bilimleri Başarılarını Etkileyen Faktörler. *Eğitim ve Bilim*, 34(152), 87-100.

Atar, B. (2010). Basit Doğrusal Regresyon Analizi ile Hiyerarşik Doğrusal Modeller Analizinin Karşılaştırılması. *Eğitimde ve Psikolojide Ölçme ve Değerlendirme Dergisi*, 1(2), 78-84.

Bayraktar, Ş. (2010). Uluslararası Fen ve Matematik Çalışması (TIMSS 2007) Sonuçlarına Göre Türkiye'de Fen Eğitiminin Durumu. *Selçuk Ü. Ahmet Keleşoğlu Eğitim Fakültesi Dergisi*, (30), 249-270.

Büyüköztürk, Ş., Çakan, M., Tan, Ş. ve Atar, H. (2014). *TIMSS 2011 Fen ve Matematik Ulusal Raporu*. Ankara: Millî Eğitim Bakanlığı Yenilik ve Eğitim Teknolojileri Genel Müdürlüğü.

Ceylan, E. (2009). PISA 2006 Sonuçlarına Göre Türkiye'de Fen Okuryazarlığında Düşük ve Yüksek Performans Gösteren Okullar Arasındaki Farklar. *Yüzcüncü Yıl Ü. Eğitim Fakültesi Dergisi*, 6(2), 55-75. doi:10.23891

Ceylan, E. ve Berberoğlu, G. (2007). Factors Related with Students' Science Achievement: A Modeling Study. *Eğitim ve Bilim*, 32(144), 36-48.

Çeken, R. (2011). "Bu Benim Eserim" Öğrenci Projelerinin Okul Türü Bakımından Değerlendirilmesi. *Mehmet Akif Ü. Eğitim Fakültesi Dergisi*, (22), 1-14.

Delil, A. ve Tetik, B. (2015). 8. Sınıf Merkezi Sınavlardaki Matematik Sorularının TIMSS-2015 Bilişsel Alanlarına Göre Analizi. *Celal Bayar Ü. Sosyal Bilimler Dergisi*, 13(4), 165-184. doi:10.18026/cbusos.87313



- Deniz, K. (2007). Psikolojik Ölçme Aracı Uyarlama. *Ankara Ü. Eğitim Bilimleri Fakültesi Dergisi*, 40(1), 1-16.
- Duru, S. ve Çöğmen, S. (2017). İlkokul ve Ortaokul Öğrencileri ile Velilerin Ev Ödevlerine Yönelik Görüşleri. *İlköğretim Online Dergisi*, 16(1), 354-365.
- Erkuş, A. (2012). Varolan Ölçek Geliştirme Yöntemleri ve Ölçme Kuramları Psikolojik Ölçek Geliştirmede Ne Kadar İşlevsel: Yeni Bir Öneri. *Eğitimde ve Psikolojide Ölçme ve Değerlendirme Dergisi*, 3(2), 279-290.
- Ersoy, Y. (2018). *TIMSS-2007: Uluslararası Matematik ve Fen Araştırması-II: Başarıyı Etkileyen Örtük Değişkenler ve Genel Eğilimler*. Fen/Fizik Eğitimi Etkinlikleri.
- Harlow, A. & Jones, A. (2004). Why Students Answer TIMSS Science Test Items the Way They Do. *Research in Science Education*, 34(2), 221-238.
- Kan, A. (2008). Psikolojik Değişkenleri Ölçmek İçin Kullanılan Ölçekleme Yaklaşımları Üzerine Bir Karşılaştırma. *Eğitimde Kuram ve Uygulama Dergisi*, 4(1), 2-18.
- Karamustafaoğlu, O. ve Sontay, G. (2012). *Bir TIMSS Sınavının Ardından: TIMSS 2011'e Katılan Öğrenci ve Uygulayıcı Öğretmenlerin Görüşleri*. UFBMEK-10 (s.1-11). Niğde Üniversitesi.
- Kelly, D. (2002). The TIMSS 1995 International Benchmarks of Mathematics and Science Achievement: Profiles of World Class Performance at Fourth and Eighthgrades. *Educational Research and Evaluation*, 8(1), 41-54. doi:10.1076/edre.8.1.41.6918
- Milli Eğitim Bakanlığı [MEB]. (2017). *Fen Bilimleri Dersi Öğretim Programı (3, 4, 5, 6, 7 ve 8. sınıf) Tanımı Öğretim Programı Tanıtım Sunusu*. https://tegm.meb.gov.tr/meb_iys_dosyalar/2017_06/09163104_Fen_Bilimleri_Dersi_Y_Yretim_ProgramY_Karşılaştırmalar.pdf adresinden 11.11.2017 tarihinden edinilmiştir.
- Milli Eğitim Bakanlığı [MEB]. (2018). *Fen Bilimleri Dersi Öğretim Programı (İlkokul ve Ortaokul 3, 4, 5, 6, 7 ve 8. Sınıflar) Öğretim Programı*. Ankara: Devlet Kitapları Basım Evi.
- Olkun, S. ve Aydoğdu, T. (2003). Üçüncü Uluslararası Matematik ve Fen Araştırması (TIMSS) Nedir? Neyi Sorgular? Örnek Geometri Soruları Ve Etkinlikler. *İlköğretim Online*, 2(1), 28-35.
- Öztürk, D. ve Uçar, S. (2010). TIMSS Verileri Kullanılarak Tayvan ve Türkiye'deki 8 Sınıf Öğrencilerinin Fen Başarısına Etki Eden Faktörlerin Belirlenmesi ve Karşılaştırılması. *Çukurova Ü. Sosyal Bilimler Enstitüsü Dergisi*, 19(3), 241-256.



