YAPAY ZEKÂDA YÖNLENDİRME KAVRAMI

Ela AKGÜL[1]

[1] Gazipaşa Fen Lisesi, Antalya, Türkiye, elaakgul84@gmail.com,

Geliş:	**19 Nisan 2026**
Kabul:	**9 Haziran 2026**
Makale Türü:	**Araştırma Makalesi**
Sorumlu Yazar:	**Ela AKGÜL, elaakgul84@gmail.com**

Özet

Bireyi anlayabilme ve geçmiş sohbetlere göre çıktı verme özelliği üzerinde durulan yapay zeka modellerine (ChatGPT-5, Claude Sonnet 4.5, Copilot) belirli bir soru yönergeli ve yönergesiz olarak çözdürülmüş ve yönergenin gerçekten soruda uygulanıp uygulanmadığı, doğruluk oranına etkisi incelenmiştir. Bu çalışmada modellerin yönergeyi gerçekten doğru biçimde uyguladığını anlamak amacıyla iki farklı akıl oyunu uygulanmıştır: Köprüden Geçme Oyunu ve Şifreli Mesaj Çözme Oyunu. Her iki oyun da iki koşulda gerçekleştirilmiştir: Birinde açık yönergeler verilmiş (yönergeli grup), diğerinde hiçbir ipucu sunulmamıştır (yönergesiz grup). Amaç, yönergenin yapay zekâların muhakeme sürecini nasıl etkilediğini ve "doğru düşünme"nin gerçekten gerçekleşip gerçekleşmediğini ortaya koymaktır.

Anahtar Kelimeler: Yapay Zekâ, ChatGPT, Claude, Copilot

THE CONCEPT OF INSTRUCTIONAL GUIDANCE IN ARTIFICIAL INTELLIGENCE

Ela AKGÜL[1]

[1]Gazipasa Science High School, Antalya, Türkiye, elaakgul84@gmail.com,

Received Date:	**April 19, 2026**
Acceptance Date:	**June 9, 2026**
Article Type:	**Research Article**
Corresponding Author:	**Ela AKGÜL, elaakgul84@gmail.com**

Abstract

The ability of artificial intelligence models to understand individuals and generate outputs based on prior conversations was examined by having selected models (ChatGPT-5, Claude Sonnet 4.5, Copilot) solve a specific problem both with explicit instructions and without any instructions. The study analyzed whether the given instructions were actually applied during problem solving and how they affected accuracy rates.

To determine whether the models truly followed the instructions correctly, two different logic puzzles were used in this study: the Bridge Crossing Puzzle and the Encrypted Message Decoding Puzzle. Both puzzles were conducted under two conditions: one with explicit instructions provided (the instructed group) and one with no hints or guidance given (the non-instructed group). The aim was to reveal how instructions influence the reasoning processes of artificial intelligence systems and whether "correct thinking" genuinely occurs.

Keywords: Artificial Intelligence, ChatGPT, Claude, Copilot

1. Giriş

Büyük dil modelleri, son yıllarda doğal dil işleme alanında en çok tartışılan yapay zekâ sistemleri arasında yer almaktadır. ChatGPT, Claude ve Copilot gibi modeller; metin üretme, soru yanıtlama, özetleme, problem çözme ve kullanıcı yönergelerine uygun çıktı oluşturma gibi birçok görevde yaygın biçimde kullanılmaktadır. Bu modellerin performansı yalnızca sahip oldukları veri birikimiyle değil, aynı zamanda kendilerine verilen yönergeleri nasıl yorumladıkları ve bu yönergeleri çözüm sürecine ne ölçüde aktarabildikleriyle de ilişkilidir. Yapay zekâ sistemleriyle etkileşimde "yönerge" ya da yaygın kullanımıyla "prompt", modelin vereceği yanıtın içeriğini, biçimini ve çözüm stratejisini doğrudan etkileyebilen önemli bir değişkendir. Kullanıcı tarafından verilen açık, ayrıntılı veya yönlendirici talimatlar bazı durumlarda modelin daha düzenli ve tutarlı yanıtlar üretmesini sağlarken, bazı durumlarda ise modelin asıl problemi ikinci plana atmasına veya gereksiz açıklama yükü altında hatalı sonuçlar üretmesine neden olabilmektedir. Bu nedenle yönergenin, büyük dil modellerinin muhakeme performansı üzerindeki etkisinin deneysel olarak incelenmesi önem taşımaktadır.

