

# MAKİNELER ÂŞIK OLABİLİR Mİ?

Doç. Dr. Enis DOKO<sup>1</sup>

**Yayın Bilgisi:** *Yapay Zekâ: Güncel Yaklaşımlar ve Uygulamalar*, Prof. Dr. Mustafa K. Yılmaz, Doç. Dr. N. Öykü İyigün (Editörler), Beta Kitap: İstanbul, (2021), s. 345-368

## ÖZET

*Bu makalemizde Yapay Zekâ'nın hakkındalık, öznellik, özgür irade gibi gerçek zihin hallerine sahip olup olmadığı incelenecek, lehte ve aleyhteki felsefi argümanlar değerlendirilmeye çalışılacaktır. Birinci bölümde Yapay Zeka zihin hallerine sahip olmamasına göre sınıflandırılacaktır. Daha sonra Yapay Zekâ'nın zihne sahip olup olmadığı ile ilgili en ünlü test olan Turing testi tanıtılacak ve temel eleştirileri ele alınacaktır. Üçüncü bölümde zihnin nasıl sistemlerde ortaya çıktığı ile ilgili felsefi görüşler özetlenecek ve bu görüşlerin Yapay Zekâ ile bağlantısı kurulacaktır. Sonraki iki bölümde Yapay Zekâ'nın zihin hallerine sahip olamayacağı aleyhindeki en ünlü iki argüman, Çin Odası Düşünce Deneyi ile Gödelci Argüman incelenecek ve bu argümanların eleştirileri değerlendirilecektir. Daha sonra farklı özgür irade yaklaşımları ışığında Yapay Zekâ'nın özgür irade sahibi olup olamayacağı sorusu ele alınacaktır. Son bölümde ise bu sefer Yapay Zekâ'nın mümkün olduğu iddiası lehinde David Chalmers'in Solan Öznellik düşünce deneyi değerlendirilecektir. Sonuç olarak ele aldığımız soruya cevap vermek için yeterli bilgiye sahip olmadığımız savunulmaya çalışılacaktır.*

## GİRİŞ

Yapay Zekâ (YZ) zihinlerin yapabildiği, onlara özgü eylemleri yapabilen bilgisayar ya da makineler üretmeyi hedefleyen disiplindir. Diğer bir deyişle YZ, yapay “kişi” ya da “hayvanlar” üretmeye çalışır. Bu da normal şartlar altında kişilere ya da hayvanlara<sup>2</sup> attığımız özelliklerin makineler tarafından sahip olunmasını gerektirir. Bu özelliklerden bazıları, mesela hesaplama, öngörülebilirlik gibi zekâ gerektirirken, diğer bir grup özellik mesela görmek zekâ gerektirmez. Bundan dolayı YZ ile ilgili daha doğru tanımlama, hayvan ve kişilerle ilişkilendirdiğimiz psikolojik eylemlerin (tanıma, hesaplama, görme, planlama vs.) bilgisayar ve makineler tarafından gerçekleştirilmesini sağlamayı hedefleyen disiplin olmasıdır (Boden, 2016: 1).

YZ çeşitli açılardan felsefe ile içi içe bir disiplindir. Her şeyden önce YZ'de kullanılan en temel yaklaşım ve teknikler arasında sayılabilecek önermeler ve yüklem mantığı dâhil çok sayıda geleneksel mantık sistemleri, inançlara ilişkin akıl yürütmeyle ilgili bir mantık türü olan doksastik mantık, yükümlülük ve izin gibi kavramlarla ilişkili olan deontik mantık, tümevarımsal mantık, Bayesçi doğrulama kuramı ve diğer olasılıksal akıl yürütme araçları felsefeciler tarafından geliştirilmiş ve hala felsefe içerisinde yoğun bir şekilde kullanılmaktadır. Hatta YZ'yi ortaya çıkaran çalışma olarak bilinen Alan Turing'in “Computing Machine and Intelligence” makalesi bir felsefe dergisi olan *Mind*'de yayınlanmıştı. Ancak temel araçlardaki ortaklık dışında felsefe aynı zamanda düşünmek nedir, bilinç nasıl sistemlerde ortaya çıkabilir, kişi olmak ne demektir, insana ait zihin süreçlerinin hangileri fiziki

<sup>1</sup> İbn Haldun Üniversitesi, İnsan ve Toplum Bilimleri Fakültesi, Felsefe Bölümü, enis.doko@ihu.edu.tr

<sup>2</sup> Bazı YZ uzmanları insan düşüncesinin anlaşılıp yapay bir şekilde oluşturulmak için fazla karmaşık olduğunu düşünür. Onlar bu yüzden böcekler gibi daha basit canlılar üstünde çalışarak, onların davranışlarını yapay bir şekilde oluşturmayı denerler.

sistemler tarafından taklit edilebilir gibi YZ ile ilgili çok sayıda temel sorularla ilgilenen disiplindir. Bu anlamda iki disiplin birbirleri ile etkileşim içindedirler.

Felsefeciler için YZ ortaya çıkardığı etik sorular açısından da ilginçtir. Hangi koşullarda bir varlık ahlaki haklara sahip olabilir? Bilinç sahibi olmayan ancak her koşulda insan gibi davranan bir varlık herhangi bir hukuki hakka sahip olabilir mi? Bir otonom makine yanlışlıkla suç işlediği zaman cezalandırılması gereken kimdir? Farklı işlemci hızına sahip YZ'ler hapis cezası yatacaksa bunların yatacağı süre ne olmalıdır? Tüm bu sorular doğaları gereği felsefecilerin cevap vermesi gereken ve gelecekte YZ'nin yaygınlaşması durumunda hayati öneme sahip olacak sorulardır.

Biz bu bölümde YZ ile ilgili en temel sorulardan biri olan, "Makineler düşünebilir mi?" sorusuna verilmeye çalışılan soruya odaklanacağız. Daha doğru bir tanımlama ile "Makineler bütün zihinsel süreçlere sahip olabilir mi?" sorusuna cevap arayacağız<sup>3</sup>. Çok sayıda felsefeci makinelerin gerçek anlamda düşünemeyeceği kanaatindedir. Bu makalemizde bu kanaatte olan felsefecilerin görüşlerini özetleyip değerlendirmeye çalışacağız. Böylesi bir değerlendirme yapmak için YZ'leri yeteneklerine göre sınıflamakta fayda var. Dolayısıyla önce YZ'leri sınıflamakla başlayacağız.

## 1. YZ'İN TÜRLERİ

Bazı YZ ve makinelerin insanlar tarafından yapıldıkları zaman zekâ gerektiren bazı faaliyetleri yapabildiği apaçıktır. Mesela hesap makinesi trigonometrik fonksiyonları hesaplama, integral alma gibi zor matematiksel işlemleri yapabilmekte ya da Deep Blue gibi yazılımlar satrançta en usta oyuncularını yenebilmektedir<sup>4</sup>. İntegral almak ya da satranç oynamak zekâ gerektiren faaliyetlerdir. Her ne kadar bu YZ'ler zekâ gerektiren bu faaliyetleri insanlardan başarılı yapabilseler de bu faaliyetler dışında faaliyetleri yapamazlar. Deep Blue Kasparov'u yenmiş olabilir ama en basit bir matematiksel işlemi bile yapamaz, benzer şekilde bir hesap makinesi de satrançta bir bebeği bile yenemez. Bu YZ'ler uzmanlaştıkları alan dışında hiçbir zekâ gerektiren faaliyeti yapamazlar. Bu tarz YZ'lere *lokal YZ'ler* diyeceğim. Günümüzde çoğu geliştirilen YZ'ler ve bu alanda yapılan AR-GE'ler farklı lokal YZ'ler geliştirmekle ilgilidir. Lokal YZ'lerin mümkün olduğu herkesin kabul ettiği, yüzlerce örneği ile gösterilmiş bir hakikattir<sup>5</sup>.

Diğer taraftan insan zihninin yapabildiği gibi çok sayıda farklı zeka gerektiren faaliyeti yerine getirebilen YZ'lere global YZ diyeceğiz. Global YZ'leri de incelememiz açısından iki farklı kategoriye bölmekte fayda var. Birinci gruba *zayıf global YZ* diyeceğiz<sup>6</sup>. Bu gruptaki global YZ'ler zeka gerektiren davranışlar sergileyebilir, çok karmaşık problemleri çözebilirler. Ancak bu YZ'lerin zihin durumları yoktur. Diğer bir deyişle bu YZ'ler zekâ gerektiren işleri yapsalar da bu işleri *anlayamazlar*. Bu anlamda Deep Blue programı zayıf lokal YZ'dir, çünkü satranç oynayabilse de satranç oynadığının farkında değildir, satranç hamlelerinin ne anlama geldiğini anlayamamaktadır. Zayıf YZ gerçek anlamda bilince sahip değildir<sup>7</sup>. Anlama ya da

<sup>3</sup> Makale başlığımızın "Makineler aşık olabilir mi?" sorununun olmasının sebebi budur.

<sup>4</sup> Deep Blue 1997 yılında efsanevi Dünya Şampiyonu Satranç oyuncusu Gary Kasparov'u yendi.

<sup>5</sup> Tabii lokal bir YZ aynı problemi insanlardan farklı yollarla çözebilir, hatta genelde öyle çözer. Dolayısıyla zihni bir faaliyetle ilgili lokal bir YZ geliştirmek, aynı faaliyeti çözerken insan zihninin nasıl yol izlediğini anlamada bize yardımcı olmayabilir.

<sup>6</sup> Zayıf YZ terimi literatürde farklı anlamlarda kullanılabilir. YZ uzmanları genelde Zayıf YZ kavramından benim yukarıda lokal YZ olarak tanımladığım şeyi anlarlar. Ancak bu kavramın mucidi felsefeci John Searle bu kavramı benim metinde tanımladığım anlamda kullanır (Searle, 1980).

