

YAPAY ZEKÂ DESTEKLİ DENEY TASARIM SÜRECİNDE FEN BİLİMLERİ ÖĞRETMEN ADAYLARININ TASARIMLARINA YÖNELİK CHATGPT DEĞERLENDİRMELERİNİN İNCELENMESİ

AN EXAMINATION OF CHATGPT'S EVALUATIONS OF PRE-SERVICE SCIENCE TEACHERS' DESIGNS IN THE AI-SUPPORTED EXPERIMENT DESIGN PROCESS

Suat TÜRKOĞUZ

Prof. Dr., Dokuz Eylül Üniversitesi, Buca Eğitim Fakültesi, İzmir, Türkiye

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

suat.turkoguz@gmail.com

Dilara Melis EREN

Araş. Gör., Dokuz Eylül Üniversitesi, Buca Eğitim Fakültesi, İzmir, Türkiye

ORCID ID: <https://orcid.org/0009-0009-2071-379X>

dilarameliseren@gmail.com

Received: November 8, 2025

Accepted: January 24, 2026

Published: January 31, 2026

Suggested Citation:

Türkoğuz, S., & Eren, D. M. (2026). Yapay zekâ destekli deney tasarım sürecinde fen bilimleri öğretmen adaylarının tasarımlarına yönelik ChatGPT değerlendirmelerinin incelenmesi. *International Journal of New Trends in Arts, Sports & Science Education (IJTASE)*, 15(1), 9-31.



Copyright © 2026 by author(s). This is an open access article under the [CC BY 4.0 license](https://creativecommons.org/licenses/by/4.0/).

Öz

Bu çalışmanın amacı yapay zekâ destekli deney tasarım sürecinde fen bilimleri öğretmen adaylarının tasarımlarına yönelik ChatGPT değerlendirmelerinin incelenmesidir. Nitel araştırma yöntemlerinden durum çalışması deseniyle yürütülen çalışmanın örneklemini 2024 – 2025 eğitim-öğretim yılında İzmir’de bir devlet üniversitesinin Eğitim Fakültesinde öğrenim gören 32 fen bilimleri öğretmen adayı oluşturmaktadır. Çalışmanın verileri yaklaşık bir saat süren bireysel uygulama oturumları sırasında gönüllülük esasına dayalı olarak toplanmıştır. Veri toplama sürecinde araştırmacılar tarafından geliştirilen Microsoft Excel makroları ile yapılandırılmış dijital formlar kullanılmıştır. Öğretmen adaylarının özgün deney tasarımları, ChatGPT dönütleri doğrultusunda gerçekleştirilen düzenlemeler ve katılımcıların ChatGPT ile olan etkileşimlerine ait metinlerden elde edilen veriler MAXQDA 24 yazılımı kullanılarak içerik analiziyle çözümlenmiş, kodlama süreci uzman görüşü doğrultusunda gözden geçirilerek araştırmacılar arası uyum sağlanmıştır. İki kodlayıcı tarafından yapılan kodlamalar arasındaki puanlayıcılar arası güvenilirlik katsayısı 0.91 olarak belirlenmiştir. Çalışmanın bulguları öğretmen adaylarının ChatGPT’yi deney tasarım sürecinde danışma, doğrulama ve geliştirme amacıyla kullandıklarını göstermekte ve öğretmen adaylarının deney tasarımının değişken belirleme, uygulama basamaklarını belirleme ve STEM entegrasyonu sağlama aşamalarında ChatGPT değerlendirmelerinden yararlanmakta olduklarını ortaya koymaktadır.

Anahtar Terimler: ChatGPT, deney tasarımı, fen bilimleri öğretmen adayları, nitel araştırma, yapay zekâ araçları.

Abstract

The aim of this study is to examine ChatGPT’s evaluations of pre-service science teachers’ designs within an AI-supported experiment design process. The study was conducted using a case study design, one of the qualitative research methods. The sample consisted of 32 pre-service science teachers enrolled in the Faculty of Education at a public university in İzmir during the 2024 – 2025 academic year. Data were collected on a voluntary basis during individual sessions lasting approximately one hour. In the data collection process, digitally structured forms developed by researchers using Microsoft Excel macros were utilized. Data obtained from the participants’ original experiment designs, the revisions made in line with ChatGPT feedback, and the textual records of their interactions with ChatGPT were analyzed through content analysis using MAXQDA 24 software. The coding process was reviewed based on expert feedback, and inter-researcher agreement was established. The inter-rater reliability coefficient between the two coders was determined to be 0.91. The findings of the study reveal that pre-service science teachers used ChatGPT for consultation, validation, and refinement during the experiment design process, and that they benefited from ChatGPT’s evaluations particularly in the stages of identifying variables, determining experimental procedures, and ensuring STEM integration in experiment design.

Keywords: ChatGPT, experiment design, pre-service science teachers, qualitative research, artificial intelligence tools.

GİRİŞ

Geçmişte gerçekleşen çeşitli sosyal olaylar ve teknolojik gelişmeler eğitimde dijitalleşmeyi ve teknoloji kullanımını önemi gün geçtikçe artan bir durum haline getirmiştir. Bununla birlikte sanal öğrenme ortamları ve içerikleri, teknolojik araçların eğitim öğretim ortamlarında kullanımı, günlük hayatta kullanılan çeşitli uygulamaların eğitime entegrasyonu gibi bu gelişmelerle doğrudan bağlantılı uygulamaların öğrenme ortamlarında nasıl kullanılabileceği toplumun her kesimini ilgilendiren önemli bir merak ögesi haline gelmiştir. Bu teknolojik gelişmeler kapsamında yapay zekânın ve ChatGPT gibi büyük dil modellerinin günlük hayatta önemli yer tutan yapılar arasında yer almakta olduğu görülmektedir. Yapay zekâ; öğrenme, uyum sağlama, sentezleme, kendini geliştirme ve çok yönlü işlem görevleri için veri kullanımı gibi insan zekâsının yeteneklerine benzer görülen çeşitli eylemleri gerçekleştirecek ve benimseyecek biçimde geliştirilmekte olan makineleri içeren bir kavramdır (Jia, Sun, & Lool, 2024; Salas-Pilco, Xiao & Hu, 2022; Popenici & Kerr, 2017; Kok, Boers, Kusters, Van der Putten, & Poel, 2009; Dobrev, 2005; Simmons & Chappell, 1988). OpenAI tarafından geliştirilmiş, doğal dil yanıtlarını işlemek ve üretmek için derin öğrenme teknikleri doğrultusunda kullanıcı girdisine dayalı olarak insan benzeri metinler üreten bir büyük dil modeli olan ChatGPT, özellikle GPT-4 üzerine inşa edilmiş olan GPT (Üretken Önceden Eğitilmiş Dönüştürücü - Generative Pre-trained Transformer) mimarisinin katkısıyla insan anlama yeteneğini simüle etmekte ve gerçek akıl yürütme yeteneği bulunmamasına rağmen çeşitli konular hakkında bağlamsal olarak uygun ve faydalı olarak değerlendirilebilecek yanıtlar sunmaktadır (Zeng, Gan, Wang, & Yu, 2025; de Souza, Serrano, & Roazzi, 2024; Fijačko vd., 2024; Playfoot, Quigley, & Thomas, 2024; Suleiman vd., 2024; Javaid, Haleem, Singh, Khan, & Khan, 2023; Song & Wang, 2023).

Bilgi okuryazarlığı ve içerik oluşturma gibi dijital yetkinliklerin geliştirilmesinde rol oynayan, teknolojinin gelişmesi ile birlikte birçok farklı araçla hayatımıza giren yapay zekâ öğrencilerin dijital dünyada yer alabilmeleri için gerekli dijital becerilerini ve problem çözme yeteneklerini bu süreçte dahil ederek eğitim öğretim sürecinin önemli bir parçası haline gelmektedir (Galindo-Domínguez, Delgado, Campo, & Losada, 2024; Li, Lowell, Wang, & Li, 2024). Literatürde yer alan çalışmalarda yapay zekânın öğretme ve öğrenme süreçlerine entegre edilerek ders planı tasarımı ya da deney tasarımı gibi çeşitli materyaller tasarlanmasını, oyunlaştırmayı, etkileşimli ve ilgi çekici öğrenme deneyimlerinin yaratılmasını desteklemesi ile öğrencilerin motivasyonunu ve tasarım süreçlerine katılımını artırdığı görülmektedir (Zhang & Huang, 2024). Akıllı öğretim platformları gibi çeşitli yapay zekâ destekli sistemler eğitim içeriğini ve geri bildirimi bireysel öğrenci ihtiyaçlarına göre uyarlayıp katılımı ve öğrenme sonuçlarını iyileştirirken aynı zamanda öğretimin hızını, zorluk derecesini ve stilini ayarlama imkanı sunmaktadır (Kovari, 2025; Madanchian & Taherdoost, 2025; An, Yang, Xu, Zhang, & Zhang, 2024; Li & Fu, 2024). Ancak yapay zekâdan eğitim öğretim sürecinde alınacak destek ve bu desteğin boyutu ile veriminin büyük oranda kullanıcılara bağlı olduğu görülmektedir (Jayawardena, Gunathilake, & Ihalagedara, 2025; An vd., 2024; Galindo-Domínguez vd., 2024).

Literatürde yer alan öğrenme öğretme süreçlerinde ve özellikle öğretmen eğitiminde ChatGPT kullanımı ile ilgili çalışmalar incelendiğinde kullanıcıların danışma amacıyla ChatGPT'den yararlandıkları görülmektedir. Cai, Han, Sun, Li, ve Wong (2025) tarafından yürütülen çalışmada ChatGPT destekli ders planı tasarımlarının öğretmen adaylarının ders planlama becerileri üzerindeki etkisini incelemek amaçlanmıştır. Yarı deneysel desen biçiminde yürütülen çalışmanın bulguları ChatGPT değerlendirmelerinin katılımcıların ders planlama becerilerinin bilişsel ve duyuşsal döngülerinde anlamlı fark görüldüğünü ancak üstbilişsel döngüde anlamlı bir fark görülmediğini göstermekte ayrıca ChatGPT değerlendirmelerinin öğretmen adaylarının ders planı tasarımlarında öğretim hedefleri, kullanılan pedagojik yaklaşımlar, öğrenme süreci etkinlikleri ve materyalleri arasında daha bütüncül ilişki kurmalarına katkı sağladığı görülmektedir (Cai, Han, Sun, Li, & Wong, 2025). Chellappa ve Luximon (2024) tarafından ürün tasarımı ve kullanıcı deneyimi tasarımı alanı öğrencilerinin ChatGPT'ye yönelik algılarını ve algıların öğretim programı, cinsiyet ve akademik düzeye göre nasıl değiştiğini incelemeyi amaçlayan nicel çalışmasının bulgularına göre katılımcılar ChatGPT'yi kullanımı kolay ve ilgi çekici bir araç olarak görmektedir. Ayrıca çalışmanın katılımcıları ChatGPT'nin destekleyici bir öğrenme kaynağı olarak bağlama uygun, iyi yapılandırılmış, detaylı

yanıtlar sunduğunu düşünmekte ancak ChatGPT yanıtlarının doğruluk düzeyine ilişkin algının orta doğruluk düzeyinde olduğunu belirtmektedir (Chellappa & Luximon, 2024). Zhu, Li, Yao, Li, & Zhu (2025) tarafından yürütülen çalışmada ise ChatGPT geri bildiriminin katılımcıların geri bildirimi anlama, kullanma ve öğrenmeye dönüştürme becerileri üzerindeki etkilerini inceleme amaçlanmıştır. Karma çalışma olarak yürütülen araştırmanın bulguları katılımcılara göre ChatGPT geri bildiriminin açık, yapılandırılmış ve bağlama uygun görülürken bireyselleşme ve derinlik açısından sınırlı bulunduğunu ve katılımcılar tarafından ChatGPT geri bildiriminin yüzeysel biçimde kullanıldığını göstermektedir (Zhu, Li, Yao, Li, & Zhu, 2025).

Fen eğitimi bağlamında deney tasarımının temeli bilimsel ve günlük hayat gözlemleri sonucunda oluşan merak duygusuna dayanmaktadır. Bu merak olgusu göz önünde bulundurularak zihinde onunla ilgili oluşan gözlem sonuçlarını ve tahminleri çeşitli bilimsel bilgiye ulaşma adımlarını takip ederek test etmek için laboratuvar ortamında veya gerçek hayat şartlarında çeşitli sına uygulamaları yapılabilir. Bir fen kavramını sınamayı ya da keşfetmeyi sağlayabilen deneyler kısaca veri toplama ortamlarının, objelerinin ve takip edilecek olan uygulama prosedürünün sistematik olarak detaylıca planlanması ve yapılandırılması olarak açıklanabilir. Bu sına veya keşfetme sürecinde gözlemlenmesi planlanan parametrelerle ilgili değişikliği tanımlayan verileri elde etmeyi amaçlayan deney tasarımı prosedürü dikkatli bir planlama sürecini gerektirmektedir. Literatürde yer alan çalışmalar incelendiğinde deney tasarımı sürecinin içerdiği ana basamakların oluşturulan hipotezleri değerlendirmek için gerekli verileri üretmek amacıyla uygulanacak yöntemlerin belirlenmesi, incelenen değişkenlerin, kontrollerin, yapılacak denemelerin içeriğinin sayısının ve adımlarının tanımlanması, sonuçların değerlendirilmesi için kullanılacak ölçüt ya da araçların tasarlanması olarak sıralanabildiği görülmektedir (Millar, 2004; Hudson, 1988; Hofstein & Lunetta, 1982).