Tavşancıl, E. (2002). *Tutumların Ölçülmesi ve SPSS ile Veri Analizi*. Ankara: Nobel Yayın Dağıtım.

Tekin, H. (1991). *Eğitimde Ölçme ve Değerlendirme*. Ankara: Yargı Yayınları.

Uzun, N., Gelbal, S. ve Öğretmen, T. (2010). TIMSS-R Fen Başarısı Ve Duyuşsal Özellikler Arasındaki İlişkinin Modellenmesi ve Modelin Cinsiyetler Bakımından Karşılaştırılması. *Kastamonu Eğitim Fakültesi Dergisi*, 18(2), 531-544.

Yıldırım, H., Yıldırım, S., Ceylan, E. ve Yetişir, M. (2013). *Türkiye Perspektifinden TIMSS 2011 Sonuçları*. Ankara: Türk Eğitim Derneği Tedmem Analiz Dizisi I.

Comparison of 2016 İzmir and 2011 Turkey Data by TIMSS 2011 Science Test

Introduction

Countries have been participating in international exams since the early 1990's to assess their own education systems, make international comparisons and build long-term data bases (Abazaoğlu etc., 2014). TIMSS, PISA and PIRLS can be examples to international exams in the measurement and assessment studies. International tests show that Turkey is not at the expected level in educational programs. TIMSS (Third International Mathematics and Science Study) is an international study that uses multidimensional and diverse components to present problems in Mathematics and Science (Ersoy, 2016). Item response theory models are used in analyzing the latent skills of students with a Hierarchical Linear Model (HLM) and evaluating the interlocking hierarchical data in international exams (Anıl, 2009; Atar, 2010; Uzun etc., 2010). Nowadays, studies using HLM become widespread and can be combined with Structural Equation Model for deeper evaluations. Reliability coefficient is an important in the investigation of error sources. Because the HLM model is important in international exams, Alpha Chronbach and KR-20/21 reliability coefficients are not considered much in research.

This study aimed to compare TrDt₁₁/TIMSS₁₁FBT₈ (Turkey Data of 8th Science Test of TIMSS in 2011) and İzDt₁₆/TIMSS₁₁FBT₈ (İzmir data of TIMSS₁₁FBT₈ in 2016) at the learning domains (Biology-Physics-Chemistry-Geology) and the cognitive domains (Knowing-Appling-Reasoning). It also aimed to compare the reliability of TrDt₁₁/TIMSS₁₁FBT₈ and the reliability of İzDt₁₆/TIMSS₁₁FBT₈.

Method

This study was a descriptive research model. TIMSS₁₁FBT₈ data were quantitatively and qualitatively examined. In this study, documents and numerical data are composed of documents related to TIMSS₁₁FBT₈. TIMSS



presents its findings after international research on the web (<http://timssandpirls.bc.edu/>).

Participants

In the TIMSS survey in 2011, 7286 students in the 8th grade in the primary school attending 259 schools in 67 cities of Turkey participated. TIMSS₁₁FBT₈'s 80-item multiple-choice question answered by 5441 8th grade student. In this study, 5441 respondents of TIMSS₁₁FBT₈ were accepted as participants. In 2016, secondary schools in İzmir were randomly selected from groups, grouped by cluster sampling. Eighth grades on the school were tested. The data collection process was carried out İZDt₁₆/TIMSS₁₁FBT₈. In 2016, 775 elementary school 8th grade students from 20 randomly selected schools in İzmir were included in the sample of the study.

