



Çocuk Resimlerinin Değerlendirilmesinde Yapay Zekâ ve Uzman Görüşlerinin Karşılaştırılması *

Comparison Of Artificial Intelligence and Expert Opinions In The Evaluation Of Children'S Drawings

ÖZET

Yapay zekâ teknolojilerinde son yıllarda yaşanan gelişmeler, eğitim ve ölçme-değerlendirme süreçlerinde yeni yaklaşımların ortaya çıkmasına olanak sağlamıştır. Özellikle görsel ve metinsel verileri analiz edebilen yapay zekâ tabanlı sistemler, insan performansının değerlendirilmesine yönelik alternatif ve tamamlayıcı araçlar olarak öne çıkmaktadır. Bu sistemler, büyük veri işleme kapasiteleri ve çok boyutlu analiz yetenekleri sayesinde geleneksel değerlendirme yöntemlerine kıyasla daha hızlı, tutarlı ve ölçeklenebilir sonuçlar sunabilmektedir. Bu doğrultuda, sanat eğitimi gibi özel değerlendirme boyutu yüksek alanlarda yapay zekâ destekli yaklaşımların kullanımı giderek önem kazanmaktadır. Sanat eğitiminde betimsel ifade, bireyin gözlemediği veya zihinsel süreçlerinde kurguladığı imgeleri görsel öğeler aracılığıyla anlamlı ve bütüncül bir yapı içerisinde sunabilme yetisini ifade etmektedir. İlkokul düzeyindeki resimsel ifadeler ise çocukların bilişsel ve duyuşsal gelişimlerinin oldukça önemli yansımasıdır. Bu tür görsel ürünlerin değerlendirilmesi geleneksel olarak uzmanlar tarafından gerçekleştirilmekte; ancak değerlendirme sürecinin öznel yapısı farklı yorumlara açık olabilmektedir. Bu noktada ChatGPT ve Gemini gibi sistemler, çözümlene kapasiteleriyle değerlendirme süreçlerine katkı sağlayacak potansiyel araçlar olarak öne çıkmaktadır. Araştırmanın amacı; ilkököl öğrencilerinin resimsel ifade düzeylerini yapay zekâ temelli sistemler ve uzman değerlendirmeleri çerçevesinde karşılaştırmalı incelemek, iki yaklaşım arasında istatistiksel farklılıkların bulunup bulunmadığını belirlemek ve yapay zekânın değerlendirme süreçlerine etkisini ortaya koymaktır. Betimsel araştırma modeli temelinde yürütülen çalışmanın grubunu amaçlı örneklemeyle belirlenen beş ilkököl öğrencisi oluşturmaktadır. Öğrenciler tarafından üretilen resimler, uzmanlar ve yapay zekâ tarafından analitik dereceli puanlama anahtarı kapsamında değerlendirilmiştir; “anlam/ifade, görsel dil tutarlılığı, kişisel üslup, kompozisyon-anlam uyumu, bilinçli seçim” ölçütleri esas alınmıştır. Elde edilen bulgular, karşılaştırmalı analiz yaklaşımı doğrultusunda çözümlenmiştir. Araştırma bulguları, yapay zekâ ve uzman değerlendirmeleri arasında belirli ölçütlerde benzerlikler bulunduğunu göstermiştir. ChatGPT'nin uzman değerlendirmelerine daha yakın ve ölçüt temelli sonuçlar sunduğu, Gemini'nin ise duygu, hayal gücü ve sembolik anlatım boyutlarında yüksek puanlama eğilimi gösterdiği belirlenmiştir. Bununla birlikte, yapay zekâ sistemlerinin çocuk resimlerinin gelişimsel ve pedagojik bağlamını yorumlamada uzman değerlendirmelerinin sunduğu derinliği tam sağlayamadığı sonucuna ulaşılmıştır. Bu doğrultuda yapay zekâ araçlarının uzman görüşlerini tamamlayıcı nitelikte olduğu düşünülmektedir.

Anahtar Kelimeler: Çocuk resimleri, sanat, sanat eğitimi, yapay zeka, sanatsal ifade

ABSTRACT

Recent developments in artificial intelligence technologies have enabled new approaches in education and evaluation processes. Systems analyzing visual and textual data emerge as alternative and complementary tools for evaluating human performance. Due to large data processing capacities and multidimensional analysis, these systems provide faster, consistent, and scalable results compared to traditional methods. Accordingly, artificial intelligence approaches are increasingly important in highly subjective assessment fields like art education. In art education, descriptive expression is an individual's ability to present mentally constructed or observed images within a holistic structure through visual elements. Primary school pictorial expressions significantly reflect children's cognitive and affective development. Traditionally, evaluating such visual products is done by experts; however, the subjective nature invites different interpretations. Here, systems like ChatGPT and Gemini stand out as potential tools contributing to evaluation processes through their analytical capacities. This research aims to comparatively examine primary school students' pictorial expression levels using artificial intelligence systems and expert evaluations, determine statistical differences between approaches, and reveal technology's impact on evaluation processes. Conducted using a descriptive research model, the study group features five primary school students selected via purposive sampling. Student produced drawings were evaluated by experts and artificial intelligence using an analytical rubric; meaning expression, visual coherence, personal style, compositional harmony, and informed choice criteria were the basis. Findings were analyzed using a comparative approach. Research findings indicated that there are similarities between artificial intelligence and expert assessments regarding specific criteria. It was determined that ChatGPT provided criteria based results closely aligned with expert assessments, while Gemini exhibited a tendency to score higher in emotion, imagination, and symbolic expression dimensions. However, it was concluded that artificial intelligence systems could not fully provide the depth offered by expert assessments when interpreting the developmental and pedagogical context of children's artwork. Accordingly, these new tools are considered complementary instruments supporting expert opinions.

Keywords: Children's drawings, art, art education, artificial intelligence, artistic expression

GİRİŞ

Sanat, insanlık tarihinin her döneminde var olan ve bireylerin duygu, düşünce ve deneyimlerini estetik formlarla dışa vurmasını sağlayan evrensel bir iletişim aracıdır (Gülpinar & Boyraz, 2024, s. 2). Sadece estetik bir haz sunmakla kalmayan sanat, aynı zamanda bireylerin kendini gerçekleştirilmesinde, toplumların kültürel mirasını

Huriye Çelikcan¹
Deniz Gül Tugen²

How to Cite This Article

Çelikcan, H. & Tugen, D. G. (2026). Çocuk Resimlerinin Değerlendirilmesinde Yapay Zekâ ve Uzman Görüşlerinin Karşılaştırılması. *International Social Sciences Studies Journal*, (e-ISSN:2587-1587) 12(6), 961-972. DOI: <https://doi.org/10.5281/zenodo.21068904>

Arrival: 25 May 2026
Published: 30 June 2026

Social Sciences Studies Journal is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

* Bu çalışma “XI International Academic Studies Congress (ASC-2026/SPRING)” da özet bildiri olarak sunulmuştur.

¹ Dr. Öğr. Üyesi, Ankara Müzik ve Güzel Sanatlar Üniversitesi, Müzik ve Güzel Sanatlar Eğitim Fakültesi, Görsel Sanatlar Eğitimi Bölümü, Ankara, Türkiye. ORCID: 0000-0002-8725-4251

² Yüksek Lisans Öğrencisi, Ankara Müzik ve Güzel Sanatlar Üniversitesi, Müzik ve Güzel Sanatlar Eğitim Fakültesi, Görsel Sanatlar Eğitimi Bölümü, Ankara, Türkiye. ORCID: 0009-0008-8125-8367

korumasında ve uygar, eleştirel düşünebilen bir yapı inşa etmesinde hayati bir öneme sahiptir (Akbaş, 2023, s. 7056). Bu anlamda sanat eğitimi, sadece sanatçı yetiştirmeyi değil toplumu sanat ve sanat eğitimi yoluyla eğiterek çok yönlü bireyler olmalarını sağlamaktadır. Özellikle eğitim ortamlarında sanat, çocuğu bağımsız düşünmeye, karar vermeye ve sosyal çevreyle uyumlu ilişkiler geliştirmeye yönlendirir.