Fen bilimleri alanlarında tasarlanan deneylerin ölçmeyi amaçladıkları şeyi ölçen, tutarlı sonuçlar üreten, tekrarlandığında benzer sonuçlar ortaya çıkan yani geçerli, güvenilir ve tekrarlanabilir olması deney sonucunda elde edilen bulguların güvenilirliği için son derece önemlidir. Deney tasarımı sürecinde gözlemlenen etkilerin deneysel müdahaleden kaynaklanması, deney sonucu elde edilen bulguların belirli bir popülasyona genellenebilmesi ya da istatistiksel olarak anlamlı sonuçların elde edilmesi gibi çeşitli iç geçerlik, dış geçerlik, istatistiksel sonuç geçerliği ve yapısal geçerlik olarak sınıflandırılabilir geçerlik ile ilgili güçlüklerle karşılaşılabilir (Aguinis & Lawal, 2012). Bu güçlükler ek olarak literatürde yer alan çalışmalar incelendiğinde deney tasarımı sırasında katılımcıların problemin tanımlanması, deney tasarlanması amaçlanan sınıf düzeyine uygun prosedürel adımların, malzemelerin ve teorik arka planın oluşturulması sırasında kayda değer biçimde zorlandıkları bunlarla birlikte çoklu değişkenlerin etkileşiminin yönetilmesi gibi çeşitli yalnızca deneyin gerçekleştirildiği sırada meydana gelebilecek zorlukların da deney tasarımı sürecinde göz önünde bulundurulması gerektiği ve deney tasarımı sürecinde karşılaşılabilecek durumlara göre değişiklik ya da güncelleme yapmaya hazırlıklı olunması gerektiği vurgulanmaktadır. Literatürde yapay zekâ destekli eğitim teknolojisinin öğrenme süreçlerini kolaylaştıran ve tasarım etkinliklerini destekleyen işbirlikçi bir araç olarak kavramsallaştırılmakta olduğu görülmektedir. Yapay zekâ araçları deney tasarımı öğrenme hedefleriyle uyumlu ve bilinçli bir şekilde kullanılmalı ve deney sürecine yapay zekâ araçlarının entegrasyonu, bilinçli yapay zekâ kullanımı yoluyla öğrenme deneyimlerini desteklemekle ilgili olmalıdır (Huang, Chen, Lee, Sandnes, & Wu, 2025; Li & Fu, 2024).

Kullanıcıların herhangi bir kavram ile ilgili akıllarına takılan şeyleri danışıp tartışabileceği, kavramları görselleştirebileceği, deney tasarımı gibi öğretim sürecinde kullanılabilecek materyalleri oluşturabileceği, değişkenleri güvenli bir ortamda kontrollü olarak manipüle edebileceği etkileşimli öğrenme ortamları ve deneyimleri sağlayan ChatGPT gibi üretken yapay zekâ araçları öğretmen eğitiminde teknolojinin geleneksel öğrenme için yardımcı araç olarak kullanılmasına ek olarak sürekli gelişim ve yaşam boyu öğrenme için destek sağlayan bir araç olarak öğrenme sürecine ve günlük hayata entegre edilip öğretmen adaylarının yansıtıcı düşünme becerilerini desteklerken aynı zamanda problem çözme gibi üst düzey bilişsel yetenekleri destekleyebilir (Liu, Sun, Sun, Wang ve Yu, 2025; Wei, Wang, Koszalka, Lee, & Liu, 2025; Fischer, Sommerhoff, & Keune, 2023; Hovardas, Ter Vrugte, Zacharia, & de Jong, 2023; Hämäläinen & Oksanen, 2012; Martin-Villalba, Urquía &

Dormido, 2012; Alvarez, Alarcon, & Nussbaum, 2011). ChatGPT gibi üretken yapay zekâ modelleri kullanıcıya sunduğu anlık geri bildirimlerle öğrenme süreçlerine yeni bir yaklaşımın entegre edilmesi fırsatını sunmaktadır. Literatürde yer alan çalışmalar incelendiğinde ChatGPT'nin öğretmen adaylarının özgün tasarımlarına yönelik değerlendirmelerini merkeze alan ve bu değerlendirmelerin tasarımın hangi yönü ile ilişkili olduğunu inceleyen sınırlı sayıda çalışmanın yer aldığı ve bu durumun ChatGPT değerlendirmelerini deney tasarımı süreci bağlamında incelenmesi ihtiyacını ortaya çıkardığı görülmektedir. Bu doğrultuda mevcut çalışma fen bilimleri öğretmen adaylarının deney tasarımına yönelik ChatGPT değerlendirmelerini incelemeyi hedeflemektedir.

Problem Cümlesi

Bu çalışmanın amacı yapay zekâ destekli bir sohbet robotu olan ChatGPT desteği ile çeşitli fen bilimleri alt disiplinleriyle ilgili deney tasarımı yapan fen bilimleri öğretmen adaylarının deney tasarımı sürecinde ChatGPT ile olan etkileşimlerinin incelenmesi ve öğretim çıktısı içeriği ile sınıf seviyesi dikkate alınarak bu deney tasarımlarıyla ilgili ChatGPT görüşünün incelenmesidir. Bu doğrultuda çalışmanın alt problemleri şu şekilde sıralanabilir:

1. Fen bilimleri öğretmen adayları özgün deney tasarlama sürecinde hangi aşamalar hakkında ChatGPT'ye soru sormakta ve hangi aşamalarda deney tasarımlarını geliştirme desteği almaktadır?
2. ChatGPT'nin fen bilimleri öğretmen adaylarının uygulama sırasında oluşturmuş oldukları özgün deney tasarımları hakkındaki görüşleri ve geliştirme önerileri nelerdir?

YÖNTEM

Araştırmanın Modeli

Bu çalışmada nitel araştırma yöntemlerinden durum çalışması kullanılmıştır. Durum çalışmaları sınırlı sayıdaki olayın veya koşulun ve bunların ilişkilerinin ayrıntılı bağlamsal analizini vurgulayan, derinlemesine araştırmaya olanak sağlayan nitel bir araştırma yöntemidir (Savin-Baden & Major, 2023; Frankael & Wallen, 2012). Durum çalışmaları esnek doğası gereği geniş yelpazedeki yaklaşımlardan beslenir ve bu tip çalışmaları yürüten araştırmacıların hem belirli bir olguya odaklanmasını hem de farklı bakış açılarını inceleyerek literatürde bütüncül yaklaşım olarak belirtilen yaklaşımın gerçekleştirilmesini sağlar (Savin-Baden & Major, 2023).

Çalışma Grubu / Katılımcılar

Mevcut çalışma katılımcıların uygulama sırasında tesadüfi bir öğretim çıktısına yönelik özgün bir deney tasarımı yapılmasını gerektirdiğinden çalışmanın örnekleme amaçlı örnekleme yöntemi ile seçilmiştir. Amaçlı örnekleme, araştırma katılımcılarının araştırma problemine uygun olarak belirli bir amaç doğrultusunda belirlenmesi olarak tanımlanabilir. Araştırmacıların daha önceki bilgilere dayanarak çalışma kapsamında elde edilmesi hedeflenen verileri sağlayacağını düşündükleri örnekleme seçmek için kendi yargılarını kullandıkları amaçlı örnekleme yöntemi ile belirlenmiş olan mevcut çalışmanın katılımcılarını İzmir'de bir devlet üniversitesinin Eğitim Fakültesinde 2024-2025 eğitim öğretim yılında Fen Öğretimi Laboratuvar Uygulamaları 2 dersini almış öğretmen adayları oluşturmaktadır (Frankael & Wallen, 2012).

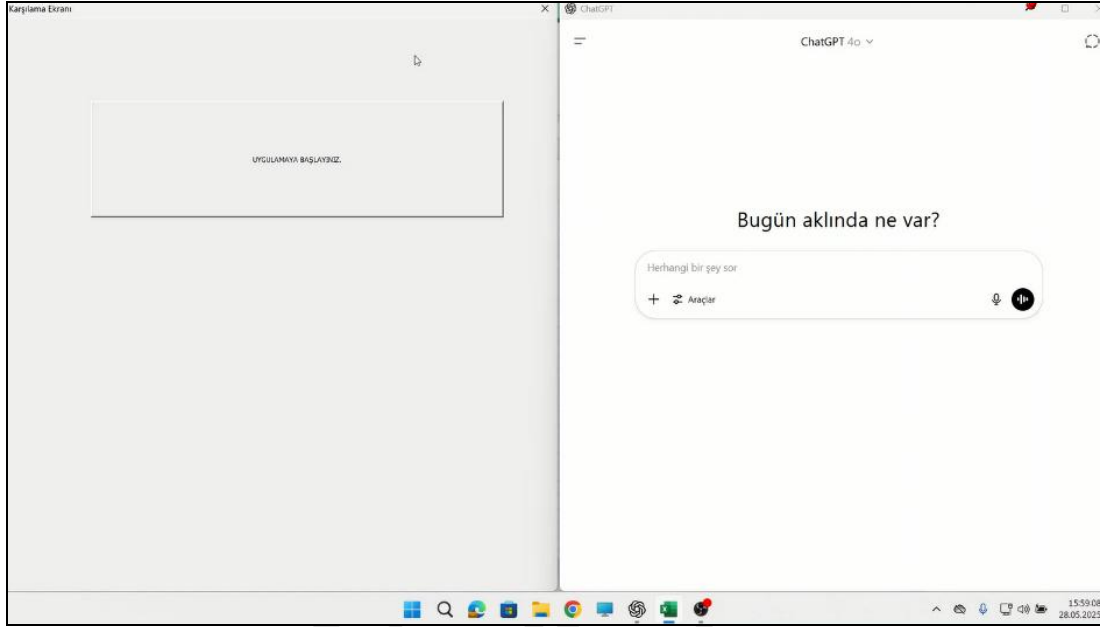
Veri Toplama Araçları ve Veri Toplama Süreci

Bu çalışmada veri toplama aracı olarak araştırmacılar tarafından geliştirilen Yapay Zekâ Araçları ile Deney Tasarlama Formu, S-FeTeMM Formu ve katılımcıların ChatGPT ile olan konuşmalarına ait metinler kullanılmıştır.

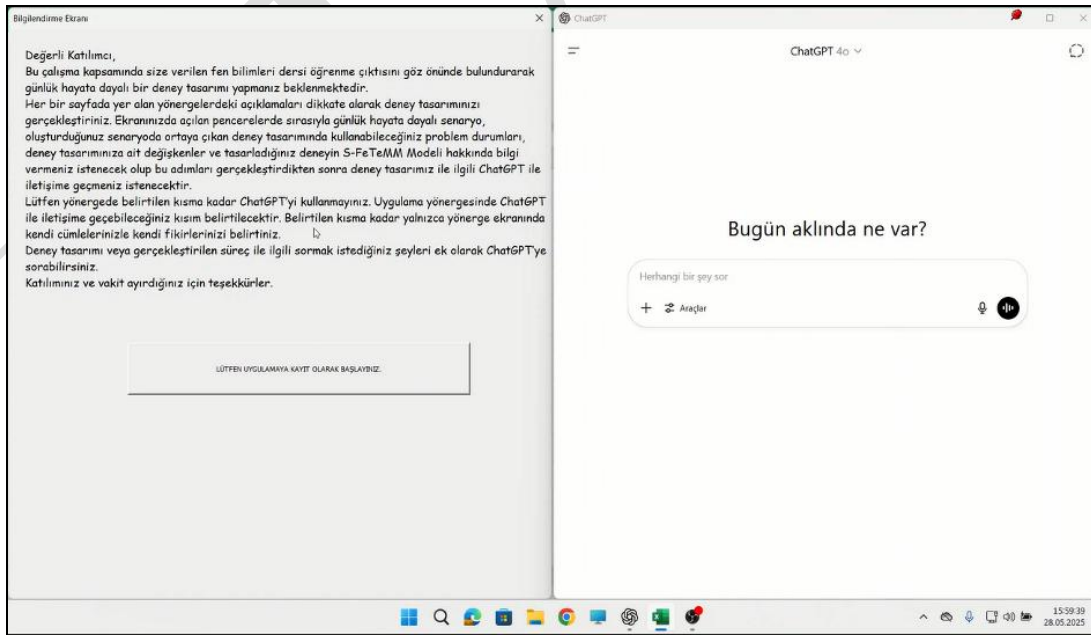
Yapay Zekâ Araçları ile Deney Tasarlama Formu ve S-FeTeMM Formu öğretmen adaylarının özgün deney tasarımlarını, özgün deney tasarımlarının STEM entegrasyonunu ve katılımcıların özgün deney tasarımları üzerinde ChatGPT dönütlerine göre yaptıkları değişiklikleri incelemek amacıyla Microsoft Excel üzerinde geliştirilen makrolar aracılığıyla tasarlanmış dijital formlardır. Katılımcıların derinlemesine görüşlerini ifade edebilmelerini sağlamak amacıyla açık uçlu olarak geliştirilen bu

formlar öğretmen adaylarının yanıtlarını doğrudan hücrelere kaydetmekte ve makrolar aracılığıyla verileri otomatik olarak düzenlemekte, kontrol etmekte ve analiz için uygun formatta saklamaktadır.

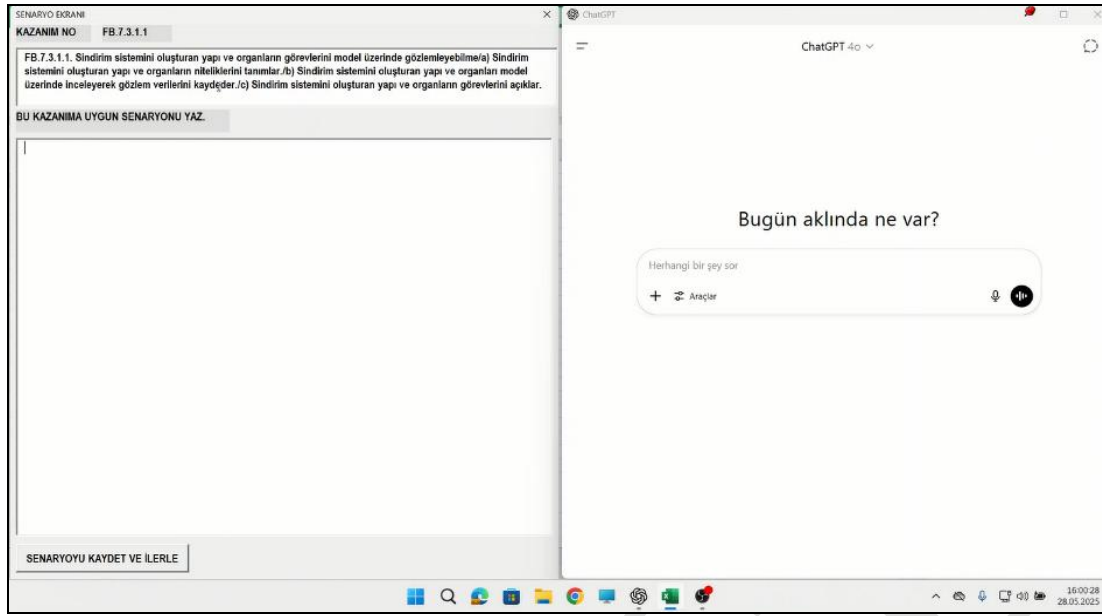
Yapay Zekâ Araçları ile Deney Tasarlama Formu sırasıyla; günlük hayat senaryosu tasarlama, problem durumu ya da durumlarını belirleme, deney tasarımının bağımlı değişkenlerini belirleme, deney tasarımının bağımsız değişkenlerini belirleme ve deney tasarımının kontrol edilen değişkenlerini belirleme basamaklarından oluşmaktadır. S-FeTeMM Formu ise sırasıyla; uygulama basamakları, deney tasarımının fen bilimleri entegrasyonu, teknoloji entegrasyonu, mühendislik entegrasyonu ve matematik entegrasyonu adımlarından oluşmaktadır.