<sup>7</sup> Bilinç kavramı gündelik hayatta çok farklı anlamlarda kullanılabilir. New Age ile uğraşanlar bilinci genelde canlılığı taşıyan bir çeşit akışkan ya da hayat enerjisi olarak yorumlarlar. Bilince popüler kültürde verilen bir diğer anlam da onun ruha eşdeğer olduğudur. Felsefeciler bu anlamlardan farklı olarak bilinci kafamızın

felsefecilerin deyimi ile *hakkındalık* bilincin en önemli özelliklerinden biridir. Hakkındalık düşüncelerle düşündüğümüz şey arasında bağlantı kuran zihnimizin özelliğidir. Zekâ gerektiren bir problemi çözmek için hakkındalığa sahip olması gerekmez. Algoritmalar problemleri anlamaya gerek kalmaksızın çözüm yöntemi sunarlar, bundan dolayıdır ki bir öğrenci yöntem ezberleyerek anlamadan bir sınavı geçebilir. Bir diğer önemli zihin halleri felsefecilerin *öznellikler (qualia)* dediği deneyimin zihnimizde bıraktığı izlerdir. Mesela kırmızı bir nesneye baktığımız zaman bilincimizde algı olarak belirir. Kırmızıyı dalga boyundan tanıyan bir cihaz düşünelim, bu cihaz kırmızıyı tespit etmesine rağmen pekâlâ kırmızı algısını deneyimleyemeyebilir. Bu cihazın öznellik/qualia özelliğinden yoksun olduğunu söyleyebiliriz. Zayıf YZ de renkleri tanıyıp ayırt edebilirken, öznellik özelliğinden yoksun bir YZ olacaktır. Yazımıza başlığını veren âşik olma deneyimi de böylesi bir öznellik deneyimidir. Dolayısıyla zayıf YZ, âşik olma taklidi yapsa da gerçek anlamda âşik olamaz.

*Güçlü global YZ*ler zekâ gerektiren davranışlar sergilemelerinin yanında öznellik ve hakkındalık gibi zihin hallerine sahip olan YZ'lerdir. Bu makinelerin gerçek anlamda düşündüğü ve bilinçli olduğu söylenebilir. Yazımızın konusu olan makinelerin âşik olmasının imkânı sorusu, güçlü global YZ mümkün müdür sorusunun özel bir halidir. YZ kavramının ortaya çıkmasından beri çok sayıda felsefeci güçlü global YZ'nin imkânsızlığını savunan argümanlar geliştirdi. Bu yazımızda bu argümanları ele alıp değerlendirmeye çalışacağız. Yazımızın ilerleyen bölümlerinde güçlü YZ derken, aslında güçlü global YZ'leri kastedeceğim.

Güçlü global YZ'nin imkânı sorusu çeşitli açılardan çok önemli bir sorudur. Sadece felsefecileri değil aslında mühendisler ve işletmeciler açısından da önemli bir sorudur. Çoğu insan bilinçli canlılara, bilinç seviyesi ölçüsünde ahlaki sorumluluk hisseder. Acı duyan, bu acının bilincinde olan, farkındalığı yüksek bir varlığa zarar vermek, varlıklarını durdurmak ("öldürmek") ya da onları iradeleri dışında çalıştırmak çoğu insan tarafından ahlaki olarak yanlış görülecektir. Dolayısıyla güçlü YZ geliştirecek şirketler "YZ köleliğini" desteklemekle suçlanma tehlikesi ile karşı karşıya kalabileceklerdir. Bunun sonucu olarak YZ'lerde bilincin ortaya çıkıp çıkamayacağı, çıkarsa hangi şartlarda çıkacağı sorusu YZ çalışan şirketler açısından son derece önemlidir. Ev işlerinde çalışacak bir YZ'nin bilinçsiz olması istenilen bir şey olabilir belki ama bazen de makinelerin bizle zarar vermemesi için bizimle empati kurabilmesinin istendiği durumlar ortaya çıkabilir. Gene YZ'lerin özgür irade sahibi olup olamayacağı sorusu da güvenliğimiz açısından son derece önemlidir. Sonuç olarak güçlü global YZ'lerin imkânı sadece felsefeciler için değil, genel olarak YZ için can alıcı öneme sahip bir sorudur. Bu soruya vereceğimiz cevap aynı zamanda insan zihnini anlamada bize son derece yardımcı olacaktır.

Güçlü YZ aleyhindeki argümanları incelemeden önce, tartışmada önemli bir yere sahip olmasında dolayı Turing Testini ele alacağım.

## 2. TURING TESTİ

Peki, bir makinenin gerçekten düşünüp düşünmediğini nasıl anlayabiliriz? YZ çalışanlar arasında soruya en yaygın cevap Turing Testidir (TT). Bu test Turing'in 1950 yılında yayınladığı klasik makalesi "Computing Machines and Intelligence"ta geçmektedir (Turing, 1950). Turing aslında adını taşıyan testi makinelerin düşünüp düşünmediğinin bir testi olarak sunmaz. Turing makalesinde "Makineler düşünebilir mi?" sorusuna cevap aramak istediğini ancak bu sorunun cevap verilemeyecek kadar muğlak olduğunu düşündüğü için soruyu bir

---

içindeki düşünce algısı olarak anlarlar. Anlama, renk ya da sesi deneyimleme, anlam vermek hepsi bilincin parçasıdır. Felsefede en zor sorulardan biri bilincin doğasını anlamaktır. Nasıl oluyor da 2 kilodan hafif bir sıvı yapı olan beyin bilinç gibi ilginç bir fenomene yol açabiliyor?

oyunla deęiřtirir. Turing makalede taklit oyunu dedięi bir oyunu tarif eder. Oyun bir kadın, bir erkek ve bir cinsiyeti önemli olmayan toplamda üç oyuncu tarafından oynanır. Kadın bir odaya, erkek de başka bir odaya geçer. Üçüncü oyuncu, ki ona sorgucu diyelim, iki oda ile email<sup>8</sup> aracılığı ile irtibat kurmaktadır. Oyun kurallarına göre iki oyuncu da kendini kadın olarak tanıtır. Kadın oyuncu sorgucuyu doğru söylediğine, erkek de kadın olduğuna ikna etmeye çalışır. Sorgucu dilediği soruları sorabilir ve gelen cevaplara göre kimin kadın olduğunu tahmin etmeye çalışır.

Turing “taklit oyunu” dediği bu oyundaki oyunculardan birinin bilgisayar olduğu bir durum hayal eder. Eğer bilgisayar erkek oyuncunun yerini alıp, sorgucuyu ikna etmede insan kadar başarılı olursa Turing’e göre genel halk “makine zekâsından” bahsetmeye başlayacaktır. Turing 2000 yılına geldiğimizde YZ’lerin bu taklit oyunu testini başarı ile geçeceklerini öngörmüştür. Turing 2000’li yıllarda bilgisayarların sahip olacağı hesap gücünü doğru bir şekilde öngörse bile<sup>9</sup>, YZ’nin taklit oyununu başarı ile oynayacağı öngörüsü ne yazık ki gerçekleşemedi. 1990 yılından beri her yıl TT’yi geçmek amacıyla 100.000 dolarlık Loebner ödülü için YZ’lar yarışmaktadır. Şimdiye kadar bu ödülü alan olmadı. Her yıl bu yarışmada en çok başarılı olan YZ 2000-3000 arası bir ödül almaktadır.

TT yapay zekânın başarılı bir testi olabilir mi? Kanaatimce bu sorunun cevabı hayırdır. Bu sadece benim değil çok sayıda felsefeci ve YZ uzmanının kanaatidir. Her şeyden önce bir YZ’nin başarısını insanlarla kıyaslayarak ölçmek doğru olmayabilir, elimizdeki en iyi YZ insan değil de bir hayvana daha yakın olabilir. İkincisi TT’yi geçme yarışı zekâyı ölçme çabasından çok insanı kandırma çabası gibi gözükmemektedir. Üçüncüsü gerçek anlamda zeki sayılmak için hatta Global YZ olmak için sadece epostaya cevap vermekten fazlasını yapmalı ve bütün davranışları ile insana eşdeğer olmalıdır. Bunun için ise YZ’nin vücuda ve hareket kapasitesine sahip olması gerekmektedir. Bundan hareketle Harnad Total TT ismini verdiği yeni bir YZ testi ortaya atmaktadır (Harnad, 1991). Bu testte YZ TT’yi geçmenin ötesinde fiziksel bir çevrede de insandan ayırt edilemeyecek davranışlar sergilemelidir<sup>10</sup>.

Kanaatimce TT’nin YZ’nin zekâyı hele de zihin hallerine sahip olduğunu gösteremeyeceği lehindeki en ciddi itiraz Ned Block tarafından geliştirilen Blockhead (Ahmak) argümanıdır (Block, 1981). Blockhead TT’yi geçebilecek özel bir bilgisayar programının adıdır. Block’a göre dilin doğası gereği bilgisayar ile sorgucu arasında konuşmayı başlatacak sonlu sayıda cümleler kümesi vardır. Bu cümlelere göre ilkece biz insanlara sorup ya da insan konuşmalarına bakarak uygun cevaplar bulabiliriz. Benzer şekilde cevabımıza verilecek sonlu sayıda olası cevap vardır, gene bu cevaplara uygun cevaplar insanlara sorup bulunabilir. Bu şekilde olası bütün diyaloglar yazıya dökülebilir. Blockhead bütün bu olası diyalogları alfabetik sırası ile yazılı olarak içermektedir. Bu yazılımın yaptığı tek şey kendisine eposta ile gelen cümleyi listede tespit edip, ona uygun olarak belirtilen cevabı eposta ile göndermektir. Bu program bariz bir şekilde TT’yi geçecektir. Ancak diğer taraftan Block’un dikkat çektiği gibi basit bir arama programı olan bu program zihin durumlarına ve zekâyı sahip değildir. Bu programın herhangi bir süpermarkette barkod okuyup ürünü tanıyıp fiyat veren bir barkod okuyucusundan bir farkı yoktur. Dolayısıyla Blockhead argümanı doğruysa TT başarılı bir şekilde zekâ ve zihin durumlarını tespit eden bir test olamaz.

TT savunucusu Blockhead’in imkânının zekâyı ve zihin hallerini TT ile tanımlamanın hatalı olduğunu kabul edip, ancak pratikte böyle bir program yazılamayacağı için TT’nin hala etkili

<sup>8</sup> O zamanlarda email olmadığı için Turing makalesinde teleprinter isimli cihazla iletişimden söz eder.

