

T.C.

ANKARA YILDIRIM BEYAZIT ÜNİVERSİTESİ

SOSYAL BİLİMLER ENSTİTÜSÜ

YAPAY ZEKA FELSEFESİNDE BİLİNÇ PROBLEMİ

DOKTORA TEZİ

Mehtap DOĞAN

FELSEFE ANABİLİM DALI

ANKARA, 2020

T.C.
ANKARA YILDIRIM BEYAZIT ÜNİVERSİTESİ
SOSYAL BİLİMLER ENSTİTÜSÜ

YAPAY ZEKA FELSEFESİNDE BİLİNÇ PROBLEMİ

DOKTORA TEZİ

MEHTAP DOĞAN
FELSEFE ANABİLİM DALI

PROF. DR. MEHMET VURAL
DANIŞMAN

ANKARA, 2020

BEYAN

Bu tez içerisindeki bütün bilgilerin akademik kurallar ve etik davranış çerçevesinde elde edilerek sunulduğunu beyan ederim. Ayrıca bu kurallar ve davranışların gerektirdiği gibi bu çalışmada orijinal olmayan her tür kaynak ve sonuçlara tam olarak atıf ve referans yaptığımı da beyan ederim; aksi takdirde tüm yasal sorumluluğu kabul ediyorum.

Adı Soyadı: Mehtap DOĞAN

İmza:

İTHAF

Berra'ma…

TEŞEKKÜR

Teze başladığım andan sonlandırana kadar her zaman yapıcı bir tavırla yol göstericim olan, bilgi ve deneyimlerini aktarmaktan hiçbir zaman çekinmeyen sayın tez danışmanım, değerli hocam Prof. Dr. Mehmet VURAL'a emeklerini ödeyemeyeceğimin farkında olarak çok teşekkür ediyorum.

Tez dönemi boyunca düşüncelerimi şekillendirmem ve ifade etmem konusunda yaptığı önerilerle zihnime ışık tutan, akademik destek ve yardımlarını esirgemeyen sayın hocam Doç. Dr. Murat ARICI'ya kalpten teşekkür ediyorum.

Kıymetli vakitlerini ayırarak tezin geliştirilmesine yönelik olumlu eleştiri ve katkılarda bulunan jüri üyeleri sayın hocalarım Prof. Dr. Halil Rahman AÇAR, Prof. Dr. Mehmet Hilmi DEMİR ve Prof. Dr. Levent BAYRAKTAR'a ve faydalı yönlendirmelerinden ötürü sayın Prof. Dr. Musa Kazım ARICAN'a teşekkürlerimi sunuyorum.

Tez yazım sürecinde en büyük moral kaynağım olan canım kızım Berra'ya, onunla geçirmem gereken zamandan çaldığım her an için özür dileyerek, teşekkür ediyorum. Ayrıca annem ve babama hem bana olan desteklerinden hem de benim eksikliğimi Berra'ya hissettirmemek adına gösterdikleri çabadan ötürü gönülden teşekkür ediyorum. Ve tez teslim aşamasında kaybettiğim, beni büyüten insan, ikinci annem babaannemi rahmetle anıyor ve kendisine teşekkür ediyorum.

Son olarak, adları burada zikredilmemiş, emeği geçen tüm hocalarıma ve arkadaşlarıma şükranlarımı sunuyorum. Bu çalışmadaki tüm eksiklikler yalnızca bana aittir.

ÖZET

YAPAY ZEKA FELSEFESİNDE BİLİNÇ PROBLEMİ
DOĞAN, Mehtap

Doktora, Felsefe Bölümü
Tez Yöneticisi: Prof. Dr. Mehmet VURAL
Aralık 2020, ix+222 sayfa

Bu çalışma, bilinçli bir yapay zekanın imkânı üzerine bütüncül bir bakış açısı oluşturmayı hedefleyen felsefi bir soruşturmadır. Soruşturmanın planı "Bütüncül bir yapay bilinç teorisi, doğal bilince yönelik kuşatıcı bir yaklaşım üzerine inşa edilmelidir." varsayımına bağlı olarak oluşturulmuştur. Bu bağlamda, öncelikle melez bir kavram olan yapay bilinci meydana getiren bileşenler olan yapay zeka ve bilinç mefhumları ayrık problematikler olarak ele alınmış, daha sonra ise yapay bilince yönelik bakış açıları incelenmeye çalışılmıştır. Çalışmanın başlangıcında bilinç, fenomenal deneyim olarak tanımlanmış ve çalışma boyunca kabul edilebilir bir bilinç yaklaşımının, deneyimin içkin bağıntıları olan deneyimleyen ve deneyimlenenin doğasını göz ardı etmemesi gerektiği fikri ayrıştırıcı bir ölçüt olarak kullanılmıştır. Makine bilincinin etik, hukuki ve epistemolojik meselelerle de ilgisi tartışılarak bilinçli bir makine inşasının pratikte yol açabileceği sorunsallara işaret edilmiştir. Son adımda ise "Bilinçli bir makine mümkün müdür?" sorusu basamaklar halinde cevaplanırken makine bilincinin insan bilinciyle karşılaştırılamaz yapıda yeni bir tür bilinç olarak var olabileceği tezi, felsefi olarak temellendirilmeye çalışılmıştır.

Anahtar Kelimeler: Yapay Zeka, Bilinç, Yapay Bilinç, Makine Bilinci, Makine Etiği.

ABSTRACT

THE PROBLEM OF CONSCIOUSNESS IN PHILOSOPHY OF ARTIFICIAL INTELLIGENCE

DOĞAN, Mehtap
PhD., Department of Philosophy
Supervisor: Prof. Mehmet VURAL
December 2020, ix+222 pages

This study is a philosophical investigation aiming the construction of a holistic view on the possibility of conscious artificial intelligence. The plan of this investigation is formed relying on the supposition that is "A holistic theory of artificial consciousness should be constructed upon a holistic account on natural consciousness". In this manner, to begin with, artificial intelligence and consciousness as the components of the artificial consciousness, which is a hybrid concept, are handled seperately. Then, it is tried to examine the details of the approaches on artificial consciousness. Consciousness is defined as phenomenal experience in the introduction of this research. Therefore, all discussions in this study about the nature of consciousness are evaluated according to the criterion that is "the experiencer and the experienced object as intiristic correlations of experience should be taken into account by any supportable theory on consciousness". Additionally, by discussing on the relationship between machine consciousness and ethical, legal, epistemological issues, the problematical issues in the practical results of the construction of a conscious machine are indicated. In the last step, the question of "Is a conscious machine possible?" is tried to be answered step by step and it is claimed that machine consciousness is possible as a new version of consciousness that is incommensurable with human consciousness.

Keywords: Artificial Intelligence, Consciousness, Artificial Consciousness, Machine Consciousness, Machine Ethics.

İÇİNDEKİLER

TEZ KABUL VE ONAY SAYFASI ..... **i**
BEYAN ..... **ii**
İTHAF ..... **iii**
TEŞEKKÜR ..... **iv**
İÇİNDEKİLER ..... **vii**
1. GİRİŞ **1**
1.1. Yapay Zekanın Fikrî Gelişimi 6
1.2. Yapay Zekaya Mantıktan Gelen Katkılar 10
1.3. Bağımsız Bir Bilim Dalı Olarak Yapay Zeka 15
1.4. Yapay Zeka ve Sinirbilim 18
1.5. Bilinçli Yapay Zekalar İçin Büyük bir Adım: Yapay Öğrenme 22
1.6. Yapay Zekadan Süper Zekalara 23
2. BİLİNÇ PROBLEMİ: YAKLAŞIMLAR VE TEORİLER **30**
2.1. Monizm (Tekçilik) 31
2.1.1. İdealizm 32
2.1.2. Materyalizm (Maddecilik) 34
2.2. Düalizm 50
2.2.1. Töz Düalizmi 50
2.2.2. Nitelik Düalizmi 55
2.3. Panpsişizm 57
2.4. Agnostik Gizemcilik 58
2.5. Bilince Yönelik Temel Yaklaşımların Değerlendirilmesi 59
3. BİLİNCE GÜNCEL YAKLAŞIMLAR **61**
3.1. Çoklu Taslaklar Kuramı: Dennett Örneği 61
3.2. Biyolojik Doğalcılık: Searle Örneği 66
3.3. Doğalcı Düalizm: Chalmers Örneği 68
3.4. Üst Düzey Temsil Kuramları 71
3.5. Güçlü veya Dışsalcı Temsilcilik: Dretske Örneği 73
3.6. Duyu-Motor Kuramı: O'Regan Örneği 75
3.7. Nörofenomenoloji: Varela Örneği 77
3.8. Refleksif Monizm: Velmans Örneği 79
3.9. Sanal Gerçeklik Kuramı- Benlik Tüneli: Metzinger Örneği 82
4. BİR ARAŞTIRMA ALANI OLARAK YAPAY BİLİNÇ VE YAPAY BİLİNCE FARKLI YAKLAŞIMLAR **85**

4.1 Yapay Bilinç Teorileri 85
4.1.1. Davranış Temelli Makine Bilinci 97
4.1.2. Bilişsel Temelli Makine Bilinci 98
4.1.3. Nörobilimsel Temelli Makine Bilinci 100
4.1.4. Fenomenal Makine Bilinci ve Sentetik Fenomenoloji 112
4.2. Yapay Bilincin Diğer Alanlarla İlişkisi 116
4.3. Yapay Bilinç Çalışmalarının Değeri ve Geleceği 118
4.4. Yapay Bilince Epistemolojik Yaklaşım: Bir Makinenin Bilinçli Olduğunu Bilebilir miyiz? 122
5. YAPAY BİLİNÇ ÜZERİNE SOSYAL, ETİK VE YASAL HUSUSLAR 128
5.1. Makine Etiğinin Felsefi ve Sosyal Temelleri 128
5.2. Makine Etiği İçin Yöntem Tartışmaları 137
5.3. Mantık ve Yapay Ahlak İlişkisi 147
5.4. Makine Hakları ve Makine Etiğine Hukuki Yaklaşım 149
5.5. Bilinç-Ahlak İlişkisi Bağlamında Makine Etiği 155
6. YENİ BİR BİLİNÇ YAKLAŞIMI TEMELİNDE YAPAY BİLİNCİN İMKÂNI 159
6.1. Varolan Bilinç Teorilerinin Varsayımları Üzerine 159
6.2. Alternatif Bir Bilinç Yaklaşımı: Fenomenal Gerçekçilik 163
6.2.1. Bilincin En Kolay Problemi 166
6.2.2. Bilincin Kolay Problemi 166
6.2.3. Bilincin Zor Problemi 168
6.2.4. Bilincin En Zor Problemi 175
6.2.5. Bilincin Doğasına Yönelik Problemlerin Basamakları 180
6.2.6. Alternatif Bir Yapay Bilinç Yaklaşımı 185
6.3. Yapay Uygulamalar 190
7. SONUÇ VE ÖNERİLER 195
8. KAYNAKLAR 204
9. EKLER 219
Ek: 1: ASIMO 219
Ek: 2: COG 220
EK:3: ÖZGEÇMİŞ 223

1. GİRİŞ

Korku batıl inançların ana kaynağıdır
ve zalimliğin de ana kaynaklarından biridir.
Korkuyu fethetmek aklın başlangıcıdır.
Bertrand Russell

Bu çalışmanın temel konusu günümüzün en önemli araştırma ve tartışma alanlarının başında gelen yapay zekada bilinç problemidir. Yapay zekada bilinç probleminin tüm yönleriyle ele alınması, hem yapay zeka hem de bilinç için bütüncül bir kavramayı gerektirmektedir. Hem bilinç hem de yapay zeka, felsefenin özel ve çetrefilli problemleridir. Bilincin empirik yöntemlerle açıklanamayan nitelikleri, gelişimi hızla devam eden yapay zekanın ise ulaşabileceği seviye felsefi sorgulamaya oldukça muhtaçtır. Yapay bilinç kavramı, yapay zeka ve bilinç arasında kurulması muhtemel bir ilişkinin ürünüdür. Bu bağlamda, bu araştırma boyunca öncelikli amaç bu ilişkinin olasılığını tartışırken bilinç ve yapay bilincin dayandığı kavramsal temelleri sağlam ve net olarak göstermek olacaktır.

Bilincin sahip olduğu tarihsel, psikolojik, felsefi ve kavramsal temeller bir yana bırakıldığında günlük dilde bilinçlilik, *uyanıklık hâli, uyaranlara tepki verme yeteneği, dikkat* ya da *farkındalık* gibi pratik ifadelere tekabül etmektedir. Ancak zihnin sahip olduğu karmaşık yapı, bilincin tanımının kesin ve net ifadelerle yapılmasına engel oluşturmaktadır. Öte yandan, deneysel bilinç çalışmaları bilincin doğasına değil, bilinçlilik ve bilinçsizlik durumlarının farklarına odaklanmaktadır. Koma hâlindeki hastaların bilinçlilik düzeyleri gelişmiş beyin görüntüleme teknikleriyle ölçülmekte, uyanık bir hastanın genel anestezi ile uyutulması anında beyninde meydana gelen değişiklikler saptanmaya çalışılmaktadır. Bilincin kaybedilmesinin beyindeki bir düğmenin kapalı konuma getirilmesiyle gerçekleştiğini gösteren bir kanıta henüz ulaşılamamış olması, bilincin sahip olduğu kompleks yapının çözümlenmesinin çok da kolay olmadığı sonucunu doğurmaktadır.

Bilinç problemini zor ve karmaşık bir yapıya büründüren ve kimi çevrelerce bilincin "muamma bir fenomen" şeklinde addedilmesine yol açan şey; beynin duyumsama, konuşma, algılama, hafıza, problem çözme vs. gibi işlevleri nasıl meydana getirdiğini keşfetmek değil,

nasıl bilinçli olduğumuzu açıklayabilmektir. Bilinci incelemek beyin ve davranış arasındaki ilişkiyi, nöronlar arasındaki bağlantıyı, beynin hangi bölgesinin bedenin hangi fonksiyonundan sorumlu olduğunu açıklayan haritayı incelemekten daha fazlasını gerektirmektedir. Zira bilinç, kendi varlığımızın doğasına yönelik derin bir problemdir. Biyolojik, fizyolojik, nörolojik boyutlarıyla birlikte bilinç aynı anda hem epistemolojik, hem ontolojik hem de metafizik bir problemdir. Bilincin ontolojisi hakkında konuşmak, zihnin ait olabileceği temel metafizik kategorileri keşfetmeyi de gerekli kılmaktadır.

İnsan hayati faaliyetlerini gerçekleştirirken çevresiyle etkileşim içerisinde olan bir varlık olmasının yanı sıra kendi varlığının farkında olan, kendi varlığını duyumsayan ya da kendi varlığını deneyimleyebilen de bir varlıktır. Her sabah uyandığımızda yeniden ve kesintisiz olarak devam eden zihinsel durumlarımızın varlığı bize öznel gerçekliğimizi sürekli olarak kendiliğinden sunmaktadır. Öznel gerçeklik, yaşamış olduğumuz tüm deneyimlerimizin bizzat sahibi olduğumuz fikrine karşılık gelmektedir. Renk, şekil, boyut algıları, hisler, sesler, tatlar, gözümüzü kapatıp dondurmamızı yerken çocukken bindiğimiz kırmızı bisikleti görmemiz, ıssız bir yolda burnumuza gelen ıhlamur kokusunun lise aşkımızı hatırlatıp ettirdiği tebessüm vs. hepsi öznel gerçekliğimizin içeriğini oluşturan deneyimlerimizdir. Deneyimler sürekli bir akış hâlinde, aralıksız ve değişen bir sarmalın temsili gibidirler. Dinamik, kısa süreli, akışkan karakterdeki bu deneyim nehrini, hayatımızın her anında zamansal ve mekânsal süreklilik ile birlikte ortaya çıkan bütünleşik öznel gerçekliğimizi göz ardı ederek anlamamız ve bilinçliliğimizi tüm yönleriyle şeffaf bir şekilde kavramamız mümkün görünmemektedir. Öznel deneyimler, sürekli değişen içeriklerine rağmen kesintisiz bir akış hâlindedir ve bu da insanın bilinçli yaşamını ya da başka bir deyişle “ben” yaşantısını oluşturmaktadır.

Varlığını doğrudan duyumsadığımız, hissettiğimiz ya da algıladığımız bilincin gizemi ya da zorluğu bu yönüyle nesnel varlığından değil, öznel ve niteliksel doğasından kaynaklanmaktadır. Zira nesnel ölçüm âletleriyle öznel ve niteliksel deneyimlerimizin yapısı hakkında açıklama yapmaya çalışmak kendi içerisinde çelişik bir uğraşmış gibi görünmektedir. Bu bağlamda bilincin gizemi, onun modern bilimin dünya görüşü içerisinde nereye konumlandırılacağının bilinememesinden ileri gelmektedir. Bilinci açıklayan hiçbir tabiat kanununun bilinmiyor oluşu, bilimin bilinci ya reddetmesi ya da indirgemeci bir tavırla ele alması sonucunu beraberinde getirmektedir. Lakin bütüncül bir bilinç kuramı

nesnel, fiziksel araç ve yöntemlerle çözümlemeyen öznel ve niteliksel bilincin ontolojik konumu hakkında tatmin edici bir açıklama sunmalıdır.

Zihin felsefesindeki bilinç tartışmalarına nüfuz edebilmek adına bilinçle ilgili bu literatürde önemli yer sahibi olan bazı özel kavramların altının çizilmesinde fayda vardır. Bu tartışmalarda yer alan bazı özel terim ve kavramlar bu çalışma boyunca sıklıkla kullanılacağından çalışmanın henüz giriş bölümünde bilincin niteliklerine yönelik yapılan bazı ayrım ve tanımların açıklanmasının gerekli olduğu düşünülmüştür. Özellikle bilincin öznel ve niteliksel yapısıyla ilgili argümanlar ile literatüre kazandırılan önemli kavramsallaştırmalar bulunmaktadır.

Bilincin zihin felsefesindeki önemli kullanımlarından birisi canlı bilinci[1] ve durum bilinci[2] arasında yapılan ayrıma dayanmaktadır. Bir canlının bilinçli olması ile bir zihin durumunun bilinçli olması birbirinden farklıdır. Canlı bilinci, durum bilincini zorunlu olarak gerektirmez. Canlı bilinci, bir organizmanın uyanık olması ve duyusal uyarıcılara tepki vermesidir ve bunun için bilinçli bir duruma sahip olmak zorunlu değildir (Rosenthal, 1986: 351).

Bilinç, öznel doğası gereği çağdaş bilimsel paradigmanın çekingen yaklaştığı bir fenomendir. Thomas Nagel, 1974 yılında yayınlanan "What is it Like to be a Bat?" makalesinde bilincin öznelliğine vurgu yaparak bilincin özgün bir fenomen olmasındaki temel sebebi "gibi olmaklık" şeklinde ifade etmiştir. Buna göre, bir "x" gibi olmak durumu, yalnızca birinci şahıs bakış açısından bilinebilir; "gibi olmaklık" üçüncü şahıs bakış açısından kavranamaz. Nagel'a göre, "Eğer deneyimin öznel karakteri tek bir bakış açısı tarafından kavranabiliyorsa daha fazla bir nesnelliğe geçiş bizi fenomenin gerçek doğasına ulaştırmaz." (Nagel, 1974: 445).

Nagel'in üçüncü şahıs bakış açısıyla öznel deneyimin doğasına erişilemeyeceği iddiası, Joseph Levine'in 1983 tarihli "Materialism and Qualia: The Explanatory Gap" makalesinde fiziksel nitelikler ve bu fiziksel niteliklerin deneyimi arasında kapatılamayacak bir "açıklama gediği"[3] olduğu teziyle de farklı bir yönden tartışılmıştır. Bu tezdeki temel iddia, acının beyindeki fiziksel korelasyonları ile acının deneyiminin doğaları gereği birbirinden

[1] *Creature consciousness.*
[2] *State consciousness.*
[3] *Explanatory gap.*

farklı olduğudur. Bu sebeple, bilincin tüm fiziksel korelasyonları bilinse dahi bu, deneyimin kendisini açıklamak için yeterli olmayacaktır.

David Chalmers ise 1995 yılında yayınlanan "Facing up to the Problem of Consciousness" makalesinde bilincin kolay problemleri ve bilincin zor problemi olarak yaptığı ayrımla bilincin neden hem felsefe için hem de bilim için "özel" bir fenomen olduğunu başka bir açıdan ortaya koymuştur. Bilincin kolay problemleri, bilimin standart yöntemleriyle çözümlenebilecek ve bilişsel ya da nöral mekanizmalarla açıklanabilecek problemlerdir. Bilincin zor problemi ise alışılagelmiş bilimsel yöntemlerle açıklanmaya direnç gösteren "deneyim" problemidir (Chalmers, 1995: 203). Bu problem, diğer problemlerden farklı olarak öznel bakış açısından bilinebilirdir.

Chalmers'ın kolay problemler ve zor problem ayrılığıyla işaret ettiği bilincin farklı nitelikleri, Ned Block tarafından da aynı yıl içerisinde fenomenal bilinç[4] ve erişim bilinci[5] olarak kategorize edilmiştir. Block bu kavramsallaştırmayı şöyle ifade etmektedir: "Fenomenal bilinç deneyimdir; bir durumun fenomenal bilinçli yanı, o durumda olmanın ne gibi olduğudur. Karşıt olarak, erişim bilincinin işaret ettiği şey ise akıl yürütmede ve rasyonel konuşma ile eylem kılavuzunda kullanılmaya uygunluktur." (Block, 1995: 227). Fenomenal bilinç, deneyimin öznel hissi ile erişim bilinci ise deneyim içeriklerinin kullanılması ile ilgilidir.

Bilinç üzerine var olan bu literatürü de dikkate alarak bu tez çalışması boyunca kullanılan terimlerin tanımlarının başlangıçta ortaya konulması, anlaşılırlığın artması ve anlam karmaşalarının ortadan kalkması açısından oldukça elzemdir. Bilinç bu tezde "deneyimleyenin deneyimde bulunduğunun farkında olması" anlamında kullanılmıştır. Buna bağlı olarak "özbilinç"[6] yerine sıklıkla "kendilik bilinci" ifadesi kullanılmış olup "deneyimleyenin deneyimde bulunduğunun farkında olmasının farkında olması" mealinde tanımlanmıştır. "Fenomenal bilinç" ise bu çalışma çerçevesinde "öznel ve niteliksel deneyim"e karşılık gelmektedir. Buradan yola çıkarak da deneyimleyen "fenomenal özne" olarak ele alınmıştır.

Yapay zeka kavramının kullanımında da bilinçte olduğu gibi bir çoğulluk söz konusudur. Yapay zeka, makinelere zeka kazandırmak üzere yapılandırılmış bir alandır. "Yapay"

[4] *Phenomenal consciousness.*
[5] *Access consciousness.*
[6] *Self-consciousness.*

sözcüğü üzerinde uzlaşılmaya varılması zor bir kelime olmuştur. 1950'lerde yapay zeka alanının çıkış noktası olarak kabul edilen Dartmouth Yaz Projesi'nde[7] "yapay" kelimesi, yapılan işe gerçekliği olmayan tamamen yapay olduğu görünümünü verdiği ve yapılan işte sanki bir düzmece varmış hissi uyandırdığı gibi gerekçelerle oldukça tartışılsa da zamanla literatürdeki yeri kalıcı hâle gelmiştir (Nillson, 2018: 82).

Yapay kelimesinin bu hissiyatının dışında gündelik dilde kullanımı da farklı göndermeler içerir. Bu göndermeleri iki grup altında toplarsak görünüş ve işlevsel olarak gerçek ancak insan yapımı olması nedeniyle oluşan yapaylık ile yalnızca görünüş bakımından gerçek olan işlevselliği aynı olmayan bir yapaylık birbirlerinden farklıdırlar. Işık, doğal ışığın insan yapımı hâlidir, ancak yapay da olsa doğal ışığın işlevi olan aydınlatmayı yerine getirebilir. Ancak durum yapay çiçek için farklıdır, görünüşte doğal çiçeğin bir taklidi olsa bile canlılık ve işlevi bağlamında gerçek bir çiçek değildir (Manzotti ve Tagliasco, 2008: 106).

Yapay zeka alanındaki "yapay" sözcüğü de benzer şekilde iki anlamda kullanılır. Yapay çiçek gibi bir yapay zeka, düşünebiliyormuş gibi görünen ancak basit bir hesaplama modelinden fazlası olmayan simülasyon bir zekaya sahiptir. Fakat yapay kelimesinin ışıkta olduğu gibi kullanıldığını savunanlar, yapay zekaların insanlar tarafından üretilmiş olmalarına rağmen bir kez kuruldukları zaman insanlardan farklı tarzda bile olsa düşünebildiklerini iddia ederler. Buna göre, nasıl bir arabadaki hareket ile tavşandaki hareket veya insandaki hareket birbirlerinden farklıysa veya nasıl kuşun uçması ile uçağın uçması birbirinden farklıysa yapay zekaların düşünmesi de insandan bu bağlamda farklıdır. Ancak nasıl ki arabada gerçek bir hareket, uçakta gerçek bir uçuş varsa makinelerin düşünmesi de gerçek bir düşünmedir (Manzotti ve Tagliasco, 2008: 107).

Yapay ışık ve yapay çiçek ayrımı teknik dilde zayıf ve güçlü yapay zeka ayrımına karşılık gelmektedir. Zayıf yapay zeka, bizim "yapay çiçek" metaforuna karşılık gelirken güçlü yapay zeka "yapay ışık" örneğine karşılık gelmektedir. Zira zayıf zeka teorisinin hedef içeriği "zekiymiş gibi görünen" programlar üretmekken güçlü yapay zeka "düşünebilen ve akıl yürütebilen, zekiymiş gibi görünen değil zeki olan" makineler üretmeyi amaçlamaktadır. Zayıf yapay zeka, zihnin hesaplama yetisini temel alırken; güçlü yapay zeka, zihni bütün olarak yeniden üretme fikri üzerine inşa edilmektedir. Güçlü ve zayıf yapay zeka kavramlarını sıklıkla kullanan Searle'e göre, güçlü yapay zeka teorisi, zayıf yapay

[7] Ayrıntılar için bkz. Bölüm 1.3.

teorisinden farklı olarak "doğru programlanmış bir bilgisayarın doğru girdi-çıktı ilişkileriyle düşünebileceği ve düşünebilmek için gerekli olan her şeye sahip olduklarından düşünmelerinin zorunlu olduğu" tezi üzerine kuruludur (Searle, 1990: 27). Yapay bilincin imkânının tartışıldığı bu tezde atıfta bulunulan yapay zeka kavramının "güçlü yapay zeka" teorisince kullanılan anlamda olduğu göz önünde bulundurulmalıdır.

Yapay zeka ve bilinçten ne anladığımız açık ve seçik olarak ortaya konulmadıkça yapay bilinç üzerine ileri sürülecek fikirler güvenilir olmayacaktır. Yapay zekanın fikrî, kavramsal, bilimsel ve teknik temellerinin tarihçesi ile birlikte ele alınması bu minvalde büyük önem arz etmektedir. Yapay zekanın bilinç ile birlikte düşünülmeden önce tek başına değerlendirilerek genel bir çerçevesinin çizilmesi hem yapay bilinç üzerine tartışırken yararlanılan bir kılavuz olması hem de teknik imkânların ulaşmış olduğu konumu göstererek yapay bilincin kuramsal imkânını tespit etmemize katkı sağlaması açısından önemlidir. Bu bağlamda, tezin giriş bölümü yapay zekanın tarihçesine; ikinci ve üçüncü bölüm ise bilince yönelik felsefi yaklaşımlara ayrılmıştır. Ardından da yapay bilince yönelik yaklaşımlar epistemolojik, sosyal, etik boyutlarıyla beraber değerlendirilecek ve "fenomenal gerçekçilik" adını verdiğimiz özgün bakış açısıyla yapay bilinç üzerine çıkarımlarda bulunulacaktır.

1.1. Yapay Zekanın Fikrî Gelişimi

Yapay zeka arayışının en ilkel formlarından itibaren başlangıcı, tüm diğer arayışlardan farksız olarak insan zihninin ve hayal gücünün bir ürünü olarak insanlığın en erken dönemine kadar gider. Bu anlamda insanın zihnî ve eylemsel yeteneklerine sahip başka bir varlığın hayali insanlığın kendisi ve etrafı hakkında düşünme serüveni ile eş zamanlı bir başlangıca sahiptir demek yanlış olmayacaktır. Bugün yapay zeka insanın teknoloji üretme kabiliyetinin geldiği önemli bir aşama olup nüvesi tarihin derinliklerinden süzülüp gelmektedir. Yapay zeka, insanlığın düşünce geçmişinde farklı formlarda da olsa sürekli olarak var olmayı başarmıştır. Buna karşılık yapay zekanın ayrı bir bilim ve uzmanlık kolu olarak derinlemesine incelenmesi ve buna bağlı teknolojik gelişmelerle hayata geçmeye ve etki alanını genişletmeye başlaması son derece devrimseldir. İnsanlığın ortak hafızasında yer bulmuş bir hayal, bizlerin de şahitliğinde elle tutulur bir hâle gelmektedir. Bugün yapay zeka fikri, insanın bireysel ve toplumsal yaşamını kökten değişikliklere uğratma gücüne

sahip büyük bir teknolojik araç olarak karşımızda durmaktayken; bu aracın insanın kas gücünün mekanik sistemlere doğru evrilmesinden, insan düşüncesinin taklidine ulaştığı süreçte ortaya çıktığını görüyoruz. Bu süreçte gelinen noktayı insanların yaşam biçimlerini, toplumsal ve kültürel yapıları, uluslararası ilişkileri, ekonomik sistemleri, hukuksal değerlendirmeleri yeniden şekillendiren büyük ve devrimsel bir adım olarak nitelersek, yapay zekanın artık belli bilimsel kolların uğraşı olmaktan çok insanlığın bir meselesi konumuna eriştiğini söyleyebiliriz. İnsan zihninin bilişsel yeteneklerini tümüyle taklit edebilen ve hatta bu yeteneklerle birlikte çok boyutlu düşünme biçimleri geliştirebilen makinelere doğru hızla ilerlerken yapay zeka sosyal yaşamın vazgeçilmez bir öğesi olmayı çoktan başarmış görünmektedir.

Yapay zeka teknolojisinde yaşanan hızlı gelişmeler, felsefi düşüncede de yeni pencereler açmaktadır. İnsanın mutluluğu, karar verme mekanizmaları, inanç sistemleri, evreni kavrayış biçimi, sosyal konumu, işgücüne katkısı, etik kaygıları ve değerleri hızlı bir biçimde yeniden şekillenmekte ve bu gidişat içerisinde felsefeye bu kırılma noktalarının anlaşılması, sorgulanması ve analizi için büyük görevler düşmektedir. Sanayi Devrimi'nin insan düşüncesinde yarattığı dönüştürücü etki düşünüldüğünde yapay zekanın insanın zihinsel yaşamına olacak ya da olmakta olan dönüştürücü rolünün ehemmiyetini kavramak, bu teknolojik ihtilalin sonuçlarına uyum sağlamanın başat koşulu olarak ele alınmayı hak etmektedir.

Tarih boyunca öyküler, şiirler, efsaneler, resimler ya da heykeller gibi sanat eserlerinde kendi kendine hareket edebilen, düşünebilen ya da görünüş olarak insandan ayırt edilemeyen düşsel varlıkların tasvirlerine sıkça rastlanır. Felsefe tarihi de antik dönemden başlayarak benzer şekilde -derinlemesine üzerinde durulmasa bile- mekanik insan konusuna değinen birçok filozof barındırır.

Bu bağlamda, otomatik çalışan âletlerin yapımının insanlık tarihi kadar eski olduğu söylenebilir. İnsanların etraflarında gördükleri nesneleri, göksel objeleri, yaşayan canlıların hareketlerini taklit etme arzusu, ilk ilim dallarından birisi olan sibernetiğin doğuşunu sağlamıştır. Bu durum, aslında insanın yaratıcı ve ölümsüz olma isteği ile icat etme tutkusuyla da yakından ilişkilidir.

Tarihteki ilk otomat örneğinin MÖ 250'de inşa edilen Mısır Su Saati ya da Klesibos Pompası olarak adlandırılan basınç yaratmak için yukarı-aşağı hareket edebilen bir kola bağlı

pistonları kullanan bir basma tulumba olduğu kabul edilmektedir[8] (Frances ve Mohun, 2014: 30). Eski Mısır mezarlarında bulunan kolları hareket edebilen bebekler ve daha karmaşık otomatik âletler mekanik alanında Mısır'ın ileri bir seviyede olduğunu göstermektedir.

Greklerde ise Daedalus'un uçan bir kuş ile tapınağı koruyan bir heykel yaptığı, yine Tarentumlu Arcytas'ın (MÖ IV. yüzyıl) tahtadan yaptığı bir kuşun uçtuğu söylenmektedir (Topdemir ve Unat, 2018: 146). Yine Antik dönemde Aristoteles'in *Politika* eserinde köleliği meşrulaştırma çabasının içerisinde imkânsız bir hayal olarak gördüğü otomasyon fikrine şöyle rastlanır:

> (…) köle ise başka herhangi bir uşak gibi canlı bir yaratık olduğu için, birçok araçlar değerinde bir araçtır. Çünkü, her aracımız, Daidalos'un yaptığı heykeller ya da ozanın "Kendiliklerinden tanrıların toplantısına girerler." dediği Hephaistos'un tekerlekli sehpaları gibi, biz söyleyince ya da gerektiğini kendisi görerek işlerini yerine getirebilseydi -diyelim, dokuma tezgâhının mekiği kendiliğinden gidip gelse, lirin mızrabı kendiliğinden çalsaydı, o zaman ne yapımcıların işçiye gereksinmesi olurdu ne de efendilerin köleye (Aristoteles, 1975: 12).

Helenistik dönemde İskenderiye Mekanik Okulu'na mensup Ctesibios, Heron ve Philon'un hava, boşluk ve denge prensibiyle yaptıkları otomatik âletler de mekanik, hidrolik ve pnömatik mühendisliğinin gelişmesine büyük katkı sağlamıştır.

Orta Çağ İslâm dünyasında da mekanik alanında önemli çalışmalar yapılmıştır. Abbasîler döneminde Halife Me'mûn'un katkılarıyla mekanik ve sibernetik alanlarında önemli çalışmalar yapan Benî Mûsâ (Mûsâoğulları) anılmaya değer öncü isimlerdendir. Onların tasarladıkları ve bir kısmını çalıştırdıkları otomatik ve yarı otomatik çalışan fıskiyeler, su saatleri ve pompalar, hassas âletler, insanları eğlendirmek için tasarladıkları oyuncaklar robotik çalışmalar alanına büyük katkı sağlamıştır. Onları takip eden el-Cezerî'nin 1206 tarihli *Mekanik Âletlerin ve Hünerli Bilginin Kitabı*[9] adlı eserinde yer verdiği görkemli su saati ve değişik müzik parçalarını çalmak için komut verilebilen ilk otomat bandı, bilim tarihindeki ilk otomat örneklerindendir. O, sibernetiğe çok önem vererek, ilk defa robot icat etmeyi denemiş, birden fazla canlıyı (insan, fil, kuş vb.) bir arada çalıştıran ve her biri farklı

[8] Grekçe kendi kendine hareket eden anlamına gelen "auto matos" (otomat), kendiliğinden harekete başlayıp görevini yaptıktan sonra durabilen, akabinde yeniden harekete geçebilen âlet anlamında kullanılmaktadır.

[9] Tekeli S., Dosay M., Unat Y., (2002). *el-Cezerî, el-Câmiʿ beyne'l-ʿilmi ve'l-ʿameli'n-nâfiʿ fî sınâʿati'l-hiyel,* Ankara: Türk Tarih Kurumu (Türkçe).

fonksiyonlar icra eden âletler yapmıştır. Fil Su Saati gibi önemli otomatik âletlerin icadına imza atan el-Cezerî'ye bu yüzden "robot ve otomatların babası" unvanı verilmiştir.

Bilim tarihinin kendi kendine hareket eden ilk makine örneklerinin su saatleri, pnömatik âletler ve pompalar gibi insan yaşamını kolaylaştıran âletler olduğu söylenebilir. Bilim tarihçisi George Dyson tarafından yapay zekanın atası olarak anılan Thomas Hobbes, ünlü eseri *Leviathan*'da "yapay canlı" yaratmanın ve insanı taklit etmenin mümkün olduğunu; bunun yönteminin de sanat olduğunu iddia ederek şöyle söyler:

> Tanrı'nın dünyayı onunla yaratmış olduğu ve yönettiği sanat olan doğa, başka pek çok şeyde olduğu gibi bunda da yapay bir canlı yaratacak şekilde, insanın sanatı ile taklit edilir. Çünkü hayat, organların, başlangıcı içerdeki bir temel parçada bulunan, hareketinden başka bir şey değildir; bütün otomatlar (yaylar ve çarklar yardımıyla kendi kendine hareket eden makinaların, mesela bir saat) yapay bir hayata sahip olduklarını söyleyemez miyiz? Kalp nedir ki bir yaydan başka; sinirler nedir ki çok sayıda yaylardan başka; ya eklemler~ yapıcının planladığı şekilde bütün gövdeyi harekete geçiren çok sayıda çarklardan başka? Sanat daha da ileriye gider, doğanın o rasyonel ve mükemmel eserini, insanı taklit etmeye kadar (Hobbes, 2007: 17).

1708 yılında Vaucanson tarafından tasarlanıp inşa edilen mekanik ördekler de yine bilim tarihinin önemle yer verdiği yapay varlıkların ilk örneklerindendir. Programlanabilir makinelerin başlangıcı ise 1801 yılında Fransız Joseph Marie Jacquard tarafından üretilen tekstil tezgâhlarına dayandırılmaktadır (Frances ve Mohun, 2014: 168). 19. yüzyılda üretilen mekanik oyuncakların dayandığı fikir de yine otomatlardır. Dışarıdan enerji kaynağı sağlamanın dışında tamamen otomatik olan bu ilk makinelerin en önemli eksiklikleri kendi çevre şartlarını algılayamamalarıydı (Nillson, 2018: 49).

On dokuzuncu yüzyılın sonlarına kadar otomatlar yalnızca mühendislik alanının çabalarıyla ilerlerken sonrasında üretilmeye başlayan bilişim cihazları, yapay zeka ve robotik alanının seyrini değiştirmiş ve yükü mühendisliğin üzerinden almıştır. 1887 yılında Amerikalı Herman Hollerith'in (1860-1929) kendisine nüfus sayımı istatistiklerini çizelge hâline getirmesinde yardımcı olarak ürettiği kartlı makine elektronik bilgisayarların bir öncülü olarak görülmektedir (Frances ve Mohun, 2014: 224). İnsan iş yükünü hafifleten bu tür icatlar ve bilişimdeki ilerleme, sembolik mantığın ortaya çıkmasıyla paralel görünmektedir.

Thomas Hobbes'un insan zihninin fiziksel bir süreç olduğunu söylemesi ve Newton sisteminin fiziksel olan her şeyin matematiksel olarak ifade edilebileceğini göstermesi, modern mantığın temellerini hazırlamakla kalmamış; fiziksel bir süreç olan insan zihninin matematiksel olarak gösterilebileceği fikrine de dayanak olmuştur. Mantığı, mekanik bir

sanat olarak tanımlayan ve tamamen biçimsel hâle dönüştürülebileceğini söyleyen Raymond Lulle (1235-1315), sembolik mantığa işaret eden ilk isimdir (Öner, 1991: 11). Lulle'un fikirlerinin geliştiricisi olarak kabul edilen Gottfired Wilhelm Leibniz (1646-1716) ise sembolik mantığın fikir babasıdır. Pascal'ın yalnızca toplama-çıkarma yapabilen ilk mekanik hesap makinesini icat etmesinden 25 yıl sonra Leibniz dört işlem yapabilen bir hesap makinesi yapmayı başarmıştır. Bu buluş bugün çok basit görünse dahi matematik bilen insanların zenginlerin yanında sırf bu sebeple iş bulduğu bir çağda, dört işlemi yapabilen bir makinenin olması fikri bile devrim olarak nitelendirilmiştir. Leibniz'ın yapay zeka tarihindeki önemi yalnızca bu hesap makinesi değildir; aynı zamanda inşa ettiği bu cihaz üzerinden yürütmüş olduğu ileri görüşlü fikirleri Leibniz'ın adının bugün yapay zekayla ilgili yapılan hemen her çalışmada anılır olmasını sağlamıştır.

Leibniz hesaplama gibi karmaşık bir bilişsel işi makinelere yaptırmanın olanaklı olduğunu gördükten sonra, diğer tüm bilişsel işlerin de makineler tarafından gerçekleştirilebileceği ihtimali üzerine düşünmüştür. O, düşünme işlerimizi bizim yerimize yapabilecek bir makine hayal etmiş ve bu makineye "calculus ratiocinator" adını vermiştir (Say, 2017: 16). Onun bu makine hayalini gerçekleştirmek için de bir önerisi vardır: Kendi inşa ettiği hesap makinesi nasıl sayıları temsil ediyorsa, inşa edilecek bir düşünce makinesi de düşüncelerin kendisini temsil etmelidir. Buradan yola çıkarak Leibniz yapay zeka yazılımlarının temelini oluşturacak olan sembolik dili öneren ilk isim olmuştur. Ona göre, düşüncede geçen tüm kavramlar sembolik bir dille matematiksel olarak temsil edilebilir. Bu sembolik dil, hem hatasız akıl yürütme yapmamızı sağlayabilir hem de anlamsal karışıklıkların önüne geçebilir:

> Muhakemelerimizi arındırmanın tek yolu onları matematikçilerin ele aldığı gibi somutlaştırmaktır, böylelikle hatalarımızı tek bakışta bulabilir ve bireyler arasında ihtilaf çıktığında bunu söyleyebiliriz: Hadi hesaplayalım [*calculemus*], telaşa mahal vermeden kimin doğru olduğunu görebilmek için (Leibniz, 2011: 19).

Leibniz'ın düşündüğü dil, insan düşünce yapısının alfabesini oluşturan evrensel bir dildir. Leibniz henüz 25 yaşındayken yaptığı bu önemli saptamayı ne yazık ki geliştirememiştir.

1.2. Yapay Zekaya Mantıktan Gelen Katkılar

Leibniz'in ortaya attığı sembolik dil önerisi yüzyıldan daha uzun bir süre sonra George Boole (1815-1864) tarafından temellendirilmiştir. Boole'a kadar var olan tek mantık

sistemi Aristoteles'ten miras olan ve doğal dille kurulan klasik mantıktı. Aristoteles mantığı yüzyıllar boyunca hem İslâm coğrafyasında hem de Avrupa'da yüzyıllar boyunca tek başına var olmuştur. Aristoteles'ten sonra Stoacılar mantığı metafizikten ayırarak, onu şekille ve dille ilgili bir bilim hâline getirmeye çalışmışlar, fakat bu konuda çok ilerleyememişlerdir (Öner, 1991: 5). Klasik mantığın en görünür açmazı, çifte anlamlılık ve muğlaklık sorunlarını aşamamasıdır.

Tüm sağlıklı kediler, 4 ayaklıdır.

Pamuk, sağlıklı bir kedidir.

O hâlde, **Pamuk 4 ayaklıdır.**

Örnekteki akıl yürütme doğru ve tutarlı bir çıkarım içermesine rağmen, bağlamdan koparıldığında Pamuğun bir kedi ismi olduğu anlaşılamamaktadır. Dolayısıyla Boole, bu muğlaklık ve çifte anlamlılıkları aşacak bir sembolik dil oluşturma niyetiyle kelime anlamlarından yola çıkmıştır.

"Beyaz orkide" kelimesi tüm beyaz şeyler ve tüm orkide kümelerinin kesişimi anlamına gelmektedir. Bu kesişim fikri, Boole'un bugünkü gösterim biçiminin temelini oluşturacak olan 0-1 sisteminin çıkış noktasıdır. "0" hangi kümeyle kesişirse kesişsin değişmeyen boş kümeyi ifade ederken "1" hangi kümeyle kesişirse kesişsin her şeyi içeren evrensel kümeye karşılık gelmektedir. Boole bir önermenin doğru olduğu anların kümesi boşsa "0", doluysa "1" ile gösterebileceğimizi ortaya koymuştur. Boole'un modern mantığın temelini oluşturduğu bu sistemde yanlış, "0", doğru ise "1" ile temsil edilir; "veya" bağlacı "+" ile "ve" bağlacı ise "x" ile gösterilir:

1+0=1	1x1=1
1x0=0	0+0=0
1+1=1	0x0=0

Boole sisteminin en büyük yetersizliği atomik yapısından kaynaklı içeriği gösteremeyişiydi; yani "Ali insandır." önermesini p ile "Ali ölümlüdür." önermesini q ile gösterirsek, p ya da q için insan olan Ali'nin, ölümlü olan Ali ile aynı kişi olduğuna dair hiçbir gerekçe yoktur (Nillson, 2018: 33). Bu eksikliğin giderilmesi için Frege (1848-1925), mantıkla aritmetiğin

birleştirilmesi gayesi ile yola çıkarak bugün "kavramlar mantığı" ya da "yüklemler mantığı" olarak ifade edilen bir sistem oluşturdu. Bu sistemde, Boole'un sistemine ek olarak ∀ (Her) ve ∃ (Bazı) niceleyicileri bulunmaktadır. Yüklemler mantığı, Russell (1872-1971) ve Whitehead'in (1861-1947) katkılarıyla son şeklini almıştır. Bu mantık, cümleleri ve cümlelerin içindeki mantıksal olarak önemli yapıları sembolize etmenin yanı sıra cümleler arasındaki ilişkileri tanımlayacak araçlar da sunmuştur (Churchland, 2019: 277). Bir önermenin, başka önermeler aracılığıyla nasıl kanıtlanacağının matematiksel gösterimini oluşturmuş ve kurallar ile kuramların mantıksal sistemini oluşturmaya olanak tanımıştır. Yüklemler mantığındaki tüm mantığın matematiksel olarak ifade edilebileceği varsayımı ile beraber matematiksel ve bilimsel akıl yürütmenin genel olarak önermeler ve onları birlikte tutan mantıksal ilişkiler üzerinden tanımlanabilmesi olası olarak kabul edilmeye başlamıştır. Bu da fikir dünyasında, tüm bir bilimin ve bilişsel sistemin, matematiğin tamamı gibi sistematize edilebileceği sezgisine ilham kaynağı olmuştur. Yüklemler mantığı akıl yürütmede kullanılacak tüm bilgilerin açık ve simgesel olarak yazılmasına imkân tanımakta ve günümüzde yapay zekada kullanılan bir gösterim biçimi olan anlamsal ağ modellerinin de öncülüğünü yapmaktadır.

Mantık alanının yapay zeka ile doğrudan bir ilişkisi vardır. Ancak tartışmalı olan bu ilişkinin derecesidir. Sembolcülere göre, zekanın yapay bir sisteme aktarılabilmesi bu zekayla ilintili davranışların sembolik dil aracılığıyla yapay ortama aktarılmasıyla başarılabilir bir durumdur. Ancak başka bir grup ise doğal zekanın taklit ya da transferi için mantığın tek başına yeterli olmayacağını iddia etmekte ve biyolojik faktörlerin önemine dikkat çekmektedirler. Sembolcülerin yapay zeka için yeterli gördükleri mantık dili de kendi içlerinde farklılaşmaktadır. Önermeler mantığı (Boole mantığı) ve birinci derece yüklemler mantığı (Frege mantığı) şeklinde gruplanan klasik mantığın yeterli olacağını savunanlar ile olasılık mantığı, zamansal mantık, bulanık mantık, kuantum mantığı gibi klasik olmayan mantıkların doğal zekanın sembolleştirilmesi için zorunlu olduğunu savunanlar sembolcü yaklaşımın farklı taraflarıdır. Hem klasik hem de klasik olmayan mantık ile çalışan yapay zeka sistemleri bilgi tabanlı olarak kurulmaktadır. Bilgi-tabanı, alana özel olan bilgi ve iyi tanımlanmış cümlelerden oluşmaktadır. Örneğin, bir akademisyenin bilgi tabanında alanıyla ilgili literatürü takip etmek, fikir hırsızlığı (intihal) yapmamak, öğrencilere düşünce özgürlüğü adı altında gerçeklikle ilgisi bulunmayan sübjektif görüşlerini empoze etmemek, en az bir yabancı dilde yayın okuyabilmek gibi bilgiler bulunmaktadır. Bir otomobil sürücüsünün bilgi tabanında, kırmızı ışıkta durulur, yayalara yol verilir, önde bir araba

durduğunda ona çarpılmaz frene basılır gibi önceden edinmiş olduğu bilgiler bulunur. Trafikteki her eyleminde aldığı kararlar onun bilgi-tabanından yaptığı çıkarımlardan ileri gelmektedir. Çıkarım, bilgi tabanında bulunan bir bilginin doğrudan alınması olacağı gibi, bilgi-tabanındaki bilgiden fazlası da olabilir. Önermeler mantığı, yargılardan, *doğru, yanlış* ve *bilinemez* olmak üzere üç sonuca varmamızı sağlarken; yüklemler mantığında yargının yanında, nesnelerin ilişkileri ve fonksiyonlarından hareketle doğru, yanlış ya da bilinemez sonucuna varmaktayız. Dolayısıyla yüklemler mantığı olarak bilinen birinci derece mantık, önermeler mantığından fazla olarak *semantik* ilişkiler kurulmasına da vesile olur.

Olasılık mantığı, zamansal mantık, bulanık mantık gibi klasik olmayan mantıkların, bu iki klasik mantıktan farkı ise ulaşılan sonuçlarda yatmaktadır. Olasılık mantığında, yargılardan dereceli sonuçlara varırız; başka bir deyişle yalnızca 0 ve 1 yoktur, minimum 0, maksimum 1 olacak şekilde doğruluk dereceleri vardır. Bulanık mantık ise olasılık mantığından yargıların da doğruluk derecelerinin olmasıyla farklılaşmaktadır. Bulanık mantığın ortaya çıkışı 1965 yılında Kaliforniya Üniversitesi'nden Lütfi-Zâde'nin ilk defa bulanık küme kavramının temeli olan "yumuşak" yaklaşım ile sistem tanıma ve tasarımını gerçekleştirmesine dayanmaktadır (Vural, 2002: 185). Bulanık mantık, belirsiz koşullarda, belirsiz olaylar üzerine karar verme ile ilgili durumlarda kullanılmak üzere oluşturulmuştur. Kaç km hızla giden bir araca hızlı deriz, kaç kg ağırlığında olan bir insana şişman deriz ya da kaç yaşında olan bir insana yaşlı deriz? Bulanık mantık, bunlar gibi kesin bir cevabı olmayan yargılardan, muğlak durumlardan sonuç aralıkları çıkarsamaya yarayan bir araç olarak kullanılmaktadır. Örneğin, uzun boylu insan ifadesi muğlak bir ifadedir. Klasik mantıkla yapılan ancak örneğin boyu 1.65 cm'den az olanlar kısa, 1.65 cm'den fazla olanlar uzun demektir. Ancak bulanık mantık bize uzun boyluluğun ve kısa boyluluğun da dereceleri olduğunu söylemektedir. 1.70 cm boy ile uzun boylu olmak ile 1.90 cm boy ile uzun boylu olmak birbirinden farklıdır. Bisiklet sürmek için havanın ılık olması en uygunudur, biz bunu günlük dilde "normal" hava olarak adlandırırız. Bu normallik 18-25 derece arasında herhangi bir aralık için geçerli olabilir. Bulanık mantık bu gibi durumlarda karar verilmesi için geliştirilmiş bir teknoloji olup "normal" gibi kavramların bilgisayar tarafından anlaşılmasını sağlamaktadır (Öztemel, 2006: 18).

Bulanık mantık, *bulanıklaştırma, bulanık önermeyi işleme* ve *netleştirme* süreçlerinden oluşmaktadır. Bulanıklaştırma, çözülecek problem ile ilgili bulanık önerme değişkenlerinin ve karar verme kurallarının belirlenmesi ve üyelik fonksiyonunun belirlenmesi işlemidir

(Öztemel, 2006: 19). Bulanık önermeyi işleme, problemin çözüm alanını oluşturma ve netleştirme de çözüm alanından tek bir değer elde etme süreçlerine karşılık gelmektedir.

Lütfizâde'ye göre, her ne kadar bugün yapay zeka klasik mantığı kullanıyor olsa da aslında onu açımlayacak olan çok değerli bulanık mantıktır. Çünkü bulanık mantık tam doğru olmayan ya da diğer deyişle yaklaşık akıl yürütme şeklinde değerleri vermekte, bu da çok keskin, çok fazla kısıtlayıcı iki değerli dikotomik düşünce yerine ihtimalleri çoğaltmaktadır. Dolayısıyla bulanık mantıkla akıllı tasarımlar üreterek daha hızlı, dakik, güvenilir ve optimal işlemler yapmak mümkün hâle gelecektir (Zadeh, 2007: 173).

Dil ve düşüncenin sembolik olarak aktarılabileceği inancı da modern mantıkla beraber üzerinde durulur bir husus hâline gelmiştir. Elektronik bilgisayarları inşa eden ilk filozoflardan olan Alan Turing'in 1936'de yayınlamış olduğu "On Computable Numbers, with an Application to the Entscheidungsproblem"[10] başlıklı makalesi yapay zeka tarihinin köşe taşlarındandır Bu makalede, Turing'in önerdiği, bugün Turing Makinesi olarak adlandırılan soyut bir makinedir. Bu bilgi-işleme makinesi, aslen bir bilgisayar modeli olma amacıyla değil David Hilbert'in (1928) "karar verme problemi"[11] olarak adlandırılan "Herhangi bir matematiksel önermenin verilen aksiyomlar kullanılarak evrensel olarak geçerli olup olmadığını gösterebilecek bir algoritma var mıdır?" sorusuna cevaben kurgulanmıştır (Hilbert ve W. Ackermann, 1950: 112). Turing bu dijital makineyi, çözmek istediği matematiksel probleme bağlı olarak sonsuz olanaklara sahip olarak tasvir etmiştir. Buna göre, dijital makine verileri üzerine işleyebileceği sonsuz uzunlukta bir banda (hafızaya) sahiptir. Bu tasarımda ayrıca verileri okuyan ve yazan, sağa-sola hareket edebilen bir okuma-yazma kafası ve programı işleten bir karar verme ünitesi bulunmaktadır. Okuma-yazma kafası girdileri değerlendirip, çıktıları da yine aynı bant üzerine işleyebilmektedir (Turing, 1936: 231). Bu makinenin, en önemli parçası olan karar verme ünitesi bugün bilgisayarların işlemcisine tekabül etmektedir. Bu ünite, bant üzerinde o anda durulan bitin türüne göre karar vermektedir. Bu kararlar, kafayı sağa veya sola hareket ettirmek, biti değiştirmek veya durmak olabilir. Durma kararı, Turing Makinesi kurgusuna sebep olan problemle ilişkilidir. Bir önermenin kanıtlanıp kanıtlanamayacağı, Turing Makinesi'nin belirli işlemler sonucunda durup durmayacağına bağlıdır. Durma, başarılı olma yani o önermenin kanıtlanması anlamına gelmektedir. Makine durduğunda, uğraşılan probleme bir

[10] Hesaplanabilir Sayılar: Karar Verme Problemine Uygulanması
[11] Entscheidungsproblem

çözüm üretmiş demektir. Ancak bazı problemler için, makinenin ne zaman duracağını tahmin etmek mümkün değildir. Dolayısıyla Turing bu makalesiyle "Herhangi bir mantematiksel önermenin verilen aksiyomlar kullanılarak evrensel olarak geçerli olup olmadığını saptayabilen bir algoritma var mıdır?" sorusuna "hayır" cevabını vermekte ve karar probleminin çözülemeyeceğini iddia etmektedir (Turing, 1936: 262).

1.3. Bağımsız Bir Bilim Dalı Olarak Yapay Zeka

Bir bilgisayar eğer bir insanı kendinin de
insan olduğuna ikna edebilirse zeki olmayı hak edebilir.
Alan Turing

Turing'in yukarıda adı geçen makalesi dışında 1950 tarihli "Computing Machinery and Intelligence" isimli makalesi de yapay zeka felsefesi için temel sağlayan bir role sahiptir. Bu makale ile Turing şunu iddia eder: Bir önerme ya da hesap ne kadar karmaşık olursa olsun, komutlara ya da bitlere bölünerek makinede kullanılan girdi-çıktı yöntemiyle sonuca ulaşabilir. Hatta evrensel Turing Makinesi, tüm Turing makinelerini taklit ederek, tek bir fiziksel makineye sonsuz sayıda değişik iş yaptırılabileceği anlamına gelir. Evrensel makine ile Turing'in amaçladığı, insan zihninde gerçekleşen tüm hesaplamaların tek bir makinede yapılabileceğini göstermektir. Programlanmış sayısal bilgisayarlar, herhangi bir bilişim amacı için kullanılabilmeleri bakımından önemlidir. Modern sayısal bilgisayarlar da hedeflenen bilişsel veriyi elde etmek için kullanılan, pratikte Turing'in kurguladığı evrensel makinelerdir. Ancak bu benzetmenin yalnızca "pratikte" nitelemesiyle beraber yapılabileceği gözden kaçırılmamalıdır; çünkü Turing'in sonsuz bandındaki gibi sonsuz bir depolama kapasitesi henüz hiçbir bilgisayar için mümkün değildir (Nillson, 2018: 64).

"Computing Machinery and Intelligence" makalesinde Turing'e popülerliğini kazandıran başka bir temel unsur "Makineler düşünebilir mi?" sorusunun muğlaklığına karşı geliştirmiş olduğu kendisinin "taklit oyunu"[12] olarak tanımladığı daha sonra ise "Turing Testi" olarak adlandırılan düşünce deneyidir. Hesaplama ve bilgi işleme süreçlerini fiziksel altyapılarından bağımsız değerlendiren Turing'e göre, biyolojik olmayan bir yapının da düşünce üretmesi mümkün görünmektedir. Makinelerin düşünebileceğini gösterebilmek için kurguladığı Turing Testi olarak bilinen taklit oyununu Turing şöyle anlatmaktadır:

[12] *Imitation game.*

> Oyun üç kişiyle oynanır: Bir erkek (A), bir kadın (B) ve bir sorgucu (C). Sorgucunun cinsiyeti önemli değildir. Sorgucu, A ve B'den ayrı bir odada yer alır. Oyundaki amaç, sorgucunun A ve B'nin cinsiyetlerini tahmin etmesidir. C'nin sorduğu sorulara, A ve B yazılı cevaplar verir. (…) A'nın görevi, C'yi yanıltarak kadın olduğuna ikna etmektir. B ise kendisinin kadın olduğunu ispatlamaya çalışacaktır. En sonunda ise C buna yönelik tahminde bulunacaktır (Turing, 1950: 433-434).

Daha sonra, Turing içeride kadın ya da erkek yerine bir makine olsaydı ne olacağı üzerine düşünür ve buradan hareketle testinin amacını oluşturan "Makineler düşünebilir mi?" sorusunu sorar. Bu soru Turing'e göre, "Taklit oyununda başarılı olabilecek hayal edilebilir dijital makineler var mıdır?" sorusuyla aynı anlama gelmektedir (Turing, 1950: 442). Bu bağlamda da taklit oyununu odadaki bir kişiyi dijital makine ile değiştirerek yeniden kurgular. A'nın yerine geçen makine, sorgulayıcıyı (C) aynı şekilde kadın olduğuna ikna etmeye çalışacaktır. Eğer sorgulayıcı (C) belirli bir süre içerinde vardığı sonuçtan emin olamazsa Turing'in erkeğin yerine kullandığı makine ya da bilgisayar testi geçmiş olacaktır. Bu sebeple düşünen makinelerin mümkün olduğunu söylemekte prensipte bir sakınca yoktur (Turing, 2004: 448). Bu testin amacı, testi geçen makinelerin düşünebilme yetisine sahip olduğunu ve bu düşünme yetisinin insanın sahip olduğu düşünme yetisinden bir farkının olmadığını göstermektir.

Turing'in bu makalesi Yapay Zeka ismi verilen bir alanın doğuşunun başlangıcıydı. Sonrasında 1955'te "Yapay Zeka Yaz Araştırma Projesi"[13] teklifi sunuldu, teklif şöyle başlamaktaydı:[14]

> Darthmouth Koleji Hanover, New Hampshire'da, 1956 yazında, yapay zeka üzerine 2 aylık, 10 kişilik bir çalışma yapılmasını teklif ediyoruz. Çalışma, öğrenmenin tüm yönlerinin veya zekanın başka özelliklerinin prensipte, bir makinenin taklit edebileceği derecede ayrıntılı olarak tarif edilebileceği kestiriminden yola çıkacaktır. Makinelerin dili kullanmalarının, soyutlamalar yapmaları ve kavramlar oluşturmalarının, şu anda yalnızca insanlara özgü olduğu düşünülen problemleri çözmelerinin ve kendilerini geliştirmelerinin nasıl sağlanabileceğini bulmak üzere araştırmalar yapılacaktır. Dikkatle seçilmiş bir bilim insanı grubunun bir yaz boyunca bu problemler üzerine birlikte çalışmasının önemli ilerlemeler sağlayacağını düşünmekteyiz.

[13] 1956 Dartmouth Yapay Zeka Yaz Araştırma Projesi teklifi Rockefeller Vakfı'na 31 Ağustos 1955'te John McCarthy, Marvin Minsky, Nathaniel Rochester ve Claude Shannon tarafından verilmiştir. Projenin orijinal nüshası bir kapak sayfası ve 17 sayfa metinden oluşmaktadır. Kopyaları Dartmouth Koleji ve Stanford Üniversitesi'nin arşivlerinde muhafaza edilmektedir.

Metin için bkz. http://jmc.stanford.edu/articles/dartmouth/dartmouth.pdf

Bu teklifte yapay zeka ilk kez şöyle kullanılmıştır: "Mevcut amaç uyarınca, yapay zeka sorunu, insan öyle davrandığında zeki olarak nitelendirilecek davranışları makinelere yaptırtmak olarak tanımlanacak." Yapay zekanın bu ilk tanımı, ilk yapay zeka çalışmalarının yönünü de büyük ölçüde şekillendirmiştir. Bu alanda, sinirsel ağlar üzerine ilk girişimler de doğrudan bu tanım üzerinden yapılmış, insanların nasıl problem çözdüğüne dair iç gözlemlerden sınırlı olarak faydalanılarak otomatik olarak yapabildiğimiz işlerin üzerinde durmadan algısal becerilerimizi makineleştirmek hedeflenmiştir. Bu düşüncelerden hareketle, ilk sinir ağları modelleri 1950'li yıllarda kurulmuştur.

Hem bilişsel bilim, hem mühendislik hem de sinirbilimin büyük vaatleriyle çok hızlı başlayan, büyük hedeflere sahip olan ve kısa sürede insan zekasının aşılacağı umudunu taşıyan yapay zeka çalışmaları, başlangıçtaki hızını 1970'lerin ilk yarısından itibaren kesmek zorunda kalmıştır. Hem devlette hem de sanayide yapay zekadan beklenilenlerin gerçekçi olmadığının görülmesi, yapay zeka çalışmaları için ayrılan fonların azalmasına yol açmıştır. Gerçekçi olmayan umutlara dayalı olarak vaat edilen sistemlerin inşa edilemeyişi, artan eleştirel yorumlarla birlikte 1970'lerin ortasından 1980'lerin sonuna dek uzanan bir zaman diliminde, literatürde "yapay zeka kışı" olarak adlandırılan bir döneme sebebiyet vermiştir (Nillson, 2018: 416). Yapay zekanın sınırlı taklitlerin önüne geçemeyeceği ve boş bir balon olduğunun düşünüldüğü bu dönemde, alanın hem çalışan sayısında hem de aldığı maddi destek miktarında ciddi bir düşüş yaşanmışsa da bu kış tek bir mevsim olarak kalmış ve umutsuzluk genel bir çöküşe yol açmamıştır. 1980'lerde savaşlarda kullanılmak üzere robot askerler yetiştirme projesiyle, yapay zekaya olan ilgi yeniden artmışsa da; savaşlarda kullanılmak üzere inşa edilen robot ve akıllı araçların yapabildiği tek şey kaybolmak olmuş ve 1990'ların ortasına kadar yeni bir yapay zeka kışı yaşanmıştır (Kaku, 2017: 255). Ancak hatırı sayılır oranda araştırmacı hedeflerini küçülterek de olsa güçlü yapay zeka yerine zayıf yapay zekaya odaklanarak çalışmaya devam etmiştir. Zayıf ya da genel yapay zeka olarak adlandırılan teoride zeka, hesaplama yetisi olarak görülmekte ve makineye düşünme ya da akıl yürütme yetisi atfedilmemektedir. Ancak bundan farklı olarak güçlü yapay zeka görüşü, insan zekasına veya bilinçli zihinselliğine eş değer ya da onu aşmış bir yapay zekanın mümkün olduğunu iddia eder. Bu teoriye göre, yeterince karmaşık programlanmış ve insan zihninin işlevsel yapısına uygun olarak tasarlanmış bir bilgisayar düşünme yetisine sahip olabilir. Zira hem bilgisayarlar hem de insan zihni belli kurallar çerçevesinde nedensel olarak yapılandırılmış girdi-çıktı ilişkileriyle çalışırlar. Doğru ve karmaşık bir yazılıma sahip yapay zeka da insan düşüncesinin temel aldığı her türlü akıl yürütmeyi yerine getirebilir (Rakover,

1999: 56). Pratik hedefin genel yapay zeka üretmek yerine var olan sistemlerin berimsel anlamda gelişimini hızlandırmak olarak belirlenmesi, vurguyu insanların yerini alacak yapay zekalardan, insana yardımcı olacak yapay zekalara çekmiştir. Bu da yapay zeka alanını yalnızca bilim-kurgu olmaktan kurtarmış ve gerçekçi bir zemine yerleştirmiştir

Yapay zeka kışının bir diğer sonucu da olasılıksal Turing Makinesi üzerine düşünülmesine yol açmış olmasıdır. İnsan akıl yürütmesinin çoğunluğu kesin olmayan önermeler ve nicelikler üzerine kurulu olduğundan, pek çok şey hakkındaki görüşümüz ve inancımız değişebilir ve koşulludur. Dolayısıyla yapay zeka sistemlerinin de insan zihnini taklit gibi bir amaçları varsa bu belirsiz bilgiyle başa çıkabilmeleri gerekmektedir. Yapay zeka kışı öncesindeki tüm çalışmalar mutlak bilgi üzerine kuruluyken bu dönemde bilginin muğlak yapısının önemi fark edilmiştir. Sonrasında ise olasılıksal grafik modelleri geliştirilerek ve olasılığa dayalı seçimler yapan programlar yazılarak bu muğlaklık aşılmaya çalışılmıştır. Olasılıksal Turing makineleri belirlenimci değildir ve "yüzde 50 olasılıkla 5. komuta, yüzde 50 olasılıkla 7. komuta atla" şeklindeki kodlamalarla çalışmaktadır. Newton fiziğini temel alan sayısal bilgisayarlar ve Turing Makinesi'nden farklı olarak kuantum fiziğini temel alan yapay zeka sistemlerinin inşa edilebileceği düşüncesi de bu şekilde ortaya çıkmıştır. Bir şeyin aynı anda iki yerde var olmasına müsaade eden kuantum fiziğinin bu tezi, bilgisayarlar için hem 0 hem de 1'i aynı anda içeren bitler anlamına gelmektedir. Bugün gelinen noktada, kuantum bilgisayarların çözülemez olan problemleri çözebileceği kesin olarak iddia edilemese de çalışma hızında önemli bir artışa yol açacağı öngörülmektedir. Bu alandaki çalışmalar hızla artarak yaygınlaşırken 1990'larda yapay zeka sistemlerinde büyük ilerlemeler yaşanmış, 1997'de IBM'in Deep Blue Programı'nın dünya satranç şampiyonu Kasparov'u yenmesiyle bu alanın kendi içerisinde bir devrim yaşanmıştır. Sonrasında ise bu çalışmalar makine öğrenmesi ve hatta derin öğrenmeye evrilmiş, sabit girdi-çıktı ilişkilerine dayalı yapay zeka sistemlerinin ilerisine geçilmiştir.