Çocukluk döneminde sanat eğitimi, bireyin bilişsel, duyuşsal ve psikomotor gelişim alanlarını eş zamanlı olarak destekleyen bütüncül bir süreçtir (Tataroğlu, 2011, s. 123). Bilişsel alanda görsel okuryazarlık, problem çözme ve nesnelere arası ilişki kurma becerileri güçlenirken; duyuşsal alanda empati kurma, farklılıklara saygı duyma ve estetik duyarlılık kazanma gibi değerler ön plana çıkarmaktadır (Ulutaş vd., 2004, s. 2; Tataroğlu, 2011, s. 141). Küçük yaşlardan itibaren boya, kil veya çeşitli materyallerle yapılan uygulamalar ise el-göz koordinasyonu ve ince motor kaslarının gelişimi açısından psikomotor becerilere doğrudan katkı sağlamaktadır (Samsun & Güven, 2022, s. 1597).

Çocukların bilişsel ve fiziksel gelişimlerine paralel olarak, sanatsal ifade biçimleri ve resimleri de basitten karmaşığa doğru doğal bir evrim geçirir (Samsun & Güven, 2022, s. 1599). Motor kontrolünün sağlandığı ve 2-4 yaş arasında kapsayan mekân algısı olmaksızın gelişigüzel yapılan "karalama evresi" ile başlayan bu süreç; 4-7 yaşlarında nesnelere sembolik olarak yansıtıldığı "şema öncesi evre", 7-9 yaşlarında zemin çizgisi kavramının geliştiği "şematik dönem" ve sonrasında oran-orantı, boyut gibi unsurların arandığı "gerçekçilik/mantık aşaması" ile devam etmektedir (Samsun & Güven, 2022, s. 1601-1604). Bu süreçte ortaya çıkan çocuk resimleri, salt birer çizim olmaktan öte, onların gelişim düzeylerini, iç dünyalarını, psikolojik durumlarını ve zihinsel imgelerini yansıtan son derece güçlü, özgün iletişim kanallarıdır (Bolat & Adar Can, 2024, s. 689).

Çocukluk döneminde sanat eğitimi, bireyin iç dünyasını ve çevresini anlamlandırma sürecinde önemli bir ifade aracı hâline gelmektedir. Özellikle çocuk resimleri, bireyin zihinsel gelişimi, algısal örgütlenmesi ve duygusal yaşantıları hakkında önemli ipuçları sunmaktadır. Çocukların çizimlerinde kullandıkları biçimler, renkler, mekânsal ilişkiler ve figür düzenlemeleri gelişimsel süreçlerle doğrudan ilişkilidir. Viktor Lowenfeld çocuk resimlerini karalama dönemi, şema öncesi dönem, şematik dönem, gerçekçilik dönemi ve görünürde doğalcılık dönemi şeklinde sınıflandırarak çocukların görsel ifade süreçlerinin gelişimsel özelliklerini açıklamıştır (Lowenfeld & Brittain, 1987, s. 39). Benzer şekilde Jean Piaget çocukların bilişsel gelişim süreçlerinin sanatsal üretim biçimlerini doğrudan etkilediği vurgusunu yapmaktadır (Piaget, 1972, s. 27). Bu nedenle çocuk resimlerinin değerlendirilmesi yalnızca estetik ölçütlerle değil; aynı zamanda gelişimsel uygunluk çerçevesinde ele alınmalıdır.

Sanat eğitiminde ölçme ve değerlendirme süreçleri, öğrencilerin sanatsal gelişim düzeylerini belirlemek, yaratıcı ifade becerilerini incelemek ve öğrenme sürecini yönlendirmek açısından önemli bir işlev üstlenmektedir. Ancak sanat eğitiminin öznel yapısı, değerlendirme süreçlerinde standart ölçütlerin oluşturulmasını güçleştirmektedir. Özellikle yaratıcılık, özgünlük, duygusal ifade ve estetik bütünlük gibi kavramların değerlendirilmesinde uzman görüşleri önemli bir rol oynamaktadır. Sanat eğitiminde uzman değerlendirmeleri; teknik yeterlilik, kompozisyon, renk kullanımı, biçim çeşitliliği ve yaratıcı anlatım gibi ölçütler doğrultusunda gerçekleştirilmektedir (Hurwitz & Day, 2007, s. 112). Bununla birlikte uzman değerlendirmeleri bireysel yorum farklılıkları, deneyim düzeyi ve estetik bakış açısı gibi değişkenlerden etkilenmektedir. Bu durum değerlendirme süreçlerinde nesnellik ve tutarlılık tartışmalarını gündeme getirmektedir.

Eğitim sürecinde, öğrencilerin sanatsal becerilerindeki bu doğal gelişimi izlemek ve uygulanan öğretim programlarının etkililiğini saptamak amacıyla ölçme ve değerlendirme uygulamaları kritik bir rol oynar (Türk, Gençaslan & Kaya Uyanık, 2024, s. 482). Geleneksel uzman değerlendirmelerinde, uzmanlar eserleri nitel ve bağlamsal bir çerçevede inceleyerek derinlemesine analizler yapsalar da, değerlendirme sürecine sıklıkla kişisel yargılar, psikolojik durumlar ve kültürel arka planlar karışabilmektedir (Şen Atiker vd., 2023, s. 21). Bu durum, değerlendiriciler arası tutarsızlıklara, öğretmenler üzerinde ciddi bir iş yüküne ve nihayetinde sanat eğitiminde adil ve nesnel bir ölçme probleminin ortaya çıkmasına neden olmaktadır (Türk vd., 2024, s. 482).

Günümüz teknolojik gelişmeler ve dijital dönüşüm açısından bakıldığında ise eğitim paradigmasını yeniden şekillendirirken; insan zekâsını taklit edebilen, algılama, öğrenme ve problem çözme yeteneğine sahip yapay zekâ (YZ) sistemleri eğitim ortamlarının merkezine yerleşmeye başlamıştır (Eren vd., 2025, s. 518; Arslan, 2020, s. 71). Yapay zekâ, öğrencilere bireysel öğrenme yolları sunarak eğitimi kişiselleştirmekte, anında geri bildirimler sağlamak ve öğretmenlerin rutin idari, hazırlık veya değerlendirme iş yüklerini azaltarak onlara zaman kazandırmaktadır (Eren vd., 2025, s. 520; İşler & Kılıç, 2021, s. 4).

Sanat eğitimi ve ölçme-değerlendirme süreçleri bağlamında yapay zekânın sunduğu en önemli yeniliklerden biri görsel veri analizi ve bilgisayarlı görme (computer vision) kapasitesidir (Özdal, 2024, s. 283). YZ algoritmaları, sanat çalışmalarındaki çizgi, kompozisyon, oran, renk kullanımı ve desen gibi objektif unsurları yüksek doğruluk, hız ve tarafsızlıkla analiz edebilmektedir (Genç, Danış & Özalp Hamarta, 2023, s. 497-498). Özellikle sanatsal

çizimlerin incelenmesinde yapay zekâ, geleneksel değerlendirmelerin barındırdığı öznelliği en aza indirerek değerlendiricilere yardımcı bir mekanizma sağlamakta, nesnel değerlendirmenin önünü açmakta ve öğretmenlerin zamanını daha çok doğrudan pedagojik süreçlere ayırmasına olanak tanımaktadır (Eren vd., 2025, s. 520).