<sup>9</sup> Bu gerçekten çok şaşırtıcı bir öngördür zira Turing bu makaleyi ele aldığı tek gördüğü mekanik bilgisayar kod kırmakta kullandığı bir oda büyüklüğündeki Manchester makinesiydi.

<sup>10</sup> TT ve TTT dışında çok sayıda başka YZ testi önerilmiştir. Bunlar için bakınız: (Bringsjord, 1995).

bir zekâ testi olduğunu iddia edebilir. Ancak böylesi bir yaklaşım mühendisler için yeterli olsa da felsefeci açısından bu testin kavramsal olarak sorunlu olduğu gerçeğini değiştirmez.

### 3. ZİHİN NE OLABİLİR?

YZ gerçek anlamda bilinç ve diğer zihin hallerine sahip olabilir mi sorusu, nasıl şeyler zihne sahip olabilir sorusu ile yakından ilgilidir. Dolayısıyla YZ ile ilgili felsefe literatüründeki farklı argümanları ele almadan önce zihnin nasıl sistemlerde ortaya çıkabileceği ile ilgili olası farklı alternatifleri incelemekte fayda görüyorum.

Zihnin madde ile ilişkisi, felsefe literatüründe zihin-beden ilişkisi sorunu olarak anılır. Bu konuda felsefe literatüründe çok sayıda farklı görüş mevcuttur. Bir kitap bölümünde, bu görüşlerin hepsini hakkı ile ele almak, hele de bunlar lehindeki/aleyhindeki delilleri incelemek mümkün değildir. Bundan dolayı biz bu bölümde bu görüşleri kabaca farklı alternatifler olarak tasnif edip bunların YZ ile ilişkisine göz atmakla yetineceğiz.

Zihin halleri nasıl tözde açığa çıkabilir? Diğer bir deyişle zihin maddi cisimlerde mi açığa çıkar? Bu soruya verilebilecek ilk cevap zihnin maddi olmayan, *ruh* adını verebileceğimiz tözlere özgü olduğudur. Zihin hallerinin ruha ait olduğu görüşü genelde dinlerle bağdaştırılan, tarihi çok eskilere giden, halk arasında en yaygın kabul gören görüştür. Buna göre bilinç ve zihin hallerimiz bedenimizin dışında maddi olmayan bir varlık olan ruh tarafından taşınır. Ruh görüşü felsefede farklı düşünce okulları tarafından savunulmuştur. Descartes'le meşhur olan *kartezyen* ya da *etkileşimci dualizm*, Leibniz'in savunduğu *paralelizm*, Huxley'in savunduğu *epifenomenalizm*, *vesilecilik (okazyonalizm)* ve *idealizm*'in farklı türleri ruh görüşüne örnek olarak verilebilir. Kartezyen dualizme göre biz ruh ve bedenden oluşmaktayız ve bu iki farklı töz birbiri ile etkileşmektedir. Kartezyen dualizm genelde etkileşim problemi ile eleştirilmiştir. Bu problem maddi bir tözün, maddi olmayan, uzay-zamanda bulunmayan bir tözle nasıl etkileşebildiği sorusudur. Bu sorudan kaçmak isteyen felsefeciler, aslında ruh ve bedenin birbiri ile etkileşmediği, bu izlenimin iki tözün birbirinden bağımsız çalışan saatler gibi senkronize çalışmasından kaynaklandığını savunan paralelizm, bedenin ruhu etkilediğini ancak ruhun beden üstünde bir etkisinin olmadığını savunan epifenomenalizm görüşünü savunmuşlardır. Üçüncü bir yaklaşım olan okazyonalizm'e göre ise ruh ile madde aslında nedensel ilişkiye doğrudan girmemekte, bu ilişkiyi tüm neden-sonuçların asıl kaynağı olan Tanrı sağlamaktadır. İdealist düşünce okulları bütün bu saydığımız dualist okullardan farklı olarak, maddenin varlığını reddetmekte sadece ruhların var olduğunu savunmaktadır.

Ruhçu görüşler doğruysa o zaman makinler ruh sahibi olmadıkları için zihin hallerine sahip olamazlar. Dolayısıyla güçlü YZ gerçekleştirilemez<sup>11</sup>. Bu görüşler günümüzde beynin zihin hallerinde oynadığı rolün anlaşılması ve fizikteki korunum yasaları ile çeliştikleri gerekçesi ile fazla popüler değildirlir.

Ruhçuluk görüşünün tam karşısında *panpsizizm* yer almaktadır. Bu görüşe göre zihin halleri tüm maddelere ait görüşlerdir. Diğer bir deyişle evrendeki her madde parçası zihin hallerine sahiptir, elektron, atom, masalar, güneş, dünya her maddi cisim değişik seviyelerde zihin hallerine sahiptir. Bu görüş geçmişte hem batı, hem de uzak doğu düşüncesinde savunulmuştur. Özellikle 19. yüzyıl Avrupa düşüncesinde altın çağını yaşamıştır. İslam düşüncesi içinde, özellikle Kuran'da ve hadislerde dağ ya da kütük gibi cansız maddi cisimlere zihin hali atfedilmesinden hareketle bazı tasavvuf okulları bu görüşü

<sup>11</sup> Turing ruhçu görüşlerin de güçlü YZ'nin imkanı ile çelişmediği kanaatindedir. Ona göre yeterince karmaşık bir makine ruh kazanabilir. Bu her ne kadar mantıksal mümkün olsa da, robotun bir anda uzay zaman dışındaki bir ruhla irtibat kurduğu senaryosu sezgisel olarak pek makul gözükmemektedir.

benimsemişlerdir. 20. yüzyılda bu görüş büyük oranda terk edilse de, günümüzde yeniden canlanmakta, David Chalmers gibi etkili felsefeciler tarafından savunulmaktadır.

Panpsişizm doğru ise her maddi varlık zihin hallerine sahip olduğu için, zayıf YZ gerçekleştirilmesi halinde, güçlü YZ de otomatik olarak gerçekleştirilecektir. Dolayısıyla Ruhçu görüşlerden farklı olarak panpsişizm güçlü YZ'yi mümkün görmektedir.

Üçüncü görüşler ailesi *zuhurculuktur*. Zuhurculuk görüşü de zihin hallerinin maddi tözler tarafından taşındığını savunur. Ancak Panpsişizm'den farklı olarak her maddi varlığın zihin sahibi olduğunu savunmaz. Bu görüşlere göre zihin halleri, beyin gibi karmaşık yapılarda zuhur ederler. Bu görüş *özellik düalizmi* adı ile de anılır. Zuhurculuk indirgemecilik karşıtı bir yaklaşıma dayanır ve sistemler karmaşıklaştıkça en temel parçalarda olmayan yeni özelliklerin çıktığı varsayımına dayanır. Zuhur etme fenomenleri doğada yaygındır, mesela atomların kendisinin katı ya da gaz olmasından söz edilemez. Diğer taraftan atomlar bir araya geldiğinde parçalarda olmayan yeni bir özellik katılık açığa çıkar. Katılık çok sayıda atomdan oluşan sistemlerde zuhur eder. Bilinç katılık benzeri karmaşık sistemlerde açığa çıkabilen bir şeydir ve panpsişizm'den farklı olarak parçalara bilinç atfetmez.

Zuhurculuk güçlü YZ sorusunu açık bırakır. Eğer zihin hallerini ortaya çıkaran karmaşık yapılar makinelerde oluşturulabilirse, bu durumda güçlü YZ mümkündür. Diğer taraftan zihnin ortaya çıkmasını sağlayan karmaşık yapılar makinelerde oluşturulamazsa o zaman Güçlü YZ mümkün değildir.

Dördüncü görüşler ailesi *canlılıkçılıktır*. Canlılıkçılık zuhurculuğun çok özel bir alt kümesi olarak da okunabilir. Bu görüşe göre zihin halleri sadece canlılarda, diğer bir deyişle biyolojik sistemlerde ortaya çıkabilir. Zihin hallerini beyin halleri ile özdeşleştiren *beyin-zihin eşitliği* teorisi, John Searle'ın *biyolojik doğalcılığı* bu görüş ailesi içinde yer alır. Canlılıkçılar genelde zihnin, mide salgıları gibi biyolojik sistemlerin oluşturduğu bir yapı olduğu kanaatinde idirler. Biyokimya zihnin olmazsa olmaz şartıdır.

Canlılıkçılık doğru ise zihin halleri sadece canlılara özgü olduğu için güçlü YZ mümkün değildir. En azından biyolojik bir malzeme kullanmadan güçlü YZ yapmak mümkün değildir.

Beşinci görüş ailesi *fonksiyonalizmdir*. Fonksiyonalizm diğer görüşlerden farklı olarak zihnin belli bir töze bağlı olduğunu reddeder. Bazı cisimler işlevleri diğer bir deyişle nedensel ilişkileri ile tanımlar. Saat ya da kalem bu şekilde tanımlanır mesela. Saatin hangi malzemeden yapıldığının, şeklinin bir önemi yoktur. Saat zaman tutma işlevi ile tanımlanır, zaman tutan her alet iç yapısı ya da malzemesi ne olursa olsun saattir. Kalem de benzer şekilde tanımlanır, elektronik, tahta, metal ya da tüy kalemden söz etmek mümkündür, yazı işlevini gören her şey kalemdir. Fonksiyonalistlere göre zihin halleri de benzer şekilde, işlevleri yani oynadıkları nedensel roller üstünden anlaşılmalıdır. Dolayısı ruhlar, beyinler, uzaylılar, bilgisayarlar ve de başka sistemler de doğru işlevlere yani doğru nedensel ilişkilerin parçası olurlarsa zihin sahibi olabilirler. Zihin bilgisayarda çalışan programa benzetilebilir.