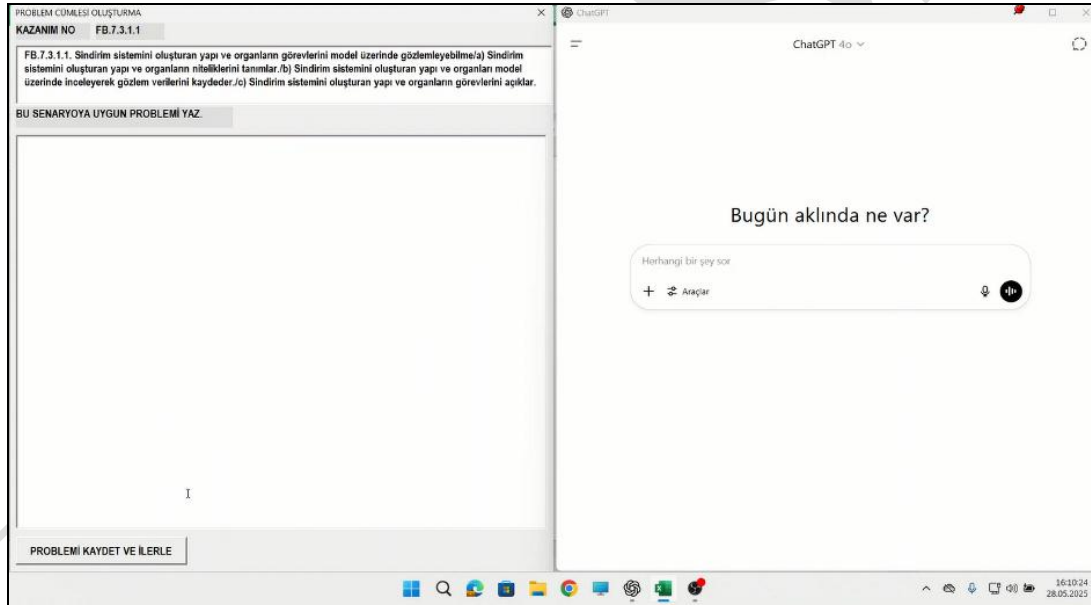
Görsel 1. Veri toplama aracıda yer alan giriş ekranına ait ekran görüntüsü.



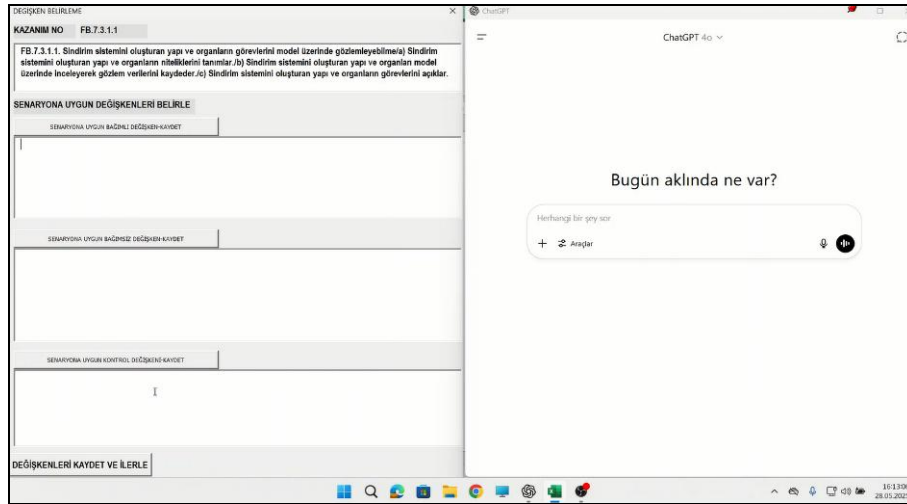
Görsel 2. Veri toplama aracıda yer alan uygulama yönergesine ait ekran görüntüsü.



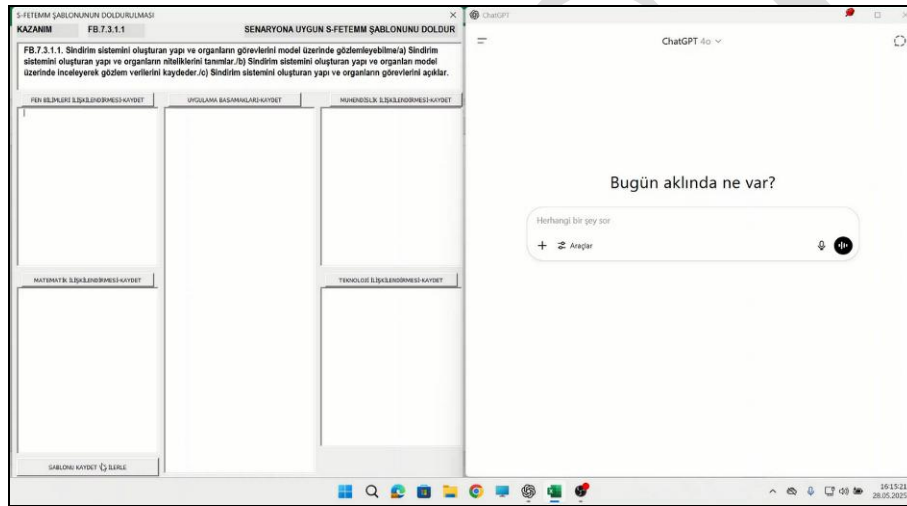
Görsel 3. Veri toplama aracında yer alan günlük hayat senaryosu oluşturma aşamasına ait ekran görüntüsü.



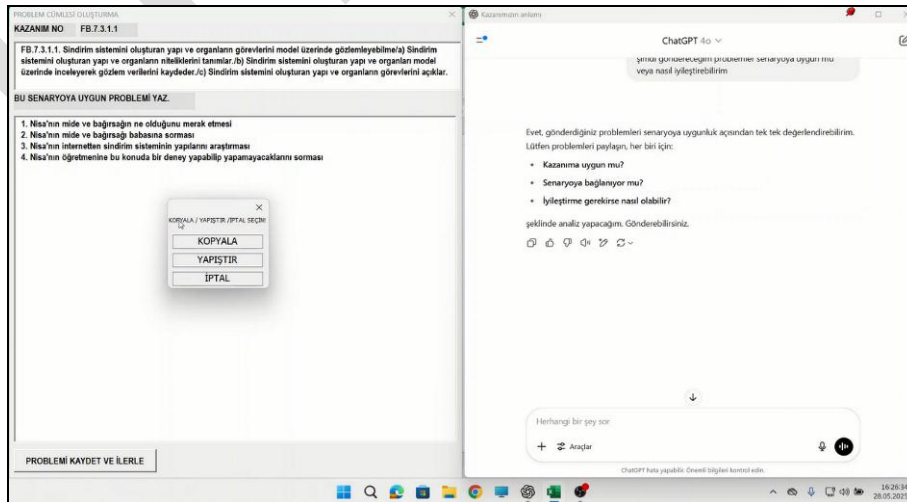
Görsel 4. Veri toplama aracında yer alan problem cümlesi oluşturma aşamasına ait ekran görüntüsü.



Görsel 5. Veri toplama aracında yer alan deney değişkenlerini oluşturma aşamasına ait ekran görüntüsü.



Görsel 6. Veri toplama aracında yer alan deney tasarımının uygulama basamakları ve STEM entegrasyonunu sağlama aşamasına ait ekran görüntüsü.



Görsel 7. Yapay zekâ destekli deney tasarım sürecine ait ekran görüntüsü.

Araştırmanın verileri nitel olarak 2024-2025 eğitim öğretim yılı bahar dönemi boyunca her bir katılımcı ile bireysel gerçekleştirilen yaklaşık bir saat süren oturumlar sırasında bilgisayar ortamında çevrimiçi olarak yazılı bir şekilde toplanmıştır. Uygulama oturumları sürecinde öğretmen adayları öncelikle 2024 Fen Bilimleri Dersi Öğretim Programı (MEB, 2024)’nda yer almakta olan ve bilgisayar ortamında tesadüfi olarak katılımcıya atanmış olan bir öğretim çıktısını göz önünde bulundurarak özgün deney tasarımlarını gerçekleştirmişlerdir. Uygulama oturumlarının ikinci aşamasında ise katılımcılar deney tasarımlarının öğretim çıktısı ile olan uygunluğunu ChatGPT’ye çeşitli sorular sorarak değerlendirmiş ve ek olarak ChatGPT’ye deney tasarımı ile ilgili kendi belirlediği soruları sorup tasarımında gerekli gördüğü değişiklikleri yapmıştır. Katılımcıların uygulama sürecinde etkileşimde bulunduğu ChatGPT versiyonu çalışmanın veri toplama sürecinin gerçekleştiği dönemde güncel ChatGPT versiyonu olan ChatGPT 4o versiyonu plus üyelik ile kullanılmıştır.

Verilerin Analizi

Nitel araştırmalar derinlemesine çalışma ile değişkenlerin incelenmesini içerir ve nitel analizler sosyal durumların, olayların ya da fenomenlerin nasıl meydana geldiğini anlamamıza yardımcı olan teknikler olarak tanımlanabilir (Savin-Baden & Major, 2023; Fraenkel & Wallen, 2012). Bu doğrultuda mevcut çalışmanın veri analizi kapsamında yürütülen içerik analizi süreci katılımcı sorularına verilen ChatGPT yanıtlarının kodlanması ile başlamıştır. Çalışma kapsamında veri toplama araçları ile elde edilen veriler bilgisayar destekli nitel veri analizi yazılımı olan MAXQDA 24 üzerinden içerik analizi kullanılarak çözümlenmiştir. Fen bilimleri eğitimi alan uzmanı iki araştırmacı tarafından bağımsız olarak MAXQDA 24 üzerinden gerçekleştirilen içerik analizi süreci sonrası araştırmacılar arası uyum incelenmiştir ve uzmanı görüşü doğrultusunda içerik analizine son hali verilmiştir. İki kodlayıcı tarafından yapılan kodlamalar arasındaki puanlayıcılar arası güvenilirlik katsayısı 0.91 olarak belirlenmiştir. Nitel veri analizi sürecinde verilerin kodlanması araştırmacının verileri anlamlı bölgelere ayırması ile başlamıştır. Ardından her bir bölgenin kavramsal olarak ifade ettiği anlamın elde edilen diğer bölgelere ve kavramlara göre ifade edilmeye çalışıldığı bu süreçte verilerin kodlanması tümevarım yöntemi ile ortaya çıkan kodların sınıflandırılması ve elde edilen verilerin araştırmacı tarafından uygun biçimde derlenip ifade edilmesi ile gerçekleşmiştir. (Savin-Baden ve Major, 2023).

BULGULAR

Öğretmen adaylarının deney tasarımlarına dair ChatGPT değerlendirmelerine yönelik içerik analizi ile elde edilen bulgular tablolar halinde sunulmuştur. İlk olarak öğretmen adaylarının bireysel olarak gerçekleştirdikleri deney tasarımı sürecinin ardından ChatGPT’ye özgün tasarımlarına dair yönelmiş oldukları “Mevcut deney tasarımının günlük hayat senaryosu öğretim çıktısına uygun mudur, geliştirilmesi gereken herhangi bir şey var mıdır?” sorusuna yönelik ChatGPT değerlendirmesi Tablo 1’de sunulmuştur.

Tablo 1. Günlük hayat senaryosu tasarlama adımı ChatGPT’nin katılımcıların deney tasarımına yönelik değerlendirmelerinin dağılımı

ChatGPT Değerlendirmesi	f	%
Yetersiz	2	6,25
Kısmen yeterli, öneri var	24	75
Yeterli	8	25

Tablo 1’de ChatGPT’nin öğretmen adaylarının özgün deney tasarımlarında yer alan günlük hayat senaryolarını kısmen yeterli bulduğu ancak çeşitli geliştirme önerileri sunduğu görülmektedir.

Katılımcı tasarımını yetersiz olarak değerlendiren ChatGPT görüşlerinden biri “Paylaştığınız senaryo başlangıç aşaması olarak kazanıma potansiyel olarak uygun olsa da şu anki haliyle eksik görünüyor

çünkü hedef kazanımın tüm davranış düzeylerini kapsamıyor.” şeklindedir (Öğretmen Adayı 28 tasarımına ait ChatGPT görüşü).

Katılımcı tasarımını kısmen yeterli olarak değerlendiren ChatGPT görüşlerinden biri “Senaryonun ana fikri güzel ancak güvenlik, ölçüm aracı ve birim açısından eksikleri var. Yukarıdaki gibi yeniden yazarak hem pedagojik uygunluk hem de kazanım tamlığı sağlanabilir.” şeklindedir (Öğretmen Adayı 7 tasarımına ait ChatGPT görüşü).

Katılımcı tasarımını yeterli olarak değerlendiren ChatGPT görüşlerinden biri “Bu deney senaryosu hem içerik hem uygulama hem de pedagojik yaklaşım açısından FB.7.3.1.1 kazanımıyla tamamen uyumludur.” şeklindedir (Öğretmen Adayı 8 tasarımına ait ChatGPT görüşü).