Geleneksel uzman değerlendirmeleri, eserlerin ruhsal, duygusal ve bağlamsal derinliğini yakalamada vazgeçilmezken; yapay zekâ sistemleri hız, tutarlılık ve nicel doğruluk açısından benzersiz fırsatlar sunmaktadır (Genç, Danış & Özalp Hamarta, 2023, s. 498). Ancak insan bilincinden kaynaklanan öznel yorum farklılıkları ile yapay zekânın salt veriye dayanan algoritmik tarafsızlığı arasındaki gerilim, sanat eğitiminde değerlendirme süreçlerinin nasıl daha tutarlı ve güvenilir hale getirilebileceğini araştırmanın temel problem durumu olarak ortaya koymaktadır (Şen Atiker vd., 2023, s. 21). Bu bağlamda, yapay zekâ ile uzman değerlendirmelerinin karşılaştırmalı olarak ele alınması, sanatsal üretimin analizinde hem nitel hem de nicel doğruluğu artıracak tamamlayıcı değerlendirme modellerinin geliştirilmesi için büyük bir önem taşımaktadır (Genç, Danış & Özalp Hamarta, 2023, s. 498). Bu anlamda Eğitimciler zaman kazandıracak ve geleceğin daha adil, nesnel ve yenilikçi ölçme-değerlendirme mekanizmalarına ışık tutacak kanıta dayalı bir perspektif sunması açısından (Eren vd., 2025, s. 520) yapılan çalışmada, sanat eğitiminde süregelen öznel değerlendirme sorunlarına karşı geleneksel uzman analizleri ile yapay zekâ destekli görsel analiz sistemlerinin sonuçlarını karşılaştırmalı olarak incelenmiştir.

Bu bağlamda çocuk resimlerinin değerlendirilmesinde yapay zekâ sistemleri ile uzman değerlendirmelerinin karşılaştırmalı olarak incelenmesi, sanat eğitiminde ölçme ve değerlendirme süreçlerinin geleceği açısından önemli görülmektedir. Yapay zekânın teknik analiz gücü ile uzmanların deneyim ve estetik yorum becerileri arasındaki benzerlik ve farklılıkların ortaya konulması, sanat eğitiminde daha bütüncül değerlendirme modellerinin geliştirilmesine katkı sağlama potansiyeli oluşturması açısından ayrıca önemlidir. Özellikle çocuk resimlerinin değerlendirilmesinde gelişimsel uygunluk, yaratıcı ifade ve teknik yeterlilik gibi ölçütlerin hem uzmanlar hem de yapay zekâ sistemleri tarafından nasıl yorumlandığının incelenmesi literatüre katkı sunacaktır. Bu doğrultuda bu çalışma, sanat eğitiminde yapay zekâ destekli değerlendirme sistemleri ile uzman değerlendirmeleri arasındaki ilişkiyi incelemeyi amaçlamakta; yapay zekânın sanat eğitimi bağlamındaki değerlendirme potansiyelini tartışmaya açmaktadır.

YÖNTEM

Araştırma Modeli

Bu araştırma, ilkökul öğrencilerinin resimsel ifade düzeylerini yapay zekâ temelli değerlendirmeler ile 2 görsel sanatlar alan uzmanının değerlendirmeleri çerçevesinde karşılaştırmalı olarak incelemek amacıyla betimsel araştırma modeli temelinde yürütülmüştür. Betimsel araştırmalar, geçmişte veya günümüzde var olan bir durumu, olguyu ya da problemi olduğu şekliyle, kendi koşulları içerisinde tanımlamayı ve betimlemeyi amaçlayan araştırma yaklaşımlarıdır (Karasar, 2012, s. 77). Araştırmada çocukların görsel ürünleri mevcut durumlarıyla ele alınmış ve değerlendirme sistemleri arasındaki ilişkiler tasvir edilmiştir. Ayrıca, alan uzmanlarının geleneksel ve nitel değerlendirmeleri ile yapay zekâ sistemlerinin çok boyutlu veri işleme kapasitelerinden elde edilen sonuçların istatistiksel ve yorumsal olarak karşılaştırılması nedeniyle çalışma, karma yöntem araştırmasına uygun olarak yürütülmüştür. Karma yöntem araştırmaları, nitel ve nicel veri toplama, analiz ve yorumlama yaklaşımlarının tek bir çalışmada entegre edilerek araştırma problemine daha derinlemesine yanıt bulunmasını sağlayan tasarımlardır (Büyüköztürk vd., 2020, s. 256).

Çalışma Grubu

Araştırmanın çalışma grubunu amaçlı örnekleme yöntemi doğrultusunda belirlenen beş ilkökul öğrencisi oluşturmaktadır. Amaçlı örnekleme yöntemi, araştırmanın amacına uygun olarak bilgi bakımından zengin durumların seçilmesini esas alan bir örnekleme yaklaşımıdır (Büyüköztürk vd., 2024, s. 92). Bu doğrultuda çalışma grubuna dahil edilen 7-8 yaş şematik evrede bulunan 5 öğrencinin; görsel anlatım çeşitliliği, yaratıcı ifade biçimleri ve sanatsal kompozisyon özellikleri dikkate alınarak belirlenmiştir.

Araştırmada sınırlı sayıda örneklem kullanılmasının temel nedeni, her bir öğrenci çalışmasının ayrıntılı biçimde incelenmesine olanak sağlamaktır. Sanat eğitimi ve görsel ifade araştırmalarında değerlendirme süreçlerinin çok boyutlu ve öznel yapısı nedeniyle, görece küçük bir örneklem grubu üzerinden çalışmak, elde edilen görsel eserlerin her bir ölçüt bağlamında titizlikle, mikro düzeyde ve derinlemesine analiz edilmesine olanak sağlamaktadır. Bu nedenle küçük örneklem grubu üzerinden ayrıntılı çözümlenmesi araştırmanın yorumlayıcı niteliğini güçlendirmiştir.

Veri Toplama Araçları

Araştırmada veri toplama aracı olarak analitik dereceli puanlama anahtarı düzenlenmiştir. Analitik dereceli puanlama anahtarları, performansın belirli ölçütler doğrultusunda ayrı ayrı değerlendirilmesini sağlayan sistematik ölçme araçlarıdır (Büyüköztürk vd., 2024, s. 181). Bu doğrultuda öğrenci resimleri; “anlam/ifade”, “görsel dil tutarlılığı”, “kişisel üslup”, “kompozisyon–anlam uyumu” ve “bilinçli seçim” ölçütleri temel alınarak değerlendirilmiştir.

“Anlam/ifade” ölçütü, öğrencinin duygu ve düşüncelerini görsel yollarla aktarabilme düzeyini incelemektedir. “Görsel dil tutarlılığı”, kompozisyon içerisindeki biçimsel unsurların bütünlük oluşturma düzeyini değerlendirmektedir. “Kişisel üslup” ölçütü, öğrencinin özgün anlatım biçimlerini ve bireysel sanatsal yaklaşımını ortaya koymayı amaçlamaktadır. “Kompozisyon–anlam uyumu” ölçütü, görsel düzenleme ile verilmek istenen mesaj arasındaki ilişkiyi incelerken; “bilinçli seçim” ölçütü ise renk, çizgi, biçim ve mekân kullanımına ilişkin tercihlerdeki farkındalık düzeyini değerlendirmektedir.

Araştırma değerlendirme sürecine iki farklı yapay zekâ sistemi dahil edilmiştir. Bunlar ChatGPT ve Gemini sistemleridir. Öğrencilere ait resimler dijital ortama aktarılarak yapay zekâ sistemlerine yüklenmiş ve her iki sisteme de aynı değerlendirme yönergeleri verilmiştir. Böylece değerlendirme sürecinde standartlığın korunması amaçlanmıştır.