Fonksiyonalizm doğru ise bu durumda güçlü YZ mümkündür. Hatta çağdaş YZ destekçilerinin hemen hemen hepsi fonksiyonalizmin özel bir türü olan *hesaplama fonksiyonalizmi* savunurlar. Bu görüşe göre zihin halleri hesaplamalardır. Diğer bir deyişle beyin bilgisayar, zihin halleri ise programlardır. Nasıl ki aynı program çok farklı cihazlar tarafından çalıştırılabiliyorsa, zihin de farklı sistemlerde açığa çıkabilir. Dolayısıyla eğer zihin hallerinin ne tür hesaplama algoritmaları olduğu anlaşılır ve bu algoritmaları çalıştıracak kadar güçlü bir makine yapılabiliirse güçlü YZ pratikte de gerçekleştirilebilir.

Son olarak bir grup felsefeci zihnin yanılsama olduğunu, gerçek anlamda var olmadığını iddia ederler. Bu görüş *elemecilik (eliminativizm)* olarak bilinir. Bu görüşün en önemli temsilcisi Daniel Dennett'tir. Bu görüş doğruysa o zaman güçlü YZ ile zayıf YZ ayrımı anlamsızdır.

Bu bölümde zihnin mahiyeti ile ilgili farklı felsefi yaklaşımları kaba ve hızlı bir şekilde ele aldık. Bütün bu görüşler lehinde ve aleyhinde ciddi argümanlar sunmak mümkündür. Bu argümanların hepsini incelemeyen güçlü YZ ile ilgili kanaat geliştirmek yerinde olmayacaktır. Elbette böyle geniş kapsamlı bir çalışmayı buradaki kısıtlı alanda yapmak mümkün değildir. Zihin felsefesine yaptığımız bu kaba yolculukta amacımız zihin felsefesinin sunduğu geniş perspektife dikkat çekmektir. Detaylı bir analiz için ilgili okuyucular (Kim, 2010) ya da (Heil, 2019)'e bakabilirler. Makalemizin geri kalan bölümünde güçlü YZ aleyhinde ve lehindeki önemli birkaç argümana göz atacağız.

#### 4. ÇİN ODASI DÜŞÜNCE DENEYİ

Bilgisayarların gerçek anlamda bir bilince sahip olamayacağı ile ilgili ortaya konmuş en ünlü argüman 1980 yılında John Searle'ın ortaya attığı Çin Odası<sup>12</sup> düşünce deneyidir (Searle, 1980). Bu düşünce deneyi son 50 yıl içinde ortaya atılan düşünce deneyleri içinde en çok tartışılanıdır ve Roger Penrose'un çok satan kitabı *Kral'ın Yeni Akli* ile popüler kitleler tarafından da tanınmıştır (Penrose, 1989). Searle'ın argümanı yayınlanır yayınlanmaz 27 bilişsel bilimci argümanı eleştirmiş, o da bu eleştirilere cevap vermiştir. Düşünce deneyi 1990 yılında *Scientific American* dergisi bu düşünce deneyi ve eleştiren felsefecilerin yazılarına yer vermiş böylece deney popüler bilim çevrelerinin de gündemine girmiştir.

Düşünce deneyi özetle şöyledir. John Searle'ü kapalı bir odanın içinde hayal edelim. Searle Çince bilmemektedir. Oda içerisinde verilen herhangi bir Çince karakterlere hangi Çince karakterle cevap verilebileceği bilgisini içeren kitaplar mevcuttur. Odanın dışında Çince konuşan Çinliler mevcuttur ve kapı altından odaya çeşitli Çince karakterlerle yazılmış sorular göndermektedirler. Searle karakterleri anlamasa da kitaplardaki talimatları takip ederek, uygun Çince karakterleri kâğıda yazarak Çinlilere cevap vermektedir. Hep anlamlı cevaplar aldıkları için Çinliler bir süre sonra odadaki kişinin Çince bildiğini sanacaklardır. Hatta bu Çin odası TT'yi başarı ile geçecektir. Ancak Searle'a göre tüm bu başarısına rağmen Çin odasındaki kişinin Çince anlamadığı açıktır, bu kişi anlamadığı bir algoritmayı takip ederek sembol manipülasyonu yaparak cevap vermektedir. Searle'a göre bilgisayarların yaptığı tam olarak budur onlar sembol manipülasyonu yaparak algoritmayı takip etmektedirler. Ancak sembol manipülasyonundan yani sözdiziminden, anlam çıkarılamaz. Bütün bilgisayarlar bu şekilde çalıştıkları için hiçbir bilgisayar anlam üretemez, yani yaptığı işi anlayamaz. Çince konuşmuş gibi gözükken bir program yazabilirsiniz elbette ama bu program Çince anlamayacaktır. Dolayısıyla zayıf YZ mümkünken, güçlü YZ imkânsızdır. Çünkü bilgisayara hakkındalık kazandırmak imkânsızdır.

John Searle'ın düşünce deneyinde ortaya koyduğu argüman şu şekilde özetlenebilir:

1. Eğer güçlü YZ doğru ise, o zaman öyle bir Çinceye özel bilgisayar programı vardır ki bu programı çalıştıran bilgisayar Çince anlar.
2. Çince anlamadan bu Çinceye özel program çalıştırılabilir.
3. Demek ki güçlü YZ doğru değildir. (Searle, 1980)

Searle Çin Odası düşünce deneyini ikinci öncülü savunmak için ortaya atmaktadır.

<sup>12</sup> Aslında Searle bu düşünce deneyine "Chinese Room", yani Çince Odası düşünce deneyi demektedir. Ama Türkçe literatürde bu düşünce deneyi sebebini bilmediğim bir neden Çin Odası olarak geçmiştir.

Çin Odası düşünce deneyine çok sayıda farklı cevaplar verilmiştir. Biz burada kanaatimce en önemli olanlarını ele alacağız. Çin Odasına karşı getirilen en yaygın cevap *sistem cevabı* olarak bilinen yaklaşımdır. Jack Copeland (Copeland, 2002), John Haugeland (Haugeland, 2002), Ray Kurzweil (Kurzweil, 2002) ve Georges Rey (Rey, 1986) gibi çok sayıda alanın ünlü düşünürü bu cevabı savunmuştur. Bu yaklaşımı benimseyen kişi Çin Odası içindeki kişinin Çince anlamadığını kabul eder, ancak bunun sorun olmadığını savunur. Odadaki kişi sistemin sadece küçük bir parçasıdır, sistem kitaplar ve oda gibi başka birleşenler içerir. Parçanın, yani odadaki kişinin Çince anlamaması odanın yani sistemin tamamının Çince anlamadığı anlamına gelmez. Nitekim bir Çinlinin beyninden seçeceğimiz herhangi bir parça da Çince anlamayacaktır. Hatta belki de anlam beyinde değil kişi seviyesinde ortaya çıkmaktadır. Diğer bir deyişle anlama sistem seviyesinde ortaya çıkan bir özelliktir. Dolayısıyla evet Searle haklıdır odadaki kişi Çince anlamamaktadır, ama bu güçlü YZ'nin imkânsızlığını göstermez çünkü sistemin kendisi Çince anlamaktadır.

John Searle sistem cevabına karşı yeni bir düşünce deneyi ile karşı çıkar (Searle, 1980). Odadaki adam ilkece bütün algoritmayı ezberleyebilir. Bu durumda kitaplara ve odaya ihtiyaç kalmaz. Searle ezberlediği algoritma sayesinde kendisine verilen herhangi bir Çince karakterden oluşmuş soruya cevap verebilmektedir. Dolayısıyla artık adam sistemin tamamını oluşturmaktadır ve hala sistem manipülasyonundan başka bir şey yapmadığı için hala Çince anlamamaktadır. Searle'ün bu yaklaşımına hala kişinin bütün sistemi oluşturmadığı şeklinde bir cevap verilmeye çalışılabilir. Ancak kanımca bu *ad hoc* bir yaklaşım olacaktır.

İkinci bir cevap da *robot* cevabıdır. Bu cevap da gene çok sayıda felsefeci tarafından zaman zaman savunulmuştur. Bu düşünürlere örnek olarak Margaret Boden (Boden, 1988: 238-251), Tim Crane (Crane, 1996) ve Jerry Fodor (Fodor, 1991) verilebilir. Bu yaklaşımda odadaki adamın da sistemin de Çince anlamadığı kabul edilir, ancak buradan Güçlü YZ'nin imkânsızlığının savunulamayacağı iddia edilir. Bunun nedeni odanın dış dünyadan bilgi almaması ve onunla aktif bir şekilde etkileşememesidir. Bu görüş savunucusuna göre bir şeyi anlamak için o şeyle etkileşmek gerekir, dolayısıyla gerçek anlamda bir güçlü YZ ancak duyularla dış dünyadan bilgi alan, vücut sahibi ve gezebilen yani vücut sahibi bir robotta açığa çıkabilir. Çin odası böyle bir sistem olmadığı için elbette ki güçlü YZ olamaz. Bu görüş insanlara bakarak da motive olabilir. Mesela biz kahve kavramının ne olduğunu nasıl anıyoruz? Muhtemelen onu görerek, tadına bakarak, onunla etkileşerek vs. Bu ise vücudumuz sayesinde mümkündür. Dolayısıyla YZ bir kavramın ne olduğunu anlayacaksa, vücut sahibi bir YZ yani robot olmalıdır. Bu YZ çeşitli mikrofon, kamera yardımı ile dışarıdan bilgi almalı, tekerlek ya da yapay ayaklarla çevrede dolaşmalı ve kanca ya da yapay eller aracılığı ile cisimlerle oynamalıdır. Ancak böyle bir YZ anlayabilir ve Çin odası böylesi bir YZ'nin anlayamayacağını gösteremez.

Searle bu cevabı da beğenmez (Searle, 1980). Ona göre dış dünyadan bilgi almak için eklenecek mikrofon, ya da kamera gibi sensörler makineye sadece ek bazı sembol girdiler sağlayacaktır. Ancak ona göre ek sembol manipülasyonları anlam üretmede aynı derecede başarısız olacaktır. Bu durumu göstermek için Searle gene yeni bir düşünce deneyi ortaya atar. Çin odamıza geri dönelim. Varsayalım ki odaya bir bant aracılığı ile 0 ve 1'lerden oluşan bir kod girişi yapılmaktadır. Adam Çin sembollerine cevap verirken artık bu banttaki rakamlara da bakmalıdır. Kitaplar hem giren karakter, hem de bu ek banttaki sayılara göre nasıl cevap verilmesi gerektiğinin bilgisini içerir. Adam farkında değildir ama banttaki 0 ve 1'ler kamera, mikrofon ve diğer sensörlerden gelen dijitalleştirilmiş bilgilerdir. Adam bu yeni senaryoda hala Çince anlamayacaktır, daha fazla sembol manipülasyonu eklemek hiçbir şekilde yardımcı olamaz.