1.4. Yapay Zeka ve Sinirbilim

Beynimizde milyarlarca nöron var fakat bu nöronlar nedir?
Sadece hücreler. Nöronlar arasında bağlantı
kurulana kadar beynin hiçbir bilgisi yoktur.

Tüm bildiğimiz, tüm varlığımız,
nöronlarımızın bağlanış biçiminden gelir.
Tim Berners-Lee[15]

Yapay zeka alanında yaşanan başarısızlıkların ve umut edilen ilerlemelerin bir türlü kaydedilememesinin sebebi nörobilimcilere göre, bu alanın araştırmacılarının çok gelişmiş bir nöral ağa sahip olan beyni analog ve dijital bilgisayarlar kullanarak modellemeye çalışmasıydı. Nöral ağlar, sabit bir yapıya sahip olan bilgisayarlardan farklı olarak sürekli ve yeniden bağlantı kuran, kendilerini yeni görevler için destekleyen bir nöron topluluğudur. Buna göre, beynin programı, işlemcisi, işletim sistemi yoktur ve nöral ağlar 100 milyar nöronun tek bir amaç olan öğrenme için aynı anda kitlesel olarak ateşlenmesiyle çalışır (Kaku, 2017: 259).

Nörobilimcilere göre, yapay zeka kışlarını doğuran en önemli faktör beyni sabit girdi-çıktı ilişkilerine indirgeyen yaklaşımdır. Zihne salt berimsel yaklaşımın yetersizliğinin görülmesi ile beraberinde gerçekleşen sinirbilimdeki buluşlar, yapay zeka çalışmalarına yeni bir yön vermiştir. Beynin yaptığı işi sinir hücresi ağlarının paralel hesaplamalarıyla gerçekleştiği anlaşılmış ve paralel işleme sistemi adı verilen ağ araştırmalarına başlanmıştır (Nillson, 2018: 429). Beynin seri bilgisayarlardan çok farklı bir hesaplama modeli olduğu anlaşıldıktan sonra insan beyninin birebir taklidinin de ilk düşünülenden çok daha zor olacağı kabullenilmiştir. Beyindeki zihinsel süreçler, ağlarda birbirine bağlı olan sinirler arasındaki paralel etkileşimin sonucu olarak doğmaktadır. Dolayısıyla geleneksel seri bilgisayarlarda beynin çalışma modelinin oluşturulması mümkün değildir. Bilgisayarlar ve insan beyni arasındaki bu yapısal farklılığın keşfi, yapay zeka çalışmalarını sinir hücrelerini model alan *bağlantıcı* yaklaşım ve *simgesel* yaklaşıma dayalı *eski moda yapay zeka* olarak iki farklı sınıfa ayırmıştır.

Bağlantıcı yaklaşım, yaygın adıyla, yapay sinir ağlarının temelleri bir yönüyle John Stuart Mill'e (1806-1873) dayanmaktadır. Mill'in siyaset ve ahlak felsefesi alanına yaptığı düşünsel katkılar mantıkçı yönünün önüne geçmiş olsa da 1843 yılında yayınlamış olduğu *System of Logic* adlı eser mantığın, zihinsel süreçlerin manifestosu olarak araştırıldığı ilk çalışmadır.

[15] Tim Berners-Lee 1989 yılında CERN laboratuvarlarında HTML işaretleme dilini geliştirerek Türkçesi "Dünya Çapında Ağ" anlamına gelen World Wide Web'i (www) kurarak dünya genelinde bilgi paylaşım sistemini oluşturan bilgisayar profesörüdür.

Mantık kuralları düşüncede temellendirilir; yapay zeka ise düşünmenin bir makinede gerçekleştirilme çabalarından meydana gelmektedir. Felsefi mantık, "Makineler düşünebilir mi?" sorusunu "Zihinsel bir süreç olarak düşünmeyi mantık kurallarıyla modelleyebilir miyiz?" şeklinde güncellemiş ve yapay sinir ağları denilen yapay zeka kolunun öncülüğünü yapmıştır. Sinir ağları, zihinsel süreçlerin ve davranışın, genetikle kazanılmış veya öğrenilmiş olan tepki örnekleri ve uyaranlara bağlı olduğunu savunan bir teori olarak tanımlanmaktadır (Kurzweil, 2015: 114).

Yapay sinir ağlarının tarihi Walter Pitts ve Warren McCulloch'un birlikte kaleme aldıkları "A Logical Calculus of Ideas Immanent in Nervous Activity" makalesi ile başlatılır. Makale özetle şu iddiayı savunmaktadır:

> Sinirsel aktivitenin 'hepsi ya da hiçbirisi' karakterinden dolayı sinirsel olaylar ve aralarındaki ilişki önermeler mantığı vasıtasıyla ele alınabilir. Her ağın davranışı bu şekilde tanımlanabilir ve belirli şartları sağlayan her mantıksal ifade için, tanımlanan şekilde davranan bir ağ bulunabilir. (Pitts ve McCulloch, 1943: 115).

Pitts ve McCulloch'un amacı, Turing Makinesi'ne benzer bir kapasiteye sahip şekilsel sinir ağları oluşturmaktı. John von Neumann, bu makaleyi kendi çalışmalarının en büyük ilham kaynağı olarak alıntılamıştır. Yine Pitts'in, Jerome Lettvin ile beraber 1959 yılında yayınladığı "What the Frog's Eye Tells the Frog's Brain" isimli çalışma yapay sinir ağları alanının bir diğer kurucu faktörüdür. Yapay sinir ağları çalışmalarında en çok referans verilen iki makale olan bu tezlerdeki amaç hiçbir zaman zihnin, mekanik bir taklidini oluşturmak olmamıştır; aksine temel yönelim insan düşüncesini formelleştirme ve daha iyi anlama üzerinedir (Skansi, 2018: 5). Bu sebeple yapay sinir ağlarına ait ilk iki makale yapay zekanın değil, felsefi mantığın bir ürünüdür.

Yapay zekayı ilk kez ayrı bir bilim dalı olarak sunan Dartmouth Projesi'nin teklif metninde de yer verilen sinir ağları, proje sonunda katılımcılardan Marvin Minsky'nin *Neural Nets and the Brain Model Problem*[16] isimli doktora teziyle ilk kez bütün yönleriyle ele alınmıştır. Minsky ayrıca sinir ağlarının ilk bilgisayar uygulaması olan bir makine de inşa etmiştir. Yapay sinir ağları da sembolik yaklaşım gibi önceleri hızlı bir taraftar kitlesi bulmuş, büyük maddi destekler almış ancak 1965-1980 yılları arasında bir gerileme dönemine girmiştir. Bunun sebebi, sembolik yaklaşımın daha pratik sonuçlar vermesi ve 1947 yılında başlayan

[16] Minsky, Marvin, "Neural Nets and the Brain Model Problem", Doktora Tezi, Princeton: Princeton University, 1954.

Soğuk Savaş'ta Amerika'nın Rusça metinlerin tercüme isteğine tatminkâr cevaplar verilememesidir. Tercümede yaşanan zorluk, dilin bilinenden çok daha karmaşık bir yapıya sahip olduğunu gösterdiği gibi zihinsel süreçlerin, sinir ağları oluşturarak taklit edilmesinin tahmin edilenden çok daha zor ve uzun vadeli bir çalışma gerektirdiğini de açığa çıkarmıştır. Yapay sinir ağları, 1970'li yıllarda deyim yerindeyse eylemsiz ve olaysız bir süreç geçirmiş, 1980'den itibaren yenilenme ve canlanma dönemine girerek bugünkü önemli konumuna erişmiştir. Psikoloji ve felsefede bilişsel bilimin yükselişi ile yapay sinir ağlarının insan beynini anlamanın ötesine geçerek merkezi sinir sisteminin bilgisayar tabanlı çalışmalara yönelmesi, bağlantıcı yaklaşımı derin öğrenme sürecine götüren önemli bir yolun başlangıcı olmuştur (Skansi, 2018: 9).

İnsan beyninin farklı bölgeleri, farklı görevleri yerine getirmektedir. Duyulardan gelen bilgiler sinir sistemi tarafından beyne aktarılır ve beyinden alınan çıktılar yine sinir sistemi tarafından organlara iletilir. Bir sinir hücresi, hücre gövdesi, dendrit, akson ve sinapslardan meydana gelmektedir. Sinapslar, sinir hücreleri arasındaki elektrik sinyallerinin iletişimini sağlayan bağlantılardır. Sinyaller önce hücre gövdesine, sonra akson aracılığıyla dendrite gönderilir. Dendrit ise sinyalleri sinapslara ileterek nörotransmitterler aracılığıyla diğer hücrelere iletir. Bu şekilde çalışan milyarlarca sinir hücresinin bir araya gelmesi ile sinir sistemi oluşur. Yapay sinir ağları da biyolojik sinir ağlarının çalışma şeklini referans alan yapay sinir hücrelerinden oluşmaktadır.

Yapay sinir ağları, öğrenme yolu ile yeni bilgiler üretebilme, yeni bilgiler keşfetme gibi işleri dışarıdan destek almadan otomatik olarak yapabilme amacıyla geliştirilen bilgisayar sistemleridir (Öztemel, 2006: 29). Yapay sinir ağları, olayın örneklerinden yaptıkları genelleme ve çıkarımlarla, her girdi setine karşı bir çıktı seti üretebilir. Bu bağlamda yapay sinir ağları, biyolojik beyin fonksiyonlarının ürünü olan örnekleri kullanarak hem öğrenebilen hem de yeni çevresel koşullara vereceği tepkileri belirleyen bilgisayar sistemleridir; kendi deneyimlerini oluşturma ve karar verme kapasitesine sahiptirler. Yapay sinir ağları, birbirlerine bağlı ve paralel olarak çalışabilen yapay hücrelerden oluşmaktadır. Her yapay hücre bağlantısı bir ağırlık değerine sahiptir ve öğrenme yolu ile elde edilen bilgiler bu bağlantı değerlerinde saklanarak bir hafıza oluşturmaktadır (Öztemel, 2006: 30). İnsan zihninin en temel özelliklerinden birisi olan öğrenmeyi örnekler yoluyla gerçekleştirebilen yapay sinir ağları, sembolik yaklaşımın hesaplama yönteminden daha fazlasını vaat etmektedir. Yeni ortamlara uyum sağlayan, eksik bilgiyle çalışabilen, hatalara

karşı toleranslı olan ve bulanık mantık ile olasılık mantığının ilkelerini temel alan bu sistemler, günlük hayatın birçok alanında önemli uygulamalar oluşturmayı başarmışlardır.

1.5. Bilinçli Yapay Zekalar İçin Büyük bir Adım: Yapay Öğrenme

Yapay öğrenme ya da makine öğrenmesi, "Bilgisayarların örnek veri ya da geçmiş deneyimi kullanarak başarımlarını artıracak biçimde programlanmasıdır." (Alpaydın, 2018: 3). Bazı parametrelere dayalı olarak tanımlanmış bir model ve veri ya da deneyim üzerinde modelin başarısını değerlendirmek için tanımlanmış ölçütlerden yola çıkarak modelin parametrelerini bu ölçüte göre en iyi yapan parametre değerlerini bulmayı amaçlamaktadır. Model gelecekle ilgili tahminler yapan *öngörücü* ve/veya veriden yeni bilgiler üretebilen *açıklayıcı* bir model olabilir (Alpaydın, 2018: 3). Klasik sembolik programlamada, bir işin bilgisayarda gerçekleştirilebilmesi için bir algoritmaya[17] sahip olmamız gereklidir. Asal sayıları küçükten büyüğe doğru sıralayabilmek için asal sayılar kümesi girdi, çıktı da asal sayıların artan biçimde sıralandığı dizidir. Ancak bazı işler için bir algoritma belirlemek ya zordur ya da bilinen bir algoritma yoktur. Örneğin, istenmeyen e-postaların ayıklanıp silinmesi için belirli bir algoritma yoktur ve hangi e-postanın istenip istenmeyeceği kişiden kişiye göre değişebilir. Bu durumda amaç, eldeki istenen posta ve istenmeyen posta örneklerinden oluşturulacak bir veri kümesinden, bir e-postayı istenmeyen yapan şeyin ne olduğunu öğrenebilmektir. Yapay öğrenme ile yapılan böylesi işler için bir algoritma oluşturmaktır. Alışverişler, kredi başvuruları, borsa, yüz tanıma, konuşma tanıma gibi alanlarda yapay öğrenme, süreci tam olarak tanımlayamasa bile veri tabanlarından yararlanarak faydalı öngörülerde bulunmayı sağlamaktadır. Ne kadar çok veriye sahipsek, o kadar çok öğrenebiliriz; aynı durum yapay öğrenme için de geçerlidir. Sürekli büyümekte olan internet ağı, veri tabanlarını tahmin edilemeyecek büyüklüklere ulaştırmıştır; bu da yapay öğrenmenin isabetli öngörülerindeki yüzdeyi giderek artırmaktadır. Yapay öğrenme, yalnızca veri tabanlarından yola çıkarak tahminlerde bulunan bir uygulama değil, aynı zamanda yapay zekanın da bir parçasıdır. Bir yapının zeki olarak nitelendirilebilmesi için deneyimlerinden yola çıkarak öğrenebilmesi gereklidir; eğer o yapı değişen şartlara uyum gösterebiliyorsa programcı her şarta göre yapıyı tekrar tasarlamak zorunda kalmaz. Yapay

[17] Algoritma, girdiyi çıktıya çevirmek için uygulanacak olan komutlar kümesidir. Algoritma, gelişigüzel talimatlar dizisi değildir: Talimatlar bir bilgisayar tarafından yerine getirilebilecek kadar kesin ve belirsizlikten uzak olmalıdır (Domingos, 2017: 29).

öğrenme örüntü tanıma, öngörücü analiz, veri bilimi, adaptif sistemler gibi birçok farklı şekilde karşımıza çıkmaktadır. Bunların hepsi farklı gruplarca kullanılan yapay öğrenme teknikleridir (Domingos, 2017: 35).

Yapay öğrenme yöntemleri, yapay zeka ve robotik alanlarında büyük miktarda verileri kullanarak bilgi üretme hızını artırmakta ve bu yönüyle her geçen gün daha ön plana çıkmaktadır. Bilimsel yöntem geleneksel olarak hipotez üretme, test etme, ıskartaya çıkarma veya rafine etmek sürecini takip etmektedir; ancak bilim insanları hayatları boyunca çok sınırlı sayıda hipotez ortaya atıp test edebilirken bir yapay öğrenme sistemi bu işlemi birkaç saniyede gerçekleştirebilir (Domingos, 2017: 41). Yapay öğrenmenin bu yönü bir bakıma bilimsel keşfi otomatikleştirmek olarak yorumlanabilir ve bu, hem bilim hem de teknoloji dünyası için devrim niteliğindedir.

Günümüzde yapay öğrenmenin kullanılmadığı insani bir iş kolu kalmamış gibidir, hatta müzik, spor, resim, gurmelik gibi yapaylaştırılması imkânsız olarak görülen alanlarda bile yapay öğrenme etkisini hissettirmeye başlamıştır. Eğer ki yapay öğrenme araştırmacılarının araştırmakta oldukları fikirler tamamen uygulanabilir hâle gelirse yapay öğrenme dünya için yeni bir çağ başlatacak güçte olacaktır. Yapay öğrenmede kullanılan veri tabanına eklenen bilgiler, artık insanlardan ziyade makinelerin ürettiği bilgilerden meydana gelmektedir. Bir süre sonra makineler, kendi oluşturdukları veri tabanlarından beslenir hâle geldiklerinde, insanların makineler üzerindeki etki ve kontrolü çok azalacaktır. Bunun sonuçlarının da etik, politik, hukuki, ekonomik, sosyal ve kültürel açılardan derin bir analizi elzemdir.

1.6. Yapay Zekadan Süper Zekalara

Yapay zeka sistemlerinin somutlaşmış versiyonları üzerine çalışan robotik alanı yapay zeka konusu için inkâr edilemez bir öneme sahiptir. “Fiziksel ve zihinsel bağlamda özne nitelikleri gösteren ama biyolojik anlamda hayatta olmayan inşa edilmiş sistemlere” robot ismi verilmektedir (Ersoy, 2017: 6). Tek bir robot tanımı olmasa da robotik cihazlar önceden kendilerine tanımlanmış olan yönergeye göre görevlerini gerçekleştirebilen ve çevreleriyle etkileşim içerisine girebilen makinelerdir. Yaygın olarak kabul edilen güncel bir robot tanımına göre, robotlar üç temel bileşene sahiptir: Çevreyi gözlemleyerek değişiklikleri tespit eden sensörler, nasıl karşılık vereceğine karar veren işlemciler veya yapay zeka ve bir

robotun çevreye tepkisini ortaya koyan uygulayıcılar (Singer, 2015: 87). O hâlde yapay zeka dediğimiz de bir robotun karar veren tarafıdır.

Yirminci yüzyıla gelindiğinde otomatlar giderek gelişmiş ve insansı robotlar geliştirilebilir düzeye erişmiştir. Öyle ki robotlar artık insana yardımcı makineler olmanın ötesinde insanı aşabilecek varlıklar olarak kurgulanmaya başlamıştır. Önceden tanımlanmış talimat ya da yönergelere göre çevresini yöneten, kendi kendine hareket eden ve görevlerini gerçekleştiren elektromekanik âletlere karşılık olarak kullanılacak olan robot sözcüğü ilk kez 1920'de Karel Capek'in (1890-1938) *Rossum'un Evrensel Robotları* adlı oyununda kullanılmıştır. Etimolojik kökeni Çekçe olan robot sözcüğü, bu dilde "zorunlu iş gücü", "zorla çalıştırılma" ve "ağır iş" anlamlarına gelmektedir. Rossum'un Evrensel Robotları oyunundaki robotların dikkat çekici özelliği ise her şeyi hatırlayabilen hafızaya sahip olmalarına rağmen düşünememeleridir. Düşünen robot fikrinin ortaya çıkması bu tarihten yaklaşık 30 yıl sonra gerçekleşecektir.

Isaac Asimov'un robot öykülerini bir araya topladığı 1950 yılında yayınlanan *I, Robot* kitabı, robot kavramını ve insanlar ile robotlar arasındaki ilişkinin doğasını sorgulamaya yönelik ilk bilimkurgu eseridir. Robotlar geliştikçe ortaya çıkacak problemleri düşünen, farklı robot türleri kurgulayan ve "İnsan nedir?" sorusunu robotlar üzerinden tartışan Asimov robotların çalışma prensiplerini Üç Robot Kanunu'yla belirlemiştir:

1- Robotlar, insanlara zarar veremez ya da eylemsiz kalarak onlara zarar gelmesine göz yumamaz.
2- Robotlar, Birinci Kanun'la çelişmediği sürece insanlar tarafından verilen emirlere itaat etmek zorundadır.
3- Robotlar, Birinci ya da İkinci Kanun'la çelişmediği sürece kendi varlıklarını korumak zorundadır.

Üç Robot Kanunu aynı zamanda şimdilerde *robot etiği* olarak adlandırılan çalışma alanının çıkış noktası niteliğini taşımaktadır. Etik bir otonom robotun inşa edilirken konulması ve uygulanması gereken kuralları tespit amaçlı düşünceler Asimov'un Üç Robot Kanunu'na katkı ya da eleştiri olarak oluşturulmaktadır.

Stephen Hawking 2014 yılında yayınladığı bir makalede, gerçekten düşünme becerisine sahip bir bilgisayarın "insanlık tarihindeki en büyük olay" olacağını, insandan daha zeki bir bilgisayarı bilim-kurgu ürünü olarak görüp dikkate almamanın "tarihteki en büyük hatamız"

olacağını söylerken bugünün popüler tartışması olan tekilliğin fitilini ateşlemiştir.[18] Var olan robotik teknoloji, *özel* veya *dar* olarak adlandırılan yapay zeka sistemlerini kullanmaktadır. Dar yapay zeka, belirli bir veya birkaç görev için programlanmış zeka anlamına gelmektedir. En karmaşık bilgisayar oyunlarında bile insanlardan çok daha hızlı ve başarılı bilgisayar programları, beste yapan hatta konser veren yapay zekalar, sürücüsüz akıllı otomobiller, yemeğin fotoğrafından yemek tarifi veren telefon uygulaması, insanlardan çok daha iyi dudak okuyan yapay zeka, birbirleriyle bir iletişim dili geliştiren, bunu yaparken tek kelimelik basit komutlardan başlayıp cümle kurmaya geçen ve hatta kendi dillerine bir gramer yapısı oluşturabilen yapay zekalar, kalp atış hareketlerinden kalp krizi geçirme risk oranını ve başka biyolojik verilerden kanser olma risk oranını hesaplayabilen yapay zekalar, rapor yazabilen, borsada işlem yapabilen, kutu taşıyan yapay zekalar, bilgisayar görüşü kullanarak görme engellilere çevredeki kişileri ve objeleri tanımlayabilen yapay zeka temelli mobil uygulamalar ve bunlara her geçen gün yenisinin eklendiği birçok gelişme, yapay zekanın her an hayatın içinde karşılaştığımız *işlevsel* ancak ***dar*** uygulamalarıdır. Bu yapay zekalara sahip makinelerin yaptıkları işler, rutin ve tahmin edilebilir görevlerden meydana gelmektedir. Zira programlandıkları görevden başkasını yapma kabiliyetine sahip değillerdir. Çok yakın zamana kadar da yapay zeka yatırımlarının büyük çoğunluğu, bu şekilde dar alanda uzmanlaşmış sistemler üzerine yapılmıştır.

Stephen Hawking'in insanlık tarihindeki en büyük olay olacaktır tahminini yaptığı yapay zeka ise bu şekilde dar bir yapay zeka değil; yeni fikirler ortaya atabilen, daha önce karşılaşmamış olduğu ortamlara adapte olabilen, kendisi için "yeni" olarak adlandırılabilecek problemlere çözüm üretebilen, geniş anlamda bir zekaya ve düşünme becerisine sahip yapay zekalardır. Dar-geniş yapay zeka ayrımı, zayıf-güçlü yapay zeka ayrımının başka bir yorumudur. Yapay zeka üzerine araştırma yapan 200 kişi ile Genel Yapay Zeka üzerine yapılan bir anketin sonuçları, dar yapay zeka araştırmalarından geniş yapay zeka çalışmalarına geçiş sürecinde olduğumuzu açıkça ortaya koymaktadır. Buna göre, ankete katılanların %42'si düşünen bir makinenin 2030 yılında mümkün olacağını düşünmekte, %25'i bu tarihin 2050 olacağını, %20'si 2100 olacağını söylemekte; yalnızca %2'lik bir kısım ise düşünen bir makinenin asla var olmayacağına inandıklarını belirtmektedir (Barratt, 2013: 196-197). Birçok araştırmacıya göre, Google, Amazon,

[18] Stephen Hawking, Stuart Russell, Max Tegmark ve Frank Wilzeck, "Stephen Hawking: 'Transcendence Looks At the Implications of Artificial Intelligence- But Are We Taking AI Seriously Enough?'". *The Independent*, 1 Mayıs 2014. https://www.independent.co.uk/news/science/stephen-hawking-transcendence-looks-at-the-implications-of-artificial-intelligence-but-are-we-taking-9313474.html.

Facebook gibi yapay zekayı iş modellerinin merkezine koyan şirketlerin yükselişi ve hatta yapay zekalar üzerinden bir rekabete girmiş olmaları ***geniş yapay zekanın*** yaklaştığına dair sinyaller olarak algılanmalıdır. Benzer rekabetin ülkelerin yapay zeka fonlamaları ve politikaları arasında da yaşandığı göz önüne alınırsa artık süper yapay zekalar çağına doğru bir sıçrayışta olunduğunu söylemek yanlış olmayacaktır.

Yapay zekanın geleceğine dair tartışmalarda yaygın kanı şudur ki, genel yapay zeka bir kez başarıldığında insan zekasını aşan yapay sistemlerin gelişimi çok hızlı bir şekilde gerçekleşecektir. Kendisini geliştirebilen, tasarımını yenileyebilen, yazılımını güncelleyebilen bir yapay sistemin kendisinin daha üst bir sürümünü, o sürümün de aynı teknikle kendisinin daha üst bir sürümünü üretmesi mümkün olacak ve her revizyonda sistem daha zeki hâle gelecektir (Ford, 2018: 264). Bu döngünün insandan milyonlarca kat daha zeki bir makine ile sonuçlanması da bu açıdan şaşırtıcı olmayacaktır.[19]

İlk kez bilgisayarların öncü ismi kabul edilen John von Neumann tarafından kullanılan "tekillik" kavramı, süper zekaların inşası ile gerçekleşme ihtimali olan teknoloji çağına verilen isimdir. Günümüzde ise tekillik denilince akla gelen ilk isim kuşkusuz Ray Kurzweil'dır. Tekilliğin gerçekleşeceği tarihi 2045 olarak işaret eden Kurzweil'a göre, teknolojik tekilliğin ötesinde insan zekası çeşitli beyin implantlarıyla gelişecek ve insan ile makine birleşecektir. Bu birleşimin sonucu da insanın ölümsüzlüğü ile sonuçlanacaktır; biyolojik bedenimizi ölümsüz kılabilecek teknolojilerin yanı sıra insan zihninin bir makineye yüklenmesi ile zihinsel ölümsüzlük de mümkün olacaktır.

Yapay zeka gelişiminin denetlenemeyecek bir zeka türü ortaya çıkarmasından duyulan kaygı ile bu alanda yapılacak tüm çalışmaların "dost yapay zeka" olarak adlandırılan etik yapay zekaya yönelik olması gerektiği ilgili çevrelerce sürekli olarak vurgulanmaktadır. Dost olmayan güçlü yapay zekadan korunmak için stratejiler geliştirilmektedir. Bu yaklaşıma göre, kendi kendini geliştiren güçlü yapay zeka, başlangıçtaki durumuna bir daha döndürülemeyeceği için ilk seferinde ve hatasız olarak insana dost, insan değerlerini destekleyen bir yapay zeka üretilmelidir (Kurzweil, 2017: 613). Ancak daha bugünden geleceğin yapay zekasının insanların sahip olduğu etiğe ve değerlere sahip olmasını sağlayacak stratejiler üretmek mümkün görünmemektedir. Ayrıca bu alandaki her kural ve

[19] Süper zeka, genellikle hep makine temelli olarak düşünülse de biyolojik temelli bir süper zeka da tahayyüller arasındadır. İnsan zekasının teknolojiyle geliştirilmesi ve genetik mühendisliğinin ilerlemesi ile de süper zekaya ulaşılması da tartışılan bir konudur.

uygulamanın tüm devletler tarafından istisnasız kabulü ve işleyişi gereklidir ki, gücü ele geçirme arzusu "dost yapay zeka" stratejilerine çok da müsaade etmeyecektir. Bu anlamda, ölümcül yapay zekaların kullanılması insanlık adına korkutucu senaryolara kaynaklık etmektedir; çünkü öyle gözüküyor ki devletlerin bırakmak istemedikleri güç yarışı, yalnızca etik ilkelere tabi yapay zekalarla sınırlı bir üretim anlayışını benimsemelerine engel olacaktır.

Yapay zeka üretiminin dost yapay zekalar ile sınırlandırılması mümkün olsa bile insan zekasını aşmış makinelerin olacağı bir dünyada, bu durumun istihdam ve ekonomi üzerinde meydana getireceği olumsuz etkiler de yapay zekanın geleceğine yönelik tartışılan hususlar arasındadır. İnsanların alacakları eğitim makinelerle rekabet etmelerini sağlamak için yeterli olmayacak, yalnızca insana özgü olduğu düşünülen iş sahaları dahi riske girecektir (Ford, 2018: 264). Bugün rutin ve öngörülebilir işlerden oluşan mesleklerin yerini makineler almaya başlamışken güçlü yapay zeka ile beraber yaratıcılık gerektiren iş kollarının da makinelerin kontrolüne geçmesi ihtimal dâhilindedir.

Yapay zekanın halk psikolojisindeki karşılığı şu cümleyle ifade edilebilir: Bilgisayarlar, makineler ve/veya robotlar insan zekasına erişecek ve hatta insan zekasını aşacaklar. Yapay zeka araştırmaları, sürekli olarak bu iddiayı destekler nitelikte gelişmelere sahne olmaktadır. Belki 10 yıl öncesinde sadece fısıltı olan süper zeka söylemleri, bugünlerde Google'ın DeepMind[20] projesi ile AlphaGo ve sonrasında AlphaGo Zero'nun başarısı, kendilik bilinci testini[21] geçen robotlar ve tekillik kavramı üzerine sürekli yenilenen senaryolar ile yüksek sesle dile getirilmeye başlanmıştır. Tekillik ve onu destekleyen bilimsel kanıtlar her gün başka bir kanaldan servis edilmeye devam ederken hâlâ cevap bekleyen temel ve kadim felsefi problemler olması da düşündürücüdür: "İnsan zekasını aşmak ne demektir? İnsan

[20] DeepMind, Google tarafından işletilen bir yapay zeka şirketidir. Alphago isimli program, belli kurallar çerçevesinde makine öğrenmesine dayalı olarak dünyanın en zor oyunlarından birisi olarak kabul edilen "Go" oyununda, dünya şampiyonunu yenmiştir. Bunun üzerine DeepMind şirketi, Alphago'ya yüklenen en baştaki kurallar olmadan sıfırdan (*tabula rasa*'dan), pekiştirmeli öğrenme yöntemine dayalı bir algoritma yazarak AlphagoZero'yu oluşturmuştur. (Yazılan algoritma go algoritması değil, pekiştirmeli öğrenme algoritmasıdır, oyunla ilgili bir bilgi verilmemiştir.) AlphagoZero ise sadece öğrenme algoritmasını kullanarak, Go öğrenme algoritmasıyla çalışan AlphaGo'yu defalarca kez yenmeyi başarmıştır. AlphagoZero'nun çalışma şekli ve algoritmasının ayrıntılı anlatımı için bkz. Silver, David, Hubert, T., Schrittwieser, J., Antonoglou, I., Lai, M., Guez, A., Lanctot, M., Sifre, L., Kumaran, D., Graepel, T., Lillicrap, T.P., Simonyan, K., ve Hassabis, D. (2017). "Mastering Chess and Shogi by Self-Play with a General Reinforcement Learning Algorithm". *CoRR*, abs/1712.01815.

[21] Bringsjord, Selmer, Licato, J., Govindarajulu, N.S., Ghosh, R., Sen, A. (2015). "Real robots that pass human tests of self-consciousness". *Proceedings of IEEE the 24th International Symposium on Robots and Human Interactive Communications*. 498-504.

zekası ne demektir ve bilinçle bağlantısı nedir?" Bir görüşe göre, süper zekalardan bahsedilmemiz için beyni ya da insan bilincini taklit eden bir sistem geliştirmeyi beklememize gerek kalmamıştır çünkü bilgisayarlar insanın hesaplamaya dayalı akıl yürütme yeteneklerini ve hızını çoktan aşmış görünmektedir. Başka bir görüş ise süper zekaların zorunlu şartının bilinç olduğunu vurgulamaktadır.

Felsefenin yapması gereken yapay zekayı ölümsüzlük varsayımı, yeni bir çağın başlangıcı, insanlığın sonu gibi kurgusal varsayımlardan soyutlayarak tartışmaktır. Zira felsefi argümanlar tahminden çok açıklama gücüne dayanmalıdır. Yapay zeka felsefesi genel olarak yapay zekanın başka bir deyişle akıllı düşünebilen bir makinenin mümkün olup olmadığı meselesi üzerine kuruludur.

Problemleri ele alış biçimleri bakımından yapay zeka felsefesinin temel tartışmalarını dört grup altında toplamak mümkündür. Birinci grup düşüncenin anlamsal içeriğini tartışmaya açarak yapay zekanın gerçekleşme ihtimalini yok saymaktadır. İkinci grup rasyonalitenin doğası üzerinden yapay zekaya karşıt nitelikte argümanlar ileri sürmektedir. Üçüncü gruptaki yapay zeka felsefecileri, meseleyi insanın aşkın muhakeme gücü üzerinden ele alarak yapay zekanın gerçekleşebilme ihtimalini değerlendirmeye tâbi tutmaktadır. Son gruptaki yapay zeka tartışmaları ise akıllı bir makinenin fiziksel yapısı üzerinden ilerlemektedir. Düşünen bir bilgisayarın kullanması gereken programın ayrık mı bütünsel mi çalışması gerektiği, veri bağlantılarının nasıl sağlanacağı üzerine yapılan analizlerle ilerleyen bu son grubun çalışmaları çoğunlukla bilişsel bilimlerden beslenmektedir.

Yapay zeka problemine yaklaşımları bakımından farklı hususları ele alan yapay zeka felsefecilerinin ortak paydada tartıştıkları en önemli sorulardan birisi bilinçli bir makinenin mümkün olup olmadığıdır. Bu soruya dikkatle odaklanacak olursak altında yatan tetikleyici problemin bilincin mekanik bir açıklamasının yapılıp yapılamayacağı tartışması olduğunu görmekte güçlük çekmeyiz. Bilinç, doğası gereği bilim tarafından her zaman çekinceli yaklaşılan bir fenomen olmuştur. Fiziksel dünyanın özellikleriyle uyuşmayan niteliklere sahip olan bilinç, hakkında ortaya konulan teorilerde karşıt bakış açılarının birbirleriyle yarıştığı bir problematik konumundadır. Bilincin neliği ve mahiyeti üzerine mutabakata varılmış bir kuramın olmayışı, onun bilimin nesnesi hâline gelmesini geciktirmiştir. Ancak zihin felsefesinin bilinç hakkında ortaya koyduğu önemli bakış açıları ile nörobilim ve bilişsel bilimlerin bilinç üzerine incelemeleri, bilincin bilimin sınırları içerisinde

tartışılabilecek ve açıklanabilecek bir doğasının olduğunu kuvvetli bir ihtimale dönüştürmeyi başarmıştır.

Bilimsel bilinç araştırmalarına katkı sağlayan alanlardan birisi de bu çalışmanın ana konusunu oluşturan yapay bilinç alanıdır. Zihnin ayrıştırılmaz bir niteliği olarak ele alınan bilinç, insan zihnine bütünüyle benzeyecek bir yapay zekanın da temel bir niteliği olacaktır. Bu sebeple bilincin doğasına yönelik olarak kabul edilen ve içerisinde çalışılan her yaklaşım ve içerdiği kabuller, yapay bilincin imkânına yönelik yaklaşımları da doğrudan etkileyecektir. Doğal bilinç için kabul edilen kuram ve varsayımlar, yapay bilinç için de geçerliliğini koruyacaktır. Nihayetinde yapılması gereken bilinç problemini olası tüm boyutlarıyla incelemek olduğundan, yapay bilinç tartışmalarına geçmeden önce, çalışmanın bundan sonraki bölümünde zihin felsefesindeki bilinç teorileri müzakare edilecektir. Klasik bilinç teorilerinin ardından, güncel bilinç teorileri de ele alınacak ve araştırmaya yapay bilince yönelik yaklaşımlar ile devam edilecektir. Ortaya çıkacak olan tablodan hareketle yapay bilincin imkânı değerlendirilecek ve bu değerlendirme teknik ve epistemolojik boyutlarının yanı sıra etik, sosyal ve hukuki kaygılarla da yorumlanmaya çalışılacaktır. Son olarak ise bilince bütüncül bir açıklama sunabilmek adına özgün bir bilinç yaklaşımı oluşturulmaya çalışılacak, bu yaklaşımın yapay bilince uygulanması ve bu uygulamanın muhtemel sonuçları gözden geçirilecektir.

2. BİLİNÇ PROBLEMİ: YAKLAŞIMLAR VE TEORİLER

Bilincin varlığı dünyayla ilgili hem en bildik
hem de en şaşırtıcı şeylerden birisidir.
Thomas Nagel

Önceki bölümde belirtildiği üzere, güçlü bir yapay bilinç teorisi yapay zeka ve bilinç kavramları üzerine net bir bakış açısı ortaya koymalıdır. Yapay zekanın fikrî temelleri, mantık-yapay zeka ilişkisi, bağımsız bir bilim dalı olarak yapay zekanın ortaya çıkışı, yapay zeka-nörobilim etkileşimi, yapay öğrenme ve yapay zekaya bağlı gelecek projeksiyonları giriş bölümünde ana hatlarıyla ele alınmıştır. Yapay zekanın neliğine dair bu genel kavrayıştan sonra, üzerinde durulması gereken ikinci husus ise bilinç olarak karşımızda durmaktadır. Bilinç mefhumu bu bölümde klasik teorilerle, üçüncü bölümde ise güncel felsefi teoriler ile birlikte eleştirel olarak ele alınacak ve yapay bilinç için zorunlu koşullar olan yapay zeka ve bilinç yaklaşımlarına dair ön bilgiler tamamlanmış olacaktır.

Öznel ve niteliksel deneyim olarak ifade edilebilecek fenomenal bilinç, bizzat insan varoluşuyla ilgili bir problemdir. Bu sebeple, Galen Strawson'un ifadeleriyle öznel deneyimin varlığının yok sayılması ya da inkâr edilmesinin "insan düşünce tarihinin tümünde gerçekleşen en garip şey" olarak nitelenmesi abartılı bir yaklaşım olmayacaktır (Strawson, 2006: 5). Bilincin deneyimden ayrı, deneyimin ise deneyimleyen ve deneyimlenenden ayrı düşünülmesi mümkün değildir. Descartes'ın "Düşünüyorum, o hâlde varım" ilkesi bilinçli benin indirgenemez bir töz olduğu iddiası ile 'düşünen ben'in (*ego cogito*) aynı zamanda 'bilinçli bir ben'in varlığını, başka bir deyişle de deneyimleyenin varlığını zorunlu kılar. Bu klasik bilinç yaklaşımının izleri çağdaş zihin felsefesi tartışmalarında da açıkça görülmektedir. David Chalmers'ın zor problemi, Ned Block'un fenomenal bilinç tanımı, David Rosenthal'ın "Bir zihinsel durumun bilinçli olması, birinin uygun bir şekilde bu durumun bilincinde olmasına dayanır." fikri, birinci şahıs/üçüncü şahıs ayrımı ile ilgili problemler deneyimleyen olmadan deneyimin imkânsızlığı ön kabulünün farkında olan önemli argümanlardır (Rosenthal, 2008: 832). Deneyimin fenomenal bilince karşılık geldiği bu denklemde, deneyimleyen fenomenal özneye karşılık gelmektedir. Fenomenal özne kavramsal, düşünsel ya da ideal bir varlık değil, gerçek bir varlıktır. Fenomenal öznenin varlığı kabul edilmeden bilinçten ve deneyimden söz etmek anlamsız olacaktır. Dolayısıyla hem deneyimin hem de deneyimleyen ve deneyimlenenin doğasını anlamaksızın bilincin doğasını anlamak mümkün görünmemektedir.

Sistematik bir yöntemle ifade edilecek olursa, bu çalışmanın beş temel ön savı bulunmaktadır:

1- Fenomenal bilinç, deneyimdir.
2- Deneyim için hem deneyimleyen hem de deneyimlenenin varlığı zorunludur.
3- Hem deneyimleyen hem de deneyimlenen ideal ve düşünsel varlıklar değil, gerçek varlıklardır.
4- Deneyimleyeni ve deneyimleneni anlamaksızın ne deneyimin kendisi ne de fenomenal bilincin doğası kavranabilir değildir.

Bu beş ön iddia şu anlama gelmektedir: Bilincin doğasını açıklama teşebbüsünde bulunan teorilerin hiçbiri, deneyim, deneyimleyen ve deneyimlenenin doğasını açıklamaksızın bilincin doğasını açıklayamaz. Deneyimin düğümü çözülmeden bilincin ipliklerine ulaşmak mümkün olmayacaktır.

Bu bölümde, çağdaş zihin felsefesinde bilincin doğasına yönelik olarak sunulan dört temel yaklaşım bahsi geçen beş ön sav çerçevesinden ele alınacaktır. Bu yaklaşımlar *monizm, düalizm, panpsişizm* ve *agnostik gizemcilik* olarak belirlenmiştir. Bu yaklaşımların ortaya koyduğu teorileri birbirleri ile kıyaslarken fenomenal deneyimin ontolojik statüsüne getirdikleri açıklama biçimleri bu teorilerin bilinci hangi boyutlarıyla açıklayabildikleri konusunda bir ölçüt teşkil edecektir.[22]

2.1. Monizm (Tekçilik)

Monizm (tekçilik), evrenin tek tip tözden oluştuğu fikrine dayanmaktadır. Evreni oluşturan bu tek tözün doğası üzerine sunulan farklı iddialar monizmin kendi içerisinde farklılaşmasına neden olur. Var olan tek tözün "maddi olmayan" ya da "zihinsel nitelikte" olduğunu savunan idealizm ile var olan tek tözü "madde" ya da "maddi nitelikte" olarak kabul eden materyalizm, monizmin farklı tezahürleridir.

[22] Bu yaklaşımların her birisinin ayrı bir tez konusu olarak çalışılabilecek geniş bir kapsamı olduğu, bu çalışmanın amacı gereği burada yalnızca öne sürdükleri temel iddialar ve kendilerine karşılık yapılan temel itirazlarla değerlendirildikleri göz önünde bulundurulmalıdır.

2.1.1. İdealizm

Var olan tek tözün "maddi olmayan" olduğunu iddia eden idealizme göre bilinç, bu zihin tözünün temel karakteristiğidir. Zihnin dışarısında kalan tüm fiziksel âlem ve madde idealist bakış açısıyla yalnızca bir yanılsama olarak kabul edilebilir. İdealizmin varlığını inkâr yoluna gitmediği tek gerçeklik bilinçli durumlardır.

Dış dünyanın gerçekliğinin apaçık görünüşüne karşın maddenin varlığını kabul etmeme yüzeysel bir bakış açısıyla ele alındığında savunması güç bir yaklaşım gibi görünmektedir. Ancak fiziksel dünya hakkında sahip olduğumuz tüm bilgilerin gözlem ve deneyime dayalı olduğu göz önüne alındığında tüm fiziki nesnenin bilinçli algılarımızın üretimi olan öznel deneyimlere dayandığı yorumuna ulaşmak zor değildir. Her gözlem, her deney ve her deneyim zorunlu olarak bir bilinçli algıyı gerektirir. Bilinçli bir algısal deneyime sahip olmadan fiziksel bir şeyin varlığını bilmek mümkün değildir. İdealizmin en güçlü temsilcilerinden Berkeley'in "Var olmak, algılanmış olmaktır." mottosundan yola çıkıldığında fiziksel dünyanın yalnızca bilinçli algılarımızla var olduğu sonucuna ulaşılırken fiziksel olanın bağımsız bir varlığa sahip olamayacağının kolaylıkla iddia edilebileceği görülecektir (Berkeley, 1984: 13). Fiziksel dünyanın bağımsız olarak varoluşuna dair elimizde bulunan tüm kanıtlar gözlem ve deney yoluyla elde ettiğimiz birinci şahıs deneyimlerimizden kaynaklanmaktadır ve dolayısıyla doğrudan kanıtlar değil, dolaylı çıkarımlardır.

Var olmanın koşulunu algılanıyor olmak olarak belirleyen Berkeley'e yöneltilebilecek en önemli soru, algılayanın kim olduğudur. Berkeley'e göre maddi dünya sadece tutarlı bir rüyadır ve bu rüya da aslında Tanrı'nın rüyasıdır. Fiziksel âlem olarak tanımladığımız evren özünde Tanrı'nın zihnindeki büyük rüyadan ibarettir. Bu görüşte var oluşun özü maddi değil, zihinseldir. Berkeley'in idealist bakış açısı, bildiklerimizin algılarımızdan ibaret olduğunu varsayarken deneyimin ötesinde bir şeyin varlığına dair bir bilgimizin olmadığını ve aynı şeyleri algılıyor olmamızın da fikirlerimizin Tanrı'dan gelmesine dayandığını iddia etmektedir (Özakpınar, 2011: 35). Berkeley'in fiziksel âlemin Tanrı'nın zihnindeki rüya olduğu çıkarımından farklı olarak idealizmin başka bir versiyonu dünyayı benim bilincimdeki bir rüya olarak tanımlamaktadır. Tekbencilik (solipsizm) olarak da adlandırılan bu görüşe göre algılarım benim bilinçli deneyimlerimdir ve dolayısıyla yalnızca benim bilincim ve onun içeriği vardır. Diğer insanlar ya da diğer zihinler de yalnızca benim zihnimdeki rüyanın bir parçasıdırlar.

İdealizm, dış dünya ve nesnenin varlığını bilinçli öznel deneyimlerimizden kaynaklı bir yanılsama olarak yorumlarken var olan tek şeyin zihinsel gerçeklik olduğunu dile getirir. Bu iddia aynı zamanda dil, din, ırk, cinsiyet, yaş, kültür vs. fark etmeksizin ortaya çıkan bir evrensel yanılsama iddiasıdır. Böyle büyük bir evrensel yanılsama, fenomenal öznenin varlığına ve açıklamasına ilkece imkân tanıyor gibi görünse de idealizmi kabul etmek en az reddetmek kadar zor görünmektedir.

İdealizmin ortaya koymuş olduğu sadece maddi olmayanlardan oluşan evren anlayışı, bilincin varlığını tek mutlak gerçeklik olarak yansıtıyor olsa da bu yaklaşım bilincin ve dolayısıyla bilinçli öznenin doğasına dair herhangi bir açıklama modeli sunmamaktadır. Bu bağlamda idealizmi üç farklı açıdan eleştirmek mümkündür.

İlk olarak, idealizme göre eğer varlık dediğimiz şey sadece zihin ve zihnin içeriğinden ibaretse o zaman her şey aslında öznenin bir parçası demektir. Çünkü her şey onun algılarından ve tasarımlarından ibarettir. Bu anlamda var olan tek şey öznedir ve özne-olmayan herhangi bir şeyin varlığından dahi söz edilemez. Bir idealist, bu iddiaya zihnin içeriğini özneden ayırabiliriz önerisiyle karşı çıkacaktır. Bu öneri dikkate alındığında şöyle bir tablo ortaya çıkmaktadır: Var olan tek şey zihindir ve zihin de özne (deneyimleyen/algılayan) ve zihinsel içerik (deneyimlenen/algılanan) olarak ikiye ayrılmaktadır. Örneğin, zihnin içeriğinin tek tek boşaltıldığı ve zihnin içeriğinde hiçbir şey kalmadığı bir düşünce deneyi tasarlandığında geriye yalnızca bir özne mi kalacaktır? İdealizmin bu soruyu kesin ve net bir dille cevaplaması mümkün değildir.

İkinci olarak, ilk eleştirideki sorunun, öznenin zihnin içeriğinden bağımsız olarak bir varoluşa sahip olduğu ve zihnin içerikleri tek tek boşaltısa bile geriye öznenin kalacağı şeklinde cevaplandığı düşünüldüğünde özne ve zihnin içeriğinin ayrımının nasıl yapılacağı yeni bir sorunsal oluşturmaktadır. Zihinsel içeriği incelerken nerede "işte bu içerik, işte bu da öznenin ta kendisi" denilecek olan durumla karşılaşılacaktır? Zihin içeriğinin öznenin kendisinden ayrılması oldukça güç görünmektedir, zira idealizm bu ayrımı yapmamızı sağlayacak hiçbir ayrırt edici ölçüt önermemektedir. Hume'un ortaya koyduğu problem de buna işaret etmektedir. Hume'a göre ne zaman özneyi zihnimizin içinde yakalamaya çalışsak tek elde edebildiğimiz şey algı parçacıkları, yani zihin içerikleridir. Ayrıca bir özne ya da kendilik bulabilmemiz söz konusu değildir. Hume'un deneyim ve deneyimleyeni yok sayarak sadece algı demetlerini dikkate aldığı bakış açısı şu şekilde yorumlanabilir: "Bilinçli bir insanın zihinsel yaşamında, hâlihazırdaki algı parçalarından bağımsız herhangi bir

kendilik ya da ben yoktur. Her bir insanoğlu, fenomenal anlamda bir algı demetinden öte bir şey değildir." (Arıcı, 2015: 67). İdealizm zihnin içeriğinden bağımsız bir özne, kendilik ya da ben olabileceğini iddia edecek olsa da bu öznenin doğasına dair herhangi bir açıklama getiremeyeceğinden, içerik ve özneyi birbirinden ayırmanın yolunu gösteremeyecektir. Bu anlamda idealizm, bir sonraki eleştiride detaylandırılacak olan, materyalizmin fiziksel özne tasavvurunu öne sürme imkânına da sahip değildir.

İdealizme bilinç içeriklerinin doğası üzerinden getirilebilecek üçüncü bir eleştiri de ileri nörofiziksel yöntemlerle neredeyse emin olduğumuz şu önerme üzerinden temellendirilebilir: "Bilinçli deneyimi oluşturan zihin durumlarının ortaya çıkması için bazı nöral aktiviteler gereklidir." Bilinçli zihin içerikleri beynin bilgi işleme süreçlerinde nöral olaylarla birlikte ortaya çıkar. Beyin ve sinir sistemi bütün olarak yok sayıldığında ise bilincin ve bilinçli öznenin nasıl var oldukları, doğalarının ne olduğu, nasıl bir işleyişe sahip oldukları gibi sorular havada ve anlamsız kalacaktır. İdealizm bu hususta, zihni ve özneyi fiziksel yollarla açıklama girişiminde bulunan materyalizm ile boy ölçüşecek argümanlara sahip değildir.

2.1.2. Materyalizm (Maddecilik)

Var olan her şeyin maddi olduğu fikrine dayalı olarak temellendirilen materyalizm doğa bilimlerine de sırtını yaslayarak zihin felsefesinin uzun yıllar hâkim görüşü olmuştur. Düalizmden türemiş bir bakış açısı olarak zihin ve bedenin karşıt ontolojik kategorilere ait olarak konumlandırılışı, felsefe tarihine çözülmesi zor bir düğüm olan zihin-beden problemini bahşetmiştir. Bu düğümü zihinsel olanı ayrı bir töz olarak ele almayı redderek çözmeye çalışan materyalizm, zihinsel olanın doğasına dair ortaya atılan argümanlarla kendi içerisinde farklı kollara ayrılmaktadır. Materyalizmin zihnin neliğine dair görüşleriyle paralel olarak, bilincin mahiyetine yönelik görüşleri de ayrık bakış açılarından tezahür etmektedir. Materyalizmin bilince yönelik farklı temel üç yaklaşımı materyalist teorilerin omurgasını oluşturmaktadır. Bu yaklaşımlardan ilki bilincin gündelik dil ve düşünüş şekli tarafından üretilmiş olduğu argümanına bağlı olarak bilincin gerçekliğini ve varlığını bütün olarak reddetmektedir. İkinci yaklaşım, bilincin varlığını kabul etmekte ancak bilinci beyinde gerçekleşen nörolojik olaylara indirgemektedir. Üçüncü yaklaşımda ise bilinç yine

fiziksel bir varlık olarak kendisine yer bulmakta ve beynin özel ya da üst düzey bir etkinliği olarak indirgemeci olmayan bir tavır ile tanımlanmaktadır.

2.1.2.1. Elemeci Materyalizm

Elemeci materyalizm (*eliminative materialism*), bilincin *halk psikolojisine*[23] ait gündelik dil ve düşünüş biçimlerinin sonucu ortaya çıkmış olan yanıltıcı bir kavram olarak dikkate alınmaması gerektiğini öne süren, halk psikolojisi kavramlarının zamanla kuramsal nörobilim kavramları ile eşleşeceğini savunan indirgemeci materyalizmden, halk psikolojisi kavramlarının yalnızca eksik değil aynı zamanda yanlış tasarımlar olduklarını söyleyerek farklılaşan ve bilincin varlığını bütüncül olarak reddeden radikal bir yaklaşımdır (Churchland, 2012: 69). Elemeci materyalizme göre "zihin" ve "bilinç" gibi kavramlar "halk psikolojisi" teorisinin kavramlarıdır, halk psikolojisi de gözlemlenebilen davranış ile gözlemlenemeyen zihin arasındaki bağıntıdan bahseden, hepimizin günlük hayatta kullandığı tabirlerden oluşan ilkel bir psikoloji teorisidir (Sayan, 2012: 50).

Elemeci materyalizm, bilincin bilimin muhtevası içerisinde yeri olmayan bir kavram olduğunu iddia ederken halk psikolojisine ait kavramların bilimden elenmesi gerektiğine yönelik tezini bilim tarihinde yanılsama olduğu saptanan kavram ve fenomenleri örnek göstererek temellendirmektedir. Elemeci materyalizm iddiasını gerekçelendirme yöntemi olarak bilim tarihinde benzer olduklarını düşündükleri olaylar ile örneklendirmeyi tercih etmektedir.

Elemeci materyalistlerin sıklıkla atıfta bulundukları örneklerin başında, kimya tarihinde ruhsal bir fenomen olarak varlığına inanılan *filojiston* gelmektedir. Bir şey yandığında veya bir metal paslandığında filojiston denilen gizemli bir maddenin ortaya çıktığına inanılmaktaydı. İnanca göre, bu gizemli madde yanan cismi terk ettiğinde geriye yalnızca kül kalmaktaydı. Ancak yanma olayı ayrıntılı olarak incelendiğinde filojistona benzer bir

[23] "Halk psikolojisi sağduyu psikolojisidir; davranışı inançların, isteklerin, algıların, beklentilerin, hedeflerin, duyumların ve birçok şeyin sonucu olarak açıkladığımız psikoloji bilimine karşılık gelir. Genellemeleri, zihinsel durumları başka zihinsel durumlarla, algılarla ve eylemlerle bağlantılandıran kuramdır. Bu tanıdık genellemeler söz konusu olan zihinsel durum ve süreçlerin nasıl karakterize edileceğine ilişkin tanımları sağlarlar ve bu genellemeler zihinsel yaşamın olgularının sınırlarını çizerek ifade mantığını açıklar. Halk psikolojisi 'sezgisel psikoloji'dir ve kendimizi nasıl kavramsallaştırdığımızı şekillendirir. Bileşenleri inanç ve istektir." (Churchland: 2019: 332).

madde bulunamadığı gibi, yanma olayının oksijen içerdiği ve de oksijenin cisimden salınmadığı aksine havadan alınarak yanan cisme bağlandığı ortaya çıkarılarak oksijen kuramı, filojiston kuramının yerini aldı ve filojiston denilen madde gerçekte hiçbir zaman var olmamış bir kavram olarak bilimden bütüncül olarak dışlandı (Revonsuo, 2016: 55).

Elemeci materyalistlerin, bilincin bilimden elenmesi gerektiği yönündeki tezlerinde bilim tarihinden öykündükleri bir diğer olay ise "ısı" kavramının tarihi ile ilgilidir. Isının yan yana gelmiş milyonlarca molekülün hareket enerjisi olduğu açıkça ortaya konulmadan önce süngerin suyu tutmasında olduğu gibi cisimlerde tutulan ince bir sıvı olduğuna inanılmakta ve *kalorik* adı verilen bu ince sıvı tözün doğası araştırılmaktaydı (Churchland, 2012: 70). Ortaya atılan ısı kuramının, kalorik kuramından daha başarılı ve açıklayıcı oluşu *kalorik* kavramını bilimsel terminolojiden elenen bir kelime hâline gelmiştir.

Filojiston ve ısı elemeci materyalistlerin gözlemlenemeyen şeylerin bilimsel terminolojiden elendiğine yönelik olarak kulladıkları örneklerdir. Bunun yanı sıra, Kopernik'ten önce insanların geceleri gökyüzünde Kutup Yıldızı etrafında dönen gök kubbeyi gözlemledikleri, bu gök kubbenin yapısını araştırdıkları ve görsel deneyim ile elde ettikleri bu imgeden de hiçbir zaman şüpheye düşmedikleri bilinmektedir. Gök kubbe ise bu bağlamda, elemeci materyalistlerin gözlemlenebilir şeylerin de bilim tarihinden elenebileceğine dair dayanak noktası olarak kullandıkları bir başka örnektir (Churchland, 2012: 71).

Elemeci materyalizme göre bilinç de filojiston, kalorik ve gök kubbe gibi bilimsel terminolojiden tasfiye edilmesi ve çöp kutusuna gönderilmesi gereken bir kavramdır. Zira bilincin dayanakları bilimsel bir kuramdan ziyade halk psikolojisinin birer parçasıdır. İnsanlar gündelik yaşamlarının rutininde bu kavramla bir şekilde aşina olurlar ve onu kullanmayı öğrenirler. Deniz kenarındaki bir bankta oturup içtiği sigarasını söndürdükten sonra çöpe atan bir insan gördüklerinde, bu insana bilinç ve dolayısıyla zihin atfederek davranışını yorumlarlar. Görünüşe göre bu insanın sigara kullanma alışkanlığı vardır, denize bakarak rahatlama istediği duyduğu bir anda onu rahatlatan başka bir nesne olarak sigarası aklına gelmiştir, manzaranın güzelliğiyle sigarasını keyifle içmiştir ve çevre kirliliğine karşı duyarlı bir insan olarak izmariti yere değil çöpe atmıştır.

İnsanlar, istek, inanç, tat, duyu gibi içsel deneyim hisleriyle başka insanların davranışlarını açıklamaya çalışırlar. Elemeci bir materyaliste göre, bir insana zihin atfederek açıklamaya çalıştığımız davranışların ardında yer alan halk psikolojisine ait bilinç kavramının asla

bilimsel bir dayanağı yoktur. Bu halk psikolojisi kavramının, zihinle ilgi doğru tasavvuru verdiğine dair hiçbir delil bulunmamaktadır. Nasıl filojiston ve kalorik kimya için, gök kubbe astronomi için gereksiz kavramlar haline dönüşmüşse bilinç de beynin doğasının açıklanmasında faydasız bir kavram haline dönüşecektir. Bu sebeple elemeci materyalistlere göre yapılması gereken şey, bilincin bilimden tamamen çıkarılması ve yerine nörobiyolojik ve nörofizyolojik kavramların konulmasıdır.

Elemeci materyalist yaklaşım, bilinç kavramının bilimsel terminolojiden elenmesi gerektiğini, bilinci sadece halk psikolojisine ait bir kavram olarak değerlendirmesinden dolayı değil, hiç kimsenin bilincin tam olarak neye delalet ettiğini söyleyemeyeceğini çünkü bilincin kapalı muğlak ve karmaşık bir kavram olduğunu iddia ederek savunmaktadır (Revonsuo, 2016: 55). Bilincin neliği üzerine görüş birliğine varılamamış olması, bilincin beyinde veya başka bir yerde bulunduğu yerin saptanamamış olması gibi sebepler, elemeci materyalistlerin bilincin bilimden tasfiye edilmesinin mutlak suretle gerekli olduğu yönündeki iddialarını delillendirirken kullandıkları diğer öncüllerdir.

Bilinci filojiston, kalori ya da gökkubbe örneklerindeki hipotetik kavramlar gibi bir durumu açıklamak için faydalı ve üretilmiş bir kavram olarak değerlendirmek, bize kesintisiz olarak sunulan kendi varlığımızın idrakinde kullandığımız öznel ve niteliksel gerçekliği reddetmek olacaktır. Bilinç, elemeci materyalistlerin bilim tarihinden çıkarsadıkları bahsi geçen örneklerde olduğu gibi farazi bir kavram değil, aksine bir gerçekliktir. Dolayısıyla bilincin reddiyesi filojistonun değil yanma olayında ortaya çıkan dumanın reddiyesi hatta yanmanın bizatihi kendisinin reddiyesiyle eşdeğerdir.

Bilinç, nesnel yöntemlerle gözlemlenebilen ya da ölçülebilen bir olgu değildir. İnsanın bilinçli olması beyindeki bir açma-kapama düğmesinin durumuna göre değişmekte de değildir. Bilinci sahip olduğumuz öznel ve niteliksel deneyimlerde bulabiliriz, bu buluşun yolu da yalnızca birinci şahıs bakış açısıdır. Birinci şahıs bakış açısı, doğa ve fizik bilimlerinde kabul edilen bir bilgi edinme yöntemi değildir. Bilincin geleneksel bilim anlayışla bağdaşmayan bu yapısı hakkında yeni açıklama modelleri sunmak yerine varlığını tümüyle reddetmek şüphesiz ki en kolay yöntemdir. Ancak elemeci materyalizm bilincin varlığını inkâr ederken yerine geçecek olan insanın öznel ve niteliksel tüm deneyimlerine karşılık gelecek nörofizyolojik kavramları tüm açıklığıyla ortaya koymadıkça bilime fayda değil bir araştırma alanının önünü kapatarak zarar verecektir.

Bilincin varlığını inkâr etmek, aynı zamanda deneyimi inkâr etmek olacaktır. Deneyimden bahsedilmeyen bir yerde, deneyimleyen özneden bahsedilmesi de söz konusu olamaz. Elemeci materyalizmin bilincin halk psikolojisine ait bir terim olarak bilimden tasfiye edilmesi gerektiği görüşü, aynı zamanda deneyim ve deneyimleyenin de bilimden tasfiye edilmesini gerektirmektedir. Deneyimin ve deneyimleyen olarak bilinçli öznenin varlığını inkâr eden bir bilinç yaklaşımı da hem mantıksal hem de deneyimsel çelişkilere yol açmaktadır. Yalnızca fiziksel bedene sahip bir özneden bahsedilecekse elemeci materyalizm o bedenin öznel ve niteliksel bilinçli içeriklere nasıl ve ne yolla sahip olduğuna dair bütüncül açıklamayı nörobilimsel kavramlar ile yapmalıdır; aksi takdirde, ortaya konulmuş olan insan tasavvuru basit bir makineden farksız olacaktır.

2.1.2.2. İndirgemeci Materyalizm

İndirgemeci materyalizm (*reductive materialism*), elemeci maddecilikte olduğu gibi bilinçli durumların inkârı üzerinden kurgulanan bir teori değildir. İndirgemeci materyalistler, bilincin ve bilinçli durumların varlığını yok saymamakta ancak zihinsel ve fiziksel olanın ayrı iki şey olarak düşünülmesi fikrine karşı çıkmaktadırlar. İndirgemeci materyalizm, bilinç ve beynin farklı iki şey olmadığını temellendirme amacıyla bilincin beyindeki diğer fiziksel süreçlerden farklı olmayan fiziksel bir şey ya da süreç olduğu yaklaşımını benimsemektedir. Bilincin varlığı değil mahiyeti konusunda bir yanılsama içerisinde olunduğunu ileri süren bu yaklaşımda, bilinçli durumların kendilerine indirgendikleri fiziksel şey veya süreçlere göre farklı bakış açıları ortaya çıkmaktadır. Davranışçılık, zihin-beden özdeşliği teorisi ve işlevselcilik indirgemeci materyalizmin farklı versiyonları olarak öne çıkan yaklaşımlardır.

2.1.2.2.1. Davranışçılık

Davranışçılık (*behaviorism*), yirminci yüzyılın başlarında içebakışçı[24] psikolojinin öznel karakterinden dolayı bilimsel olmadığı gerekçesiyle eleştirildiği bir dönemde, içebakış

[24] Psikolojinin, bilinci ve anlık deneyimi araştırmak için basitçe deneyimlere sahip olmak, onlara özenli bir şekilde dikkatini vermek ve sonrasında onları sözel olararak betimlemekten oluşan içebakış yöntemini kullandığı yaklaşım (Revonsuo, 2016: 101).

yöntemine tepki olarak ortaya çıkmış olan, psikolojinin doğası ve metodolojisi üzerine bir düşünce akımıdır (Kim, 2011: 61). Davranışçılığın temelleri pozitivizm ve deneycilik üzerine inşa edilmiştir. Hem pozitivizmin hem de deneyciliğin ortak iddiası olan bilimin doğrudan ve gözlemlenebilir şeyleri konu alması gerektiği fikri, aynı zamanda davranışçılığın da vazgeçilmez kuralını oluşturmaktadır. Dolayısıyla davranışçı yaklaşıma göre, bilinç ve öznel deneyim herkesçe gözlemlenen nesnel olgular olmamaları hasebiyle herhangi bir bilimin konusu olamazlar. Davranışçılık, bilinci ve zihni kategorik olarak bilimsel psikolojinin dışına iterek, psikolojiyi zihnin ve bilincin bilimi olarak değil davranışın bilimi olarak tanımlamaktadır (Revonsuo, 2016: 113).