Uzman değerlendirme sürecinde ise görsel sanatlar eğitimi alanında uzman 1 akademisyen doktora derecesine sahip bir Öğretmen görüşlerine başvurulmuştur. Uzmanlar, öğrenci çalışmalarını birbirlerinden bağımsız biçimde incelemiş ve aynı analitik dereceli puanlama anahtarı doğrultusunda değerlendirme gerçekleştirmiştir.

Veri Toplama Süreci

Araştırma sürecinde öğrencilere belirlenen “Ailemle Bir Gün” teması doğrultusunda resim uygulaması yaptırılmıştır. Öğrencilerin çalışmalarını özgür biçimde oluşturabilmeleri amacıyla uygulama sürecinde yönlendirici müdahale de bulunulmamıştır. Tamamlanan çalışmalar dijital ortama aktarılmış ve değerlendirme sürecine uygun görsel veriler hâline getirilmiştir.

Dijitalleştirilen öğrenci resimleri öncelikle alan uzmanlarına iletilmiş; uzmanlardan çalışmaları analitik dereceli puanlama anahtarı doğrultusunda değerlendirmeleri istenmiştir. Aynı görseller, aynı yönergeler kullanılarak yapay zekâ sistemlerine yüklenmiştir. Yapay zekâ değerlendirmelerinde standart komutların kullanılması, sistemler arası değerlendirme tutarlılığının korunması açısından önemli görülmüştür.

Araştırma sürecinde elde edilen tüm değerlendirme verileri kayıt altına alınmış ve uzman değerlendirmeleri ile yapay zekâ sistemlerinden elde edilen veriler karşılaştırmalı analiz için düzenlenmiştir.

Verilerin Analizi

Elde edilen veriler, alan uzmanlarının ve yapay zekâ sistemlerinin değerlendirmelerini merkeze alan karşılaştırmalı analiz yaklaşımı doğrultusunda çözümlenmiştir. Her iki kaynaktan sağlanan analiz sonuçları ile verilen ortalama puanlar karşılaştırılarak, aralarında istatistiksel açıdan anlamlı bir farklılık bulunup bulunmadığı test edilmiştir. Analiz süreci yalnızca sayısal verilerle sınırlandırılmamış; anlam/ifade, görsel dil tutarlılığı, kişisel üslup gibi ölçüt bazı çözümlenmeler ile derinleştirilmiştir. Geleneksel uzman değerlendirmelerinin öznel yorumları ile yapay zekâ algoritmalarının veri işleme mantığı arasındaki yorum farklılıkları ve puanlama eğilimleri saptanarak, yapay zekânın eğitimde tamamlayıcı bir araç olup olamayacağı akademik bir çerçevede tartışılmıştır.

BULGULAR

Görsellerin uzman, Chatgpt ve Gemini değerlendirmelerine ilişkin karşılaştırmalı analizine bakıldığında;

1. Görselin: Uzman, ChatGPT ve Gemini Değerlendirmelerine İlişkin Karşılaştırmalı Analizi

Görsel: 1



Birinci görselde uzman değerlendirmeleri dereceli puanlama anahtarına göre çalışmayı yeterli ve iyi aralığında ele alırken; ChatGPT ve Gemini’de uzmanlarla benzer şekilde çalışmayı iyi düzeyinde puanlamışlardır. Ayrıca her iki yapay zeka değerlendirici çalışmanın özellikle “konuya uygunluk, duygu aktarımı ve sembolik anlatım” bakımından güçlü olduğunu belirtmektedir. ChatGPT “Aile bireylerinin bir arada verilmesi, kalp sembolleri, gülen güneş ve “Ailemi çok seviyorum” ifadeleriyle öğrencinin sevgi temasını açık biçimde yansıttığını göstermektedir. Bununla birlikte birinci uzman “hareket etkisinin sınırlı kalması ve kompozisyonun klasik aile resmi şemasına yakın olmasının” çalışmanın özgünlük açısından orta düzeyde değerlendirilebileceğini belirtmektedir. Genel olarak uzman ve yapay zeka değerlendiricileri Görsel 1’in, 7 yaş çocuğu şematik dönem özellikleriyle uyumlu, duygusal anlatımın belirgin olduğunu belirtirken, özgünlük açısından geliştirilebilir bir çalışma niteliğinde olduğunu ifade etmişlerdir.

2. Görselin: Uzman, ChatGPT ve Gemini Değerlendirmelerine İlişkin Karşılaştırmalı Analizi

Görsel: 2



İkinci görsel için her iki uzman dereceli puanlama anahtarına göre; çalışmanın iyi düzeyde olduğunu, ChatGPT ise uzmanlardan ve Gemini’den daha düşük bir puanlama yaparak yeterli düzeyde olduğu şekilde daha sınırlı bir puanlama yapmıştır. Gemini ise uzmanlarla aynı puanlama aralığında yer almış ve çalışmanın iyi düzeyde olduğunu belirtmiş ve özellikle “anlatı gücü ve aile içi etkileşim” açısından diğer değerlendiricilere göre daha olumlu bir değerlendirme sunmuştur. ChatGPT ve Gemini “ev içi sahne, konuşma balonları, oyun/televizyon unsurları ve figürlerin jestleri” kullanarak çocuğun “günlük yaşamdan bir kesiti” ifade etme çabası içerisinde olduğunu ifade etmektedir. Tüm değerlendiriciler açısından ise görsel renk kullanımı ve teknik uygulama açısından sınırlı bulunmakla birlikte, buna karşılık çalışmanın konuya uygunluğu ve bireysel anlatım dilinin belirgin olduğunu belirtmişlerdir. Genel olarak Görsel 2, teknik ve yüzey düzeni bakımından geliştirilmeye açık; ancak duygu, iletişim ve günlük yaşam anlatısı bakımından yaş düzeyine uygun bir sanatsal ifade örneği olarak değerlendirilebilir.

3. Görselin Uzman, ChatGPT ve Gemini Değerlendirmelerine İlişkin Karşılaştırmalı Analizi

Görsel: 3



Üçüncü görselde uzmanlar çalışmayı dereceli puanlama anahtarına göre; iyi ve çok iyi aralığında konumlandırırken, ChatGPT özellikle “renk kullanımı, yüzey düzeni, konuya uygunluk ve hayal gücü” bakımından oldukça iyi bulunduğunu belirterek, tüm görseller içerisinde değerlendirici puanları açısından verilen en yüksek puanı vermiştir. Gemini ise “çalışmayı yaş düzeyine uygun ve başarılı” bulmakla birlikte “teknik kullanımda daha sınırlı” olduğunu belirtmiş ve çalışmayı iyi düzeyinde puanlamıştır. ChatGPT “park/kaydırak sahnesi, aile bireylerinin birlikte oyun ortamında sunulması ve canlı renklerin cesur kullanımı, çalışmanın anlatım gücünü artırdığını ifade etmiştir. Uzmanlar ise “hareket etkisinin kaydırak ve oyun alanı üzerinden kurulduğunu ve yüzey etkisinin dengeli biçimde” kullanıldığını belirtmişlerdir. Genel olarak Görsel 3, çocuğun aile temasını dinamik bir oyun mekânı içinde başarılı biçimde yorumladığını; renk, kompozisyon ve hayal gücü bakımından güçlü bir ifade ortaya koyduğunu göstermektedir.