Öğretmen adayları bireysel olarak gerçekleştirdikleri deney tasarımı sürecinin ardından ChatGPT’ye özgün tasarımlarına yönelik “Mevcut deney tasarımının problem durumu öğretim çıktısına uygun mudur, geliştirilmesi gereken herhangi bir şey var mıdır?” sorusunu yönelmiştir. Katılımcıların özgün deney tasarımlarına yönelik değerlendirmeler Tablo 2’de sunulmuştur.

Tablo 2. Problem durumu tasarlama adımı ChatGPT’nin katılımcıların deney tasarımı yönelik değerlendirmelerinin dağılımı

ChatGPT Değerlendirmesi	f	%
Yetersiz	4	12,50
Kısmen yeterli, öneri var	19	59,38
Yeterli	8	25
Bir fikir belirtmeyip ChatGPT’den öneri istenmiş	5	15,63
Bununla ilgili soru sorulmamış	2	6,25

ChatGPT’nin öğretmen adaylarının özgün deney tasarımlarında yer alan problem durumlarını kısmen yeterli bulduğu ancak çeşitli geliştirme önerileri sunduğu görülmektedir. Katılımcıların herhangi bir fikir belirtmeyip doğrudan ChatGPT’den bu basamak hakkında fikir sunmasının istendiği ve katılımcıların bir kısmının özgün deney tasarımlarının problem durumu hakkında ChatGPT’ye herhangi bir soru sormadığı görülmektedir.

Katılımcı tasarımını yetersiz olarak değerlendiren ChatGPT görüşlerinden biri “Neden uygun değil? Çünkü bu haliyle sadece ne olduğu anlatılıyor, ama neden olduğu sorgulanmıyor. İyi bir problem cümlesi, öğrenciye düşünme alanı bırakmalı ve çözüm arayışına yönlendirmelidir.” şeklindedir (Öğretmen Adayı 20 tasarımına ait ChatGPT görüşü).

Katılımcı tasarımını kısmen yeterli olarak değerlendiren ChatGPT görüşlerinden biri “Evet, ‘Enes vücudundaki bu sistemin nasıl çalıştığını anlayabilecek mi?’ sorusu problematik (problem cümlesi) olarak ele alındığında kazanıma uygundur, ancak daha açık ve hedefe odaklı hâle getirilebilir.” şeklindedir (Öğretmen Adayı 2 tasarımına ait ChatGPT görüşü).

Katılımcı tasarımını yeterli olarak değerlendiren ChatGPT görüşlerinden biri “Evet, verdiğin 4 problem sorusu senaryoya çok uygun ve doğrudan FB.7.5.3.2 kazanımının alt kazanımlarını karşılıyor.” şeklindedir (Öğretmen Adayı 16 tasarımına ait ChatGPT görüşü).

Deney tasarımının problem durumuna ait herhangi bir içerik belirtmeyip ChatGPT’den tasarım önerisi istenmesi durumunda ChatGPT’nin katılımcıya verdiği dönütlerden biri “Elbette! Aşağıda, senaryoya tamamen uygun, kazanıma dayalı bir problem durumu hazırladım. Bu problem, öğrencinin hem kavramsal anlayışını hem de bilimsel çıkarım yapma becerisini ölçmeyi amaçlar.” şeklindedir (Öğretmen Adayı 4 tasarımına ait ChatGPT görüşü).

Öğretmen adayları bireysel olarak gerçekleştirdikleri deney tasarımı sürecinin ardından ChatGPT’ye özgün tasarımlarına yönelik “Mevcut deney tasarımının bağımlı değişkeni öğretim çıktısına uygun

mudur, geliştirilmesi gereken herhangi bir şey var mıdır?” sorusunu yöneltmiştir. Katılımcıların özgün deney tasarımlarına yönelik değerlendirmeler Tablo 3’te sunulmuştur.

Tablo 3. Bağımlı değişken tasarlama adımı ChatGPT’nin katılımcıların deney tasarımına yönelik değerlendirmelerinin dağılımı

ChatGPT Değerlendirmesi	f	%
Yetersiz	6	18,75
Kısmen yeterli, öneri var	6	18,75
Yeterli	9	28,13
Bir fikir belirtmeyip ChatGPT’den öneri istenmiş	11	34,38
Katılımcı sormadan ChatGPT öneri vermiş	1	3,13
ChatGPT’den direkt sunulan şablonun geliştirilmesi istenmiş	1	3,13
Bununla ilgili soru sorulmamış	2	6,25

Katılımcıların deney tasarımının bağımlı değişken tasarlama aşamasında herhangi bir fikir belirtmeyip ChatGPT tarafından fikir önerisine başvurdukları görülmektedir. Bununla birlikte deney tasarımının bağımlı değişkeni hakkında ChatGPT görüşüne başvuran öğretmen adaylarının özgün deney tasarımlarında yer alan bağımlı değişkenlerin bir kısmının yeterli bulduğu büyük bir kısmının ise kısmen yeterli bulunup ChatGPT tarafından geliştirme önerisi sunulduğu ya da direkt olarak yetersiz bulunduğu görülmektedir.

Katılımcı tasarımını yetersiz olarak değerlendiren ChatGPT görüşlerinden biri “Hayır, ‘uygulanan kuvvet’ bu senaryo için bağımlı değişken değildir, bu bir bağımsız değişkendir.” şeklindedir (Öğretmen Adayı 5 tasarımına ait ChatGPT görüşü).

Katılımcı tasarımını kısmen yeterli olarak değerlendiren ChatGPT görüşlerinden biri “Senin yazdığın bağımlı değişken tanımı doğru yöne işaret ediyor, fakat biraz daha bilimsel ve net ifadelerle yazılırsa daha sağlam olur.” şeklindedir (Öğretmen Adayı 32 tasarımına ait ChatGPT görüşü).

Katılımcı tasarımını yeterli olarak değerlendiren ChatGPT görüşlerinden biri “Evet, ampul parlaklığı bu kazanım kapsamında bağımlı değişken olmaya kesinlikle uygundur.” şeklindedir (Öğretmen Adayı 29 tasarımına ait ChatGPT görüşü).

Deney tasarımının bağımlı değişkenine ait herhangi bir içerik sunmayıp ChatGPT’den tasarım önerisi istenmesi durumunda ChatGPT’nin katılımcıya verdiği dönütlerden biri “Tabii! Aşağıda senaryoya uygun olacak şekilde, ‘yoğunluğa bağlı olarak değişen’ farklı bağımlı değişkenler geliştiriyorum.” şeklindedir (Öğretmen Adayı 27 tasarımına ait ChatGPT görüşü).

Öğretmen adayları bireysel olarak gerçekleştirdikleri deney tasarımı sürecinin ardından ChatGPT’ye özgün tasarımlarına yönelik “Mevcut deney tasarımının bağımsız değişkeni öğretim çıktısına uygun mudur, geliştirilmesi gereken herhangi bir şey var mıdır?” sorusunu yöneltmiştir. Katılımcıların özgün deney tasarımlarına yönelik değerlendirmeler Tablo 4’te sunulmuştur.

Tablo 4. Bağımsız değişken tasarlama adımı ChatGPT’nin katılımcıların deney tasarımına yönelik değerlendirmelerinin dağılımı

ChatGPT Değerlendirmesi	f	%
Yetersiz	7	21,88
Kısmen yeterli, öneri var	6	18,75
Yeterli	11	34,38
Bir fikir belirtmeyip ChatGPT’den öneri istenmiş	11	34,38
ChatGPT’den direkt sunulan şablonun geliştirilmesi istenmiş	1	3,13
Katılımcı sormadan ChatGPT öneri vermiş	1	3,13
Bununla ilgili soru sorulmamış	3	9,38

Katılımcıların deney tasarımının bağımsız değişken tasarlama aşamasında herhangi bir fikir belirtmeyip ChatGPT tarafından fikir önerisine başvurdıkları görülmektedir. Bununla birlikte deney tasarımının bağımsız değişkeni hakkında ChatGPT görüşüne başvuran öğretmen adaylarının özgün deney tasarımlarında yer alan bağımsız değişkenlerin bir kısmının yeterli bulunduğu büyük bir kısmının ise kısmen yeterli bulunup ChatGPT tarafından geliştirme önerisi sunulduğu ya da direkt olarak yetersiz bulunduğu görülmektedir.

Katılımcı tasarımını yetersiz olarak değerlendiren ChatGPT görüşlerinden biri “Hayır, ‘su ve yağ’ ifadesi bu haliyle bağımsız değişken olarak eksik veya hatalıdır. Doğru ifade: Kullanılan sıvı türü bağımsız değişkendir. Bu sıvı türleri örnek olarak ‘su’ ve ‘yağ’ olabilir.” şeklindedir (Öğretmen Adayı 32 tasarımına ait ChatGPT görüşü).

Katılımcı tasarımını kısmen yeterli olarak değerlendiren ChatGPT görüşlerinden biri “‘Çok su içmesi’ bağımsız değişken olarak kazanıma uygun, ancak modelle desteklenen bir etkinlik ya da gözlem süreciyle ilişkilendirilirse daha etkili olur.” şeklindedir (Öğretmen Adayı 2 tasarımına ait ChatGPT görüşü).

Katılımcı tasarımını yeterli olarak değerlendiren ChatGPT görüşlerinden biri “Evet, ‘lambda sayısı’ bu kazanım kapsamında deneyin bağımsız değişkeni olarak uygun, anlamlı ve öğretici bir seçimdir.” şeklindedir (Öğretmen Adayı 1 tasarımına ait ChatGPT görüşü).

Deney tasarımının bağımsız değişkenine ait herhangi bir içerik sunmayıp ChatGPT’den tasarım önerisi istenmesi durumunda ChatGPT’nin katılımcıya verdiği dönütlerden biri “Bağımsız değişken: Ali’nin sıcak tencereye ya da buhara temas etmesi. Çünkü bu değişken, sonucu (elinin yanıp yanmaması) etkileyen dış etkidir. Ali bu davranışı yapınca sonuç değişiyor.” şeklindedir (Öğretmen Adayı 3 tasarımına ait ChatGPT görüşü).

Öğretmen adayları bireysel olarak gerçekleştirdikleri deney tasarımı sürecinin ardından ChatGPT’ye özgün tasarımlarına yönelik “Mevcut deney tasarımının kontrol edilen değişkeni öğretim çıktısına uygun mudur, geliştirilmesi gereken herhangi bir şey var mıdır?” sorusunu yöneltmiştir. Katılımcıların özgün deney tasarımlarına yönelik değerlendirmeler Tablo 5’te sunulmuştur.

Tablo 5. Kontrol edilen değişken tasarlama adımı ChatGPT’nin katılımcıların deney tasarımına yönelik değerlendirmelerinin dağılımı

ChatGPT Değerlendirmesi	f	%
Yetersiz	5	15,63
Kısmen yeterli, öneri var	11	34,38
Yeterli	12	37,5
Bir fikir belirtmeyip ChatGPT’den öneri istenmiş	11	34,38
Katılımcı sormadan ChatGPT öneri vermiş	2	6,25
Bununla ilgili soru sorulmamış	4	12,50

Katılımcıların deney tasarımının kontrol edilen değişkeni hakkında ChatGPT görüşüne başvuran öğretmen adaylarının özgün deney tasarımlarında yer alan kontrol edilen değişkenlerinin önemli bir kısmının yeterli bulunduğu büyük bir kısmının ise kısmen yeterli bulunup ChatGPT tarafından geliştirme önerisi sunulduğu ya da direkt olarak yetersiz bulunduğu görülmektedir. Bununla birlikte çeşitli öğretmen adaylarının deney tasarımının kontrol edilen değişken tasarlama aşamasında herhangi bir fikir belirtmeyip ChatGPT tarafından fikir önerisine başvurdıkları görülmektedir.

Katılımcı tasarımını yetersiz olarak değerlendiren ChatGPT görüşlerinden biri “‘Kullanılan asit ve baz maddeleri’ ifadesi, kontrol değişkeni olarak uygun değildir — çünkü bu ifade deneyde sabit tutulan değil, değiştirilen bir unsuru ifade ediyor. Yani bu cümle aslında bağımsız değişkeni tarif eder” şeklindedir (Öğretmen Adayı 21 tasarımına ait ChatGPT görüşü).

Katılımcı tasarımını kısmen yeterli olarak değerlendiren ChatGPT görüşlerinden biri “Harika bir noktaya değindin! Yazdığın kontrol değişkenleri temel olarak doğru ama bilimsel netlik, tekrar

edilebilirlik ve anlamlılık açısından biraz daha iyileştirilebilir.” şeklindedir (Öğretmen Adayı 32 tasarımına ait ChatGPT görüşü).

Katılımcı tasarımını yeterli olarak değerlendiren ChatGPT görüşlerinden biri “Evet, devre anahtarı bu deneyde kontrollü değişken olmaya kesinlikle uygundur. Çünkü deney süresince sabit tutulan, test edilmeyen ama etkili olabilecek bir değişkendir.” şeklindedir (Öğretmen Adayı 29 tasarımına ait ChatGPT görüşü).

Öğretmen adayları bireysel olarak gerçekleştirdikleri deney tasarımı sürecinin ardından ChatGPT’ye özgün tasarımlarına yönelik “Mevcut deney tasarımının uygulama basamakları öğretim çıktısına uygun mudur, geliştirilmesi gereken herhangi bir şey var mıdır?” sorusunu yöneltmiştir. Katılımcıların özgün deney tasarımlarına yönelik değerlendirmeler Tablo 6’da sunulmuştur.

Tablo 6. Uygulama basamakları tasarlama adımı ChatGPT’nin katılımcıların deney tasarımı yönelik değerlendirmelerinin dağılımı

ChatGPT Değerlendirmesi	f	%
Yetersiz	2	6,25
Kısmen yeterli, öneri var	21	65,63
Yeterli	5	15,63
Bir fikir belirtmeyip ChatGPT’den öneri istenmiş	3	9,38
ChatGPT’den direkt sunulan şablonun geliştirilmesi istenmiş	1	3,13
Genel bir şablon verip ChatGPT’den uygulama basamaklarını yazması istenmiş	2	6,25
Bununla ilgili soru sorulmamış	5	15,63

ChatGPT’nin öğretmen adaylarının özgün deney tasarımlarında yer alan uygulama basamaklarının büyük bir kısmını kısmen yeterli bulduğu ve çeşitli geliştirme önerileri sunduğu görülmektedir. Katılımcılar tarafından doğrudan ChatGPT’den özgün uygulama basamaklarını geliştirip sunmasının istendiği ve bazı katılımcıların özgün deney tasarımlarının uygulama basamakları hakkında ChatGPT’ye herhangi bir soru sormadığı görülmektedir.