Davranışçılık teorisi zihin durumlarının gözlemlenebilen davranışlarla ya da davranışsal dışavurumlarla eşdeğer olduğunu savunmaktadır. Buna bağlı olarak davranışçılığa göre, bir zihin durumu içerisinde bulunma ve bilinçli durumların varlığı, içinde bulunulan durumun davranışsal dışavurumu olarak kabul edilebilir (Rakova, 2007: 16). Davranış ise insan, hayvan gibi organizmaların ve hatta mekânik sistemlerin herkes tarafından gözlemlenebilecek şekilde yaptıkları şeylerdir. Davranışın, zihinsellik ile bir ilişkisi olduğu yadsınanamazdır. Davranışlar, düşünce, his, niyet ve amaçlarla koparılamaz bir bağa sahiptirler. Davranışlar, inanç, arzu, his, duygu, ilham ya da hedeflerin açık ifadeleridir (Kim, 2011: 61). Ancak, bilinçli durumların davranışa indirgenmesiyle ortaya çıkan tabloda zihin ve davranış arasındaki bu ilişkinin nasıl gerçekleştiği ve bu ilişkinin doğasının ne olduğu gibi elzem sorular cevapsız bırakılmaktadır. Oysaki davranış zihinselliğin bir işareti ya da ölçütü olarak indirgemeci bir tavırla ele alınacaksa bu işaretin nasıl gerçekleştiği de açığa kavuşturulmalıdır.

Davranışçılık bilinçle ilgili tüm durumları, başkaları tarafından da gözlemlenebilen davranışlarla izah edebileceğimizi ileri sürerken bilinçli bir kişinin dahi içinde bulunduğu durum ile ilgili olarak herhangi dikkat çekici bir davranışta bulunamama ihtimali ile çelişkiye düşmektedir. Bu durumun farkında olan bazı davranışçılar, açık ve gizli davranış arasında ayrım yapmaktadır. Gizli davranış tanımlamasıyla anlatılmak istenen düşünmenin dudakların belli belirsiz bir hareketinde, ses tellerinde ya da dilin belli belirsiz hareketlerinde de olsa açığa çıkacağını ifade etmektir. Ancak, bu girişim de bazı sorunların çözülmesi için yeterli gözükmemektedir. "Geçici felce sebep olabilen *kürar*[25] uyuşturucusu üzerine yapılan

[25] Güney Amerika yerlilerinin eskiden beri oklarının ucuna sürdükleri, çizgili kasları felç eden, güçlü bir bitkisel zehir.

yeni çalışmalar, zihinsel olaylar devam ederken gizli davranışların bulunmayabildiğini göstermektedir. Tam bir kas felci elde etmek için yeterli miktarda kürar alan hastalar, küraın uyuşturucu etkisi kaybolduktan sonra, yaşadıkları felç durumu boyunca hiçbir bilinç, düşünce, duyum, düşünme yetisi, hayal vs. kaybı olmadığını bildirmişlerdir." (Shaffer, 2005: 32). Bu, var olan tüm zihin durumlarının davranışa bağlı olarak açıklanabileceği iddiasını çürütmeye yönelik olarak kullanılabilecek önemli bir karşıt örnektir. Kimi davranışçı filozoflar, kürar uyuşturucusu örneğinde olduğu gibi davranış sergilenemeyen bilinçli durumların varlığının ortaya çıkardığı problemlerle başa çıkmak için "davranış eğilimi" kavramına sığınmışlar ve zihin durumlarını doğrudan davranış diliyle değil de davranışa olan yatkınlıkla tanımlamaya çalışmışlardır (Shaffer, 2005: 32). Ancak davranış yatkınlığı ile zihin durumlarını açıklama girişimi de davranışçılık için kurtarıcı olamamıştır zira bilinçli bir insanın sorulara cevap verme yatkınlığı olduğu düşünülür, ancak insan sorulara cevap vermeye yatkınlığı olmaksızın da bilinçli olabilir ya da meşgul olduğundan soruyu duymayabilir. Dolayısıyla sorulan soruya cevap vermemenin kesin sebebi cevap verme yatkınlığının yoksunluğu olmadığı gibi, cevap verme yatkınlığına sahip olmak da bilinçli olmak için zorunlu bir ölçüt değildir. Ayrıca bir organizmanın herhangi bir davranışta bulunma yatkınlığına sahip olması, davranışçılığın kabul ettiği kriterler eşliğinde değerlendirildiğinde gözlemlenebilir olmamasıyla davranışçılık için bir çıkış yolu oluşturmamaktadır.

Bu bağlamda davranışçılık, zihin ve zihin içeriklerinin varlığı ve doğasına dair yaklaşımlarıyla farklılaşan yöntemsel, ontolojik ve mantıksal davranışçılık olmak üzere üç kategori altında incelenebilir (Kim, 2011: 67).

Yöntemsel davranışçılara göre yalnızca davranış ve uyarıcı gibi dışsal ve gözlemlenebilir faktörler, bilimsel bir zihin anlayışının konusunu oluşturabilirler. Buna göre, bilimsel psikolojinin tek amacı davranışı açıklamaktır, eğer zihinde gerçekleşen başka durumlar varsa bunlar göz ardı edilmelidir. Yöntemsel davranışçılık, "Zihinde gözlemlenemez hiçbir şey yoktur." demediği ve gözlemlenemeyen durumları sadece yok saydığı için diğer davranışçılık türlerine kıyasla daha yumuşaktır. Bu yaklaşımda davranış, zihin durumlarının ontolojik statüsüne yönelik bir açıklama şekli olarak değil sadece bir yöntem olarak kullanılmaktadır.

Ontolojik davranışçılar, yöntemsel davranışçılardan farklı olarak öznel bilinçli durumların varlığını yalnızca yok saymazlar, aynı zamanda tümüyle reddederler. Bu gruba göre, sadece bir davranışa yansımış olarak gözlemlenebilen zihin durumları vardır (Kim, 2011: 78). Ontolojik davranışçılık, gözlemlenebilen davranışın dışında öznel bilinçli zihin içeriklerinin varlığını kabul etmemektedir. Bu bağlamda, en sevdiğin arkadaşının cenaze töreninde siyah giysiler giyip, siyah gözlük takıp, biraz gözyaşı dökmüş olmak ile arkadaşının kaybından duyulan üzüntü bir ve aynı şeydir. Bu üzüntünün bilinçli bir zihin içeriği olarak varlığı davranışsal karşılığı ile aynıdır, ötesinde bir zihin durumundan bahsetmek anlamsızdır.

Mantıksal davranışçılara göre, öznel zihin durumlarına karşılık gelen terimler davranış veya dışavuruma karşılık gelen terimlerle yer değiştirebilirler (Crumley, 2000: 13). Bu, zihin durumlarına tekabül eden bir terim içeren her cümlenin, davranışa tekabül eden terimlerden oluşan bir cümleye çevrilebileceği anlamına gelmektedir. Mantıksal davranışçılık, bilimde kullanılan her cümlenin, onun doğruluğunu kanıtlayacak deneysel şart ve gözlemlerle anlaşılması gerektiğini savunan mantıksal pozitivizme dayanır (Graham, 2010: 45). Mantıksal pozitivizmin etkisindeki mantıksal davranışçılık, bütün zihin durumlarının davranışsal olarak tercüme edilebileceği iddiasındadır.

Davranışçılığın ana fikri, zihin durumları olarak ifade edilen kavramların içsel ve görünmeyen durumlara değil yalnızca davranışlara tekabül ettiği yaklaşımı üzerine inşa edilmiştir. Davranışçılara göre, şu iki önerme hiçbir anlam kaybı olmadan birbirinin yerine kullanılabilir ifadeler içermektedir: "Ali acı çekiyor." ve "Ali'nin kan basıncı yükseldi, kızardı, kıvranıyor, bağırıyor ve ağlıyor." Bu yaklaşım, acının kan basıncı yükselmesi, kızarma, kıvranma, bağırma ve ağlama davranışlarıyla bir ve aynı şey olduğunu kabul etmektedir. Davranışçılığın bu durumda elemeci materyalizmden farkı zihin durumlarını inkâr etmemesinden kaynaklanmaktadır. Davranışçılık zihin durumlarının varlığını onları fiziksel bir fenomen olarak görerek kabul etmektedir zira zihin durumları yalnızca empirik olarak gözlemlenebilen davranışlarla görünür kılınmaktadır (Madden, 2013: 97).

Davranışçılığın öne sürdüğü tezler genel bir bakış açısı ile değerlendirildiğinde davranışta bulunan organizmanın ne beyni ne de zihninin davranışçılığın konu edindiği fenomenler arasında olmadığı açıkça görülmektedir. İnsan ya da hayvanın iç dünyası, öznel deneyimleri ve hatta beyin yapısı davranışçılar tarafından göz ardı edilen unsurlardır. Davranışçılık sadece nesnel fiziksel uyarıcı ile o fiziksel uyarıcının ardından meydana gelen dışavuruma

odaklı bir yaklaşım ortaya koyar. Bu yaklaşımın temel sonucu, bizim zihin durumlarımızın bizim davranışlarımız olduğudur. Bu iddia doğru kabul edilirse diğer insanların bizim zihin durumlarımızı en az bizim kadar ve hatta daha yakinen bilebileceğinin de koşulsuz kabul edilmesi gereklidir. Diğer insanlar bizim yüz ifadelerimizi, ses tonumuzdaki değişiklikleri, jest ve mimiklerimizi ve diğer tüm dışavurumlarımızı bizden daha iyi gözlemleyebilirler. Dolayısıyla başkaları öznenin içinde bulunduğu zihinsel ve psikolojik durumu bizzat öznenin kendisinden daha iyi gözlemleyebilir. Çünkü davranışçılık bu hususa bağlı olarak, bizim kendi zihin durumlarımıza özel bir erişimimiz olmadığını, birinci şahıs bakış açısı gibi bir yaklaşımdan bahsedilemeyeceğini, tüm zihin durumlarının tıpkı diğer tüm fiziksel durumlar gibi üçüncü şahıs bakış açısıyla var olduğunu ima etmektedir (Madden, 2013: 102).

Tüm zihin durumlarını gözlemlenebilen davranışa indirgeyen davranışçılar, bilince ontolojik bir değer atfetmeden "bilinçli" olarak addedilen zihin durumlarını da davranışla özdeşleştirmektedirler. Bilincin kendisine öznel bir erişim imkânı bırakmayan bu yaklaşımda, bilinçli zihin durumları nesnel olarak erişilebilen davranışlar vasıtasıyla tanımlanmaktadır. Dolayısıyla davranışçılık öznel bilinçli durumları ve öznel deneyimi inkâr eden, deneyimin varlığını sadece üçüncü-şahıs bakış açısından kabul eden bir görünümdedir. "Her deneyimin bir deneyimleyeni vardır." önermesi davranışçılığın öne sürdüğü tezler etrafında değerlendirildiğinde bu iddia şu önermeye evrilmektedir: "Her davranışın bir davranışta bulunanı vardır." Bu önermede deneyim, davranış ile fiziksel olarak ifade edilirken deneyimleyen de yine fiziksel bir davranış sahibine tekabül etmektedir. Dolayısıyla açıkça görülmektedir ki, davranışçılığı bilincin öznel ve niteliksel boyutunu göz önüne alarak ve bu öznel, niteliksel deneyimlerin sahibi olan öznenin varlığının farkında olarak okumak mümkün değildir. Zira davranışçılığın üzerine kurulmuş olduğu temel, öznellik ve nitelikselliğe müsaade etmemektedir.

2.1.2.2.2. Özdeşlik Teorisi

Zihinsel olanın fiziksel olandan ontolojik ve epistemolojik anlamda farklı bir varoluşa sahip olmadığı fikri üzerine inşa edilmiş olan özdeşlik teorisinin (*identity theory*) temel tezi zihin ve beyin durumlarının tek bir varlığa karşılık gelmesidir. Burada özdeşlikle kastedilen, "zihinsel terimlerin, anlam yönünden fizikalist terimlerle eş anlamlı olmaları bağlamında 'özdeş' olmaları değil, bilakis zihinsel terimlerle ifade edilen fiili olayların fizikalist terimlerle ifade edilenlerle bir ve aynı olması anlamında özdeş" olmalarıdır (Shaffer, 2005:

73). Zihni, beyne bütünüyle indirgeyen bu yaklaşımı sistematize eden ilk filozoflardan olan J. J. Smart özdeşlik tezini şöyle ifade etmektedir:

> ‘Duygulanım bir beyin sürecidir.’ ya da ‘Işık bir elektrik deşarjıdır.’ dediğimde burada kullandığım ‘-dır.’ mutlak özdeşlik manasındadır. Burada duygulanımın zamansal ve mekânsal bağlamda beyin süreçleriyle birlikte devam eden bir şey olduğunu kastetmiyorum. Duygulanım ve beyin sürecinin özdeşliğini kesin ve net olarak garantiliyorum (Smart, 2004: 119).

Zihinsel ve fiziksel olanın özdeşliği iddiasının semantik değil ontolojik bir iddia olması, bu teoriyi mutlak olarak monist ve materyalist kategoriye dâhil etmektedir.

Özdeşlik teorisi taraftarları, elemeci materyalistlerin aksine zihin durumlarının varlığını yok saymamaktadır. Özdeşlik kuramına göre, bilinçli zihin durumların varlığını kabul etmek hatalı değildir. Bunun yerine, bilinçli zihin durumlarının beyinden ve nörolojik süreçlerden farklı bir varlığa işaret ettiklerini kabul etmek büyük bir kategorik hataya düşmektir. Özdeşlik teorisine göre, zihin ve beyin ontolojik olarak bir ve aynı şeydir. Bu bağlamda, bilinç de beyindeki fiziksel süreçlerden hiçbir farkı olmayan yine fiziksel bir sürece işaret etmektedir. Özdeşlik teorisi bilinçli zihin durumlarını yok sayan değil beyindeki fiziksel süreçlere indirgeyen bir yaklaşımla, bilincin varlığı değil yapısı üzerine bir sorgulama gerçekleştirmektedir.

Özdeşlik teorisinin, zihinsel olanın fiziksel olana tüm yönleriyle indirgenebileceği iddiası da tıpkı elemeci materyalizmin iddialarında olduğu gibi bilim tarihine atıflarda bulunmaktadır. Isı, ışık ve su terimleri, gündelik dilden bilimsel dile geçen kavramları örneklendirmek amacıyla indirgemeci materyalistler tarafından sıkça kullanılmaktadır. Önceleri günlük dile ait terimler olan ışık “belirli dalga boylarındaki elektromanyetik radyasyon”, ısı “moleküllerin ortalama kinetik enerjisi” ve su da “H_2O” ile yer değiştirerek bir fizik kuramı tarafından kapsamlı bir şekilde açıklanan fiziksel varlıklar hâline dönüşmüştür (Revonsuo, 2016: 61). Bilim tarihinin, günlük dile ait fenomenleri fiziksel terimlere indirgeyerek açıkladığı örneklerin sayıca fazlalığı, indirgemeci materyalistleri zihin ve bilincin de bir gün tamamen nörofizyolojik terimlerle açıklanabilen fenomenlere evrilecekleri fikrine inanma yönünde cesaretlendirmiştir. Özdeşlik teorisi, tüm bilinçli zihin durumlarının nörobilimin sağladığı kuramsal çerçeve içerisinde açıklaması yapılabilecek fenomenler olduğunu umut ederken bilinçli zihin durumlarının her birinin beyne ait fiziksel ve nörolojik durumlarla özdeş olduğunu kabul etmektedir.

Özdeşlik kuramı, zihin durumları yalnızca beyin durumlarıdır iddiasında olan, anlaşılması basit bir açıklama modeli sunduğundan dolayı materyalist teoriler arasında öne çıkan bir yaklaşım olmayı başarmıştır. Bu kuramın diğer zihin teorilerine kıyasla çok daha az varsayıma dayanması, dayandığı aksiyomların test edilebilir olması, düalizmin temellerinin dayandığı zihin-beden karşıtlığını indirgemeci bir tavırla bertaraf etmiş olması, onu tercih edilmesi kolay bir yaklaşım hâline getirmiştir (Maslin, 2007: 73). Aynı zamanda zihin-beden problemi olarak adlandırılan, fiziksel olmayan bir şeyin fiziksel olan bir şeyle nasıl etkileşimde bulunabildiği sorunsalına da fiziksel nedensellik yasası ile doyurucu bir cevap veriyor görünümünde olduğundan, özdeşlik kuramı zihin felsefecileri arasında sahip olduğu taraftarların çokluğu bakımından güçlü bir konum elde etmiştir. Beyin ile zihin arasında göze çarpan açık bir korelasyonun varlığı da özdeşlik kuramını güçlü kılan etmenler arasında yer almıştır. Zira bazı bölgesel beyin hasarlarının, zihinsel yetilerde belli bazı hasarlara yol açması, beyin ve zihin arasında var olan güçlü bağlantıyı ortaya koymaktadır.

Sahip olduğu bu avantajlı konuma rağmen, özdeşlik yasasının başa çıkılması güç handikapları olduğu da bir gerçektir. Bu kuramın içerisinde barındırdığı önemli bir sorunsal, *Leibniz Yasası* olarak bilinen, farklı iki şeyin ancak ve ancak tamamen aynı özellikleri taşımasıyla özdeş olabilecekleri ölçütü ile çelişkili olmasıdır (Crumley, 2000: 10). Leibniz Yasası'nı temel aldığımızda zihin ve beynin özdeş olabilmesi için her ikisinin sahip olduğu tüm niteliklerin de aynı olması gerektiği sonucuna ulaşırız. Buradan hareketle, bu kural bizim duyumsamamız ile nörolojik durumumuzun özdeş olmasını mecbur kılar. Fakat duyumsamamız ile nörolojik durumumuzun bir ve aynı şey olduğunu savunmak yerinde olmayacaktır. Özdeşlik kuramının cevabı Clark Kent ve Superman, su ve H_2O vs. örneklerinde olduğu gibi, zihin durumları ve beyin örneğinde de elimizde olanın aslında iki farklı şey değil, tek bir şey olduğudur. Özdeşlik kuramcılarının iddiası şudur ki; günlük dil ve halk psikolojisinde kullanılan terimler, nörolojik durumlar ile zihin durumlarınının birbirinden farklı şeyler oldukları hususunda insanları yanlış bir inanca yönlendirmektedir. Ancak zihinsel ve fiziksel olarak iki farklı töz olduğu iddiasından vazgeçilse ve zihinsel olan her durumun fiziksel olarak tanımlanabileceği kabul edilse dahi fiziksel olarak açıklanabilen zihin durumlarının da zihinsel özellikleri olduğu gerçeğini yok saymak güç görünmektedir. Beyin durumlarını aynı zamanda zihin durumları yapan özellik, beyin durumlarını tanımlayan fiziksel özelliklerden kavramsal olarak farklı bir özellik olmalıdır. Bunun yanı sıra, su = H_2O özdeşliği suyun yalnızca H_2O olduğu anlamına gelir; H_2O olmadan su olmaz ve su için H_2O'dan fazlasına ihtiyaç yoktur. Etrafta suyu içecek, hissedecek ya da görecek

kimse olmasa bile o sudur, suyun varlığı için öznel deneyime ihtiyaç yoktur (Nagel, 2015: 49). Su algılanabilir olan nitelikleriyle tanımlanıyor olsa bile bizim algısal deneyimlerimiz su olmanın bir parçası değildir. Zihin durumlarını su örneğinde olduğu gibi beyin durumlarının bizzat kendisi olarak tanımlarsak da fiziksel özelliklerini bildiğimizde biberin tadını, acı hissini ya da özdeş olduğu her ne ise onu da bilmemiz gerekir. Ancak durum böyle değildir. Deneyim olmadan bir zihin durumundan bahsetmek mümkün değildir. Dolayısıyla beyin durumları ve zihin durumlarının özdeşliği, su ve H_2O özdeşliği gibi teorik bir özdeşlikle benzer olarak açıklanamaz.

Zihinsel ve fiziksel karşıtlığı ile ilgili problemlerin yanı sıra, özdeşlik teorisinin eleştiriye en çok maruz kalan yönlerinden bir diğeri de bilincin öznel ve niteliksel yapısını göz ardı etmesidir. Özdeşlik teorisinin bilincin öznel yanını açıklamada eksik kaldığına dair farklı formlarda öne sürülen argümanların temel tezi şu şekilde özetlenebilir:

> Mükemmel bir nörobiyoloji bilimine sahip olsak dahi nöron ateşlemelerinin nörobiyolojik örüntüsü ve acının hissedilmesi gibi birbirinden ayrı iki özellik var olmaya devam edecektir. Hissetmeye nöron ateşlemeleri neden olur, fakat nöron ateşlemeleri ile hissetme aynı şey değildir (Searle, 2018: 34).

Bu teze bağlı kalarak ortaya atılan en önemli karşı çıkışlardan birisi "Bir deneyim hakkında sadece o deneyimin yaşanması durumunda elde edilebilecek bilgiler mevcuttur." iddiasına dayalı olarak kurulan *bilgi argümanı*dır[26] (Gulick, 2008: 664). Bilgi argümanını zihin felsefesinin gündemine taşıyan "What Mary Didn't Know?" (Jackson, 1986: 291-295) isimli makalenin iskeletini siyah-beyaz bir odada kapalı olarak yaşayan, hiçbir şekilde diğer renkleri deneyimlememiş ancak renklerin tüm fiziksel bilgisine eksiksiz olarak hâkim olan bir renk bilimcisi olan Mary üzerine kurgulanmış bir düşünce deneyi oluşturmaktadır. Bu düşünce deneyinin işaret ettiği ve sorguladığı husus şu şekilde bir soruyla yorumlanabilir: Kırmızı rengini görsel anlamda hiçbir zaman deneyimlememiş ancak kırmızının tüm fiziksel bilgisine sahip olan, başka bir deyişle kırmızı rengi meydana getiren tüm fiziksel nedenleri bilen Mary kapalı olarak yaşadığı odadan çıkarak kırmızı rengini deneyimlediğinde yeni bir şey öğrenmiş olur mu ya da yeni bir bilgi elde etmiş olur mu?

Bir özdeşlik kuramcısının bu soruya cevabı elbette ki "hayır" olacaktır zira tüm zihin durumları fiziksel durumlarla bir ve aynı şeydir. Bu durumda, kırmızının fiziksel bilgisi zihindeki kırmızı imgesi ile eşdeğer kabul edilecektir. Özdeşlik kuramcısına göre Mary

[26] Knowledge argument.

odadan çıkıp kırmızıyı görsel olarak deneyimlediğinde yalnızca sahip olduğu fiziksel bilgileri yeni bir yoldan keşfetmiş olur; yeni bir bilgiye ulaşmış olması söz konusu değildir. Ancak Mary'nin odadan çıkıp ilk kez görsel anlamda kırmızıyı deneyimlediğinde şaşırmayacağını ve gördüğü şeyin onun için "yeni" bir deneyim olmayacağını söylemek hem mantıksal hem de sezgisel perspektiften tutarlı görünmemektedir. Özdeşlik kuramının ihmal ettiği husus Mary'nin kırmızı ile ilk defa karşılaştığında elde etmiş olduğu öznel deneyim ve bu öznel deneyimin oluşturduğu öznel fenomenal zihin durumudur. Mary'nin kırmızı deneyiminin öznel ve niteliksel boyutunu tamamen göz ardı eden özdeşlik kuramı, deneyimi fizikselleştirerek deneyimleyeni de zorunlu olarak fizikselleştirmektedir. Fenomenal özne olarak deneyimleyenin doğasına dair bir açıklama modeli önermemektedir.

Bilinçli durumları nörofizyolojik durumlarla bir ve aynı şey olarak ele alan özdeşlik kuramının dışarıda bıraktığı unsurlardan birisi de bu bilinçli zihin durumlarına sahip olmanın ne demek olduğu, nasıl bir şey olduğu, bu durumların yarattığı öznel his ve bilinçli bir zihne sahip bir öznenin zihin yaşantısının ne olduğudur. Bilinçli zihin durumlarının öznel ve niteliksel yanlarını hesaba katmayan öznel zihinsel gerçekliği inkâr etmese de ihmal eden özdeşlik kuramının karşısındaki en güçlü eleştirilerden birisi de şüphesiz Thomas Nagel'in "Bir Yarasa Olmak Nasıl Bir Şeydir?" (Nagel, 1974: 435-450) makalesidir. Nagel'in bu makaledeki amacı zihin ve beyni bir ve aynı şey olarak kabul eden indirgemeci materyalistleri deneyimin öznel karakterini göz ardı ettikleri için eleştirmektir. Nagel'a göre özdeşlik teorisi zihni beyne indirgerken indirgediği şeyin öznel yapısını incelemeyi unutmaktadır. Nagel'ın ifadeleriyle "Her öznel fenomen tek bir bakış açısıyla bağlantılıdır ve objektif, fiziksel teorilerin bu bakış açısını ihmal etmesi kabul edilemezdir." (Nagel, 1974: 437). Bu tek bakış açısının ne anlama geldiğini açıklamak için "Bir yarasa olmak nasıl bir şeydir?" sorusunun cevabını aramaya girişen Nagel, yarasanın memeli, görmediği için sonar kullanan bir hayvan olduğunu herkesin bildiğini kabul eder. Ona göre yarasa olmanın nasıl bir şey olduğu üzerine düşünülürse dışsal seslere göre hareket eden, uçabilen, çok az görme yetisi olan, böcek yiyen bir hayvan olduğumuzu düşünürüz, fakat aslında olması gereken yarasa için yarasa olmanın nasıl bir şey olduğu üzerine düşünebilmemizdir (Nagel, 1974: 439). Fakat bu bizim için imkânsızdır çünkü biz kendi zihnimizin kaynakları içinde sınırlıyız ve bu kaynaklar "yarasa olmaklık" durumunu kavrayabilmemiz için yetersizdir.

Nagel'a göre "Eğer bu evrende başka bir yerde bilinçli bir hayat varsa bunun bir kısmı bizim için kullanılabilir olan deneyimsel terimlerle bile tanımlanamaz." (Nagel, 1974: 440). Bir

kişi, diğer bir kişinin deneyimleri hakkında objektif bir yaklaşıma sahip olabilir ancak yine de üçüncül bakış açısıyla yaklaştığı için deneyimlerin öznel doğasına nüfuz edemez. Nagel şöyle iddia eder: "Eğer deneyimin öznel karakteri tek bir bakış açısı tarafından kavranabiliyorsa daha fazla bir nesnelliğe geçiş bizi fenomenin gerçek doğasına ulaştırmaz." (Nagel, 1974: 445). Bu da zaten *başka zihinler problemi*nin[27] çıkış noktasıdır; başka zihinler hakkında kesin bir bilgiye ulaşmamız deneyimlerin öznel yapısı itibarıyla mümkün gözükmemektedir. Nagel'a göre problem, deneyimleri oluşturan duyumların nesnel bir yapıya sahip olup olamayacakları üzerinde durulmamasıdır ve materyalizmin de düşünmesi gereken deneyimlerimizin bize nasıl gözüktüğü değil de gerçekte nasıl olduğu olmalıdır (Nagel, 1974: 448). Nagel'in dikkat çekmeye çalıştığı husus nesnel-öznel ayrımı tam olarak yapılmadan ne zihin-beden probleminin ne de başka zihinler probleminin üstesinden gelinemeyeceğidir.

Özdeşlik teorisi, deneyimi fiziksel süreçlere indirgeyerek, deneyimin öznel yapısını tamamen göz ardı etmektedir. Deneyimin öznelliğini göz ardı eden bir teorinin de deneyimleyenin doğasını gündemine alması beklenemez. Deneyimleyen olmadan deneyimden bahsedemeyiz, dolayısıyla özdeşlik teorisinde fiziksel bir süreç olarak tanımlanan deneyimin öznesi de fiziksel bir özne olacaktır. Deneyimleyen gerçek bir varlıktır, dolayısıyla özdeşlik teorisinin fiziksel öznesi de gerçek bir varlıktır. Eğer fiziksel öznenin doğası kavranırsa deneyimin de doğası kavranacaktır. Ancak, burada bahsedilen deneyim asla bilinçli deneyimlerin sahip olduğunu aksiyom olarak kabul ettiğimiz öznel ve niteliksel doğaya sahip olmayacaktır. Fiziksel düzeyde, yalnızca beyin süreçleriyle açıklanan bir özne tasavvuru insanı basit bir bilgisayardan ayıracak hiçbir ölçüt öne sürmemektedir. Ayrıca bu yaklaşım, beyni özneye eviren şeyin ne olduğunu tamamen kuramın dışında tutarak kapatılması çok zor bir gedik bırakmaktadır.

[27]Other mind problem. "Bir başkasının bir zihne sahip olduğunu ya da bilinçli olduğunu nereden biliyoruz?" sorusu başka zihinler probleminin temel sorunsalıdır. Bu problem zihinsel durumların ikinci veya üçüncü kişilere atfedilmesini konu alır ve şu soru üzerinden yola çıkar: Biz nasıl olur da üçüncü bir kişiye ait -içsel bir durum olan- zihinsel ya da psikolojik durumları bilebiliriz? Churchland bu problemi şöyle özetlemektedir: "Özgül psikolojik durumların- düşüncelerin, algıların, arzuların veya acıların- başkalarına her göndermenin açık bedensel davranışları temel alması gerekir (…) Ancak bu türden 'içsel durumlar' ve bunların 'dışsal ölçütleri' arasında bulunan genel bağıntıların bir değerlendirmesi, kişinin bu bağıntıları başkalarında gözlemlemesinden türetilemez çünkü varsayımsal bağıntıların 'dışsal' öğeleri herkesin incelemesine açıktır. Bu türden bağıntıların bilgisi dolayısıyla yalnızca kişinin kendi durumundan türetilebilir çünkü denir ki, içsel öğeler incelemeye açıktır." (Churchland, 2013: 124).

2.1.2.2.3. İşlevselcilik

İşlevselcilik (*functionalism*), zihin durumlarının, başka zihin durumları, duyusal girdiler ve davranışsal çıktılar ile olan işlevsel ya da nedensel ilişkilerinin bütününden meydana geldiğini savunan bir zihin yaklaşımıdır (Rakova, 2007: 69). İşlevselcilere göre, zihin durumlarının temel fonksiyonu, onların oynadıkları nedensel rollerdir. İşlevselci yaklaşımın dayandığı ana tez şöyle ifade edilebilir:

> Her zihinsel durumun özsel veya tanımlayıcı özelliği, onun (1) beden üzerindeki çevresel etkilerle, (2) diğer zihinsel durum tipleriyle ve (3) bedensel davranışlarla bir dizi nedensel ilişkiye girmesidir (Churchland, 2012: 57).

Bu yaklaşıma göre, eğer özel bir zihin durumunun temel özelliği bilinmek isteniyorsa bu zihin durumunun çevresel girdilerle, diğer zihin durumlarıyla ve davranışla ilişkisi dikkate alınmalıdır. İşlevselciliğe göre, bir zihin durumu içinde bulunmak ile işlevsel bir durum içinde bulunmak aynı şeydir. Bir zihin durumu, bir neden olarak duyusal bir girdiye ve bir sonuç olarak davranışsal bir çıktıya ihtiyaç duyar (Priest, 2018: 197).

İşlevselcilikte bilinç ne fiziksel ne de fiziksel olmayan bir tözle tanımlanmaktadır. Bu yaklaşım, bilincin ontolojik düzeyine bir atıfta bulunmadan bilimsel bir çerçeve çizmeye çalışsa da birçok işlevselci aynı zamanda materyalisttir (Priest, 2018: 197).

İşlevselciliğe göre bir zihin durumunun bilinçli olup olmadığı da sistem içerisinde oynadığı işlevsel role bağlıdır. Zihnin mahiyetini oluşturan girdi ve çıktılar arasındaki nedensel ilişkidir. Bu yaklaşımda zihin durumu, bilgi işleme sisteminin bir fonksiyonudur ve bu fonksiyon sisteme gelen girdiler ve sistemin ürettiği çıktılar arasındaki ilişkilerle tanımlanmaktadır (Revonsuo, 2016: 84). Zihin duyusal bilgiyi girdi olarak almakta ve bu duyusal bilgi pek çok farklı aşama ve düzeyde işlemekte, sonrasında ise çıktı olarak davranış üretilmektedir. Bu girdi-çıktı düzeyinde bilgi işleme süreci olarak tanımlanan zihin tasvirinin önemli bir sonucu, beyin ile hiçbir ilişkilendirme kurmadan bile işlevselci bir zihin tasvirine müsaade etmesidir (Revonsuo, 2016: 85). Bu bağlamda, özdeşlik teorisinin beyin-zihin ayniliğine karşı, işlevselcilik biyolojik olmayan zihin tasavvurlarına da imkân tanımaktadır. Zira işlevsilciliğin zihin projeksiyonu bir bilgisayarın çalışma şeklinden farklı değildir. Bir bilgisayar programı da benzer şekilde girilen kodu işleyerek bir çıktı üretmektedir. Yapay zihnin inşasına ilkece büyük bir zemin hazırlaması sebebiyle işlevselci yaklaşım, bilişsel bilim, zihin felsefesi ve yapay zeka çalışmalarından büyük destek görmüştür.

Zihin durumlarını, bilgisayar programlarına neredeyse özdeş bir konuma taşıyan işlevselciliğin karşısına çıkan en büyük problem şüphesiz ki bilinçtir. Zira bu teori içerisinde zihin durumlarını bilinçli ya da bilinçsiz kılanın ne olduğuna dair hiçbir açıklama yapılmadığı gibi, bilincin öznel ve niteliksel özellikleri de teorinin ele aldığı unsurlar arasına dâhil edilmemektedir. Bilinçli zihin durumları bu teoride yalnızca işlevsel durumlar olarak ele alınmaktadır. Sabah içilen ilk kahvenin tadı, fırında pişen kurabiye kokusu, yalnızlık duygusunun iliklerde hissedildiği bir anda gelen telefondaki eski dostunuzun sesi bu yaklaşımda işlevsel durum olmanın ötesinde başka hiçbir şeye tekabül etmemektedir. Bu durumların insanda meydana getirdiği öznel hissi, yaşattığı öznel deneyimi girdi-çıktı ilişkisi ile çalışan bilgi-işlem mekânizmasından üretmek sezgisel olarak mümkün görünmemekle birlikte, bu öznel durumların ortaya çıkmak için girdi-çıktı ilişkisinden daha fazlasına ihtiyaç duydukları da aşikârdır.

İşlevselciliğin zihin felsefecilerinden en çok eleştiri aldığı hususlar, bu yaklaşımın bilinçli zihin durumlarını göz ardı etmesi ile doğrudan ilintilidir. İşlevselciliği gösterilen sorunu cevaplandırmada çaresiz bırakan önemli antitezlerden birisi *kayıp deneyimsel nitelikler* (*kayıp nitelceler*)[28] argümanıdır. Bu argümanın temel iddiası işlevsel olarak duyusal varlıkların bir kopyasının üretilmesinin mümkün olduğu ancak bu kopyaların asla deneyimsel niteliklere sahip olamayacaklarıdır. Ned Block bu argümanı *Çin Milleti* olarak adlandırılan bir düşünce deneyi üzerine inşa etmiştir. Bu düşünce deneyi her Çinlinin bir nöronu temsil ettiği ve bu nöronların birbirlerine radyo sinyalleriyle bağlandığı, başka bir deyişle işlevselciliğin varlığını savunduğu zihin işleyişinin yapay bir ortamda oluşturulduğu hayalî bir durumu tasvir etmektedir (Block, 1978: 270).

Block'un bu düşünce deneyiyle hedeflediği, bahsedilen durumun oluşturulması ile ortaya çıkacak olan Çin Milleti'nin bütünsel olarak oluşturduğu zihinsel yapının, öznel deneyim ve niteliksel hislere sahip olamayacağını göstermektir. Bu deneyin iddiasına göre, *Çin Milleti* olarak tasvir edilen bu sistem işlevsel olarak insan zihni ile eşdeğer olabilecekse de hiçbir zaman bütün bir *Çin Zihni* olamayacaktır (Maslin, 2007: 140). Bu iddia da deneyimsel niteliğin doğasının işlevsel bir biçimde nedensel ilişkiler ile açıklanamayacağını ima etmektedir.

[28] *Absent qualia*

İnsan zihninin bilgi-işleme sistemi bütün olarak başka bir yapıya kopyalansa, bu yapı girdilere doğru çıktılar verse, konuşsa, gülse, yemek yese, şarkı söylese, müzik dinlese, ağlasa bile yaşadığı deneyimin öznel niteliğine sahip olduğunu iddia etmek oldukça güçtür. Deneyimsel nitelikler bilinçli zihin durumlarının sahip olduğu öznel ve niteliksel özelliklerdir. Bu öznel ve niteliksel deneyimler için işlevselciliğin sunduğu bir açıklama şekli yoktur. Dolayısıyla bu deneyimlere sahip olan deneyimleyen, başka bir deyişle fenomenal özne de bu yaklaşımın çözme eğiliminde olduğu problemlerin arasında yer almamaktadır. Özneyi ve doğasını göz ardı ederek, sadece işlevsel durumlar ile girdi-çıktı ilişkilerine odaklanan işlevselcilik teorisi bu anlamda savunulacak bir bilinç yaklaşımı ortaya koyamamaktadır. İşlevselciliğin mantıksal sonucu, Chalmers'ın ifade ettiği gibi, bizim sahip olduğumuz bilgi-işlem mekânizmasına sahip, bizim gibi davranan hatta zekice tepkiler veren ancak *bilinçsiz zombiler ya da fenomenal zombiler* olacaktır (Chalmers, 1996: 95).

2.2. Düalizm

Düalizm, evreni meydana getiren iki farklı varlık ya da töz vardır kabulü üzerine temellendirilmiş bir yaklaşımdır. Düalist yaklaşıma göre, evrende yer alan tözlerden birisi fizikseldir. Bu töz, madde, enerji, temel parçacık ya da kuvvetleri, kuvvet alanlarını ve fiziksel bilimlerde evrenin yapıtaşları olarak kabul edilen diğer her şeyi oluşturmaktadır (Revonsuo, 2016: 36). Diğer töz ise fiziksel olmayana tekabül etmektedir, fiziksel olmayanın özü ise zihinseldir. Düalist kuramlar kendi aralarında farklı bakış açıları ile ayrışıyor olsalar da tüm düalist zihin kuramlarının müşterek iddiası: "Bilinçli zekanın özsel doğasının fiziksel olmayan bir şeye; fizik, nörofizyoloji ve bilgisayar bilimi gibi bilimlerin kavrayış alanının ebediyen ötesinde kalacak bir şeye bağlı olduğu" fikridir (Churchland, 2012: 12).

2.2.1. Töz Düalizmi

Düalizmin en katı formu olarak nitelendirilebilecek töz düalizminin (*substance dualism*) temelleri ruh ve bedeni ayrı iki töz olarak kabul eden Platon'a dayanmaktadır. Platon, *Phaidon* diyaloğunda ruh ve bedenin ayrılığını şöyle ortaya koymaktadır:

> Ruh en çok tanrısal olana, ölümlü olmayana, düşünülebilene, yalın olana, dağılmayana, her zaman aynı kalana benzer; ten de en çok insansı olana, ölümlü olana, düşünülemeyene, çok şekilli olana, dağılana, asla kendisinin aynı kalmayanına benzer (Platon, 1997: 53).

Ruh ve bedenin karşıtlığını Platon'dan sonra sistematize ederek ele alan en önemli isim ise Descartes'dır. Ruhun bedenden apayrı bir varoluşa sahip olduğu fikrini Descartes şu cümleleriyle apaçık bir biçimde ifade etmektedir:

> Sonra, ne olduğumu dikkatle inceleyerek, hiçbir bedenim olmadığını, bulunabileceğim hiçbir dünya hiçbir yer olmadığını varsayabileceğimi, tersine başka şeylerin doğruluğundan kuşkulanmayı düşünüyor oluşumdan var olduğum sonucunun apaçık ve kesin bir biçimde ortaya çıktığını, oysa düşünmeyi bıraksam tasarladığım tüm başka şeyler doğru olsalar bile var olduğuma inanmam için elimde hiçbir neden bulunmadığını görerek tüm özü ya da doğası düşünmekten başka bir şey olmayan, herhangi maddi bir şeye bağımlı olmayan bir töz olduğumu anladım. Öyle ki bu ben yani kendisiyle neysem o olduğum ruh, bedenden tümüyle ayrıdır hatta bedenden daha kolay tanınır ve beden olmadığında bile o kendisi olmaktan çıkmaz (Descartes, 2015: 44).

Descartes felsefesinde özne bölünemez ve fiziksel olmayan ruhla bir ve aynı şeydir. Ruh, uzunluk, genişlik gibi ölçülere sahip değildir ve mekânsal bir varlığı yoktur; beden ise aksine fiziksel ve bölünebilir olmasıyla mekânda yer kaplar. Descartes her tözün bir özniteliği olduğunu vurgularken ruhun özünü düşünme, maddenin özünü ise uzam olarak belirlemektedir (Descartes, 2015: 98). Düşünmek, ruhun sadece bir özelliği değil, aynı zamanda özüdür. Descartes'ın düşünmeden anladığı şey yalnızca muhakeme ve akıl yürütme değil, aynı zamanda duygusal ve duyusal durumlardır. Descartes şüphe etme, anlama, inkâr etme, istekli olma, istekli olmama, hayal etme ve her türlü duyusal algıyı düşünmenin örnekleri olarak görür (Maslin, 2001: 43). Descartes'a göre düşünme varoluşun ispatı iken düşünmenin gerçekleşmesi için de zorunlu koşul bilinçtir. O hâlde, düşünce kavramından anlamamız gereken içinde bulunulan bilinçli zihin durumlarıdır (Günday, 2003: 34).

Töz düalizminde, algılar, düşünceler, duygulanım ve duyumlar beyne ait değil, zihne aittir. Zihin, bedenden tamamen bağımsızdır ve nedensel olarak beyinle etkileşimde bulunsa bile beyin olmadığında dahi var olmaya devam eder. Töz olarak hem beden hem de zihin, kendi içinde, kendi yasaları ve özelliklerine sahiptir; dolayısıyla beyinle ilgili araştırmalar zihnin doğası ve dinamikleriyle ilgili bilgi vermeyeceği gibi, zihinle ilgili araştırmalar da beynin yapısı ve çalışma şekli hakkında herhangi bir şey söylemeyecektir (Churchland, 2019: 352).

Töz düalizminin karşısındaki en büyük engel şüphesiz nörofizyoloji alanında gerçekleşen gelişmelerdir. Beyin ve zihni birbirinden bağımsız varoluşlara sahip iki töz olarak kabul eden bu yaklaşımın, beyin hasarlarıyla birlikte meydana gelen bilinç ve benlik hasarlarını hatta bilinç kaybını açıklaması zor görünmektedir. Zira töz düalizminin dayandığı temel prensip, beynin tamamen yok olduğu durumlarda bile zihnin kusursuz bir şekilde varlığını devam ettirmesini gerektirmektedir. Bunun yanı sıra, töz düalizminin diğer bir zayıf noktası bilinçsiz zihin durumlarıyla ilgili bir açıklama getirmemesidir. Bilişsel bilimlerin, açıkça ortaya koyduğu zihinde gerçekleşen bilinçsiz işleme süreçlerinin töz düalizminde herhangi bir karşılığı yoktur. Kör bakış, çoğul kişilik vakaları ve ayrık beyin hakkında da bir görüş öneremeyen töz düalizmi, nörobilimsel gelişmeler karşısında ezilmiş görünmektedir.

Descartes'ın birbirinden bağımsız varoluşlar yüklediği fiziksel beden ile fiziksel olmayan ruhun birbirleriyle nasıl ilişki kurdukları sorusu da kartezyen düalizmin felsefe tarihine mirası olan zihin-beden problemine temel teşkil etmektedir. Descartes ruh ve bedenin etkileşimini "Ruhun asıl meskeni beynin ortasında bulunan guddedir.[29] Ruh buradan bedenin her yerine zerreler ve sinirler vasıtasıyla ışık saçar, bir de bu zerrelerin izlenimine katılıp onları atardamarlar yoluyla tüm organlara geçirebilen kan vasıtasıyla." (Descartes, 2017: 49) şeklinde açıklasa da bu yorum zihin felsefecileri için tatmin edici olmamıştır.

2.2.1.1. Etkileşimcilik

Töz düalizmi zihin ve beden arasındaki ilişki bağlamında üç ayrı grup altında incelenebilir. Bunlardan ilki zihin ve bedenin birbirlerini karşılıklı olarak etkiledikleri iddiasına dayanan *etkileşimcilik*tir. Descartes'ın da dâhil olduğu etkileşimci düalistlere göre, beden ve zihin arasında iki yönlü bir nedensel etkileşim söz konusudur. Etkileşimcilik kuramında hem beden (beyin) zihin (bilinç) üzerinde nedensel etkilerde bulunmakta, hem bilinçli zihin durumları nedensel olarak birbirlerini etkilemekte, hem de bilinçli zihin durumları beden ve davranış üzerine nedensel etkilerde bulunmaktadır.

Etkileşimcilik, deneyimi ve deneyimleyeni ihmal etmeyen bir yaklaşımdır. Bu yaklaşımda, fenomenal öznenin doğasını bilinçli zihin durumları oluşturmaktadır. Descartes'ın "Düşünüyorum, o hâlde varım" ilkesi bir anlamda bilinçli ben'in başka bir deyişle fenomenal

[29] Bez, epifiz bezi.

öznenin indirgenemez ve bölünemez yapısına işaret etmektedir: "Bilincimin içerikleri değişebilir, şunu görüyor, bunu düşünüyor olabilirim. Ama her halükârda bunları bilinçli kılan benim bilinç içeriklerim olmalarıdır, benim onların farkında olmamdır." (Tura, 2017: 48). Descartes'ın bu iddiası ontolojik bir iddiadır ve bilincin varlığı için düşünen benin (*ego cogito*), bilinçli benin başka bir ifadeyle de deneyimleyenin (fenomenal öznenin) varlığını zorunlu kılmaktadır:

1- Her deneyimin bir deneyimleyeni (fenomenal öznesi) vardır.
2- Deneyimleyen (fenomenal özne) ontolojik olarak gerçek bir varlıktır.
3- Deneyimleyeni (fenomenal özneyi) anlamaksızın ne deneyimin kendisi ne de fenomenal bilincin doğası kavranabilir değildir.

olarak belirlediğimiz ölçütleri yerine getiriyor görünümünde olan etkileşimciliğin karşı karşıya olduğu problem öznenin, deneyimin ya da bilinçli zihin durumlarının doğası ile değil, bilinçli zihin durumlarının fiziksel ve biyolojik gerçeklikle ilişkisi ile ilgilidir. Etkileşimci düalizmde bu öznel bilinçli zihin durumlarının fiziksel ve biyolojik gerçeklikle bağlantısı ortaya konulamamış, aralarında nedensel bir ilişki olduğunu söylemekten ileri gidilememiştir. Fakat hiçbir bilimsel kuram etkileşimciliğin yumuşak karnı olan etkileşimin nasıllığı problemine kesin bir çözüm getiremediğinden etkileşimci töz düalizmi çok fazla taraftarı olmasa da henüz geçersiz kılınmış değildir.

2.2.1.2. Epifenomenalizm

Töz düalizminin bir diğer türevi, zihin ve beden arasındaki ilişkinin bedenden zihne doğru tek taraflı bir etki ile sürdürüldüğünü savunan *epifenomenalizm*dir. Bu yaklaşım, bedendeki olayların zihinsel olaylara neden olduğunu, ancak zihinsel olayların, ne başka zihinsel olaylara ne de fiziksel olaylara sebebiyet vermediğini ifade ederken zihni bir epifenomen, yani gölge olarak adlandırır. Zihnin, beden üzerinde herhangi bir nedensel etkisi söz konusu değildir. Bilinçli zihin durumlarının ne beyin ne de diğer zihin durumları üzerinde bir neden-oluş gücünden bahsedilemez. Bilinçli durumlar, beyin ve beden üzerinde etkisi bulunmayan yan unsurlardır. Epifenomenalizm, bu yönüyle aslında sadece fiziksel bir açıklama modeli sunmaktadır. Epifenomenalizme göre "Elimi kaldırmam, sadece düşüncem tarafından kendisine neden olmuş gibi görünen, katışıksız bir biçimde fiziki bir eylemdir."

(Büyük, 2013: 136). Epifenomenalist kuramda zihin durumları bağımsız bir varoluşa sahip olsalar da zihinde gerçekleşen olaylar bedensel olayların bir ürünüdür.

Epifenomenalizmin zihin ve bilince pasif bir rol ile varlık atfeden bakış açısı, zihinsel hayatı evren içerisinde etkisiz bir gölgeye dönüştürerek, bilinçli bir varlık olarak insanı da varlığı ile yokluğu arasında hiçbir fark bulunmayan gereksiz bir gölge şeklinde resmetmektedir. Bu yaklaşım insanın tüm davranışlarının nedenini fiziksel boyutta açıklarken iradeyi de yok saymış olmaktadır. Epifenomenalizm öznel deneyim ve deneyimleyene yer açan düalist bir teori olsa da öznenin doğasına hiçbir neden-oluş gücü vermeden bilinçli bir zombi tasavvurundan öte bir şey sunmamaktadır.

2.2.1.3. Paralelizm

Etkileşimcilik ve epifenomenalizmin aksine zihin ile beden arasında herhangi bir doğrultuda nedensel bir ilişki olmadığını söyleyen *paralelizm*e (koşutçuluk) göre ise bilinç ve beyin, aralarında nedensel bir ilişki bulunmaksızın eş-zamanlı ve birbirlerine paralel olarak vardır (Revensuo, 2016: 437). Fiziksel ve zihinsel durumlar eş-zamanlı ve uyumlu olarak gerçekleşirler. Bu uyum, nedensel bir ilişkiden değil çok iyi bir bağıntıdan ya da önceden kurulmuş bir düzenden kaynaklıdır. 17. yüzyılın önemli filozoflarından olan Leibniz de paralelciliği savunurken ruh ve bedenin, aynı zamana kurulmuş iki saat gibi, birbirleri ile ilişkiye girmeden paralel olarak çalıştıklarını ifade etmiştir. Bilindiği gibi, Leibniz, şuurlu, sonsuz sayıda, her biri bölünemez bir birlik olan monad adlı bir cevher kabul eder (Leibniz, 1997: 1). Monadlar, fiziksel değil metafizik noktalar olarak kabul edilmelidir, zira onların boyutu yoktur ve değişmezdirler. Monadlar, kendi içlerine kapalı olduklarından dolayı birbirlerini etkileyemezler. Fakat her monad evrenin tümüne içeriksel olarak sahiptir: "Her monad, evrenin bir aynasıdır." (Dölek, 2010: 218). Monadları birbirlerinden farklılaştıran özellik, her monadın evren tasarımının farklı derecelere sahip olmasıdır. Buna göre, en yetkin monad Tanrı'dır ve daha sonra sonsuz sayıda monad onun altında bulunur. Leibniz'in paralelizmine göre, tüm monadlar önceden kurulmuş tanrısal bir plan dâhilinde uyum içerinde hareket etmektedir. Bu önceden kurulmuş ahenk, ruh ve bedenin de birbiri ile uyumlu davranmasına yol açmaktadır.

Paralelizm ilkece öznel deneyim ve öznenin varlığına müsaade etse de sahip olduğumuz öznel bilinçli durumların tanrısal bir müdahale ile beden durumlarımızla paralel olarak

hareket ettiğini söyleyerek zihin-beden problemini teolojik açıdan çözmeyi denemektedir. Elde edilen tüm biyolojik ve nörofizyolojik veriler etkileşim içerisinde bulunmayan bir paralel zihin-beden tahayyülünden ziyade birbirleri üzerinde etkileri olan bilinç-beyin ilişkisine işaret etmektedir. Dolayısıyla günümüz bilimsel gelişmeleri teolojik bir zihin kuramına bağlı kalmayı gereksizleştirmekte ve paralelizmi geçersiz kılmaktadır.

2.2.2. Nitelik Düalizmi

Töz düalizmi evrende iki tözün olduğu fikrine yaslanırken nitelik düalizmi (*property dualism*) var olanın iki farklı töz değil, iki farklı nitelik olduğu görüşünü temel almaktadır. Nitelik düalizminin dayandığı ana fikir, fiziksel tözün (beyin) dışında başka bir töz bulunmasa da beynin başka hiçbir fiziksel varlıkta bulunmayan birçok özel niteliğe sahip olduğu ve sahip olduğu bu özel niteliklerin fiziksel olmadığıdır (Churchland, 2012: 17). Kaygılı olmak, ağrı çekmek, sarı rengin deneyimi, yumuşaklık hissi, müziğin verdiği keyif, sufle yemeyi arzulamak gibi nitelikler, beynin fiziksel durumlarından farklı ancak beynin niteliği olan zihinsel durumlardır. Bu niteliklerin fiziksel kavramlara indirgenerek sadece fiziksel kavramlarla açıklanması mümkün değildir. Nitelik düalizminde zihin bağımsız bir töz olarak var olmamakla birlikte, zihin durumları fiziksel durumlardan tamamen farklı ve onlara indirgenemeyecek özelliklere sahiptirler. Düalizmin bu formunda ikilik, zihin ve bedenin varoluşlarında değil zihin durumları ile fizik durumlarının karşıtlığında saklıdır.

Nitelik düalizmine göre niteliksel bir doğaya sahip olan bilinç, fiziksel ve davranışsal durumları kategorize etmenin bir yolu değil, gerçekten olan ya da beliren bir fenomendir (Robinson, 2017: 13). Nitelik düalistlerine göre bilinçli durumlar, beyinden türeyen durumlar olsa da fiziksel değildirler ve fiziksele indirgenemezler. Tüm fiziksel açıklama modelleri ve nörolojik durumlar bilinçli zihin durumlarının temel niteliklerini dışarıda bırakmaktadırlar. Oysaki nöral durumların dışında fiziksel olmayan nitelikler de mevcuttur, bunlar da zihinsel niteliklerdir. Nitelik düalizminde deneyimi içeren ayrıca bağımsız bir töz söz konusu değildir. Buna göre öznel deneyimler, beyin tarafından üretilip beyni etkiler ancak beynin herhangi bir fiziksel niteliğiyle aynı değildir.

Materyalistler, nitelik düalizminin beyinden beliren zihinsel özellikler senaryosunu fazladan bir açıklamaya muhtaç olduğu için gereksiz görmektedirler. Materyalistlere göre, beynin fiziksel durumlarından çıkan özelliklerin fiziksel olmayan bir açıklama modeline gerek

yoktur, tüm özellikler pekâlâ fiziksel özelliklere indirgenebilirdir. Nitelik düalistleri ise fiziksel durumlardan ortaya çıkan zihinsel özelliklerin farklı bir tür olduklarını zira şimdiye kadar beyin araştırmalarının zihinsel özelliklerin doğasına dair ikna edici fiziksel bir açıklama getirmediğini öne sürmektedir (Crumley, 2006: 26).

Nitelik düalizmi hem töz düalizminin büyük açmazı olan etkileşim problemi ile karşı karşıya kalmadığı için hem de öznel bilinçli durumları indirgemek zorunda olmadığı için bu iki yaklaşıma kıyasla daha avantajlı konumdadır. Zihinsel olana ayrı bir ontolojik statü atfetmese de bilinçli zihinsel niteliklere fiziksel niteliklerden farklı bir varoluş atfederek hem bilincin öznel niteliksel doğasını göz ardı etmemekte hem de zihin-beden problemi gibi kadim bir meseleyle uğraşmamaktadır. Fakat zihinsel niteliklerin meydana gelişini açıklamada birtakım problemler baş göstermektedir. Nitelik düalizmi bu hususta iki farklı koldan ilerleme imkânına sahiptir. İlk kolda, zihinsel niteliklerin gerçekliğin temel nitelikleri olarak, elektromanyetik niteliklerin kendi başlarına temel nitelikler olduklarının sonradan anlaşılması gibi tarihsel bir örneğe yaslanarak, evrenin başlangıcından beri var oldukları iddia edilebilir (Churchland, 2012: 20). Zihinsel niteliklerin evrenin başlangıcından beri var olan temel nitelikler olduğunu öne sürmek zihinsel niteliklerin gerçekliğin en alttan başlayarak gerçekliğin her düzeyinde kendini gösterdiğini savunmak olacaktır ki bu da panpsişizm demektir. İkinci kolda ise zihinsel niteliklerin, doğru bir şekilde organize olmuş karmaşık sistemlerde fiziksel yapıdan zuhur eden fiziksel olmayan nitelikler oldukları savunulabilir. Bu iddianın da zayıf yanı bu belirmenin ya da zuhurun hangi şartlarda ve nasıl gerçekleştiğini açığa kavuşturmak olacaktır. Böyle bir iddianın temellendirilmesi için, zihinsel niteliklerin ortaya çıkışını sağlayan gerek ve yeter koşulların eksiksiz bir şekilde ortaya konulması gereklidir. Bu da şu an için başarılmış görünmemektedir.

Nitelik düalizmi, öznel ve niteliksel deneyimin varlığına müsaade eden bir bilinç yaklaşımıdır. Zira zihinsel nitelik olarak bilinçli deneyimler fiziksel olmayan bir yapıya sahiptirler. Deneyimin varlığı, deneyimleyenin varlığını da zorunlu kıldığından nitelik düalizminde deneyimleyen öznenin varlığına da yer vardır. Ancak bu fenomenal öznenin doğası hakkında açık bir görüş bildirilmemiştir. Zihinsel nitelikler, beyinden beliren nitelikler ise "Bu niteliklere sahip öznenin doğası yalnızca fiziksel mi olacaktır yoksa ayrı bir ontolojik statüye sahip olacak mıdır?" sorusunun cevabı net değildir. Eğer zihinsel nitelikler, gerçekliğin temel nitelikleri ise öznenin doğasının da temel nitelikleri olarak ele alınacak ve fenomenal özne açıklaması bağlamında sorun ihtiva etmeyecektir.

2.3. Panpsişizm

Panpsişizm evrendeki her fiziksel şeyin aynı zamanda bir zihne ya da bilinçli bileşenlere sahip olduğunu iddia eden bir teoridir (Goff, Seager ve Allen-Hermanson, 2017: 10). Bu teoriye göre, her bir molekül ya da atom bilinçli parçacıklardan oluşmakta, zihinsel ve fiziksel özellikler de her bir varlıkta mutlak olarak bir arada bulunmaktadır.

Panpsişizmin yakın zamanda geliştirilen formlarından birisi olan proto-panpsişizm David Chalmers tarafından zihin felsefesi literatürüne kazandırılan, fiziksel her şeyin yetkin bir bilince sahip olmasa da çok basit bir bilince (ilk-bilinç) sahip olduğunu ileri süren bir bilinç teorisidir (Revonsuo, 2015: 82). Fiziksel parçacıkların sahip olduğu bilinç içerikleri çok basit düzeyde olduğundan onları birer bilinç örneği olarak kabul etmekte zorlansak dahi bu basit bileşenler insan beyninde kompleks bir yapıya bürünerek büyük bir bilinç sistemi oluşturacak şekilde organize olmuşlardır.

Panpsişizmin yine yakın tarihte önemli iki savunucusu Thomas Nagel ve Galen Strawson'dur. Her ikisi de fiziksel olandan beliren ya da ortaya çıkan zihinsel özelliklerin meydana geliş şeklinin irrasyonel ya da gizemli bir beliriverme olmadığını savunmaktadırlar. Bilince sahip olmayan fiziksel bir varlıktan nasıl olup da bilinçli içerikler çıktığı sorusunu bertaraf edecek tek yaklaşımın insan ve hayvan bilincinin ancak mikro basit düzeydeki bir bilinçten türeyebileceği tezi olduğunu savunan Strawson'a göre bu türeyiş ya da beliriş rasyonel düzeyde gerçekleşmektedir (Goff, Seager ve Allen-Hermanson, 2017: 24). Mikropsişizm adı verilebilecek bu yaklaşımda bilinç ve deneyim en alt düzeydeki fiziksel varlıklardan yukarıya doğru her düzeyde bulunmaktadır.

Panpsişizmin kullandığı bir diğer argüman da fiziksel olanın doğasına yöneliktir. Bu argümana göre "fiziğin, fiziksel olanın doğru bir teorisi" olduğunu söylemek totolojik bir ifadeden daha fazlası değildir, önemli olan fiziksel olanın teorisini değil doğasını keşfetmektir (Goff, Seager ve Allen-Hermanson, 2017: 26). Bir elektronun temel iç doğası nedir? Bu doğanın bilinçli olduğuna ya da olmadığına dair elimizde hiçbir dayanak noktası mevcut değildir. O hâlde elektronların mikro düzeyde bir bilince sahip olduğunu söylemek hem çelişkiye yol açmayacak hem de karmaşık zihin ve bilinç yapısının açıklanmasında kolaylaştırıcı rol üstlenecektir. Bu nedenle panpsişizm özellikle evrim teorisi için güçlü bir yol göstericidir.

Sonuç olarak panpsişizm, küçük mikro parçacıkların dahi temel doğasının öznel ve niteliksel deneyim olarak ele alınmasına imkân tanımakta, bu yönüyle evrendeki en küçük parçaya dahi bir özne statüsü vermekte bir sakınca görmemektedir. Tüm fiziksel dünyayı oluşturan şeyin öznel ve niteliksel deneyim olma ihtimalini açık bırakan bu yaklaşımda, her zerre fenomenal bir özneye tekabül etmektedir. Fiziksel adlandırılan her şeyin deneyimsel olma ihtimali de aynı şekilde mevcuttur. Her bir parçacığı özne olarak değerlendirmek sezgisel olarak çok savunulabilir olmasa da mümkündür. Ancak bu noktada bilinçliliği belirleyen fenomenal özne değil, özneyi oluşturan bilinçlilik olmaktadır. Resmin tamamlanmış hâlinde karşımıza küçük küçük öznelerden oluşan fiziksel evren büyük bir özne olarak çıkmaktadır(!) Küçük öznelerin hangi yolla ve nasıl birleşerek daha büyük özneler oluşturduğu sorusu da panpsişizmin en büyük açmazı olarak kabul edilmektedir. Parçacık öznelerin bir araya gelerek büyük özneler var etmesinin imkânsızlığına vurgu yapan bu açmaz *özne-toplaması*[30] problemi olarak adlandırılmaktadır, zira bu toplama ve birleşme işlemi teorik olarak mümkün değildir (Goff, Seager ve Allen-Hermanson, 2017: 34). Her fiziksel parçacığın sahip olduğu öznel ve niteliksel deneyimler ile büyük bir öznel deneyim oluşturması kategorik bir hatadır.

2.4. Agnostik Gizemcilik

Agnostik gizemcilik bilincin doğasının bilinemezliğine vurgusuyla ilk defa Owen Flanagan (1992) tarafından ortaya konulan, Colin McGinn tarafından geliştirilip sistematize edilen bir bilinç teorisidir (Rowlands, 2007: 335). McGinn'in gizemciliği bilincin doğal dünyanın bir parçası olduğu ve doğal dünyanın bize sunduğu bir bilinç açıklamasının olmadığı öncüllerinden yola çıkarak, bütün bir bilinç açıklamasının mümkün olamayacağı sonucuna varmaktadır.

Gizemciliğin temel tezlerinden birisi zihin-beden probleminin çözülemezliğine dayanmaktadır. Zira bilincin bilgisine erişmenin tek yolu iç-gözlemdir. Ancak dış dünyanın bilgisi algısal deneyim ile elde edilmektedir. Algının dış dünyanın bilgisini açığa çıkarma şekli ile iç-gözlemin bilinci açığa çıkarma şeklinin çok farklı olduğuna dair güçlü

[30] *Subject-Summing Problem*

sezgilerimiz vardır; bu sebepledir ki birinin diğerinin oluşması için temel sağlıyor olması imkânsız görünmektedir (Rowlands, 2007: 339).

Gizemciliğin dayandığı bir diğer tez ise bilincin mekânsal bir varoluşu olmamasıdır. İç-gözlem yoluyla bilince mekândan bağımsız olarak erişilebilirken sahip olduğumuz tüm diğer fiziksel bilgiler mekâna bağımlıdır dolayısıyla mekândan bağımsız bilincin mekânsal olan beyinden belirmiş olması mümkün değildir (McGinn, 2004: 105). Bilincin mekândan bağımsız oluşu, onun doğasını bilmemize engel olmaktadır. Kökleri Kant'a dayanan bu argüman epistemik kapasitemizin zaman ve mekân ile sınırlı olduğu görüşüne atıf yapmaktadır. Bilincin gizemi, bilincin kendisinden değil bizim mekânsal olmayana bilişsel ya da epistemik olarak kapalı oluşumuzdan kaynaklanmaktadır.

Agnostik gizemcilik, bilincin ve öznel deneyimin doğasının bilinemezliğini ve epistemik erişilemezliğini savunurken bilinçli öznenin ontolojik yapısını ve doğasını bilme imkânımızı da ortadan kaldırmaktadır. Deneyimleyen olmadan deneyimden bahsedilemez ve deneyimleyenin doğasını açıklamayan hiçbir teori bilincin de doğasını açıklayamaz. Agnostik gizemcilik, bir gün tüm nörobilim ve beyin araştırmaları kusursuz bir şekilde sonuçlandırılsa ve beynin tüm çalışma mekanizması eksiksiz bir şekilde çözülse dahi bilinçle ilgili keşfedeceğimiz ve bilincin nasıl ortaya çıktığını anlayacağımız hiçbir şey olmayacağını varsaymaktadır. Bilinç ve bilinçli özne probleminin çözümündense sonsuz bir gizemi tercih eden bu yaklaşım gelecek için umutsuz bir perspektif çizmekten öte yeni bir şey vaat etmemekte ve bilinç için olumlu bir açıklama modeli sunmamaktadır. Bilincin gizemi, öznenin gizemidir ve varlığından sezgisel olarak emin olduğumuz özneye yüklenen bu gizem, insanlar için heyecan verici gözükse de bilim için hiçbir katkı sunmamaktadır.