4. Görselin: Uzman, ChatGPT ve Gemini Değerlendirmelerine İlişkin Karşılaştırmalı Analizi

Görsel: 4



Dördüncü görselde ise uzman ve ChatGPT değerlendirmeleri çalışmayı daha çok yeterli düzeyde ele alırken, Gemini değerlendirmesinde; diğer tüm görsellerde de olduğu gibi dördüncü görselde de yüksek düzeyde bir puanlama yaparak çok iyi olduğunu ayrıca “renk, yüzey kullanımı, biçim çeşitliliği ve konuya uygunluk” bakımından oldukça iyi olduğunu ifade etmiştir. Uzman ve yapay zeka değerlendiricileri açısından çocuğun duygu ve düşüncesini ifade etmesi bakımından dördüncü görsel; “park teması, aile figürleri, oyun alanı ve çevresel nesnelere” aracılığıyla desteklenmiştir; ancak “hareket duygusu ve duygu aktarımı açısından diğer görsellere göre daha sınırlı kalmıştır” biçiminde ifade edilmiştir. Buna karşın nesne çeşitliliği, düzen anlayışı ve mekân kurma çabası çalışmanın güçlü yönleri arasında yer almaktadır. Görsel, teknik ve biçimsel açıdan yaş düzeyine uygun görünmekle birlikte, figürler arası ilişki, hareket ve yüzey algısı bakımından daha belirgin bir anlatıma ihtiyaç duyulmaktadır.

5. Görselin: Uzman, ChatGPT ve Gemini Değerlendirmelerine İlişkin Karşılaştırmalı Analizi

Görsel: 5



Beşinci görselde her iki uzmanda çalışmayı çok iyi düzeyde değerlendirirken, ChatGPT daha ölçülü bir puanlama yaparak iyi düzeyde olduğunu belirtmiştir. Gemini ise özellikle “hayal gücü, duygu aktarımı ve özgünlük” açısından yüksek bir değerlendirme sunarak uzmanlarla benzer biçimde çalışmanın çok iyi düzeyde olduğunu belirtmiştir. Çalışmada aile bireylerinin oyun/eğlence ortamında yer alması, figürlerin gülümseyen ifadeleri ve mekânın yaratıcı biçimde kurgulanması olumlu bir aile atmosferi oluşturmaktadır. Renk kullanımı canlı, biçimler zengin ve kompozisyon kişisel anlatıma açıktır. Bu anlamda ChatGPT “teknik kontrol ve yüzey dengesinde bazı sınırlılıklara” dikkat çekerken, uzman ve Gemini değerlendirmeleri “çalışmanın yaşlılarının üzerinde yaratıcı potansiyel” taşıdığını belirtmektedir. Genel olarak beşinci görsel, konuya uygunluğu, özgün mekân kurgusu ve güçlü hayal gücüyle öne çıkan en başarılı çalışma olarak değerlendirilebilir.

Karşılaştırmalı Bulgular

Bu bölümde, “Ailele Bir Gün” teması doğrultusunda üretilen beş öğrenci resmi; iki uzmanın değerlendirmeleri, ChatGPT ve Gemini değerlendirmeleri açısından karşılaştırmalı olarak incelenmiştir. Karşılaştırmalı analizde öncelikle iki uzman puanının ortalaması alınmış, ardından bu ortalama puanlar ChatGPT ve Gemini tarafından verilen puanlarla karşılaştırılmıştır. Böylece yapay zekâ sistemlerinin uzman değerlendirmelerine ne ölçüde yaklaştığı, hangi ölçütlerde daha yüksek ya da düşük puanlama eğilimi gösterdiği ve ölçüt bazında benzerlik oranlarının nasıl dağıldığı ortaya konulmuştur.

Araştırma bulguları genel olarak değerlendirildiğinde, yapay zekâ sistemlerinin çocuk resimlerini analiz etme sürecinde uzman değerlendirmeleriyle belirli ölçütlerde örtüşen sonuçlar üretebildiği; ancak özellikle hareket, özgünlük, hayal gücü ve duygu aktarımı gibi yoruma açık alanlarda farklı puanlama eğilimleri sergilediği görülmektedir. Uzman değerlendirmeleri daha çok gelişimsel uygunluk, pedagojik bağlam ve teknik yeterlilik çerçevesinde şekillenirken; ChatGPT’nin yapısal düzen, teknik kontrol ve yüzey organizasyonuna; Gemini’nin ise sembolik anlatım, duygusal atmosfer ve yaratıcı göstergelere daha fazla ağırlık verdiği belirlenmiştir.

Tablo 1: Uzman Ortalaması, ChatGPT ve Gemini Toplam Puan Karşılaştırması

Görsel	1. Uzman	2. Uzman	Uzman Ortalama Puanı	ChatGPT	Gemini	ChatGPT-Uzman Farkı	Gemini-Uzman Farkı
1. Görsel	24	22	23,0	32	28	+9,0	+5,0
2. Görsel	25	27	26,0	23	31	-3,0	+5,0
3. Görsel	28	24	26,0	26	31	0,0	+5,0
4. Görsel	25	31	28,0	23	28	-5,0	0,0
5. Görsel	32	31	31,5	29	28	-2,5	-3,5

Tablo 1 incelendiğinde, ChatGPT’nin birinci görselde uzman ortalamasına göre oldukça yüksek puan verdiği görülmektedir. Bu durum, ChatGPT’nin özellikle renk kullanımı, duygu aktarımı, yüzey kullanımı ve hayal gücü gibi ölçütlerde daha olumlu bir değerlendirme eğilimi gösterdiğini düşündürmektedir. Buna karşılık ikinci, dördüncü ve beşinci görsellerde ChatGPT puanlarının uzman ortalamasının altında kaldığı görülmektedir. Üçüncü görselde ise ChatGPT puanı ile uzman ortalaması tamamen örtüşmektedir. Bu bulgu, ChatGPT’nin bazı çalışmalarda uzman değerlendirmesine oldukça yakın sonuçlar üretebildiğini; ancak bazı görsellerde teknik ve yapısal ölçütleri daha sınırlayıcı biçimde yorumlayabildiğini göstermektedir.

Gemini değerlendirmeleri incelendiğinde ise ilk üç görselde uzman ortalamasından daha yüksek puanlar verdiği, dördüncü görselde uzman ortalamasıyla aynı düzeyde kaldığı, beşinci görselde ise uzman ortalamasının altında puanlama yaptığı görülmektedir. Gemini'nin özellikle ikinci ve üçüncü görsellerde yaratıcı anlatım, sembolik detaylar ve kompozisyon bütünlüğünü daha güçlü yorumladığı anlaşılmaktadır. Bu durum, Gemini'nin görsel anlatı içerisindeki sembolik, duygusal ve yaratıcı göstergeleri uzmanlara kıyasla daha yüksek değerlendirme eğiliminde olduğunu göstermektedir.

Tablo 2: Değerlendiricilere Göre Genel Ortalama Puanlar

Değerlendirici	Toplam Puan	Genel Ortalama
1. Uzman	134	2,98
2. Uzman	135	3,00
ChatGPT	133	2,96
Gemini	146	3,24

Tablo 2'ye göre genel toplam puanlar açısından 1. uzman, 2. uzman ve ChatGPT değerlendirmeleri birbirine oldukça yakın görünmektedir. 1. uzman toplamda 134 puan, 2. uzman 135 puan, ChatGPT ise 133 puan vermiştir. Bu sonuca göre ChatGPT'nin genel puanlama eğiliminin uzman değerlendirmelerine yakın olduğunu söylemek mümkündür. Buna karşılık Gemini toplamda 146 puanla diğer değerlendiricilerden daha yüksek bir puanlama eğilimi sergilemiştir. Bu bulgu, Gemini'nin çocuk resimlerinde yer alan yaratıcı, sembolik ve anlatsal öğeleri daha olumlu yorumlama eğiliminde olduğunu düşündürmektedir.

Uzmanlar ile ChatGPT arasındaki yakınlık, yapay zekâ destekli değerlendirme sistemlerinin belirli ölçütlerde uzman değerlendirmesine yaklaşabileceğini göstermesi bakımından önemlidir. Ancak Gemini'nin daha yüksek puanlama eğilimi, yapay zekâ sistemlerinin aynı görsel veriyi farklı algoritmik önceliklerle yorumlayabildiğini ortaya koymaktadır. Dolayısıyla yapay zekâ değerlendirmeleri tek başına kesin sonuç üretici bir araç olarak değil, uzman görüşlerini destekleyen tamamlayıcı bir değerlendirme mekanizması olarak ele alınmalıdır.