Katılımcı tasarımını yetersiz olarak değerlendiren ChatGPT görüşlerinden biri “Bu üç maddelik uygulama basamakları temel bir sırayı vermekte, ancak kazanımın gerektirdiği ‘deney yapma, analiz etme’ hedeflerine tam olarak uygun değildir.” şeklindedir (Öğretmen Adayı 5 tasarımına ait ChatGPT görüşü).

Katılımcı tasarımını kısmen yeterli olarak değerlendiren ChatGPT görüşlerinden biri “Evet, verdiği 6 adımlık deney basamakları genel olarak anlamlı ve senaryoya uygun; ancak bilimsel doğruluk, kazanım uyumu ve deneysel mantık açısından bazı noktaları iyileştirmek gerekiyor.” şeklindedir (Öğretmen Adayı 7 tasarımına ait ChatGPT görüşü).

Katılımcı tasarımını yeterli olarak değerlendiren ChatGPT görüşlerinden biri “Bu deney, hem tümevarımsal akıl yürütmeyi destekleyen bir yapıdadır hem de günlük yaşam örüntüleri üzerinden sürtünme kuvveti kavramını somutlaştırır. Bu nedenle, belirttiğiniz kazanıma tam anlamıyla uygundur.” şeklindedir (Öğretmen Adayı 22 tasarımına ait ChatGPT görüşü).

Öğretmen adayları bireysel olarak gerçekleştirdikleri deney tasarımı sürecinin ardından ChatGPT’ye özgün tasarımlarına yönelik “Mevcut deney tasarımının fen bilimleri entegrasyonu öğretim çıktısına ve deney tasarımı uygun mudur, geliştirilmesi gereken herhangi bir şey var mıdır?” sorusunu yöneltmiştir. Katılımcıların özgün deney tasarımlarına yönelik değerlendirmeler Tablo 7’de sunulmuştur.

Tablo 7. Tasarımın fen bilimleri entegrasyonunu sağlama adımı ChatGPT’nin katılımcıların deney tasarımına yönelik değerlendirmelerinin dağılımı

ChatGPT Değerlendirmesi	f	%
Yetersiz	1	3,13
Kısmen yeterli, öneri var	13	40,63
Yeterli	15	46,88
Katılımcı sormadan ChatGPT öneri vermiş	2	6,25
Bir fikir belirtmeyip ChatGPT’den öneri istenmiş	6	18,75
Bununla ilgili soru sorulmamış	4	12,50

ChatGPT’nin öğretmen adaylarının özgün deney tasarımlarının fen bilimleri entegrasyonunun önemli bir kısmını yeterli bulduğu görülmektedir. Bununla birlikte katılımcıların deney tasarımlarının fen bilimleri entegrasyonunun önemli bir kısmının ise ChatGPT tarafından kısmen yeterli olarak değerlendirildiği ve çeşitli geliştirme önerileri sunulduğu görülmektedir. Katılımcılar tarafından doğrudan ChatGPT’den deney tasarımı fen bilimleri entegrasyonu için fikir sunmasının istendiği ve bazı katılımcıların özgün deney tasarımlarının fen bilimleri entegrasyonu hakkında ChatGPT’ye herhangi bir soru sormadığı da görülmektedir.

Katılımcı tasarımını yetersiz olarak değerlendiren ChatGPT görüşlerinden biri “Bu cümle, STEM kapsamında fen bilimleri ile disiplinlerarası ilişkilendirme amacı taşıyorsa, eksik ve yüzeysel kalmaktadır.” şeklindedir (Öğretmen Adayı 1 tasarımı ChatGPT görüşü).

Katılımcı tasarımını kısmen yeterli olarak değerlendiren ChatGPT görüşlerinden biri “Güzel bir başlangıç yapmışsın ama yazının bilimsel dili açısından biraz daha netleştirilmesi ve güçlendirilmesi gerekiyor. Yazdığın ifade, fen bilimleri ile ilişki kuruyor ama hedef kazanımı ve deneysel süreci açıklama konusunda eksik kalıyor.” şeklindedir (Öğretmen Adayı 32 tasarımı ChatGPT görüşü).

Katılımcı tasarımını yeterli olarak değerlendiren ChatGPT görüşlerinden biri “Evet, fen bilimleri ile ilişkilendirmen hem pedagojik hem de kazanım temelli olarak son derece uygun ve yerindedir.” şeklindedir (Öğretmen Adayı 27 tasarımı ChatGPT görüşü).

Öğretmen adayları bireysel olarak gerçekleştirdikleri deney tasarımı sürecinin ardından ChatGPT’ye özgün tasarımlarına yönelik “Mevcut deney tasarımı teknoloji entegrasyonu öğretim çıktısına ve deney tasarımı uygun mudur, geliştirilmesi gereken herhangi bir şey var mıdır?” sorusunu yöneltmiştir. Katılımcıların özgün deney tasarımlarına yönelik değerlendirmeler Tablo 8’de sunulmuştur.

Tablo 8. Tasarımın teknoloji entegrasyonunu sağlama adımı ChatGPT’nin katılımcıların deney tasarımına yönelik değerlendirmelerinin dağılımı

ChatGPT Değerlendirmesi	f	%
Yetersiz	4	12,50
Kısmen yeterli, öneri var	11	34,38
Yeterli	7	21,88
Bir fikir belirtmeyip ChatGPT’den öneri istenmiş	7	21,88
Katılımcı bu STEM aşamasının entegrasyonu olmadığını belirtmiş	1	3,13
Katılımcı sormadan ChatGPT öneri vermiş	2	6,25
Deney tasarımı bu STEM aşamasının entegrasyonu bulunmamaktadır	1	3,13
ChatGPT’den direkt sunulan şablonun geliştirilmesi istenmiş	1	3,13
Bununla ilgili soru sorulmamış	4	12,50

ChatGPT’nin öğretmen adaylarının özgün deney tasarımlarının teknoloji entegrasyonunun büyük bir kısmını kısmen yeterli bulduğu ve çeşitli geliştirme önerileri sunduğu görülmektedir. Katılımcılar tarafından doğrudan ChatGPT’den deney tasarımı teknoloji entegrasyonu için fikir sunmasının istendiği ve bazı katılımcılar tarafından özgün deney tasarımlarının teknoloji entegrasyonu hakkında ChatGPT’ye herhangi bir soru sorulmamış olduğu görülmektedir.

Katılımcı tasarımını yetersiz olarak değerlendiren ChatGPT görüşlerinden biri “‘Öğretmenin söylediklerini telefonuna not alma’ ifadesi, teknoloji ilişkilendirmesi açısından kazanıma sınırlı ölçüde uygundur; yani yüzeysel bir ilişki kurulmuş, ancak teknolojik düşünme, araç kullanımı veya veri işleme gibi daha derin unsurlar açısından zayıftır.” şeklindedir (Öğretmen Adayı 2 tasarımına ait ChatGPT görüşü).

Katılımcı tasarımını kısmen yeterli olarak değerlendiren ChatGPT görüşlerinden biri “‘Senin ifaden doğru yolda ve deneyinle uyumlu. Ama ben olsaydım, öğretim kazanımlarına uygun ve STEM diline entegre olması açısından şöyle yazardım: ‘Maddelerin sıcaklıklarını düşürerek donma olayını başlatmak için teknolojik araçlar (buzluk vb.) kullanır ve bu araçların deneysel gözlemlerdeki katkısını değerlendirir.’” şeklindedir (Öğretmen Adayı 6 tasarımına ait ChatGPT görüşü).

Katılımcı tasarımını yeterli olarak değerlendiren ChatGPT görüşlerinden biri “‘Değerlendirme: Tam uygun. Bu ifade STEM kapsamında teknoloji boyutunu net ve doğrudan karşılıyor. Çünkü teknoloji boyutu, öğrencilerin araç-gereç kullanarak veri toplaması, günlük yaşamda kullanılan teknolojik cihazları bilimsel amaçla kullanması, gözlemlerini destekleyecek dijital araçlardan yararlanması gibi durumları kapsar.’” şeklindedir (Öğretmen Adayı 12 tasarımına ait ChatGPT görüşü).

Öğretmen adayları bireysel olarak gerçekleştirdikleri deney tasarımı sürecinin ardından ChatGPT’ye özgün tasarımlarına yönelik “‘Mevcut deney tasarımının mühendislik entegrasyonu öğretim çıktısına ve deney tasarımına uygun mudur, geliştirilmesi gereken herhangi bir şey var mıdır?’” sorusunu yöneltmiştir. Katılımcıların özgün deney tasarımlarına yönelik değerlendirmeler Tablo 9’da sunulmuştur.

Tablo 9. Tasarımın mühendislik entegrasyonunu sağlama adımı ChatGPT’nin katılımcıların deney tasarımına yönelik değerlendirmelerinin dağılımı

ChatGPT Değerlendirmesi	f	%
Yetersiz	3	9,38
Kısmen yeterli, öneri var	16	50
Yeterli	5	15,63
Katılımcı bu STEM aşamasının entegrasyonu olmadığını belirtmiş	1	3,13
Katılımcı sormadan ChatGPT öneri vermiş	2	6,25
Bir fikir belirtmeyip ChatGPT’den öneri istenmiş	7	21,88
Bununla ilgili soru sorulmamış	7	21,88

ChatGPT’nin öğretmen adaylarının özgün deney tasarımlarının mühendislik entegrasyonunun büyük bir kısmını kısmen yeterli bulduğu ve çeşitli geliştirme önerileri sunduğu görülmektedir. Katılımcılar tarafından doğrudan ChatGPT’den deney tasarımının mühendislik entegrasyonu için fikir sunmasının istendiği ve bazı katılımcılar tarafından özgün deney tasarımlarının mühendislik entegrasyonu hakkında ChatGPT’ye herhangi bir soru sorulmamış olduğu görülmektedir.

Katılımcı tasarımını yetersiz olarak değerlendiren ChatGPT görüşlerinden biri “‘Verdiğin metin, mühendislik açısından deneyle bağlantı kurma niyeti açısından yerinde bir giriş yapıyor. Ancak bu haliyle fikirler güzel ama dağınık yapılandırılmış, mühendislik süreci (tasarım, test, iyileştirme gibi) açıkça ifade edilmemiş, teknik netlik ve dilsel bütünlük eksik kalmış.’” şeklindedir (Öğretmen Adayı 18 tasarımına ait ChatGPT görüşü).

Katılımcı tasarımını kısmen yeterli olarak değerlendiren ChatGPT görüşlerinden biri “‘İfade çok genel kalıyor. Mühendislik ile daha güçlü bir ilişkilendirme yapılabilmesi için öğrencinin bir problemi çözmek amacıyla model tasarlaması, malzeme seçimi, işlevsellik, sunum gibi süreçlere girmesi, tasarımı yaparken bilimsel bilgiyle desteklenen kararlar alması gibi yönler vurgulanmalıdır.’” şeklindedir (Öğretmen Adayı 8 tasarımına ait ChatGPT görüşü).

Bir diğer katılımcı tasarımını kısmen yeterli olarak değerlendiren ChatGPT görüşlerinden biri ise “‘Bu ifade STEM yaklaşımı açısından kabul edilebilir ve doğrudur, çünkü mühendislik bileşeni, genellikle bir sorunu çözmek veya bir sistem/araç/ortam kurmak ile ilişkilidir. Deney düzeneği kurmak da bu

kapsamda yer alır. Ancak ifade fazla genel kalıyor. Öğrencinin mühendis gibi düşünme becerisini (tasarlama, karar verme, malzeme seçme, sistem kurma) yansıtmak açısından geliştirilebilir.” şeklindedir (Öğretmen Adayı 12 tasarımına ait ChatGPT görüşü).

Katılımcı tasarımını yeterli olarak değerlendiren ChatGPT görüşlerinden biri “Tam isabet! Bu süreç mühendislik tasarım döngüsüne (planla – uygula – test et – geliştir) birebir uygundur. Test edilen farklı malzemelerle sorun çözme ve yeniden tasarlama becerisi de gösterilmiştir.” şeklindedir (Öğretmen Adayı 24 tasarımına ait ChatGPT görüşü).

Deney tasarımının mühendislik entegrasyonuna ait herhangi bir içerik belirtmeyip ChatGPT’den tasarım önerisi istenmesi durumunda ChatGPT’nin katılımcıya verdiği dönütlerden biri “Harika bir noktaya geldik! Şimdi deneyini mühendislik disipliniyle ilişkilendirelim. Mühendislik öğrencilerin gerçek dünya problemlerine çözüm üretmesini, tasarım yaparak öğrenmesini sağlar. İlişkilendirme sayesinde öğrenciler fen bilgisini problem çözmeye dönüştürür, yoğunluk kavramının gerçek dünyadaki etkisini mühendis bakış açısıyla analiz eder, hayal gücü ve eleştirel düşünmeyle somut bir ürün tasarlar” şeklindedir (Öğretmen Adayı 11 tasarımına ait ChatGPT görüşü).

Öğretmen adayları bireysel olarak gerçekleştirdikleri deney tasarımı sürecinin ardından ChatGPT’ye özgün tasarımlarına yönelik “Mevcut deney tasarımının matematik entegrasyonu öğretim çıktısına ve deney tasarımına uygun mudur, geliştirilmesi gereken herhangi bir şey var mıdır?” sorusunu yöneltmiştir. Katılımcıların özgün deney tasarımlarına yönelik değerlendirmeler Tablo 10’da sunulmuştur.