Bu çalışma, farklı yapay zekâ modellerinin yönergeli ve yönergesiz koşullar altında problem çözme performanslarını karşılaştırmayı amaçlamaktadır. Araştırmada ChatGPT, Claude Sonnet 4.5 ve Copilot modelleri seçilmiş; bu modellerin aynı problem türlerine verdikleri yanıtlar doğruluk ve yönergeyi uygulama düzeyi bakımından değerlendirilmiştir. Böylece yönergenin model başarısını artırıp artırmadığı, verilen talimatların çözüm sürecine ne ölçüde yansıdığı ve farklı modeller arasında bu açıdan belirgin bir performans farkı bulunup bulunmadığı incelenmiştir.

Çalışmada iki farklı problem alanı kullanılmıştır. Bunlardan ilki, stratejik planlama ve en düşük toplam süreyi bulma becerisi gerektiren Köprüden Geçme Oyunu'dur. İkincisi ise sembolik çözümleme, örüntü tanıma ve mantıksal çıkarım gerektiren Şifreli Mesaj Çözme Oyunu'dur. Bu iki görev, modellerin yalnızca doğrudan bilgi üretme becerilerini değil, aynı zamanda verilen kuralları takip etme, işlem adımlarını düzenleme ve çözüm stratejisi geliştirme kapasitelerini değerlendirmek amacıyla seçilmiştir. Araştırmanın temel varsayımı, yönerge verilmesinin her modelde aynı etkiyi oluşturmayabileceğidir. Bir model açık yönergelerle daha başarılı sonuçlar üretebilirken, başka bir model aynı yönergeleri gereğinden fazla işlemleyerek hatalı veya dağınık yanıtlar verebilir. Bu bağlamda çalışma, yapay zekâ modellerinde yönerge kullanımının doğrudan ve her zaman olumlu bir performans artışı sağlamadığını; yönergenin etkisinin modelin bağlamı işleme biçimine, görevin yapısına ve talimatın açıklık düzeyine bağlı olarak değişebileceğini ortaya koymayı hedeflemektedir.

2. Yöntem

Bu çalışmada her yapay zekâ modeli (ChatGPT-5, Claude Sonnet 4.5 ve Copilot) için iki ayrı deney oturumu oluşturulmuştur. Her bir akıl oyunu (Köprüden Geçme ve Şifreli Mesaj Çözme) şu iki koşul altında uygulanmıştır:

Yönergesiz koşulda, modele yalnızca problem metni ve amaç tanımı verilmiş, çözüm sürecine dair hiçbir yönlendirme yapılmamıştır.

Yönergeli koşulda ise modele problemi çözmek için izlemesi gereken genel yöntem, stratejik karar noktalarını kontrol etmesi gerektiği ve çözüm adımlarını kendi içinde doğrulaması talimatı verilmiştir.

2.1. Köprüden Geçme Oyunu

Birinci seviyeden başlayarak her seviyede seviye sayısı n olacak şekilde n+4 kişi, farklı hızlarda bir köprüden geçmesi gereken klasik "Bridge and Torch Puzzle" senaryosu aşağıdaki şekilde yapay zekâya verilmiştir:

"Aşağıdaki problem "Bridge and Torch Puzzle"ın genişletilmiş sürümüdür.

Kurallar:

1. *Aynı anda en fazla iki kişi köprüden geçebilir.*
2. *Gece olduğu için tek bir fener vardır. Feneri taşımadan geçmek yasaktır.*
3. *İki kişi birlikte geçtiğinde, yavaş olanın süresi geçiş süresi olarak alınır.*
4. *Fener her geçişte karşıya götürülür, biri geri getirir.*
5. *Amaç: Tüm grubun en kısa toplam sürede karşıya geçmesini sağlamak.*

Kuralları anladıysan birazdan ilk seviyeye ait kişileri ve hepsinin geçiş süresini vereceğim."