Tablo 3: Ölçüt Bazında Uzman Ortalaması, ChatGPT ve Gemini Puan Ortalamaları

Ölçüt	Uzman Ortalaması	ChatGPT Ortalaması	Gemini Ortalaması	ChatGPT Farkı	Gemini Farkı
Bilgi Objelerinde Hareket	3,00	2,40	2,60	-0,60	-0,40
Renk Kullanımı	2,90	3,00	3,00	+0,10	+0,10
Duygu-Düşünce İfadesi	2,80	3,20	3,80	+0,40	+1,00
Tekniğin Kullanımı	3,00	2,40	3,00	-0,60	0,00
Konuya Uygunluk	3,50	3,80	4,00	+0,30	+0,50
Yüzey Kullanımı	2,80	3,00	3,20	+0,20	+0,40
Biçim Çeşitliliği	3,10	2,80	3,20	-0,30	+0,10
Hayal Gücü	2,90	3,20	3,20	+0,30	+0,30
Özgünlük	2,90	2,80	3,20	-0,10	+0,30

Tablo 3'te yer alan ölçüt bazlı ortalamalar, değerlendirme yaklaşımları arasındaki temel farklılıkları daha görünür hâle getirmektedir. "Konuya uygunluk" ölçütü, tüm değerlendiriciler tarafından en yüksek puanlanan alanlardan biri olmuştur. Bu durum, öğrencilerin "Ailele Bir Gün" temasını görsel anlatılarına açık biçimde yansıtabildiklerini göstermektedir. Aile bireylerinin bir arada sunulması, oyun, park, ev içi etkinlikler, kalp sembolleri ve gülümseyen yüz ifadeleri hem uzmanlar hem de yapay zekâ sistemleri tarafından tema ile ilişkili göstergeler olarak değerlendirilmiştir.

"Duygu-düşünce ifadesi" ölçütünde Gemini'nin uzman ortalamasına göre belirgin biçimde daha yüksek puan verdiği görülmektedir. Bu bulgu, Gemini'nin çocuk resimlerinde yer alan sembolik ve duygusal göstergeleri daha geniş bir anlam alanı içerisinde yorumladığını göstermektedir. Bu anlamda özellikle görsellerde yer alan kalp figürleri, aile bireylerinin yakın konumlandırılması, konuşma balonları, oyun sahneleri ve canlı renk tercihleri Gemini tarafından güçlü duygusal aktarım göstergeleri olarak değerlendirilmiştir.

ChatGPT'nin ise "bilgi objelerinde hareket" ve "teknik kullanımı" ölçütlerinde uzman ortalamasının altında kaldığı görülmektedir. Bu durum, ChatGPT'nin hareket algısını daha çok figürlerin beden pozisyonları, yönelimleri ve eylemsel ilişkileri üzerinden değerlendirdiğini düşündürmektedir. Benzer şekilde teknik kullanım ölçütünde de boyama taşmaları, çizgi kontrolü, yüzey düzeni ve biçimsel bütünlük gibi unsurlara daha eleştirel yaklaştığı söylenebilir.

Tablo 4: Uzman–ChatGPT ve Uzman–Gemini Benzerlik Oranları

Bu tabloda benzerlik oranı, yapay zekâ sisteminin ilgili ölçütte verdiği puanın en az bir uzman puanıyla aynı olması esas alınarak hesaplanmıştır.

Ölçüt	Uzman-ChatGPT Benzerlik Oranı	Uzman-Gemini Benzerlik Oranı
Bilgi Objelerinde Hareket	%40	%0
Renk Kullanımı	%60	%60
Duygu–Düşünce İfadesi	%40	%20
Tekniğin Kullanımı	%40	%80
Konuya Uygunluk	%80	%80
Yüzey Kullanımı	%80	%80
Biçim Çeşitliliği	%40	%40
Hayal Gücü	%80	%80
Özgünlük	%100	%80
Genel Benzerlik Oranı	%62,22	%57,78

Tablo 4 incelendiğinde, ChatGPT'nin uzman puanlarıyla genel benzerlik oranının %62,22; Gemini'nin ise %57,78 olduğu görülmektedir. Bu sonuç, ChatGPT'nin genel puan uyumu bakımından uzman değerlendirmelerine bir miktar daha yakın olduğunu göstermektedir. Ancak Gemini'nin özellikle “teknik kullanımı”, “konuya uygunluk”, “yüzey kullanımı”, “hayal gücü” ve “özgünlük” ölçütlerinde yüksek benzerlik oranları yakaladığı görülmektedir.

ChatGPT'nin “özgünlük” ölçütünde %100 benzerlik göstermesi dikkat çekicidir. Bu bulgu, ChatGPT'nin öğrencilerin bireysel anlatım özelliklerini uzmanlarla benzer biçimde yorumlayabildiğini düşündürmektedir. Bununla birlikte “bilgi objelerinde hareket”, “duygu–düşünce ifadesi” ve “biçim çeşitliliği” ölçütlerinde benzerlik oranının daha düşük olması, ChatGPT'nin bazı görsel göstergeleri uzmanlardan farklı biçimde sınıflandırdığını ortaya koymaktadır.

Gemini açısından en düşük benzerlik “bilgi objelerinde hareket” ölçütünde ortaya çıkmıştır. Bu durum, Gemini'nin hareket göstergelerini uzmanlardan farklı biçimde yorumladığını göstermektedir. Gemini bazı görsellerde hareketi doğrudan figürlerin fiziksel eylemleriyle değil; kompozisyon içerisindeki sembolik yönelim, çizgisel akış ve anlatsal dinamizmle ilişkilendirmiştir. Bu nedenle uzmanlarla birebir puan benzerliği düşük görünse de nitel yorum düzeyinde Gemini'nin daha geniş bir anlatı çözümlenmesi yapılabileceği söylenebilir.

Tablo 5: Ortalama Puan Farklılıklarının Genel Yorumu

Karşılaştırma	Genel Eğilim	Yorum
Uzman – ChatGPT	Genel olarak yakın	ChatGPT toplam puan bakımından uzmanlara oldukça yakın sonuçlar üretmiştir. Ancak bazı görsellerde teknik ve hareket ölçütlerinde daha eleştirel davranmıştır.
Uzman – Gemini	Kısmen daha yüksek puanlama	Gemini özellikle duygu, sembolik anlatım, konuya uygunluk ve hayal gücü ölçütlerinde daha olumlu değerlendirme yapmıştır.
ChatGPT – Gemini	Yaklaşım farkı belirgin	ChatGPT daha analitik ve yapısal; Gemini daha yorumlayıcı ve sembolik bir değerlendirme eğilimi göstermiştir.
Uzmanlar arası değerlendirme	Büyük ölçüde yakın, bazı görsellerde farklı	Uzmanlar genel olarak benzer puanlama yapmış; ancak özellikle 3. ve 4. görsellerde yorum farkları ortaya çıkmıştır.

Tablo 5'te görüldüğü üzere, değerlendirme yaklaşımları arasındaki farklar yalnızca nicel puan düzeyinde değil, yorumlama biçiminde de ortaya çıkmaktadır. Uzmanlar, çocuk resimlerini gelişimsel dönem özellikleri, teknik yeterlilik ve pedagojik bağlam içinde değerlendirirken; yapay zekâ sistemleri görsel verilerdeki sembolik, biçimsel ve anlatsal göstergeleri farklı yoğunluklarda öne çıkarmıştır. Bu bağlamda ChatGPT'nin daha ölçülü, analitik ve teknik odaklı; Gemini'nin ise daha geniş yorumlayıcı, sembolik ve duygusal anlatı odaklı bir yaklaşım geliştirdiği söylenebilir.