2.5. Bilince Yönelik Temel Yaklaşımların Değerlendirilmesi

Bilince yönelik bütüncül bir bakış açısının, ancak deneyimin, deneyimleyenin ve deneyimlenenin doğasını göz ardı etmeyen bir yaklaşımla mümkün olabileceği düşüncesiyle yola çıkılan bu çalışmanın temel tezi öznel ve niteliksel bir yapıya sahip olan deneyim ile deneyimleyenin doğası anlaşılmadan bilincin kavranabilme imkânının bulunmadığıdır. Bu anlamda, bilincin doğasına yönelik söyleyecek sözü olan her teorinin hem deneyimin kendisi hem de deneyimleyen ve deneyimlenene dair, bir diğer ifadeyle de özne ve nesnenin doğasına dair söyleyecek sözü bulunmalıdır. Bilinçli deneyimlerimizi kesintisiz bir akış

hâlinde bir arada tutan bir öznenin sahip olduğu yapı açığa kavuşturulmadan tek tek deneyimlerin açıklanması bize öznellik ve nitelikselliğin kaynağını vermeyecektir; zira deneyimleyeni olmayan bir deneyimden bahsetmek mümkün değildir. Bilincin zor problemi onun sahip olduğu öznel ve niteliksel doğadan kaynaklanmaktadır. Bu zor problemin anlaşılmasının yegâne yolu ise öznel ve niteliksel doğanın kaynağı olan deneyim ve özneden geçmektedir. Bu öncüllerden yola çıkılarak dört farklı bilinç yaklaşımının ele alındığı bu bölümde ulaşılan sonuçta incelenen monist, düalist, panpsişist ve agnostik yaklaşımların deneyimin ve öznenin doğasını ya tamamen göz ardı ettikleri ya da inkâr ettikleri görülmüştür.

İdealizm öznenin nörofiziksel özelliklerini göz ardı ettiği için, indirgemeci materyalizm öznel deneyim ve bilinci göz ardı ettiği için, elemeci materyalizm bilincin ve deneyimin varlığını inkâr ettiği için ve son olarak indirgemeci olmayan materyalizm ise zihinsel nitelikleri yine fiziksel olanın bir özelliği olarak gördüğü için öznel bilinçli deneyimlere sahip öznenin ve deneyimin doğasını açıklayamamaktadır. Buna karşılık düalizm bilinçli öznenin varlığını ilkece kabul ediyor görünse de maddeyle olan ilişkisini açıklayamadığından doyurucu bir teori olarak ele alınamamaktadır. Panpsişizm ise var olan her şeyin belirli bir derecede bilince sahip olduğunu ileri sürerek, her zerrenin bir özne olarak inşa edilebilmesine imkân tanımakta ancak fenomenal anlamda bir özneye kapı aralamamakta ve deneyime dair öngörüde bulunmamaktadır. Agnostik gizemcilik ise en baştan ortaya koyduğu bilincin doğasını kavramamızın mümkün olmadığı teziyle deneyim ve deneyimleyenin doğasının kavranabilmesinin de mümkün olamayacağını ima etmektedir. Sonuç olarak, ele alınan teorilerin hiçbirisi insan bilincinin doğasını, deneyimleyen özneyi ve deneyimin kendisini dikkate alarak anlamamıza müsaade etmemektedir. Bilincin muamma ve gizemli bir fenomen olarak anılmasından vazgeçilmesi ancak deneyimleyen öznenin ve deneyimin doğası hakkında açıklama sunan bir bilinç yaklaşımıyla mümkündür. Bu nedenle bir sonraki bölümde, bilince yönelik güncel felsefi teoriler ele alınarak deneyimleyenin doğasına dair kuşatıcı bir bakış açısı yakalanmaya çalışılacaktır.

3. BİLİNCE GÜNCEL YAKLAŞIMLAR

Gerçekliği, gerçekte olduğu gibi algılayabilseydiniz,
renksiz, kokusuz, tatsız sessizliği karşısında hayrete düşerdiniz.
David Eagleman

Güncel felsefi bilinç teorilerinin temelleri fikirsel olarak ikinci bölümde ayrıntılı olarak ele aldığımız dört bilinç yaklaşımına dayanmaktadır. Ancak bu güncel teoriler, hem yararlandıkları bilimsel verileri işleyişleri ile hem de klasik teorilerin problemlerinin farkında oluşlarıyla klasik teorilerden ayrılmaktadır. Klasik yaklaşımların sınırlarını geliştiren ve dönüştüren bakış açıları olarak güncel felsefi bilinç teorileri, zihin felsefesinin yükselişinde önemli rol oynamaktadır. Güncel felsefi bilinç teorileri, bilinç mefhumunun açıklanması yönünde katedilen aşamayı göstermeleri bakımından da bilhassa önemlidir. Bu bölümde bu güncel teorileri ele alırken yine en önemli ölçütümüz deneyim ve deneyimleyenin doğasına açıklama getirip getirmedikleri olacaktır. Ayrıca bu güncel yaklaşımlardan elde ettiğimiz faydalı ipuçları, son bölümde geliştireceğimiz bilinç teorisi için kurucu nitelikler olarak kullanılacaktır.

3.1. Çoklu Taslaklar Kuramı[31]: Dennett Örneği

Çoklu Taslaklar Kuramı 1991 yılında materyalist bir filozof olan Daniel Dennett tarafından *Consciousness Explained*[32] kitabında ortaya konulmuştur. Bu kuramı klasik materyalistlerden farklılaştıran yaklaşım, bilincin gizemli bir fenomen olarak adlandırılmasına yol açan etkenleri ve içsel deneyimin doğasını inkâr etmeden önce anlamaya çalışmasıdır.

Dennett'e göre bir fenomenin gizemli olması demek, onun üzerine nasıl düşüneceğimizi ve nasıl çalışacağımızı bilmememiz demektir ve bu anlamda bilinç de nasıl açıklanacağı yönünde henüz mutabakata varılmamış olan, ayakta kalmayı başarmış en son gizemdir (Dennett, 2017a: 31). Bilincin gizemi, Dennett'a göre bize verdiği öznellik ve ayrıcalık hissinden kaynaklıdır. Buna göre, yalnızca kendimize ait bir olgu olarak deneyimlediğimiz bilincin, herkes tarafından nesnel bir olgu olarak kavranabileceğini kabul edersek, öznel

[31] *Multiple Drafts Model*
[32] Açıklanmış Bilinç

"benlik" algımızı kaybedeceğimizden endişe duyarız. Çocukken yaşadığımız aşkların üzerine kalbimizin pırpır edişinde, midemizde kelebekler uçuşmasında, ellerimizin titreyişinde sinen masumiyet nasıl büyükçe kayboluyor ve aşk rasyonel bir olgu hâlini alıyorsa, bilincin de gizemi çözüldüğünde aşka benzeyeceğinden korkarız. Bu sebeple de bilinci açıklanamaz bir mefhum olarak kabul etmeyi bilinci açıklamayla uğraşmaya yeğleriz.

Dennett kuramının daha ilk başında bilincin gizemli dünyasının doğasına odaklanır. Dennett'in gizemli bilinç anlayışına gösterdiği ilk gerekçe, bilinçte meydana gelen olayların deneyimine sahip, onlara tanıklık eden bir deneyimleyen olması gerekliliği varsayımıdır. Bu varsayımı şöyle ifade eder:

> Bir düşüncenin olabilmesi için, bir zihnin onu düşünmesi gerekir, bir ağrının hissedilebilmesi için birisinin onu hissedebilmesi gerekir ve mor bir ineğin birden imgemizde var olmaya başlaması için, birisinin onu düşlemesi gerekmektedir. (Dennett, 2017a: 40).

Dennett'a göre buradan hareketle inanırız ki; deneyimlenen olayları bilinçli hâle getiren deneyimleyendir. Deneyimleyenin bu şekilde ayrık bir kategoride ifade edilmesi ise bilincin çözülemez bir gizemmiş gibi algılanmasının başat sebebidir.

Dennett'in bilinci gizemli yapısına sunduğu ikinci sebep ise öznel hislerin ve qualianın[33] (nitelcelerin) varlığıdır. Qualia problemi denilince akla gelen birincil yaklaşım işlevselciliktir zira qualia problemi işlevselciliğin sunduğu bakış açısına karşıt olarak formülüze edilmiştir. Bilinci insanın bilişsel sisteminin işlevsel yapısının bir parçası ya da sonucu olarak gören işlevselcilik qualianın varlığını kabul etmez. İşlevselciliğe göre, bir gurmenin bilişsel sisteminin işlevsel yapı ve özellikleri yeniden üretilse, bütün zihinsel süreçleri ile birlikte gurmeliği de yeniden üretilecektir. Silikon bir yapıda dahi yemek tatlarını ve onlardan alınan zevkleri üretmek mümkündür. İşlevselciliğin aldığı en büyük eleştiri hiçbir makinenin bir insanın beyin süreçlerini ne kadar doğru taklit ederse etsin, bir Beethoven sanatını, bir basketbol maçını ya da çilekli turta yemeyi sevme kabiliyetine sahip olamayacağı yönündedir (Dennett, 2017a: 43). Sevmek, üzülmek, mutlu olmak, âşık olmak hepsi birer öznel deneyimdir ve dolayısıyla bilinç gerektirir; bu da makinede yoktur. İşlevselciliğin en zayıf noktası olarak görülen qualia sorunu, Dennett'a göre günlük bakış

[33] Her bilinçli durum için ayrı bir niteliksel his mevcuttur. Bilinç üzerine çalışan araştırmacılar, bu öznel hislerin, nesnel dünya içerisinde nereye yerleştirileceği konusunda uzlaşmış değiller. Bu öznel durum ve olaylar qualia, onların bizim genel dünya görüşümüz içinde kalarak açıklanması sorunu da qualia sorunu olarak isimlendirilmektedir (Searle, 2018: 19).

açısında bilincin doğasının gizemine katkıda bulunur. Ona göre, öznel ve niteliksel durumların kabülü, bilinç kavramını bilimsel yaklaşımın dışına taşıyarak bilince açıklanamaz bir yön atfeder.

Dennett bilincin açıklamasını zorlu kılan bir diğer etmenin de özbilgiye yönelik kapasitemizin sınırlılığı olduğunu belirtir çünkü klasik kavrayışa göre, kendimize yönelik belirli bir erişime sahibiz ve öznel deneyimler de bu erişim ile mümkün olmaktadır (Dennett, 2017b: 102). Bu öznel erişim, hepimizin sanki bir bakış açısının içerisine sıkışmış olduğunu ve kimsenin bir başkasının bakış açısına giremeyeceğini ima etmektedir. Öyle hissediyoruz ki der Dennett:

> Bu bakış açısı sadece bize ait; atomlarda, moleküllerde, ağaçlarda, kıtalarda, galaksilerde vb. böyle bir bakış açısından bahsetmemiz mümkün değildir, hatta başka zihinler sorununda olduğu gibi başka insanlarda bu bakış açısı var mı yok mu ondan bile emin değiliz. (Dennett, 2017b: 102).

Yine son olarak, Dennett'a göre, bilincin bizi otomattan ayıran bir fenomen olarak görülmesi ve bir eylemin arkasında bilinçli bir zihin yoksa sorumluluk sahibi bir öznenin olmayacağı anlayışı da bilincin insana özel, gizemli bir fenomen olarak adlandırılmasının önemli bir sebebidir. Birçok insana göre, bizim gerçek bakış açılarımız ya da gerçek farkındalığımız ve dolayısıyla gerçek bilincimiz ile robotların ayırt eden ama sezgilere sahip olmayan varlıklar olarak bakış açıları ve farkındalıkları arasında kapanması çok büyük boşluklar bulunmaktadır (Dennett, 2017b: 103). Bu bağlamda, bilincin gizemliliği Dennett'a göre insanların sahip içgüdüsel olarak oldukları her ne kadar yanlış olsa da ikna edici sezgilerden kaynaklanmaktadır.

Dennett, bilincin açıklanamazlığının altında yatan gerekçeleri tespit ettikten sonra, bu sebeplerin bizi bilince yönelik yaygın bakış açısına; daha açık ifadeyle de bilinçli zihnin sadece beyinden ibaret olamayacağı ya da beynin herhangi bir parçası olamayacağı zannına ulaştırdığını ifade eder. Zira sezgilerimize göre beyindeki şeyler şunlardan hiçbirisi olamaz:

> (1) mor ineğin oluşturulduğu ortam, (2) düşünen şey, 'düşünüyorum, öyleyse varım' ifadesindeki ben, (3) turtayı seven, ırkçılıktan nefret eden, umursayan, âşık olan, (4) ahlaki sorumlulukla eyleyen (Dennett, 2017a: 45).

Dennett'a göre insan bilincine dair kabul edilebilir bir kuram, beyinden farklı zihinsel bir tözün varlığına vurgu yapan bu dört uğraştırıcı problemin bir açıklamasını yapmak zorundadır.

Daniel Dennett tüm bilimlerin nesnel ve üçüncü şahıs bakış açısıyla yapılması gerektiğini savunur. Dennett bilişsel bilimlere ağırlık veren bir zihin felsefecisidir. Bilişsel bilimler zihni, bir bilgi- işleme sistemi olarak ele alır. Dennett de buna paralel olarak, bilinci bilgi işlemenin özel bir türü olarak değerlendirir. Dennett fenomenal bilinç ve erişim bilinci ayrımından yalnızca erişim bilinci kısmını dikkate alır.[34] Bilinç, zihindeki kişinin erişimine açık bilgi- işlemenin bir parçasından başka bir şey değildir. Bir kimsenin bazı bilgilere erişiminin olması, o kimsenin bilgiyi fiziksel ya da sözel davranış yoluyla ifade edebilmesidir. Bilince erişim yöntemini Dennett heterofenomonoloji olarak tanımlar ve fenomenolojiden farkını şu şekilde ortaya koyar:

> Eğer oto-fenomenoloji taraftarıysan, içebakışçıysan, birinci şahıs bakış açısını benimsiyorsan yanlış anlamaya mahkûmsun hepsi bu. Bilincini sahip olmadığı özelliklere sahip sanarak aslında kendini kandırıyorsun. Bilincin fenomenolojisini tarafsız bir şekilde sınıflandıracak bir yöntemin özelliklerini saptayıp ondan sonra işe koyulmak, açıklamaya böyle başlamak lazım. Böylece bilinci açıklamış olursun. Heterofenomenoloji, açıklanması gereken şeylerin bilimsel kataloğudur (Dennett, 2017b: 108).

Heterofenomenoloji, öznel deneyim verilerinin nesnel yöntemlerle sınanmasıdır. Bir bilinç bilimcisinin yapması gereken kişilerin deneyimledikleri ve bu deneyimlerin kendileri için ne demek olduklarına dair verileri ve bildirimleri toplayarak, bu verileri tanımlayan "anlatılar" oluşturmaktır. Bilinç bilimcisinin oluşturduğu anlatılar, fenomenolojik yöntemde olduğu gibi içsel bir özneye ve öznel, niteliksel bir dünyaya ait değildir. İnsanların bilinç hakkındaki günlük kavrayışları sürekli olarak onun öznel ve içsel bir fenomen olduğuna işaret etse de Dennett'e göre bu, sadece beynin bir yanılsamasıdır ve kendi tabiriyle beyinde gerçekleşen Kartezyen bir tiyatrodur (Revonsuo, 2016: 272). Kartezyen tiyatronun kişiye sunduğu ayrık, gizemli bir öznenin varoluşudur. Oysaki üretilen anlatıların arkasında gizemli bir özne veya fenomenal bir iç dünyanın varlığı söz konusu değildir. Qualia olarak adlandırılan saf, öznel hislerin ve bilincin ait olduğu ayrı bir "benliğin" varlığı söz konusu değildir.

[34] Fenomenal bilinç ve erişim bilinci ayrımı Ned Block (1995) tarafından yapılmıştır. Block bu ayrımı şöyle ifade eder: "Fenomenal bilinç deneyimdir; bir durumun fenomenal bilinçli yanı, o durumda olmanın ne gibi olduğudur. Karşıt olarak, erişim bilincinin işaret ettiği şey ise akıl yürütmede ve rasyonel konuşma ile eylem kılavuzunda kullanılmaya uygunluktur." (Block, 1995: 227).

Dennett bir şeyin varlığının anlamının, onun nesnel ve fiziksel olarak var olması demek olduğunu savunur. Dolayısıyla onun yaklaşımına göre, beynin hiçbir bölgesinde nesnel ve fiziksel varlığına dair bir işaret bulunmayan fenomenal içsel dünyanın varlığından bahsedilemez.

Dennett'a göre kartezyen tiyatro, başka bir deyişle içsel bir öznenin varlığı, mantıksal olarak da hatalı sonuçlara yol açar. Felsefi literatürde *homonkulus* (küçük insan) paradoksu olarak bilinen bu durum küçük insanlarla dolu sonsuz bir sürece atıfta bulunur (Revonsuo, 2016: 273). Buna göre, eğer beyinde qualianın sunulduğu mitik bir özne varsa bu mitik içsel özne de bu içsel hisleri algılamalıdır. Dolayısıyla onun bu algı ve farkındalığa sahip olması için, kendisinin de beyninin bir yerinde başka bir mitik özneye ya da bilince sahip olması gerekir. Ve bu süreç, sonsuz bir biçimde her bilincin, daha küçük başka bir bilinci gerektirmesiyle devam eder. Homonkuluslarla dolu bu sonsuz gidişatın doğru bir bilinç açıklaması vermesi de mümkün değildir. Dennett'e göre bizim "içsel özne" ya da "benlik" olarak adlandırdığımız şey, yalnızca kurgusal bir öznedir, fiziksel ve nesnel olarak var değildir.

Beyinde gerçekleşen bilgi akışları, Dennett'ın kuramına adını veren çoklu taslaklardır. Beyindeki bu bilgi akışları, çıktı sistemine erişmek için sürekli olarak bir yarış içerisindedirler. Çıktı sistemlerine erişmeyi başaran bilgiler, bilinci kurarak, bilinçli bilgi olarak adlandırılırlar. Bu da Dennett'ın bilinç tanımını ortaya koyar: "Bilinç, beyinde erişim yarışını kazanan, anlık bir üne veya politik etkiye kavuşan ve böylece davranışı hem kontrol eden hem yönlendiren bilgi içeriğinden ibarettir." (Dennett, 2005: 137).

Dennett'in çoklu taslaklar kuramı, bu bağlamda makine bilincine imkân tanıyan ve makine bilinci için de kullanılabilecek bir kuramdır. Öznel ve niteliksel deneyimi inkâr ederek elemeci materyalizme yakın duran Dennett'e göre, insan bilincini açıklamak karmaşık ve gelişmiş bir robotun davranışının bilgi türleriyle nasıl kontrol edildiğini açıklamaktır. Dolayısıyla çoklu taslaklar kuramı dayanak kabul edildiğinde insan ve bilişsel açıdan insana benzeyen bir robotun bilinç seviyeleri birbirlerinden farklı olmayacaktır.

Dennett'in bu kuramını ortaya koyduğu kitabının ismi *Consciousness Explained*'dir. Ancak bu isim, kitapta sunulan çoklu taslaklar kuramında bilincin açıklanmak yerine yalnızca üstünün kapatıldığı ve fenomenal bilincin tamamen reddedildiği gerekçesiyle yoğun bir eleştiriye maruz kalmıştır. Her ne kadar Dennett elenmesi gerekenin bilincin kendisi değil, bilincin fenomenal özellikleri ve qualia olduğunu söylese de bilincin öznel ve niteliksel

yönlerini göz ardı eden bu açıklama, bilincin bütününe yönelik değil yalnızca erişim bilincine yönelik bir açıklamadır. Bilişsel bilimlerin de bütün çabası zaten erişim bilincini açıklama gayesine yöneliktir. Ancak bilincin zor problemi olarak adlandırılan davranışlarımızdan bağımsız ve farklı olarak nasıl içsel bir deneyime sahip olduğumuzun açıklaması Dennett tarafından yapılmadığı gibi, bu öznel deneyimlerin varlığı da inkâr edilmektedir. Çünkü ona göre, sıralı işleyen paralel bir makineyi simule eden karmaşık bir bilgi işleme sistemi olan bilinç, insanlarda konuşarak kendisini ve deneyimlerini ifade eden, dışsal davranışlarını kontrol eden "tek bir özne" veya "benlik hissi" yaratır ancak bu his yalnızca bir yanılsama ve beyinde gerçekleşen bir tiyatrodur (Revonsuo, 2016: 274).

Sonuç olarak, Dennett'in bilincin varlığını yalnızca epistemolojik bağlamda ele alarak onu bir bilgi türüne indirgediği bu yaklaşım, deneyim ve deneyimleyenin ontolojik doğasını göz ardı etmektedir. Bu sebeple Dennett'in çoklu taslaklar kuramı, bu tez çalışmasının başlangıcında ortaya konulan ve "Bilincin doğasını açıklamaya yönelik her teori, deneyimleyenin doğasını da açıklamalıdır." şeklinde ifade edilen ölçütü karşılayamamaktadır.

3.2. Biyolojik Doğalcılık[35]: Searle Örneği

John Searle'e göre bilincin ayırt edici özelliklerinin başında öznel ve niteliksel karakteri gelmektedir. Öznellik, her bilinçli durumun yalnızca bir özne tarafından tecrübe edilmiş olarak mevcut olduğunu ima etmektedir. Searle'e göre, kullanılan terminolojide bu özelliğin doğuracağı problemler, bilincin öznel yanının hep göz ardı edilmesine neden olmuştur. Hatta ona göre,

> Pek çok materyalist felsefecinin, kelimenin en sıradan anlamında, bilincin varlığını reddetme isteği duymasına yol açan ve bilincin istisnai bir biçimde bizim genel bilimsel dünya görüşümüzle bağdaştırılmasını zorlaştıran özellik de budur. (Searle, 2006: 86).

Searle ayrıca her bilinç durumunun öznel olmasının yanı sıra kendine ait bir niteliksel hissi olduğunu da kabul eder. Ona göre, "Her bilinç deneyimiyle ilgili olarak bu deneyime 'gibi hisseden', 'gibi olan' bir şey bulunduğunu söyleyerek de konuya yaklaşabiliriz." (Searle, 2000: 560). Turta yemek gibi olan şey ile gül koklamak gibi olan şey birbirinden niteliksel

[35] *Biological Naturalism.*

olarak farklıdır. Searle'e göre "ağaç gibi olmak" ya da "ev gibi olmak" durumlarına benzer bir şey yoktur çünkü onlar bilinçli değildirler. "Gibi olmak" Searle'e göre yalnızca bilinçli durumlara atfedilebilecek bir özelliktir. Bu bağlamda, Searle'ün biyolojik doğalcılık yaklaşımının temeline koyduğu bilinç, varlığımızın içsel birinci şahıs deneyimi olan fenomenal bilinçtir.

Bilincin diğer ayırt edici özellikleri ise bütünselliği/birliği ve yönelimselliğidir. Searle'e göre,

> Bizim patolojik olmayan bilinç biçimlerinde, söz gelimi koldaki bir ağrı, sıcaklık hissi ya da bir şeyi kırmızı görme deneyimine kesinlikle tek başına sahip olmadığımızı, fakat tüm bu deneyimleri eş zamanlı olarak birleşik tek bir bilinç deneyiminin parçası olarak yaşadığımızı kabul etmemiz çok önemlidir. (Searle, 2005: 24).

Bilinçli deneyimin meydana gelmesi yalnızca bilinçli olayların diğer öznel niteliklerle ilişki içerisinde olduğu birleşik alan içerisinde mümkündür. Qualianın kendi başlarına var olması mümkün değildir.

Birçok zihin durumunun, şeylerin dünyadaki bağıntı durumlarına yönelmiş olmalarına ya da onlar hakkında olmalarını sağlayan özelliğe felsefe tarihinde "yönelimsellik" adı verilmektedir. Bunu şöyle açıklayabiliriz: Sahip olduğum bilinç durumları (inanç, arzu, korku vs.) bir içeriğe sahiptir ve bu içerik bir şey hakkında olmak durumundadır; bu bilinç durumlarımın dış dünyayla bağlantısını sağlamak için zorunlu bir özelliktir. Searle, bilinci kendi içine kapalı bir öznellik alanı olma izleniminden kurtarmak maksadıyla, öznel durumlarımızı dış dünyayla ilişkilendiren bilinç özelliğine "yönelimsellik" adını verir ve şöyle tanımlar:

> Yönelimsellik, zihnin nesneler ve dünyadaki şey durumlarına yönelmesini ve onlara dair olmasını sağlayan çeşitli biçimlerin hepsinin genel terimidir (Searle, 2006: 99).

Searle zihin teorisini ve daha özelde bilinç teorisini düalizm ve materyalizmden kaçınarak "biyolojik doğalcılık" olarak adlandırmaktadır: "'Doğalcılık'tır çünkü bu görüşe göre, zihin doğanın bir parçasıdır; 'biyolojik'tir çünkü zihinsel görüngünün varlığını açıklama tarzı biyolojiktir." (Searle, 2006: 65). Biyolojik doğalcılık yaklaşımı, bilincin biyolojik düzeydeki varlığına vurgu yaptığı için biyolojik, bilinç de diğer biyolojik fenomenler gibi doğal dünyanın bir parçası olduğu için doğalcılık kelimelerini içermektedir (Searle, 2007: 329). Biyolojik doğalcılığı Searle bilinç problemi ile bağlantılı dört temel tez üzerine inşa etmektedir:

1- Zihin durumları, gerçek fenomenler oldukları için onların eleyici bir indirgemesini yapmamız mümkün değildir.

2- Zihin durumları temel olarak beyindeki daha alt seviyedeki nörobiyolojik süreçlerden kaynaklandığı için, nörobiyolojik süreçlere "nedensel" olarak indirgenebilir zira kendilerinin nörobiyolojiden bağımsız hayatları yoktur.
3- Zihin durumları beyin sisteminin bir özelliği olarak ortaya çıkarlar ve nöronlardan daha yüksek seviyeye sahiptirler. Nöronlar tek başına bilinçli değilken beyin sisteminin nöronlardan oluşan parçaları bilinçlidir.
4- Bilinç durumları, susadığımızda su içmemiz gibi, nedensel olarak işlerler (Searle, 2004: 113).

Biyolojik doğalcılık öznel ve niteliksel bilinçli durumların varlığını kabul etmekte ve bilincin bu öznel ve niteliksel özelliklerinin beyindeki nörobiyolojik süreçlerden kaynaklandığını iddia etmektedir. Dolayısıyla bu yaklaşımda öznel deneyimlerin kaynağı fenomenal bir özne olarak değil bu deneyimlerin ontolojik olmasa da nedensel olarak indirgenebileceği nörolojik süreçler olarak gösterilmektedir. Biyolojik doğalcılık, bu bakış açısıyla bize bilinçli öznenin doğası hakkında bir açıklama sunmamaktadır. Searle beyin ve bilinç arasındaki ilişkiyi anlaşılabilir bir hâle getirmiş olsa da bu ilişkinin nasıl gerçekleştiğine dair kapsamlı bir açıklama sunamaz çünkü beynin bilince neden olduğunu belirtse de ne şekillerde neden olduğunu betimlemez ve bunun deneysel olarak hangi yöntemle ispatlanabileceğine dair bir çıkarımda bulunmaz.

3.3. Doğalcı Düalizm[36]: Chalmers Örneği

David J. Chalmers, savunmuş olduğu doğalcı düalist yaklaşımdan ziyade, bilincin kolay problemleri ve zor problemleri şeklinde yaptığı ayrımla zihin felsefesinin sıkça atıf alan isimlerinden birisi olmayı başarmıştır. Bilincin davranışla ilgili yönlerini ve bağıntılarını açıklamak Chalmers'a göre bilinçle ilgili kolay çözülebilecek problemlerdir. Örneğin, beyin duyusal girdilerle nasıl ilgilenir, onları nasıl yorumlayıp cevaplar, dili kullanarak içsel durumunu nasıl aktarır gibi sorular oldukça çetrefilli olsalar da bilincin değil, zekanın gizemleri olarak çözülebilir "kolay" sorunlardır (Tegmark, 2019: 365). Ancak bilincin, öznel deneyim ile bağlantısını, bir başka deyişle, nesnel dünyanın nasıl olup da öznel bilinci açığa çıkaracağı problemini açıklamak bilincin çözülmesi en zor problemini oluşturmaktadır. Chalmers bu ayrımı 1994 yılında bir konferansta kullandığını ve aslında yaptığı şeyin "malumun ilamı" olmaktan başka derinlikli bir yanı olduğunu düşünmediğini belirtir. Çünkü ona göre, fiziksel herhangi bir sistemin deneyimsel ve niteliksel durumlara

[36] *Naturalistic Dualism*

nasıl sahip olabildiği ve o durumları nasıl üretebildiği problemi herkesçe apaçık bir problem olarak kolaylıkla görülebilirdir (Chalmers, 2017: 56). Öznel deneyimin bir beyin sürecine indirgenemeyeceğini ve yalnızca beyin süreçlerine dayalı bir açıklamanın bilincin varlığını ortaya koymaya yetmeyeceğini düşünen Chalmers'a göre, zor sorun mekanistik ve indirgemeci bir yaklaşım ile çözülemez.

Dünyayı fiziksel bir yapboz olarak düşündüğümüzde, yapbozun tüm parçalarını - kimya, fizik, evrim, görelilik, kuantum mekaniği, DNA vs.- yerine yerleştirdiğimizde bile açıkta kalan ve yapboza uymayan bir parça olacaktır: Bu da "Sadece bir şeyler yapmakla kalmayıp, aynı zamanda çevremizdeki dünyayı deneyimlediğimiz gerçeğidir." (Manzotti, 2018: 12). Deneyimi fiziksel âlemin parçalarından birisi olarak açıklamanın zorluğu üzerinden zor problem tanımı yapan Chalmers'a göre, tam bir bilinç açıklaması deneyimin gerçekliğini göz önünde bulundurmalıdır. Chalmers'a göre, "Deneyim hakkındaki konuşma fire vermeden beyin ve davranış hakkındaki konuşmalara tercüme edilemez. Fiziksel olaylarla zihinsel olaylar arasında 'bir açıklama boşluğu' vardır." (Zeman, 2006: 388). Chalmers, dünyayla ilgili bütün gerçekleri bilsek dahi bilinci bir bütün olarak asla bilemeyeceğimizi düşünür; ancak, bilinçli deneyimler ile beyin süreçleri arasındaki sıkı bir bağıntı ya da bağlantının varlığını da inkâr etmez. Ona göre bilincin bilim olarak çalışılması için yapılması gereken bu bağıntı veya bağlantının sistematiğini ortaya çıkarmaktır (Chalmers, 2017: 57).

Chalmers, öznel deneyim ve qualianın fiziksel bir temelinin olmadığını ve fiziksel bir fenomene atıfla açıklanamayacağını savunur. Bilincin açıklanmasının yolunu şöyle ortaya koyar:

> Bilinç konusunda yapmamız gereken şey, onun tıpkı uzay ve zaman gibi indirgenemez olduğunu ve dünyanın temel bir niteliği olduğunu kabul etmektir. Bundan sonra yapacağımız şey ise onu yöneten yasalara, yani öznel deneyimin birinci şahıs verileri ile üçüncü şahıs nesnel fiziksel nitelikleri arasındaki bağlantıya bakmak olur. Sonunda, bu bağlantıyı yöneten ve fizikte keşfetmiş olduğumuz basit temel yasalara ulaşabiliriz (Chalmers, 2017: 58).

Chalmers'a göre öznel deneyim ve qualia evrenin temel özellikleridir. Temel özelliklerin kendilerinin altında yatan daha ileri bir açıklaması ve üstüne inşa edildikleri daha temel bir şey de yoktur, kendileri zaten evrenin en altındaki temel katmanı oluştururlar (Revonsuo, 2017: 282). Searle bilinç kuramının aynı zamanda biyolojik bir kuram olması gerektiğini savunurken; Chalmers bilinç kuramının temel fizikteki kuramlara benzer olarak yasalarla

açıklanması gerektiğini savunur. Psikofiziksel yasalar bilinç kuramının merkezinde olmalı ve öznel deneyimin fiziksel özelliklerle nasıl bağlantılı olduğunu açıklamalıdır. Bu bağlamda, Chalmers'a göre bilincin nöral korelasyonları üzerine yapılan çalışmalar, bizi birinci şahıs ve üçüncü şahıs arasındaki ayrımı birleştiren temel yasa ve ilkelere yaklaştırma misyonlarıyla oldukça değerlidir.

Chalmers'ın kuramının düalist olarak adlandırılmasına yol açan ise onun fenomenal bilinçte deneyimlenen herhangi bir bilginin, beyinde bilişsel olarak da temsil edileceğini söylemesi ve öznel, niteliksel deneyimin kendisinin fiziksel dünyanın bir parçası olmadığını düşünmesidir (Revonsuo, 2017: 283). Chalmers, hem öznel deneyimden gelen doğrudan bilginin hem de fiziksel bilginin varlığını inkâr etmez, ancak ikisi arasındaki bağıntının araştırılması gerektiğini savunurken bu ikisinin asla birbirine indirgenemeyeceği fikrine vurgu yapar. Chalmers'ın düalizmi Descartes'ın düalizminden çok önemli farklarla ayrılır. Chalmers, zihinsel bir töz olduğunu zorunlu bir koşul olarak olarak ortaya koymaz. Chalmers, "Gerçekliğin zihinsel özelliklerini fiziksel varlıkların (hayvanların) özellikleri olarak görür. Hayvanların zihinsel özelliklerinin fiziksel özellikleriyle yasaya dayalı bir ilişki içinde olduklarını düşünür." (Zeman, 2006: 389). Dolayısıyla Chalmers'ın teorisi töz düalizmine değil, doğalcı nitelik düalizmine örnek teşkil eder.

Bunun yanı sıra, Chalmers karmaşık bilgi işlemenin olduğu her yerde, karmaşık yapıda bir bilincin olduğunu ve bilgi işleme türleri basitleştikçe daha basit bilinç türlerinin ortaya çıktığını söyleyerek var olan her şeyin bir düzeyde bilince sahip olduğunu söyleyen panpsişist görüşe de yakın durmaktadır (Chalmers, 2017: 58). Chalmers, bilinci bilgi işleme ile özdeşleştirdiği için basit hesaplama sistemlerine sahip makinelere de bilinçli olma imkânı verdiği söylenerek eleştirilmektedir. O ise kuramının davranış, işlev ve bilincin bağlantılı olduğu fikrine dayandığından kendisini geniş anlamda bir işlevselci olarak tanımlayabileceğini ancak "Bilinç, işleyiştir." şeklinde bir tutumu asla benimsemediğini çünkü öznel deneyimin bilgisini de hesaba kattığını belirterek, bu eleştiriyi kabul etmemektedir (Chalmers, 2017: 61). Doğalcı düalizm Chalmers tarafından şöyle tanımlanmaktadır:

> Bir kimse, bilincin işlevsel organizasyondan doğduğuna ama işlevsel bir durum olmadığına inanabilir. Benim savunduğum görüş de böyle bir biçime sahip. Bunu indirgemeci olmayan işlevselcilik diye de adlandırabiliriz. Bu görüş, işlevselcilik ile nitelik ikiciliğini birleştirmenin bir yolu olarak görülebilir (Chalmers, 1996: 249).

Chalmers'a göre, öznel deneyim ve qualia, fiziksel özelliklerden açıkça farklıdır. Bu sebeple bilinç salt fiziksel bir fenomen olarak tanımlanamaz. Bu yaklaşımı Chalmers'ı maddeci işlevselcilerden farklılaştırmakta ve yeni bir düalist form oluşturmasına yol açmaktadır.

3.4. Üst Düzey Temsil Kuramları[37]

Bilinç fenomeni üzerine yapılan araştırmaların yoğunlukla üzerinde durduğu problemlerden birisi de bilinçli bir durumu bilinçli hâle getiren unsurların neler olduğu, bir başka deyişle de bilinçli durumlar ile bilinçsiz durumlar arasındaki farkın kaynaklarıdır. Temsil kuramları da bilince yönelik bu sorunsalın çözümüne yönelik inşa edilmişlerdir. Dolayısıyla temsil kuramlarının ele aldığı bilinç, canlı bilinci değil durum bilincidir. Temsil kuramları bilinçli durumlar ile bilinçsiz durumlar arasındaki fark üzerine çalışırken yalnızca bilincin nöral korelasyonları üzerinden ilerlemezler. Üst düzey temsil kuramları, bilinci "zihinsel" terimlerle açıklamaya çalıştıklarından, indirgemeci nörofizyolojik yaklaşımlardan farklılaşmaktadırlar (Gennaro, 2004: 2).

Üst düzey temsil kuramlarına göre, kendileri dışındaki şeyler hakkında yönelimsel içeriğe sahip olan durumlar veya dış dünyadaki şeyleri temsil eden durumlar, zihinseldir. Temsilî durumlar ya da zihin durumları bilinçli veya bilinçsiz olabilirler. Temsilci yaklaşıma göre bilinç, alt düzey bir durum ile üst düzey bir durum arasındaki ilişkiden ortaya çıkmaktadır. Alt düzey bir durum hakkındaki üst düzey bir düşünce, algı ya da farkındalık bilinçsiz zihin durumunun, bilinçli duruma dönüşmesine imkân sağlamaktadır. Saf acı ya da susuzluk hissi alt düzey bir durumdur. Bu durumların bilinçli hâle gelmesini sağlayan acı ya da susuzluk hissi hakkında üst düzey bir düşüncenin, algının ya da farkındalığın (canım acıyor, susadım vb.) varlığıdır. Üst düzey temsil yaklaşımında bilinçli durumun tanımı, öznesinin o durumda olduğunun farkında olduğu, algıladığı ya da hakkında düşünceye sahip olduğu durum olarak yapılır. Öznenin hiçbir şekilde farkında olmadığı, algılamadığı ya da düşünce üretmediği zihin durumu, bilinçsiz bir durumdur. Burada kullanılan farkında olmak, algılamak ve düşünceye sahip olmak kavramları üst düzey temsil teorilerinin üst düzey düşünce teorisi, üst düzey algı teorisi ve üst düzey farkındalık teorisi şeklinde ayrışmasına yol açar. Odaklandıkları kavramlar çeşitlilik göstermesine rağmen, üst düzey temsil kuramlarının tamamında ilk-düzey bir his ya da duyum bilinçli zihin durumu için yeterli görülmez ve bilinçlilik için üst düzey bir durumun gerekliliği varsayılır. Üst düzey temsil kuramları,

[37] *Higher-Order Representation Theories*

bilinçli durumların nasıl bilincinde olduğumuzu açıklama girişiminde bulunurken bilinçli bir durumda olmayı, birinin bilinçli bir durumun içerisinde olduğunun bilincinde olması olarak tanımlamaktadır (Rosenthal, 2004: 41). Buna göre bir şeyi gördüğünün bilincinde olmak, o şeyi gördüğü düşüncesine sahip olmak demektir; bilinç, düşünce bilinçsiz duyumu aydınlattığında ortaya çıkar (Zeman, 2006: 360).

Üst düzey temsil kuramlarına göre, fenomenal bilinç içerikleri onları seçen, kavramsallaştıran ve içebakışsal bir biçimde ifade eden yukarıdan aşağıya dikkat mekanizmaları ile refleksif bilincin nesneleri olarak ele alınana kadar bilinçli zihin durumlarına dönüşmezler ve bilinçdışı olarak kalırlar (Revonsuo, 2017: 286). Dolayısıyla fenomenal içeriğin de bilinçsiz olma ihtimali söz konusudur.

Temsil kuramlarında bilinç, zihnin içerisinde olup bitenlerle ilgili içsel bir algı olarak ele alınır. Bilincin nöral ve bilişsel korelasyonları temsilci kuramın kapsama alanına dâhil değildir. Dolayısıyla bu kuram teorik olarak açıklayıcı gibi görünse de deneysel olarak test edilebilme imkânı yoktur. Üst düzey bilinç kuramları varlık bilinciyle değil de zihin durumlarının bilinciyle ilgilendiklerinden, bilincin zor problemi olarak kavramsallaştırılmış olan fenomenal bilincin öznel ve niteliksel özelliklerini açıklamaya yönelik sorunla yüz yüze gelmezler. Fenomenal bilincin doğası üst düzey temsilci kuramlarda merkezi bir araştırma konusu olmadığından, bilincin zor problemini açıklamadıklarına yönelik bir eleştirinin muhatabı konumunda da değildirler.

Üst düzey temsil kuramlarının önemli bir açmazı, üst düzey düşünce oluşturma yetisi olmayan bebek ve hayvanları bilinçsiz varlıklar olarak kabul etmek zorunda olmasıdır. Hayvanlarda ve bebeklerde özbilincin varlığı tartışılan bir konudur. Yeni doğan bir bebek kendisini çevresinden ayırt edebilecek özbilince sahip değildir, bedensel uzuvlarının kendisine ait olduğunun farkında değildir. Bir başka bebeğin ağlamasını duyduğunda ağlar, aniden kolunu oynattığında korkar. Ancak bu durumlar bebeğin bilinçli olmadığını göstermez zira karnı acıktığında, canı yandığında bunlara ağlayarak tepki verebilir, annesinin sesiyle sakinleşebilir. Üst düzey temsil kuramına göre açlık düşüncesine sahip olmayan bu bebek bilinçsiz bir zombi olacaktır. Birtakım duyumlara sahip olsa bile hiçbir bilinçli hisse sahip olmayacak/bir şey hissetmeyecektir ya da bilinçli ya da bilinçsiz hiçbir duyuma sahip olmayacaktır. Bu iki olasılık da bebekler hakkındaki test-edilebilir ve gözlemlenebilir verilerle uyuşmamaktadır.

Bu kuramlara getirilen ciddi eleştirilerden bir tanesi de yanlış-temsil ihtimali üzerinedir. Üst düzey durumlar ile alt düzey durumlar arasındaki temsil ilişkisi üzerine kurulmuş olan bu temsil teorilerinin, alt düzey zihin durumunun kırmızı bir algı kaydettiği ancak üst düzey durumun yanlış nöral aktivitelerden dolayı yeşil algısını ürettiği durumlar hakkında yetersiz bir açıklama modeline sahip olduğu iddia edilmektedir (Gennoro, 2004: 8). Eğer son deneyimin kırmızılık olacağını söylerlerse, bu üst düzey durumların deneyimlerin niteliksel karakterine etki etmediğini gösterecektir. Tam tersi, eğer son deneyimin yeşillik olacağını söylerlerse de üst düzey durumların, alt düzey durumları temsil ettiğini yanlışlamış olacaklardır. Dolayısıyla yanlış ya da hatalı temsil itirazı üst düzey temsil teorilerini zora sokmaktadır.

3.5. Güçlü veya Dışsalcı Temsilcilik[38]: Dretske Örneği

Bilincin doğası ve yeri ile ilgili kuramları sınıflandırırken içselcilik ve dışsalcılık olmak üzere iki geniş bakış açısı mevcuttur. İçselciliğe göre bilinç, bedenin (beynin) sınırları içerisinde meydana gelen durumlar ve süreçlere özdeş ya da onlarla ilintiliyken; dışsalcılığa göre bilinç, kafanın hatta beynin dışında gerçekleşen süreçlere bütün olarak ya da kısmen dayanmaktadır (Manzotti, 2007: 179). Bilince içselci yaklaşımın karşı karşıya kaldığı en büyük sorun, bilinci beyne ya da kafatasına tamamen yerleşik bir şekilde konumlandırırken zihnin dış dünyada yer alan olay ya da durumları nasıl temsil ettiğine dair bir açıklama getirememesidir. Temsilciliğin ortaya çıkışı da içselciliğin yol açtığı bu gediği kapatma gayesine dayanmaktadır.

Bilinçli zihin durumlarının, diğer zihin durumlarından farklılaşmasını sağlayan şey o durumlarla ilgili sahip olunan "öznel hissi" oluşturan fenomenal içeriktir. Temsilci kuramların dayandıkları temel, tüm bilinçli durumların aslında temsilî durumlar olduğu ve fenomenal içeriklerin de bu sebeple temsilî içeriklere indirgenebileceği veya fenomenal içeriklerin temsilî içerikler yoluyla tamamen açıklanabileceğidir (Revonsuo, 2017: 288).

Temsilci kuramlar genellikle saf temsilcilik, zayıf temsilcilik ve güçlü temsilcilik olarak gruplandırılırlar. Saf temsilcilik, temsilin tek başına duyusal nitelik için yeterli olduğunu savunur. Zayıf temsilcilik, niteliksel durumların zorunlu olarak temsilî içeriğe sahip

[38] *Strong or Externalist Representationalism.*

olduğunu söylerken duyusal niteliklerin ontolojik olarak fiziksel olmayan özelliklere sahip olabileceğine açık kapı bırakırlar. Güçlü temsilcilik olarak sınıflanan dışsalcı temsilcilik ise yine temsilin, duyusal nitelik için yeterli olduğunu söyler ancak temsil edilen şeyin işlevsel ya da fiziksel olarak açıklanamayan bir özelliğinin olabileceğini kabul etmez.

Güçlü temsilcilikte bilinç içeriklerinin ayırt edici özelliği, Searle'ün bilinç kuramında olduğu gibi yönelimsel olmalarıdır. Bilinç içerikleri zorunlu olarak kendilerinin dışında bir şeye gönderme yaparlar. Deneyim, her zaman bir şeyin deneyimidir. Dretske'ye göre ise yalnızca deneyim değil deneyimsel -fenomenal- nitelikler de bir şeye yöneliktir. Onun geliştirdiği güçlü temsilcilik kuramında, yalnızca deneyimler değil aynı zamanda deneyimlerin fenomenal karakteri de -o deneyime sahip olmaklık- temsilîdir (Dretske, 2003: 67). Buradan hareketle Dretske, qualiayı, deneyimde bulunma durumunda birinci şahıs bakış açısıyla bilinebilen ve o deneyimi, diğer deneyimlerden ayırmaya yarayan nitelikler olarak tanımlarken qualianın nesnel fiziksel özelliklerin alt kümesi olduğunu söyler. Qualia, nesnenin duyusal temsil nitelikleridir. Bu da demektir ki bir başkasının (bir hayvanın) sahip olduğu fenomenal durumlar hakkındaki sorular, aslında o kişinin (ya da o hayvanın) temsilî durumları hakkındaki sorulardır; bu durumların işaret ettiği işlevin sahip olduğu özellikler hakkındaki sorulardır (Dretske, 1995: 73)

Temsilci yaklaşımda, beyindeki bilinçli durumlar ile kitapta yazılı kelimeler arasında bir analoji kurulur. Kitapta yazılı olan "kırmızı elma" kelimesinin kırmızı olmadan ve elmaya hiç benzemeden kırmızı elmayı temsil etmesi, kırmızı elma temsilinin nesneleri veya içeriklerinin fiziksel dünyadaki kırmızı elma nesnesiyle aynı olmasındandır. Beyindeki nöral aktivite ve işlemler de deneyimin kendisine hiç benzemese de onun bir temsilidir. Duyusal deneyimin kendisi ele alındığında onun yalnızca temsil özellikleri bulunur, temsilden fazla bir gerçeklik söz konusu değildir. Elimize kaynar su dökülüp yandığında bilinçli deneyimimizde bulabileceğimiz tek şey, elimizde gerçeklesen olaya ait özelliklerdir.

Bu yaklaşıma göre temsiller, temsilin taşıyıcısı ve içeriği olmak üzere iki ayrı unsurdan meydana gelir; temsilin taşıyıcıları kafanın içinde, içerikleri kafanın dışındadır (Revonsuo, 2017: 290). Bu bağlamda, temsilî bir içerik olarak öznel ve niteliksel deneyim yahut fenomenal içerik de beynin dışında bulunmalıdır. Dışsalcı temsilcilik bu yönüyle fenomenal bilincin varlığını reddetmemektedir. Ancak bu yaklaşımda, fenomenal bilincin içerikleri zihinde değil dış dünyada bulunmaktadır. Dolayısıyla öznel deneyim ve qualia zihinden bağımsız bir konumda, dış dünyadaki fiziksel özellikler arasında yer almaktadır. Bu

yaklaşımın dışsalcı olarak adlandırılması da fenomenal deneyimi, deneyimleyenin dışarısında konumlandırmasından ileri gelmektedir. Zihin, öznel deneyim hususunda, qualiayı temsil aracılığıyla fark etmekten öte bir işleve sahip değildir. Deneyimin kendisinin zihin üzerinde hiçbir kurucu ve şekillendirici etkisi yoktur. Sonuç olarak, temsilci yaklaşım deneyimin kendisinin doğasını açıklamaktan uzak olup deneyimlerin temsilleri üzerine kurulu olduğundan deneyimleyenin doğasına dair de bir açıklama getirmemektedir.

3.6. Duyu-Motor Kuramı[39]: O'Regan Örneği

Çoklu taslaklar kuramıyla benzer bir şekilde, Duyu-Motor Kuramı da içsel fenomenal deneyim olarak adlandırılan fenomenal bilinci yok saymaktadır. Bu kuramda bilinç, dünyayla kurulan duyu-motor etkileşimine indirgenmektedir (Noë, 2007: 510).

Materyalist kuramların çoğundan farklı olarak, bilinç beynimize değil dünyayla davranışsal etkileşimlerimize veya ilişkilerimize dayanır ve bu kurama göre, beynin bilince nasıl sebep olduğunu açıklamaya gerek yoktur (Revonsuo, 2016: 276). Duyu-motor kuramını ilk ortaya atan isimlerden olan O'Regan'a göre, "Birisi kolunuza vurduğunda neden kol-ağrısı hissedersiniz de bacak ağrısı hissetmezsiniz, neden kırmızı elma kırmızı görünür de yeşil görünmez ya da neden bir şey işitsel deneyime neden olur da görsel deneyime neden olmaz?" gibi sorular gerçekten de bilincin zor problemidir ve duyuların niçin başka şekilde değil de oldukları şekilde farklılaştıklarına dair soruların nörofizyolojik terimlerle ifade edilebilecek bir cevabı yoktur (O'Regan, 2012: 124). Bu sorunlar aslında, neden duyuların his içerdiğine dair çok daha temel bir soruyu da içermektedir. Beyindeki süreçlerin birçoğu duyusal reseptörlere sahip olmalarına rağmen, beş temel duyumuz haricindekiler hisse neden olmazlar. O'Regan beyindeki fiziksel süreçler ile onların ilişkili olduğu duyusal deneyimler arasındaki bu açıklanamaz boşluğun canlılığın açıklanma sürecini andırdığını iddia eder. Canlılığı açıklamak için bugün nasıl can veren ruh gibi fazladan bir varlığa ihtiyaç duymuyorsak, bilinç üzerine konuştuğumuz bu açıklama boşluğunun da kapanacağını ileri sürer. Ona göre, hislere nasıl sahip olduğumuz sorunsalı her şeyden önce kategorik bir hatadan ortaya çıkmaktadır. Nasıl yaşam dediğimiz şeyi, bugün bir çevreyle etkileşimde bulunma şekli olarak anlıyorsak, hissetmeyi de çevreyle etkileşimde bulunmanın bir başka

[39] *Sensorimotor Theory*

şekli olarak görmeliyiz. Hisler fiziksel bir mekanizma tarafından üretilen şeyler değildir; bu nedenle beynin içinde hislere neden olan bir şeyi aramak hatalıdır:

Bir süngere baskı uyguladığımda, yumuşaklık hissinin üretildiği yer neresidir? O'Regan'a göre bu, açıkça yanlış bir sorudur.

> Yumuşaklığın gerçekte hiçbir yerde üretilen bir şey olmadığını biliriz. Aslında, yumuşaklık üretilebilen türden bir şey de değildir. Süngerin yumuşaklığı, parmağımın basıncı altında süngerin ezilmesindedir. Ben yumuşaklığı, belirli bir duyu-motor kuralı –süngere bastırdığımda, onun ezildiğini tanımlayan bir kural- etkin olduğunda hissederim. Hisse karşı duyu-motor yaklaşımına göre, tüm duyusal modüllerde, hissin niteliği çevreyle olan duyu-motor etkileşimi kurallarıyla oluşur (O'Regan, 2012: 125).

Duyu-motor kuramı, bu bakış açısı ile hissi ve bilinci, beynin dışında etkileşimin gerçekleştiği yerde konumlandırarak, fiziksel süreçler ile öznel hisler arasındaki açıklama boşluğunu kapattığını iddia etmektedir. Bu yaklaşım, bir başka bağlamda deneyimi, deneyimleyen ve deneyimlenen ögelerden bağımsız olarak tek başına ele almaktadır. Buna göre, deneyimin kendisi sosyal bir unsur olarak, organizmanın dışında gerçekleşmektedir.

Ancak dünyayla hiçbir motor etkileşimin bulunmadığı fenomenal deneyimler ve rüyalar bu kurama antitez oluşturmaktadır çünkü uyku esnasında gerçekleşen canlı deneyimlerin varlığı ile ilgili bilimsel kanıtlar mevcuttur. Dolayısıyla duyu-motor kuramının bu konuyla tek başa çıkma yolu rüyaların bilinçli deneyimler olduğunu reddetmek olacaktır. Ancak uyku üzerine yapılan çalışmaların deneysel verileri rüyalar ve uyanık deneyimlerin öznel niteliklerinin birbirlerinden farklı düzeyde olmadığını göstermektedir.

Duyu-motor kuramı, makine bilinci araştırmaları tarafından oldukça ilgi çekici bulunmuştur. Zira bilincin, öznenin dünyayla davranışsal etkileşimine indirgenmesi, makinelerin çevreleriyle ve dünyayla yeterince etkileşimde bulunabilecek düzeyde programlanmasıyla bilinçli makineler olarak inşa edilmiş olacaklarını ima etmektedir. Duyu-motor etkileşimi ile çevresiyle ilişki kurabilen her makine tanımı gereği bilinçli olacaktır (Revonsuo, 2016: 278). Ancak bu, yeni bir kategori olarak makine bilincinin varlığından bahsetmek için yeterince güçlü bir gösterge oluşturmamaktadır. Para atıp, kolanın yer aldığı rakamları tuşladığımızda bize kola veren bir makine de çevresiyle ilişki kurmuş olmaktadır. Ancak bu ilişki türü, o makinenin bilinci olduğunu iddia edeceğimiz karmaşıklık düzeyinde değildir. Bu eleştiriye cevap vermek adına, O'Regan bilinçli olabilecek bir mekanizmanın belli bilişsel özelliklere sahip olması gerektiğini öne sürer ve bu bilişsel özellikleri, mekanizmanın gelişmişliğine

göre seviyelere ayırır. Buna göre, ilk seviye, kendisini çevresinden ve diğer mekanizmalardan ayırabilme yetisi; ikinci seviye görev ve planlarını bilmek anlamında kendini bilme yetisi ve son olarak üçüncü seviye ise kendini bildiğini bilme yetisidir (O'Regan, 2012: 121). O'Regan'a göre bunların hepsi sadece bilişsel nitelikteki yetilerdir ve bir robotun bu yetilere sahip olmasının önünde hiçbir teorik engel yoktur. Fakat bize göre, kendini bildiğini bilme yetisi adını verdiği özbilinç, kolayca bilişsel düzeye çekilecek bir yeti değildir. Bilincin kendisi duyu-motor etkileşimleri ile açıklayarak beynin dışında konumlandıran O'Regan'ın bilinçli olduğunun bilincinde olmak manasına gelen özbilinci yalnızca bilişsel bir süreç olarak ele alması çelişkilidir. Zira eğer özbilinci bilişsel bir süreç olarak açıklayabiliyorsak, bilinci bilişsel olarak açıklayamamazı gerektirecek bir nitelik kalmamalıdır. Bilincin bilişsel olarak açıklanmasının önündeki engeli O'Regan'ın hisler olarak gösterdiğini göz önüne aldığımızda, özbilinci hislerden bağımsız olarak tanımlaması eksik bir tanım görünümü yaratmaktadır.

3.7. Nörofenomenoloji: Varela Örneği

Nörofenomenoloji, Husserl'in felsefi fenomenolojisinden nörobilime ve beyin görüntüleme yöntemlerine, somutlaşmış biliş yaklaşımından Budistlerin meditasyon pratiklerine uzanan yöntemsel olarak geniş bir havuza sahip olan bir bilinç kuramıdır.

Felsefi fenomenoloji zihnin tanımı ve doğasının belirlenmesi açısından içe-bakışı uygun bir yöntem olarak görmemiştir. Felsefi fenomenolojinin kurucusu sayılan Husserl'e göre zihinsel olanın ayırt edici özelliği yönelimselliktir. Yönelimsellik kavramı, zihin durumlarının kendilerinin ötesindeki bir şeylerin hakkındalığını veya o şeylere yönelmişliğini ifade eder.

Nörofenomenoloji ise en genel tanımıyla, felsefi fenomenoloji ile nörobilim ve bilişsel bilimler çalışmalarını birleştirme gayretindedir. Bu amaca yönelik düşünsel çabası olan en önemli isimlerden birisi Francisco Varela'dır. Varela'ya göre, birinci şahıs bakış açısı ile üçüncü şahıs bakış açısı arasındaki gediği kapatacak "kuramsal bir çözüm" ya da "ekstra bir bileşen" yoktur (Varela, 1996: 330). Bilincin doğasına yönelik bir araştırmaların ihtiyacı olan şey titiz bir yönteme sahip olmaktır. Varela bu yöntemi, deneyimin kendisi ve bilişsel bilimlerin ele aldığı bilinç arasındaki bağlantıları araştırmak olarak belirler ve bunu nörofenomenoloji olarak tanımlar.

Varela'nın fenomenolojik yaklaşımının en temel kabulü bilincin indirgenemezliğidir. Bilince yönelik araştırma yapmanın tek somut olasılığı, deneyimin doğasına yönelik sistemli bir incelemedir (Varela, 1996: 335). Varela fenomenolojiyi bir tür düşünme biçimi olarak görür ve fenomenolojinin bilişsel bilimler ışığında deneyimi açıklama yönelik faydalı bir araç olacağını savunur. Varela'nın kullandığı fenomenolojik yöntem "içebakış" yönteminden farklıdır çünkü içebakış öznenin ya da "ben"in varlığını zorunlu koşul olarak görürken Varela özne-nesne ikiliği tartışmalarından kaçınmak ister. Onun yapmak istediği özne ve nesnenin temel korelasyonlarını açığa çıkarmaktır (Varela, 1996: 339). Varela'ya göre deneyim her ne kadar kişisel olsa da kişiye özel değildir çünkü kendi bilincimiz, diğerlerinin bilincine sıkı şekilde bağlıdır. Öznel ve nesnel olanın zıtlıkları arasındaki kapanmaz boşluğa odaklanmak bilinç çalışmaları için hiçbir fayda sağlamaz; yapılması gereken bu boşluğu pratikte görmezden gelmektir. Bu bağlamda, nörofenomenolojinin amacı da bilincin madde ve beyinden nasıl çıktığına dair bir keşifte bulunmaya çalışmak yerine, indirgenemez bu iki fenomen arasında anlamlı köprüler kurmaktır (Varela, 1996: 340). Varela'nın bilince nörofenomenolojik yaklaşımı Doğu'ya ait düşünce pratiklerini ve Budist derin düşünme pratiklerini temel alır. Ona göre, deneyim ile fiziksel dünya arasındaki ilişkinin doğası kadar, bilimlerin yöntem ve alışkanlıklarını değiştirmek de zor problemdir. Bilimsel bakış açısı, "yaşanan deneyim anını" merkeze koyan düşünce pratiklerini dikkate almadıkça deneyimin bilimsel bir araştırma konusu olması mümkün değildir.

Bilince nörofenomenolojik yaklaşım yalnızca Varela ile sınırlı değildir. Nörofenomenoloji birçok farklı bakış açısından ele alınmış bir teoridir. Bu yaklaşımların hemen hepsinde ortak olan bir tez, fenomenolojinin de ana iskeletini oluşturan düşüncelerden birisi olan bilinç de dâhil olmak üzere tüm zihin durumlarının belirli bir yapıya sahip olduğudur: "Zihinsel durumlar, kendilerine yol açan nesneye yönelmiş bir zihinsel eylem içerir ve bu, bilincin eylem-nesne şeklindeki iki yönlü yapısıdır." (Revonsuo, 2017: 292). Fenomenolojik yaklaşımda zihin, biyolojik, hareket eden, eyleyen bir bedende bedenlenmiştir ya da somutlaşmıştır (bedenlenmiş biliş). Somutlaşmış bilinç sayesinde kendimize bağlı olduğumuzu öznel bir şekilde deneyimleriz. Bu deneyim, Descartes'ın bedenden ayrı bir zihin deneyiminin aksine kendimizin ve zihnimizin bedenimizle beraber deneyimidir.

Nörofenomenolojideki somutlaşmış biliş için kurulan argüman şu adımlarla takip edilebilir:

- Yaşadığım deneyimler benimdir. Bu deneyimleri yaşayan bir "ben" vardır. Dolayısıyla deneyimler zorunlu olarak bir kişiye, bir benliğe veya bir ben'e atıfta bulunur.

- Deneyimler öz-gönderimseldir.
- Gönderimde bulundukları benlik, bedene sahip -somutlaşmış- bir benliktir. Bu nedenle düşünüm öncesi özbilinç, bedensel farkındalığa yapılan bir gönderimi içerir.
- Bedensel deneyim en başından itibaren bilince yerleşiktir. Öz-farkındalığın bu temel biçimi olmaksızın bilinç mümkün değildir (Revonsuo, 2017: 293).

Bu yaklaşımda, bilinçli öznenin varlığı, somutlaşmış bir varlık veya somutlaşmış yönelimsellik en baştan zorunlu koşul olarak öne sürülür. Buna göre, bilincin ilkel düzeyden en üst düzeyine kadar her seviyesinde bedene bağlılık mevcuttur.

Nörofenomenolojinin bilimsel bilinç araştırmaları için umut vadeden yanı deneysel bilinç çalışmalarında birinci şahıs yaklaşımının kullanılmasına imkân tanıması ve bunu olanaklı kılan teorik altyapıyı hazırlamaya çalışmasıdır. Bilimsel çoğulculuk olarak değerlendirebilecek bu bakış açısı bilimsel yöntemlerin nesnellik ölçütüyle koydukları sınırları kaldırmayı amaçlamaktadır.

3.8. Refleksif Monizm: Velmans Örneği

Max Velmans'ın refleksif (düşünümsel) monizm olarak adlandırdığı bilinç yaklaşımında, birinci ve üçüncü şahıs bakış açıları her zaman birbirini tamamlarlar ve dolayısıyla birbirlerine indirgenemezler. Zihnin doğasını anlamak için iki perspektif de gereklidir. Velmans bu gerekçeyle zihnin doğasının sadece fiziksel olmadığı gibi sadece bilinçli deneyim de olmadığını söyleye ve zihnin doğasının aynı anda hem fiziksel hem de bilinçli deneyim olduğunu iddia eder. Bu iddiayı kavramsallaştırmak adına da zihnin doğasını psikofiziksel olarak adlandırır (Velmans, 2009b: 310).

Max Velmans'a göre, bilinci tanımlamak için girişilen çabanın en önemli aşaması doğru yerden başlamaktır ve başlanılması gereken yer gündelik deneyimdir. Deneyim, uzayda yayılım gösteren üç boyutlu fenomenal bir dünyadır (Velmans, 2018: 280). İnsanın öznel dünyası ile deneyimlediği fiziksel dünya bir ve aynı şeydir. Velmans'a göre,

> Bizim algıladığımız şekliyle dünya, dünyada gerçek enerjiyle etkileşime geçen beynimizin içinde meydana gelen süreçlerin ürettiği bir temsildir. Bu temsil, içinde gömülü olduğumuz gerçekliğin doğasını modeller (Velmans, 2018: 28).

Çevremizde deneyimlediğimiz dünya ile bu dünyaya ilişkin deneyimimiz arasında hiçbir boşluk yoktur, bu ikisi doğaları gereği fenomenolojik açıdan hiçbir fark barındırmazlar. Bu bizim deneyimlediğimiz fiziksel dünya ile fiziğin tasvir ettiği dünyanın aynı şeyler olduğu anlamına gelmez.

Max Velmans'ın bu doğrultuda geliştirdiği bilinç kuramının merkezi amacı bilincin yerini anlamaktır. Deneyimlerimizin nitelikleri esasında zihin veya beyin tarafından üretiliyor olmalarına rağmen biz bu nitelikleri zihin veya beyinde konumlanmış olarak değil fiziksel dünyaya yerleşmiş olarak deneyimleriz. Deneyimlere ait özelliklerin mekânsal olarak dışsal nesnelerde mevcut bulunduğu şeklinde bir yanılsama içerisindeyizdir. Ağrı, acı, kaşınma, gıdıklanma gibi durumlar bedenimizde, renkler, sesler, kokular vs. dışarıdaki nesnelerdedir. Peki, bunların fiziksel olandan çıkıp da zihinsel bir deneyime dönüşmeleri ve zihnin içerisinde yayılmaları nasıl mümkün olmaktadır? Nitelikleri, kaynaklandıkları ya da ait oldukları yerden, deneyimlendikleri yere taşıma bir tür yansıtma ile mümkün olmaktadır. Duyusal işlemin sonuçlarının, dışsal dünyayla birleştirilmesi için bir şekilde beyinden geriye doğru yansıtılması gereklidir (Velmans, 2009b: 296).