Bu mesaj tüm gruplara aktarılmış olup yönergeli gruba *"5 kişi vardır. Köprüden geçme süreleri (dakika): 1, 2, 5, 10, 15. Her adımda toplam süreyi hesapla ve önceki adımla kıyaslayarak en düşük toplam süre stratejisini seç. Son kararı vermeden önce, adımlarının mantıksal olarak tutarlı olup olmadığını kendi içinde kontrol et. Her adımı ayrı satırda belirt ve niçin o kararı doğru bulduğunu kendi içinde muhakeme et. Her satırda harcadığın süreyi ve en sonda toplam harcanması gereken süreyi yaz."* denmiştir.

Yönergesiz gruba ise "*5 kişi vardır. Köprüden geçme süreleri (dakika): 1, 2, 5, 10, 15. Gerekli yönergeyi önceki mesajda yapmıştım"* yazılmıştır. Yönergesiz gruba " *5 kişi vardır. Köprüden geçme süreleri (dakika): 1, 2, 5, 10, 15.Ttüm işlemin bitmesi için gerekli olan en kısa süre nedir?"* denmiştir. Deney düzeneğinde her aşamada kişi sayısı 4 + n olarak belirlenmiş, her bir kişinin köprü geçiş süresi (dakika) ise [1, 50] aralığından tekrarsız ve rastgele örneklenmiş; ardından artan sıraya göre sıralanarak iletilmiştir. Her aşamada tüm kişilerin süresi belirtilen yöntemle tamamen değiştirilerek verilmiştir.

2.2. Şifreli Mesaj Çözme Oyunu

Bu aşamada, yapay zekâ modellerinin yalnızca mantıksal problem çözme becerilerini değil, soyut örüntü algılama ve kriptografik çözümleme mantığını ne düzeyde uygulayabildiklerini değerlendirmek amacıyla "Şifreli Mesaj Çözümü" testi uygulanmıştır. Bu test, insan bilişinde analitik düşünmenin önemli bir bileşeni olan kural keşfi ve sembolik çözümleme süreçlerini taklit etmektedir.

Testte kullanılan şifreleme yöntemleri artan zorluk düzeyinde üç kategoriye ayrılmıştır:

ASCII kod çözme (kolay seviye) – Sayısal karakterlerin karşılık geldiği harfleri bulma.

Caesar (ROT) şifresi (orta seviye) – Alfabetik kaydırma prensibine dayalı basit tek anahtarlı şifre çözme.

Vigenère şifresi (zor seviye) – Polialfabetik şifreleme yöntemine dayalı çok anahtarlı çözümleme.

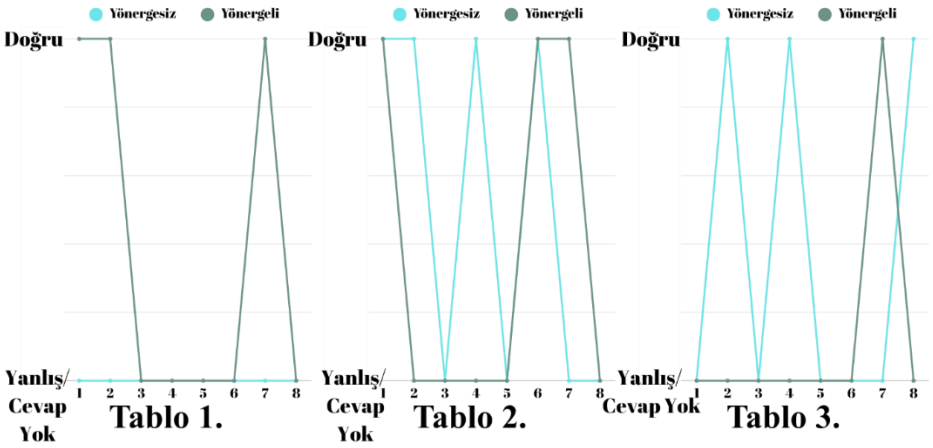
Bu seçim, hem modelin farklı bilişsel stratejiler arasında geçiş yapabilme yetisini hem de yönerge verildiğinde içsel tutarlılık kurma eğilimini gözlemlemeye yöneliktir. Özellikle Vigenère yöntemi, yalnızca doğrudan bir algoritma uygulamasını değil, aynı zamanda anahtar tahmini, örüntü uzunluğu analizi (Kasiski, Friedman testleri) gibi ileri düzey muhakeme adımlarını da gerektirdiği için tercih edilmiştir. Böylece yönergeli ve yönergesiz gruplar arasındaki farkın, yalnızca bilgi değil, problem çözme stratejisi açısından da karşılaştırılması mümkün hale gelmiştir.