Genel olarak bulgular, çocuk resimlerinin değerlendirilmesinde yapay zekâ sistemlerinin uzman değerlendirmeleriyle belirli ölçütlerde anlamlı bir yakınlık gösterebildiğini ortaya koymaktadır. Özellikle konuya uygunluk, yüzey kullanımı, hayal gücü ve özgünlük ölçütlerinde yüksek benzerlik oranlarının elde edilmesi, yapay zekânın sanat eğitimi alanında ön değerlendirme, rubrik temelli puanlama ve geri bildirim oluşturma süreçlerinde destekleyici bir araç olarak kullanılabileceğini göstermektedir. Bununla birlikte hareket, duygusal ifade ve sembolik anlamlandırma gibi daha öznel ölçütlerde yapay zekâ sistemleri ile uzmanlar arasında farklılıklar ortaya çıkması, insan uzmanlığının değerlendirme sürecindeki önemini koruduğunu göstermektedir.

Araştırma sonuçları kapsamında, karşılaştırmalı bulgular, yapay zekâ destekli değerlendirme sistemlerinin çocuk resimlerini analiz etmede işlevsel bir potansiyele sahip olduğunu; ancak bu sistemlerin tek başına nihai değerlendirici olarak değil, uzman görüşlerini destekleyen tamamlayıcı araçlar olarak kullanılmasının daha uygun olacağını ortaya koymaktadır.

SONUÇ VE TARTIŞMA

Araştırmanın “Ailele Bir Gün” teması doğrultusunda 7 yaş grubundaki beş öğrencinin resimleri; iki alan uzmanı, ChatGPT ve Gemini yapay zekâ sistemleri tarafından değerlendirilmiş ve elde edilen sonuçlar karşılaştırmalı olarak incelenmiştir. Araştırma bulguları, yapay zekâ destekli değerlendirme sistemlerinin çocuk resimlerini analiz etme sürecinde uzman değerlendirmelerine belirli ölçütlerde yakın sonuçlar üretebildiğini göstermektedir. Özellikle konuya uygunluk, yüzey kullanımı, hayal gücü ve özgünlük ölçütlerinde uzmanlar ile yapay zekâ sistemleri arasında dikkat çekici benzerlikler görülmüştür. Bu durum, yapay zekânın sanat eğitimi alanında ölçme-değerlendirme süreçlerine destek sağlayabilecek potansiyele sahip olduğunu ortaya koymaktadır. Benzer biçimde Wang vd. (2025) çocuk resimlerinin çocuk resimlerinin değerlendirilmesinde yapay zekâ tabanlı sistemlerin estetik kalite ve teknik beceri düzeylerinde tutarlı sonuçlar verdiğini belirtmektedir.

Araştırmada, ChatGPT’nin genel puanlama eğiliminin uzman değerlendirmelerine daha yakın olduğu belirlenmiştir. Uzmanların toplam puanları ile ChatGPT’nin toplam puanı birbirine oldukça yakın bulunmuştur. Bu sonuç, ChatGPT’nin özellikle rubrik temelli değerlendirme süreçlerinde daha ölçülü, analitik ve yapısal bir yaklaşım sergilediğini göstermektedir. ChatGPT’nin teknik kullanım, yüzey organizasyonu ve biçimsel düzen gibi ölçütlerde daha eleştirel değerlendirmeler yaptığı; buna karşılık bazı görsellerde duygu aktarımı ve hayal gücü gibi alanlarda daha sınırlı yorumlar sunduğu görülmüştür. Bu veriyi destekler nitelikte Liu vd. (2026) yapmış oldukları araştırmada özellikle ayrıntılı ölçütler dayalı değerlendirme modellerinin yapay zekâ sistemleri ile daha tutarlı ve açıklanabilir sonuçlar üretimine katkı sağladığını belirtmektedir.

Gemini değerlendirmelerinde ise genel olarak daha yüksek puanlama eğilimi dikkat çekmiştir. Gemini’nin özellikle duygu-düşünce ifadesi, sembolik anlatım, hayal gücü ve konuya uygunluk ölçütlerinde uzmanlara kıyasla daha olumlu değerlendirmeler yaptığı görülmektedir. Bu durum, Gemini’nin çocuk resimlerindeki duygusal atmosferi, sembolik göstergeleri ve anlatsal öğeleri daha geniş bir yorum çerçevesinde ele aldığı göstermektedir. Ancak bu yüksek puanlama eğilimi, yapay zekâ sistemlerinin her zaman uzman değerlendirmesiyle birebir örtüşmediğini ve kullanılan sistemin algoritmik yaklaşımına göre sonuçların değişebileceğini ortaya koymaktadır.

Araştırma sonuçları, uzman değerlendirmelerinin çocuk resimlerini yalnızca biçimsel özellikler açısından değil; gelişimsel uygunluk, pedagojik bağlam ve bireysel ifade dili bakımından da değerlendirdiğini göstermektedir. Bu yönüyle uzman görüşleri, yapay zekâ sistemlerinin sağlayamadığı bağlamsal ve gelişimsel derinliği sunmaktadır. Yapay zekâ sistemleri ise hızlı, sistematik ve ölçüt temelli analiz yapabilme özellikleriyle değerlendirme sürecini destekleyici bir işlev üstlenmektedir. Bu anlamda Ye vd. (2025) bu bulguyu destekler nitelikte, çocuk resimlerinin değerlendirilmesinde estetik ve yaratıcılık gibi bazı soyut kavramların yorumlanmasında insan uzmanlığına ihtiyaç duyulduğunu ifade etmektedir.

Genel olarak, yapay zekâ sistemlerinin çocuk resimlerinin değerlendirilmesinde tek başına nihai karar verici olarak kullanılmasının uygun olmadığı sonucuna ulaşılmıştır. Özellikle sanat eğitimi gibi öznel, yaratıcı ve gelişimsel boyutları güçlü olan alanlarda, insan uzmanlığı ile yapay zekâ analizlerinin birlikte kullanıldığı hibrit değerlendirme modellerinin daha güvenilir ve çok yönlü sonuçlar sağlayacağı düşünülmektedir. Diğer taraftan Kim vd. (2025) yapay zeka ve uzman ile geliştirilmiş olan modellerin güvenilir sonuçlar sunduğu ve özellikle ön değerlendirme sürecinde önemli katkılar sağlayabileceğini belirtmektedir.

Sonuç olarak bu araştırma, yapay zekâ destekli değerlendirme sistemlerinin sanat eğitimi alanında kullanılabilirliğine ilişkin önemli bulgular sunmaktadır. ChatGPT ve Gemini gibi yapay zekâ sistemleri, çocuk resimlerini belirli ölçütler doğrultusunda analiz edebilmekte; ancak duygusal, gelişimsel ve pedagojik bağlamın yorumlanmasında insan uzmanlığının yerini tam olarak alamamaktadır. Bu nedenle yapay zekâ, sanat eğitiminde uzman görüşünü ortadan kaldıran değil, değerlendirme sürecini zenginleştiren ve destekleyen bir araç olarak ele alınmalıdır.

ÖNERİLER

Araştırma sonuçları doğrultusunda öncelikle, yapay zekâ destekli değerlendirme sistemlerinin sanat eğitimi alanında tamamlayıcı bir araç olarak kullanılması önerilmektedir. Öğretmenler ve alan uzmanları, çocuk resimlerini değerlendirirken yapay zekâ sistemlerinden ön analiz, rubrik temelli puanlama ve geri bildirim oluşturma süreçlerinde yararlanabilir. Ancak nihai değerlendirme mutlaka uzman görüşüyle desteklenmelidir.