Searle'ün bu cevabına sistem cevabı ile robot cevabının kombinasyonu ile cevap verilmeye kalkılabilir. Nitekim Georges Rey muhtemelen böyle bir yaklaşım geliştirmektedir. Evet yukarıdaki örnekte oda içindeki adam Çince anlamıyor olabilir, ancak sistemin kendisi dış dünyadan bilgi almaya başlayınca anlamaya başlayabilir. Searle böylesi bir itiraza muhtemelen iki cevabı birleştirerek cevap vermeye kalkabilir. Ama kanaatimce bu durumda bu cevap pek güçlü olmayacaktır. Çünkü kişi algoritmayı ezberleyip Çince sembollere cevap verirken Çince anlamayacaktır, ama duyuları aracılığı ile çevre ile etkileştiği için belki de bir noktadan sonra Çince anlamaya başlayacaktır. Elbette kişi bilgisayardan/robottan farklı olarak zaten hakkındalığa yani zihne sahip olduğu için Çince anlamaya başladı denebilir. Ancak bu yaklaşımın sezgisel olarak önceki düşünce deneyleri kadar güçlü olmadığı açıktır. Dolayısıyla bence Robot cevabı sistem cevabı ile birleştirildiğinde güçlenmektedir.

Üçüncü önemli bir cevap *beyin simülasyon* cevabıdır. Bu cevap da Çin odasının Çince anlamadığını kabul eder, ancak buradan güçlü YZ'nin imkânsız olduğunun çıkarılmayacağını savunur. Bu görüşü savunanlara göre Çin odasının anlayan gerçek bir Çinliden farkı çalışma biçimidir. Bu düşünce deneyinin gösterdiği şey semboller dizisi ile çalışan bir bilgisayar programının güçlü YZ olamayacağıdır, ancak bu gerçek anlamda beynin çalışmasını simüle eden bir programın güçlü YZ olamayacağını göstermez. Çince anlayan kişinin beynini düşünelim. Buradaki her nöronun ateşlemesini ve sahip olduğu bağlantıları simüle eden bir program düşünelim. Bu programın bilgiyi işleme biçimi beynin işleme biçimine benzeyecek ve Çin odasından farklı olacaktır. Bu görüş savunucusu böylesi bir programın Çince anlayacağını savunacaktır. Bu görüş Paul ve Patricia Churchland (Churchland ve Churchland, 1990) tarafından savunulmuştur. Churchland çifti Searle'in neden hatalı olduğunu ortaya atmak için *Işık Odası* isimli bir düşünce deneyinden söz ederler. Bir odanın içinde elinde mıknatıs tutan birini düşünelim. Bu kişi Maxwell'in "Yapay Işık Teorisine" inanmamaktadır. Bu teoriye göre ışık elektromanyetik bir dalgadır ve elektrik ya da manyetik alanın dalgalanmasından ortaya çıkmaktadır<sup>13</sup>. Odadaki kişi karanlık bir odada eline mıknatıs alıp hızlıca sağlamaktadır, ancak salladığı mıknatıs ışık yaymamaktadır. Bundan hareketle bu kişi yapay ışık teorisinin yanlış olduğunu söyler. Ancak elektromanyetik teori bilen herkes bu düşünce deneyinin hatalı olduğunu anlayacaktır. Mıknatısın ışık yayması için çok daha hızlı sallanması gerekmektedir, bu hız ise insanın eli ile elde edebileceği hızların üstündedir. Işık odasındaki kişinin hatası düşünce deneyine yeterince bilim katmamasıdır. Churchland çiftine göre Searle de aynı hatayı işlemekte ve düşünce deneyinde nörolojiyi göz önüne almamaktadır. Evet, Çin odası Çince bilmemektedir, ama bunun nedeni beyin bağlantılarının onda olmamasıdır. Benzer bir yaklaşım Andy Clark tarafından da savunulmaktadır. Ona göre Searle'un argümanı klasik algoritmalarla çalışan bilgisayarlar için geçerli olsa da, yapay nöral ağlarla çalışan YZ için geçerli değildir.

Searle beyin simülasyonu karşısında da başka bir düşünce deneyi önerir (Searle, 1980). Odanın devasa bir boru sistemi içerdiğini düşünelim. Buradaki her boru beyindeki bir bağlantıyı temsil etmektedir. Odadaki adam gelen girdiye göre kitaba bakıp çeşitli supaplar aracılığı ile borulardan su akmasını sağlamaktadır. Buradaki su akışı Çince anlayan bir adamın beynindeki nöron aktivasyonlarını paralel şekilde ilerletmektedir, yani oda birebir Çinli bireylerin beynini simüle etmektedir. Searle'a göre buradaki kişinin ve borulardan oluşan sistemin Çince anlamadığı açıktır, dolayısıyla ona göre bu yaklaşım da çalışmamaktadır.

Searle'ın bu cevabının çekiciliği olsa da bu düşünce deneyi önceki kadar sezgisel değildir. Bu düşünce deneyinde beyni simüle eden sistemin tamamı olduğu için her şeyden önce odadaki adamın Çince anlamamasının önemsiz olduğu açıktır. Buradaki önemli soru

<sup>13</sup> Bu Maxwell'in elektromanyetik kuramının bir sonucudur ve doğrudur.

borulardan oluşan karmaşık sistemin Çince anlayıp anlamadığıdır? Sezgilerimiz bu durumda da hakkındalık ortaya çıkmadığını ima etse de, özellikle önceki bölümde ele aldığımız zuhur etmeci yaklaşımı savunan biri bu kadar karmaşık bir sistemde ortaya çıkabilecek özelliklerle ilgili sezgilerimize güvenemeyeceğimizi savunacaktır. Dolayısıyla böylesi bir sistemin gerçek anlamda farkındalığa sahip olmadığını iddia etmek pek kolay değildir.

Ele alacağımız son cevap *sanal zihin* cevabı olarak bilinen cevaptır. Bu cevap savunucusu da odadaki kişinin ve odanın Çince anlamadığını kabul eder. Ancak bu görüşü savunanlara göre bu süreç sonunda ortaya Çince anlayan bir sanal zihin ortaya çıkmaktadır. Bu sanal zihnin varlığı sisteme bağlı olsa da, sistem dışında bir varlıktır. Bu görüşü savunanlara göre YZ programını çalıştıran bir makine için bu makine Çince anlıyor demek hatalıdır, doğru ifade makinenin çalıştırdığı programın Çince anlamayı ortaya çıkardığıdır. Nitekim insanlar da söz konusu olduğunda beynin ya da vücudun Çince anladığını söylemiyoruz, kişi olarak gördüğümüz aslında vücut ve bedenden bağımsız bir varlığa anlama özelliğimizi atfediyoruz.

Sanal kişiliklere örnek olarak Apple'ın Siri'si gibi dijital kişisel asistanları ya da bilgisayar oyunlarındaki kişilikleri verebiliriz. Kişisel asistanlar ya da oyun karakterleri kendi kişilik ve özelliklerine sahiptirler. Bunlar hem bilgisayardan hem de programdan bağımsızdırlar, zira program birden fazla kişiyi içerebilir. Bu cevap Sloman ile Croucher (Sloman ve Croucher, 1980), David Chalmers (Chalmers, 1996: 326) ve David Cole (Cole, 1991) tarafından savunulmuştur.

Chalmers Çin odasını içine cin girmiş bir bedene benzetir (Chalmers, 1996). Ona göre odadaki adam Çince konuşan bir bedenine içine girmiş cin gibidir. Onun Çince anlamaması odada Çince anlayan ikinci bir kişilik olmadığını göstermez. Cole Chalmers'ın ortaya attığı görüşü destekleyecek önemli bir noktaya dikkat çeker (Cole, 1991). Çin odasına kişisel sorular sorarsak, mesela boyunu, sevdiği yemeği, nerede doğduğu gibi; bu sorulara verilecek cevapların içerikleri düşünüldüğünde cevapların oda ya da içerideki adamın olmayacağı açıktır. Cevaplar algoritmaya kodlanan kişinin, yani odadaki kişi ve odanın cevapları olacaktır. Yani cevap veren kişi oda ve Searle'den başkası olacaktır. Cole görüşünü temellendirmek için Kore-Çin odası adını verebileceğimiz diğer bir argüman daha sunar. Çin-Kore odası düşünce deneyi Çin odası düşünce deneyi ile neredeyse aynıdır. Tek fark odada iki kapı vardır ve birinden Çince diğerinden de Korece sorular gelmektedir. İçerideki kişi gelen sorulara gene kitaplara bakarak uygun cevapları yazmaktadır. İlkece Çince sorulara verilen cevaplarla, Korece sorulara verilen cevaplar birbirinden farklı olabilir. Soruyu soran kişinin hem Korece hem de Çince bildiğini düşünelim. Çince ve Korece soruların cevapları birbiri ile uyumlu olmadığı için soruları soran kişi içeride biri Korece, biri de Çince bilen iki farklı kişi olduğu sonucuna varacaktır. Cole'a göre Çin-Kore odasının gösterdiği şey, odanın ya da içerideki kişinin değil, farklı zihinlerin cevap verdiğiidir. Dolayısıyla sistem iki farklı zihin ortaya çıkarmıştır, bunlar iki farklı sanal zihindir.

Sonuç olarak kanaatimce John Searle'un Çin Odası argümanı sanal bir zihnin ortaya çıkamayacağını, çevre ile etkileşen bir sistemde ya da beyni tam olarak simüle eden bir sistemde zihnin zuhur edemeyeceğini gösterememektedir. Bundan dolayı Çin odası argümanı güçlü YZ'nin imkânını ortadan kaldıramamaktadır.