Data collection tools and analysis

TIMSS₁₁FBT₈ is a 172-items test consisting of 4-choice multiple-choice questions for the concepts of biology, chemistry, physics, earth sciences and open-ended questions that measure conceptual meaning. TIMSS₁₁FBT₈ measure both cognitive and learning domains. İZDt₁₆/TIMSS₁₁FBT₈ were gathered by 80-items test consisting of 4-multiple-choice questions of TIMSS₁₁FBT₈. Multiple-choice questions are four choices and there is a correct answer. The correct answer for every question in the multiple-choice test is 1 point. Wrong answers do not affect correct answers.

Findings

In 2011, TIMSS₁₁FBT₈ was divided into 6 parts and applied in Turkey. For this reason, the reliability of the test was obtained by averaging 6 parts. Therefore, the reliability of TrDt₁₁/TIMSS₁₁FBT₈ was found to be 0.70. 80-items of TIMSS₁₁FBT₈ were reapplied in İzmir in 2016. Therefore, the reliability of İZDt₁₆/TIMSS₁₁FBT₈ was found to be 0.93.

The biology score of İZDt₁₆/TIMSS₁₁FBT₈ were found to be lower than TrDt₁₁/TIMSS₁₁FBT₈. Likewise, a similar finding can be said within the geology score of İZDt₁₆/TIMSS₁₁FBT₈. However, the physics score of İZDt₁₆/TIMSS₁₁FBT₈ were found to be slightly higher than those TrDt₁₁/TIMSS₁₁FBT₈. Likewise, a similar finding can be said within the chemistry score of İZDt₁₆/TIMSS₁₁FBT₈.

The knowing score of İZDt₁₆/TIMSS₁₁FBT₈ were close to each other regarding TrDt₁₁/TIMSS₁₁FBT₈. The applying and reasoning scores of İZDt₁₆/TIMSS₁₁FBT₈ were found to be lower than TrDt₁₁/TIMSS₁₁FBT₈.

Results and Discussion

It can be said that the collected data are reliable because the reliability coefficients of İZDt₁₆/TIMSS₁₁FBT₈ are more than TrDt₁₁/TIMSS₁₁FBT₈. The higher reliability coefficient of İZDt₁₆/TIMSS₁₁FBT₈ may occur as a result of



applying as a single dimension structure of the test in the same sample group. The value of the confidence coefficient is directly related to the number of items in the test. Therefore, as the number of items increases, the value of the reliability coefficient also increases. TIMSS₁₁FBT₈ was applied in gathering TrDt₁₁/TIMSS₁₁FBT₈ with different samples in six dimensions and the reliability coefficient of the test was calculated with less substance content. For this reason, it is natural that the reliability coefficient of the test is lower. Because the 80-item test was obtained from six dimensions and six samples. Since there is an error value in the reliability coefficient calculation for each dimension, the error rate is mixed 6 times for all of the test.

The biology and geology scores of İzDt₁₆/TIMSS₁₁FBT₈ are lower than TrDt₁₁/TIMSS₁₁FBT₈. However, the chemistry and physics scores of İzDt₁₆/TIMSS₁₁FBT₈ are more advanced than TrDt₁₁/TIMSS₁₁FBT₈. In interviews with teachers and students; It has been understood that the students and teachers participating in the TIMSS exam have been preparing for the exam before the original examination. In addition, according to the opinions of the students, it was seen that the students were inadequate to answer in the sense of asking in the TIMSS exams (Karamustafaoğlu etc.; 2012). Student sample in İzDt₁₆/TIMSS₁₁FBT₈ did not receive any training on the TIMSS exam in advance. Despite the lack of training, scores close to 2011 were obtained. It can be said that when the students were educated before the TIMSS exams, İzmir province had a slight increase in this test.

According to the results of İzDt₁₆/TIMSS₁₁FBT₈, the applying and reasoning scores are lower than TrDt₁₁/TIMSS₁₁FBT₈. However, the knowing score of İzDt₁₆/TIMSS₁₁FBT₈ data is similar to TrDt₁₁/TIMSS₁₁FBT₈. The reasoning score of TrDt₁₁/TIMSS₁₁FBT₈ is lower than İzDt₁₆/TIMSS₁₁FBT₈. Mental activities in knowing domain are more basic and easier to operate than applying and reasoning domains (Yıldırım etc., 2013). According to these findings, it is seen that the students in both Turkey and İzmir are more successful at the level of knowing, and that this level is decreased at the level of reasoning and applying. It can be said that the open-ended examination system in the direction of these findings is very effective in measuring the level of students' applying and reasoning.