Tablo 10. Tasarımın matematik entegrasyonunu sağlama adımı ChatGPT’nin katılımcıların deney tasarımına yönelik değerlendirmelerinin dağılımı

ChatGPT Değerlendirmesi	f	%
Yetersiz	2	6,25
Kısmen yeterli, öneri var	16	50
Yeterli	4	12,5
Katılımcı sormadan ChatGPT öneri vermiş	2	6,25
Deney tasarımında bu STEM aşamasının entegrasyonu bulunmamaktadır	1	3,13
Bir fikir belirtmeyip ChatGPT’den öneri istenmiş	7	21,88
Bununla ilgili soru sorulmamış	8	25

ChatGPT’nin öğretmen adaylarının özgün deney tasarımlarının matematik entegrasyonunun önemli bir kısmını kısmen yeterli bulduğu ve çeşitli geliştirme önerileri sunduğu görülmektedir. Katılımcılar tarafından herhangi bir tasarım fikri sunulmadan doğrudan ChatGPT’den deney tasarımının matematik entegrasyonu için fikir sunmasının istendiği ve bazı katılımcılar tarafından özgün deney tasarımlarının matematik entegrasyonu hakkında ChatGPT’ye herhangi bir soru sorulmamış olduğu görülmektedir.

Katılımcı tasarımını yetersiz olarak değerlendiren ChatGPT görüşlerinden biri “‘Asit ve bazların pH değerlerini periyodik cetvelde belirleme’ ifadesi, matematiksel süreçler açısından doğrudan uygun değildir. Çünkü $pH \neq$ periyodik cetvel ilişkisi doğrudan değildir. Matematiksel olarak pH değeriyle işlem yapılabilir ama bunu periyodik cetvel üzerinde göstermek fen kavramıdır, matematik süreci değil.” şeklindedir (Öğretmen Adayı 21 tasarımına ait ChatGPT görüşü).

Katılımcı tasarımını kısmen yeterli olarak değerlendiren ChatGPT görüşlerinden biri “Güzel bir düşünce ortaya koymuşsun! Matematikle ilişkilendirme denemen özünde doğru bir yaklaşım içeriyor, çünkü deneydeki sıvı-katı ilişkisi üzerinden oran, miktar ve işlem kavramlarına dayalı bir matematiksel bağ kuruluyor. Ancak ifadeni hem daha netleştirebiliriz hem de matematiksel tutarlılığı ve akademik dili güçlendirebiliriz.” şeklindedir (Öğretmen Adayı 9 tasarımına ait ChatGPT görüşü).

Bir diğer katılımcı tasarımını kısmen yeterli olarak değerlendiren ChatGPT görüşü ise “Metnin niyeti oldukça doğru; sayısal verilerin anlaşılmayı kolaylaştırdığı, ses frekansı ölçümünün matematiksel karşılığı olduğu gibi noktalar yerinde. Ancak şu anki haliyle metin dil açısından biraz karmaşık ve

tekrar içeriyor, matematiksel işlem veya kavram içeriği net belirtilmemiş, öğrencinin matematiksel becerisiyle neyi geliştirdiği doğrudan yazılmamış.” şeklindedir (Öğretmen Adayı 18 tasarımına ait ChatGPT görüşü).

Katılımcı tasarımını yeterli olarak değerlendiren ChatGPT görüşlerinden biri “Matematik entegrasyonu: Tam uygun. Zaman ölçme ve karşılaştırma içerdiği için uygun.” şeklindedir (Öğretmen Adayı 25 tasarımına ait ChatGPT görüşü).

Deney tasarımının matematik entegrasyonuna ait herhangi bir içerik belirtmeyip ChatGPT’den tasarım önerisi istenmesi durumunda ChatGPT’nin katılımcıya verdiği dönütlerden biri “Matematik Kazanımı: Hâl değişim süreçleri ile ilgili sıcaklık ve süre verilerini toplar, tabloya aktarır, karşılaştırma ve çıkarım yapar. Termometre ile sıcaklık ölçümü, süre takibi, tablo oluşturma ve yorumlama içerir.” şeklindedir (Öğretmen Adayı 6 tasarımına ait ChatGPT görüşü).

TARTIŞMA, SONUÇ ve ÖNERİLER

Araştırmada öğretmen adaylarının ChatGPT’yi deney tasarımlarının çeşitli basamaklarına dair danışma, doğrulama, geliştirme, öneri sunma ve düzeltme gibi amaçlarla yoğun biçimde kullandıkları görülmektedir. Katılımcıların değişkenlerin belirlenmesi, uygulama basamaklarının ChatGPT önerisi doğrultusunda güncellenmesi ve deney tasarımının STEM entegrasyonunun sağlanması ya da güçlendirilmesi başta olmak üzere deney tasarımlarının en az bir bileşeninde ChatGPT değerlendirmesi doğrultusunda değişiklik yapmış olduğu saptanmıştır. Katılımcıların tasarımına ait ChatGPT değerlendirmeleri incelendiğinde katılımcıların çoğunluğunun kontrol edilen değişken tasarımının diğer bileşenlere göre daha yüksek oranda ChatGPT tarafından yeterli bulunduğu görülmektedir. Öğretmen adaylarının özgün deney tasarımlarının STEM entegrasyonuna dair ChatGPT değerlendirilmeleri incelendiğinde ise tasarımların fen bilimleri entegrasyonunun yüksek oranda ChatGPT tarafından yeterli bulunduğu tespit edilmiştir. Bununla birlikte ChatGPT’nin katılımcıların özgün deney tasarımlarının günlük hayat senaryosu, problem durumu, uygulama basamakları, deney tasarımının teknoloji entegrasyonu, deney tasarımının mühendislik entegrasyonu ve deney tasarımının matematik entegrasyonu adımlarını kısmen yeterli olarak değerlendirdiği bu doğrultuda tasarımlara yönelik çeşitli geliştirme önerileri sunduğu görülmektedir. Ayrıca öğretmen adaylarının deney tasarımının bağımsız değişken belirleme basamağının ChatGPT tarafından en sık yetersiz olarak değerlendirilen basamak olduğu dikkat çekmektedir. Katılımcıların önemli bir kısmının bağımlı değişken, bağımsız değişken ve kontrol edilen değişken belirleme basamaklarında herhangi bir fikir sunmadan direkt ChatGPT önerisine başvurmuş oldukları tespit edilmiştir. Katılımcıların ChatGPT görüşüne en az oranda başvurdukları basamağın ise deney tasarımlarının matematik entegrasyonunu sağlama adımı olduğu görülmekte, katılımcıların tasarımlarının matematik entegrasyonu hakkında ChatGPT’ye herhangi bir soru yöneltmedikleri ve bu aşama için ChatGPT görüşüne diğer adımlara göre daha az düzeyde başvurmuş oldukları dikkat çekmektedir.

Cai ve diğerleri (2025) tarafından vurgulanan, yüksek ön bilgi düzeyine sahip kullanıcıların ChatGPT desteğiyle daha nitelikli bir gelişim sergilediği sonucu, bu araştırmanın sonuçlarıyla teorik bir uyum içerisindedir. Araştırma kapsamında, öğretim çıktısına dair taslak fikirlerini sisteme sunan öğretmen adaylarının, yapay zekanın değerlendirmelerini mevcut tasarımlarını pedagojik ve bilimsel açıdan iyileştirmek için birer geliştirme fırsatı olarak kullandıkları görülmüştür. Buna karşın, sürece dair herhangi bir içerik sunmayan adayların; değişkenlerin belirlenmesi, mühendislik ve matematik entegrasyonu gibi teknik uzmanlık gerektiren aşamalarda doğrudan hazır tasarım talep ederek yapay zekayı birincil üretici rolünde konumlandıkları saptanmıştır. Bu durum, öğretmen adaylarının sunduğu girdilerin niteliği ve derinliği arttıkça ChatGPT’nin "sıfırdan içerik üreten bir kaynaktan", mevcut tasarımı teknik hatalar (örneğin; bağımlı-bağımsız değişken karmaşası veya fen kavramlarının matematiksel süreçlerle hatalı ilişkilendirilmesi) açısından rafine eden bir "eleştirmen ve geliştiriciye" dönüştüğünü kanıtlamaktadır. Sonuç olarak, yapay zekanın tasarım süreci üzerindeki dönüştürücü etkisi, kullanıcının sürece dahil ettiği akademik hazırlık düzeyi ve sorgu niteliğiyle doğrudan ilişkilidir.

Chellappa ve Luximon (2024) tarafından yürütülen çalışmanın bulguları ChatGPT yanıtlarının katılımcılar tarafından bağlama uygun yapılandırılmış orta doğruluk düzeyinde yanıtlar olarak değerlendirildiğini bununla birlikte ChatGPT'nin öğrenmeye etkili destek sağlayabilecek bir tamamlayıcı araç olarak görüldüğünü ifade etmektedir. Mevcut çalışmada katılımcıların deney tasarımı sürecinin bağımlı, bağımsız ve kontrol edilen değişken belirleme adımlarında herhangi bir fikir belirtmeden ChatGPT önerisine deney tasarımı sürecinin diğer adımlarına göre daha fazla başvurdukları tespit edilmiştir. Bununla birlikte öğretmen adaylarının deney tasarımının uygulama basamaklarını belirleme aşamasında diğer aşamalara göre daha az ChatGPT önerisine başvurmuş oldukları saptanmıştır. Deney tasarımının STEM entegrasyonu gerçekleştirilirken ise en az matematik bileşeni hakkında ChatGPT'ye soru sorulmuş olması dikkat çekmektedir. Öğretmen adaylarının ChatGPT önerisine başvurdukları ve başvurmadıkları tasarım aşamaları göz önünde bulundurulduğunda katılımcıların literatürde yer alan çalışmalara uyumlu biçimde ChatGPT'yi tamamlayıcı bir araç olarak kullanma eğiliminde oldukları görülmektedir.

Araştırmadan elde edilen veriler, öğretmen adaylarının deney tasarımı süreçlerinde sergiledikleri pedagojik performans ile Zhu ve ark. (2025) tarafından vurgulanan "yapay zeka destekli gelişim" ihtiyacı arasında güçlü bir bağ olduğunu göstermektedir. Adayların hazırladığı özgün senaryoların büyük bir çoğunluğunun ChatGPT tarafından çeşitli revizyon önerileriyle karşılanması, adayların henüz bağımsız materyal üretiminde tam bir pedagojik yetkinliğe ulaşamadıklarına dair önemli bir göstergedir. Bu durum, Zhu ve ark. (2025) çalışmasında ifade edilen; öğretmen adaylarının sınıf içi pratik eksikliği ve değerlendirme okuryazarlığındaki sınırlılıklar nedeniyle geri bildirim süreçlerinde yaşadıkları güçlükleri bizzat uygulama düzeyinde kanıtlamaktadır. Yapay zekanın adaylara sunduğu güvenlik, birim kullanımı ve kazanım uyumu gibi teknik müdahaleler, literatürde savunulan "ölçüt temelli geri bildirim" mekanizmasının işlevselliğini ortaya koymaktadır. Özellikle adayların raporlarındaki matematiksel tutarlılık ve akademik dil eksikliklerinin ChatGPT tarafından raporlanması, yapay zekanın adayın bilişsel süreçlerinde fark edemediği kör noktaları aydınlatan bir "pedagojik iskele" işlevi gördüğünü doğrulamaktadır. Öte yandan, bazı adayların kendi fikirlerini sunmak yerine doğrudan yapay zekadan çözüm talep etme eğilimi, literatürde "yapay zekaya aşırı bağımlılık" olarak tanımlanan ve adayın öz-yeterlik gelişimini tehdit edebilecek bir risk faktörü olarak öne çıkmaktadır. Sonuç olarak bu bulgular, öğretmen eğitiminde yapay zekanın stratejik kullanımına yönelik hedefli eğitimlerin gerekliliğini vurgulayan güncel literatürle (Zhu vd., 2025) tam bir uyum sergilemekte ve teknolojinin eleştirel bir süzgeçten geçirilerek mesleki sürece dahil edilmesinin önemini pekiştirmektedir.

Katılımcılar deney tasarımlarının değişken belirleme aşamalarında ChatGPT önerilerine sıklıkla başvurmuşlardır. Öğretmen adaylarının kontrol edilen değişken tasarımları ChatGPT tarafından genellikle yeterli olarak değerlendirilirken bağımsız ve bağımlı değişken tasarımlarının kısmen yeterli olarak değerlendirildiği ve ChatGPT tarafından katılımcılara tasarımlarını geliştirmeye yönelik önerilerin sunulduğu görülmektedir. Öğretmen adaylarının özgün deney tasarımlarının değişkenleri hakkındaki ChatGPT görüşleri ve önerileri incelendiğinde katılımcıların bir deneyin değişkenlerini belirlerken değişkeni doğru sınıflandırmakta yeterli olmadıkları görülmektedir. Ateş (2005) tarafından yürütülen çalışmanın bulguları öğretmen adaylarının deneysel çalışmalarda değişkenleri tanımlama, değişkenler arası ilişkileri kurma ve diğer değişkenleri süreç boyunca kontrol etme konusunda yetersiz olduklarını göstermektedir. Ateş (2005) tarafından yürütülen çalışmanın bulgularına benzer biçimde mevcut çalışma ile elde edilen bulgular öğretmen adaylarının bağımlı değişken, bağımsız değişken ve kontrol edilen değişken kavramları hakkında yanlışlara sahip olduklarını, deney tasarımı sürecinde özellikle bağımlı değişken ve bağımsız değişken kavramlarını tanımlamakta ve bu kavramlar arasında ayırım yapmakta güçlük yaşadıklarını göstermektedir. Bu durum değişkenleri belirleme ve kontrol etme becerilerinin temel süreç becerilerine göre daha karmaşık ve üst düzey bilişsel beceriler gerektirdiğine işaret etmektedir (Ateş, 2005; Brotherton & Preece, 1995). Kılıç, Yardımcı ve Metin (2009) tarafından gerçekleştirilen çalışmanın bulguları öğretmen adaylarının en çok kontrol edilen değişken belirleme aşamasında hata yapmakta olduklarını belirtmektedir. Kontrol edilen değişkeni ifade etmeleri gereken tasarım aşamasında çoğunlukla deneyin bağımlı değişkenini belirten öğretmen adaylarının kontrol ifadesini günlük hayatta kullandıkları haline göre algılayıp

deney boyunca değişken üzerine etkisinin ne olduğu incelenen değişken olduğu göz önünde bulundurularak tasarlandığı dikkat çekmektedir. Aynı çalışmanın bulguları öğretmen adaylarının çoğunlukla bağımlı değişken tasarımları gereken tasarım aşamasında bağımsız değişkeni tasarlamış olduklarını göstermektedir. Bunun nedeninin bağımlı değişken ifadesinin bir şeye bağlı olmak anlamıyla algılanarak öğretmen adayları tarafından bu aşamada gözlenen değişimin neye bağlı olacağının ifade edilmesi istendiğinin düşünülmüş olabileceği ifade edilmektedir. Bu çalışmanın bulgularına benzer biçimde mevcut çalışmanın bulguları da öğretmen adaylarının deneysel süreçlerde değişkenler arasındaki ilişkileri bilimsel bağlamda yapılandırmakta zorlandıklarına işaret etmektedir (Kılıç, Yardımcı, & Metin, 2009).