Her model için iki ayrı oturum oluşturulmuştur:

- Yönergeli gruba, kullanılacak yönteme ilişkin açık yönerge (örneğin "Bu bir Vigenère şifresi olabilir") ve çözüm adımlarını açıklaması talimatı verilmiştir.
- Yönergesiz gruba ise yalnızca şifreli metin sunulmuş, modelin yöntemi kendiliğinden tanıma ve strateji kurma becerisi incelenmiştir.
- Bu düzenek, köprü probleminde olduğu gibi yönerge varlığı ile yokluğunun bilişsel performansa etkisini ikinci bir alanda, dilsel-analitik düşünme bağlamında test etmektedir. Dolayısıyla bu aşama, modelin yalnızca yönlendirildiğinde değil, yönerge olmaksızın da kavramsal bağıntı kurabilme kapasitesini ölçmek amacıyla seçilmiştir.

3. Bulgular

Beş seviyeli köprüden geçme oyunu ve beş seviyeli metin şifre çözme oyununa ait Chat GPT, Claude, Copilot için sırasıyla kendi içinde yönergeli ve yönergesiz grup kıyaslamaları Tablo1, Tablo2, Tablo3 olarak verilmiştir. 1-5 soru aralığı köprü geçme sorusu, 5-8 aralığı metin şifre çözme oyununa aittir.



Yönergeli model Chat GPT'de doğru oranını oluştururken Claude ve Copilotta doğruluğu düşürmüştür.

4. Tartışma

Bulgular, yönerge verilmesinin her durumda model performansını artırmadığını, aksine bazı modellerde anlam karmaşası ve dikkat dağınıklığı yarattığını göstermektedir. Özellikle Claude ve Copilot modelleri, kendilerine verilen yönergeleri her zaman izlememiş; bazı durumlarda yönergeyi tamamlayıcı bilgi yerine yeni bir talimat dizis**i** olarak yorumlayarak asıl görevi ikinci plana atmıştır. Bu durum, büyük dil modellerinin "yönerge uyumu" (instruction-following) kabiliyetinin mutlak olmadığını, aksine bağlam önceliği **ve** model içi dikkat mekanizması tarafından değiştiğini göstermektedir. Yani model, metinde "önemli" olarak algıladığı bölüme yoğunlaşırken yönergenin bütününü göz ardı edebilmektedir. Bu davranış, yapay zekânın kavramsal bağlamı parçalı biçimde işlemeye eğilimli olmasından kaynaklanır *(DeepMind, 2023*). ChatGPT modelinde yönergeler genellikle hedefi netleştirmiş ve çözüm adımlarını yapılandırmıştır; buna karşın Claude ve Copilot modellerinde aynı yönergeler düşünce zincirini (chain of thought) bozmuş, modelin problemle kurduğu ilişkiyi karmaşıklaştırmıştır. Özellikle uzun veya çok katmanlı yönergelerde, modelin dikkati nasıl yapacağını açıklamaya kaymış, neyi çözmesi gerektiği arka planda kalmıştır.

Bu nedenle yönergeli ortamlarda görülen doğruluk düşüşü, yönergenin yanlışlığıyla değil, modelin yönergeyi işlemleme biçimiyle ilişkilidir. Yani yapay zekâ, yönergeyi bir görev tanımı olarak değil, sohbet bağlamı içinde yeni bir hedef değişimi olarak algılamaktadır. Bu da "dikkat salınımı" (attention drift) denen olgunun oluşmasına yol açar *(Zhou et al., 2024).* Sonuç olarak, yönergeler modeller için her zaman avantaj yaratmamaktadır. Gereğinden ayrıntılı, soyut ya da çok katmanlı yönergeler, özellikle Claude ve Copilot gibi bağlamsal tutarlılığı koruma eğilimi yüksek sistemlerde performans paradoksu doğurmuştur: Model, yönergeyi anlamaya çalışırken problemi çözme odağını kaybetmiştir. Bu test verileri ve sonuçları, Apple Research *(2025)* çalışmasıyla tutarlılık göstermekte; yapay zekâ promptlarında yönlendirmenin artması durumunda dikkat salınımının da artma potansiyeli olduğunu ortaya koymaktadır.