## 5. GÖDELİ ARGÜMAN

Güçlü YZ aleyhinde getirilen en meşhur argümanlardan biri de Gödel'in birinci eksiklik teoreminden hareketle getirilen argümandır. Gödel'in birinci eksiklik teoremine göre aritmetik içerecek kadar zengin olan bir biçimsel sistem ya tutarsızdır ya da eksiktir (Gödel, 1931).

Teorimi biçimsel sistemlere aşına olmayanlar için açıklamakta fayda var.<sup>14</sup> Biçimsel sistemler sınırlı sayıda aksiyomdan hareketle, çeşitli çıkarım yasaları ile teoremlerin ispatlandığı sistemlerdir. Mantık, Aritmetik, Cebir, Bilgisayar bilimleri biçimsel sistemlerdir. YZ'nin algoritması da bir biçimsel sistemdir. Biçimsel sistemler tutarlı ya da tutarsız olabilir. Tutarlı biçimsel sistemlerde bir önermenin hem kendisi, hem de bu önermenin olumsuzu aynı anda doğru olamaz. Diğer bir deyişle bu sistemlerde çelişki bulunmaz. Diğer taraftan tutarsız sistemler çelişkili sistemlerdir, bundan dolayı bu sistemlerde hem bir önermeyi hem de bu önermenin tersini ispatlamak mümkündür. Tutarsız sistemlerde herhangi bir önerme doğru olacağı için bu sistemler pek yararlı değillerdir. Biçimsel sistemler ayrıca tamamlanmış ya da eksik olabilirler. Tamamlanmış sistemlerde kurulan herhangi bir cümlenin doğruluğu ya da yanlışlığı aksiyomlardan hareketle gösterilebilir. Diğer bir deyişle bu sistemlerdeki bütün doğru önermeler aksiyomlar yardımı ile gösterilebilir. Diğer taraftan eksik sistemler içerisindeki bütün doğru önermelerin doğruluğunu aksiyomlardan hareketle ispatlamak, yani bu önermeyi aksiyomlardan çıkarsamak mümkün değildir.

Gödel'in birinci teoremine göre aritmetik ya da onu içeren daha üst sistemler ya tutarsızdır, ya da eksiktir. Aritmetiğin tutarlı olduğunu varsayarsak o zaman aritmetik eksik olmalıdır. Buna göre aritmetik içinde doğru ama ispatlanamayan önermeler vardır. Bu önermelere Gödel cümleleri denir. Gödel cümlelerinden biri şöyledir: "Ben S biçimsel sistemi içinde ispatlanamam." (S ilgilendiğimiz biçimsel sistemin adıdır, Gödel için bu aritmetiktir). Gödel bu cümleyi aritmetik dilinde yazmayı başarır. Bu cümlenin doğru olduğunu ispatlarsak o zaman aritmetik tutarsızdır çünkü cümle ispatlanamayacağını iddia etmektedir. Yani cümle hem doğru hem yanlıştır. Diğer taraftan bu cümle ispatlanamazsa, o zaman doğrudur. Dolayısıyla elimizde doğru ama ispatlanamamış bir cümle vardır.

Felsefeci G.R. Lucas (Lucas, 1961) ve fizikçi R. Penrose (Penrose, 1989) Gödel'in birinci eksiklik teoreminden hareketle güçlü YZ'nin imkânını eleştirmişlerdir<sup>15</sup>. Lucas'a göre Gödel gibi bazı matematikçiler bazı Gödel cümlelerinin doğru olduğunu bilmektedirler. Diğer taraftan YZ bir çeşit biçimsel sistem olarak düşünülebileceği için bu Gödel cümlelerinin doğruluğunu aksiyomlarından çıkartamadığı için bu cümlelerin doğru olduğunu bilemez. Dolayısıyla insanın yapıp, YZ'lerin yapamayacağı şeyler vardır (Gödel cümlelerinin doğruluğunu bilmek gibi). Bundan Lucas'ın direk çıkar sonuç da insan zihninin bir çeşit yazılım olmadığıdır.

Gödel teoremini iyi bilmeyenler bu argümana şu şekilde itiraz edebilirler. Diyelim ki makine gerçekten Gödel cümlesini çıkaramıyor olsun, pekâlâ Gödel cümlesini makineye aksiyom olarak veririm. Bu durumda YZ Gödel cümlesini doğru bir cümle olarak algılayacaktır. Ancak bu cevap işe yaramaz zira söz konusu Gödel cümlesinin aksiyom olduğu bu yeni biçimsel sistemde başka Gödel cümleleri ortaya çıkacaktır. Bu şekilde Gödel'in birinci teoreminden kaçmak mümkün değildir.

Bu argüman hakkında ne söyleyebiliriz. Klasik eleştirileri incelemeden önce kendi kanaatimi paylaşmak istiyorum. Gödelci argüman başarılı olsa bile bence bu Güçlü YZ'nin imkansızlığını göstermez. Çünkü bu argümanın tek gösterebileceği şey insan zihninin olası bir makine zihninden farklı olduğudur. Ancak buradan güçlü YZ'nin imkânsızlığı gösterilemez. Böylesi bir iddia bilincin sadece insana benzeyen zihinlerde ortaya çıkabileceği varsayımı ile gösterilebilir. Ancak bu antroposentrik iddiayı temellendirmek kolay değildir. Bir makinenin Gödel cümlelerinin doğruluğu bilememesi, onun öznelik ya da hakkındalık gibi zihni

<sup>14</sup> Gödel teoreminin anlaşılır bir açıklaması için bakınız: (Nagel ve Newman, 2008)

<sup>15</sup> Muhtemelen Gödel'in kendisi de teoreminin YZ aleyhinde imaları olduğu kanaatini taşımaktaydı. Bakınız: (Gödel, 1995: 310).

özelliklere sahip olmadığı ya da olamayacağı anlamına gelmez. Gödelci argümanın gösterebileceği tek şey insan zihninin algoritmik YZ'lerden farklı olduğudur.

Lucasçı argümanla ilgili en genel eleştiri tutarsızlık cevabıdır. Tutarsızlık cevabının iki farklı türü vardır. Birinci tutarsızlık cevabını savunanlara göre insanlar tutarsızdırlar. Aynı anda çelişkili şeylere inanabilirler. Ancak bu durumda insanların yapıp, makinelerin yapamadığı bir şey yoktur. Gödel'in birinci eksiklik teorisine göre bir biçimsel sistem ya eksik ya da tutarsızdır. İnsanların tutarsız olmasından dolayı bu teoriye maruz kalmadıkları söylenemez, nitekim YZ tutarsız yapılırsa o da Gödel cümlelerine doğru diyebilir. Bu cevap fonksiyonalizmin savunucularından H. Putnam tarafından Lucas'a daha o makalesini yayınlamadan verilmiştir. İkinci tutarsızlık cevabı ise, insanların tutarsız olduklarını iddia etmek yerine tutarsız olup olmadığını bilmediğimizi söyler. Eğer insanların tutarlı olduğunu iddia etmezsek argüman çalışmaz. Ancak bu görüş savunucusuna göre tutarlı olduğumuzu göstermek mümkün değildir. Nitekim Gödel'in ikinci teoremine göre aritmetik içeren bir sistemin tutarlılığını kendi içinden göstermek mümkün değildir. Eğer biz bir biçimsel bir sistemle düşünüyorsak o zaman kendi kendimizin tutarlı olduğumuzu göstermemiz mümkün değildir. Öyle ise Gödelci argüman insanın tutarlılığını varsaydığı için başarılı olup olmadığı bilinemez.

Lucas bu itirazlardan ilkinin argümanı ortaya attığı makalede ele alır (Lucas, 1961). Lucas'a göre insan zihni tutarsız bir makine olamaz. O zaman zaman bizim tutarsız inançlara sahip olduğumuzu kabul eder, ancak ona göre bunlar hatalarımızdan kaynaklanır. Biz sistematik olarak tutarsız olamayız, böyle olsaydı herhangi bir önermeye inanır, tutarsız olduğumuz gösterildiğinde inançlarımızı düzeltmeye çalışmaz bu durumu böyle kabul ederdik. Nitekim tutarsız bir biçimsel sistemde bütün öncüller çıkarılabilir, dolayısıyla eğer biz tutarsız biçimsel bir sistem olsaydı bütün önermelere inanırdık. Dolayısıyla Lucas'a göre zihnimizin bir çeşit tutarsız programla çalıştığını iddia etmek makul değildir. Lucas'ın cevabı ilk bakışta makul gözükse de itirazı tamamen devre dışı bırakmaz. Zira biz sınırlı sayıda tutarsızlığa izin veren *tutarsızlık ötesi mantık*<sup>16</sup> ile çalışan programlar olabiliriz. Bu durumda bazı inançlarımız tutarsız olsa da, genel olarak tutarsızlıktan kaçınılırız. Tabii bu durumda Gödelci argüman başarısız olur.

Lucas ikinci itiraza iki farklı yöntemle cevap vermeye çalışır. Birincisi, Lucas Gödel'in ikinci teorisinin gösterdiği şey tutarlılığın içerden gösterilemeyeceğini, ancak başka sistemlere atıfla yani dış sistemler yardımı ile tutarlılık gösterilebileceğini hatırlatır (Lucas, 1990). Ancak bu hatırlatma yerinde olsa da sorunu çözmez, zira birincisi bu sefer tutarlılık ispatında kullanılan dış sistemin tutarlı olduğu gösterilmesi gerekir. İkincisi de insanın kendi tutarlılığını, kendi dışına çıkıp göstermesi pek mümkün gözükmemektedir. Dolayısıyla bu yaklaşımla sorunu çözmek pek makul gözükmemektedir. Lucas'ın ikinci yaklaşımı da düşünmek ve felsefe yapmak için belki de tutarlı olduğumuzu varsaymamız gerektiğini iddia etmektir (Lucas, 1976). Eğer anlamlı bir felsefe ya da bilim yapacak tutarlı çıkarımlar yaptığımızı düşünmemiz gerekir. Ancak felsefe yapmak ya da düşünmek için böyle bir varsayım yapmak zorunda olmamız, bu varsayımın gerçek olduğunu göstermez. Belki de yanılıyoruz. Dolayısıyla bu ikinci cevap stratejisi de pek makul gözükmemektedir.