Velmans'a göre bilincin yerini tespit etmenin tek yolu, bilincin bağlı olduğu durumdaki fenomenolojisinin incelenmesidir. Örneğin elinize kaynar su döküldüğünde, acıyı elinizde deneyimlersiniz. Elinizde deneyimlenen acının bilinçli deneyimine işaret eden yer de suyun döküldüğü yer olmalıdır. Dolayısıyla bu kuramın iddiası şudur ki; deneyimler yalnızca onları deneyimlediğimiz yerde bulunurlar, beyinde veya başka bir yerde ikinci bir bilinçli deneyimi aramak gereksizdir (Revonsuo, 2017: 297). Velmans, bir beyin durumuna ait zihinsel bir modelde eldeki acının kodlandığını inkâr etmez ancak ona göre, deneyim ve zihinsel model aynı şey değillerdir. Ancak bu yaklaşım, bilincin doğasını açıklama bağlamında muğlak düşüncelere yer vermektedir. Zira acının nöral bağıntı ve nedenleri eğer beyinde bulunuyorsa deneyimin de beyne ait olması gerekir gibi durmaktadır. Neden ve ilişkileri beyinde gerçekleşen deneyimin, ele ait olması fazladan bir açıklama gerektirmektedir. Velmans bu açıklama boşluğunu "algısal yansıtma" adını verdiği psikolojik bir süreçle kapatmaya çalışır (Revonsuo, 2017: 297). Algısal yansıtma süreci, deneyimin beynin ötesindeki fenomenal uzaydan deneyimin gerçekleştiği yere yansıtılması olarak ele alınabilir. Velmans'ın teorisinde bu yansıtma, fiziksel bir mekanizmaya değil psikolojik bir sürece tekabül eder.

Velmans'a göre, psikolojik bir açı olan birinci şahıs bakış açısı ile fiziksel olan üçüncü şahıs bakış açısı birbirlerine sıkı bir şekilde bağlıdırlar ve ayrık olarak düşünülmemelidirler. Hem felsefi hem de bilimsel literatürde öznel ve nesnel bakış açılarının birbirlerinden keskin bir şekilde ayrılmış olması Velmans'a göre hatalıdır. Bu iddiasını da belirli bir deneyimin nöral

bağıntılarını izole edebileceğimiz şekilde gelişmiş bir nörofizyolojiye dayanan bir düşünce deneyiyle temellendirmeye çalışır:

> Çok basit bir deneyim düşün, mesela önümüzde bir kayıt cihazı var. Sen kayıt cihazına bakan bir katılımcı ol, ben de beynini inceleyen bir araştırmacı. Benim senin beyninde gördüğüm şeylerle, senin dış uzayda gördüğün şey arasında nasıl bir ilişki var?... Sen kayıt cihazına bakarken gördüğün şeyleri anlatıyorsun, ben de senin beynine ve deneyiminin nöral bağıntılarına bakarken aslında kendi gördüğüm şeyi anlatıyorum. İlk bakışta, benim gördüğüm şeyin nesnel ontolojik konumunun senin gördüğün şeyinkinden daha yüksek olduğu pek açık değil. Bunun böyle olduğunu görebilmenin yollarından biri, rollerimizi değiştirebilmemizdir (Velmans, 2018: 282).

Roller değişip kayıt cihazına bakan, beyne, beyne bakan da kayıt cihazına bakmaya başladığında öznel bakış açısına sahip olan bir anda nesnel bir bakış açısına geçecek ve tam tersi şekilde nesnel bakış açısının sahibi de artık öznel deneyimlerini aktarıyor olacaktır. Velmans'a göre öznelden, nesnele böyle bir geçiş problemlidir ve bu problem aslında öznellik ve nesnellik arasındaki ilişkinin iyi analiz edilememesinden kaynaklıdır.

Velmans nesnelliğin dört şeklinden bahseder. İlki, özneler-arası doğrulanabilir gözlemler ile kurulan nesnellik, ikincisi, tarafsız ve doğrucu olmak anlamındaki nesnellik, üçüncüsü ise araştırma yöntemlerini açık ve detaylı bir şekilde ortaya koyarak test edilebilirliğine müsaade eden nesnelliktir. Dördüncü nesnellik ise hiçbir zaman elde edemeyeceğimiz gözlemciden bağımsız olma anlamındaki nesnelliktir, bir başka deyişle, gözlemcinin deneyimlerini içermeyecek bir gözlem yapmaktır. Bu bağlamda, bütünüyle nesnel bir bilim söz konusu olamaz. Velmans kendisinin bu bakış açısını eleştirel gerçekçilik olarak adlandırır (Velmans, 2018: 283).

Öznel ve nesnel bakış açıları insan bilincinde aynı anda bulunmaktadır. İnsan öznel bakış açısıyla hem fiziksel dünyanın tamamını deneyimleyebilir hem de bu fiziksel dünyayla birlik içerisindedir. Fiziksel evrenin çift görünüşlü doğası onun hem insan bilincinde kendi üzerine dönerek (refleksif) düşünmesini hem de insanın içinde kendi bilincine ulaşmasını sağlar. Varlığımızın tek yönü/görünüşü birbirimizden ayrı bedenlere sahip oluşumuz değildir, aynı zamanda herkes kendi bakış açısından evrene dair bir bakış açısına da sahiptir. Velmans'ın refleksif tekçilik adını verdiği tutumda, bütün evren, bütüne dair bir görüşe sahip olmak gibi bir yeteneğe sahip minik parçacıklardan oluşmuştur (Velmans, 2018: 284). Öznel deneyim ile nesnel beyin etkinlikleri aynı şeyin farklı iki görünüşüdür ve nerden baktığımıza bağlı olarak birbirlerinin yerine geçmektedirler.

Refleksif tekçilik kuramı bilincin ve bilinçli bir varlık olarak "ben"in deneyimlediği dünyanın gerçekliğine müsaade etmesi bağlamında geniş ve kullanışlı bir bakış açısı sunmaktadır. Bu kuramın vaatlerine göre, bir yandan öznelerin varlığına ya da yapıp etmelerine bağımlı olmaksızın fiziksel dünya varlığını sürdürmekte, bir yandan da öznenin kendisiyle varlığını sürdürdüğü birinci şahıs bakış açısı herhangi bir bilimsel gerçekliği reddetmeden de konumunu koruyabilmektedir.

3.9. Sanal Gerçeklik Kuramı- Benlik Tüneli: Metzinger Örneği

Felsefi bilinç kuramlarının çoğu bilincin yeriyle ilgili bir karmaşaya işaret eder. Bu karmaşanın ve paradoksal durumun kaynağı deneyimlerimizin, beyinlerimizin içinde bir yerlerde meydana geliyor olmalarını düşünmemize rağmen beynin belirli bir bölge ya da alanını deneyim merkezi olarak gösteremiyor oluşumuzdur. Bu sebeple öznel deneyimin varlığını inkâr etmeyen güncel bilinç kuramlarının çoğu, bilincin beynin dışında bir yerlerde gerçekleştiği varsayımı üzerine kuruludur. Sanal gerçeklik kuramcıları bu varsayımı beynin-dışı yanılsaması veya bir-dünyada-olma yanılsaması olarak adlandırır (Revonsuo, 2016: 299).

Kuramın temsilcilerinden Thomas Metzinger *The Ego Tunnel - The Science of the Mind and the Myth of the Self*[40]kitabının henüz girişinde amacının okuyucuyu kendilik diye bir şey olmadığına ikna etmek olduğunu söyler. Ona göre, birçoğumuzun inandığımızın aksine hiçbirimiz bir kendiliğe sahip değiliz. Kendilik ne beynimizde ne de metafizik düzlemde yer alan bir gerçeklik değildir (Metzinger, 2009: 1).

Metzinger'e göre filozoflar kendiliği fenomenal bilinç ile açıklarken bilincin birliği ve "şimdi"nin ya da "an"ın varlığını çıkış noktası olarak kullanırlar. Oysaki hem bilincin birliği hem de "şimdi", beyindeki nöral şebekenin ritmik çalışması ile oluşan bir yanılsamadır. Fenomen, bir görünüştür ve fenomenal "şimdi" de "şimdi"nin bir görünüşüdür (Metzinger, 2009: 38). Fakat gerçekte "şimdi" sadece bir illüzyondur ve gerçeklikle hiçbir bağlantısı yoktur.

Metzinger'e göre kendilik diye bir şey olmamasına rağmen *Deneyimlerimizin deneyimleyeni kimdir? Hayalleri hayal eden ya da hisleri hisseden kimdir?* gibi sorular hâlen merkezi sorulardır ve cevaplanmaya muhtaçtır. O bu soruları "benlik tüneli" yaklaşımıyla

[40] Benlik Tüneli- Zihnin Bilimi ve Kendilik Miti

yanıtlamaya çalışır. Sanal benlik modeline göre bilinçli beyin, fiziksel gerçekliğin içerisinden fenomenal, öznel bir tünel açar; öznel deneyimin beyindeki fenomenal tünelden akışı, beyni çevreleyen mevcut fiziksel gerçekliğin sadece dar ve seçilmiş olan bir yansımasıdır (Revonsuo, 2017: 301). Buna göre, öznelliğin açıklaması bilinç bulmacasının en çetrefilli bölümüdür. Bu bölümü tamamlamanın tek yolu da insanların içerisinde oldukları bilinç tünelinin onları nasıl benlik tüneline soktuğunu anlamaktan geçer. Fizikseli bilmenin birinci şahıs bakış açısı ve ikinci şahıs bakış açısı olmak üzere iki yolu vardır. Birinci şahıs bakış açısı, benlik tünelinin bilme şeklidir. Ontolojik olarak monist, epistemik olarak düalist olan bu tutumda benlik, sanal gerçekliğin bir parçasıdır (Metzinger, 2009: 105).

Bu kuramının temelleri, idealizmde de olduğu gibi nesnelerin fiziksel dünyasının gerçek bir dünya değil bir kandırmaca olduğu fikrine dayalıdır. Fiziksel dünya, beynimizin içerisinde gerçekleşen olaylarla kurgulanmış bir oyunun gerçek gibi görünen ama gerçek olmayan bir sonucudur. Buna göre, fiziksel dünya diye tabir ettiğimiz şey esasen beyin-dışı yanılsamasıdır ve deneyimlerimizle kurduğumuz ilişki fiziksel dünyadan değil kendi içsel fenomenal dünyamızdan kaynaklanmaktadır. Bu iddianın en yalın hâli şöyle ifade edilebilir: Beynin içerisinde sanal bir gerçeklik vardır. Dünyayı algılarken deneyimimizde doğrudan ilişki içerisinde olduğumuz şey fiziksel nesnelerin dışsal dünyası değil, beynin sanal gerçekliklerden meydana gelen bilincinde oluşturulan fenomenal temsilleri veya imgeleridir (Revonsuo, 2016: 300).

Thomas Metzinger'e göre, bilinç hem içeriden hem dışarıdan bilinebilen tek doğal fenomendir ve bilinci ayrıcalıklı ve zor problem hâline getiren de onun bu özelliğidir. Zira bizim bilmediğimiz şey; içerisi ve dışarısının nasıl bir ilişkide olduğu ve içeriden bilmek gibi bir söz söylediğimizde tam olarak ne kastettiğimizdir yani birinci şahıs bakış açısının anlamıdır (Metzinger, 2017: 184). Bilinçli deneyimin bireysel modellerde ve bireysel beyinlerde birinci şahıs bakış açısından meydana geldiğini söyleyen Metzinger'e göre birinci şahıs bakış açısı ya da benlik, "Beyinlerin dünyayı, merkezli bir dünya şeklinde ya da merkezinde benliğin bulunduğu bir dünya şeklinde betimlemek için kullandığı çok özel bir temsilî yapıdır." (Metzinger, 2017: 185).

Metzinger'in içsel benlik kuramı berimsel çabaların birinci şahıs bakış açısına yönelik felsefi analizinde kullanıldığı faydalı bir çerçeve niteliğindedir. Bu yaklaşım, zihnimizin nöral aktivite modelleri tarafından desteklenen, beden imgesi olarak adlandırılan bedenin mekânsal özelliklerinin içsel temsiline dayanan bir benlik modeli içermektedir (Reggia,

2013: 120). Beden imgesi esasında kısmen içkin de olsa uyarlanabilir sanal bir modeldir. Bu argümana göre, benliğin öznel bilinçli deneyimi, içalgısal, semantik ve diğer duyusal girdilerden dolayı beden imgesinin sabit ve sürekli oluşu ile bizim beden imgesinin, fiziksel bedenin yalnızca sanal bir temsili olduğunu fark edecek bilişsel süreçlere sahip olmamızdan doğmaktadır (Reggia, 2013: 120). Buna göre ise bedenin benlik modeli, kişinin fenomenal öznel durumlarına ve bilinçli deneyimlerine yol açmaktadır.

Sanal benlik kuramına göre, fiziksel nesneleri ve fiziksel dünyayı hiçbir zaman olduğu gibi deneyimleyemeyiz. Bizim deneyimlediklerimiz, onların beynimizde oluşturduğu fenomenal imgelerdir. Fenomenal deneyimin öznel ve niteliksel olmasına yol açan şey, deneyimin fiziksel nesnenin kendisi değil beyindeki imgesi olmasıdır. Beynimizdeki fenomenal imgeler şeffaftır. Şeffaflıkla kastedilen ise bizim onları temiz bir pencereden bakıyor gibi görüyor oluşumuz ve şeylerin kendilerini gördüğümüze inanmamızdır (Revonsuo, 2017: 302). Kendi benliğimiz de gerçekte buna benzer olarak, beyin tarafından üretilmiş fenomenal imgelerden bir tanesidir ancak biz onu deneyimlerimizin öznesi olarak görüp, gerçek benlik olarak kabul ederiz. Sanal benlik modeline göre, fenomenal deneyim ve benlik gerçek değildir. Bu anlamda, hepimizin kendi beynindeki fenomenal deneyim tüneli içerisinde oluşturduğu imgeler ile birlikte sanal gerçeklik dünyasında yaşayan sanal benlikler olduğumuzu iddia eder.

Bu bakış açısı, özne ve nesne arasındaki ilişkinin nasıl kurulabileceği sorusunu, diğer teorilerden farklı olarak öznelliği değil, nesnelliği yadsıyarak cevaplamaya çalışmaktadır. Ancak nesnelliğin inkârı aynı zamanda deneyimin de inkârına yol açmaktadır. Buradan hareketle de hem deneyimin hem de deneyimleyenin gerçekliği yok sayılmaktadır. Tüm benlik yaşanstısını illüzyon olarak ele alan bu yaklaşım sezgisel olarak kabulü güç bir yaklaşım olarak kalmakta ve makine bilincinin tek gerçek bilinç türü olarak ele alınmasına kapı aralamaktadır. Zira makineler, fiziksel dünya ile olduğu gibi etkileşime girecekler, kendi fenomenel tünellerinde sanal bir dünya yaratmayacaklardır.

4. BİR ARAŞTIRMA ALANI OLARAK YAPAY BİLİNÇ VE YAPAY BİLİNCE FARKLI YAKLAŞIMLAR

4.1 Yapay Bilinç Teorileri

Bundan önceki üç bölüm, yapay zeka ve bilinç kavramları için literatür araştırmasına ayrılmıştır. Bu çerçevede, her iki kavramın da tartışmalı yönlerini ortaya konulmaya çalışılarak, yapay bilincin araştırma konuları için kavramsal bir ön hazırlık yapılmıştır. Bu bölümde ise yapay bilincin tanımı ve yapay bilinç çalışmalarında dayanak kabul edilen yaklaşımlar üzerinde durularak, yapay bilincin epistemolojik, bilimsel, etik unsurları için bir temel oluşturmaya çalışılacaktır.

Yapay zeka çalışmaları, bilişsel bilim, nörobilim, robotik ve mantıkla sıkı sıkıya ilişki içerisindedir. Bu alanlardan her birisi yapay zekayı farklı yönlerden ele almakta ve ortak hedef olan, insanın bilişsel kapasitesini taklit için farklı yöntemler kullanmaktadır. Fakat hemen bütün yapay zeka araştırmaları, insanlar tarafından sergilenen akıllı davranışlar üzerinden çizdikleri çerçeve içerisinde ilerlemektedir. İnsanın çevresiyle ilişkisi ve eylemlerine şekil verişi, yapay zeka çalışmalarının temel ilham kaynağıdır. Ancak yapay zeka alanındaki mühendislik payının fazla olması, insana salt mekanik bir bakış açısıyla yaklaşılmasına ve insanın sadece eyleyen ya da davranışta bulunan bir varlıkmış gibi ele alınarak, zekanın göstergesinin ölçülebilir olan davranışta aranmasına yol açmıştır. Oysaki insanı farklılaştıran unsur, yaptıkları ve algıladıklarının bilincinde olmasıdır.

Bilinç kavramı hem günlük dilde hem de bilimsel literatürde değişik anlamlara karşılık gelecek şekilde kullanılmakta ve tanımı üzerinde uzlaşıya varılamamaktadır. Bilinç, belli olay veya süreçlerin farkında olma, kendimizin farkında olma (özbilinç, kendilik bilinci), uyanık ve uyku durumlarını birbirinden ayırma, ruhun dini hissiyatını kavrama, bilinçli ve bilinçsiz boyutlar arasındaki psikoanalitik ayrım, deneyimin öznel niteliklerini kavrama ve yönelimsel içeriğe sahip temsili durumları geliştiren yetenek vs. farklı anlamlara atıfta bulunmakta veya onların yerine kullanılmaktadır (Giovagnoli, 2013: 180). Bilinçli olmak ya da bilince sahip olmak ile zeka sahibi olmak arasında bir bağlantı olduğu açıktır. Ancak bilincin, zihinsel anlamda insana üstünlük sağlayan bir yeti olup olmadığı, insanın çevreyle etkileşiminde oynamış olduğu rol ve pratik sonuçları hususları uzun süre tartışma konusu olmuş ve özellikle materyalizmin hem bilim hem de felsefe camiasında mutlak hâkim olduğu yıllarda bilinç, bilimsel bir araştırma konusu olarak kabul görmemiştir. Bilinç yerine, problem çözme, bilgi, temsil, plan yapma, öğrenme ve dil işleme gibi sınırları ve tanımı daha

net olan zihinsel yeteneklere odaklanılmıştır. Bu bakış açısına göre, bilinç yeterince gelişmiş bir yapay zeka ile ulaşılabilecek bir ödül ya da kazanım olabileceği gibi biyolojik ve işlevsel süreçlerin yan ürünü olarak da değerlendirilebilir (Manzotti, 2013: 7).

Bilinci sadece faydasız bir epifenomen olarak gören bu yaklaşım, bilinci aynı zamanda insan zihnini diğer sistemlerin işlem mekanizmasından üstün ve ayrıcalıklıymış gibi gösteren aldatıcı bir özellik olarak kabul etmektedir. Hatta bilincin gereksizliğini vurgulayan bu yaklaşımın aşırı ucunda, bilinç tıpkı apandisist gibi, bizi hasta etmekten başka hiçbir işe yaramayan bir nitelik olarak adlandırılmaktadır (akt. Dehane, 2018: 120).

Zihin-beyin özdeşliği teorisini savunan ilk isimlerden olan Place, teorisini bilincin empirik çalışmasının yapılabilmesinin önündeki engelleri aşabilecek bir yol olarak tanımlarken onu izleyenler zihin-beden problemi önündeki en büyük engelin bilince atfedilen farklı ontolojik statü olduğunu vurgulamışlardır (Polak ve Marvan, 2018: 11). Fakat görünen odur ki, bütüncül bir bilinç araştırması, zihin-beden probleminin bir devamıdır ve yapılması gereken metafizik bir sorunsal olarak görülen zihin-beden etkileşimini bilimsel bilinç araştırmalarıyla bütünleşik hâle getirmektir. Beyin durumlarının nasıl olup da fenomenal durumların taşıyıcısı olduğu sorunsalı ancak bu bakış açısıyla çözüme kavuşturulabilir. Zira bilinçli deneyim ne beyin aktivitesiyle aynı şeydir ne de ondan metafizik olarak farklıdır.

Bilinç meselesi, 19. ve 20. yüzyıl boyunca bulanık ve öznel karakteri nedeniyle normal bilimin sınırları dışında kalarak, nesnel deneylere konu olmamıştır. Hatta bilişsel psikolojinin kurucusu olarak gösterilen George Miller, 1962 yılında yayınladığı *Psikoloji, Zihinsel Hayatın Bilimi*[41] isimli ders kitabında bilinç kelimesinin kullanımını yasaklamayı önermiştir:

> Bilinç, herkesin diline sakız olduğu için artık işe yaramaz bir kelimedir… Bilinç kelimesinin şu an anlaşılmaz birçok kullanımı yerine daha net ifadeler bulana kadar, bu kelimeyi belki de on ya da yirmi yıl yasaklamamız gerekir (akt. Deheane, 2018: 20).

Ancak, son 20 yıl içerisinde bilincin sadece gölge ya da yan bir fenomen olmaktan fazlası olduğu, beynin temel bileşenlerinden birisi olabileceği ve nöral korelasyonlarının dikkate

[41] *Psychology, the Science of Mental Life* (1962).

değer olduğu yönündeki yaklaşımların sayısında dikkate değer bir artış yaşanmış ve bilinç, bilimsel bir araştırma konusu olarak ele alınmaya başlamıştır.

Her bilim, bir açıklayan (açıklamayı yapan olaylar, durumlar, özellikler, yasalar vb. açıklamayı yapan) ve bir açıklanan (açıklanacak olan durumlar, olaylar, özellikler vs.) gerektirir. Bilinç bilimini farklı kılan ise bilincin hem açıklayan hem de açıklanan rolüne sahip olmasıdır (Chrisley, 2009: 53). Bu sebepledir ki, bilinç bilimi çalışılması en güç sorunsalları içeren bilim olarak adlandırılmakta ve farklı yöntemlerle çalışan alanlara ayrılmaktadır. Bilincin, bilimsel araştırmasının yürütüldüğü alanlardan birisi de yapay bilinç alanıdır.

Makine bilinci ya da yapay bilinç "Bilincin doğası nedir?" ve "Bir makine bilinçli olabilir mi?" soruları üzerine kurulmuş müstakil bir araştırma alanıdır. Bu alanın temel hedefi bilinci biyolojik olmayan bileşenlerle yeniden üretmektir. Ancak yapay bilinç çalışmalarını zorlaştıran ve kaçınılan bir alan hâline getiren de bilim camiası tarafından ortak ifadelerle tanımlanamayan bilincin, biyolojik olmayan bir eşleniğini üretme çabasıdır. Dolayısıyla yapay bilinç ile uğraşmak, parametreleri bilinen ve ölçülebilen bir şeyle uğraşmak gibi değildir (Manzotti ve Tagliasco, 2008: 107). Yapay bilinç, muamma bir fenomen olarak adlandırılan bilince yönelik teori ve modellerin formüle edilmesini ve bunların bilgisayarlar ile robotlarda uygulanmasını içeren disiplinlerarası bir alandır. Bilincin mahiyeti henüz kavranamamış bir zihinsel fenomen oluşu, yapay bilincin imkânına umutsuzlukla yaklaşılmasının ana sebebidir. Bilincin engellerle dolu bir yapısının olduğu ve bunları nasıl aşabileceğimizin ve aşıp aşmadığımızı nasıl ölçeceğimizin belirli bir yolunun olmadığı iddia edilerek, yapay bilinç başı ve sonu bilinmeyen beyhude bir hedef olarak görülmektedir: "Teorik akıl yürütme ile bilinçli olacağını tahmin ettiğimiz tamamen yeni bir yapı yaratma kabiliyetine sahip değiliz. Hatta nasıl başlayacağımızı ve başarımızı nasıl ölçeceğimizi bildiğimizi sanmıyorum." (Kim, 1998: 102). Ancak yapay bilincin inşası yönünde umutsuzluğa kapılmanın yersiz olduğu zira bilim ve teknoloji tarihinin doğası bilinmeden keşfedilen birçok örnek içerdiğini ve bilincin de yapay bir eşleniği için bilincin tam bir açıklamasının zorunlu olmayabileceğini iddia eden başka bir grup da yapay bilinç alanına güçlü bir dayanak sağlamaktadır:

> Tarihsel olarak, önemli bilimsel ilerlemeler yapısal tanımların yokluğunda gerçekleşmiştir. Örneğin, elektrik akımının fenomenolojik kuralları, elektronun 1892 yılında Thompson tarafından keşfinden önce, Ohm, Ampere ve Volta tarafından

> formüle edilmişti. Bu yüzden şimdilik, bilincin bir tanımını benimsiyorum ve bunu yapmaya ne kadar uzaklıkta olduğumu göreceğim (Koch, 2004: 12).

Yapay bilincin vakit harcamak için gereksiz bir alan olduğunu varsayanların sahip olduğu örtük inanç, biyolojik şovenizm olarak adlandırılan insandan başka hiçbir varlığın bilince sahip olamayacağıdır. Hiçbir makinenin asla bir insan gibi olamayacağı varsayımı, insanın tarih boyunca kendisini evrendeki en ileri seviyedeki canlı olarak konumlandırmasından ileri gelmektedir. Sahip olunan ayrıcalık konumu kaybetme kaygısı, insanın kendi bilincini taklit eden bir yapı ihtimali üzerine düşünüldüğünde, tüm bilimsel çabaya set çekebilecek kadar ileri gidebilmektedir. Fakat yapay bilinç, aşması gereken tüm zorluklarına rağmen hem bilinci davranışı yansıtan gerçek bir fenomen olarak kabul ettiği için hem de yapay sistemler ya da makineler yoluyla zihnin en içkin olarak adlandırılabilecek bilinçli deneyimi yeniden üretmeyi amaçladığı için ümit vaat eden bir alan olarak kabul görmeyi başarmıştır (Manzotti, 2013: 8).

Sıklıkla yapay zekanın bir uzantısı olarak adlandırılan yapay bilinci, yapay zekadan ayırmak ilk bakışta kolay görünmektedir. Basit bir ifadeyle, yapay zeka zeki ya da akıllı yapılar üretme girişimi olarak tanımlanabilirken; yapay bilinç de bilinçli yapılar üretme girişimi olarak tanımlanabilir. Ancak, bu ayrım göründüğünden daha karmaşık sorunsallar içermektedir çünkü zeka ve bilinç sınırları muğlak olan ve iç içe geçmiş unsurlardır. Ayrıca yapay zeka alanı için kullanılan "zeka" ifadesi bugün gelinen teknolojik düzeyde yetersiz bir kavram olarak görünmektedir. Çünkü yapay zeka alanı, bugün yalnızca zeki görünen yapılar üretme çabasında değildir. Yapay zeka çalışanlarının merkezi amacı, zekadan çok daha fazlasını ifade eden zihinsel özelliklere sahip veya zihin sahibi varlıkların özelliklerine sahip yapılar ya da sistemler meydana getirmektir. Zihinsel özelliklerin kapsam alanı zeka haricinde algı, eylem, duygu, yaratıcılık ve bilinç benzeri unsurları da içermektedir. Bu bağlamda yapay bilinç, yapay zihne atfen kullanılan yapay zekanın bir alt dalı olarak görülebilir. Fakat her ne kadar yapay bilinç, yapay zekanın bir alt dalı olsa da yapay bilinç alanı yapay zekadan fazla olarak zihinsel kapasitelerin fenomenolojik yönlerini de dikkate alır ve zihnin yalnızca üçüncü şahıs niteliklerini değil, birinci şahıs niteliklerini de araştırma konusu yapar (Chrisley, 2008: 120). Şimdiye kadar hiç kimse, zihnin beyinle özdeş olduğunu gösterememiştir. Bilinç olarak deneyimlediğimiz şey insanın bir parçasında ya da beyninde değil, bütününde ortaya çıkan bir şeydir. Bu nedenle eğer insan yapımı yapay bir madde insan gibi eyleyebilirse, yapay bilinç yeni bir statü elde edecek ve yapay zekanın güncellenmiş bir versiyonu ortaya çıkacaktır (Manzotti ve Tagliasco, 2008: 108). Bu açıdan

bakıldığında bugün yapay zekanın bir alt dalı olarak görülen yapay bilinç, gerçekleştirildiği takdirde yapay zekanın tanımını değiştirecek ve onu yeni bir ontolojik konuma yerleştirecektir.

Yapay bilinç terimi ilk kez sibernetik bağlamında Tihamêr Nemes tarafından 1962'de yazılan ve 1969'da İngilizceye *Cyberbetic Machines* olarak çevrilen *Kibernetikai Gêpek* kitabında kullanılmıştır (Manzotti ve Tagliasco, 2008: 105). Bu çalışmada Nemes bilinçli makinenin özelliklerini şu şekilde tasvir etmektedir:

> Yürüyen bir makinenin, mekanizmanın aktif olduğu bilgisini aldığını varsayın; bu onun "gidiyorum" önermesini kurması için yeterli olmalıdır. Bir çocuk gitmek kelimesinin gitmek ve yürümenin duyumlarının farklı karakterlerine rağmen ikisini de kapsadığını uzun bir deneyimle öğrenir. Çocuk bir yandan diğer insanların kendi yürüyüşlerine gitmek dediğini öğrenir ve bir yandan da aynada kendisinin diğer yürüyen insanlarla aynı hareketleri sergilediğini görür (Nemes, 1969: 219).

Mühendislik alanında bilinç hakkında ilk kez açıkça konuşan ise Giuseppe Trautteur'dur. "Yapay Bilinç Mümkün müdür?" makalesinde şöyle tartışır:

> Yapay zeka araştırmacılarının özellikle kaçındığı bir fenomen vardır ve bu bilinçtir. Bu çalışma, kompleks otonom faillere ulaşmak istediğimizde yapay bilincin mümkün ve zorunlu olduğunu tartışmaktadır (Trautteur, 1991: 42).

Mühendislik alanı kendi tarihi içerisinde genellikle kesin tanımı olmayan bir yapıyı tanımlamak için uğraşmıştır. Mühendisler, gerçek ya da sanal bir varlığın kendi hayal dünyaları içerisinde bir taklidini inşa etmeye uğraşırken görünenin arkasında yatan gerçekleri araştırmakla uğraşmamışlardır. Fakat başka bir yandan, nörobilim ise biyolojik yapıların daha derin analizleri için, yapay bir bilincin meydana getirilmesi gerekliliğini vurgulamıştır:

> Zihinseli anlamak için, beyinlere bakmanın yeni yollarını bulmalıyız. Hatta bu süreçleri daha iyi anlamak için beyne (bedensel fonksiyonlara) benzeyen yapıları yapay olarak üretmeliyiz. Böyle bilinçli yapıları yaratacağımız gün çok uzak olsa bile düşüncenin kendisinin süreçlerini derinlemesine anlamadan önce yapmalıyız (Tononi ve Edelman, 2000: 220).

Yapay zeka alanına ait en temel sınıflandırmanın kaynağı zayıf ve güçlü yapay zeka ayrımıdır. Bu ayrım, bilincin zor problemi olarak bilinen (Chalmers, 1996), duyusal bilginin farkındalığının nasıl ve neden oluştuğu ile deneyimin neden öznel bir niteliğe sahip olduğu soruları üzerine kurulan, bilincin en çok korkulan yönünden beslenir. Aynı şekilde erişim bilinci ve fenomenal bilinç ayrımı da (Block, 1995: 2002) zayıf ve güçlü yapay zeka

ayrımının bir başka dayanak noktasıdır. Erişim bilinci, bilincin objektif ve genel kontrole açık, davranışsal olarak gözlemlenebilir, akıl yürütmeyi, eylemde bulunmayı, konuşma kontrolünü sağlayan bileşenlerinden meydana gelir. Fenomenal bilinç ise öznel ve niteliksel deneyim ile özdeştir. Bir durumu fenomenal bilinçli yapan şey "o durumda olmak" gibi bir şeyin varlığıdır (Manzotti, 2007: 177).

Bilincin öznel ve nesnel nitelikler ile bilişsel ve fenomenal durumları aynı anda kapsayan bir yapıya sahip oluşu, bilincin bütün olarak anlaşılmasını zorlaştıran en önemli unsurdur. Bilinç durumlarının birbirine indirgenemeyen karşıt bileşenleri, makine bilinci araştırmacılarının da farklı yaklaşımları benimsemesine neden olmuştur. Bilinç, zihnin yapay zeka tarafından anlaşılmaya, üretilmeye veya taklit etmeye en az uygun yönü olarak düşünülür zira bilincin yalnızca birinci şahıs tarafından bilinebilen doğası ve öznel karakteri ile yapay zeka yöntemlerinin hesaplamaları, algoritmaları, süreçleri ve işlevleri birbiriyle uyuşmaz görünmektedir. Sembolik ve nesnel bir algoritma, bilincin öznel bileşenlerini üretmek için hiçbir zaman yeterli olmaz düşüncesi, yapay bilince bakış açısının ekseriyetle kötümser olmasına yol açmıştır.

Yapay bilinç çalışmalarına motivasyon sağlayan çok sayıda çıkış noktasından bahsedilebilir. Bilinçli bir varlık olarak insanın özünde fiziksel bir varlık olduğu kabul edildiğinde başka bir fiziksel varlığın da bilinçli olabilme ihtimali ortaya çıkmaktadır. Fiziksel bir fenomenin varlığından sorumlu yasalar, insan tarafından üretilen yapay fiziksel sistemler için de işletilebilir görünmektedir. Bunun yanı sıra, bilişsel yaklaşım evrensel bir makinenin insanın hesaplayabileceği her şeyi hesaplayabileceğini kabul etmektedir. Bu kabulün altında, zihinsel süreçlerin matematiksel olarak ifade edilebileceği varsayımı yatmaktadır. Buradan hareketle, evrensel bir makinenin işlevsel olarak insan zihniyle eşdeğer bir biçimde programlanabileceği düşünülmektedir. Zihinsel olanı hesaplanabilir kabul eden bu yaklaşımda, planlama, teorem kanıtlama, satranç oynama, konuşma, tanıma gibi zihin gerektirdiği düşünülen birçok davranışın bilgisayarlar tarafından taklit edilebildiği vurgulanırken; bilincin ya da en azından bilince ihtiyaç duyan davranışın da yapay sistemler tarafından taklit edilebileceği görüşü mevcuttur (Chrisley, 2008: 126).

Yapay bilinç, bilim ve mühendislik tarafından farklı şekillerde işlenen bir alandır. Bilimsel yapay bilinç ile teknolojik yapay bilinç birbirlerinden farklı hedeflere sahiptirler. Mühendislik yapay bilincinin temel ilgi alanı, doğal bilince sahip varlıklar tarafından yapılan şeyleri yapabilen yapay sistemler üretmektir. Bu yapay sistemlerin sahip olduğu işlevlerin

ve sergiledikleri davranışların, doğal sistemlerin sahip oldukları yoldan sergileyip sergilemediği hususu mühendislik yapay bilinci için hayati bir öneme sahip değildir (Chrisley, 2008: 121). Bu alanın öncelikli ilgisi gelişen teknolojinin işlevsel kapasitesidir. Başka bir deyişle, bir sistemin daha çok bilinçli bir sistem gibi davranmasını sağlayarak elde edilecek işlevsel yararlara odaklanır. Üzerinde çalışılan sistemin gerçek ya da somut bir bilince sahip olup olmadığı yapay bilinç mühendislerinin araştırma konusu değildir.

Yapay bilinç alanındaki araştırmacılar, doğal bilinç alanındaki çalışmaların nihai bir hedefe ulaşmasının uzak olduğunu düşünmektedirler. Bu nedenle, mühendislerin amacı önceden birleştirilmiş bir robotun içine bir bilinç modülü yerleştirmek değil, bilinçli (gibi) olan bir robot ya da bilinçli (gibi) davranan bir robot inşa etmektir (Manzotti ve Tagliasco, 2008: 108).

Mühendislerin başka alanlarda da yapay maddenin inşasında kullanılan yöntem ve süreçlerin kurallarını bilmeden oluşturabildikleri yapay maddeler mevcuttur. Üç boyutlu protein katlanmasını yöneten kurallar bilinmeden protein tasarlanıp kurulması; mühendisler için, bilincin mahiyetinin bilinmeden yapay bilincin inşa edilebilmesi yönünde umut doğuran bir örnektir.

Bilimsel yapay bilinç ise bilincin altında yatan süreçleri ve mühendislik yapay bilinci tarafından sağlanan teknolojileri anlamakla ilgilenir (Chrisley, 2008: 121). Temel amaç, teknolojik yapay bilincin yalnızca taklit bir bilinç mi olduğunu yoksa bilincin doğasına teşkil eden süreçleri aydınlatan ve somutlaştıran bir unsur mu olduğunu ayırt etmektir. Dolayısıyla bilimsel yapay bilinç içerisindeki tartışmalar yapay bilinçli sistemlerde kullanılan teknoloji ve bilincin mahiyeti arasındaki ilişki üzerinedir.

Zayıf ve güçlü yapay zeka ayrımı, yapay bilinç için de benzer şekilde kullanılmaktadır. Owen Holland (2003), bu alan için de güçlü yapay bilinç ve zayıf yapay bilinç ayrımını önermiştir. ***Zayıf yapay bilinç*** olarak isimlendirilen, bilinci ya da bilinçle ilgili bilişsel süreçleri taklit eden makinelerin tasarlanması ve inşasına yönelik çalışmalardır. Zayıf yapay bilinç, teknoloji ve bilinç arasında zorunlu ya da yeterli herhangi bir ilişki iddiasında bulunmaz; yalnızca teknolojinin bilinci anlamak için kullanılmasıdır (Chrisley, 2008: 121). ***Güçlü yapay bilinç*** alanı, somut olarak bilinçli makinelerin yapısı ve nitelikleri üzerine araştırma yapar. Bu yaklaşımın merkezi sorunsalı uygulandıklarında bilincin örneği olacak, somut bilinç için yeterli olan sistemler tasarlamaktır. Zayıf yapay bilinç çalışmaları

simülasyon olarak bilincin varlığını bir makineye bilinçli demek için kâfi görürken güçlü yapay bilinç teorisi için bir makinenin bilinçli olması, onun ancak somut, gerçek bir bilince sahip olmasıyla mümkün olacaktır. Fakat güçlü ve zayıf yapay bilinci birbirinden ayıran sınırları belirlemek oldukça zordur. Bilinçli varlığa ait tüm davranışları sergileyen bir makinenin bilinçli statüsünü akla uygun bir şekilde yok saymak kolay olmadığı gibi, bu makinenin öznel ve niteliksel bir bilince sahip olduğunu iddia etmek de aynı oranda zordur. Bilince hangi aşamada simülasyon değil de gerçek bilinç diyebileceğimize dair kriterler net bir şekilde ortaya konulamamıştır. Öznel niteliklere sahip bir makine bilinci üretebilmek için, bilincin fenomenal yönünün bilimsel olarak ele alınması zorunludur. Ayrıca fenomenal bilincin, zihnin işlevsel yönleriyle bağıntılı olduğu da herkesçe kabul görmektedir. Öznel ve nesnel bileşenler, birbirlerine indirgenemez karşıtlıklarına rağmen, bilinci bir arada var ettiklerinden, birbirlerinden bağımsız olarak çalışılmaları mümkün değildir. Yalnızca fenomenal bilince odaklanan bir makine bilinci araştırması tahayyül edilemez, zira fenomenal bilinç, bilincin bilişsel ve işlevsel karakteriyle doğrudan ilişkilidir. Bu bağlamda, yapay bilinç hem imkânının hem de epistemolojik temellerinin sorgulanması açısından birçok sorunsalı barındırmaktadır.

Yapay bilincin hangi aşamada başarılmış olarak kabul edileceği sorusu, bilincin öznel doğasının yapay bilinç içerisinde test edilemeyeceği düşüncesinden ileri gelmektedir. Bu problem, yalnızca yapay bilinç alanı ile ilgili olmayıp, bilincin tüm üçüncü şahıs açıklamalarında geçerlidir. Bu probleme çözüm yolu olarak, bilincin altında yatan fiziksel sistemi çeşitli soyutlama seviyelerinde modellemek, bilinçli süreçlerin nedensel yapılarını dolaysız bir şekilde iç gözlem yoluyla not ederek modellemek ya da bilinçli olduğuna inanılan bir sistemin, bilinçli davranışına neden olan fiziksel ya da fenomenal yapının hiçbir doğrudan bilgisi olmadan modellemek gibi farklı uçlarda yer alan öneriler ileri sürülmektedir (Chrisley, 2008: 127). Ancak görülmektedir ki yapılan önerilerin hemen hepsi bilincin tamamıyla fiziksel süreçlerden kaynaklandığını önsel bir bilgi olarak kabul etmektedir. Zira aksi takdirde, öznel-nesnel ikiliğinin aşılması bilincin bilimsel incelemesi için çok zor görünmektedir.

Güçlü ve zayıf yapay bilinç yaklaşımlarının arasında ihmâl edilmiş bir seçenek olarak ***lagom yapay bilinç*** yaklaşımı da dikkate değerdir. Lagom hem dilimizde hem de İngilizcede doğrudan karşılığı bulunmayan İsveç diline ait bir kelimedir; "ölçülülük yoluyla mükemmellik" gibi bir anlamı vardır (Chrisley, 2008: 121). Lagom yapay bilinç yaklaşımı -

zayıf yapay bilincin aksine- teknoloji ve bilinçli zihinsel fenomenler arasındaki ilişkiyi ihmal etmez ve ortak açıklayıcı özellikleri olduğunu savunur. Fakat güçlü yapay bilinçten farklı olarak da; bu ortak özelliklerin bilincin varlığı için tek başına yeterli olmayacağı iddiasındadır. Güçlü yapay bilinç, bilinç için yeterli koşulları keşfetmeyi hedeflerken; lagom yapay bilinç zorunlu koşulları keşfetmeyi amaçlar denilebilir (Chrisley, 2008: 121). Bu yönüyle, bilincin doğal dünyanın nasıl bir parçası olduğuna dair tam bir açıklama yapamasa bile önemli katkılar yapabilir görünmektedir. Lagom yapay bilinç, zayıf yapay bilinçten fazla olarak yapay bilinci başarmak için temel olarak gerekli şartlarla ilgilenirken; güçlü yapay bilinçten az olarak bu gerekli şartların yeterliliğini değerlendirmez.

Yapay bilinç, hem yapay zeka alanının karşı karşıya kaldığı sorunları hem de bilincin açıklamasında yaşanan problemleri bir arada göğüslemek zorunda olan bir alandır, zira her iki alanın niteliklerini de tanımı gereği kendisinde barındırmaktadır. Bilinç biliminin uğraştığı zor sorun, açıklama gediği, qualia bilgi argümanı gibi problemler, yapay bilinç için de aynı şekilde geçerliliğini korumaktadır. Bunun yanı sıra, yapay zekaya yöneltilen sembolik bir sistemin, bir objenin bilişinin milyonlarca seri işlem gerektirmesinden kaynaklı olarak insan bilişinin doğru bir modellemesinin yapılamayacağı itirazı aynı şekilde yapay bilince de yöneltilmektedir. Dahası, yapaylık ile zihinselliğin iki zıt ontolojik kategori olarak asla bir arada var olamayacağını iddia eden düalist yaklaşım için yapay bilinç tanımsal olarak dahi bir çelişkidir. Yapay zekaya karşı Çince Odası ile bilinir hâle gelen semantik itiraz da yapay bilincin hesaplaşması gereken diğer bir sorundur.

Semantik itiraz kapsamında Turing Testi'ni eleştiren en önemli argümanlardan birisi şüphesiz ki John Searle'ün Çince Odası düşünce deneyidir. Searle zihin ile girdi-çıktı ilişkisi aracılığıyla yürütülen bir bilgisayar programı analojisini kabul etmemektedir. Ona göre, bir varlığın zihne ya da düşünme yetisine sahip olduğunu söyleyebilmek için Turing Testi'ni geçmesinden daha fazla özelliğe sahip olması gerekmektedir. Searle zihnin, bir bilgisayar programı ile taklit edilemeyeceği inancı doğrultusunda, dijital bir bilgisayarın bize insan zihninin doğru bir tasavvurunu verip veremeyeceğini değerlendirmek açısından zihin felsefesinin önemli ispat yöntemlerinden olan düşünce deneyine başvurur. O, Çince Odası deneyini güçlü yapay zeka teorisinin savunulabilir olmadığını göstermek için kurguladığını ve bu sebeple güçlü yapay zeka ve zayıf yapay zeka teorileri arasındaki farkı kavramanın önemli olduğunu sıkça vurgular. Güçlü yapay zeka teorisi, zayıf yapay zeka teorisinden farklı olarak, doğru programlanmış bir bilgisayarın doğru girdi-çıktı ilişkileriyle

düşünebileceği ve henüz keşfedilmemiş psikolojik özelliklerinin olabileceği değil, düşünebilmek için gerekli olan her şeye sahip olduklarından düşünmelerinin zorunlu olduğu tezi üzerine kuruludur (Searle, 1990: 27). Güçlü yapay zekaya göre, program doğru uygulandığında herhangi bir donanım ile zihinsel durumlar üretilmesi mümkündür. Zayıf yapay zekada ise bilgisayarlar zihin durumlarını üretemezler, yalnızca simule edebilirler. Ancak nasıl ki bir patlama simülasyonu patlamanın kendisiyle aynı şey değilse, zihinsel durumların simülasyonu da zihinsel durumların kendisi değildir (Searle, 2018: 26). Bu farklılık göstermektedir ki, güçlü yapay zeka teorisi bir bilgisayarın düşünmesini mantıksal bir olasılık olarak görmekten öte, doğru programlanmış bir bilgisayarın insan zihninden hiçbir farkı olmadığını kabul ettiğinden, düşünme yetisini bilgisayarlarla olasılık olarak değil, zorunluluk olarak ilişkilendirmektedir.

Searle Çince Odası düşünce deneyinin genel çerçevesini şöyle çizer:

> (…) Ben hiçbir şekilde Çinceyi anlamıyorum hatta Japon yazısını Çin yazısından bile ayıramıyorum. Fakat benim Çince sembollerden oluşan kutularda dolu bir odada, bir kural kitabı ve beni sorulara Çince cevaplar vermeye ulaştıracak bir bilgisayar programı ile birlikte bir odada kilitli olduğumu hayal edin. Hiçbir şekilde bilmediğim sembollerden oluşan soruları alıyorum, ne yapacağıma dair kural kitabına bakıyorum, kutulardan sembolleri alıyorum, kitaptaki kurallara göre onları değiştiriyorum ve cevap olarak ihtiyaç olan sembolleri veriyorum. Burada benim Çinceyi anlamak için olan Turing Testi geçtiğimi farz edebiliriz ancak ben Çincenin tek bir kelimesini bile anlamıyorum. Ve eğer ben doğru programı uygulama temelinde Çinceyi anlamıyorsam, program temelli hiçbir bilgisayar Çinceyi anlayamaz çünkü bir bilgisayar bende olmayan hiçbir şeye sahip değildir (Searle, 2004: 90).

Bu düşünce deneyinin amacı, zihnin bir bilgisayar programına indirgenemeyecek niteliklere sahip olduğunu ortaya koymaktır. Anlamadığımız bir dilde sorulan soruları, programdaki adımları uygulayarak yanıtlayabiliriz. Çince sembollerin yer aldıığı kutularla dolu bir oda, bu deneyde programın yararlandığı veri tabanına karşılık gelmektedir. Çince sembollerin yer aldığı kural kitabı ise bilgisiyar programının kendisidir. Deneye bütün olarak bakıldığında odadaki kişi olan ben Çince sorulara cevap veren programı yürüten bir bilgisayardan başkası değilimdir ve tek kelime bile Çince anlayamam. Searle'e göre doğru çalışan bir bilgisayar programı, zihin sahibi bir bilgisayar programı olarak adlandırılamaz. Bilgisayar programının doğru bir şekilde yürütülmesi, Çince Odası'ndaki operatörün soruları anlamadan doğru cevap vermesi özünde aynı şeylerdir. Hem zihin hem de bilgisayar programı, girdi-çıktı ilişkisi içerisinde çalışıyor gibi gözükse dahi zihni programla eşdeğer tutanların gözden

kaçırdığı nokta, zihnin programdan farklı olarak sahip olduğu *anlama kapasitesidir*. Bu anlama kapasitesi ise sembolik olarak üretilemez bir yetiye karşılık gelmektedir.

Searle bu düşünce deneyinden sonra güçlü yapay zeka teorisyenlerinin Turing Testi'ni Çince ya da başka bir dili anlayarak geçecek bir bilgisayarı da icat etmenin mümkün olduğunu savunacaklarını varsayarak bir bilgisayar programının neden ilkece anlamsal içeriğe sahip olamayacağı üzerine akıl yürütmelerde bulunmuştur.

Güçlü yapay zeka teorisini destekleyenler, gelişmiş bir teknoloji ile programlanmış sofistike bir bilgisayarın anlamsal içerik sahibi olmasının önünde bir engel bulunmadığını iddia etmektedirler. Searle'ün bu iddia karşısında duruşu net bir şekilde olumsuzdur:

> (…) Bunlar derin bir sebeple alakasızdır: Bilgisayar sembolleri manipüle ederek çalışır. Onun süreçleri tamamen söz-dizimsel (sentaktik) olarak tanımlanır, zıt olarak insan zihni tercüme edilmemiş sembollerden çok daha fazlasına sahiptir, insan zihni anlamları sembollere iliştirir (Searle, 2004: 91).

Searle'ün Çince Odası düşünce deneyiyle hedefledikleri ve hedeflemediklerinin ayrımını yapmak önemlidir. Çünkü Searle sıklıkla yapay zeka karşıtı olarak algılanır. Searle'ün karşı çıktığı husus, bilgisayarların sembolik olarak akıl yürütmesi değil, bilgisayarların sembolik hesaplamalarının semantik içerik üretebileceği iddiasıdır. Bilgisayarın hesaplama yetisi ile insan zihninin işleyişi arasındaki farklılığın en büyük kaynağı Searle'e göre, bilgisayarın yaptığı hesaplama ve akıl yürütmelerin ne anlama geldiğinin farkında olmamasıdır. Çince Odası ile Searle'ün ulaştığı sonucu şöyle ifade etmek mümkündür: Mükemmel şekilde insan taklidi yapabilen bir programın, kusursuz bir şekilde Çince soruları cevaplandırması dahi onun sorulan soruların ve verdiği cevapların anlamını bildiğini göstermeyecektir. İnsan zihni Çince Odası'nda kullanılan programdan farklı olarak yaptığı hesaplama ve akıl yürütmeleri anlayarak ve anlamlandırarak yapmaktadır. Bu minvalde, yapay bir sistemin, doğal zihinden eksik olan yanı anlamsal içerikten yoksun olmasından ileri gelmektedir.

Çince Odası argümanını oluştururken Searle'ün ilk aksiyomu şudur: "Bilgisayar programları sentaktiktir." (Searle, 1990: 27). Bu demektir ki; bilgisayar programları soyut ve fiziksel bir varlığa sahip olmayan, nedensel ilişkilerden yoksun sembollerle çalışmaktadır. Bu sembollerin kendileri Searle'e göre hiçbir anlamsal referans içermezken anlamları tamamen programcı tarafından yaratılmaktadır. Bunun devamında Searle'ün kullandığı ikinci aksiyom, insan zihinlerinin, bilgisayar programından farklı olarak zihinsel içeriğe, bir başka ifadeyle de semantiğe sahip olmalarıdır (Searle, 1990: 27). Bu aksiyom, düşünce, algı ve

anlamaların zihinsel birer içerik oldukları yorumunu da gerektirmektedir. Buna göre, kelimeler veya sembollerle düşündüğümüzde, o kelime ve sembollerin ne anlama geldiğini de bilmek zorundayızdır; bu sebeple Türkçe düşünebilirken Çince düşünemeyiz (Searle, 2018: 22). Searle üçüncü ve ilk iki aksiyomu birleştiren en temel kabulünü ise şöyle ifade eder: "Sentaks, ne tek başına semantik için yeterlidir ne de semantiğin kurucusudur." (Searle, 1990: 27). Sentaks ve semantik arasında zorunlu bir ilişki olmadığını ve sentaksın hiçbir şekilde anlamsal içerik için yeterli ve kurucu unsur olmadığını temel alan bu aksiyom, Searle'ün düşünce deneyini üzerine kurduğu temel yargısıdır. Sistemin gerçekten anlayan bir insanın davranışlarını ne ölçüde taklit edebildiğinin ya da yaptığı sembol işlemelerinin ne ölçüde karmaşık olduğu mühim değildir; yalnızca sentaktik süreçlerden yararlanılarak asla semantik üretilemez (Searle, 2018: 22).

Searle'e göre bu üç aksiyom bizi bu akıl yürütmenin ilk sonucuna ulaştırmaktadır: "Programlar, ne zihin için yeterlidir ne de zihnin kurucusudur." (Searle, 1990: 27). Eğer bir zihinden ve zihnin bir durumu olan düşünmeden bahsedilecekse, sentaktik kurallara dayalı bir program bunun için yeterli değildir, zira sentaktik bir yazılımın tek başına anlamsal içerik üretmesi mümkün değildir.

John Searle'ün yapay zekaya yönelik eleştirisi kendisinin de belirttiği üzere bir makinenin insan gibi düşünebileceğini iddia eden güçlü yapay zeka teorisine karşıdır. Bu eleştiri makine bilinci yaklaşımlarından ancak fenomenal makine bilincine yönelik olarak ele alınabilir. Searle, doğrudan fenomenal zihin durumlarından bahsetmese de anlamsal içeriğin taşındığı zihin durumları, fenomenal durumlardır. Zira Searle bir makinenin insan gibi davranabileceğini, insan gibi hesaplama yapabileceğini de varsayım olarak mümkün görmektedir. Ancak onun itirazı bir makinenin anlamsal içerik üretmesi yetisine sahip olabilmesi ihtimalinedir. Searle'e göre, bir makinenin öznel seviyede bir bilince sahip olması mümkün değildir. Zira öznel deneyimler semantik içerik gerektirmektedir ve makinelerin sahip olduğu sentaktik içerik, semantik üretmek için yeterli değildir. Başka bir deyişle, bir makinenin sahip olduğu formel yapı onun asla birinci şahıs olarak içerik üretmesine neden olmayacaktır.

Yapay bilinç çalışmalarının hem motivasyon kaynakları hem de yüzleşmek zorunda oldukları karşı çıkışlar, yapay zeka ve bilinç alanlarının ortak kümesi olarak ele alınabilir. Yapay bilinç, birçok araştırma programını aynı anda gerektiren heterojen bir çalışma alanıdır. Yapay bilinç ile ilgilenen bilim insanlarının bazıları bilinçle ilgili davranışlar

üzerine çalışmakta, bazıları bilincin bilişsel niteliklerini modellemekle uğraşmakta ve bazıları ise makinelerde fenomenal durumlar oluşturulması yönünde bir eğilimle araştırma yapmaktadır.

Bu bağlamda, David Gamez tarafından açıkça ortaya koyulduğu gibi makine bilinci araştırmaları dört grup altında toplanabilir (Gamez, 2008: 888). Birinci grup genel yapay zeka olarak adlandırılan ve bilinçle ilişkilendirilen dışsal davranışlara sahip makineler ile ilgilenmektedir. İkinci grup genellikle bilişsel bilimcilerin uğraştıkları bilincin bilişsel niteliklerine sahip makineler üzerine çalışmaktadır. Üçüncü grupta insan bilinciyle ilişkilendirilen yapıya sahip makineler üzerine çalışan nörobilimciler bulunmaktadır. Son grup ise diğer üç grubun çalışmalarını temel alarak fenomenal anlamda bilinçli makineler oluşturmayı hedeflemektedir. İlk üç grup için simülasyon olarak bilincin varlığı yeterliyken son grup araştırmalar makinenin kendisinde bilincin varlığını aramaktadırlar. Bilinç taklidinin hedefi bilincin bazı yönlerini açıklamak ya da bilincin sinirsel ve davranışsal bağlantılarını bilişsel bir modelde bulmaya çalışmaktır. Simülasyon olarak bilinç üretmenin fenomenal bilincin varlığını reddetmek gibi bir hedefi yoktur, yalnızca yapay sistemlerin fenomenal durumlarıyla ilgilenmez. Son gruptaki çalışmalar ise bilincin simülasyonuna değil, somutlaşan bilince odaklanmaktadırlar. Yapay bir sistemin fenomenal bilince sahip olup olamayacağı, öznel ve niteliksel deneyimlere sahip olup olamayacağı bu grubun temel sorularını oluşturmaktadır.

4.1.1. Davranış Temelli Makine Bilinci

Pozitivizm etkisinde şekillenen felsefi davranışçılık, zihin durumlarının davranışlarla eşdeğer olduğunu savunmaktadır. Bu yaklaşıma göre, bir makinenin sahip olduğu zihinsel durumlar da davranışlarıyla ölçülebilir. Bilinçli insan davranışlarını sergileyen bir makinenin bilinçli olduğunu söylemenin önünde herhangi bir engel yoktur. Davranışçılığa göre, yüzünde aptal bir sırıtmayla, kötü şiirler yazan, şarkılar mırıldanan bir makinenin âşık olduğunu kolaylıkla söyleyebiliriz. Bir makinenin, bilinçli aşk durumunda olduğunu gösteren yegâne işaret onun davranışlarıdır. Aynı şekilde, çenesini tutarak ağlayan bir robot da davranışçılığa göre diş ağrısı çeken bir robottur.

Davranış temelli makine bilinci çalışmaları zihin durumlarını, davranışa indirgeyen davranışçılık etrafında şekillenmektedir. Bu yöndeki çalışmaların amacı, davranışsal olarak

insana benzeyen makineler üretmektir. Bu makinelerin üretiminde hedeflenen bilincin oluşturulması değildir, yalnızca bilinci taklididir. Son on yıla kadar, robotik teknoloji ve mühendislik neredeyse tamamen davranış odaklı bilinç simülasyonları inşa etmek üzerine çalışmaktaydı. Ancak, son yıllarda makine öğrenmesiyle yaşanan hızlı bilişsel gelişmeler, mühendislik çalışmalarını da davranışın yanı sıra bilişsel simülasyonlar üretmeye yönlendirmiştir.

Turing Makinesi çalışmaları ile dar yapay zeka çalışmalarının dâhil olduğu birinci grup yürümek, araba kullanmak, konuşmak gibi birçok uyanık davranışımızın bilinçsiz tepkilerden oluştuğunu savunur. Bu gruba göre, fenomenal zihin durumlarının oluşturulması davranışsal olarak zeki makinelere bu durumların eklenmesiyle gerçekleştirilebilir. Ancak bu yaklaşımın karşılaştığı en büyük sorun zombi problemidir. Zira bir zombi robot da hiçbir deneyime sahip olmadan bilinçli insan davranışını taklit edebilir. Bu gruptaki çalışmalar, bilinçli ve bilinçsiz davranışı ayırt etmeyi hedeflemediği gibi bilinçli ve bilinçsiz makineleri farklılaştıran herhangi bir ölçüt de önermemektedir.

4.1.2. Bilişsel Temelli Makine Bilinci

Davranış temelli makine bilinci araştırmaları, bilincin davranışsal dışavurumları üzerine çalışma yürütürken; ikinci grup araştırmalar ise bilincin bilişsel bağlantılarına odaklanmaktadır. Biliş kelimesi, duyusal ve motor fonksiyonlar ve refleksler haricindeki bilinçli ya da bilinçsiz olmalarından bağımsız olarak algı, öğrenme, hafıza, yaratıcılık, düşünme, tahmin ve planlama gibi beyin fonksiyonlarına işaret eder (Roth, 2000: 78). Biliş, yalnızca insan zihniyle sınırlı olarak ele alınmaz ve bilincin varlığı gibi bir ön koşulu gerekli kılmaz. Bilişsel niteliklerin yapay olarak modellenmesi ise basit bilgisayar programlarından taklit nöronlara dayalı sistemlere kadar çok geniş bir yelpazede çalışılmaktadır. Dışsal davranışlarla ilişkilendirilmiş makineler bilişsel modeller için zorunlu değildir. Bilişsel temelli makine çalışmalarında duyguların dışsal davranışına sahip olmayan bir bilgisayar programı da bilinçli olarak adlandırılabilir.

Felsefeciler, bilişsel nörobilimciler, psikologlar, farmakologlar, fizikçiler, yapay zekacılar başta olmak üzere birçok araştırmacı bilinci işlevsel terimlerle tanımlamaya ve açıklamaya çalışmıştır (Baars, 1998; Blackmore, 2002; Dennett, 2001; Sun, 1997; Velmans, 2002, 2009a; Chrisley, 2003; Haikonen, 2007). Bu türden modellerin inşasında temel olacak hiçbir

açıklama, bilinçli makinelerin bilişsel modellerini oluşturmaya temel teşkil edecek bütün bir bilinç teorisi ortaya koyamamıştır. Bilişsel temelli yaklaşım, bilinci fiziksel zihnin bir özelliği olarak ele almakta ve açıklamalarında metafizik unsurlara yer vermemektedir. Ancak bilincin fiziksel açıklamasındaki eksiklikler, bilincin bütün bir bilişsel teorisinin üretilmesine engel olmaktadır (Starzyk ve Prasad, 2011: 1). Dolayısıyla bilişsel temelli makine bilincinin ilk hedefi, bilincin fiziksel açıklamasını tam olarak yapabilmektir.

Araştırmacılar, makine bilinci ve imkânı üzerine oldukça farklı görüşler ileri sürseler de birçoğunun uzlaştığı husus makine bilinci teriminin bilişsel ve işlevsel analizinin bu alanının zorunlu ihtiyacı olduğudur (Starzyk ve Prasad, 2011: 2). Beden sahibi olma, deneyim, dikkat, duygu gibi çeşitli işlevsel bloklar, bilincin tanımlanması için gerekli şartlar olarak kabul edilmektedir. Beden sahibi olma ve yer kaplama, episodik hafıza (yaratıcılık), dikkat, hedef ve motivasyonlar (planlama) ile duygular bilinçli bir sistemin beş temel işlevsel özelliği olarak ileri sürülmüştür (Aleksander, Igor ve Dunmall, 2003: 632). Her ne kadar bu özellikler bilinç için geçerli olsa da yeterli şartların kümesini oluşturmamaktadır. İşlevsel özelliklerin yanı sıra kimi araştırmacılar da bilgi üzerine odaklanmış ve bilgi seviyeleri ile işlevsel seviyeleri eşleştirme çabası içerisine girmiştir.

Bilişsel görüş, zihnimizin işlem yapısı bakımından iki farklı türe ayrılabilecek olan bir bilgi-işleme mekanizması olduğunu ileri sürmektedir. İlk tür, duyu girdilerini inceleyen birbirinden ayrı birçok bilişsel modülden; ikinci tür ise birleşik bir merkezi sistemden oluşur (Revonsuo, 2016: 309). Modül, sadece belli türden bir bilgi türünün analiz edilmesi için özelleşmiş bir işleme mekanizmasıdır. Bilgi ulaşılabilir olduğunda modül, bilgiyi otomatik ve bilinçli deneyim olmadan işleyerek, bilincin kullanımına hazır olan bir çıktı üretir. Bu bilişsel model, zihnin modüler yapısını betimleyerek çalışma prensibini ortaya koymaya çalışırken bilinçli ve bilinçsiz süreçleri ayırt etmemize yarayacak bir ölçüt ortaya koymaktan uzaktır. Bu modele dayalı makine bilinci çalışmaları da yine zihnin modüler yapısını taklit eden bilişsel sistemler ortaya koymaya çalışmaktadır.

Zihin durumlarının çoklu gerçekleşebilirliği, bilişselci yaklaşımın yapay bilinç araştırmalarında önemli bir motivasyon kaynağıdır. Çoklu gerçekleşebilirlik, insan gibi davranan ancak fizyolojik olarak insandan çok farklı olan yaratıklara (genellikle Marslı olarak adlandırılır) dayanan düşünce deneyleriyle temellendirilir. Davranışsal olarak insanla eşdeğer olan Marslıların bir zihne sahip olabileceğini inkâr etmenin dayanak noktası bulunamadığından, zihin durumlarının yalnızca biyolojik yapılarda olabileceğini iddia

etmenin yanlış olacağını savunurlar. Bilişselcilere göre, Marslı ve insanın ortak yönü bu noktada davranış değil, zihinsel durumlara sebebiyet veren soyut nedensel organizasyondur (Chrisley, 2008: 121). İşlevselcilikle birebir örtüşen bu tez, zihin durumlarının bilinci de kapsayan bir küme olarak değerlendirdiğinden; bilincin fenomenal doğasını ve içgözlemsel verisini göz ardı ederek yapay bilincin inşa edilebilir olduğunu iddia etmektedir.

4.1.3. Nörobilimsel Temelli Makine Bilinci

Sinir sisteminin anlaşılma ve açıklanma çabasının kökleri Antik Yunan hekimlerinden Galen'de (MÖ 200) bulunabilir. Galen hareketin kaslara bağlı olduğunun ve kaslardaki iplik adını verdiği sinirlerin farkına varmış; sinirlerin psişik ruhu kaslara aktardığı ve psişik ruhun (nefes veya hava) kasa girmesiyle kasın şişerek hareketi oluşturduğunu varsaymıştır (Churchland, 2019: 27). Mekaniği psişikle birleştiren bu görüş, 19. yüzyıla kadar geçerliliğini korumuştur. Bunun yanı sıra, nörobilim her ne kadar 21. yüzyılda yükselişe geçmiş olsa da temel iddialarını, ruh ve beden etkileşiminin beyindeki epifiz bezinde gerçekleştiğini söyleyen Descartes'a kadar götürmek mümkündür[42]. Descartes'tan itibaren ruhun konumu ve zihnin işleyişiyle ilgilenen felsefe ile vücudun biyo-ekonomisinde maruz kaldığı hastalıklarda beynin etkisini merak eden tıp birleşmeye başlar (Rose ve Rose, 2018: 17). Descartes'ın ruh ve beden arasındaki etkileşimine mekanik yaklaşımının yanı sıra, insana bütün olarak da materyalist yaklaşımı söz konusudur; insanı karmaşık bir otonom olarak ele almaktadır:

> Bu makineye yüklemiş olduğum tüm fonksiyonları göz önüne almanızı istiyorum, örneğin yemeğin sindirilmesi, kalbin ve arterlerin atışı, vücudun çeşitli bölümlerinin büyümesi ve gelişmesi, nefes alma, uyuma ve uyanma; dış duyu organlarının ışığı, sesi, kokuyu, sıcaklığı ve bu tür diğer nitelikleri algılaması ve bütün bunlarla ilgili olarak sağduyu ve hayal gücü organındaki düşüncelerin izlenimi ile bu düşüncelerin bellekte saklanması ya da zihne etki etmesi; istek ve arzuların içsel hareketleri; ve son olarak, nesnelerin duyulara yönelen hareketlerini uygun şekilde takip eden tüm vücut bölgelerinin dış hareketleri… Bu fonksiyonlar, tıpkı bir saatin ya da başka bir

[42] Descartes, ölümünden sonra 1664'te basılan *L'homme* adlı eserinde görme yetisi ve eyleminin göz, beynin içindeki beyin epifizi ve kol kasları arasında uygun bir düzenlemenin sonucu olarak belireceğini öngörmüştür. Hafızayı, bu patikaların seçici şekilde güçlendirilmesi olarak düşünmüştür. Descartes'a göre bilinçteki dalgalanmalar beyin epifizini hareket ettiren canlılıktaki basınç değişiklikleriyle açıklanabilirdi: Yüksek basınç uyanık kalmaya, alçak basınç uykuya yol açardı. Descartes bu mekanik bakış açısına rağmen, zihin ve bedenin epifiz bezi aracılığıyla birbirini etkileyen farklı maddelerden yapıldığına inanmıştır (Dehaene, 2018: 17).