Öğretmen adaylarının özgün deney tasarımlarının fen bilimleri entegrasyonunun genel olarak ChatGPT tarafından yeterli bulunduğu görülmektedir. Çalışmanın bulguları ChatGPT'nin öğretmen adaylarının özgün deney tasarımlarının matematik entegrasyonunu matematiksel işlem veya kavram içeriğinin net belirtilmesi, öğrencinin matematiksel becerisiyle neyi geliştirdiğinin doğrudan yazılması, matematiksel tutarlılığın eksiliği ve zayıf akademik dil kullanılması gibi açılardan güçlendirilmesi gerektiğine dikkat çekmektedir. Bu bulguların STEM uygulamalarında matematik entegrasyonunun yüzeysel veya örtük biçimde yer aldığını belirten çalışmalarla uyumlu olduğu görülmektedir. Fitzpatrick ve Leavy (2025) tarafından yürütülen çalışmanın bulguları öğretmen adaylarının mühendislik tasarım sürecine odaklanırken matematiği daha çok ikincil veya tesadüfi bileşen olarak ele alıp matematiksel öğrenme hedeflerini açıkça yapılandıramadıklarını göstermektedir. Ayrıca mevcut çalışmada katılımcıların matematik entegrasyonuna yönelik ChatGPT değerlendirmeleri öğretmen adaylarının matematiksel işlem, kavram ve beceri hedeflerini net biçimde tanımlayamadıklarını göstermektedir. Bu bulgu literatürde yer alan Çibik ve Boz-Yaman (2025) ile Kewalramani, Devi ve Ng (2025) tarafından yürütülen çalışmaların STEM uygulamalarında matematik entegrasyonunun açık öğrenme hedefleri açısından eksik ve diğer bileşenlere göre daha zayıf biçimde temellendirilen bir bileşen olduğu bulguları ile uyumludur.

Öğretmen adaylarının deney tasarımlarının STEM entegrasyonu tasarlama aşamalarında teknoloji ve mühendislik entegrasyonunu sağlama adımlarında ChatGPT değerlendirmelerine başvurdukları ve ChatGPT'nin tasarımın bu bileşenleri ile ilgili önemli önerilerde bulunduğu görülmektedir. ChatGPT deney tasarımının STEM entegrasyonunun teknoloji bileşenini verilerin teknolojik araç-gereç kullanarak toplaması, günlük yaşamda kullanılan teknolojik cihazların bilimsel amaçla kullanılması ve gözlemleri destekleyecek dijital araçlardan yararlanılması gibi durumları kapsayan bir bileşen olarak tanımlamaktadır. Bu bulgu öğretmenlerin teknoloji içerik ve öğretim tasarımını bütünleştirmekte zorlandıklarını ve teknoloji entegrasyonunu sağlamaya dair yeterli güvene sahip olmadıklarını ortaya koyan Ali, Younis, Ahmad, Saba ve Ullah (2025) tarafından yürütülen çalışma ile elde edilen bulgularla örtüşmektedir. Dockendorff ve Zaccarelli (2025) tarafından yürütülen çalışma sonucunda elde edilen bulgular dijital teknolojinin kavramsal öğrenmeyi destekleyen ve gerçek yaşam bağlamlarıyla öğrenme süreci arasında ilişki kurmayı sağlayan bir araç olarak kullanıldığında anlamlı öğrenme çıktıları sağladığını vurgulamaktadır. Bu bağlamda ChatGPT'nin öğretmen adaylarının deney tasarımlarına yönelik değerlendirmeleri literatürde önerilen dijital teknoloji entegrasyonu anlayışını yalnızca kuramsal düzeye ek olarak uygulamaya dönük somut önerilerle görünür kılmakta ve öğretmen adaylarının STEM entegrasyonunu yapılandırma süreçlerini destekleyici bir rol üstlenmektedir.

STEM entegrasyonunun mühendislik bileşeni ile ilgili değerlendirmelerinde ise ChatGPT'nin bu bileşeni literatür ile örtüşen biçimde genellikle bir sorunu çözmek veya bir sistem, araç ya da ortam kurmak ile ilişkili olarak tanımladığı ve bileşenin amacını deney tasarımını uygulayan bununla birlikte öğrencinin mühendis gibi düşünme becerisini tasarlama, karar verme, malzeme seçme ve sistem kurma gibi becerileri uygulaması ve geliştirmesi yoluyla yansıtmayı sağlamak olarak açıkladığı görülmektedir. Bununla birlikte öğretmen adaylarının özgün deney tasarımlarının mühendislik entegrasyonunu değerlendirme aşamasında ChatGPT'nin katılımcıların tasarımlarının bu aşamasını planla, uygula, test et, geliştir adımları takip edilerek uygulanan mühendislik tasarım döngüsüne uygunluğu ile test edilen farklı malzemelerle sorun çözme ve yeniden tasarlama becerilerine göre değerlendirmiş olduğu görülmektedir. ChatGPT tarafından mühendislik

entegrasyonunun yeterli olarak görülmesi için tasarımın öğrencinin bir problemi çözmek amacıyla model tasarlaması, malzeme seçmesi, tasarımın işlevselliğini değerlendirmesi, bilimsel bilgiyle desteklenen kararlar almasının gerekmesi gibi yönlerin vurgulanması gerekliliğini belirttiği görülmektedir. Öztürk, Aydın-Günbatır ve Roehrig (2025) tarafından yürütülen çalışmanın bulguları test etme ve yeniden tasarlama gibi mühendislik tasarım döngüsünün ileri aşamalarının buna yönelik uygulama süreci sonrasında öğretmenler tarafından gerçekleştirilen tasarımlarda yer almaya başladığını dolayısıyla entegrasyonun bu boyutunun zaman ve deneyim ile gelişen bir doğaya sahip olduğunu göstermektedir. Ayrıca mevcut çalışmayla elde edilen ChatGPT'nin mühendislik entegrasyonunu yeterli olarak değerlendirebilmesi için öğrencinin bilimsel bilgiye dayalı kararlar almasını ve mühendislik tasarım döngüsünü deneyimlemesini beklemesi bulguları García-Carmona, Muñoz-Franco ve Cruz-Guzmán (2025) tarafından yürütülen çalışmanın mühendislik entegrasyonuna yönelik bulgularıyla paralellik göstermektedir. Zou, Kuek, Feng ve Cheng (2025) tarafından yürütülen çalışmanın bulguları dijital araçların geleneksel öğretim yöntemleriyle birlikte kullanıldığında öğrenme deneyimlerini anlamlı biçimde dönüştürme ve geliştirme potansiyeline sahip olduğunu göstermektedir. Çalışmanın bulgularına göre sanal öğretim asistanları ve otomatik değerlendirme sistemleri gibi yapay zekâ destekli araçlar eğitimcilerin kişiselleştirilmiş öğretime daha fazla odaklanmalarına olanak tanımaktadır. Bu çalışmaların bulgularına paralel olarak mevcut çalışmanın bulguları öğretmen adaylarının deney tasarımı sürecinde ihtiyaç duydukları aşamalarda ChatGPT desteğine başvurmuş olup uygulama sürecini kişiselleştirmiş olduklarını göstermektedir.

Bu çalışmada fen bilimleri öğretmen adaylarının öğretim çıktılarını göz önünde bulundurarak hazırlamış oldukları deney tasarımları hakkındaki ChatGPT değerlendirmeleri incelenmiştir. Gelecek çalışmalarda ChatGPT değerlendirmelerinin alan uzmanı değerlendirmeleriyle benzer ve farklı yönlerinin neler olduğu incelenebilir. ChatGPT ve alan uzmanı görüşleri karşılaştırılarak yapay zekâ temelli geri bildirimin ne kadar geçerli ve güvenilir olduğu incelenebilir.

Çalışmanın STEM entegrasyonunun mühendislik bileşeni ile ilgili bulguları öğretmen adaylarının bu bileşeni fen öğrenimini destekleyecek biçimde mühendislik adımlarını bilimsel kavram ve ilkelerle ilişkilendirmesi gerektiğini ortaya koymaktadır. Bununla birlikte öğretmen adaylarının deney tasarımlarında farklı malzemeleri karşılaştırmaya, başarısız sonuçları tekrar değerlendirmeye ve tasarım iyileştirmeye yönelik etkinliklere daha fazla yer vermelerinin faydalı olacağını göstermektedir. ChatGPT gibi yapay zekâ destekli araçların hazır yanıt alma amacıyla kullanımı yerine var olan tasarımı sorgulama, geliştirme ve alternatifleri tartışma gibi amaçlarla kullanmaları önerilmektedir. Bu doğrultuda gelecek çalışmalarda öğretmen adaylarının STEM bileşenleri arasındaki entegrasyonu sağlama ile ilgili yaşadıkları zorlukların giderilmesine yönelik öğretimsel ve uygulama temelli çalışmalar tasarlanabilir.

Yapay zekâ geri bildirimlerinin bazen yüzeysel kalabileceği veya bireyselleşme/derinlik açısından sınırlı olabileceği göz önünde bulundurulmalı, bu araçlar tamamen öğretmenin yerini alan değil, süreci destekleyen işbirlikçi birer yardımcı olarak konumlandırılmalıdır.

Öğretmen eğitiminde yapay zekâ kullanımı, adayların yansıtıcı düşünme ve problem çözme gibi üst düzey bilişsel yeteneklerini desteklemek amacıyla günlük hayata ve öğretim süreçlerine entegre edilmelidir.

Etik ve Çıkar Çatışması

Bu çalışma, Dokuz Eylül Üniversitesi Eğitim Bilimleri Enstitüsü Matematik ve Fen Bilimleri Eğitimi Anabilim Dalı Fen Bilgisi Öğretmenliği yüksek lisans programında yürütülmekte olan “Yapay Zekâ Araçlarıyla Deney Tasarlayan Fen Bilimleri Öğretmen Adaylarının Göz Hareketlerinin ve Yapay Zekâya Yönelik Tutumlarının İncelenmesi” başlıklı yüksek lisans tezi doğrultusunda IV. Uluslararası Bilim, Eğitim, Sanat ve Teknoloji Sempozyumu’nda sunulan “Yapay Zekâ Araçlarıyla Deney Tasarlayan Fen Bilimleri Öğretmen Adaylarının ChatGPT ile Etkileşimlerinin İncelenmesi” başlıklı bildiriden üretilmiştir. Çalışmanın yürütülmesi sürecinde katılımcılardan gönüllü onam alınmış tüm etik ilke ve kurallara uyulmuştur. Yazarlar araştırmanın tüm süreçlerinde etik kurallara uygun davranıldığını ve yazarlar arasında herhangi bir çıkar çatışmasının olmadığını beyan eder.

KAYNAKÇA

- Aguinis, H., & Lawal, S. (2012). Conducting field experiments using eLancing's natural environment. *Journal of Business Venturing*, 27(4), 493–505. <https://doi.org/10.1016/j.jbusvent.2012.01.002>
- Ali, Z., Younis, S., Ahmad, N., Saba, F., & Ullah, N. (2025). Teachers' perspective of technology integration effects on students learning at university level. *Journal of Rehabilitation Research and Current Updates*, 1(1), 1–6.
- Alvarez, C., Alarcon, R., & Nussbaum, M. (2011). Implementing collaborative learning activities in the classroom supported by one-to-one mobile computing: A design-based process. *Journal of Systems and Software*, 84(11), 1961–1977. <https://doi.org/10.1016/j.jss.2011.07.011>
- An, Q., Yang, J., Xu, X., Zhang, Y., & Zhang, H. (2024). Decoding AI ethics from users' lens in education: A systematic review. *Heliyon*, 10(18), <https://doi.org/10.1016/j.heliyon.2024.e39357>
- Ateş, S. (2005). Öğretmen adaylarının değişkenleri belirleme ve kontrol etme yeteneklerinin geliştirilmesi. *Gazi Üniversitesi Gazi Eğitim Fakültesi Dergisi*, 25(1), 21–39.
- Brotherton, P. N., & Preece, P. F. W. (1995). Science process skills: Their nature and interrelationships. *Research in Science & Technological Education*, 13(1), 5–12. <https://doi.org/10.1080/0263514950130101>
- Cai, H., Han, B., Sun, J., Li, X., & Wong, L. H. (2025). Harnessing AI for teacher education to promote inclusive education: Investigating the effects of ChatGPT-supported lesson plan critiques on the development of pre-service teachers' lesson planning skills. *The Internet and Higher Education*, 64, <https://doi.org/10.1016/j.iheduc.2025.101022>
- Chellappa, V., & Luximon, Y. (2024). Understanding the perception of design students towards ChatGPT. *Computers and Education: Artificial Intelligence*, 7, <https://doi.org/10.1016/j.caeai.2024.100281>
- Çibik, N. F., & Boz-Yaman, B. (2025). The effect of a cross-curricular course on pre-service teachers' sustainable development attitudes and mathematical modeling self-efficacy beliefs. *International Journal of Science and Mathematics Education*, 23(4), 1033–1056. <https://doi.org/10.1007/s10763-024-10497-9>
- de Souza, B., Serrano de Andrade Neto, A., & Roazzi, A. (2024). The generative AI revolution, cognitive mediation networks theory and the emergence of a new mode of mental functioning: Introducing the Sophotechnic Mediation scale. *Computers in Human Behavior: Artificial Humans*, 2(1), <https://doi.org/10.1016/j.chbah.2024.100042>
- Dobrev, D. (2005). Formal definition of artificial intelligence. *International Journal "Information Theories & Applications"*, 12(3), 277–285.
- Dockendorff, M., & Zaccarelli, F. G. (2025). Successfully preparing future mathematics teachers for digital technology integration: A literature review. *International Journal of Mathematical Education in Science and Technology*, 56(5), 948–979. <https://doi.org/10.1080/0020739X.2024.2309273>
- Fijačko, N., Creber, R., Abella, B., Kocbek, P., Metličar, Š., Greif, R., & Štiglic, G. (2024). Using generative artificial intelligence in bibliometric analysis: 10 years of research trends from the European Resuscitation Congresses. *Resuscitation Plus*, 18, <https://doi.org/10.1016/j.resplu.2024.100584>
- Fischer, F., Sommerhoff, D., & Keune, A. (2023). Perspectives on learning from the learning sciences. In *International Encyclopedia of Education*. Elsevier. <https://doi.org/10.1016/B978-0-12-818630-5.14007-2>
- Fitzpatrick, M., & Leavy, A. (2025). Reciprocal interplays in becoming STEM learners and teachers: Preservice teachers' evolving understandings of integrated STEM education. *International Journal of Mathematical Education in Science and Technology*, 1–30. <https://doi.org/10.1080/0020739X.2025.2472260>
- Fraenkel, J. R., & Wallen, N. E. (1990). *How to design and evaluate research in education*. McGraw-Hill.
- Galindo-Domínguez, H., Delgado, N., Campo, L., & Losada, D. (2024). Relationship between teachers' digital competence and attitudes towards artificial intelligence in education. *International Journal of Educational Research*, 126, <https://doi.org/10.1016/j.ijer.2024.102381>
- García-Carmona, A., Muñoz-Franco, G., & Cruz-Guzmán, M. (2025). Integration of engineering practices into primary science classrooms: What does educational research tell us? *Science & Education*, 1–30. <https://doi.org/10.1007/s11191-025-00616-5>
- Gilbert, A., Suh, J., & Choudhry, F. (2025). Exploring the development of preservice teachers' visions of equity through science and mathematics integration. *International Journal of Science and Mathematics Education*, 23(2), 489–514. <https://doi.org/10.1007/s10763-024-10467-1>
- Hämäläinen, R., & Oksanen, K. (2012). Challenge of supporting vocational learning: Empowering collaboration in a scripted 3D game – How does teachers' real-time orchestration make a difference? *Computers & Education*, 59(3), 856–869. <https://doi.org/10.1016/j.compedu.2012.01.002>
- Hodson, D. (1988). Experiments in science and science teaching. *Educational Philosophy and Theory*, 20(2), 53–66. <https://doi.org/10.1111/j.1469-5812.1988.tb00144.x>