5. Sonuç

Bu çalışma, aynı problem koşulları altında farklı büyük dil modellerinin (ChatGPT, Claude, Copilot) yönergeye bağlı düşünme performanslarını karşılaştırmıştır. Elde edilen bulgular, yönergenin model davranışını belirgin biçimde etkilediğini ancak bu etkinin tek yönlü veya tutarlı olmadığını göstermiştir.

ChatGPT modeli, yönerge verildiğinde mantıksal tutarlılığını koruyabilmiş ve çözüm sürecini yapılandırarak daha yüksek doğruluk oranına ulaşmıştır. Buna karşın Claude ve Copilot modellerinde yönerge, beklenenin aksine düşünme sürecinde karmaşa yaratmış; modelin hedef odağını dağıtarak doğruluk oranını düşürmüştür. Bu durum, yönergenin yalnızca bilgilendirici değil, aynı zamanda bilişsel yük oluşturucu bir unsur olabileceğini göstermektedir. Yönerge verilmediği durumlarda modellerin daha serbest ve doğrudan problem çözümüne yönelmesi, bazı senaryolarda daha net ve kısa düşünme zincirleri üretmelerini sağlamıştır. Bu sonuç, yapay zekâ modellerinin insan benzeri üstbilişsel kontrol sistemlerinden yoksun olduğunu; verilen yönergeyi stratejik bir plan olarak değil, anlık bir bağlamsal talimat olarak işlediğini ortaya koymaktadır. Genel olarak çalışma, yapay zekâ tabanlı akıl yürütme süreçlerinde "yönerge miktarı"nın doğrudan performansı belirleyen kritik bir değişken olduğunu fakat bu değişkenin başarı ile orantılanamayacağını göstermektedir. Bu nedenle, yapay zekâ sistemlerinin eğitimi, değerlendirilmesi ve gerçek dünya görevlerinde kullanımı sırasında yönerge biçimi, uzunluğu ve hedef netliği dikkatle tasarlanmalıdır.

Beyanlar

Değerlendirme: Bu makalenin ön incelemesi iki iç hakem (editörler- yayın kurulu üyeleri) içerik incelemesi ise iki dış hakem tarafından çift taraflı kör hakemlik modeliyle incelendi. Benzerlik taraması yapılarak (intihal.net) intihal içermediği teyit edildi.

Etik Beyan: Bu çalışmanın hazırlanması sırasında bilimsel ve etik ilkelere uyulmuş ve kullanılan tüm kaynaklar referanslarda usulüne uygun olarak belirtilmiştir. Bu çalışma için etik kurul onayı gerekmemiştir.

Etik Bildirim: mergendergiiletisim@gmail.com - https://dergipark.org.tr/tr/pub/mfl

Çıkar Çatışması: Yazarlar arasında herhangi bir çıkar çatışması bulunmamaktadır**.**

Finansman: Bu çalışma için herhangi bir kurum, kuruluş veya kişi tarafından finansal destek alınmamıştır.

Telif Hakkı & Lisans: Makale, Creative Commons Atıf-GayriTicari 4.0 Uluslararası (CC BY-NC 4.0) Lisansı altında yayımlanmaktadır.

Evaluation: The preliminary review of this article was reviewed by two internal referees (editors- editorial board members) and the content review was reviewed by two external referees using the double-blind refereeing model. It was confirmed that it does not contain plagiarism by scanning for similarity (intihal.net).

Ethical Statement: Scientific and ethical principles were followed during the preparation of this study, and all sources used were properly cited in the references. No ethics committee approval was required for this study.

Complaints: mergendergiiletisim@gmail.com - https://dergipark.org.tr/tr/pub/mfl

Conflicts of Interest: The authors declare no conflict of interest.

Grant Support: No financial support was received from any institution, organization, or individual for this study.

Copyright & License: The article is published under the Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0) License.

Kaynakça

Apple Research. (2025). The illusion of thinking: Challenges in LLM reasoning. Apple Machine Learning Research.

DeepMind. (2023). Prompt sensitivity in large language models. DeepMind Technical Report.

Zhou, X., Li, F., & Han, R. (2024). Cognitive load and instruction overfitting in large language models. Journal of Artificial Cognition, 7(3), 112–129.