Gödelci argümana getirilen bir diğer cevap da insanların da makinelerin Gödel cümleleri ile karşılaştıkları sorunlarla benzer sorunlara karşılaştıklarını iddia etmektir. Bu görüş Whiteley (Whiteley, 1962) ve Hofstadter (Dennett ve Hofstadter, 1981) tarafından savunulmuştur. Bu cevabı anlamak için Whiteley cümlesi olarak bilinen cümleye göz atalım: "Lucas tutarlı bir şekilde bu cümleyi ortaya atamaz." Eğer bu cümle doğruysa, ve Lucas bu cümleyi ortaya atarsa o zaman Lucas tutarsızdır. Eğer Lucas bu cümleyi ortaya atamazsa o zaman Lucas

<sup>16</sup> Tutarsızlık ötesi mantık için bakınız: (Priest, 2003)

eksiktir ve Gödel cümlesi karşısında biçimsel sistemlerin düştüğü duruma düşer. Bu cevap Martin ve Engleman tarafından eleştirilmiştir, Hofstadter'in argümanının daha detaylı analizi için makalelerine bakılabilir (Martin ve Engleman, 1990).

Lucas'ın argümanı idealize biçimsel sistemlere dayanmasından hareketle çok sayıda eleştiriye tabi olmuştur. Yer sınırlamamızdan dolayı bu idealize biçimsel sistemlerin insan ve makineleri tam yansıtmadığı şeklindeki eleştirileri bu yazımızda ele alamıyoruz<sup>17</sup>. İlgili okuyucu bu itirazlar ve onlara Lucas'ın getirdiği cevaplar için (Lucas, 1990)'a bakabilir.

## 6. YZ ÖZGÜR İRADE SAHİBİ OLABİLİR Mİ?

Güçlü YZ aleyhindeki en eski ve en yaygın argümanlardan biri makinelerin ve yazılımların doğaları gereği determinist oldukları ve bundan dolayı özgür irade sahibi olamadığı iddiasıdır. Diğer taraftan biz insanlar ise özgür iradeye sahibiz, dolayısıyla insanlar ve YZ birbirinden temel olarak farklıdır. Bu argüman YZ'nin öncülerinden Ada Lovelace (Lovelace 1842) tarafından savunulmuştur. YZ'lerin özgür irade sahibi olup olmadığı pratik açıdan çok önemlidir, zira bir YZ eylemlerinden sorumlu tutulacaksa özgür irade sahibi olmalıdır. Geleneksel olarak ahlaki ve hukuki sorumluluk için özgür irade önemli bir ön koşul olarak kabul edilir.

Determinizm bir sistemin geleceğinin o sistemi yöneten yasalar ve başlangıç ile sınır koşulları tarafından belirlendiğini iddia eden görüştür. Bu itiraz savunucusuna göre geleneksel olarak YZ'ler yapılar gereği onu kontrol eden algoritma ve girilen parametreler tarafından kontrol edilirler. Dolayısıyla YZ'ler deterministik sistemlerdir ve determinizm özgür irade ile çeliştiği için de YZ'ler özgür irade sahibi olamazlar.

Yukarıdaki argümandaki iki temel varsayım da eleştiriye açıktır. Birinci varsayım, yani YZ'lerin deterministik olduğu varsayımı genel olarak doğru olsa da ilkece bir YZ kuantum mekaniksel bir sisteme bağlanarak indeterminist yapılabilir. Ancak böylesi bir hamle YZ'yi irade sahibi yapmayacaktır, çünkü popüler kanaatin aksine indeterminizm doğrudan özgür iradeyi sisteme getirmez. Kuantum mekaniksel sistemler rastgele çalışır ve rastgelelik iradenin tam tersidir. Kuantum mekaniğinin yasaları gereği rastgele bir zamanda parçalanmış bir atom çekirdeğinin parçalanması deterministik bir süreç sonucunda gerçekleşmez, ancak bu parçalanmanın atomun iradesi ile gerçekleştiği iddia edilemez. Eğer eylemlerim benim kontrolüm dışında, rastgele gerçekleşirse kimse bu eylemlerimin benim irademle gerçekleştiğini söyleyemez.

Yukarıdaki argümandaki kritik varsayım kanaatimce determinist bir sistemin özgür irade sahibi olamayacağı varsayımdır. Bu varsayımı kabul etmeyen çok sayıda felsefeci vardır. "Determinizm doğru ise özgür irade ortadan kalkar mı?" sorusuna verilecek cevaba göre felsefeciler iki gruba ayrılır:

1. *Bağdaştırıcılık (Compatibilism)*: Bu görüşe göre, özgür irade determinizmle uyumludur. Determinizm doğru olsa bile biz özgür olabiliriz. Bir bireyin davranışları belirlenmiş olmasına rağmen, birey hala özgür olabilir.
2. *Bağdaşmazcılık (Incompatibilism)*: Bu görüşe göre, özgür irade determinizm ile uyumlu değildir. Eğer determinizm doğru ise, bireylerin özgür iradesi yoktur.

Dolayısıyla bağdaştırıcılık doğru ise yukarıdaki argüman başarısız olur ve pekâlâ YZ irade sahibi olabilir. Bu felsefecilere göre kişi yaptığı eylemden başkasını yapma gücü olmasa da özgür olabilir. Genelde, bağdaştırıcılar özgür olmayı kişinin iradesinin olayın ana kaynağı

---

<sup>17</sup> Şahsen bu eleştirilerin güçlü olmadığı kanaatini taşıdığım için bu yönde bir tercih yapıyorum. Ayrıca Gödelci argümana diğer bazı ünlü iki itiraz için bakınız (Benacerraf, 1967) ve (Lewis, 1969). Lucas'ın Lewis'e cevabı için bakınız: (Lucas, 1970).

olması olarak tanımlarlar. Diğer bir deyişle eylemlerin temel kaynağı kişi ise o zaman o kişi özgürdür görüşü benimsenir. Bu görüşü savunanlara göre özgürlük birden fazla alternatif olmasıyla ilgili değildir.

Bu ikinci yaklaşım genelde Frankfurtçu örnekler olarak anılan düşünce deneyleri ile motive edilir. Bu örnekler ismini Harry Frankfurt isimli felsefeciden alır. Frankfurt vakaları, bir kişinin bir eylemi yapması kaçınılmaz olduğu durumlarda bile sezgisel olarak özgür olduğunu göstermeye çalışır. Bir tane örnek verelim:

Enis Hegel'i vurmaya planlamaktadır. Ahmet de Hegel'den nefret eder, o yüzden Enis'in beynine özel bir çip yerleştirir. Eğer Enis Hegel'i öldürmekten vazgeçerse, çip devreye girip Enis'in Hegel'i öldürmesini sağlayacaktır. Diğer taraftan eğer Enis, karar değiştirmez ve Hegel'i öldürürse, çip hiç devreye girmez. Enis diyelim ki gitti ve Hegel'i öldürdü ve çip devreye girmedi. Enis özgür iradeye sahip miydi? Yaptığı suç muydu ve bundan sorumlu muydu? Bu soruların cevabı evetmiş gibi duruyor. Ancak diğer taraftan Enis'in Hegel'i öldürmekten başka çaresi yoktu. Dolayısıyla bu düşünce deneyi doğrusa o zaman kişi yaptığı eylemden başkasını yapma gücü olmasa da özgür olabilir.

Eğer Frankfurtçu örnekler başarılı ise bağdaştırıcılık makuldür ve YZ irade sahibi olabilir. Bağdaştırıcılık ve bağdaşmazcılık tartışması çok çetrefillidir. Tabii insanın da özgür iradesi olmadığını iddia eden düşünce okulları da mevcuttur. Dolayısıyla bu argümanın başarılı olup olmadığına karar vermek için insanın özgür iradesi olup olmadığı, varsa bunun determinizm ile çelişip çelişmediğine karar vermek lazım. Bu konuda ilgili okuyucular (Beebee, 2013)'e bakabilir. Ancak YZ özgür iradeye sahip olmasa bile bu genel olarak Güçlü YZ'nin imkânsız olduğunu göstermez. Pek ala bilinçli ama iradesiz YZ olabilir, bu durumda bu YZ'lere ahlaki sorumluluk yükleyemsek de, bunlara karşı bizim ahlaki sorumluluklarımız olabilir.

## 7. SOLAN ÖZNELİK DÜŞÜNCE DENEYİ

Bu bölüme kadar Güçlü YZ'nin aleyhinde verilen argümanları inceledik. Makalenin bu son bölümünde Güçlü YZ'nin imkânı lehinde geliştirilen bir argümanı ele almak istiyorum. Bu bölümde ele alacağımız düşünce deneyi David Chalmers tarafından savunulmuştur ve solan öznelik düşünce deneyi olarak bilinmektedir (Chalmers, 1995).

Birebir nöronları taklit eden silikon çipler ürettiğimizi varsayalım<sup>18</sup>. Bu çiplerin yapısı nöronlardan farklı olsa da aynı elektrik sinyaline maruz kaldığında nöronlarla aynı sinyal çıktısını verdikleri için işlevsel olarak nöronlarla aynıdırlar. Ahmet isimli birini düşünelim. Ahmet'in beyindeki nöronları teker teker yavaşça bu silikon çiplerle değiştirdiğimizi varsayalım. Çiplerin elektrik sinyallerine tepkileri nöronlarla aynı olduğu için diğer nöronlar bu değişimin farkında olmayacaktır. Ayrıca vücudumuz da beyinden gelen elektrik sinyalleri tarafından yönetildiği için, nöronların çiplerle değişmesi kişinin davranışlarını değiştirmeyecektir. Dolayısıyla dışarıdan Ahmet'in davranışlarında herhangi bir değişiklik fark edilemeyecektir. Süreç sonucunda Ahmet'in bütün nöronlarının değiştiğini varsayalım, bu durumda Ahmet'in beyni tamamen silikondan yapılmış bir YZ'ye dönüşecektir. Zira nöronları çiplerle değiştirmek yerine silikon beyni doğrudan oluşturabiliriz.