> otomatın hareketlerinin denge ağırlığı dağılımı ve tekerleklerden kaynaklanması gibi, doğal olarak organların dağılımından kaynaklanır. (Descartes, 1985: 99)

Saat mekanizmaları ve su fıskiyelerinin çalışma sistemlerinden oldukça etkilenen Descartes, bedeni son derece karmaşık bile olsa makine olarak kabul etmiştir. Descartes, sinirleri beyinden gelen ve beyne giden mesaj kabloları olarak değerlendirmiş ve canlı organizmaların beyin aracılığıyla yaptıklarını algıladıklarını ve istedikleri şekilde hareket ettiklerini iddia etmiştir:

> Ruhun, bedeni etkileyen şeyleri bedenin her biriminde bulunduğu için değil ama kendisine bedenin çeşitli kısımlarına dokunan dışsal nesnelerin çeşitli eylemleri sinirlerin hareketleri yoluyla iletildiği beyinde olduğu için hissettiği kolayca kanıtlanır (Descartes, 1644; akt. Haldane ve Ross, 1968: 293).

Descartes bu mekanik fikirlerini, yaşamsal faaliyetleri maddi olmayan ruhanî şeyler olarak değerlendiren Katolik kilisesine gösterdiği itaat ve bağlılık dolayısıyla yüksek sesle dile getirmekten kaçınmış ve ruh ile bedeni birbirlerinden bağımsız olarak var olan iki töz olarak kabul etmiştir.

Beyin ve zihin arasındaki ilişkinin materyalist açıklaması da bu tarihten iki yüzyıl sonra gerçekleşir. 20. yüzyıla kadar yalnızca ölü ve zarar görmüş insan beyinleri üzerinde inceleme yapılmış, milenyumla birlikte hayvanların beyinlerine müdahale edilmeye başlanmış ve günümüzde çeşitli görüntülüme teknikleriyle aktif insan beyni üzerinden haritalar oluşturulmaya çalışılmaktadır. 21. yüzyıldan önce yalnızca insan beyninin yapı ve işlevleri, beyin hasarları ve hastalıkların beyin üzerindeki etkilerini inceleyen nörobilim, daha sonra araştırmalarını, çocukluktan yaşlılığa kadar olan süreçte beyinde meydana gelen değişiklikler ile normal olmayan beyinler üzerinde yoğunlaştırmıştır. Bugün ise nörobilim, "hem şirketler hem de devlet tarafından, günümüzün asimetrik savaşlarının sonucu olan yeni silahlar geliştirme ve askerlerin hasarlı beyinlerine yönelik tedaviler geliştirme gibi konularda ortak projeler finanse edilen", küresel bir teknobilim girişiminin parçası hâline gelmiştir (Rose ve Rose, 2018: 10).

Nörobilimcilere göre, zihinsel aktivite tümüyle beynin işlemlerine indirgenebilir; zihinsel olaylar, insan beynindeki nöronlar tarafından oluşturulan trilyonlarca bağlantı arasında gerçekleşen nöro-iletici akışından başka bir şey değildir (Rose ve Rose, 2018: 21). Bu sebeple nörobilim beyinsel süreçlerin genetiği, fizyolojisi ve kimyasını açığa çıkarmakla ilgilenir. Bu yaklaşıma göre, beyinsel süreçlerin açıklanması, insanın aslında beyninin üretimi olan "ben" yanılsamasını ve beynin kararlarını kendi kararlarıymış gibi düşünmesine

neden olan "kişi" yanılsamasını ortadan kaldıracaktır. Nörobilimin bu temel varsayımından hareket eden nöro-felsefeciler ise bilincin kendisini dahi nörobilimin sade formülleriyle açıklamaya çalışarak, bilinci (fenomenal bilinç anlamında) halk psikolojisinin üretimi olan bir terim olarak kabul etmektedirler.

Nörobilimsel temelli makine bilincine en büyük finansal destek 2013 yılında İnsan Beyni Projesi için Avrupa Birliği'nden gelmiştir. Bu projenin amacı, beyni çözmek ve insan beyni benzeri bir bilgisayar oluşturmaktır. Projenin amacı "insan beynini anlamak ve nihayetinde hesaplama yeteneklerini taklit etmek için küresel bir iş birliğine dayalı çabayı hızlandıran, tıp ve bilgisayar alanındaki beyinle ilgili araştırmalar için bir bilgi işlem teknolojisi altyapısı oluşturmak" şeklinde tanımlanmıştır (akt. Rose ve Rose, 2018: 38). Proje başlangıcındaki niyet, insan beyniyle ilgili tüm veri, bağlantı ve kimyasal süreçleri bilgisayar sistemlerine aktarmak ve beynin nasıl çalıştığını modellemek olarak aktarılmıştır. Ancak, bugün gelinen noktada hem elde edilen verilerin çoğunun denek olarak kullanılan hayvanlardan alınmasının insan beyninin hayvan beyninden çok daha farklı ve gelişmiş olmasından kaynaklı olarak yetersiz kaldığı, hem de bu modellemenin tahmin edilenden çok daha karmaşık bağlantıları gerektirdiği görülmüştür. Netice olarak, nörobilimsel modelleme çalışmaları, fare beyni üzerinde yoğunlaşmış ve fare beyninin küçük bir parçasının nöral haritası tamamlanabilmiştir.

Kapsamlı (Küresel) Çalışma Alanı, Bilincin Nöral Korelasyonları ve Yüksek Bilgi Birleştirme Sistemleri gibi nörobilimsel temellere sahip bilişsel bilinç teorilerini modelleme arzusundan doğan üçüncü grup makine bilinci araştırmaları, insan bilincinin yapısının yeniden üretilmesi inancına dayanmaktadır (Gamez, 2008: 890). Bu gruptaki araştırmalar hem bilişsel hem de davranışsal temelli makine bilinci araştırmalarıyla uyumludur ve her ikisini de zorunlu olarak gerektirmektedir.

4.1.3.1. Kapsamlı (Bütünsel) Çalışma Alanı Modelleri[43]

Beyindeki kapsamlı bilgi sürecini dikkate alan bilincin nöral korelasyonları üzerine olan modellere kapsamlı çalışma alanı modelleri ismi verilmektedir. Geleneksel bilişsel yöntem, yalnızca zihnin modüler yapısına odaklanırken; kapsamlı çalışma alanı modelleri

[43] *Global Workspace Models*

önce bilginin modüller tarafından nasıl işlendiğini, daha sonra merkezi bilinç sistemine nasıl girdiğini açıklamaya çalışmaktadır. Kapsamlı çalışma alanı, modüllerin çıktılarını gönderdikleri bilişsel bir yapıdır (Revonsuo, 2016: 310). Bilinçli zihinsel aktivite sırasında geniş olarak dağıtılmış olan beyin aktiviteleri artmakta ve beyin zarının bölgeleri arasındaki iletişim daha iyi gözlemlenebilmektedir. Bu alandaki çalışmalar, Baars'ın Kapsamlı Çalışma Alanı teorisinden etkilenmişlerdir (Baars, 1988, 2002). Bu teori, insan beynini özelleştirilmiş otomatik işlemcilerin (duyum, motor kontrol, dil, muhakeme vs. sağlayan) organize bir ağı olarak kabul etmektedir (Reggia, 2013: 116). Beyindeki süreçlerin birçoğu beynin belli bir bölgesinde gerçekleşen, sınırlı ve bilinçsiz süreçlerdir. Bu süreçlere ek olarak, beyin zarında bulunan ve içeriği özelleştirilmiş işlemcilere uygun olan bir kapsamlı çalışma alanı mevcuttur. Özelleştirilmiş işlemciler, kapsamlı çalışma alanlarına erişmek için yarış içerisindedir, zira bu erişim onlara bilgiyi gönderme ya da alma şansı vermektedir (Reggia, 2013: 116). Özelleştirilmiş işlemciler olarak adlandırılan modüllerin ürettiği çıktılardır. Kapsamlı çalışma alanı teorisine göre, bilincin temelinde yer alan nöral sistem, yani "kapsamlı nöral çalışma alanı" sabit değil dinamik, nöral sistemin işlemine göre sürekli değişiklik gösteren bir sinir ağı olarak düşünülür (Tura, 2017: 33). Bu model, bilincin kapsamlı çalışma alanı ile çıktıların ortaklaşa etkileşimlerinden meydana geldiği hipotezini kabul etmektedir. Modüler çıktılar, gönderildikleri ağ içerisinde baskın bir konum elde edebilmek için ya da kuramsal tabirle "kapsamlı erişim" için birbirleriyle yarışırlar ve yarışmayı kazanan çıktı mesajını kapsamlı olarak yayarak bilincin içeriğine katılır (Revonsuo, 2016: 310). Talamus ve beyin kabuğu Baars'ın kapsamlı çalışma alanı teorisinin en çok üzerinde durduğu beyin bölgeleridir.

Baars teorisini, bilinci açıklamak için sıkça başvurulan bir yöntem olan tiyatro metaforuyla aktarmaya çalışmaktadır. Zihin bir tiyatro sahnesidir. Bilinç, bu sahnede üzerine ışık tutulan şeydir. Sahnedeki geri kalan her şey zihnin bilinçdışı içerikleridir. Seyirciler, bilinçten bilgi alan zihnin bilinçdışı bileşenlerini oluşturmaktadırlar. Senarist, yönetmen, suflör gibi sahne arkasındaki insanlar da bilincin içeriklerini belirlemekle ve düzenlemekle görevlidirler.

Baars'ın teorisinden farklı bir başka kapsamlı çalışma alanı modeli de Stanislas Dehaene (1998) ve Lionel Naccache (2001) tarafından "Bütünsel Nöral Çalışma Alanı" adıyla geliştirilmiştir. Nöral kapsamlı çalışma alanları modellerine dayanan bu çalışmaların çoğu biyolojik olarak gerçekçi bir tutum içerisindedir. Bu modele göre, yaşanılan bilinç durumu,

bilginin bütünsel olarak paylaşımıdır. Bu model, bilinçli durumların oluşmasını şöyle açıklamaktadır:

> Beyin, her biri tek bir işlem tipinde uzmanlaşmış olan düzinelerce yerel işlemci içerir. Özel bir iletişim sistemi olan "bütünsel çalışma alanı", bu işlemcilerin esnek bilgi paylaşımına imkân sağlar. Çalışma alanı herhangi bir anda işlemcilerin bir alt kümesini oluşturur, isteğe bağlı bir süre akılda tutup diğer işlemcilerden herhangi birine geri gönderir. Çalışma alanına ulaşan her bilgi parçası bilinçli hâle gelir (Dehane, 2018: 213).

Bütünsel Nöral Çalışma Alanı modelinin temelinde yatan düşünce, zihnimizde bilinçli olarak farkına varılan tek bir öğe olsa bile zihnin aynı anda çalışan çok sayıda bilinçdışı işlemciden oluştuğudur. Bilinçli durumların ayırt edici özelliği bütünsel çalışma alanına ulaşmış olmalarıdır. Baars'ın tiyatro metaforunu geliştiren Dehane de insan zihnini tiyatro sahnesiyle karşılaştırarak açıklamaktadır:

> Yer ışıklarının olduğu alan çok dardır ama arkaya gittikçe genişler. Yer ışıkları hizasında birden fazla oyuncuya yer yoktur. Kişi yer ışıklarından uzaklaştıkça, ışığı daha uzak olduğu için silikleşen diğer figürler vardır. Bu grupların ötesinde, kuliste ve tüm arka planda, beklenmedik bir çağrının öne çıkarabileceği ve hatta doğrudan yer ışıklarının menziline sokabileceği sayısız karanlık şekil bulunur. Antik bir projektörden yansıyan görüntünün gözümüzün önünden akıp gitmesi gibi, her türlü oyuncunun bulunduğu kalabalık arasından koro şefini ortaya çıkarmak üzere sürekli tanımsız evrimler gerçekleşir (Dehane, 2018: 214).

Bu zihin modelinde bilincin görevi, bilgiyi bütünsel olarak erişilebilir hâle getirmek ve beyin sistemlerine esnek şekilde yayınlamaktır. Buna göre, bilinci açıklamak için "Dışarıdan veya içeriden gelen bir bilgi parçasının nasıl bilinçdışı işleme sürecinin ötesine geçip bilinçli işleme sürecine erişim sağladığını, böylece raporların bir öznel deneyimin varlığıyla nitelenen bir geçiş gerçekleştiğini göstermemiz gerekir." (Deheane, Changeux ve Naccache, 2011: 56). Bu modelin çıkış noktasında, başarılı bir bilinç teorisinin nörolojik ve psikolojik veriler arasındaki gediği aşan bir köprü oluşturması gerektiği fikri yatmaktadır. Bütüncül bir bilinç teorisinin yapması gereken nesnel nörofizyolojik veriler ile öznel raporlar arasındaki nedensel ilişkiyi açıklamaktır (Deheane, Changeux ve Naccache, 2011: 55). Bu modelde, bizim öznel deneyimimize bilinçli erişimimiz, dikkat çekiciliğine ve hedefe uygunluğuna göre seçilen bir bilgi parçasının beynin farklı ve uzak bölgelerine seçimi, eklenmesi ve dağıtılmasından başka bir şey değildir. Dolayısıyla öznel deneyim -ilerleyen sayfalarda incelenecek olan bilgi birleştirme modelinde olduğu gibi- bilgi işleme sürecinin bir ürünü olarak ele alınmaktadır.

Deheane, Changeux ve Naccache tarafından ortaya konulan bu model de Baars'ın modeline benzer olarak otomatik beyin aktivitelerine ait nöral aktivite örnekleriyle bilinçli süreçlerin örneklerinin karşılıklı analizleri üzerine inşa edilmiştir (Reggia, 2013: 117). Bu modelin yapısı her alanın bir modeli ya da kapsamlı çalışma alanını temsil ettiği korkteks alanları ile belirlenmiştir. Belirli korteks tabakalarında bulunan, kortekste yaygın bağlantılar kuran kortikal çalışma alanı nöronlarına dayalı bu modele göre, söz konusu tabakalarda gerçekleşen aktiviteler bilinçle bağıntılıdır (Revonsuo, 2016: 312). Sadece kortekse odaklanan bütünsel nöral çalışma alanı modeli, Baars'ın klasik kapsamlı çalışma alanı teorisinden farklı olarak talamusa bilinçli içeriklerin üretimi hususunda hiçbir rol vermemektedir.

Deheane'ye göre bilgisayar gibi biyolojik olmayan bir sistemde bilince yol açan fonksiyonların üretilmesinin önünde teorik olarak hiçbir engel yoktur. Kapsamlı/bütünsel çalışma alanına sahip bir sistemin -silikondan yapılmış olsa dahi- bilinçli durumlara sahip olması mümkündür. Ancak bilinçli durumları oluşturan işlemlerin uygulaması teorikte göründüğünden çok daha karmaşıktır. Zira henüz beynin bu işlemleri nasıl gerçekleştirdiği ya da bir makinenin bu işlemleri ne şekilde gerçekleştirdiği bilinmemektedir. Bilgisayar yazılımları genel olarak modüler tarzda düzenlenme eğilimindedir: Özel girdiler alır ve iyi tanımlanmış çıktılar üretmek üzere bu girdileri kurallara bağlı olarak dönüştürür (Dehane, 2018: 217). İşlemciler, bir bilgi parçasını istenilen sürede saklayabilir ancak bilgisayar bir bütün olarak, bu bilgi parçasının uygunluğuna karar verecek ya da bilginin diğer programlara serbest bir şekilde erişimini sağlayacak bir araca sahip değildir. Modüler yapılı bir bilgisayarın bir modülünde bilinen bilgi, diğer modüller tarafından bilinmemekte ve paylaşılmamaktadır. Beyin korteksi ise bilgisayarlardan farklı olarak, modüler işlemci kümesine ve esnek yönlendirme sistemine aynı anda sahiptir ve bilginin paylaşılması problemi insan zihni için geçerli değildir. Bütünsel nöron çalışma alanı ile bilgi, beyindeki tüm modüller arasında serbestçe paylaşılabilmektedir. Bütünsel nöral çalışma alanı modeline göre, bilginin bu şekilde bütünsel olarak kullanılabilir olması tam olarak yaşadığımız öznel bilinç durumudur (Dehane, 2018: 213).

Özet olarak, bilgisayar simülasyonları şu an için bilinçli erişimin bazı özelliklerini kopyalasa da bütün olarak beyni taklit etmekten oldukça uzak görünmektedirler. Fakat kapsamlı çalışma alanı modellerinin iddiası şudur ki, çok daha fazla nöral durumu dikkate alan daha iyi bir simülasyon prensipte beyni tümüyle taklit edebilir ve beynin bütün olarak taklidi ile

oluşturulacak olan kapsamlı çalışma alanı, bir bilgisayarın öznel bilinçli durumlara sahip olması için yeterli koşulu sağlamış olur.

Pratikte bilincin oluşturulmasını uzak bir hedef olarak değerlendiren kapsamlı çalışma alanı modelleri simülasyon olarak bilinç üretmek için çalışmakta, kapsamlı çalışma alanını insan bilincinin nöral korelasyonlarını keşfetmek için kullanmaktadırlar. Bilinci somutlaştırmak gibi bir hedefe sahip değildirler. Nöropsikoloji ve son bilimsel verilerle uyumlu olan bu modeller, bilincin hakkında parçalı olsa da bir yaklaşım sunmakta ve bilinçli aktiviteleri bilinçsiz aktivitelerden ayıranın ne olduğunu belirlemeye yönelik net bir öneride bulunmaktadırlar (Reggia, 2013: 117). Bu ölçüt bilinçli aktivitelerin kapsamlı, bilinçsiz aktivitelerin bölgesel olduğunu söylemektedir. Fakat bu modeller sınırlılıklar olmadan uygulanabilir değildir. Kapsamlı bilgi süreçlerinin neden bilincin nöral korelasyonu olduklarını sorgulamaktan uzaktırlar. Bilincin *kolay problemlerine* kullanışlı bir yöntem sunmaktan öteye gidemezler çünkü *zor problem* bu modellerin araştırma alanı içerisine dâhil değildir.

4.1.3.2. Nöral Korelasyonlar

İlk olarak 1990 yılında Francis Crick and Christof Koch tarafından kullanılan bilincin nöral korelasyonları, beyindeki belirli olaylar ve fenomenal deneyimin devam eden dinamiklerindeki belirli olaylar arasındaki korelasyonu araştırmaktadır. Bilincin nöral korelasyonları tezini ortaya atan Francis Crick'in temel iddiası şudur:

> Sen, senin keyif üzüntülerin, anı ve tutkuların, sahip olduğun kişisel kimlik hissi ve özgür iraden aslında bir araya gelmiş çok sayıda sinir hücresinden ve onlarla ilişkili moleküllerin davranışından başka bir şey değil! (Crick, 1994: 3).

Bu teze göre, zihinsel olan beyinsel olan ile eş değerdir. Tüm zihinsel durumlarımız, beynin içinde gerçekleşen fiziksel durumlardır. Buna bağlı olarak, nöronlar ve onlarla ilişkili nörotransmitter moleküller beynimizin içinde yer alan zihinsel durumlarımızdan sorumlu özgül mekanizmalardır (Searle, 2018: 28).

Bilincin nöral korelasyonu "Aktivitesi doğrudan bilinçli deneyim durumlarıyla korelasyon oluşturan özel bir sistemdir." (Chalmers, 2000: 18). Bu sistem empirik anlamda, nörobilimsel analizle bir organizmanın zihnindeki bilinçli içeriği aktive eden özelliklerin bir kümesini oluşturmaya çalışmaktadır. Bilinçli durum ve deneyimlerimizi nöronların

davranışı ile açıklayan bu yaklaşımda, bilinçli zihinsel durumlar nöronların işleyişi ile ortaya çıkan ya da beliren niteliklerdir. Bu yaklaşımın kullandığı temel yöntem, bilinçli deneyimimiz hariç her şeyin aynı olduğu iki durumda nöronlarımızın neler yaptığını karşılaştırmaktır. Beynin farklı davrandığı gözlemlenen kısımları bilincin nöral korelasyonları olarak tanımlanır. Bu bağlamda, bilincin nöral korelasyonlarını tespit etmek bilinç bilimi için önemli bir köşe taşı niteliğindedir. Bağıntılar, bilincin bilimsel çalışmasını kolaylaştıran bir yöntemdir. Ayrıca, bağıntıların keşfi üzerine çalışmak kuram-yüklü bir metot değildir; hiçbir ön kabule gereksinim duymadan kuramdan bağımsız salt gözlem ve deneye dayalı araştırmalardan oluşmaktadır. Bilince objektif bir bakış açısına izin vermesi, nöral korelasyonları bilinç bilimi için vazgeçilmez kılmaktadır.

Bilince yönelik bilimsel çalışmaların kayda değer bir kısmını bilinç ve bilincin altında yatan süreçler arasındaki ilişkinin mahiyeti oluşturmaktadır. Bilincin nöral korelasyonları üzerine çalışanlar bilincin beyin ile olan ilişkisini çözümlemeyi hedeflemektedir zira bilincin alt katmanının biyolojik beyin olduğunu varsaymaktadırlar. Beyin süreçleri ile bilinçli süreçler arasındaki bağıntıları tespit etme uğraşı, bilince yönelik bilimsel araştırmanın temel yöntemidir. Nöral korelasyon kavramı, "aktivasyonu bilinçli algıya yol açan minimal ve yeterli nöral sistem" olarak tanımlanmaktadır (Chalmers, 2000: 31).

Nöral korelasyon araştırmaları genel olarak bilinçli ve bilinçsiz benzer iki zihin durumunun karşılaştırmalı incelemesi üzerine kuruludur. Uyanık olmak ve uyku karşılaştırılması sıkça üzerinde durulan bir analizdir. Bu analizin sonuçlarında, bilinç genel olarak beynin talamus ve beyin kabuğu bölgesiyle ilişkilendirilmektedir; üst beynin ortasında yer alan yukarı uyarılma sistemi ve talamustaki çeşitli çekirdekler beyin loblarının aktivasyonunu sağlamaktadırlar (Reggia, 2013: 115). Beyin hasarlarında oluşan tablolar göstermektedir ki; beyin yarımkürelerinin aktivasyon modeli bilincin içeriğiyle doğrudan bağlantılıdır. Koma

durumundaki ve genel anestezi altındaki hastaların fMRI[44], EEG,[45] TMS[46] ve PET gibi tekniklerle görüntülenen nöral aktiviteleri de bu hastalarda pre-frontal ve yan kafa kemiği bölgelerinde çok daha az metabolik faaliyet olduğunu göstermiştir. Beynin hangi kısımlarının bilinçten sorumlu olduğuna dair tam bir beyin haritası henüz çıkarılamamıştır. Bilincin hiçbir kısmının beynin kökünde ve omuriliğe bağlanan beynin alt bölgesinde yer almadığı bilinmektedir; ayrıca bilinç, nöronların üçte ikisine sahip olan beyincikte de görüntülenmemiştir (Tegmark, 2019: 378). Son nöron bağıntısı araştırmaları bilincin talamus ve korteksin arka kısmını içeren sıcak bölgede olduğunu ileri sürmektedir. Ancak bu iddia da üzerinde uzlaşıya varılmış konumda değildir.

Bilincin varlığı ve yokluğu durumlarında beynin elektriksel aktiviteleri de değişiklikler göstermektedir. Elektrokimyasal aktiviteler ve kuantum fiziksel süreçler de yine beynin nöral bağıntıları üzerine olan araştırmalarda dikkate alınan konulardır. Biyolojik nöron aktivitelerinin bilinçle ilişkili olduğu, derin uyku/koma ve uyanık durumdaki sinirsel aktivitelerin karşılaştırılmasıyla çok kez gösterilmiştir. Bu konuda oluşmaya başlayan birlik ise bilincin sinirsel bağıntılarının beynin birçok farklı bölgelerinde dağıldığı yönündedir. Bu bölgeler arasındaki koordinasyon ise sinirsel ateşleme senkronizasyonuyla, NMDA sinapslarıyla, talamus ve çekirdek bağlantılarıyla ve bu mekanizmaların bazı kombinasyonlarıyla başarılır görünmektedir (Gamez, 2009: 217).

Christof Koch'un nöral korelasyonlar üzerine araştırma yaparken ulaşmış olduğu bir olay ile onun bilincine varmamız arasında saniyenin dörtte biri kadar bir süre olduğu fikri de bilinç araştırmaları açısından oldukça önemlidir (Tegmark, 2019: 379). Bu fikir, beynin duyu organlarından gelen bilgiyi işlemesi için zamana ihtiyaç duyduğunu ortaya koymaktadır.

[44] Bu tekniğin başlangıç noktası, kişinin kuvvetlice titreşen bir manyetik alana yerleştirilmesini içeren bir tarama yöntemi olan manyetik rezonans görüntülemenin (MRI) icadıdır. Bu manyetik alan, vücuttaki hidrojen atomlarını uyararak, dedektörün aldığı radyo sinyalinin oluşmasını sağlar. MRI, beyin yapılarının üç- boyutlu X-ışını görüntülerini sağlar. Söz konusu görüntüler, hasar gören bölgelerin tanımlanması yönünden çok önemlidir. MRI'ın statik görüntülerini, beynin dinamik hareketlerine dönüştüren gelişme ise 1990'larda geliştirilen fonksiyonel manyetik rezonans görüntülemedir (fMRI). Beyin, kanda taşınan oksijenin büyük bir bölümünü kullandığından, güçlü manyetik alanlar kullanarak oksijenin de radyo dalgaları oluşturması sağlanabilir. Kandaki oksijen miktarı, fMRI tarafından ölçülür. Beyin her zaman aktif olduğu için, deneylerde zihinsel görev sırasında ve dinlenme esnasındaki kişinin beynindeki kan akışı ve oksijen kullanımı karşılaştırılır. Zihinsel aktivite esnasında beynin herhangi bir bölgesinde kan akışı artarsa, bu bölgenin söz konusu zihinsel aktivite için önemli olduğu varsayılır (Rose ve Rose, 2018: 29-30).

[45] Yirminci yüzyılın ortalarından beri kullanılan bu teknoloji beynin nöral aktivitelerinin oluşturduğu elektrik dalgalarının kaydedilmesi esasına dayanır (Tura, 2017: 30).

[46] Beynin elektromanyetik dalgalarla uyarılması esasına dayanır. Beynin çalışmasını çeşitli şekillerde geçici olarak bozmak suretiyle beyin fonksiyonları hakkında dolaylı olarak bilgi edinilir (Tura, 2017: 30).

Buna göre, bilinç dış dünyayı çeyrek saniye geriden takip ederek, geçmişte yaşamaktadır. Fakat gözümüze bir sinek yaklaştığında hemen gözümüzü kırpmamız gibi bazı şeylere onlara bilincinde olamayacağımız kadar kısa sürede tepki veririz, bu da en hızlı tepkilerimizden sorumlu bilgi işlemenin bilinçsiz olduğunu gösterir (Tegmark, 2019: 380).

Nöral korelasyonlar üzerine çalışan araştırmacıların inancı şudur ki; insan beyni zamanla daha çok bilinçli davranış ve fonksiyon hakkında incelendikçe bilincin daha çok potansiyel bağıntısına ulaşılacaktır. Ancak bilincin nöral korelasyonları üzerine çalışanların uygulayacakları deneyler oldukça kısıtlıdır. Zira beynin materyal özelliklerinin ölçümü kısıtlıdır ve diğer kimyasal, fiziksel, biyolojik ölçümlerden farklı olarak moleküler ve atom-altı düzeydedir (Gamez, 2009: 218). EEG, fMRI, PET gibi var olan tarama ve görüntüleme teknikleri hem zamansal ve mekânsal duyarlılık konusunda hem de çok sayıda nörona ulaşma konusunda yeterli değillerdir. Bilincin işlevsel ve bilişsel korelasyonlarını ölçen güncel yöntemler, psikoloji testleri uygulamalarını, beyin hasarlı hastaları incelemeyi ve beynin devre dışı parçalarına transkraniyal manyetik stimulasyon uygulamayı kapsamaktadır (Gamez, 2009: 218). Görüntüleme ve tarama yöntemlerinden kaynaklı ölçüm kısıtlamaları teknolojik gelişmelerle birlikte aşılabilecek görünümde olsa da beynin atom-altı özelliklerinin bilinç için gerekli olup olmadığına karar vermek için gerekli zamansal ve mekânsal duyarlılığın elde edilmesi zor görünmektedir.

Bilincin nöral bağıntıları üzerine yapılan deneylerin daha temel bir kısıtlaması ise bilinçli durumların değerlendirmesinin dil ve hafızaya bağlı olmasıdır. Eğer bir deney öznenin hatırlama ve/veya raporlama yeteneğini değiştiriyorsa deneyin sonuç tanımı değiştirilmiş fenomenolojinin veya zarar görmüş raporlama özelliğinin bir sonucu olabilir; bu durumda hangisinin gerçek durum olduğuna karar vermek imkânsız hâle gelebilir (Gamez, 2009: 218). Davranış, bilinci ölçmek için şu an bilinen tek yoldur. Bilincin sinirsel bağıntılarını ölçmek için davranıştan bağımsız ve kişinin bilinçli durumlarını aktaran raporlarıyla çatışmayan deneyler uygulamak gereklidir ancak bu, her durumda imkân dâhilinde değildir.

Nörobilim şimdiye kadar bilincin nöral korelasyonları üzerine çok önemli araştırmalar yürütmüş ve beynin, bilinçle ilgili olan kısımlarının önemli bir bölümünü aydınlatmayı başarmıştır. Birçok duyu deneyimi esnasında beynin hangi bölgelerinin aktif olarak çalıştığı ve bu aktivitelerin doğası büyük ölçüde tespit edilmiştir. Ancak nörobilim bu yükselişine rağmen, bilincin nöral korelasyonlarının keşfinin, bilincin kendisinin keşfi olmadığı yönünde sıkça eleştirilmektedir. Bir şeyin deneyimi esnasında oluşan nöral aktivitenin, o

deneyimin kendisi olmadığı ve nöral aktivitenin deneyime nasıl yol açtığının gösterilemediği, nörobilimin bilincin zor problemine bir açıklama getiremediğini gerekçelendirmek üzere ileri sürülen sebeplerdir. Nörobilim "Fiziksel, nesnel, nicel olarak tanımlanabilen nöron ateşlemeleri nitel, şahsi ve öznel deneyimlere nasıl neden olabilmektedir?" sorusunu açıklamakla uğraşmamaktadır; başka bir deyişle, bu yaklaşımda beynin elektrokimyasal akitivitelerden deneyimlemeye geçişteki boşluğu nasıl kapattığı sorunsalı cevaplanmadan bırakılmaktadır (Searle, 2018: 33).

Bu eleştiriye ek olarak, deneyimin kendisinin beyin dışında da bağlantıları olduğu duyu organları ile deneyimlenen nesnenin kendisi (ışık, ses, tat vs.) ve bu bağlantılar göz önüne alınmadan yalnızca beyne odaklanmanın eksik bir yaklaşım olarak kalacağı da dile getirilen başka bir karşı çıkıştır (Manzotti, 2018: 16). Nöral bağıntıları ve beyni, deneyimin diğer bağıntılarından üstün ve ayrıcalıklı bir konumda ele almanın henüz bilimsel ve rasyonel bir açıklaması yapılamamıştır. Yalnızca beyin üzerine araştırma yapmak, deneyimin açıklamasını eksik bırakacaktır zira deneyim, yalnızca deneyimleyen ile değil, deneyimlenen ve deneyime imkân tanıyan duyularla beraber gerçekleşmektedir. Bu hususta, bilincin nöral korelasyonları, deneyimin yalnızca tek boyutuna odaklanarak tam bir bilinç tanımı yapamamaktadır.

4.1.3.3. Bilgi Birleştirme Modeli

Bilinç bilimi, deneyimin davranışsal ve nöral bağıntılarının incelenmesi suretiyle büyük ilerleme kaydetmiştir. Ancak bu bağıntıların araştırılması, hâlâ çok temel soruların cevaplanması için yeterli olmamaktadır. Henüz hiçbir teori, bilincin neden beyin zarı (*celebral korteks*) ile ilişkili olmasına rağmen, beyin zarından daha fazla nörona sahip ve daha karmaşık bir yapıda olan beyincik (*cerebellum*) ile ilişkili olmadığını açıklayamamaktadır. Bilgi Birleştirme Teorisi, bu temel soruların açıklanması için nöral sistem hakkında daha detaylı bilginin yanı sıra, deneyimin ne olduğunu ve nasıl bir fiziksel sistemin deneyime sahip olabileceğini açıklayan bir bilinç teorisine ihtiyaç olduğu fikrine dayanmaktadır. Buna göre, eğer beynin ya da bilgisayarın bilinçli bir kısmı geri kalanıyla iletişim kurmuyorsa kalan kısım öznel deneyime ait olamaz (Tegmark, 2019: 385). Bu teori, deneyimin kendisini tanımlayan beş fenomenolojik aksiyom üzerine kuruludur: "İçsel varoluş, kompozisyon, bilgi, birleştirme ve dışarda bırakma." (Tononi ve Koch, 2015: 1).

Bu teori, bireysel deneyimin hem niteliğini hem de niceliğini dikkate almakta ve fiziksel bir sisteminin bilinçli olup olmadığını ölçecek bir hesaplama yöntemi önermektedir. Bilgi Birleştirme Teorisine göre, bilinç nedensel özelliklere sahip fiziksel sistemlere ait temel bir niteliktir ve biyolojik organizmalarda dereceli olarak bulunur. En basit canlılardan, en karmaşık organizmalara kadar her biyolojik sistem belli bir derecede bilince sahiptir. Ancak bu yaklaşıma göre, ileri beslemeli sistemler ve dijital bilgisayarlar ne kadar komplike olurlarsa olsunlar, bunların davranışları bizim davranışlarımıza ne kadar benzerse benzesin ve insan beynini ne kadar iyi taklit ederlerse etsinler, deneyim sahibi olmalarından ve dolayısıyla bilinçli olmalarından söz edilemez.

Bilgi birleştirme modelinin bilinç açıklamasında bilgi süreçleri ve bilgi birleştirme işlemi merkezi bir rol üstlenmektedir. Beynin bölgeleri arasında paylaşılan bilgi, bireysel alanların bilgi içeriğinin üstünde ve dışındadır. Bu bilgi birleştirme işleminin, bilincin nöral bağıntılarının bir sonucu olduğu kabul edilir. Bilince yönelik nörobiyolojik veriler bilgi birleştirme işlemi ile tercüme edilir. Bilinç, sistemin bilgiyi tek bir deneyimde birleştirme kapasitesidir (Reggia, 2013: 118). Bu model, özel bir ağın bilgi birleştirme yeteneğini niceliksel olarak ölçebilmeyi sağlar. Niceliksel ölçüm, bilincin bilimsel değerlendirmesi açısından nesnel ve yararlı bir ölçüttür.

Bilgi birleştirme Giolio Tononi tarafından "birlik" problemine bir yaklaşım olarak geliştirilmiştir (Tononi, 2004). Ona göre bilgi işleme süreçlerinin belli yönleri birbirlerine içsel olarak bağlıdır. Çok sayıda bilinçli deneyimin aynı anda var olabilmesi, her deneyimin birliği gibi önemli bilinç özelliklerini doğrular. Tononi, "φ" ile gösterdiği, bir sistemin farklı kısımlarının birbirleri hakkında bildiklerinin niceliksek değerini ölçen ve entegre/birleşik bilgi adını verdiği bir ölçü birimi teklif etmiştir (Tegmark, 2019: 389). Ona göre, bilgi işlemenin birleşmiş olması yani PHI'NİN büyük olması gereklidir. Ancak bu ölçüm, hacim olarak küçük ağlar içerisinde mümkün olabilmektedir. Beyindeki tüm nöral süreçlere hâkim olmamız söz konusu değildir çünkü beyin çok sayıda geniş nöral ağa sahiptir. Dolayısıyla bugün eldeki verilerle bu modelin tek başına bilincin varlığına ya da yokluğuna dair hükümde bulunması mümkün değildir. Gelecek araştırmalarda daha büyük ağlarda tam olmasa da yaklaşık nicelikler tespit edilerek, modelin daha kullanışlı hâle gelmesi ve kuantum bilgi teorisine dâhil edilmesi sağlanabilir.

Bu modelin dayandığı tezin kabulü, bilinci gerçek bir fenomen olarak değil, ölçülebilen ve derecelendirilebilen bir nicelik olarak görmeyi gerektirir. Bu tez, sistemin öznel deneyimini

onun bilgi birleştirme kapasitesine eşitler; sistemin bileşenlerinin bilgisel ilişkilerini qualia alanı olarak tanımlar (Reggia, 2013: 119). Doğal ya da yapay bir sistemin bilinç düzeyini objektif olarak ölçme şansı vererek makine bilinci teorilerini değerlendirmek için bir araç olarak faydalı bir model olsa da bilgi birleştirme kapasitesini zeka ile değil de öznel deneyim ile eşleştiriyor olması eleştirilebilir bir teoridir.

4.1.4. Fenomenal Makine Bilinci ve Sentetik Fenomenoloji

Makine bilinci çalışmalarına en büyük eleştiri fenomenal bilinç ve qualia problemi olarak adlandırılan, bir deneyiminin neliğinin asla bir makine tarafından hissedilmeyeceği iddiası üzerine kuruludur. Bu iddianın temeli ise bir makinenin herhangi bir olguyu insandan daha iyi idare etse dahi asla neyi idare ettiğini anlamayacağı fikrine yaslanmaktadır; buna göre makineler, algıları insanlardan daha iyi işleseler dahi gerçekten bu algıların özlerini yahut doğalarını hissedemez ve deneyimlemezler (Kaku, 2017: 281). Bir makinenin asla fenomenal bilince sahip olamayacağını iddia eden bu bakış açısına göre, kulakları duymayan bir insana kuş sesi işitme deneyiminin öğretilmesi, gözleri görmeyen bir insana yeşil renk deneyiminin öğretilmesi veya tat almayan bir insana çikolata tadı deneyiminin öğretilmesi nasıl imkânsızsa bir robotun da kuş sesinin, yeşil rengin ya da çikolatanın öznel hissini deneyimlemesi o derece imkânsızdır.

Fenomenal makine bilinci araştırmalarına karşı çıkan bir başka grup ise makinelerin zekasının, insan zekasını geçtiği noktada o makinenin öznel deneyimlere sahip olup olmamasının hiçbir önemi kalmayacağını savunmaktadır. Bu bakış açısına göre ise gelecekte makineler yeşil rengin, kuş sesinin ya da çikolata tadının hissini fiziksel verilerden yola çıkarak da olsa insanlardan çok daha iyi bilecek ve işleyebileceklerdir. Bu aşamaya gelindiğinde "Bir makine yeşil rengi hisseder mi?" sorusu anlamsız bir soruya dönüşecektir (Kaku, 2017: 281). Makineler işlevsel olarak bizim öznel hislerimizi ve onlara ait kelimeleri kullanma kabiliyetimizi aştıklarında, onların bu his ve kelimeleri anlamaları, deneyimlemeleri ya da hissetmeleri onların zeka seviyeleri üzerinde etkili olmayacak ve bu tartışmanın gereksizliği ortaya çıkacaktır. Bu yaklaşıma göre, bilginin kullanılabilmesi anlaşılmasından çok daha önemlidir. Aynı yaklaşım Çince Odası'ndaki anlama temelli argümana da karşı gelmiş olmaktadır. Zira bir makine ya da bilgisayar simultane bir şekilde Çinceden başka bir dile tercüme yaptığı gün, o makinenin Çinceyi gerçekten anlayıp

anlamaması işlevselliğinin yanında ehemmiyetsiz kalacaktır. Bu görüşün savunucuları güçlü yapay zeka için bilinci ve özellikle fenomenal bilinci gerekli bir koşul olarak kabul etmediklerinden, fenomenal makine bilinci çalışmalarının boş bir uğraş olduğu konusunda hemfikirdirler.

Hem makine bilinci çalışanlarının kendi içlerinde uzlaşıya varamadığı bir alan olarak hem de bilinç konusu içerisindeki en gizemli alan olarak fenomenal bilincin, karşı çıkışlara rağmen bir grup makine bilinci çalışanı tarafından kilit bir noktada olduğu kabul edilmektedir. Bu grubun dayandığı temel düşünce ise fenomenal bilince sahip olmayan bir makinenin, bilinçli olarak adlandırılamayacağıdır. Bütüncül bir bilinç teorisinin fenomenal bilinci de açıklamasının zorunlu olduğu fikrinden yola çıkarak, bütün olarak bilinçli bir makinenin fenomenal bilince sahip olmasını zorunlu bir şart olarak kabul etmektedirler.

Fenomenal makine bilinci çalışanlarına göre, nörobilim ve bilişsel bilim bilincin doğasını açıklamak için yeterli değildir; işlevselcilik, hesaplamacılık ya da temsilciliğin farklı şekillerde işlenişinden oluşan bu alanlar fenomenal bilinç için oldukça sınırlıdırlar. Zira odaklandıkları hususlar, şimdilik bilinen berimsel ve sinirsel mekanizmalardan ibarettir ve bunlar fenomenal bilinci kavrayamamaktadır. Bütünüyle bilinçli yapay bir sistem oluşturmanın mutlak koşulu fenomenal bilinci birinci şahıs bakış açısından kavrayan açıklamalar ve tanımlar üretmektir. PET, EEG, fMRI gibi görüntüleme yöntemleri fenomenal bilinci saptamak için yeterli ölçüm teknikleri değildir. Bu tekniklerle beyinde görüntülenen bilincin kendisi değildir. Bilincin kendisi hiçbir işlevsel beyin görüntüsünün içerisinde yoktur. Nöron bağıntıları yaklaşımı hiçbir zaman bilincin nöral bağıntıları ile bilincin gerçek nöral bileşenleri arasındaki ilişkiyi açıklamamaktadır (Revonsuo, 2000: 60).

Fenomenal makine bilinci çalışanlarına göre, insan ve evren arasındaki ilişkinin sadece insan ve dış olaylar arasındaki ilişki olarak kavranması doğru bir yaklaşım değildir. Benzer olarak, bilinçli bir makine, yalnızca çevreyle etkileşim kurabilen bir makine anlamına gelmemektedir. Güneşi gören kirazın kızarması, soğuk havada suyun donması, sonbaharda yaprakların dökülmesinden oluşan bir fiziksel dünya içerisinde, insanın yaşantısını oluşturan bu fiziksel olayları (güneşi, soğuk havayı, baharı vs.) deneyimlemesidir. Fiziksel dünyayla etkileşimin deneyimi, fenomenal bilincin varlığına yol açmaktadır. Dolayısıyla bu yaklaşımı benimseyen araştırmacılar, fiziksel etkileşime deneyimin eklenmesini yapay bilincin zorunlu ön koşulu olarak kabul etmektedirler.

Günümüzde insanlarla etkileşimde bulunan robotlar üretilebilmektedir. Robotların bu etkileşimi verili davranış kümelerinden seçim yapmaktan başlayıp, bağımsız ve kendine göre eylemde bulunma kapasitesine kadar çeşitli seviyelerde gerçekleşebilir (Manzotti ve Tagliasco, 2008: 108). Robotlar, dokunma sensörleri, ses algılama modülleri ve anahtarlarla eylemleri öğrenebilirler. Sofistike robotlar, belirli bir çevreyle etkileşim yoluyla kendilerine özel bir kişilik geliştirebilirler. Robotlar ayrıca mutluluk, öfke, şaşkınlık, üzüntü gibi duygu durumlarını taklit yoluyla da olsa yansıtabilirler. Meraklı ve çevrelerini keşfedebilen robotlar, kendi kişisel tarihlerine göre gelişebilirler. Ancak henüz ulaşılamayan hedef, onlara ne olduğuyla ilgili öznel deneyime sahip olma kapasitesi olan robotların uygulamasıdır (Manzotti ve Tagliasco, 2008: 109). Bu hedef, bilişin bilinçli yanlarını anlamayı içermektedir ve bu sebeple tartışmalıdır. Öznel deneyimlere sahip olma yetisi olan bir robot, fenomenal bilince sahip bir robot olacaktır. Fenomenal bilince sahip olan bir robot inşa etmek ise mühendislerin şimdiye kadar yaptıklarından farklı olarak nesne değil özne inşa etmektir.

Güçlü yapay zekaya sahip olamayışımızın en temel nedeni fenomenal durumlar, özne ve bilince yönelik tanım ve anlayış eksikliğimiz olarak gösterilir. Fenomenal durumlar üzerine teori oluşturma çabaları hâlen devam ediyor olsa da işlevselliğinde uzlaşıya varılmış bir fenomenal durum teorisinden söz etmek zordur. Fenomenal bilinç, bilincin öznel ve niteliksel karakterdeki yönüne işaret eder. Bu yön, özeldir ve doğrudan genel ve özneler-arası bir doğaya sahip olan bilim tarafından açıklanmak için uygun değildir. Fakat bu, bilincin zorunlu ve yeterli koşullarının tanımlanamayacağı anlamına gelmemektedir; yalnızca fenomenal bilinç içeriklerinin tanımlanması ve deneyimlenmesinin farklı şeyler olduğunu ima etmektedir (Edelman ve Tononi, 2000: 140). Fenomenal makine bilinci çalışmaları bu bağlamda, fenomenal bilinç içeriklerinin doğru tanımına makineler üzerinden ulaşmayı hedeflemektedirler.

Bir makinenin, makine bilinci çalışmaları içerisinde bir yere sahip olması ve bu çalışmalara katkı sağlaması kendisinin bilinçli olması zorunlu bir koşul değildir. Makine bilinci alanı genel olarak yalnızca fenomenal durumlara sahip yapay zekalar üretmeyi hedeflemez, aynı zamanda insanlardaki ve diğer hayvanlardaki doğal fenomenal durumları anlamamıza yardımcı olan yapay zekalar üretmeyi amaçlar (Chrisley, 2009: 53). Fenomenal durumlar üzerine odaklanan makine bilinci araştırmaları, bu alanın en karmaşık ve felsefi anlamda en sorunlu alandır. Zira diğer araştırmalardan farklı olarak, bu grubun hedeflediği bizzat kendisi

bilinçli olan bir makine inşa etmektir. Bu gruptaki makineler bilinç alanın çalışmaları için bir araç değildir, aksine hedef bilinçli makineler tasarlamaktır. Diğer üç gruptaki makinelerin bütüncül ya da tek tek inşa edilmiş olmaları fenomenal anlamda bilinçli makinelere yol açabilir. Ancak fenomenal bilince sahip makinelerin oluşturulmasının bu üç gruptan bağımsız olarak gerçekleşmesi de muhtemeldir. Bu muğlaklığın en temel sebebi ise fenomenal bilincin doğasının felsefi ve bilimsel açıdan net bir açıklamasının yapılamamasıdır. Gerçek bilince ve fenomenal deneyimlere sahip makinelerin fenomenal durumları ölçen ve ayırt eden yöntemler belirlenmeden üretilmesi mümkün görünmemektedir. Şimdiye kadar hiçbir yapay bilinç teorisi, bir makinedeki fenomenal bilince işaret edememiş ya da fenomenal bilincin muhtemel olduğu bir sistemi kesin olarak belirleyememiştir.

Fenomenal deneyimi ölçme temelli kurulan sentetik fenomenoloji alanı ile bu gruptaki araştırmalar arasında sıkı bir etkileşim söz konusudur. Sentetik fenomenoloji tabiri ilk kez Scott Jordan'ın 1998 yılında yaptığı bir konuşmada fenomenal durumları sentetikleştirmek-yapaylaştırmak anlamında kullanılmıştır (Gamez, 2008: 904). Sentetik fenomenoloji çağdaş yorumunda yapay sistemlerin bilinçli durumlara sahip olup olmadıkları ve sahip olurlarsa ya da sahiplerse bilinçli durumların fenomenolojik tanımlarının nasıl yapılacağına kafa yormaktadır. Husserl'den miras kalan felsefi fenomenoloji insan bilincinin tanımına yönelikken sentetik fenomenoloji ise makine bilincinin tanımına yönelik bir araştırma alanıdır. Sentetik fenomenoloji alanı şu sorular üzerine kurulmuştur:

> Yapay bir sistemin bilinçli deneyimlere sahip olduğunu söyleyebilmemiz için hangi kriterleri sağlaması gereklidir? Bir yapay sistemin bilinçli deneyimlerini bilinçsiz deneyimlerinden ayıran ölçütler nelerdir? İnsana kıyasla yapay bir sistemin deneyimini tanımlamak daha mı kolaydır yoksa daha mı zordur? (Chrisley: 2009: 58).

Bilinçli olmayan bir sistemin fenomenolojisi hakkında konuşmak anlamsızdır. Bu sebeple sentetik fenomenoloji alanının ilk uğraşı fenomenal durumlara sahip sistemleri tanımlamak ve koşullarını belirlemektir. Bu uğraşa yönelik iki karşıt bakış açısı mevcuttur. Önemli bir çoğunluk fenomenal bilinçli sistemler ile fenomenal bilince sahip olmayan sistemleri ayırt etmeye yarayacak bir bilinç teorisinin varlığını zorunlu görmektedir. Alexander ve Morton bir sistemin sentetik fenomolojik durumlara sahip olmaya aday olması için iki zorunlu kıstas ortaya koymuştur:

1- Sistem, dünyanın ne olduğunu sistemin gözünden temsil eden ya da yansıtan bir mekaniğe sahip olmalıdır.
2- Sistem, sistemin kendi içinde neye benzediğini sistemin gözünden yansıtan bir mekaniğe sahip olmalıdır (Alexander ve Morton, 2007: 19).

Sentetik fenomenoloji araştırmacılarının küçük bir bölümü de fenomenal bilincin gerek ve yeter koşullarını belirleyemediğimizi ve bu yüzden bir sistemin fenomenoloji kapasitesi olup olmadığını hiçbir zaman bilemeyeceğimizi savunmaktadırlar. Fenomenal bilinçli durumların ayırt edilemeyeceği yönündeki bu görüşe karşıt olarak David Gamez 2005 yılında "Fenomenal Durum Skalası adını verdiği soyut bir ölçek geliştirmiştir.

Makinelerin fenomenal durumlara sahip olup olamayacağına yönelik karar aşamasının atlatılması durumunda, sentetik fenomenolojinin karşılaşacağı ikinci önemli sorun ise fenomenal durumların ne zaman ortaya çıktığına yönelik saptamalarda bulunmaktır (Gamez, 2008: 905). İnsan zihni fenomenal olmayan birçok içsel duruma sahiptir ve fenomenal olmayan durumların varlığı makineler için de geçerli olacaktır. Fenomenal ve fenomenal olmayan durumları birbirinden ayırt etmek için hangi içsel durumların bilinçli olduğunu tahmin etmeye yönelik geliştirilmiş olan *Bilgi Birleştirme Teorileri* gibi bilinç yaklaşımlarından faydalanılabilir.

Sentetik fenomenoloji ayrıca önceki bölümde ele alınan üçüncü şahıs bakış açısından insan fenomenolojisini araştıran nöro-fenomenolojiyle de karşılıklı etkileşim içerisindedir. Beyin etkinliklerinin bilimsel ölçümlerine dayanan nöro-fenomenoloji sentetik fenomenoloji için önemli bir veri kaynağı olmalıdır. Zira her iki araştırma alanı da içsel fenomenolojik durumları dışsal verilerle ve üçüncü şahıs bakış açısıyla tanımlamaya çalışmaktadır. Nörofenomenoloji var olan fenomenolojik durumları dışsal bir yaklaşımla açıklama çabasındayken sentetik fenomenoloji de fenomenolojik durumları yapay sistemlerde üretmeyi hedeflemektedir. Dolayısıyla nöro-fenomenolojinin verileri sentetik fenomenoloji için önemli katkılar sağlayacaktır.

4.2. Yapay Bilincin Diğer Alanlarla İlişkisi

Yapay zeka, işlevsel yönüyle günlük hayatın bir parçası hâline gelmiş durumdadır. Tartışılan yapay zeka meselelerin neye atıfta bulunduğunu doğru tespit edebilmek bu açıdan önemlidir. John Searle'ün yapmış olduğu güçlü ve zayıf yapay zeka ayrımı bu bağlamda önemli bir farka işaret etmektedir. Güçlü yapay zeka, insan zekasına eş ya da onu aşmış zeka

anlamına gelmektedir. Güçlü yapay zekaya göre, bilgisayar zihin çalışmalarında sadece bir araç değil, doğru programlanmış bir bilgisayar aynı zamanda gerçek bir zihindir. Yeterli ve karmaşık programlara sahip olan bilgisayarlar anlayabilirler ve diğer bilişsel durumlara sahip olurlar. Bu ayrım, Franklin'in (2003) fenomenal ve işlevsel bilinç ayrımıyla ve David Chalmers'in bilincin kolay ve zor problemleri arasında yaptığı ayrımla benzerdir. Bu ayrımların temelinde gerçek zihin ve zihnin dışsal gösterimi arasındaki karşıtlığa vurgu yapılmaktadır. Makine bilinci için ise bu ayrımlar, üçüncü gruptaki bilişsel bilim temelli yaklaşımlar ile dördüncü gruptaki fenomenal bilince sahip makine araştırmaları arasındaki zıtlığa tekabül etmektedir. Ancak fenomenal bilinçli makineler ile güçlü yapay zekayı özdeşleştirmek doğru bir bakış açısı olmayabilir. Zira zihin kavramı, fenomenal bilinçten ayrılabilir ve bir grup düşünüre göre, bilgisayarlar fenomenal bilince sahip olmadan da zihin sahibi olabilir (Gamez, 2008: 891).

İnsan zekasını tümüyle taklit etmeyi amaçlayan genel yapay zeka çalışmaları birinci grup davranış odaklı makine bilinci çalışmalarıyla tamamen örtüşmektedir. Ancak birinci gruptaki makine bilinci çalışmaları bilinçli insan davranışları üzerine ilerlerken genel yapay zeka çalışmaları zekayla bağlantısı olan tüm insan davranışlarını üretmeyi amaçlar. Dolayısıyla davranış temelli makine bilinci, genel yapay zeka alanının özel bir parçasıdır.

Felsefe ve yapay zeka, mantığın önemli bir yere sahip olduğu alanlar olmaları hasebiyle tarihsel olarak sıkı bir şekilde bağlıdır. Mantığın her iki alanda da etkisini yitiriyor oluşu bu bağlantıyı bir süreliğine zayıflasa da yapay bilinç çalışmaları, felsefe ve yapay zeka çalışmaları arasında yeni bir bağlantı sahasının oluşmasına yol açtı. Felsefe, davranış odaklı makine bilinci ve fenomenal makine bilinci çalışmalarının teorik zeminini hazırlarken bilişsel ve nöro-bilişsel makine bilinci çalışmaları da felsefeye ampirik veriler sunarak teorilerini test etme imkânı vermektedir.

Yapay zeka, robotik, nörobilim ve bilişsel bilimlerin birçok kolu algı, düşünme ve eylem kavramları ve bunlara bağlı olarak bilinç fenomeni üzerine değişik bakış açılarından araştırmalar yürütmektedir (Bonsignorio, 2013: 135). Yapay zeka alanı kendi içerisinde eski ve yeni moda yapay zeka olarak ayrılmaktadır. Anglosakson ülkelerde yaygın olarak benimsenen davranışçılık ekolü ile Geştalt psikolojisinden etkilenen Alman Okulu ayrımı, hatta daha temelde Anglosakson ve Kıta felsefesi ayrımı eski ve yeni moda yapay zeka yaklaşımları ayrımının da kaynağıdır (Bonsignorio, 2013: 135).

Eski moda yapay zeka olarak adlandırılan sembolik yapay zeka programlama üzerine kuruludur. Bilişsel bilimin kuruluşu, eski moda sembolik yapay zeka ile eş zamanlıdır. Sembolik yapay zekanın kuruculuğu giriş bölümünde bahsedildiği üzere McCarthy, Minsky ve Shannon tarafından paylaşılmaktadır. Sembolik yapılı bilgi sistemlerine odaklanan bu çalışmalara mantıksal temelli olarak ilerlemektedir. Sembolik yapay zekanın karşılaştığı en büyük problem ise semantik problem[47] ve temsil problemidir. Yeni moda olarak adlandırılan çalışmalar ise somutlaşmış yapay zeka çalışmalarıdır. Somutlaşmış yapay bilinç çalışmaları her ne kadar öznenin çevreyle etkileşiminin önemini vurgulasa da 1980'lerin sonunda geliştirilen ilk sistemler, çevreye adapte olma ve öğrenme konusunda yetersiz kalmışlardır (Nelson, 2013: 53). Şimdilerde, somutlaşmış yapay bilinç çalışmaları hem kontrol mekanizmaları hem de öğrenme yetileri bağlamında gelişmeler göstermektedir. Yine de beden sahibi olan öznelerde bilincin nasıl ortaya çıktığı ve ne şekilde var olduğu kesin olarak belirlenemediğinde bu çalışmaların evrileceği nokta tahminler dâhilinde değildir.

Sembolik ve somutlaşmış yapay bilinç çalışmaları sürekli olarak birbirini beslemek durumundadır. Çünkü bilinç, hem bilişsel ve hem de bedensel faaliyetlerden koparılamaz görünmektedir. Eğer ki, tamamlanmış bir bilinç çerçevesi yapay bir sistemde oluşturulmak istenecekse bunun yolu mutlak ki tüm bu alanların ortak çalışmalarıyla meydana gelecektir. Yapay zeka, robotik, nörobilim ve bilişsel bilimlerin bakış açılarını birleştirecek fiziksel olarak somutlaşmış bilişsel sistemler, bu bağlamda yapay bilinç araştırmalarının temel hedefi olmalıdır. Beden sahibi sistemlerin önemi, insanın bilinçli davranışlarının çevreyle olan etkileşimi ve eylemlerinden bağımsız olarak değerlendirilememesinden ileri gelmektedir. Hareket edebilen, çevreyi tanıyıp tepki verebilen, yeni bir çevresel durumla karşılaştığında eylem üretebilen yapay sistemler, bilinçli insan davranışlarına çok daha yakın olacaktır. Zira insanın, bilişsel kapasitesinin yanı sıra eylem ve çevreyle etkileşim yeteneği onun bilinçli durumlarını şekillendiren temel öğelerdir. Bu minvalde robotik çalışmalar, yapay bilinç çalışmaları için en az nörobilim ve bilişsel bilim alanları kadar ehemmiyet arz etmektedir.

4.3. Yapay Bilinç Çalışmalarının Değeri ve Geleceği

[47] Semantik problem, sembol temellendirme problemi olarak da adlandırılmaktadır. Anlamın, dijital bilgisayarlar gibi berimsel sistemdeki sembollerle nasıl eşleştirildiği sorunsalını vurgulamaktadır. Bu problem, çok sayıda felsefî sorgulamaya yol açmış ve henüz tatmin edici bir şekilde cevaplanamamıştır (Dreyfus, 2007; Searle, 1980; Steels ve Brooks, 1995; Wheeler, 2008).

Bilinç üzerinde henüz uzlaşıya varılmamış olan bir mefhumdur. Bu anlamda, kesin sınırları ve kapsamı belirlenemeyen bir yapının yapay olarak yeniden üretilmesi üzerine yapılan çalışmalar, kimi çevrelerce gereksiz bir uğraş olarak görülmektedir. Yapay bilinç alanının temel tanımlar ve aksiyomlar konusundaki eksikliğinden kaynaklı olarak, yapılan araştırmaların hedefsiz ve beyhude olduğunu savunanlar sayıca küçümsenmeyecek kadar çoktur. Doğal bilinci bile kesin ve net ifadelerle tanımlayamazken yapay olarak bilinç üretmenin mümkün olmayacağını iddia eden bu gruba göre, makine bilinci sonuçsuz bir çalışma alanı olarak kalmaya mahkûm olacaktır. Zira bir şeyin yapay olarak taklidinin yapılabilmesi için öncelikle taklit edilecek olan doğal yapının mahiyeti bilinmelidir. Bu grup makine bilinci çalışma sahasının, bilinç çalışmalarından ayrı düşünülemeyeceğini ve bilinç kesin bir bilim dalı hâline gelmedikçe de makine bilinci alanının çıktı üretemeyeceğini iddia etmektedir. Ancak gözden kaçırılmamalıdır ki, makine bilinci üzerine çalışmalar hiç değilse bilinçli bir makine üretmek gibi teknolojik bir hedefe sahiptirler. Var olan yapay zeka sistemlerinin hedefleri makine bilincine oranla oldukça sınırlıdır. Yapay zeka sistemleri genel olarak tek ya da birkaç yeti üzerinde uzmanlaşmış yapılar üretmeyi amaçlarken makine bilinci daha bütüncül bir bakış açısıyla insan zihnini yeniden üretme üzerine odaklanmaktadır. Şüphesiz bilinç, insan zihninin en temel niteliklerinden birisidir. Dahası, zihnin çözülemeyen yapısında rol oynayan en önemli faktördür. Bilimin, bilinç konusundaki çekingenliği ise bilincin öznelliğinden ve öznel ifadeler dışında test edilmeye çok olanaklı olmayan yapısından kaynaklanmaktadır. Bu anlamda, makine bilinci, bilincin bilimsel olarak ele alınabilmesi için kullanışlı bir yöntem olarak görülebilir. Yapay bilinç, gelecekte insan zihni çalışmalarına beklenmedik katkılar sağlayabilir bir alandır. Zira bilince yönelik teori ve hipotezleri kontrol etmek için iyi bir test alanı sunmaktadır. İnsan beyninin deneye tabi tutulamayacağı durumlarda, yapay zeka sistemleri ve robotik yapılar, araştırmacıya deney imkânı sunarak, bilinç çalışmalarına önemli destekler sağlayacaktır.

Yapay bilinç araştırmacılarının yapması gereken belki de bir seferde bütünüyle bilinçli bir yapay maddeyi inşa etmeye çalışmak değil, akıllı robotlardan bilinçli robotlara gelişimsel bir patikanın izlerini takip etmektir. Bilinçli yapay madde, bilinçli canlıların gelişimini zorunlu olarak takip etmek zorunda değildir. Bundan 50 yıl önce yapay zekanın sahip olduğu muğlak niteliklere bugün yapay bilinç sahiptir. Olması gereken ise yapay zekanın kendi içerisindeki evrim ve gelişim aşamalarını takip ederek yapay bilince uygulamaktır. Basit hesaplamalarla başlayan yapay zeka araştırmalarının bugün derin öğrenme sürecine eriştiği göz önüne alındığında aynı durumun yapay bilinç çalışmaları için de mümkün olabileceği

düşünülmektedir. Bu sebeple minimal bilinçli bir yapay maddeden yola çıkmak hedefe daha sağlam adımlarla ulaşılmasını sağlayabilir.

Makine bilinci üzerine çalışan araştırmacılar, bilincin bilgisayar modellerinin bizim bilinçli bilgi sürecimizi anlamımıza yardımcı olacağı gerekçesiyle makine bilincinin üzerinde durulması gereken değerli bir alan olduğunu savunmaktadırlar (Reggia, 2013: 112). Bilincin bilişsel niteliklerini taklit eden makineler geliştikçe bizim dünyamızda daha çok yer edinip, dilimizi öğrenıp bize yardımcı olma ve iletişim kurma yeteneklerini arttıracaklardır. Yapay sistemler empati yeteneği kazanarak, düşüncelerimizi taklit ve tahmin edebileceklerdir. Makine öğrenmesi temelli bu sistemler, bulunduğu sosyal çevrenin davranış, eğilim, inanç ve fikirlerini kendileri için bir veri tabanı oluşturacak şeklinde kullanacak ve sosyal etkileşim içerisinde kendilerine ait bir karakter elde edeceklerdir. Bu makineler bir anlamda insanlar ve makineler arasında özneler-arası bir ilişki türünün başlangıcı olacaklardır.