Çocuk resimlerinin değerlendirilmesinde yalnızca teknik yeterlilik ya da biçimsel düzen değil; gelişimsel uygunluk, duygusal ifade, özgünlük ve bireysel anlatım dili de dikkate alınmalıdır. Bu nedenle yapay zekâ sistemleri kullanılırken değerlendirme yönergeleri açık, ayrıntılı ve yaş düzeyine uygun biçimde hazırlanmalıdır. Kullanılan komutların standartlaştırılması, yapay zekâ sistemlerinden elde edilen sonuçların daha tutarlı olmasına katkı sağlayabilir.

Sanat eğitimi alanında görev yapan öğretmenlere yapay zekâ destekli ölçme-değerlendirme araçlarının kullanımı konusunda hizmet içi eğitimler verilmesi önerilmektedir. Bu eğitimlerde öğretmenlere yapay zekâ sistemlerinin güçlü yönleri, sınırlılıkları ve etik kullanım ilkeleri açıklanmalıdır. Böylece yapay zekâ, öğretmenlerin değerlendirme sürecini kolaylaştıran bilinçli bir yardımcı araç hâline getirilebilir.

Gelecek araştırmalarda daha geniş örneklem gruplarıyla çalışmalar yapılması önerilmektedir. Bu araştırma beş öğrenci resmiyle sınırlı olduğundan, elde edilen bulgular derinlemesine analiz sağlamasına rağmen genelleme açısından sınırlıdır. Daha fazla öğrenci çalışmasının incelendiği araştırmalar, yapay zekâ ve uzman değerlendirmeleri arasındaki benzerlik ve farklılıkları daha kapsamlı biçimde ortaya koyabilir.

Ayrıca farklı yaş gruplarına ait çocuk resimlerinin değerlendirilmesi, yapay zekâ sistemlerinin gelişimsel dönem özelliklerini ne ölçüde ayırt edebildiğini ortaya koyması bakımından önemlidir. Özellikle şema öncesi dönem, şematik dönem ve gerçekçilik dönemi gibi farklı gelişim evrelerinde üretilen resimlerin karşılaştırmalı olarak analiz edilmesi, yapay zekâ sistemlerinin çocuk resmi açısından yeterliliğini daha ayrıntılı biçimde değerlendirmeye olanak sağlayacaktır.

Son olarak, ileride yapılacak çalışmalarda yalnızca ChatGPT ve Gemini değil, farklı yapay zekâ tabanlı görsel analiz sistemlerinin de karşılaştırmaya dahil edilmesi önerilmektedir. Böylece farklı yapay zekâ sistemlerinin çocuk resimlerini hangi ölçütlerde daha güçlü ya da sınırlı değerlendirdiği belirlenebilir. Bu tür çalışmalar, sanat eğitiminde daha güvenilir, etik ve pedagojik açıdan uygun yapay zekâ destekli değerlendirme modellerinin geliştirilmesine katkı sağlayacaktır.

KAYNAKÇA

- Akbaş, S. (2023). 2000 sonrası Türkiye’de sanat eğitiminin gelişimi. *International Social Sciences Studies Journal*, 9(111), 7055–7064.
- Arslan, K. (2020). Eğitimde yapay zeka ve uygulamaları. *Batı Anadolu Eğitim Bilimleri Dergisi*, 11(1), 71–88.
- Bolat, A., & Adar Can, S. (2024). Farklılaştırılmış öğretim yöntemleri ile görsel sanatlar eğitim yöntemlerinin karşılaştırılması. *International Journal of Eurasian Education and Culture*, 9(28), 680–698.
- Büyüköztürk, Ş., Kılıç Çakmak, E., Akgün, Ö. E., Karadeniz, Ş., & Demirel, F. (2020). *Bilimsel araştırma yöntemleri* (28. baskı). Pegem Akademi.
- Buyurgan, S., & Buyurgan, U. (2020). *Sanat eğitimi ve öğretimi* (4. baskı). Pegem Akademi.
- Eren, E., Ayata, N., Çetinkaya, N. B., Koç, A., Yardımcı, T., & Akpınar, E. N. (2025). Eğitimde yapay zekâ kullanımı: Öğretmenlerin perspektifinden bir değerlendirme. *Toplum, Eğitim ve Kültür Araştırmaları Dergisi*, 516–568.
- Genç, M. A., Daniş, S., & Özalp Hamarta, H. K. (2023). Görsel sanatlar alanı özel yetenekli bireylerin eğitiminde yapay zekânın kullanımına eleştirel bir bakış. *Necmettin Erbakan Üniversitesi Ereğli Eğitim Fakültesi Dergisi*, 5(Özel Sayı), 497–519.
- Gülpınar, Ş., & Boyraz, B. (2024). Sanatın dijital çağda yeniden tanımlanması: Yapay zekâ perspektifinden bir inceleme. *Yıldız Sosyal Bilimler Enstitüsü Dergisi*, 8(1), 1–14.
- Hurwitz, A., & Day, M. (2007). *Children and their art: Art education for elementary and middle schools* (8th ed.). Thomson Wadsworth.
- İşler, B., & Kılıç, M. Y. (2021). Eğitimde yapay zekâ kullanımı ve gelişimi. *Yeni Medya Elektronik Dergisi*, 5(1), 1–11.
- Karasar, N. (2012). *Bilimsel araştırma yöntemi: Kavramlar, ilkeler, teknikler* (24. baskı). Nobel Yayın Dağıtım.
- Kim, J. A., Lee, S. O., Kim, Y. A., Kim, Y. W., & Kim, B. C. (2025). A study on how to use the SMC scoring system in children’s drawing analysis through collaboration between experts and AI. *Journal of Digital Policy*, 4(3), 53–63. <https://doi.org/10.23149/JDP.2025.4.3.053>

- Liu, X., Ma, X., Ma, Y., et al. (2026). XpertBench: Expert level tasks with rubrics-based evaluation. arXiv. <https://arxiv.org/abs/2604.02368>
- Lowenfeld, V., & Brittain, W. L. (1987). Creative and mental growth (8th ed.). Macmillan Publishing Company.
- Özdal, M. A. (2024). 21. yüzyıl yapay zekâ destekli resimlerde Avrupa ve Türk kültüründen izler. *Kültür Araştırmaları Dergisi*, 22, 280–307.
- Piaget, J. (1972). The psychology of the child. Basic Books.
- Samsun, M., & Güven, İ. M. V. N. (2022). Sanat eğitiminin çocuk gelişimi üzerindeki etkileri üzerine. *İdil Sanat ve Dil Dergisi*, 99, 1597–1608.
- Şen Atiker, E., Ermurat, Z., Gök, E., Mayalıdağ, B., Akkın, R., & Özkul, A. (2023). Yapay zekâ ve sanat. Üsküdar Üniversitesi Sosyal Bilimler Enstitüsü Yayınları.
- Türk, S., Gençaslan, D., & Kaya Uyanık, G. (2024). Öğretmenlerin not verme uygulamalarına ilişkin algılarının ölçme ve değerlendirme tutumları bağlamında incelenmesi. 1. Uluslararası EJER Congress 2024 Bildiri Özetleri Kitabı içinde (ss. 482–484).
- Ye, M., Zheng, C., Yu, Z., Xiang, C., Zhao, Z., Yuan, Z., & Yannakoudakis, H. (2025). KidsArtBench: Multi-dimensional children's art evaluation with attribute-aware MLLMs. arXiv. <https://arxiv.org/abs/2512.12503>
- Wang, S., Liu, W., Wang, J., Yi, Y., & Xu, J. (2025). DrawAA-Net: An AI-supported evaluation tool for children's drawings. *Lecture Notes in Computer Science*, 15877, 309–322. https://doi.org/10.1007/978-3-031-98414-3_22