Güçlü YZ'nin imkanını reddedenler Ahmet'in öznelik ve hakkındalık gibi zihin hallerine sahipken, silikon beyinli Ahmet'in bu zihin hallerine sahip olmadığını iddia edeceklerdir. Ancak bu iddia şu önemli soruyu gündeme getirecektir: Ahmet tamamen biyolojik beyinden, tamamen silikon beyne geçişte hangi noktada ve nasıl zihin hallerini, özelde de özneliğini (qualiyı) kaybetti? Chalmers'a göre bu soruya cevap olarak iki olası senaryo sunulabilir. Birincisi Ahmet'in yavaş yavaş özneliğini kaybettiği söylenebilir, bu solan öznelik durumudur.

<sup>18</sup> Aralık 2019'da Nature'da çıkan bir makalede böylesi çiplerin üretildiği duyuruldu.

Bu senaryoda Ahmet davranışları hiç değişmese de nöronları değiştikçe, renk, ses, koku algısını kaybetmektedir. Mesela bir parlak kırmızı bir cisme bakıyorsa, parlak kırmızı rengin zihinde bıraktığı resim gittikçe solmakta, en nihayetinde tamamen kaybolmaktadır. Sesler gittikçe yavaşlamakta ve nihayetinde kaybolmaktadır. Ancak Ahmet sesleri artık tam duyamasa da, renklerin tonlarını tam fark edemese de daha önceden kendisine bu sesler ya da renklerle ilgili sorulara aynı cevapları verebilmektedir. Mesela sesi zar zor duysa da, önceden sesi net duyduğunu söylüyorsa gene aynı şeyi söylemekte, duyamadığı halde sesle ilgili duyarken verdiği cevapları verebilmektedir. Chalmers'a göre bu durum mantıksal olarak mümkün olsa da fiziki olarak mümkün olmayacak derecede saçmadır.

İkinci olası senaryo ise ilk başta nöronlar çiple değişirken bilinçte bir değişiklik olmazken, belli bir nöron silikon çiple değiştirildiği sırada aniden bilincin kaybolduğunu iddia etmektir. Elbette gene bu durum dışarıdan algılanan davranışlara yansımaya, bu nöron değişim gerçekleşmeden önce ve sonra Ahmet'in davranışları aynı olacaktır. Ancak bu noktada Ahmet bilinçli bir kişiden bilinçsiz bir zombiye dönüşecektir. Chalmers'a göre bu ikinci alternatif de makul değildir, zira doğada böyle ani ve radikal değişimler gözlenemez. Dolayısıyla ona göre iki senaryo da makul değildir. Ancak iki senaryo da geçersizse o zaman silikon beyinli Ahmet'in de öznelliğe, dolayısıyla bilince sahip olduğunu kabul etmemiz gerekir. Bu da Güçlü YZ'nin mümkün olduğu anlamına gelmektedir.

Chalmers'ın solan öznellik düşünce deneyi Güçlü YZ'nin imkânı lehinde güçlü bir argüman olsa da bence en az iki açıdan eleştirilebilir. Birincisi silikon beyinli Ahmet'in zihinsel hallere sahip olmasının Güçlü YZ'nin imkânını göstermediği, zira bu Ahmet'in biyolojik beyinli Ahmet'in ruhuna sahip olduğu için zihin hallerine sahip olduğu söylenebilir. Bu itirazda biyolojik bir canlıdan elde edilmeyen bir YZ'nin zihin hallerine sahip olmadığı söylenebilir. Diğer bir deyişle düalizm doğruysa, solan öznellik düşünce deneyi Güçlü YZ'yi desteklemek zorunda değildir. İkincisi, Chalmers'ın iddiasının aksine doğada ani geçişler olduğu iddia edilerek onun imkânsız gördüğü ikinci senaryonun imkânı savunulabilir. Fizikte birinci dereceden faz geçişleri olarak bilinen durumlarda fiziksel büyüklüklerde ani değişiklikler gözlemlenebilmektedir. Faz geçişleri genelde zuhur eden özelliklerin açığa çıktığı durumlardır. Eğer zihin halleri zuhur eden özelliklerse, bunların nöronlar çiplerle değiştirilirken aniden kaybolmaları pek şaşırtıcı olmayabilir.

## SONUÇ

Bu makalemizde YZ'lerin de öznellik, irade, hakkındalık gibi zihin halleri olabileceğini savunan güçlü YZ lehinde ve aleyhindeki bazı felsefi argümanları ve bu argümanların eleştirilerini ele aldık. Her argümanların karşısında geliştirilen eleştirilerin en azından bazılarının başarılı olma ihtimali olduğunu, dolayısıyla bu hiçbir argümanın güçlü YZ'nin imkânı ya da imkânsızlığını gösteremediğini söyleyebiliriz. Bu konuda daha kesin bir kanaat geliştirmek için insan zihninin yapısını ve nasıl çalıştığını daha iyi anlamamız gerekmektedir. Bu ise nöroloji ve bilişsel bilimler alanındaki bilgimizin artması ile gerçekleşecektir. Dolayısıyla kanaatimce bu aşamada güçlü YZ'nin imkânı ile ilgili en makul yaklaşım geçici olarak bu konuda bilinmezci bir tavır takınmak ve "bekleyelim görelim" demektir.

## KAYNAKÇA

Benacerraf, P. (1967). God, the Devil, and Gödel, *Monist* 51:9-32.

Block, N. (1981). Psychologism and behaviorism, *The Philosophical Review* 90:5-43.

- Boden, M.A. (1988). *Computer Models of the Mind*. Cambridge: Cambridge University Press.
- Boden, M.A. (2016). *AI: Its Nature and Future*. Oxford: Oxford University Press.
- Bringsjord, S. (1995). Could, how could we tell if, and why should—androids have inner lives?, K. Ford, C. Glymour, ve P. Hayes (ed.), *Android Epistemology* (93–122). Cambridge, MA: MIT Press.
- Chalmers, D. (1995) Absent Qualia, Fading Qualia, Dancing Qualia, Metzinger, T. (ed.) *Conscious Experience*, Imprint Academic.
- Chalmers, D. (1996). *The Conscious Mind*. Oxford: Oxford University Press.
- Churchland, P. ve Churchland, P. (1990). Could a machine think?. *Scientific American*. 262(1): 32–37.
- Copeland, J. (2002) The Chinese Room from a Logical Point of View. Preston, J. ve M. Bishop (ed.) *Views into the Chinese Room: New Essays on Searle and Artificial Intelligence* (104–122). New York: Oxford University Press.
- Cole, D. 1991. Artificial Minds: Cam on Searle. *Australasian Journal of Philosophy*, 69: 329–33.
- Crane, T. (1996). *The Mechanical Mind: A Philosophical Introduction to Minds, Machines and Mental Representation*. London: Penguin.
- Dennett, D.C. ve Hofstadter, D. R. (1981). *The Mind's I: Fantasies and Reflections on Self and Soul*. New York: Basic Books.
- Fodor, J. (1991). Yin and Yang in the Chinese Room. D. Rosenthal (ed.). *The Nature of Mind*. New York: Oxford University Press.
- Gödel, K. (1931). Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme I, *Monash. Math. Phys.* 38: 173-198.
- Gödel, K. (1995). *Collected Works III*. New York: Oxford University Press.
- Harnad, S. (1991). Other bodies, other minds: A machine incarnation of an old philosophical problem, *Minds and Machines* 1: 43–54.
- Haugeland, J. (2002). Syntax, Semantics, Physics. Preston, J. ve M. Bishop (ed.) *Views into the Chinese Room: New Essays on Searle and Artificial Intelligence* (379–392). New York: Oxford University Press.
- Heil, J. (2019). *Philosophy of Mind*. New York: Routledge.
- Kim, J. (2010). *Philosophy of Mind*. New York: Westview Press.
- Kurzweil, R. (2002) Locked in his Chinese Room. Richards, J. W. (ed.). *Are We Spiritual Machines: Ray Kurzweil vs. the Critics of Strong AI*. (128-171) Seattle: Discovery Institute.
- Lewis, D. (1969). "Lucas against Mechanism," *Philosophy* 44:231-3.
- Lovelace, A. A. (1842). Translator's notes to L. F. Menabrea's 'Sketch of the analytical engine invented by Charles Babbage, Esq.'. Bowden, B. V. (ed.). (1953). *Faster than Thought: A Symposium on Digital Computing Machines*. New York: Pitman Publishing: 362-408.

- Lucas, J. R. (1961). Minds, Machines and Gödel, *Philosophy* 36:112-127.
- Lucas, J. R. (1970). Mechanism: A Rejoinder, *Philosophy* 45:149-51.
- Lucas, J. R. (1976). This Gödel is killing me: A rejoinder, *Philosophia* 6:145-8.
- Lucas, J. R. (1990). Mind, machines and Gödel: A retrospect. Turing Conference Brighton 6 Nisan 1990.
- Nagel, E. ve Newman, J. (2008). Gödel Kanıtlanması. İstanbul: Boğaziçi Üniversitesi Yayınları.
- Penrose, R. (1989). *The Emperor's New Mind*. Oxford: Oxford University Press.
- Priest, G. (2003). Inconsistent Arithmetic: Issues Technical and Philosophical, *Trends in Logic: 50 Years of Studia Logica*, F. Hendricks ve J. Malinowski (ed.), Dordrecht: Kluwer Academic Publishers.
- Rey, G. (1986). What's Really Going on in Searle's "Chinese Room", *Philosophical Studies*, 50: 169–85.
- Searle, J. R. (1980). Minds, brains, and programs, *Behavioral and Brain Sciences*. 3 (3): 417–424.
- Sloman, A. ve Croucher, M. (1980). How to turn an information processor into an understanding. *Brain and Behavioral Sciences* 3: 447–8.
- Turing, A. (1950). Computing Machinery and Intelligence, *Mind*, 59: 433–460.
- Whiteley, C. (1962). Minds, Machines and Gödel: A Reply to Mr. Lucas, *Philosophy* 37:61-62.