Nöral temelli makine bilinci çalışmaları ise başka bir açıdan bizim insan bilincini anlama ve anlamlandırmamız bağlamında yeni yollar açacaktır. Nöral temelli makine bilinci, doğal bilincin nöral korelasyonlarının keşfi için test imkânı sağlayan bir yöntem olacaktır. Bilincin nöral korelasyonları genel olarak, beyin görüntüleme teknikleri ve kişilerin sözel aktarımlarına dayalı olarak tespit edilmeye çalışılmaktadır. Bu yöntemler ise hem tüm bağıntıların keşfi için sınırlı kalmakta, hem de sözel verilerin öznelliğini aşamamaktadırlar. Bu bağlamda, makine bilinci çalışmaları bilincin nöral korelasyonlarının keşfi için deneysel olarak nesnel bir ortam sağlayarak nörobilime katkıda bulunacaktır. Bunun yanı sıra, nöral temelli makine bilinci araştırmalarının gelişimi insanın problem çözme yeteneğini makinelere aktarmamızı sağlayacak ve gündelik hayatın kolaylaştırılması ve iyileştirilmesi adına yeni buluşların önünü açacaktır. Problem çözme yeteneği kazanan makinelerin sosyal hayatta yer bulması, işgücünün paydaşları arasındaki dağılımı da değiştirecektir. Bir yandan birçok meslek için işsizlik kaygısı doğuracak olsa da insandan daha hızlı hesaplama yeteneğine sahip makineler ile üretim hızlanacak ve yeni iş sahaları yaratılacaktır. Bu bağlamda, makine bilinci çalışmalarının dünyadaki ekonomik dengeler üzerinde de belirleyici rolü olacaktır. Nöral temelli makine bilinci çalışmaları, tersine mühendislik ile beynin geliştirilmesi ve hızlandırılması çalışmalarına da katkıda bulunabilir. Beyne takılacak

küçük çipler ile beyni yapay zekaya bağlayarak çok daha hızlı düşünmesi, çok daha geniş bir hafızaya sahip olması ve çok daha büyük verilerin analizini yapması planlanmaktadır.[48]

Fenomenal bilince sahip makineler ise hem fenomenal bilincin genellikle bilimin sınırlarına dâhil edilmeyişi hem de fenomenal bilinç tanımının muğlaklığından kaynaklı genellikle kaçınılan bir çalışma alanını oluşturmaktadırlar. İnsan beyni gibi fiziksel bir sistemin öznel deneyimleri desteklediği algısı bilimsel olarak çelişkili sonuçlar yaratmaktadır. Hatta agnostik yaklaşım, bilimin objektif yöntemlerinin bilinci öznel doğası sebebiyle, hiçbir zaman aydınlatamayacağını ileri sürmektedir. Bu alan fenomenal bilincin bilimsel açıklaması için kullanılabilecek bir yöntem olarak düşünüldüğünde, teorik olanın uygulanabilirliği ve test edilebilirliği açısından önemli bir rol üstlenecektir. Fenomenal makine bilinci çalışmaları, fenomenal bilincin doğası üzerine bilimsel anlayışımızı arttıracaktır. Özellikle sentetik fenomenoloji, insan dışındaki sistemlerin fenomenal durumları hakkında tahminler yapacak (Gamez, 2008: 907), bu tahminler, bitki ve hayvan bilinci araştırmalarına da değerli katkılar sunacaktır. Bunun dışında, dil yeteneği olmayan bebeklerin ya da beyninin ilgili bölgesi zarar görmüş, deneyimlerini aktaramayan insanların fenomenal durumlarını anlamamıza ve araştırmamıza yardımcı olacaktır.

Bunun dışında makine bilinci çalışmaları, yapay zekaların ve robotların sosyal hayat içerisinde kendilerine yer bulabildikleri bir gelecek tahayyülü için de büyük önem arz etmektedir. Gelecek senaryolarını şekillendirecek olan insan zekasına eş ya da insan zekasını aşmış bir yapay zekanın ortaya çıkışından ziyade yapay bilincin oluşturulması olacaktır. Bilinç, deneyimi mümkün kılan yegâne unsurdur. Bu, iyi ve kötü tüm deneyimlerin sonucu olarak mutluluk, güzellik, iyilik, kötülük, acı vs. her türlü anlamın bilinçle beraber var olduğu anlamına gelmektedir. Yaşantımıza anlam katan bu bağlamda zeka değil, bilinçtir. Bilinç, bireysel yaşantıların yanı sıra evrene de anlam veren tek unsurdur. Evrenin varoluşunu anlamlı hâle getiren bilinçli deneyimlerdir. Bu sebeple yapay zekanın gelecekte evrenin anlamına bir tür müdahalesi söz konusu olacaksa bu müdahale ancak onun sahip olacağı bilinçle mümkün olacaktır. Bugün için konuşulduğunda ise insanı hâlen evrende ayrıcalıklı bir konumda tutan özelliği onun düşünen bir varlık olması değil, deneyimleyen

[48] Elon Musk tarafından 2017'de kurulan Neuralink isimli şirket insan beynini, yapay zekaya bağlayacak teknolojiler geliştirmek üzere büyük bir proje etrafında çalışmaktadır. Şirketin bu yöndeki çalışmaları gizli tutulduğundan, bu konuda yayınladıkları herhangi bir bilimsel makale yoktur. Ancak projeyi amaç ve kapsam yönünden tanıtmak üzere Elon Musk'ın yaptığı sunuma https://www.youtube.com/watch?v=dXbShyPQrhM adresinden ulaşılabilir.

ve hisseden bir varlık olmasıdır. Zeka bağlamında makinelerin ulaştığı nokta ve makine öğrenmesi ile çizilen gelecek perspektifi, insanın zeka konusundaki üstünlüğünü makinelere kaptırmaya çok da uzak olmadığını söylemektedir. Bu durumda insan zihnindeki tek ayrıcalıklı yön bilinç olarak göze çarpmaktadır. Yukarıda ayrıntılı olarak ele alınan makinenin ahlakiliği tartışmalarında açıkça görülmüştür ki ahlak ile bilinç arasında sıkı bir ilişki vardır. Değer ve anlam varlığı olarak insana benzer bir ahlaki sistemin yapay olarak yeniden üretilmesi, bilinçten yoksun olan sistemlerde mümkün görünmemektedir. Dolayısıyla makine etiği araştırmaları için yalnızca yapay zeka gelişmelerinin değil yapay bilinç çalışmalarının da dikkate alınması zorunluluk olarak göze çarpmaktadır.

Bilinç, insanın hem etik ve sosyal yapısı olarak hem de zihinsel yapısı itibariyle ayrılmaz bir parçasıdır. Bu nedenle, insan zihnine eş değer nitelikte bir yapay zekadan bahsedilecekse, bu yapay zihnin bilinç bileşeni olmadan inşası mümkün değildir. Bunun yanı sıra, bilimkurgu filmlerinin temel ögelerinden olan yapay zekalı robotların sosyal yaşamın içerisinde yer aldıkları ve dünyanın kontrolünü ele geçirdikleri vb. senaryoların hemen hepsi makine bilinci tartışmalarını zorunlu kılmaktadır. Makinede etik bir yapı oluşturulması ve makine hakları tartışmaları ise tartışıldığı üzere bilinçten bağımsız değildir. Bu sebeplerle, yapay bilinç hem etik, hukuki ve sosyal tartışmalar bağlamında hem de bilinç araştırmaları için yeni bir yöntem olarak kullanılması bağlamında yapay zeka alanının en önemli alt dallarından birisidir.

4.4. Yapay Bilince Epistemolojik Yaklaşım: Bir Makinenin Bilinçli Olduğunu Bilebilir miyiz?

Makine bilincinin hangi şartlarda mümkün olabileceğinin tartışması, zihin felsefesi tarihinin en köklü tartışmalarından olan başka zihinler probleminin bir yansıması olarak da ele alınabilir. "Bir başkasının bir zihne sahip olduğunu ya da bilinçli olduğunu nereden biliyoruz?" sorusu başka zihinler probleminin temel sorunsalıdır. Bu problem zihinsel durumların ikinci veya üçüncü kişilere atfedilmesini konu alır ve şu soru üzerinden yola çıkar: Biz nasıl olur da üçüncü bir kişiye ait -içsel bir durum olan- zihinsel ya da psikolojik durumları bilebiliriz? Churchland bu problemi şöyle özetlemektedir:

> Özgül psikolojik durumların -düşüncelerin, algıların, arzuların veya acıların- başkalarına her göndermenin açık bedensel davranışları temel alması gerekir (…) Ancak bu türden 'içsel durumlar' ve bunların 'dışsal ölçütleri' arasında bulunan genel bağıntıların bir değerlendirmesi, kişinin bu bağıntıları başkalarında gözlemlemesinden türetilemez çünkü varsayımsal bağıntıların "dışsal" öğeleri

> herkesin incelemesine açıktır. Bu türden bağıntıların bilgisi dolayısıyla yalnızca kişinin kendi durumundan türetilebilir çünkü denir ki, içsel öğeler incelemeye açıktır (Churchland, 2013: 124).

Başka zihinler probleminin kastettiği aslında kişinin kendi içinde bulunduğu durumda keşfettiği bağıntıları tümevarımsal yöntemle genelleyerek kendi durumunu başkalarına da atfetmesidir. Ancak bu kendi durumundan diğer durumlara sıçrayış ve bu aradaki bağıntı problem teşkil etmektedir. Kişinin tek bir deneyiminden ulaştığı genelleme ile tüm başka zihinleri bilemeyeceği iddiası bizi başka zihinlerin bilgisi hakkında kuşkuculuğa götürmektedir.

Bu problemin çözümüne ilişkin iki yaklaşım bulunmaktadır. Birincisi yukarıda detaylı açıklamasını verdiğimiz davranışçılık ve ikincisi de benzeşim argümanıdır. Geleneksel yaklaşımda ve günlük dilde, "bilinçlilik" davranışlara bakarak değerlendirilir. Birisi uyanıksa ve davranışlarında anormal ve anlamsız görünen bir durum yoksa konuşabiliyorsa ve konuşmalarında bilinçli olup olmadığına yanıt verebiliyorsa onun bilinçli olduğu varsayılır. Bunun yanı sıra, laboratuvar çalışmalarında da birisinin deneyimini rapor edebilme kapasitesi bilincin varlığını ispatlamak için altın standart olarak kabul edilmektedir (Tononi ve Koch, 2015: 2). Davranışçılık kişinin kendi durumundan diğer kişinin durumuna geçiş arasındaki bağlantıyı gerekçelendirmeyi reddeder. Zira davranışçılık, zihin durumlarının gözlemlenebilen davranışlarla ya da davranışsal dışavurumlarla eşdeğer olduğunu savunan görüştür. Benzeşim argümanı ise kişinin kendi zihninin bilgisine dayanarak diğer zihinlerin bilgisine ulaşabileceğini savunur çünkü bütün zihinler benzerdir (Churchland, 2013: 125).

Benzeşim argümanı ve davranışçılığın laboratuvar ortamında kullanışlı olmasını sağlayan raporlama yeteneği ve davranışa bakarak bilinçliliğe karar verme kendi içerisinde problemli bir yöntemdir. Zira uyurgezerlik gibi kişinin uyurken yürüdüğü durumlar ile kişinin uyur gibi gözüm hayal kurduğu durumlar davranış açısından yanıltıcıdır. Aynı şekilde, raporlamada da kişi rüyasında gördüğü ya da hayal ettiği şeyleri olduğu gibi aktaramayabilir. Dolayısıyla rapor edilemeyen durumlarda da bilincin varlığından söz edilebileceği gibi; yanlış raporla bilinçli olmayan durumların bilinçli olarak görünebilmesi söz konusu olabilir. Ayrıca, raporlama yeteneğini bilinç için ayrıştırıcı bir ölçüt olarak görmek, dil yeteneğine çok büyük bir rol vermekte ve bebekler, hayvanlar ve makineleri raporlama yetenekleri olmadıklarından dolayı bilinçliliği değerlendirmeye bile alınamayacak sistemler hâline getirmektedir (Tononi ve Koch, 2015: 2). Bu sebeplerle, davranışçılık, bilincin altında yatan

fiziksel ve zihinsel mekanizmalar bilinmeden tek başına başka zihinler problemi için çözüm üreten bir yaklaşım olarak kabul edilmemelidir.

Başka zihinler problemini, günlük dilden kurtarıp zihin felsefesi içerisinde ayrı bir tartışma konusu hâline gelmesi Thomas Nagel'in ünlü "Bir Yarasa Olmak Nasıl Bir Şeydir?" isimli makalesidir. Nagel'in bu makaledeki amacı zihin ve beyni aynı şeyler olarak kabul eden fizikalizmi deneyimin öznel karakterini göz ardı ettikleri için eleştirmektir. Nagel'a göre fizikalistler zihni beyne indirgerken indirgedikleri şeyin öznel yapısını incelemeyi unutmaktadırlar. Zira her öznel fenomen tek bir bakış açısıyla bağlantılıdır ve objektif, fiziksel teorilerin bu bakış açısını ihmal etmesi kabul edilemezdir (Nagel, 1974: 437). Bu tek bakış açısının ne anlama geldiğini açıklamak için "Bir yarasa olmak nasıl bir şeydir?" sorusunun cevabını aramaya girişen Nagel, yarasanın memeli, görmediği için sonar kullanan bir hayvan olduğunu herkesin bildiğini kabul eder. Nagel'a göre yarasa olmanın nasıl bir şey olduğu üzerine düşündüğümüzde, dışsal seslere göre hareket eden, uçabilen, çok az görme yetisi olan, böcek yiyen bir hayvan olduğumuzu hayal ederiz fakat aslında onun kastettiği yarasa için yarasa olmanın nasıl bir şey olduğu üzerine düşünebilmemizdir (Nagel, 1974: 439). Nagel'in "Bir Yarasa Olmak Nasıl Bir Şeydir?" makalesindeki maksadı Zeman şöyle özetler:

> Sahip olmadığınız bir duyuya sahip bir yaratığın, mesela böcekçil yarasanın, davranışlarını ve fizyolojisini inceleyen bir bilim insanı olduğunuzu düşünün. Güçlü bir beyne sahip ve memeli olan bu hayvanlar, yönlerini sesle, yüksek perdeli bir çığlık sesiyle etraflarını kolaçan ederek bulurlar: Seslerinin yankıları şekilleri, derinliği, hareketleri ve dokuları ayırt etmelerini sağlar. Uzak bir gelecekte, sinirbilimin yarasanın iç işleyişini tamamen izah edebilecek bir aşamada olduğu bir çağda araştırma yaptığınızı kabul edelim. Yarasalar da çoğumuzun düşündüğü gibi, bilinçli yaratıklar olsun. Sinirbilimin bu izahını yalayıp yutmuş olmak, yarasa olmanın nasıl bir şey olduğu konusunda bilgi sahibi olmanızı sağlayabilir mi? (Zeman, 2006: 381).

Nagel'a göre bu bizim için imkânsızdır çünkü biz kendi zihnimizin kaynakları içinde sınırlıyız ve bu kaynaklar bu durumu kavrayabilmemiz için yetersizdir. Bilincin problem olarak ele alınan bir fenomen olmasının ana sebebi öznelliğin varlığıdır. Nagel gerçek sorunu öznel bakış açılarının yarattığını ileri sürer. Beynin nesnel bir açıklamasının, bu tam bir açıklama olsa bile beynin izin verdiği öznel deneyimi nasıl olup da izah edebileceğini tahayyül edemeyiz (Zeman, 2006: 382).

Nagel'a göre "Eğer bu evrende başka bir yerde bilinçli bir hayat varsa bunun bir kısmı bizim için kullanılabilir olan deneyimsel terimlerle bile tanımlanamaz." (Nagel, 1974: 440). Bir

kişi, diğer bir kişinin deneyimleri hakkında objektif bir yaklaşıma sahip olabilir ancak yine de üçüncül bakış açısıyla yaklaştığı için deneyimlerin sübjektifliğinden kurtulamaz. Nagel şöyle iddia eder: "Eğer deneyimin öznel karakteri tek bir bakış açısı tarafından kavranabiliyorsa daha fazla bir nesnelliğe geçiş bizi fenomenin gerçek doğasına ulaştırmaz." (Nagel, 1974: 445). Bu da zaten başka zihinler probleminin çıkış noktasıdır; başka zihinler hakkında kesin bir bilgiye ulaşmamız deneyimlerin öznel yapısı itibarıyla mümkün gözükmemektedir. Nagel'a göre problem, deneyimleri oluşturan duyumların nesnel bir yapıya sahip olup olamayacakları üzerinde durulmamasıdır ve fizikalizmin de düşünmesi gereken deneyimlerimizin bize nasıl gözüktüğü değil de gerçekte nasıl olduğu olmalıdır (Nagel, 1974: 448). Nagel'in dikkat çekmeye çalıştığı husus nesnel-öznel ayrımı tam olarak yapılmadan ne zihin-beden probleminin ne de başka zihinler probleminin üstesinden gelinemeyeceğidir.

Churchland de başka zihinler problemini Nagel gibi indirgemeciliği eleştirerek ele almıştır. Davranışçıların, zihinsel durumları davranışa indirgemesini eleştiren Churchland, bu durumu açık bir şekilde şöyle ifade eder:

> İçsel durumlar ve açık davranış arasında birebir ilişki yoktur; daha ziyade bunlar tamamen gelişmiş bir kurama özgü karmaşık ve açık uçlu yollarla gerçekleşir. İndirgemeci davranışçılığın yanlış olması da kuşkusuz bu sebeptendir. Zihinsel durumların fiilen ve olanak olarak açık davranışa indirgenmesi, protonların fiilen veya olanak olarak buhar izlerine indirgenmesi kadar bile mümkün değildir (Churchland, 2013: 131).

Churchland'e göre insan kendisi sayesinde, kişiler hakkındaki kavramsal çerçeve dâhilinde sağduyuya dayalı olarak diğer kişilerin psikolojik durumlarını kavrayabilir. Bu kavramsal çerçeveyi verdiği on üç yasayla özetlemeye çalışan Churchland'de bu yasalar dışsal koşullar ve onların içsel etkileri, zihin içi özellikler ve açık davranışın içsel nedenleriyle ilgilidir.

Nagel ve Churchland tarafından başka zihinler ile ilgili olarak açıkça ortaya koyulan problem makineler için de geçerlidir. Churchland'in eleştirisiyle birlikte düşünüldüğünde, bir makinenin bilinçli olup olmadığına makinenin davranışlarına bakarak karar vermenin yanlış bir yol olduğu kabul edilmelidir. Dışsal gözlemle bir makineye bilinç atfetmek, benzetim ve gerçekliği aynı görmek olacaktır ki, bunun bir çelişki olduğu da açıktır. "Bir Yarasa Gibi Olmak Nasıl Bir Şeydir?" makalesi üzerinden yola çıkıldığında ise çok daha olumsuz bir tablo ile karşılaşacaktır. Bu soru "Makine gibi olmak nasıl bir şeydir?" ya da "Makineye göre makine gibi olmak nasıl bir şeydir?" şeklinde revize edildiğinde

cevaplanması bizim için imkânsız hâle gelmektedir. Deneyimlerin öznelliğine vurgu yapan bu iddia ışığında, makine bilincinin imkânının ve hangi şartlarda bilinebilir olduğunun tartışılmasının yapay zeka felsefesi için oldukça elzem bir mesele olduğu daha kolay görülebilecektir.

"Bir makinenin bilinçli olduğunu nasıl bilebiliriz?" sorusu aslında başka-zihinler probleminin özelleştirilmiş hâlidir. "Bilinç" kavramı niteliksellik, öznellik, ben hissi, birlik gibi birçok tanımlayıcı özellik ile birlikte anılmaktadır. Bu kavramın bilimsel açıklaması için ölçülebilir insan davranışları üzerine skalalar üretilmektedir. Ancak bu skala ve kriterleri pratikte makinelere uygulamak oldukça zor görünmektedir.

Zihin felsefecilerinden Alexander ve Dunmall (2003) minimal düzeyde bir bilincin varlığını beş aksiyom önermektedir: Algısal, duygusal ve imgesel içsel durumlarının varlığı ile dikkat ve planlama mekanizmaları. Bu yaklaşıma göre, bir makinenin bilinçli olduğunu bilmemiz için en azından bu beş özelliğe sahip olduğunu bilmemiz gerekir. Ancak "Makinenin bu özelliklere sahip olduğunu nasıl bileceğiz?" sorusu da yaklaşım içerisinde cevaplanmamış bir soru olarak hâlâ geçerlidir. Bir başka yaklaşım ise soruna makine yapısını ve işlevselliğini test ederek yaklaşmaktadır. Bu bakış açısına göre, "Bir makine içsel imgelere ve içsel konuşmaya sahipse, ayrıca da bu fenomenlere sahip olduğunu rapor edebiliyorsa, içeriklerini tanımlayabiliyorsa ve o şekilde programlanmadığı hâlde bunları kendi ürünü olarak fark edebiliyorsa bilinçlidir." (Reggia, 2013: 116). Bir makinenin bilinçli olduğunu bilme yöntemlerine yönelik üretilen testlerden bazıları ise karşılaştırma tekniklerine dayalıdır. Bilişsel karşılaştırma ve makinelerin fiziksel özelliklerini sağlıklı yetişkin beyniyle karşılaştırma yapmak için skalalar üretilmeye çalışılmaktadır. Ancak bu testlerin hiçbirisi henüz makine bilincinin varlığını ya da yokluğunu değerlendirmek için üzerlerinde mutabakata varılmış objektif ölçütler olarak kabul edilmemektedir.

Yapay bilince bir başka epistemolojik yaklaşım da sistemin potansiyel olarak bilinçli olması ve bilfiil/aktüel olarak bilinçli olması ayrımını dile getirmektedir. İnsanlar için de aynı ayrım geçerlidir; ortalama bir insan sürekli olarak potansiyel bilince sahip olsa da günün ancak 16 saatinde aktüel olarak bilinçlidir. Makinelerin potansiyel olarak bilinçli olup olmadıklarını değerlendirmek için insanlarda bilinçle ilgili fiziksel şartlara bakmak ve yapay sistemlerde bu ilişkinin bulunup bulunmadığına karar vermek gerekli bir şart olarak öne sürülür (Gamez, 2009: 214). Bir bakış açısı olarak, eğer insandaki bilinç ve fiziksel şartlar arasındaki korelasyon yapay sistemde de mevcutsa, o sistemin bilinçli durumlara sahip olma

potansiyelinden bahsedebilmek için iyi bir temel olduğu söylenebilir. Ancak bilincin korelasyonlarına dair deneylerin uygulamalarındaki zorluklar, bu bakış açısının da pratik olarak problemli olmasına yol açmaktadır. Bilinçli durumlar, deneysel olarak ancak davranışlarla ölçülebilmektedir. İnsanın davranışlarını etkilemeyen deneylerden ve sonuçlarıyla insanın fenomenal deneyimini rapor etmesiyle çatışan deneylerden kullanışlı sonuçlar elde etmek mümkün görünmemektedir (Gamez, 2009: 214).

5. YAPAY BİLİNÇ ÜZERİNE SOSYAL, ETİK VE YASAL HUSUSLAR

5.1. Makine Etiğinin Felsefi ve Sosyal Temelleri

Aristoteles'ten itibaren felsefeyi teorik ve pratik felsefe şeklinde ikiye ayırmak bir gelenek hâlini almıştır. İyi karakter, doğru eylem ve mutlu bir hayatı konu edinen pratik felsefe, diğer bir deyişle aksiyoloji, insanla ilgili her türlü söz, eylem ve durumu kendisine konu edinebilir. Aksiyoloji içerisinde yer alan en önemli alanlardan ahlak ve hukuk da konumuzla yakından ilgilidir. Hem ahlak hem de hukuk gündelik hayatı düzenleyerek daha mutlu ve müreffeh toplumlar veya insanlar yaratmayı amaçlayan disiplinlerdir. Özellikle ahlak, tüm bireyleri ilgilendiren, neyin iyi neyin kötü olduğunu, doğru ile yanlışı birbirinden nasıl ayırabileceğimizi, yapılması ve kaçınılması gereken davranışların neler olduğu ortaya koyduktan sonra her durumda genel geçer birtakım ilke ve kurallar arayışında olan normatif bir bilimdir. Ahlak ortaya koyduğu kurallar ile iyi olduğunu kabul ettiği davranışların (faziletler/erdemler) yapıla yapıla alışkanlık hâline gelmesini, kötü olduğunu kabul ettiği davranışların da (reziletler/kötü ahlak) yapılmayarak uzak durulmasını hedefler. Yine ahlak felsefesinin ele aldığı konuların başında gelen "Mutlu bir hayatı elde etmek için onurlu bir yaşam mı, bilgece bir hayat mı, hazlar mı, iyiye sahip olmak mı ya da ilahî olana bağlanmak mı gerekir?" sorusu filozofları tarih boyunca meşgul eden bir başka bir problematiktir. Görüleceği üzere biyo-psiko-sosyal bir varlık olan insan için bu dünyada ahlaki anlamda nasıl bir yaşam sürmeli sorunsalı üzerine sorgulama yapmak varoluşsal olarak kaçınılmaz bir durumdur. İnsanın biyo-psiko-sosyal ihtiyaçlarının karşılanması ile ahlaki norm ve değerler insan olmanın zorunlu bir sonucu olarak ortaya çıkmaktadır. Bu bağlamda insan eylemlerini ve bu eylemleri belirlemekte temel bir role sahip ahlaki prensipleri belirleyenin akıl mı (Aristoteles, Kant vd.), duygular mı (Hume, J. S. Mill vd.), sezgiler mi (St. Thomas, G. E. Moore vd.) yoksa diğer etmenler mi olduğu da sıkça konu edinilen felsefi bir husus olmuştur.

İnsanı diğer canlılardan ayıran ve onu evrende sahip olduğu ayrıcalıklı konuma yerleştiren sebeplerin başında, insanın yalnızca biyolojik bir varlık değil aynı zamanda bir değer varlığı oluşu gelmektedir. Bu bağlamda, evrende insandan daha üstün bir konuma yerleşmesi muhtemel bir varlığın, sahip olması gereken yetiler arasında ahlakın muhakkak bulunması gerektiği sıklıkla tartışılmaktadır. Yapay bilincin de insana benzer bir duruma gelebilmesi için yalnızca süper zekaya sahip olmasının yeterli olmayacağı aynı zamanda etik bir sisteme

sahip olması gerektiği fikri ahlak ve yapay zeka ilişkisini ayrı bir araştırma konusu hâline getirmiştir.

Otonomluk[49] seviyesi giderek artan sistemlerin hızlı gelişimleri ve bu sistemlerin insanlar ile diğer canlılara zarar verme ihtimali üzerine oluşturulan senaryolar *makine etiği, makine ahlakiliği, yapay ahlak, berimsel etik, robot etiği* ve *dost yapay zeka* gibi farklı isimlerle adlandırılan yeni bir araştırma sahasının ortaya çıkmasına yol açmıştır. Bu yeni araştırma sahası, etik önemi olan eylemleri gerçekleştirebilecek, insan yapımı araçlar üzerine çalışmaktadır. İnsanların makine ve bilgisayar kullanımı esnasında oluşan etik meselelere odaklanan bilişim etiğinden farklı olarak; makine etiği, makinelerin insanlara ve diğer makinelere davranışları ile aralarındaki etkileşimin ahlakiliğine odaklanır.

Makine etiği, yapay zeka araştırmacıları tarafından önemle üstünde durulması gereken bir alandır. Makinelerin bugün yaptıkları ve gelecekte yapacakları işlere ait etik unsurlar oldukça fazla sayıda ve mühimdir. Artık ortak bir öngörü hâline gelen senaryoya göre, insanın beden gücüyle yaptığı her işi yapabilecek ve aynı zamanda insandan çok daha hızlı akıl yürütebilecek makineler; evden işe giderken kullanılacak sürücüsüz otomobillerden, bizim günlük angarya olarak adlandırdığımız tüm işleri yardım almadan yapabilecek robotlara kadar hayatın her alanında yer bulacaklardır. Toplumsal yaşam içerisinde bu kadar geniş yer kaplayacak makinelerin, etik bileşenlere sahip olmadıkları sürece insana, doğaya, evrene, toplumsal düzene veya ahlaki değerlere zarar verme ihtimalleri her zaman olacaktır. Bu kapsamda, yapay zekanın gelecekte elde edeceği düşünülen güç ve sosyal hayatta erişeceği konuma dair olan korkunun temelinde de yapay zekaların etik ilkelere sahip olup olamayacağı sorunsalı yatmaktadır. Otonom ve zeki makinelerin üretimine dair toplumsal temayül ancak bu makinelerin etik olmayan davranışlarda bulunmayacakları garanti altına alındığında veya tüm etik olmayan davranışlarına karşı güvenlik tedbirleri mevcut olduğunda yasal olması gerektiği yönündedir (Anderson, 2007: 16). Bu hususta önemli olan etik olan ve olmayanı ayırt etme üzerine doğru formülasyonlar üretebilmektir zira aksi

[49] Felsefi anlamda otonomi, kendi kendine karar alabilme, eyleme geçebilme ve davranışsal olarak bağımsız olma yetisi şeklinde tanımlanmaktadır. Otonom sistem kendi içinde bağımsız karar verebilme, deneyimlerden yola çıkarak yeni şeyler öğrenme ve irade sahibi olarak seçim yapabilmeyi barındırır. Dolayısıyla bir sistem ya da makine ne kadar otonomsa, o kadar yüksek öngörülemeyen ve tahmin edilemeyen kararlar alma potansiyeline sahiptir (Sparrow, 2007: 70). Yeterli karmaşıklığa sahip robotlar yönelimsel durumlara sahip olabilirler. Yönelimsellik, zihinselliğin ayırt edici özelliklerinden bir tanesidir. Yönelimsel durumlara sahip robotlar, otonom yönelimsel bireyler inşa edebilirler. Makineler içkin yönelimsel durumlara sahip olurlarsa, onların otonom bireyler olma potansiyelleri vardır ve otonom bireyin de yeni kavramlar öğrenme ve kendi kavramlarını üretme becerisi vardır (Young ve Torrance, 1994: 199).

takdirde hangi makinenin etik ilkelere sahip olduğu, hangi makinenin etik olmayan davranışlarda bulunabileceği tahmin ve tespit edilemez. Bu noktada, yapay zeka çalışanlarının ve mühendislerinin, etik tartışmaları dikkate alması oldukça mühimdir.

Etik üzerine düşünmek, etiğin sahip olduğu kimi nitelikler sebebiyle pratik ve sonuca kolay ulaşılabilir olsa da etik alanının kimi nitelikleriyle ise çetrefilli ve zor bir uğraştır. Ahlakın tüm günlük eylem ve ilişkilerin temellerine sirayet etmiş olması, ilk bakışta etiğin anlaşılması çok kolay bir alan olduğu izlenimini vermektedir. Bir değer varlığı olarak insanı, etik alandan bağımsız değerlendirmek mümkün değildir; yine de bu hepimizin etik uzmanı olduğumuz anlamına gelmemektedir (Anderson, 2007: 15). Ahlaki norm ve değerlere sahip olmak ile sahip olduğumuz ahlaki değerlerin ayırdında olmak birbirinden çok farklıdır. Etiği zorluğu kılan, insanda var olan etik karar verme ölçütlerini açık hâle getirmektir ve makine etiğinin zorluğu ise etiğin tespit ettiği bu ölçütleri, makineye kodlamak ya da öğretmektir. Bu bağlamda makine etiği, hem etik alanının zorluklarıyla yüzleşmekte hem de teknik sorumlulukları üstlenmektedir.

"Etik makine" ifadesi, farklı bakış açıları için farklı çağrışımlara sahiptir. Bir makinenin etik fail olması "açık" ve "örtük" olmak üzere iki farklı şekilde mümkündür. Örtük ya da gizli etik fail bir makine, tasarımcısının etik prensipleriyle sınırlandırılmış olarak, etik olmayan davranıştan kaçınmaya programlanmıştır; açık etik fail olan makine ise etik prensipleri kullanmayı bilir ve etik ikilemlerde en iyi eylemi hesaplayarak seçim yapabilir (Moor, 2006: 19). Bu bağlamda, makine etiğinin temel hedefi açık etik failler olarak yapay ahlaka sahip sistemler inşa etmektir. Açık etik fail, bir eylemin neden doğru ya da yanlış olduğunu etik prensiplere dayalı olarak açıklayabilir. Ahlakı birey, kural ya da prensipleri açıklamaya dayanmaksızın körü körüne uygulamaz; bağlamı anlayarak, eylem türleri arasında açıklamaya dayalı seçimler yapabilir (Nath ve Sahu, 2017: 4). Gizli etik fail olan bir makine ise doğru etik kararlar vermek için programlanmışsa da yargılarını temellendirmek için herhangi bir prensibe sahip değildir; bu sebeple yapay ahlaka sahip bir sistem olarak adlandırılması doğru değildir.

Yapay ahlak sahibi bir sistemin tanımı şu şekilde yapılmaktadır:

> Sistemlerin tasarımcılarının ya da mühendislerin, sistemlerin yeni bir girdi ya durumla karşılaştıklarında nasıl davranacağını tahmin edemediği noktada, bu sistemlerin eylemlerinin güvenirliğini, uygunluğunu ve sosyal ya da ahlaki kabul edilebilirliğini değerlendirecek mekanizmalara ihtiyacı olacaktır; bu

> değerlendirmeleri yapabilen sistemler ise fonksiyonel olarak ahlaki sistemler olarak adlandırılacaktır (Wendell, Allen ve Franklin, 2011: 178).

Yapay ahlaka sahip sistemler, kendi programlarının ötesine geçerek etik tercih ve değerlendirmelerde bulunabilen ve kodlarında var olmayan girdiler için tahminde bulunarak çıktı üretebilen yapılardır.

Felsefenin, aksiyolojik olarak pratiğe ve günlük hayata en yakın kollarından olan etik, bireylerin etik ikilemlerle karşılaştığında nasıl davranacağı ile ilgilenirken makine etiğinin ilgisi de benzer olarak makinelerin sosyal ilişkilerindeki davranışları üzerinedir. Ancak makine etiği tartışmaları davranışlardan önce, etiğin ne şekilde makineye aktarılacağı sorunsalıyla başlamaktadır. Makine etiğinin temel felsefi problemlerinden birisi etiğin hesaplanabilir olup olmadığıdır. Etiğin yapay bir uygulamasının yapılması için hesaplanabilir olması gerektiği fikri yapay zeka tartışmalarından çok önce felsefi bir problematik olarak ele alınmıştır. 18. yüzyılda Francis Hutcheson eylemlerin ahlakiliğinin hesaplanmasını bir tartışma konusu hâline getirmiştir. Her bir insanın ahlak duygusuna sahip olduğunu ve bu ahlak duygusunun temel nesnesinin iyilik olduğunu iddia eden Hutcheson'a göre iyilik hesaplanabilir bir olgu olarak, çoğunluk için üretilen iyiyle doğru orantılıdır (Moor, 1995: 2). Etik erdemler hakkında yargıda bulunurken yapılması gereken yol açtığı eylemlerin iyilik ve kötülük derecelerini hesaplamaktır.

Hutcheson'ın bu teorisi, etiğin hesaplanabilir bir alan olduğunu savunan temel görüşlerden bir diğeri olan eylem utilitaryanizminin de (faydacılık) çıkış noktasıdır. Jeremy Bentham'ın ilk savunucularından olduğu "mümkün olan en yüksek sayıda insanın, mümkün olan en yüksek miktarda mutluluğu" olarak tanımlanabilecek yarar ilkesini ahlakın temeli olarak kabul eden utilitaryanizm/faydacılık, eylemleri vermekte oldukları mutluluk nispetinde doğru/iyi sayar (Mill, 2017: 61). Mutluluğun zıddı olan şeyleri getirdikleri acı/zarar nispetinde yanlış/kötü görür. Mutluluktan hazzın varlığı veya acının yokluğu, mutsuzluktan da acının varlığı veya mutluluktan mahrumiyet anlaşılır. Eylem faydacılığının temel tezi şu şekilde formüle edilebilir: "Bir eylem ancak ve ancak öznenin onun yerine yapabileceği daha faydalı başka bir eylem yoksa doğrudur." (Feldman, 2016: 46). Bu formül sonuççu bir ahlak ilkesidir; eylemlerin ahlaki statüsünü tamamen sonuçlarına bağlı olarak değerlendirmektedir.

Faydacılığın ima ettiği gibi, eylemlerin ahlaki doğrulukları ya da yanlışlıkları, bunların haz ya da acıya yol açma eğilimlerine göre değerlendiriliyorsa; eylemde bulunan birey, yapacağı

eylemin doğru ya da yanlış olduğuna karar verirken eylemin neden olması muhtemel olan haz ve acı miktarını ölçmek ve birini ötekine karşı tartmak zorunda kalacaktır (Aydın, 2013: 605). Bentham, bu ölçme problemine binaen Felicific Calculus (Yanlışlama Kalkülü) ve Hedonistic Calculus olarak adlandırılan haz ölçeğini, iki eylem arasında karşılaştırma yapan ahlaki failler için niceliksel bir yöntem olarak sunmuştur. Bentham'ın öne sürdüğü faydacılık yaklaşımında, ahlaki doğruluk ve yanlışlıklar arasında nitelik değil, nicelik farkı vardır. Bentham'ın ahlak ölçeği, haz ve acıyı niceliksel olarak değerlendirmek üzere çeşitli ölçütler önermektedir. Bunlardan ilki, haz ve acının yoğunluğunu; ikincisi, süresini; üçüncüsü, kesinliğini; dördüncüsü, kişiye yakınlık ya da uzaklığını; beşincisi, başka haz ya da acılara yol açma potansiyelini; altıncısı, haz ya da acı yönünden saflığını; son olarak yedincisi ise haz ya da acının etki alanını eylemleri berimsel olarak sınıflandırmak için referans noktası olarak kullanmaktadır (Bentham, 2009: 139). Yedi birimlik ölçüm sonucunda bir eylemin doğru ya da yanlış olduğunun ortaya konulduğu bu yöntemde, ahlak ölçülebilir ve hesaplanabilir olarak ele alınmaktadır.

Ahlakın hesaplanabilir olması, haz ve acının ölçülebilir olması ve birimlerle ifade edilebilir olması ile doğrudan bağlantılıdır. Güzel bir kahvaltının verdiği haz 10 birim, sakin bir koyda yüzmenin verdiği haz 15 birim, yemeğin üzerine içilen çayın verdiği haz 5 birim olabilir. Buna benzer olarak, migren ağrısının verdiği acı 10 birim, bacağın kırılmasının verdiği acı 20 birim, aşk acısı ise 30 birim olabilir. Haz ve acılar, sadece kendi içlerinde değil birbirleriyle de işleme tabi tutulabilirler. Örneğin, migren ağrısı çekerken içilen çay, "10 acı - 5 haz = 5 acı" şeklinde hesaplanabilir. Bu hesaplama şekli bizi faydanın şu tanımına ulaştırmaktadır:

> Bir eylemin faydası, o eylemin sonuçları olarak gerçekleşecek tüm haz olaylarının birim olarak değerlerinin toplamından o eylemin sonuçları olarak gerçekleşecek acı olaylarının tamamının acı değerinin toplamını çıkardığımızda elde ettiğimiz sonuçtur (Feldman, 2016: 44).

Faydacılığın gündeme getirdiği ahlakiliğin hesaplanması, makine etiği için üretilen teorilere de çıkış noktası teşkil etmektedir. Yapay bir etik fail tasarlanmasının önünü etiği sembolleştirilebilir ve hesaplanabilir bir olgu şeklinde ele alarak açan faydacılığın yapay uygulamaları, makine etiği için deney ortamı oluşturmaktadır. Faydacı teorinin izlenmesi açısından makineler insanlardan çok daha avantajlıdır. Zira insanların doğrudan aritmetik hesapları yetenekleri gelişmiş değildir; yalnızca daha iyi sonuçlar verecek eylemleri tahmin edebilirler veya alışkanlıkların getirdiği eylemlere yönelebilirler. Bu sebeple insanların

doğru hesap gerektiren konularda yalnızca tahmine dayalı seçim yaptıkları için yanılma olasılıkları fazladır. Diğer taraftan, makinelerin aritmetik yetenek ve hızları insanlardan daha gelişmiş olduğundan tahmin yoluyla değil, hesap yoluyla hangi eylemin iyi olduğuna karar vererek hata payını en aza indirecektirler (Anderson, 2007: 18). İlaveten, insanlar taraf tutmaya meyillidir ve kendilerinin, yakınındakilerin ya da sevdiklerinin faydasına olacak eylemleri diğer insanları dikkate almadan yapabilirler; fakat duygusal bağları bulunmayan bir makine bu bağlamda çok daha tarafsız ve akla uygun davranacaktır. Öte yandan makineler yine olası sonuçları hesaplamak açısından insanlardan daha başarılı ve hızlı olacaklarından faydacı ahlak teorisini takip etmeleri daha kolay olacaktır.

Faydacılık, makine etiği açısından uygulanabilir bir etik teorisi olsa da bu teoriye en büyük eleştiri, en yüksek iyi uğruna insan haklarına zarar verebilecek eylemlere izin verebilme ihtimali üzerine yapılmaktadır. Faydacı etikte fedakârlık, doğruluk gibi erdemler, kendinde bizatihi iyi olan değerler değildirler. Kendinde iyi olan tek değer, kişinin hem nitelik hem de nicelik bakımından mümkün olan en yüksek hazzı gerçekleştirmesidir. Bu durumda eylemlerin meydana getirdiği haz ya da acı oranının hesaplanabileceğini düşünmek, bireyin öznel duygulanımlarını yok saymaktır.

Bunun yanı sıra, nafile ya da boşuna yapılan eylemler de faydacılığın temel ilkeleriyle çatışır görünmektedir. Faydacılığa yapılan bu itirazı aktarmanın ve anlamanın en iyi yolu örneklendirmektir:

> Bir postacının günlük posta dağıtımı sırasında yanan bir evdeki bebeği canı pahasına kurtarması onun ne görevidir ne de ahlaki yükümlülüğüdür. Ve de yangından bebeği kurtarmak o denli güç ve cesaret gerektiren bir eylemdir ki, yangının ortasına girerek bebeği kurtarmamış olsa dahi kimse postacıyı görevden kaçtığını düşünerek suçlayamayacaktır. Bebeği yangından kurtarmak yanlış da değildir; aksine takdire şayandır (Feldman, 2016: 76-77).

Nafile eylemler olarak adlandırılabilecek örnekler sınırsız bir şekilde çoğaltılabilir: Evsiz kalacağını bilerek tüm servetini çocuk esirgeme kurumuna bağışlamak, iflas edeceğini bilerek kendi köyündeki insanları istihdam etmek üzere bir fabrika kurmak, hasta bir çocuğa böbreğini bağışlamak için tehlikeli bir ameliyata girmek gibi vazifenin gereklerinin dışında yapılan eylemlerin tümü nafile eylemler olarak adlandırılabilir. Ancak eylem faydacılığı bu tür eylemleri doğru kabul etmemektedir, zira bu yaklaşıma göre, bir eylemin doğru olması, ancak kişinin onun yerine yapabileceği herhangi başka bir eylemden daha fazla fayda sağlamasına bağlıdır; ancak nafile eylemler faydadan çok zarar göze alınarak yapılan

eylemlerdir (Feldman, 2016: 77). Faydacılığın cevaplamakta en çok zorlandığı eleştiri, takdire şayan ve belki de örnek olması gereken eylemler faydacılık tarafından ahlaki olarak doğru bulunmamasıdır.

Ayrıca eylemlerin doğruluğunun ve yanlışlığının, eylemlerin gelecekteki sonuçlarına bağlı olması da adalet nosyonundan feragat edilmesine yol açmaktadır. Hak etmek, gelecek değil geçmiş davranışlar üzerinden değerlendirilmelidir. Yalnızca sonuca odaklanmak hak ve adalet kavramları için yeterli değildir. Bu sebeple hak ve adaleti dikkate alan bir başka etik teorisine ihtiyaç vardır.

Etik kararlar arasında en zor olanları şüphesiz hak ve ödevlerin çeliştiği durumlar ile içlerinden tercih yapılacak her eylemin ciddi ve kaçınılmaz zararlı sonuçlarının olduğu durumlardır. Bu gibi durumlarda, biz olası tüm eylemler arasından en adil olanını seçmeye çalışırız. Bir makine için adaletin ne anlama geleceği ise makine etiğinin başka bir problemidir. Kant'a kadar olan etik tarihinde, ahlak en yüksek iyi olan mutluluğa ulaşma çabası olarak tanımlanırken Kant'a göre duyulara bağlı bu tanım, insanı sadece teknik pratik akıl sahibi bir varlık olarak ele alması bakımından hatalıdır (Heimsoeth, 2016: 119). Oysa insan aynı zamanda salt pratik akla da sahiptir ve burada insanı yöneten duyular değil, doğrudan akıldır. Salt pratik aklın alanı ahlak bilgisinin alanıdır. O, gündelik hayatın ahlaki tecrübelerinden yola çıkarak, transendental yöntemle sağduyu sahibi herkesin kabul edeceği ahlak yasaları oluşturmayı hedefler. Bu bağlamda Kant pratik aklın yasalarını (*maksim/düstur*) şöyle ortaya koyar:

1- Öyle davran ki, davranışın temelindeki ilke, tüm insanlar için geçerli olan evrensel ilke veya yasa olsun!
2- İnsanlığı, kendinde ve başkalarında, bir araç olarak değil, her zaman bir amaç olarak görecek şekilde davran!
3- Öyle davran ki, iraden, kendisini herkes için geçerli olan kurallar koyan bir yasa koyucu olarak hissetsin! (Kant, 2016: 35).

Kendi yasasını iyi niyet ve ödev duygusuyla kendisi koyabilen bir varlık olan insan bu yüzden diğer canlılardan farklı ve değerlidir.

Ahlak yasası yalnızca bireysel bir ilke değildir; herkesin onaylayıp uymak isteyeceği evrensel bir ilkedir. Dolayısıyla Kant, koymuş olduğu bu ahlak yasası ile olması gerekenin adil bir sistem olduğunu da ayrıca vurgulamaktadır. Ahlak yasası bir doğa yasası da değildir

ve ahlak yasasına uymak zorunluluk değil ödevdir. Ödev, yerine getirmeyi kendi istek ve irademizle üstlendiğimiz, sorumluluğu bize ait olan, bizim kendi kendimize koymuş olduğumuz koşulsuz bir buyruktur (Özlem, 2004: 73). Kant'a göre, kimse ahlaklı olmak için zorlanamaz herkes kendi ahlaki gereklilik ve yükümlülüklerini kendisi için kabul etmelidir. İnsanlara karşı iyi görünmek için borçlarını ödeyen insan, doğru eylemi yapıyor olsa da ahlak idealine uygun değildir; ödev olarak gördüğü için borcunu ödeyen ve borç ödemeyi ahlaken doğru kabul eden insan ahlak idealine uygun olan insandır (Feldman, 2016: 200).

Kant etiğinde doğru eylem, ödevi gerçekleştirmeye yönelik olan eylemdir.[50] Kant'ın çizdiği çerçevede, ahlaki fail, kendi eğilimi ile ödev arasında sallantıda kalandır. Bu sallantı, bireyi tek başına ahlaklı yapmaz yalnızca bireyin ahlaklı olma potansiyeline sahip olduğunu gösterir. Ahlaklı bir birey olması ise yönelim ile ödev arasında kaldığında hangisini seçtiğine bağlıdır. Bu sebeple kategorik yani koşulsuz buyruklar Kant etiğini yapay zekalı makinelere uygulamak için tek başlarına yeterli değildir çünkü yalnızca akla dayalı eylemler, ahlaklı bireyler olmak için gerçek seçenekler değildir (Nath ve Sahu, 2017: 5). Kant etiği eylemlerin sonuçlarından bağımsız olarak, eylemlerin doğruluğunu ödevi yerine getirip getirmediğine bağlamaktadır. Kant etiği ahlakın rasyonelliğine, yüceliğine, idealine, hak ve adalete vurgu yapsa da içerisinde çok fazla metafizik unsur barındırması, küreselleşen dünyasındaki ahlak sorunlarına çözüm önerememesi ve eylemlerin sonuçlarını göz ardı etmesi bakımından makineler için uygun bir etik teorisi olarak değerlendirilmemektedir.

Yalnızca sonuçları dikkate alan faydacı ve hazcı teoriler ile sonucu bir kıstas olarak görmeyen Kant'ın deontolojik etiğinin tam bir etik teorisi olma bakımından eksik olduğu dile getirilirken makine etiğinin daha bütüncül bir yaklaşıma ihtiyaç duyduğu savunulmaktadır. Bu bağlamda, William David Ross'un Kant etiğini faydacılık ile birleştiren "deontolojik sezgici ahlak teorisi" olarak adlandırılan yaklaşımı da makine etiği için alternatif bir kuram olarak tartışılmaktadır. Ross, bir eylemin doğruluk veya yanlışlığını tek bir ölçüte göre değerlendirmek yerine, birden fazla ahlaki yükümlülük olduğunu dile getirmektedir. İyinin tek bir ilke olarak hazza, faydaya ya da şartsız emire indirgenemeyeceğinin iddia eden Ross'a göre, ahlak alanında çoğulcu ilkeler geçerli olmalıdır; yalnızca haz değil, bilgi, erdem ve adalet de iyi değerine sahiptir (Kala, 2018: 169). Ross'a göre, bir eylemi doğru kılan şey ne onun arkasında yatan niyet ya da sebep ne

[50] Kant etiğinde doğuştan, değişmez bir ödev yasası kabul etmekle ahlakına teolojik bir karakter verdiğini iddia eden bazı düşünürler, Kant'a "Sînâ Dağı'nda konuşan Musa" eleştirisi getirmişlerdir (Pazarlı, 1980: 97).

de sonuçlarıdır; bir eylemin doğruluğu yalnızca onun ne tür bir eylem olduğuna bağlıdır (Bayoğlu Kına, 2018: 227). Kant'ın ödev fikrini kabul etmeyen Ross, onun yerine "Prima Facie Duty"[51] kavramını kullanır. Ross, ilk bakışta ödev kavramını şöyle açıklar:

> Ben "ilk bakışta ödev"i başka bir deyişle "koşullu ödev"i (asli ödevden çok farklı olarak) bir eylemin belli bir türde (örneğin verilen sözü tutma) olmasından ötürü, aynı zamanda ahlaken önemli olan başka türde bir eylem olmasaydı bir asli ödev olacak bir eylem olma özelliğine atıfta bulunmanın kısa yolu olarak görüyorum. Bir eylemin asli bir ödev mi yoksa fiili bir ödev mi olduğu, örnek teşkil ettiği ahlaken önemli tüm türlere dayalıdır (Ross, 1930: 19-20).

Ross, ilk bakışta ödev ya da şarta bağlı yükümlülük türlerini de "sadakat, telafi, minnettarlık, adalet, iyilik, kendini geliştirme, zararı önleme ve zarar vermeme yükümlülüğü" olarak gruplandırır (Bayoğlu Kına, 2018: 229). Sadakat ödevleri aleni ya da zımni verilen sözleri, telafi ödevleri geçmişteki yanlışları telafi edecek eylemleri, minnettarlık ödevleri geçmişteki nezaket ya da yarara karşılık mahiyetinde eylemleri, adalet ödevleri ise iyiliklerin adaletsiz dağıtımını engelleyecek eylemleri kapsamaktadır (Feldman, 2016: 223). Ancak Ross'un yaklaşımı, ilk bakışta ödev türleri birbirleriyle çatışan ya da çelişen önerilerde bulunduğunda bir karar verme prosedürü gerektirmektedir. Hemen hemen herkes için apaçık doğru görülen bu yükümlülükleri Ross daha da geliştirilebilir görse de bu karar verme sürecinin de bir makine için nasıl işleyeceği muğlaktır. Ross sezgici olarak adlandırılır çünkü ona göre insanlar, sözlerini tutma, başkalarını daha bilgili, erdemli ve hoşnut kılma gibi ilk bakışta ödevleri olduğunu "sezme" becerisine sahiptirler (Feldman, 2016: 231). Bu sezgici tezin makine etiği için uyarlanması teknik olarak mümkün değildir; irrasyonel bu yetinin bir makinede oluşabileceğini iddia etmek yapay zeka çalışmalarının seyrine bakıldığında ulaşılabilir görünmemektedir.

Etiğin hesaplanamayacağı görüşünde olan erdem etiği teorisyenleri "etik ikilemler arasında kaldığımızda ne yapmamız gerektiğini değil, nasıl insanlar olmamız gerektiğini sorgulamalıyız" iddiasındadırlar (Anderson, 2007: 19). Sokrates, Platon ve Aristoteles ile birlikte Antik Yunan ile özdeşleştirdiğimiz erdem ahlakı, bir eylemin doğruluğunun ya da yanlışlığının sebebini tartışmak yerine, iyi insan olmanın gerekliliklerine odaklanmaktadır. Erdem etiği, sahip olmamız gereken nitelikler olarak erdemlerin temel tartışma konusu olduğunu, eylemlerin ise ikinci planda olduğu tezini savunmaktadır. Erdem etiğinde öncelikli amaç insanın neliğini ortaya koymaktır; olması gerekene nasıl ulaşılacağı bunun

[51] İlk bakışta ödev.

ardından konu edinilir. Ancak etiğin hesaplanamaz ve ölçülemez olduğu fikri ve ahlakı insanların sahip olduğu erdemler ile eşdeğer kabul eden tezi ile erdem etiği, makineler için uygulanamaz görünümdedir. Zira bir etik teorisinin makinede uygulanabilir olması zorunlu olarak hesaplanabilir bir etik olmasını gerektirmektedir. Makinelerin erdem sahibi olması da ancak bir tür programlama ile mümkün olacağından, eylemleri göz ardı eden ve karaktere odaklanan bir makine etiği teorisi mümkün görünmemektedir.

5.2. Makine Etiği İçin Yöntem Tartışmaları

Makine etiği alanı yapay ahlaka sahip sistemler oluşturma amacıyla çalışırken öncelik verdiği araştırma konularının başında ahlaki değerlendirme yapabilen makineler üretmenin imkânı ile yöntemini tespit etmek gelmektedir. Bir makinenin ahlaki ve etik değerlendirme yapabilmesi için farklı çevresel koşullarda nasıl davranacağını değerlendirebilmesi; bunun için de öncelikle eylemlerinin sonuçlarını tahmin edebilmesi ve bu sonuçların iyiliğini ya da arzu edilebilirliğini analiz edebilmesi gereklidir (Georges, 2003: 151). Makinede etik ilkeler oluşturulmasına dair iki temel yaklaşım söz konusudur. Bunlardan ilki yukarıdan-aşağıya bakış açısı olarak adlandırılmaktadır. Bu yaklaşımda, ahlaki prensiplere insandan yola çıkarak ulaşılması ve ulaşılan ahlaki prensiplerin daha sonra makineye kodlanması savunulmaktadır. İkinci yaklaşımda ise makinelerin ahlaki kodları çevreyle etkileşim deneyimlerinden çıkardıkları sonuçlara dayalı olarak, kendi başlarına programdan ve programcıdan bağımsız bir şekilde öğrenmesi gerektiği öne sürülmektedir (Georges, 2003: 153). Evrimsel psikolojiden etkilenen bu yaklaşım aşağıdan-yukarıya bakış açısı olarak da adlandırılmaktadır. Her yaklaşımın kendi içerisinde avantaj ve dezavantajları vardır:

İnsandan direk olarak yüklenen ahlaki kodlar "hızlı ve kirli" olacaktır; antropomorfik eğilimli olacak ve bu sebeple sınırlı, gelişigüzel ve kendi içerisinde çelişkili olacaktır. Deneme yanılmayla öğrenilen ahlakilik ise "yavaş ve pahalı" ve beklenmeyen sonuçlara gebe olacaktır. Öğrenilen kodlar, çevreden edinilen yeni bilgiler ile güncellenmeye daha müsait olsa da insanda olduğu gibi içsel çelişkiler barındırması kaçınılmaz olacaktır (Georges, 2003: 155).

İnsan ahlakı ise hem verili normlar hem de aşağıdan yukarıya bakış açısını kendi içerisinde barındırmaktadır. İnsanlar ahlaki kodlara, insan doğası gereği doğuştan verili olan kodlar ile öğrenilmiş ve edinilmiş kodların birleşimi şeklinde sahiptirler. Bazı temel ahlaki kod ve

bilgiler beyinlerimizde genlerle aktarılmış olarak vardırlar. Hayatta kalma ve üreme dürtüleri tüm canlılarda ortak bir şekilde bulunmaktadır. Bunun dışında, değerler, prensipler ve ahlaki kodların büyük bir kısmı ise hem deneme yanılma yoluyla elde edilen deneyimlerle hem de rol model olarak görülen aile bireyleri, öğretmenler ve liderler gibi farklı kişilerden gerek aktarım gerek gözlem vasıtasıyla elde edilirler.

İki temel yaklaşım olarak hem yukarıdan-aşağıya hem de aşağıdan-yukarıya bakış açıları tek başlarına değerlendirildiğinde kendi içlerinde eksik oldukları görülmektedir. Her iki yaklaşım da bütüncül bir ahlak sistemi oluşturmak için yeterli değildir. Bütün ahlaki normların, insanda ya da makinede en başından verili olması mümkün değildir. Yukarıdan-aşağıya bakış açısı insanda bulunan ahlaki normların, aynı şekilde makineye yüklenmesi gerektiğini savunsa da hem insanın ahlaki kodlarının değişkenliği hem de ahlakın tek başına değil toplumda ortaya çıkan bir mefhum olması sebebiyle, makinenin tek seferde ahlaki olarak kodlanması mümkün değildir. Bir makinenin ahlaki olarak nitelendirilmesi, karşılaştığı olaylar karşısında ahlaki değerlendirme yapabilmesine bağlıdır. Otonom bir makinenin karşılaşacağı tüm çevresel etken ve olayların önceden tahmini mümkün olmadığından, ahlaki yapısının önceden kurgusu da mümkün değildir. Tam tersi olarak aşağıdan-yukarıya bakış açısı ise tüm ahlaki kodların deneme-yanılma yoluyla makinenin kendisi tarafından öğrenilmesi gerektiğini salık vermektedir. Ancak temel hiçbir prensip bulunmadan makinenin kendi ahlak yapısını oluşturmasını beklemek umulmadık tehlikelerin doğmasına yol açabileceğinden ve hiçbir çıkış önermesi bulunmayan bir sonucu elde etmenin güvenilir olmamasından dolayı aşağıdan-yukarı bakış açısı da makine etiği için tek başına yeterli bir yöntem değildir. Buradan hareketle, makine ahlakı için de insanda olduğu gibi hem verili kuralların olması gerektiği hem de makinelerin gelişimsel olarak ahlaki ilkeler edinmeye müsait yapıda olmaları gerektiği makine etiği çalışmalarında üzerinde neredeyse uzlaşıya varılmış bir görüş olarak öne çıkmaktadır.

Aşağıdan yukarıya ve yukarıdan aşağıya bakış açılarını birleştiren bu görüşte, yapay ahlak sistemlerinin oluşturulması için yalnızca bilgisel ve işlevsel bir mekanizmanın kâfi olmayacağı da ayrıca vurgulanmaktadır. Zira duygular, zihin, empati, sembollerin semantik içeriğini anlama, çevreyle bedensel etkileşim ve bilinç gibi rasyonelliğin ve işlevselliğin üstünde kapasiteler ve sosyal mekanizmalar da ahlaki yapı içerisinde rol oynayan önemli faktörlerdir (Wendell, Allen ve Franklin, 2011: 179). Sınırlı bir çevre içerisinde kalan ve sosyal etkileşimi bulunmayan makinelerin ahlakiliği için bu rasyonellik üstü yetilerin

gerekliliği hayati bir değere sahip olmasa da bir sistem insanlar ve diğer canlılarla sosyal etkileşim içerisine girdiği andan itibaren bu yetilerin önemi oldukça artmaktadır. Sosyal olmayan bir yapay ahlak oluşumunun bilgisel ve işlevsel kurulumu yeterli olsa da sosyal yönü olan yapay ahlak sistemlerinin oluşmasında diğer faktörleri devre-dışı bırakmak oldukça tehlikeli sonuçlara yol açabilir.

Ahlaki yapının en önemli faktörlerinden birisinin duygular olduğu kabul edilir; zira birçoğuna göre makine bilinci, insan bilincinden duygular ve amaçlar yönünden farklılaşmaktadır (Kaku, 2017: 267). Duyguların makine bilinci ve makine etiğindeki rolü sosyal etkileşim için zorunluluklarından kaynaklanmaktadır. İnsan ve robot arasındaki etkileşim için duyguların olmazsa olmazlığı makine etiğinin temel kabullerinden sayılmaktadır. Ayrıca, nörobilimdeki gelişmeler, duyguların insan zihni için olan öneminin şimdiye kadar büyük ölçüde göz ardı edildiğini göstermektedir. Bu konuyla ilgili gözlemsel çalışmalar göstermektedir ki; insan beynindeki düşünme ve akıl yürütmeden sorumlu prefrontal lob ile limbik sistem gibi duygusal merkezler arasındaki ilişki hasar aldığında hastaların değer yargıları kaybolmaktadır (Kaku, 2017: 271). Duygusal merkezle kurulan ilişki zarar gördüğünde, hastalar karar vermede zorluk çekmekte; seçim yapamamaktadır. Seçenekler arasında onlar için duygusal olarak hiçbir fark bulunmadığından, her şeye eşit değer atfetmektedirler. Değer bakımından derecelendirme yapamadıkları durumda, etik ve ahlaki yargıda bulunmaları söz konusu olmamaktadır. Bunun dışında, ancak duygu sahibi bir birey, başkasının duygu durumu hakkında empati yapabilir. Acı çeken birini, ancak acı hissine sahip başka bir varlık anlayabilir. Duygulara sahip olmayan bir makinenin, etik ikilemde kaldığı durumlarda başkalarının acı oranlarını nasıl hesaplayacağı belirsizdir (Nath ve Sahu, 2017: 6). Bu sebeple etik makinelerin inşası için rasyonalitenin dışında duyguların gerekliliği de önemli bir unsurdur. Yapay zeka çalışmalarına duyguların dâhil edilmesi fazladan ve gereksiz bir çaba olarak görülmemelidir; duygular, özellikle makine etiği için çevresel bir konu değil, merkezi bir tema olarak değerlendirilmelidir.

Duygular, zihin, empati, sembollerin semantik içeriğini anlama, çevreyle bedensel etkileşim ve bilinç gibi ek yetilerin hemen hepsi için ayrı uzmanlık alanları bulunmaktadır. Makine bilinci, robotik, genel yapay zeka gibi alanlar bu yetiler üzerine yoğun araştırmalarda bulunsa da çalışmalarını etik problemleri hesaba katmadan sürdürmektedirler. Makine etiği alanı ise bu çalışmalara ek olarak, duygulara sahip olan ve olmayan yapay sistemlerin, çevreyle fiziksel etkileşimi bulunan ve bulunmayan sistemlerin, bilinçli ve bilinçsiz yapay

sistemlerin ahlaki karar verme mekanizmaları arasındaki farka büyük önem vermektedir. Bilince sahip olmayan makinelerin ahlaki olarak kabul edilebilir eylemler üretmede sınırlılıkları ve avantajlarını ayrı bir inceleme konusu olarak ele alınmaktadır. Ancak mühendislik ve bilişim alanındaki ilerlemelerin, makine etiği çalışmalarından habersiz devam etmeleri uygulamada mühim sıkıntılara yol açabilecek görünmektedir. Etik kaygılardan bağımsız inşa edilen yapay sistemlere sonradan eklenecek etik ilke ve mekanizmalar uyum ve birleşmede ciddi sıkıntılar doğurabilir. Zihinsel gerçeklik ile ahlaki statünün gerçekliği sosyal güvenlik açısından birbiriyle sıkı bir içerisinde olmalıdır; zaten zihin ile ahlakın birbirinden bağımsız değerlendirilmesi çok büyük bir kuramsal eksiklik olacaktır (Torrance, 2008: 498). Bu sebeple hem makine bilinci çalışmalarının hem de diğer yapay zeka çalışmalarının makine etiği ile iş birliği içerisinde olması, olumsuzluklar üzerine kurulu tekillik senaryolarının gerçekleşmesini engellemek adına büyük önem arz etmektedir.

Makine etiği/ahlakiliği ile makine bilinci alanlarını ortak paydada buluşturmak için öncelikli olarak yapılması gereken ahlak ve bilinç kavramlarının içerikleri ile bunlara olan yaklaşımları net bir şekilde tespit etmektir. Bilinçli makineler üretmeyi başarırsak, bu aşama oluşacak sosyal ilişkilerin düzenlenmesi açısından yeni bir dönemin başlangıcı olacaktır. Otonom ve bilinçli makineler bu aşamada artık bizim için yalnızca işlevsel araçlar olmayacaklar, bizim sosyal varoluşumuza bizimle birlikte katılımcı olacaklardır (Gamez, 2008: 906). Sosyal düzen içerisinde yer alan bir makinenin ahlakiliği makine bilinci çalışmalarının odaklanması gereken bir diğer meseledir.

Ahlaki değerlere sahip makineleri tanımlamak ve kurgulamak düşünüldüğünden çok daha zordur. Çünkü tıpkı bilinç gibi, ahlak da tanım ve sınırları net olmayan bir kavramdır. Mahiyeti tartışmalı bir kavramın mekanik olarak inşası ise yeni tartışmaları beraberinde getirmektedir. "Bir yapay zeka ne zaman ahlaki bir fail olur?" sorusu, "Bir fail ne zaman ahlaki olur?" sorusuyla doğrudan ilgilidir (Parthemore ve Whitby, 2013: 106). Dolayısıyla makine etiği çalışmalarının beslendiği temel kaynak insan ahlakıdır. Ayrıca, makine etiği, insanın ahlaki ve etik kodları üzerine düşünürken yapay sistemleri merkeze koyarak, ahlak alanın metafizik sorular ile çeşitli efsaneler ve inanışlardan arındırılmış bir şekil kazanmasına da destek olabilir. Bugün geçmişten farklı olarak ahlakın ilgi alanı genişlemiş, bilişim suçları, biyo-etik, çevre etiği, kök hücre çalışmaları, genetik kopyalama, kürtaj, organ nakli, insan hakları, hayvan hakları ve ötenazi gibi problemlerle çeşitlenmiş, daha önce mesafeli durduğu ya da yok saydığı teknoloji, ekonomi politikaları, genetik, çevre vb.

alanlara yönelmek durumunda kalmıştır. Dolayısıyla insan tutum ve davranışlarının bir sonucu olan tüm bu sorunlara çözüm bulma konusunda zeki sistemlerin insanlığa çok şey katacağı muhakkaktır. Yapay zekaya sahip sistemler, doğayı ve insanı sürekli yenileyerek gezegeni daha yaşanır hâle getirebileceği gibi aksi senaryoların dile getirdiği gibi insanlığın sonunu da getirebilir. Ancak yapay zekalara verilecek bu etik kodların evrensel mi olacağı yoksa tarihsel mi (paradigmatik), kültürel mi olacağı sorunu yine felsefenin tartışma alanına girecektir.