- Hofstein, A., & Lunetta, V. N. (1982). The role of the laboratory in science teaching: Neglected aspects of research. *Review of Educational Research*, 52(2), 201–217. <https://doi.org/10.3102/00346543052002201>
- Hovardas, T., Ter Vrugte, J., Zacharia, Z., & de Jong, T. (2023). Simulations and games for learning. In *International Encyclopedia of Education*. Elsevier. <https://doi.org/10.1016/B978-0-12-818630-5.14074-6>
- Huang, Y., Chen, P., Lee, H., Sandnes, F., & Wu, T. (2025). ChatGPT-enhanced mobile instant messaging in online learning: Effects on student outcomes and perceptions. *Computers in Human Behavior*, 168, <https://doi.org/10.1016/j.chb.2025.108659>
- Javaid, M., Haleem, A., Singh, R., Khan, S., & Khan, I. (2023). Unlocking the opportunities through ChatGPT tool towards ameliorating the education system. *BenchCouncil Transactions on Benchmarks, Standards and Evaluations*, 3(3), <https://doi.org/10.1016/j.tbench.2023.100115>
- Jayawardena, C., Gunathilake, Y., & Ihalagedara, D. (2025). Dental students' learning experience: Artificial intelligence vs human feedback on assignments. *International Dental Journal*, 75(1), <https://doi.org/10.1016/j.identj.2024.12.022>
- Jia, F., Sun, D., & Looi, C. K. (2024). Artificial intelligence in science education (2013–2023): Research trends in ten years. *Journal of Science Education and Technology*, 33(1), 94–117. <https://doi.org/10.1007/s10956-023-10077-6>
- Kewalramani, S., Devi, A., & Ng, A. (2025). Supporting Early Childhood Preservice Teachers to Effectively Integrate STEM in Their Future Teaching Practice. *Education Sciences*, 15(2). <https://doi.org/10.3390/educsci15020189>
- Kılıç, G. B., Yardımcı, E., & Metin, D. (2009). Fen öğretiminde değişkenler nasıl adlandırılabilir? *Abant İzzet Baysal Üniversitesi Eğitim Fakültesi Dergisi*, 9(2), 13–26.
- Kok, J. N., Boers, E. J., Kusters, W. A., Van der Putten, P., & Poel, M. (2009). Artificial intelligence: Definition, trends, techniques, and cases. In *Artificial Intelligence* (pp. 1-20). Eolss Publishers.
- Kovari, A. (2025). A systematic review of AI-powered collaborative learning in higher education: Trends and outcomes from the last decade. *Social Sciences & Humanities Open*, 11, <https://doi.org/10.1016/j.ssaho.2025.101335>
- Li, B., Lowell, V., Wang, C., & Li, X. (2024). A systematic review of the first year of publications on ChatGPT and language education: Examining research on ChatGPT's use in language learning and teaching. *Computers and Education: Artificial Intelligence*, 7, <https://doi.org/10.1016/j.caeai.2024.100266>
- Li, T., & Fu, Z. (2024). Formulating a descriptive framework and tagging system for design futures practice cases: Enabling heuristic tool development for design education and creative inspiration. *Thinking Skills and Creativity*, 53, <https://doi.org/10.1016/j.tsc.2024.101536>
- Liu, J., Sun, D., Sun, J., Wang, J., & Yu, P. (2025). Designing a generative AI enabled learning environment for mathematics word problem solving in primary schools: Learning performance, attitudes and interaction. *Computers and Education: Artificial Intelligence*, 9, <https://doi.org/10.1016/j.caeai.2025.100438>
- Madanchian, M., & Taherdoost, H. (2025). Decision-making criteria for AI tools in digital education. *Digital Engineering*, 7, <https://doi.org/10.1016/j.dte.2025.100069>
- Martin-Villalba, C., Urquia, A., & Dormido, S. (2012). Development of virtual-labs for education in chemical process control using Modelica. *Computers & Chemical Engineering*, 39, 169–177. <https://doi.org/10.1016/j.compchemeng.2011.10.010>
- Millar, R. (2004). *The role of practical work in the teaching and learning of science*. National Academy of Sciences.
- Milli Eğitim Bakanlığı. (2024). *Türkiye Yüzyılı Maarif Modeli: Fen bilimleri dersi öğretim programı (İlkokul ve ortaokul)*. Ankara: Milli Eğitim Bakanlığı.
- Ozturk, N., Aydın-Günbatır, S., & Roehrig, G. H. (2025). Elementary science teachers' engineering integration after long-term in-service training program with and without curriculum material support. *Research in Science & Technological Education*, 43(1), 170–190. <https://doi.org/10.1080/02635143.2023.2260996>
- Playfoot, D., Quigley, M., & Thomas, A. (2024). Hey ChatGPT, give me a title for a paper about degree apathy and student use of AI for assignment writing. *The Internet and Higher Education*, 62, <https://doi.org/10.1016/j.iheduc.2024.100950>
- Popenici, S. A., & Kerr, S. (2017). Exploring the impact of artificial intelligence on teaching and learning in higher education. *Research and Practice in Technology Enhanced Learning*, 12(1), <https://doi.org/10.1186/s41039-017-0062-8>
- Salas-Pilco, S. Z., Xiao, K., & Hu, X. (2022). Artificial intelligence and learning analytics in teacher education: A systematic review. *Education Sciences*, 12(8), <https://doi.org/10.3390/educsci12080569>
- Savin-Baden, M., & Major, C. H. (2023). *Qualitative research: The essential guide to theory and practice* (2nd ed.). Routledge. <https://doi.org/10.4324/9781003377986>

- Simmons, A. B., & Chappell, S. G. (1988). Artificial intelligence: Definition and practice. *IEEE Journal of Oceanic Engineering*, 13(2), 14–42. <https://doi.org/10.1109/48.551>
- Song, P., & Wang, C. (2023). Can ChatGPT replace scientists? *Science Bulletin*, 68(11), 1143–1146. <https://doi.org/10.1016/j.scib.2023.08.032>
- Suleiman, A., von Wedel, D., Munoz-Acuna, R., Redaelli, S., Santarisi, A., Seibold, E., ... & Schaefer, M. (2024). Assessing ChatGPT's ability to emulate human reviewers in scientific research: A descriptive and qualitative approach. *Computer Methods and Programs in Biomedicine*, 254, <https://doi.org/10.1016/j.cmpb.2024.108313>
- Wei, X., Wang, L., Koszalka, T., Lee, L., & Liu, R. (2025). Enhancing pre-service teachers' reflective thinking skills through generative AI-assisted digital storytelling creation: A three-dimensional framework analysis. *Computers & Education*, 235, <https://doi.org/10.1016/j.compedu.2025.105356>
- Yun, M., & Crippen, K. J. (2025). Computational thinking integration into pre-service science teacher education: A systematic review. *Journal of Science Teacher Education*, 36(2), 225–254. <https://doi.org/10.1080/1046560X.2024.2390758>
- Zeng, F., Gan, W., Wang, Y., & Yu, P. (2025). Distributed training of large language models: A survey. *Natural Language Processing Journal*, 12, <https://doi.org/10.1016/j.nlp.2025.100174>
- Zhang, Z., & Huang, X. (2024). The impact of chatbots based on large language models on second language vocabulary acquisition. *Heliyon*, 10(2), <https://doi.org/10.1016/j.heliyon.2024.e25370>
- Zhu, S., Li, Q., Yao, Y., Li, J., & Zhu, X. (2025). Improving writing feedback quality and self-efficacy of pre-service teachers in Gen-AI contexts: An experimental mixed-method design. *Assessing Writing*, 66, <https://doi.org/10.1016/j.asw.2025.100960>
- Zou, Y., Kuek, F., Feng, W., & Cheng, X. (2025). Digital learning in the 21st century: Trends, challenges, and innovations in technology integration. In *Frontiers in Education*, 10, Frontiers Media SA. <https://doi.org/10.3389/feduc.2025.1562391>

EXTENDED ABSTRACT

The aim of this study is to examine the interactions of pre-service science teachers with ChatGPT, an artificial intelligence-based conversational agent, during the process of designing experiments in various sub-disciplines of science, and to analyze ChatGPT's evaluations of these experiment designs by considering the instructional outcome content and grade level. In this context, the sub-problems of the study can be stated as: 'At which stages of the original experiment design process do pre-service science teachers ask questions to ChatGPT, and at which stages do they receive support to improve their experiment designs?' and 'What are ChatGPT's evaluations and development suggestions regarding the original experiment designs created by pre-service science teachers during the implementation process?' This study employed a case study design, one of the qualitative research methods that emphasizes in-depth investigation by focusing on a limited number of events or conditions and their relationships through detailed contextual analysis. Since the study required participants to design an original experiment based on a randomly assigned instructional outcome during the implementation process, the sample was selected using purposive sampling. Accordingly, the participants of the study consisted of pre-service science teachers who were enrolled in the course Science Teaching Laboratory Applications II during the 2024–2025 academic year at the Faculty of Education of a public university in Izmir. The data collection tools of the study included the Experiment Design with Artificial Intelligence Tools Form, the S-FeTeMM Form, and the textual records of the participants' interactions with ChatGPT. The Experiment Design with Artificial Intelligence Tools Form and the S-FeTeMM Form are digital forms designed using macros developed in Microsoft Excel to examine pre-service teachers' original experiment designs, the STEM integration of these designs, and the modifications made based on ChatGPT feedback. These forms were developed as open-ended tools to enable participants to express their views in depth. Participants' responses were recorded directly into spreadsheet cells, and the macros automatically organized, checked, and stored the data in a format suitable for analysis. The Experiment Design with Artificial Intelligence Tools Form consists of the following stages: designing a daily-life scenario, identifying the problem situation(s), determining the dependent variables of the experiment design, determining the independent variables, and identifying the controlled variables. The S-FeTeMM Form consists of the following stages: implementation steps, science integration, technology integration, engineering integration, and mathematics integration of the experiment design. The data were

collected qualitatively through online, written responses in a computer-based environment during individual sessions lasting approximately one hour with each participant throughout the spring semester of the 2024–2025 academic year. During the implementation sessions, pre-service science teachers created original experiment designs based on a randomly assigned instructional outcome from the 2024 Science Curriculum. In the second stage, they evaluated the alignment of their experiment designs with the instructional outcome by posing questions to ChatGPT and made revisions when deemed necessary. The version of ChatGPT used during the implementation process was ChatGPT-4o, which was the current version at the time of data collection and was accessed via a plus subscription. The data obtained through the data collection tools were analyzed using content analysis with the computer-assisted qualitative data analysis software MAXQDA 24. Following independent content analysis conducted by two researchers who are experts in science education using MAXQDA 24, inter-rater agreement was examined, and the content analysis was finalized in line with expert opinion. The inter-rater reliability coefficient between the two coders was determined to be 0.91. The findings revealed that pre-service science teachers frequently used ChatGPT for purposes such as consultation, verification, improvement, suggestion generation, and correction across various stages of experiment design. Analysis of ChatGPT's evaluations of the participants' designs indicated that the majority of controlled variable designs and science integration components were assessed as sufficient. However, ChatGPT evaluated the daily-life scenario, problem situation(s), implementation steps, technology integration, engineering integration, and mathematics integration components as partially sufficient and provided various development suggestions. In addition, the independent variable identification stage emerged as the most frequently evaluated as insufficient by ChatGPT. It was observed that participants directly relied on ChatGPT's suggestions without proposing their own ideas during the stages of variable identification. It was also observed that during the STEM integration design stages of the pre-service teachers' experiment designs, they consulted ChatGPT evaluations particularly in the steps related to ensuring technology and engineering integration, and that ChatGPT offered significant recommendations regarding these components of the designs. In this study, ChatGPT evaluations of experiment designs prepared by pre-service science teachers by considering learning outcomes were examined. In future studies, the similarities and differences between ChatGPT evaluations and subject-matter expert evaluations may be investigated. By comparing ChatGPT and expert opinions, the extent to which AI-based feedback is valid and reliable can be examined.