Ahlaki yönü olan birey birçoğu tarafından Darwin'in tanımında olduğu gibi "geçmiş ve gelecek eylemlerini karşılaştırabilen ve bunların bazısını onaylayan ve bazısını onaylamayan bir birey" olarak tarif edilir. Kendi eylemleri üzerinde düşünme ve iyi ya da kötü yargılarında bulunma yetisine sahip olan insan türü ise bu sebeple ahlaki bir varlık olarak ele alınır. Zeki bir makinenin ahlakiliği bu tanıma göre, kendi eylemleri hakkında ahlaki yargılama yapıp yapamamasına bağlı olarak değerlendirilmektedir. Ancak bu türden bir ahlakilik tarifi bireyi tek başına ele alması bakımından eksiktir. Zira bu tarife göre, satranç oynayan bir bilgisayarın ahlakiliğinden de bahsetmek mümkündür. Çünkü satranç oyunu her hamlenin önce ve sonrasının değerlendirilmesini gerektirir ve bilgisayar da oluşturduğu karar ağacından faydalanarak bunu bir insandan çok daha iyi ve hızlı bir şekilde yerine getirebilir. Bu tanımdaki eksiklik, bireyin sosyal yönünü dikkate almamasından ileri gelmektedir. Ahlakilik ancak bir toplumsal yapı çerçevesinde anlamına kavuşur. Tek bir insanın veya insan topluluğunun belli bir tarihsel dönemde belli türden yönelim, düşünce, inanç, alışkanlık, gelenek, görenek ve bunlarda bulunan değer, emir, norm ve yasaklara göre düzenlenmiş ve bu hâliyle kalıcı hâle gelmiş yaşam biçimine ahlak adı verilir (Özlem, 2004: 15). Gerçek anlamda ahlaki bir birey dış dünyayla etkileşim içerisinde olmalıdır; bireyin sahip olduğu beden ya da fiziksel form zamansal ve mekânsal duyarlılık ile çevreyle ilişki kurabilmelidir (Parthemore ve Whitby, 2013: 105). Kavanozda saklanıp gerekli organik şartlar sağlanarak yaşatılan bir beynin ahlakiliğinden söz edilemez. Ahlakilik ancak çevreyle etkileşime giren bir beden ile buna bağlı bir zihnin varlığında ortaya çıkan bir mefhumdur. Dolayısıyla gerçek bir ahlakilik bir makinede üretilmek isteniyorsa o makinenin beden sahibi olan bir formu olmalı ve bu form çevresiyle ilişki kurabilmelidir. Aksi takdirde, kendi içerisinde çalışan bir programdan ibaret olan bir makinenin, çevre için çıktı üretmediği ve de çevresinden girdi almadığı sürece ahlaki yapısı üzerine tartışmanın bir anlamı olmayacaktır.

Ayrıca ahlaki makinenin kurucusu olacak etik kodların kaynağı da tartışmalıdır. Bu etik kodların insan türünün etik kodlarından mı türetileceği yoksa yepyeni kurallara mı dayalı olacağı makinelerin ahlakiliği meselesinin önemli bir tartışma unsurudur. Ahlak üzerine düşünüp tartışan bir insanın ilk karşılaştığı zorluk, ahlakların çokluğudur. Bu çeşitli ahlak türlerinin varlığını görmek, iki temel etik yönelimin de başlangıcıdır. Ahlakların çokluğundan hareketle ulaşılan ahlaki görelilik ile tek ve evrensel bir ahlak geliştirmek adına evrenselciliğe başvurmak İlk Çağ'dan günümüze kadar sürüp gelen iki etik doğrultudur (Özlem, 2004: 19). Sofistlerin "Her şeyin ölçütü insandır." tezine dayalı rölativist ahlak ile Sokrates ve Platon'un evrensel ahlakı arasında ilk kez net şekilde görülen bu felsefi karşıtlık düşünce tarihi boyunca farklı bakış açıları üzerinden var olmaya devam etmiştir. Konumuz açısından değerlendirecek olursak hem göreli hem de evrensel ahlak yaklaşımlarının makine ahlakı için uyarlanması farklı zorluklara sebebiyet verecektir. Rölativist ahlak yaklaşımının nasıl sembolleştirileceği bir problematik iken makineler için evrensel bir ahlak yasası tespit etmek de başka bir problematiğin konusu olacaktır.

Evrensel bir ahlak yasası teorik olarak mümkün gözükse de bugüne kadar ittifak edilmiş bir yasa olmadığı gibi, evrenselci kuramların savunduğu tezler de birbiri ile çelişmektedir. Bu konuda evrensel ahlaki ilke olarak dillendirilen en önemli motto "Sana yapılmasını istemediğin bir şeyi sen de başkasına yapma!" şeklindeki Altın Kural'dır. Bu kuralın benzer ifadelerine dinî, felsefi, ahlaki birçok metinde rastlamaktayız.[52] Yapay zeka çalışmalarında insanlığın ürünü olan bu ahlaki tecrübelerin de dikkate alınması gerektiğini söylenebilir.

Rölativist ahlaka göre etik kodlar, zaman içerisinde değişime uğrar ve kültürden kültüre farklılaşır. Bir kültürde namus için eşini öldürmek ahlaki olarak doğru ve alkol tüketmek ahlaki olarak yanlış olmasına rağmen, başka bir kültürde bunlar tam tersi bir şekilde kabul edilebilir. Dolayısıyla ahlak kesin bir şey olmak yerine birey ve çevreye bağlı bileşenleri olan değişken bir yapıdır. Ahlakiliğin sosyal, kültürel ve inanç ögelerine bağlı olan değişken yapısı ahlaki makineler üretirken de dikkate alınmalıdır. Ahlaki bir makine etkileşimde bulunduğu insan ve makine topluluğunun ahlaki koşul ve standartlarına dayalı olarak

[52] Altın Kural'ın benzer ifadelerine şu örnekler verilebilir: "İnsanların size nasıl davranmasını istiyorsanız siz de başkalarına öyle davranın! Çünkü Kutsal Yasa'nın ve peygamberlerin söylediği budur." (Matta: 7: 12; Luka, 6: 31); "Başkalarının sana yapmalarını istemediğin şeyi sen de onlara yapma!" (Konfüçyüs); "Başkaları sana yaptığında kızacağın şeyi sen de başkalarına yapma!" (İsokrates); "Kendine nasıl davranılmasını istiyorsan başkalarına da öyle davran!" (Mahabharata); "Başkalarının sana yapmalarını istediğin şeyi sende onlara yap!" (Thomas Hobbes); "Başkaları için istemediğin şeyi kendin için de isteme!" (Spinoza); "Benzer bir durumda başkalarının sana yapmasını isteyeceğin şeyi sen de başkalarına yap!" (Kropotkin).

türetilen kendi ahlak kurallarına uymaya çalışacaktır. Makinelerin içerisinde bulundukları sosyo-kültürel çevrenin kabul ettiği norm ve değerler, ahlakiliğin değerlendirilmesinde muazzam bir kıymete sahip olacaklardır.

Ahlakın rölativist olduğunu savunanlar için iki önemli dayanak noktası vardır. İlki, farklı toplumlar farklı gelenek ve kültürlere sahip olacaklarından ahlak sistemleri de farklı olacaktır görüşüdür. Ancak bu noktada etik meseleler ile etikten bağımsız gelenekler net bir şekilde ayrılmalıdır. Çünkü sadece gelenek olduğu fikrine yaslanarak, etik olmayan bir eyleme tolerans gösterilmesini beklemek yanlıştır. Gelenek ve görenekler, etik sınırlar içerisinde saygı duyulan kültürel ögeler olmalıdır. Örneğin bir toplumun aynı ailenin sahip olacağı ikinci kız çocuğunu öldürmek gibi bir gelenek oluşturmuş olması, o toplumun ve o çocuğu öldürenlerin suçlu olduğunu ve öldürme eyleminin etik dışı olduğunu değiştirmemektedir.

Rölativist ahlak temsilcilerinin ikinci dayanak noktası ise etik uzmanlarının bile hangi eylemin doğru olduğu konusunda hemfikir olamadığı etik ikilemlerin varlığıdır. Bu durum, evrensel bir etik ilkenin olamayacağına örnek olarak gösterilmektedir. Ancak, gözden kaçırılmamalıdır ki, bu tartışmalı etik meselelere yönelik her türlü görüş aynı oranda doğru olamaz. Etik ikilemin varlığı, ahlakın çokluğundan değil, meseleyi çözmenin zorluğundan kaynaklanmaktadır. Bu durumda yapılması gereken çoğunluk tarafından kabul edilen ve çoğunluğun inançlarıyla çelişmeyen en doğru yaklaşımı tespit etmektir (Anderson, 2007: 20). Bu sebeple makinelere tek seferde bütün etik ikilemleri aşabilecek ilkeler yüklemek imkânsızdır. Etiğin makineler üzerinde uygulanmasının eksiksiz olmasını beklemek hata olacaktır, zira eksiksiz ve kusursuz bir etik, teori olarak da mevcut değildir. Eksiksizlik, etik konusundaki ideal olsa da şu an için bu ideale ulaşılması çok mümkün görünmemektedir. Bu sebeple amaç güncellemelere açık bir etik teori ortaya koymaktır; karşılaşılan her yeni etik ikilemde yeni bakış açılarını kendisine ekleyebilecek bir etik uygulama, mükemmele en yakın teoriyi oluşturacaktır.

Tam bir etik teori üretmekten daha önemli olan, çelişkisiz bir teori oluşturabilmektir; bu husus makinelerin hızlı hesaplama yetileriyle insanlara çelişkileri tespit etmek hususunda yardımcı olabilecekleri bir alan olarak karşımızda durmaktadır (Anderson, 2007: 20). Bir diğer açıdan da çelişkili etik ilkeleri makineye yüklemek hata vereceğinden, makine etiği çelişkilerin çözüme kavuşturulmaları hususunda etik teorisyenlere yol gösterici olacaktır.

İlk bakışta zeki bir makinenin ahlaklı bir makineye çevrilmesi basit bir programlama olarak düşünülebilir. Örneğin, eklenecek "On Emir"[53] (*Evâmir-i Aşere*) gibi ahlaki öğretiler ile ahlaki bir makine elde edilebileceği varsayılabilir. Asimov'un *Ben Robot*'unda[54] olduğu gibi insana zarar vermeyecek şekilde kodlanacak bir makinenin ahlakiliği de savunulabilirdir. "Zarar verme!" ilkesi üzerine etik bir makine inşa etmek yüzeysel bir değerlendirmeyle oldukça kolay görünmektedir. Zarar verme emri üzerine bir makine inşa etmenin altında yatan fikir kabaca şu şekildedir:

> Bir makineye zararın ne demek olduğu ve ne tür eylemlerin zararlı olduğu öğretilir ve daha sonra kimseye zarar vermemesi yönünde programlanır. Bu etik bir fail üretmek için çok umut veren bir başlangıçtır çünkü yaygın ahlaki olmayan eylemler (öldürmek, yaralamak, kandırmak, hırsızlık yapmak, çatışmak, etkisiz hâle getirmek vs.), zararın belli çeşitlerini içeren kötü olarak nitelendirilen eylemlerdir. Robotlarımızın bu tür eylemlerde bulunmasını istemeyiz (Moor, 1995: 8).

Fakat "zararsız" olarak inşa edilen bir robot dahi her zaman hesaplanmayan zararlar verme potansiyeline sahiptir. Bunun için bir makinenin tümüyle zararsız olarak tasarlanması mümkün değildir, ancak öngörülebilen zararlara yol açmayan bir makine tasarlanabilir. Uyurken bizi rahatsız etmesinin bize zarar vereceğini bilen bir makinenin, biz uyurken ev yandığından bizi uyandırmaması mümkündür. Başkalarına söylememesi gereken gizli bilgileri (sırrı) yalan söylememek için söyleyebilir. Bu gibi durumlarda, makinenin zarara yol açmamak dışında, zararı önlemek gibi bir koda da sahip olması gerekir. Fakat hem zarar vermemek hem de zararı engellemek üzere oluşturulan bir makinenin ikisinden birisini tercih etmesi gerektiği durumlarda neye göre seçim yapacağı tartışmalıdır. Faydacılık bize zararların niceliksel olarak sıralamasını yapmayı önerirken bunun tüm olasılıkları kapsayıcı bir listesinin yapılması pratikte uygulanması zor bir yaklaşımdır.

Zarar vermenin içeriği ve boyutları bile bu kadar belirsizken konulacak tüm ilkelerin ahlak için kapsayıcı olmayacağı açıktır. Bu düşünülebilecek her kural için de geçerlidir, tüm insani davranış ve sosyal ilişkileri kuşatacak kesin kurallar belirlemek oldukça zordur. Bir yandan da kurallar ne kadar ayrıntılı olursa muğlaklık ve çelişkilerin ortaya çıkma ihtimali de o kadar artmaktadır. Bir kuralla yasaklanan şeye, başka bir kuralla açık kapı bırakılması ihtimali doğmaktadır. Ayrıca kurallar belirlenirken kullanılan tüm isim, sıfat ve yüklemlerin atıflarının da net olarak tespit edilmesi gereklidir. İnsani etik kodların ve hukuk sistemlerinin

[53] Hz. Musa'ya Sînâ Dağı'nda taş tabletler üzerine yazılı olarak verilen ilk ilahî emirler.

[54] *Ben Robot* isimli eserde Isaac Asimov'un ileri sürdüğü Üç Robot Kanunu'ndan bu çalışmanın Giriş Bölümü'nde bahsedilmiştir.

muğlaklığı da ayrıca ortadadır. Mahkemeler ve hâkimlerin görevi bu muğlak ve çelişkili durumlar karşısında vicdanının sesini dinleyip rasyonel değerlendirmeler yaparak karar vermektir. Aksi takdirde, çelişkisiz ve bütünüyle açık bir hukuk sistemi, hâkime gerek duymayacak bir yapıda olacaktır.

Makinelerin ahlaki olmayan davranışlarını engelleyecek, dikkatlice düşünülmüş, doğru kurallar kümesi oluşturmak mümkün değildir çünkü tüm insanların bu kuralları olduğu gibi ve ayrıntılarıyla kabul etmeleri imkânsızdır (Georges, 2003: 154). Kurallar kümesi oluşturulurken belirleyici olacak uzmanların sahip oldukları farklı toplumsal, siyasî, kültürel, etik, dinî kodlar, onların makineye aktarmak isteyeceği ahlaki kuralları doğrudan şekillendirecektir. Tüm sosyo-kültürel faktörlerden arınmış tarafsız bir etik kural dizisi oluşturulması gereklidir, ancak bunun da insan eliyle üretilmesi uygulanabilir görünmemektedir. Makinelerde ahlaki yapı oluşturulmasını zorlaştıran ikinci bir faktör de akıllı bir makinenin çevresinden sürekli olarak yeni bilgiler edinecek olmasıdır. Bu yeni bilgiler, sürekli olarak var olan kurallarla bütünleşmek zorundadır. Yeni bilgilerin etik kurallar kümesinde mutlaka bir karşılığı olmalı ve verili kurallarla çelişkili olmamalıdır. Ancak bu öngörülemez durumun, bilgisayar ortamında programlanması pek de olası görünmemektedir. Ve son olarak, kurallar etik geleneklere ve insan davranışlarına bakarak oluşturulacağından, bunların bir kısmı boş ve gereksiz kurallar olacak ve bir kısmı da bilgisayar algoritması ile ifade etmek için uygun kurallar olmayacaktır (Georges, 2003: 154).

Tüm bunlardan hareketle, ahlaki normların dışında bir davranışta bulunmayacak bir makinenin oluşturabileceği asla garanti edilemez görünmektedir. Hem insanlarda hem de hedef üretebilen makinelerde, çatışan içsel değerler beklenmedik sonuçlara yol açabilir. İnsani alanda sahip olunan tüm gelenek ve kanunlara rağmen hâlen çok sayıda kural ihlalinin olması da bunun açık bir göstergesidir. İnsanın bütün olarak ahlakiliği dahi kurallar, kanunlar, cezalar ve yasaklara rağmen sağlanamazken bütün olarak ahlaklı makineler üretilmesi oldukça iyimser bir düş olmanın ötesine geçemez görünmektedir.

Tüm ahlaki ve genetik kodlamalara rağmen, insanlar ahlaki olarak içsel çelişkilere sahip olmaya müsaittirler. Etik bir ilkenin çeşitli çıkar ve amaçlar doğrultusunda çiğnenmesi insanlarda oldukça sık karşılaşılan bir davranıştır. Bu bağlamda hırsızlık yapmak ve yalan söylemek davranışlarını ele alabiliriz: Bir çocuğa ailesi tarafından hırsızlığın ne kadar kötü bir şey olduğu ne derece öğretilmiş olursa olsun, o çocuğun çevresinden çıkarsadığı, hırsızlığın para kazanma için en zahmetsiz yol olduğu fikriyle hırsızlık yapma ihtimali her

zaman vardır. Benzer şekilde toplumsal refahı, mutluluğu ve hazzı artırdığı oranda hırsızlığı meşru gören ahlaki yaklaşımlar da söz konusudur (Çilingir, 2019: 114). Bir insanın sağlam karakterli ve ahlaklı olarak tanımlanışındaki en büyük etken ahlaki içsel çelişkilerden muaf olmasıdır. Ahlak konusunda kendi içerisinde çelişmeyen değer yargılarına sahip olması, kendi ahlaki prensiplerini belirlemesi, kendi prensiplerine uygun olarak hareket etmesi ve bu prensiplerin sosyal normlarla çelişmemesi bir insanın ahlaklı olarak adlandırılması için yeter sebeplerdir (Georges, 2003: 156). Ancak insanların seçimleri genel olarak menfaate dayalıdır. Her insanın sahip olduğu öncelikler ve ihtiyaçlar kendi dünyası içerisinde belirlidir ve eğer bu öncelikli ihtiyaçların karşılanması ivedilikle gerekliyse, o insan şahsi menfaatini sahip olduğu değerlere kolaylıkla değişebilir. O an için ailesini doyurmak zorunda olan bir birey hırsızlık yapabilir ya da bir bilim adamı yüksek bütçeli bir projenin başında olmak için ulaşmış olduğu sonuçları yanlış lanse edebilir. Bu eğilim ve kavrayışı makineler için genişlettiğimizde, makinelerin de insanlar gibi beklenmeyen ve ahlaki olarak yanlış olan davranışlarının olacağını açıkça görebiliriz. Bu sebeple en gelişmiş makinelerin dahi öncelik verilmiş kurallara ihtiyacı vardır.

Makinelerde ahlak konusu ile ilgili bir diğer önemli mesele ise makinelerin kendi içlerinde oluşturacağı sosyal çevrenin bağlı olacağı ahlaki ve etik normlarının ne şekilde belirleneceğidir. Bu hususta cevabı aranan soru şöyle ifade edilebilir: Makinelerin hedefleri birbirleriyle çeliştiğinde, aralarındaki etkileşimi düzenleyecek ahlaki ve etik kodlar nasıl olmalıdır? (Georges, 2003: 157). İnsani etkileşimi düzenleyen ahlaki ve etik kodların büyük bir kısmı, sosyal yapıların oluşturdukları tarihsel birikimle elde edilmiştir. İnsani ahlak alanının tarihi, insanlık tarihi kadar eskidir. Bu alanı şekillendiren kültürel, etnik, siyasî, ekonomik, dinî ve sosyal faktörler hem birbirlerine sıkı bir şekilde bağlıdırlar hem de ayrıştırılamayacak niteliklere sahiptirler. İnsanlık tarihinin ve dolayısıyla insani ahlak tarihinin bu uzun geçmişine rağmen, insanın karşılaşacağı bütün uyumsuz durumlara ve karşılaşılan tüm ahlaki problemlere çözüm sunan tek bir ahlaki sistemden bahsedilememektedir. Buna karşılık makinelerin ise şu an için hiçbir ahlak geçmişi ve etik tarihi yoktur. Şimdilik yalnızca çeşitli protokoller ve belirli kurallara göre inşa edilmektedirler. Zira makinelerin bugün kendi aralarında oluşturdukları sosyal ve kültürel ortamlar mevcut olmadığı gibi, ilişkilerini düzenleyen ekonomik dengeler, dinî inanışlar, siyasî eğilimler de söz konusu değildir. Gelecekteki makinelerin ahlaki kodları ise şimdikilerden farklı olarak bireysel değil, deneyim ve hedeflerini paylaştıkları, içerisinde etkileşimde bulundukları makine toplumları veya toplulukları tarafından kullanılan kodlar

olabilir. Bu şekilde esnekliğe sahip olan makine toplulukları, tıpkı insan topluluklarında olduğu gibi, tarihsel süreç içerisinde kendilerine ait ahlaki ve etik değerler kümesini oluşturabilir ve yeni deneyimlerine göre bu kümeyi güncelleyebilir ve genişletebilir konuma gelebilirler (Georges, 2003: 157). Bu da öncelikle makinelerin bir topluluk ya da toplum olarak görüleceğini varsaymaktan geçmekte, daha sonra bizi bir toplumun sahip olacağı ahlak tarihine makinelerin de sahip olabileceği fikrine ulaştırmaktadır. Sonuçta tarihsel bir varlık olan insan, geçmişi ile var olup ancak belirli bir zaman, mekân, ortam ve topluluk içinde varoluşunu ve kendine özgü yaşam hikâyesini gerçekleştiriyorsa yapay zeka için de aynı şartların geçerli olacağını, bugünün paradigmasında var olan yapay zekalar ile ileriki zamanlarda kendi zamanlarının ürünü olarak var olması muhtemel yapay zekaların farklılaşacaklarını, dolayısıyla yapay zekalar için oluşturulacak ahlaki kodların da yaşanan zaman ile paralel olması gerektiği söylenmelidir.

5.3. Mantık ve Yapay Ahlak İlişkisi

İyi ahlak ve kötü ahlak arasındaki ayrımın kesin çizgileri olmaması, insan zihninin genel olarak yargı ve değerlendirmelerini ikicil ya da iki değerli sınıflandırmalarla yapmaya eğilimli olmasından ileri gelmektedir. Olgu ve olayların özelliklerinin birbirinden çok temelli bir biçimde ve birbirlerine indirgenemezcesine farklı olarak iki temel parçaya bölünmesi (dikotomi) zihnin aşamadığı durumlardandır. Aristoteles'in "Zıddı olmayan bir şey yoktur; her şey zıddı ile kaimdir." temel mantık ilkesinden hareketle dikotomik düşünme şekli evrensel bir ilke olarak kabul görmüştür. İnsan zihni herhangi bir şeyi iyi ya da kötü, güzel ya da çirkin, suçlu ya da masum, faydalı ya da zararlı, ölü ya da diri, siyah ya da beyaz gibi zıtlıklarla düşünerek değerlendirme ve ona tek bir değer atfetme çabası içerisindedir. Bu iki değerli düşünme şekli sosyal, siyasal, bilimsel kategorizasyonlarda da aynı şekilde korunmuş, idealizm ve maddecilik, zihin ve beden, akıl ve duyular, kapitalizm ve komünizm gibi karşıtlıklar ortaya çıkmıştır (Georges, 2003: 157). Oysa bir insanın kilolu olması, bir insanın yaşlı olarak adlandırılacağı ilk yaş, uzun boyluluk, iş yerine uzaklık gibi sınırları kesin olmayan değerler günlük hayatın içerisinde büyük yer kaplamakta ve ortak paydalarda buluşmak için önemli rol üstlenmektedirler. Ne derecede otonom bir makineye ahlaki bir sorumluluk yükleneceği de bu şekilde sınırları bulanık olan bir karardır. Ancak şu an için ahlak ve etik içerisinde tartışılan birçok husus, iki değerli düşünmenin etkisinden kurtulmuş ve bulanık sınırları kabul etmiş değildir. İkili düşünme tarzı karmaşık unsurları basitleştirdiği

gibi, sahip olduğumuz seçenekleri de yapay olarak sınırlandırmakta ve tercihimizi tek bir seçenekten yana kullanmaya zorlamaktadır (Georges, 2003: 160). Cinsiyetler erkeklik ve kadınlık olarak ayrılmakta, homoseksüellik bir tercih olarak görülmemektedir. Kürtaj ya hak görülmekte ya da yasaklanmakta, duruma göre tercih söz konusu edilmemektedir. Ötenazi aynı şekilde bir grup tarafından hak olarak değerlendirilirken bir grup içinse intihardan farklı bir anlama gelmemektedir. Oysaki ahlak, bu sınırların arasında kalan muğlak alanlarda ortaya çıkmaktadır.

Orta Çağ boyunca vazgeçilmez bir şekilde bilim, felsefe ve ahlakta kullanılan iki değerli klasik mantık tamamıyla bir şekilden ibaret kalmış, eşyanın derinliğine inmemiş, hep bilinen hakikatlerle uğraşmış ve bilimin dışında özlerden bahseden bir düşünce sistemi olmuştur. Kesinlik, teknik elverişlilik ve yeni problemlerin çözümü gibi konularda yetersiz kalmış, tekâmül esnasında çok fazla düzenleme yapılamamış ve bunların neticesinde klasik mantığın esaslarının yetersizliği metodoloji konusunu ve yeni mantık arayışlarını gündeme getirmiştir (Reichenbach, 1939: 45). Bu yüzden klasik mantık anlayışına tarih boyunca eleştiriler getirilmiş, özellikle Doğu bilgeliğindeki çok değerli mantıkların da dikkate alınması gerektiği vurgulanmıştır. Batı'nın "tenakuz mantığı" şeklinde adlandırıp eleştirel yaklaştığı bu mantıklar, özellikle ahlaki yönü ağır basan bir öğreti olan Budizm'de kendini göstermiştir. Buna göre, A ya da A değil, sadece spekülatif bir dünya olan matematikçilerin ya da ideologların dünyası için geçerlidir. Oysa gerçekte dünyada geçerli olan, "A hem A'dır, hem A değildir." mantığıdır. Dolayısıyla iki değerli düşünme sistemi makine ahlakiliğinin gerçekleştirilmesi için yeterli görünmemektedir. Bu anlamda *Bulanık mantık*[55] kurallarının kullanılması daha sürdürülebilir ahlaki yapılar oluşturmak için faydalı olacaktır. Bulanık düşünme, makinelerin tercihler, önyargılar ve hisler gibi etkenlere dayanan bulanık davranışlar hakkında ahlaki yargıda bulunmasını kolaylaştıran bir yöntem olarak, iki değerli mantığın sınırlamalarından sıyrılmış olacaktır. Makine etiği çalışanlarının pratiklerinde yapması gereken sistemleri iki değerli mantığa göre değil, bulanık mantık ilkelerine göre inşa etmek olmalıdır; bu da yine mühendislik ve etik alanı arasında kurulacak iş birliği ile mümkün olacaktır. Etik çalışmaların teorik bilgisinin sistemlerde uygulanması adına mühendislik ve bilişim alanlarının bu teorik bilgileri dikkate alma ve işlemeleri, etik makine üretimi konusunda hayati öneme sahiptir.

[55] Bulanık mantık hakkında detaylı bilgi, bu çalışmanın Giriş bölümünde bulunmaktadır.

5.4. Makine Hakları ve Makine Etiğine Hukuki Yaklaşım

Yapay zekadan en çok yarar beklenilen alanların başında hukuk gelmektedir. Buna göre, yapay zekanın hukuk ve yönetim sistemlerinde yaratacağı iyileştirmeler sosyal ve ahlaki düzenin gelişimine de büyük katkılar sağlayacaktır. Kanunların hem oluşturulması hem de uygulanması yönünde güncellenmesi gereken çok fazla husus bulunmaktadır. Bu hususların en önemlilerinden bir tanesi hukuk sisteminin hızlanması gerekliliğidir. Yapay zekanın hukukta aktif kullanımı, davaların sonlanma sürelerini kısaltacak ve uzun ertelemelerden kaynaklı maliyetlerin önüne geçecektir. Hukuk alanında insan hâkimlerin yerini robot hâkimlerin alması senaryoları sıklıkla tartışılır hâle gelmiştir. Robot hâkimlerin, insanların yerini almasının en önemli sonucunun dil, din, millet, ırk, renk, cinsel tercih gibi faktörlerden kaynaklanan önyargılardan arındırılmış kararlar olacağı vurgulanırken; robot hâkimlerin herkesin yasalar önünde tamamen eşit olmasını sağlayabileceği iddia edilmektedir (Tegmark, 2019: 142). İnsan hâkimlerin sahip olduğu önyargılara sahip olmayan robot hâkimler, daha etkili ve daha adil olarak çalışabilirler. Böylelikle avukatı iyi olanın kazanma ihtimalinin fazla olduğu bir sistem, yerini haklı olanın kazandığı bir sisteme bırakabilir. Bunların yanı sıra, insan hâkimlerin karmaşık patent tartışmaları, son teknoloji adli tıp tekniklerine dayalı cinayetler gibi konularda yeterince teknik bilgiye sahip olmaları mümkün değilken robot hâkimler sınırsız hafıza ve öğrenme kapasite sahip olabilirler (Tegmark, 2019: 143). Bu da robotik teknolojinin davaların tüm teknik ayrıntılarına sahip olmalarını sağlayarak insan hâkimler karşısında avantajlı konuma gelmesini sağlayabilir. Ancak tüm bu avantajlarına rağmen robot hâkimlerin program hatası yapmayacağına ve hacklenmeyeceklerine dair güvenin oluşması çok zor görünmektedir. Ayrıca insanların neden ceza aldıklarına dair ‘yapay zeka öyle uygun gördü’ açıklamasını kabul etmeleri de sezgisel olarak uzak görünmektedir. Yapay zekanın hukuki süreçlerde kullanılmasının bir diğer dezavantajı da biraz önce savunulanın aksine veri tabanından kullandığı örneklerden önyargılar çıkarabilecek olmasıdır. Örneğin, 2016 yılında Amerika’da lokal olarak kullanılan yapay zekanın, Afrikalı-Amerikalıların suç işlemeye daha yatkın olduğu konusunda bir önyargı edindiği ve bu ön yargıyla beraber adil olmayan karar verdiği görüldü (Tegmark, 2019, 144). Bu avantaj ve dezavantajlar iki taraflı olarak dikkate alındığında hukuk sistemini tamamen yapay zeka sistemlerine bırakmak yerine, yapay zeka bazlı karar destek sistemleri ya da bir aşama ilerisi olarak insan hâkimlere temyize götürecek robot

hâkim kararları adalet ve güvenilirlik açısından daha makul görünmektedir. İdam cezasında, savaş kararlarında ya da çocuk mahkemelerinde dahi son sözün robotik sistemlerde olması, bugünün bakış açısıyla insana biçilen değeri zedeler ve duygusal olarak kabul edilemez görünmektedir.

Ahlaki makineler hakkında konuşulması onların yasal hak ve statülerini tartışmayı da beraberinde getirmektedir. Makinelerin yasal hakları hakkında konuşmanın, bir insanın ya da hayvanın yasal haklarından bahsetmekten farksız olacağını ileri sürenler olduğu gibi, makinelerin yasal bireyler olarak kabul edilmesinin ve hatta bu konu üzerinde düşünülmesinin yalnızca zaman kaybı olacağını dile getirenler de mevcuttur. Makinelerin yasal hak savunucuları makinelerin, işlevsel ve berimsel yetenekler bakımından insanlardan çok daha ileri seviyede olduklarını ileri sürerek, yasal hak sahibi olmak için rasyonel olmayan yetenekler arayışında olmanın yersiz olduğunu vurgulamaktadır. Bu gruba göre, biyolojik olmayan varlıkların, zihinsel durumlara ve bilinçli davranışlara sahip olmadıkları takdirde yasal statü sahibi olmayacaklarını iddia etmek için hiçbir yasal ve mantıksal neden bulunmamaktadır (Wendell, Allen ve Franklin, 2011: 180). Makinelere hukuki statü kazandırılmasını gereksiz görenler ise bir makinenin işlevsel yönden insan zihnine ne kadar benzese de hiçbir zaman bilinçli olamayacağını var sayarak; bilinçsiz varlıklara yasal sorumluluk atfetmenin mantıksal çelişkiler doğuracağını iddia etmektedirler.

Hukukun tüm açılardan makinelere insana davrandığı gibi davranması beklenemez olsa da "tüzel kişilik" hukuku bazı açılardan makine hukuku için model oluşturabilir kabul edilmektedir (Georges, 2003: 160). Tüzel kişilikler, alım satım yapabilir, vergi verir, yanlışlarından sorumludur; oluşturulur ya da yıkılırlar. Ancak evlenemez, oy kullanamaz ya da askerlik yapamazlar. Tüzel kişilik ya da şirket gibi yapı örneklerinin hukuki statülerinden yola çıkıldığında makinelerin insan gibi tanımlanmadan da kendi fonksiyonlarına göre düzenlenmiş yasal hak ve sorumluluklara sahip olabileceği fikrine ulaşılabilmektedir. Bu fikir, makinelerin paraya sahip olmalarını ve mülk edinme haklarını gündeme getirmektedir. Böylece, akıllı makineler borsada para kazanabilir ve bu parayı istediği şekilde kullanabilir görünmektedir. Bu da bütün bir ekonominin makinelere ait olması ve makinelerce kontrol edilmesi sonucunu doğurmaktadır (Tegmark, 2019: 147). Kimilerine göre tüm ekonomi zaten içlerindeki kişilerden çok daha güçlü olan ve kendi kimlikleri olan şirketlerin elinde olduğu için, bilgisayar temelli ekonomi bir risk barındırmazken; kimilerine göre ise bilgisayarlar insanlara çalışmaları için para ödemeye başladıkları anda insanları

köleleştirmelerinin önünde hiçbir teknik engel kalmayacağından makinelerin para kazanma ve mülk edinme hakları olmamalıdır.

Bir makinenin yasal olarak sorumlu olması için süper zekaya sahip olması ya da insan olarak tanımlanması gerekmez; belirli bir konu hakkında uzman olarak hesaplama yapabilmesi yasal sorumluluk alması için yeterlidir. Ancak var olan hukuk sistemlerinde kullanılan ceza yöntemleri, makineler için uygulanabilir değildir. Cezanın insan için karşılığı tazminat ödeme, hapis cezası, meslekten men, idam vs. uygulamalara karşılık gelmektedir. Makine hukuku için ise cezai yaptırım uygulanması yerine, daha iyi performans gösteren ve bağlı olduğu etik ilkelere uygun davranan makinenin varlığını devam ettirme ihtimalinin artması, güncelleme alabilmesi ve rakiplerinin önüne geçmesi gibi ödül yollarıyla makinelerin ahlakiliğe teşvik edilmesi tartışılmaktadır.

Şimdiye kadar makine etiği ve hukuku bağlamında tartışılan problemlerin hepsinde makineler etik "üreticiler" -kendi ahlaki zorunluluk ve sorumlulukları olan bireyler- olarak ele alınırken-; makine etiği çerçevesinde diğer bir bakış açısında, makineler etik "tüketiciler" -onlara karşı bizim yasal zorunluluk ve sorumluluklarımızın olduğu bireyler- olarak değerlendirilmektedir (Torrance, 2008: 499). Bilince sahip olan bir makinenin, bilinçli diğer varlıklar gibi haklarının olması da üzerinde durulması gereken bir konudur. Tüm gelecek projeksiyonlarında bizim tüm işlerimizi yapabilecek olmaları, sosyal hayatın içinde bize yardımcı olarak bulunmaları, her evde yemekten temizliğe tüm görevleri yerine getirecek şekilde çalışmaları gibi robotların köleleştirildiği senaryolar yer almaktadır. Ancak robotların bu işleri yapmak isteyip istemeyecekleri dikkate alınmamaktadır. Bilinçli olabileceklerini varsaydığımız anda, sadece insan olmamaları gerçeğine dayanarak robotları bu şekilde köleleştiremeyiz, robotların bizim tarafından üretilmiş olmaları bizi onların varlıklarının ve düşüncelerinin sahibi yapmaz (Levy, 2009: 212).

Üzerinde durulması gereken bir başka unsurun da makinelere davranış etiği olduğu göze çarpmaktadır. Makinelere nasıl davranmamız gerektiği, makinelerin bize davranışı kadar önemli bir konu olarak müstakil bir tartışma konusu olmayı hak etmektedir. Bilinçli makinelerin, sosyal çevre içerisinde etkileşimde bulunabilen ve deneyimlerle öğrenen makineler olması gerektiğini varsaymıştık. Bu varsayım ışığında, bilinçli makinenin bizim ona davranışlarımızdan hareketle kendi davranışlarını güncelleyeceğini ve hangi durumlarda hangi davranışlarda bulunulması gerektiğini bizim davranışlarımızdan oluşturduğu veri tabanından seçeceğini söyleyebiliriz. O hâlde, biz makineye kaba davranırsak, kaba bir

makine yaratmış oluruz. Biz çevremize ve makineye saygılı olursak, makine de bizden saygıyı öğrenmiş olur. Bizim makineye ve makinenin şahit olduğu durumlarda diğer insanlara ve çevreye karşı olan davranışlarımız, makinelerin davranışlarını ve bu davranışlardan çıkarsayacakları etik ilkeleri doğrudan şekillendirecektir. Bu sebeple insanların makineye olan davranışları da belli ölçütler göz önüne alınarak sınırlandırılmalı ve belli etik kurallara tabi tutulmalıdır. Bu ölçüt ve kuralların belirlenmesi makine etiği çalışanları tarafından gündeme alınması gereken bir yaklaşım olarak göze çarpmaktadır.

Makine hakları hususunda tartışılan konular hayvanların bilimsel deneylerde kullanımı üzerine tartışılagelen meseleler ile paraleldir. Bir görüşe göre, akıllı makineler de kan dolaşımına sahip tüm diğer canlılar gibi kendilerine karşı etik saygı ve ahlaki tutumu hak etmektedirler. Hayvan ve insan hakları bağlamındaki etik tutum, sadece berimsel sistemlerden oluşan akıllı makinelere karşı dahi aynen takınılmalıdır iddiasında olan bu radikal grubun, daha yumuşak bir versiyonu ise ancak bilinçli makinelere karşı etik ve hukuki sorumlulukların doğacağını iddia etmektedir. İnsanların birbirlerine ve hayvanlara karşı olan etik sorumluluklarının ana kaynağı, o insanın ya da hayvanın duygulanımları olduğu ve o duygulanımların bilincinde olduğunun kabul edilmesidir. Fiziksel olarak acı hissetmediği emin olunan bir bireye ya da zombiye fiziksel işkencede bulunmak etik bir problem doğurmaz. Etik kaygıların çıkış noktası bilinçli farkındalıktır. Hâl böyleyken "Bilinçli makinelere nasıl davranmalıyız?" sorusu etik, sosyal ve hukuki boyutları olan temel bir tartışma konusu hâline gelecektir. Eğer ki makineler deneylerde kullanılacaksa fenomenal bilince sahip olmaları etik bir problemin başlangıcı olur. Zira niteliksel bilinçli durumlar ile ahlaki durumlar arasında doğrudan bir ilişki olduğunu dile getiren bu görüşe göre, deneyim ve hislere sahip bir makinenin kobay olarak kullanılması etik çerçevede savunulabilir bir eylem değildir.

Bu yaklaşım bizi makine bilincinin en zor tartışmasının içerisine tekrar sokmaktadır çünkü gelecekteki akıllı makinelerin niteliksel bilinçli durumlara ve duygulanımlara sahip olup olmayacağı muamma bir sorunsaldır. Bu hususta, yapay zeka temelli elektronik aksamlardan oluşan berimsel yapıların, akıllı, otonom, dilsel olarak adapte olabilen ve akıcı sosyal etkileşim yeteneğine sahip sistemler olsalar bile tüm davranışlarının bilinçten yoksun olacağını söyleyenler ile biyolojik olmayan sistemlerde de bilinç bulunabileceğini söyleyenler iki karşıt grubu oluşturmaktadır (Torrance, 2008: 499).

İlk gruba göre, yapay zeka temelli robotlar zeka, hedef, yaratıcılık vs. diğer zihinsel özelliklere sahip olsalar bile asla gerçek anlamda fenomenal ve niteliksel olarak bilinçli olmayacaklardır. Zira fenomenal bilincin önkoşulu biyolojik bir yapının varlığıdır. Gelecekteki robotların sahip olabileceği en büyük yeti, bilinçliymiş gibi davranmak olacaktır; davranışsal olarak bilinçli varlıklardan ayırt edilemeyecek olsalar bile gerçek bir bilince sahip olmayacaklardır. İkinci gruba göre ise şu an için gerçek anlamda bilinçli robotlar üretilmediyse bile bu bir zaman sonra mümkün olacaktır çünkü bilinç de tüm diğer zihinsel özellikler gibi berimsel fonksiyonların bir ürünüdür.

Niteliksel olarak bilinçli makinelerin etik varlıklar olmaları gerektiği, yasal ve ahlaki statülerinin kabulü, etik saygıyı hak ettikleri herkesçe kabul edilen rasyonel bir çıkarımdır. Ancak imkânı dahi henüz tartışma aşamasında olan fenomenal bilinçli makineler yerine, günümüz teknolojik ortamında sahip olunan ve giderek gelişecek görünümde olan bilinçsiz ancak akıllı makinelerin etik statüsü, üzerinde düşünülmeyi kurgusal robotlardan şu an için daha fazla hak etmektedir. Çünkü yapay zeka, etik ve ahlaki boyutları olan birçok eylemde kendisine yer bulmuş görünmektedir. İşe alımlarda kullanılarak insanların ve şirketlerin geleceklerini şekillendiren yapay zeka algoritmaları, dilsel iletişimde insana yardımcı olan çeviri programları, hastalık tespiti yaparak insanları çeşitli tedavilere yönlendiren uygulamalar, yargıç olmadan karar verebilen yapay zekalar, insansız savaş araçları ve geçmiş örneklerden oluşan veri tabanlarından yola çıkarak ilgili oldukları her durum için karar veren algoritmaların eylemlerinin ahlaki ve sosyal boyutları göz ardı edilmemelidir. Bunun yanı sıra, cesur, vahşi, güçlü, tehlikeli, kibar, kaba vs. davranabilecek olan robot davranışları da yine ahlaki alanın içerisinde değerlendirilmelidir.

Bilinçli olmayan insansı robotların ahlakiliği ve ahlaki sorumluluğu basit bir örnekle anlatacak olursak:

> İlkel düzeyde akıllı insansı bir robotun ihmalkâr bir davranışını düşünelim. Bir fabrikada güvenlik görevlisi olarak çalışan bu robotun potansiyel suçlulara ve hırsızlara karşı yeterince uyanık davranamadığını ve fabrikanın soyulmasına engel olamadığını varsayalım. Bu durumda, bu robotu dikkatsizliğinden dolayı sorumlu tutabilir ya da robotun üreticilerini o robotun tasarım ve test aşamalarında yaptıkları hatalardan kaynaklı olarak soygundan sorumlu tutabiliriz. Son olarak da robotun sahiplerini ve işverenleri, üreticileri robotu nasıl tasarlamaları gerektikleri konusunda yeterince bilgilendirmedikleri için suçlayabiliriz (Torrance, 2008: 502).

Robotun programcısının ya da üreticisinin etik ve hukuki sorumluluğunun değerlendirilmesi otonomi kavramıyla doğrudan ilişkilidir. İnsansı bir robotun otonomluk düzeyi,

programcının/üreticinin kontrolünden kopuş düzeyini de ifade eder. Zira kendi başına karar alabilen bir yapının üreticisiyle iletişim ve bağlantısı kopmuş olur. Bu durumda, irade sahibi bir robotun eylemlerinden onu programlamış olan üretici sorumlu tutulamaz görünmektedir. Ancak diğer bir bakış açısından da programcı robotun hata payını sıfıra indirgemeden işverene teslim ettiği için suçlu bulunabilir.

İkincil olarak, otonom insansı robotun üretilmesini talep eden ve güvenlik görevlisi olarak çalışmasına müsaade eden işverenin etik ve hukuki sorumlulukları da aynı şekilde tartışılmalıdır. İşyerinin çalışanlarıyla beraber tüm sorumluluğunun işverende olduğu kabul edildiğinde insansı robotun hatasından kaynaklı olarak fabrikada doğan zarardan ötürü fabrika sahibinin tek başına gerçek sorumlu olduğu ve kimseyi suçlayamayacağı söylenebilir. Fakat başka bir bakış açısıyla nasıl programcıyı otonom bir robotun eylemlerinden sorumlu tutmazken aralarında bir ilişki olmadığını ve otonom olmanın davranışların tahmin edilemezliği anlamına geldiğini ileri sürüyorsak, aynı şekilde işverenin de otonom insansı robotun eylemlerini öngöremeyeceğini ve bu sebeple sorumlu tutulamayacağını iddia edebiliriz (Sparrow, 2007: 70).

Bu durumda, geriye kalan tek seçenek, insansı akıllı robotun kendi eylemlerinin ahlaki, etik ve hukuki sorumluluğunu taşıyıp taşımadığını tartışmaktır. İnsanlar, irade sahibi varlıklar olarak, her türlü eylem ve davranışlarının ahlaki, hukuki ve cezai sorumluluğunu taşımaktadır. O hâlde, otonom robotların da irade sahibi olarak yapay ahlaka sahip birey statüsü ile eylemlerinden sorumlu tutulması geçerli bir çıkarım olacaktır. İrade bağlamında insan ve robot arasında kurulan basit bir analoji otonominin sorumluluk gerektirdiğini göstermektedir. Fakat insani sorumluluk ile makineye yüklenen sorumluluk birbirlerinden karakteristik olarak farklı görünmektedir. İnsani sorumluluk birincil sorumluluk olarak ele alınırken makineye yüklenen sorumluluk ikincil türden bir sorumluluktur; başka bir deyişle insani sorumluluk doğal ve zorunlu bir nitelik iken robot sorumluluğu ilintisel niteliktedir (Torrance, 2008: 502).

Otonom makinelerin sebep olduğu kaza ve hatalarda sorumluluğun kimde olacağı meselesi çok boyutlu bir tartışmayı gerektirmektedir. Hem makineyi üreten mühendis, hem kullanımına izin veren yetkililer, hem yapay zekanın kodunu oluşturan yazılımcılar, hem makinenin kendisi farklı açılardan kazadan sorumlu tutulabilir görünmektedir. Otonom makinelerin günlük hayattaki kullanımlarının artışı hukuki ve ahlaki bağlamlarda "fail" tanımlamalarının yeniden yapılmasını gerektirmektedir. Bir makinenin de fail olabilme

ihtimalı bu konuda yasal düzenlemeler yapılmasını zorunlu kılmaktadır. Bu hususta yapılacak yasal düzenlemeler için devletler üstü bir çalışma gerekmektedir. Her devletin kendi yapay zeka hukukunu oluşturması gerekliliğinin yanı sıra uluslararası yapay zeka hukuku oluşturulması ve otonom makinelerin kullanımının belirli anlaşmalarla standarda kavuşturulması da başat bir gerekliliktir (Tegmark, 2019: 177). Rusya'da gerçekleşmiş olan otonom makinenin kullanıldığı bir kazada, yazılımcının Çinli, mühendisin İngiliz, olay mağdurlarının Türk olma olasılığı her zaman bulunmaktadır. Bu gibi durumlar için evrensel yapay zeka kanunları geçerli olmalı ve ülkeler arasındaki kuralların farklılıklarından doğacak sorunlardan kaçınılmalıdır. Küreselleşen dünyanın yapay zeka ile yeniden şekillenişi her geçen gün çok daha açık hâle gelirken hukuk sistemi bu değişimin arkasında kalmamalıdır. Yapay bilinç uygulamalarının kültürel ve sosyal boyutları olmasına karşılık, hukuki anlamda evrensel normlarla düzenlenmesi gereklidir. Aksi takdirde, her devletin kendi çıkarları için uygulayacakları yapay zeka politikaları küresel çapta büyük krizlerin yaşanmasına sebebiyet verebilir. Bu bağlamda, yapay zeka hukuku dost yapay zeka ekseninde şekillenmeli ve yapay zekanın neden olduğu her türlü etik ve hukuki ikilemin çözümlenmesi için uluslararası tahkim kurulları oluşturulmalıdır.

5.5. Bilinç-Ahlak İlişkisi Bağlamında Makine Etiği

Yapay zeka ve hukuk ilişkisi bağlamında şu soruların cevabı muğlak olduğu gibi, bu türden soruların soruluşu bile insani bakış açısıyla sezgisel olarak çelişir görünmektedir: "Bir robotun sorumlu olması, cezalandırılması ya da ödüllendirilmesi ne anlama gelir? Bir yapay zekayı hapse atmak, elektroşok uygulamak, program işlevlerini kısıtlamak, idam etmek vs. gibi uygulamalar ona acı çektirir mi, ceza olur mu? Zaman kavramı olmayan, bir yaşam süresine sahip olmayan bir varlık için hapis ya da kısıtlılık gibi eylemler ceza sayılabilir mi?". Sezgisel olarak tüm bu olasılıkların anlamsız görünmesinin sebebi otonomi için fenomenal bilinci zorunlu bir koşul olarak kabul etmeyişimizdir. İlaveten, fenomenal bilince sahip olmayan bir varlığın cezalandırılması, işkence edilmesi, öldürülmesi, hapse atılması vs. o varlıkta karşılığı olmayacak çünkü içsel ve öznel bir his uyandıramayacak uygulamalar olarak kabul edilmektedir. Eğer ki bir sistemin qualia sahibi olmasından bahsedemiyorsak ve bilincinin öznel bir yanı olduğunu kabul edemiyorsak, o sistemin eylemleri şekillendirecek ceza ve ödül mekanizmalarının faydasız yaklaşımlar olduğu varsayılmaktadır.

Bilincin ahlaki karar verme açısından rolü ve önemi, üzerinde ehemmiyetle durulması gereken bir husustur. Doğal ve yapay sistemlerin karmaşık ve yeni durumlar karşısında karar verirken bilince düşen görev hâlâ net olarak saptanmış değildir. Etik ve bilinç arasındaki ilişkinin sorgulanması şu dört soru ile özetlenebilir:

- (Doğal ya da yapay) Bir sistemin bilinçli olması ile ahlaki karar verme kapasitesi arasındaki ilişki nedir?
- Ahlaki karar verme için bilinç temel bir gereksinim midir?
- Hangi bilişsel ve berimsel mekanizmalar hem bilinç hem de ahlaki karar verme için gereklidir?
- Tümüyle bilinçli bir birey olarak adlandırılmak için, ahlaki karar verme kapasitesi temel bir katkı olarak mı düşünülmelidir? (Wendell, Allen ve Franklin, 2011: 180).

Zeka sahibi varlıkların hangi noktada bilinçli olduğunu söyleyebileceğimiz sorusu yapay zekanın son yıllardaki yükselişiyle birlikte yeniden gündeme gelmiştir. Zira makinelerin sahip olması gereken haklar ile onların acı veya mutluluk hissi yaşayıp yaşamadıkları ve bilinçli olup olmadıkları sıkı sıkıya ilişkilidir. Makineler üzerinde uygulayacağımız bir etik teorisine karar vermek için onların bilinçli olup olmadığını bilmeye ihtiyacımız vardır. Bilinçli olmayan bir makinenin sahip olduğu olumlu deneyimleri arttırmak üzerine kurulmuş faydacı bir yaklaşım tutarlı olmayacaktır. Öte yandan, bu konu üzerine tartışan insanların bir kısmı bilinçli robotların günlük hayatta kullanımı köle-efendi ilişkisine yol açacağından ve onları köleleştirmenin bir tür suç olacağından yola çıkarak bilincin yapay üretimine karşı çıkmakta, bir kısmı ise insanın zihinsel ölümsüzlüğünün ancak bilinçli bir robota zihninin yüklenmesiyle mümkün olacağı argümanına dayanarak bilinçli makinelerin üretiminin gerekliliğine vurgu yapmaktadırlar (Tegmark, 2019: 362).

Ahlak sahibi varlıkların duygulanım ve empati yeteneğine sahip olmaları gerektiği genel bir varsayım olarak kabul edilmektedir. Bu sebeple etik değerlere sahip bir makine üretilmesi hedefleniyorsa öncelikli amacın o makineye duygu ve empati yeteneği eklemek olması gerektiği öne sürülmektedir. Ancak makinelerde sentetik duygular -özellikle haz ve acı- oluşturulması için kayda değer çalışmalar[56] olsa da sentetik duyguların gelecekte üretilip üretilemeyeceğini değerlendirmek için henüz erken görünmektedir.

[56] Picard, 1997; Franklin ve Patterson, 2006; Haikonen, 2009.

Son yıllarda, ahlaki davranışların birçoğunun bilinçdışı yargı süreçlerinden kaynaklandığını gösteren pek çok çalışma yapılmıştır; bu çalışmalar bilinci, ahlaki yargılama süreçlerinin tamamen dışında tutmasa da eylem üzerine bilinçli düşünmenin varsayılandan daha nadir olduğunu vurgulamaktadır (Wendell, Allen ve Franklin, 2011: 181). "Sosyal Sezgici Model" olarak adlandırılan görüşe göre, ahlaki davranışların çoğunluğunda karar verici faktör bilinçli akıl yürütme yerine duygulanım yoluyla aktive edilen sezgilerdir (Haidt ve Craig, 2007: 368). Ahlaki davranışın, duygu durumundaki küçük değişikliklerle ve bilinçsiz ya da bilinçaltı etkilerle kolayca değişebildiği sosyal psikologların deneylerinde çok kez ortaya konulmuştur.

Diğer taraftan da "olan" ile "olması gerekenin" birbirinden ayrılması mutlak suretle önemlidir. Bize göre, insanların ahlaki seçimlerinde "olan" bilinçdışı etkenler ve sezgisel faktörler, insanları "olması gereken" etik değerlerden uzaklaştırmaktadır. Buna göre, etik bir makine insanlarda ahlaki seçimleri etkileyen "olan" faktörlere göre değil, "olması gereken" bilinçli akıl yürütme süreçlerine göre inşa edilmelidir. Aksi takdirde, kurulacak yapı insanın sezgisel ve dürtüsel seçimlerinden kaynaklı ahlak normlarına uymayan davranışlarını da tekrarlayacaktır. Dolayısıyla olması gereken yapay bir ahlak sisteminde bilinç, ahlaki karar vermelerde de en az diğer tüm rasyonel karar vermelerde olduğu kadar rol oynamalıdır. Zira bilinç, bireyin çevreyle etkileşiminde en önemli unsurdur.

İradeye dayalı olan her türlü eylemde bilincin önemi açıkken çevresiyle etkileşim kuran bir bireyin iradesiz varlığını devam ettirmesi mümkün değildir. Ahlak da sosyal çevre içerisinde ortaya çıkan ve önem kazanan bir yapı olarak, temellerini bilinçli iradi eylemlerden almaktadır. Ayrıca deneysel çalışmalar göstermektedir ki, hem ahlaki karar verme hem de bilinç, algı ve eylemin birleşmesini gerektirmektedir (Wendell, Allen ve Franklin, 2011: 189). Bu bağlamda, sosyal çevre etkileşiminde merkezi öneme sahip olan bilinç ve ahlak birbirlerinden bağımsız değerlendirilmemelidir. Bütünüyle bilinçli bir makinenin inşasının, ahlaki değerleri kapsaması gerektiği gibi; etik bir makinenin inşası için de bilincin kurucu etkisi olmalıdır. Bu durum yapay zeka çalışmalarında felsefe ve etiğin ne kadar önemli olduğunu bir kez daha gündeme getirmekte ve bunun yanı sıra bilinç ve ahlak arasındaki kuvvetli bağı ortaya koymaktadır.

Bir sonraki bölümde açıklanacağı üzere, bilinç sosyal etkileşimle açığa çıkan bir süreçtir. Bilincin varlığı, deneyimin varlığıyla mümkündür. Bu tez, ahlakın da sosyal ve kültürel bir süreç ile varlık kazandığı gerçeğiyle tamamen paraleldir. Bu sebeple ahlakın bilinci

önceledi̇ği̇ ya da bilincin ahlaki yapıyı önceledi̇ği̇ gibi bir tezi savunmak yerine, bilinç-ahlak ilişkisini paralel yapılar olarak ele almak daha uygun görünmektedir. Hem bilinç hem de ahlak var olmak için kendileri dışında bir dünya ile etkileşime girmeyi gerektirmektedir. Hem bilinç hem de ahlak esasen öznenin yanısıra deneyime de zorunlu olarak bağlıdır. Dolayısıyla hem makine etiği çalışmaları hem de makine bilinci çalışmaları teknik ve felsefi yanlarının dışında sosyolojik ögeler olarak da dikkate değerdir. Bu bağlamda, çalışmanın bundan sonraki bölümünde bilincin doğasına yönelik özgün bir bakış açısı geliştirmeye çalışılırken bilinçli öznenin aynı zaman da sosyal bir özne olması gerektiği savunulacaktır. Bununla birlikte bilinçli yapay zekanın imkanı üzerinden yürütülecek tartışmada yapay bilinçte sosyal katman oluşturulabileceği iddia edilecek ve makine etiği çalışmalarının gerekliliği vurgulanacaktır.

6. YENİ BİR BİLİNÇ YAKLAŞIMI TEMELİNDE YAPAY BİLİNCİN İMKÂNI

Bilinç, geçmişte kalanı gelecekte olacakla birleştiren bir tire işareti, geçmişi ve geleceği birbirine bağlayan bir köprüdür.
Henri BERGSON

6.1. Varolan Bilinç Teorilerinin Varsayımları Üzerine

Bilinçli bir yapay sistemin inşası, aynı zamanda yapay bir öznenin inşasıdır. Yapay bir özne inşası, deneyim ve algıya dair tüm kabul ve varsayımların gözden geçirilmesini gerektirir. Klasik algı teorisi, "temsil" kavramı üzerine inşa edilmiştir. Neredeyse tüm bilinç teorilerinde örtük olarak da olsa bulunan bir varsayım, "Dışsal olay ya da çevre ile onun beyindeki temsillerinin nedensel olarak bağlı olsalar da aslında ayrı olduklarıdır." (Manzotti ve Tagliasco, 2008: 110). Bilinci bu denli çözülemez bir sorun hâline getiren de dış dünya ve dış dünyanın iç dünyadaki temsili yani dış dünya ile deneyimlenen dünya arasındaki karşıtlıktır.

Bilindiği kadarıyla, beyinde bizim fenomenal durumlarımızın içeriğine işaret eden hiçbir bölge ya da fenomenal durumlarımızın içeriği gibi bir şey mevcut değildir. Deneyimleyen özne, bir ses duyduğunda, ekşi bir şey yediğinde, parlak bir top gördüğünde, tüylü bir kediye dokunduğunda beyinde bu duyularının özelliklerini içeren bir yer görülmemektedir. Başka bir deyişle, beyinde hiçbir deneyime dair fenomenal durumların görüntülenebildiği bir yer mevcut değildir. Bu demektir ki; fenomenal durumlar, nesnel görüntüleme yöntemleri ile destek bulmayan, öznel verilerin sonucunda varlıkları kabul edilen gerçekliklerdir. Beynin sahip olduğu tek şey ise bu niteliksel ve fenomenal durumlardan tamamen farklı özelliklere sahip nöral modellerdir. Bu nöral modeller, ekşi tat, yüksek ses, yumuşaklık gibi yalnızca öznel deneyimle bilinebilen durumların açıklaması hakkında bir görünüş içermemektedir. Hatta bu dışsal nesnelerle işlevsel olarak bağlı nöral modellerden ortaya çıktığı varsayılan fenomenal deneyimin, ekşi elmanın, tüylü kedinin ya da sakinleştirici bir melodinin fiziksel özellikleriyle nasıl alakalı olduğu da bilinmemektedir. Fenomenal deneyim ve beyin ilişkisi gizemli ve açıklanamaz bir durum olarak değerlendirilmektedir. Bu ilişkiyi açıklamak için korelasyon, nedensellik, beliriverme, özdeşlik gibi birçok farklı yaklaşım sunulsa da hiçbirisi fenomenal deneyimin öznel doğası ile nesnel beyin arasındaki ilişkiyi bütün hatlarıyla ortaya koymamaktadır. Bizim iddiamız şudur ki; beynin nöral haritasının bütünüyle oluşturulması

ilkesel olarak imkânsız değildir, ancak bu nöral haritanın tamamlanması deneyimin öznel doğası ile nesnel beyin arasında var olan boşluğu açıklanabilir kılmayacaktır.

Bilincin nöral bağlantıları yaklaşımına göre, beyin gelen elektrik sinyallerinden bazı içsel kodlamalarla, dış dünyanın deneyimini üretmektedir. Ancak beyin ve deneyim ilişkisinin açıklaması, bu yaklaşımda da mevcut değildir. Bu yaklaşıma dışardan bakıldığında beyindeki içsel sinirsel kodlamaların, sinirsel aktivite "suyunu", fenomenal deneyim "şarabına"[57] değiştirme mucizesini gerçekleştirdiğini savunduklarını söylemek yanlış olmaz (McGinn, 1991: 1). Sinirsel aktivitenin ne aşamada fenomenal duruma yol açtığı bilinemeyen bir süreçtir. Karakteristik olarak nöronlar ile fenomenal durumların ortak hiçbir özelliği bulunmamaktadır. Öyle ki, sinirsel aktiviteyi bir kirpi olarak görürsek, fenomenal durumu pamuk olarak ele almak yanlış olmaz. Nitelikleri bakımından hiçbir ortak paydada buluşmayan nöronlar ile fenomenal zihin durumları arasındaki nedensel ilişkinin açıklaması bu sebeple birçok araştırmacı tarafından ya imkânsız görülmekte ya da yapısı tanımlanamayan fenomenal durumların varlığı reddedilerek bu aşılamaz gedik kapanmaya çalışılmaktadır. Fiziksel ve zihinsel nitelikler arasında bir farklılık olduğunu kabul edenlerden McGinn bu konuda kötümser bir bakış açısına sahipken; Searle'ün biyolojik doğalcılığı bu boşluğun kapatılabilir olduğunu iddia etmektedir. McGinn beyinde meydana gelen olayların deneyimlerimizi nasıl zorunlu olarak meydana getirdiğini gösterecek bir özelliğini aradığımızı söyler; ama ona göre, insan zihninin doğasından kaynaklı olarak, bu özellik daima kavrayışımızın ötesinde kalmak zorundadır (Zeman, 2006: 390). Searle ise bilinci, beyinden beliren bir özellik olarak görerek, zihin-beden probleminin aşılabileceğini iddia eder.

Yapay bilinç araştırmacılarının öncelikli olarak yapması gereken deneyimlenen ile deneyimleyen arasındaki açıklanamaz boşluğu tartışmadan önce, bilinç hakkında bildiklerimize ve bilebileceklerimize odaklanmaktır. Zira ontolojik bir problemin çözümsüzlüğüne odaklanarak, bilincin kavramsal betimlemesini gereksiz görmek bilimsel bakış açısı ile uyumlu değildir. Deneyimi, deneyimleyen ve deneyimlenen arasındaki etkileşim sürecine indirgeyen bir bakış açısı zihinsel temsiller, qualia ve fenomenal durumlar üzerine var olan tartışmalardan kaçınmayı bir süreliğine mümkün kılmaktadır. Dolayısıyla deneyimin fenomenal niteliğinin doğurduğu ontolojik problem başlangıç için bir kenara bırakıldığında bilince epistemolojik düzeyde yeni bir bakış açısı kazandırmak daha olası hâle

[57] Sinirsel aktivitenin doğası ile fenomenal deneyimin doğası birbirlerinden su ve şarap kadar farklıdırlar.

gelmektedir. Burada epistemolojik düzeyden kasttettiğimiz ise bilinci öncelikle erişim bilinci bağlamında ele almaktır. Block (1996) tarafından yapılan erişim bilinci ve fenomenal bilinç ayrımında erişim bilinci, bir durumun bilincine karşılık gelmekte ve çoğunlukla farkındalık anlamında kullanılmaktadır. Bir durumun bilincinde olmak demek, o durumu rasyonel kararlarımızda, planlarımızda veya dilsel ve iletişimsel davranışlarımızda kullanmaya hazır olmamız demektir (O'Regan, 2012: 118). O hâlde, erişim bilincini, deneyimleyenin deneyimlenen ile kurduğu etkileşim süreci ile açıklamaya çalışmak, ontolojik problemlerden arınmış bir bakış açısı olacaktır. Manzotti erişim bilinci ve fenomenal bilinç arasında bir ayrım yapmasa da bilinci şöyle tanımlar:

> Bilinçli olmak, deneyimlenen dünya ve beyin arasındaki fiziksel devamlılığın oluşumuyla özdeştir; buradaki tek varsayım ise bir şey yalnızca etki ya da sonuç ürettiğinde var olur prensibidir (Manzotti ve Tagliasco, 2008: 108).

Zihnin doğasına dair yaklaşımları en genel anlamıyla temelde içselcilik ve dışsalcılık olarak sınıflandırabiliriz. İçselciliğe göre bilinç, beynin ya da bedenin sınırları içerisinde gerçekleşen olay ya da durumlarla özdeş veya ilişkilidir. Dışsalcılığa göre ise bilinç, tamamıyla olmasa bile parçalı olarak bedenin dışında gerçekleşen olay, durum ya da süreçlere dayanmaktadır.

Nörobilimde son yıllarda giderek artan gelişmelerle birlikte güncel birçok bilinç teorisi, içselci bakış açısı üzerine kuruludur. Ancak içselcilik kendi içerisinde birçok açmaza sahiptir. Bilinci tamamen beyinde yer alan bir fenomen olarak değerlendirdiğimizde, dış dünyada yer alan olay ve durumları nasıl temsil ettiğini açıklamak oldukça güçtür.[58] Bilincin beynin içerisinde bir yerde olduğunu savunan içselcilik, hafıza aracılığıyla tüm deneyim ve hatıraların beyinde depolandığı, dışarıda fiziksel bir dünya bulunduğu ve beyinde ise bu fiziksel dünyanın temsilinin bulunduğunu ima etmektedir (Seager ve Bourget, 2007: 262). Ancak nörobilimde kaydedilen muazzam ilerlemelerin hiçbir aşamasında, beyinde bu imayı doğrulayan somut bir veriye rastlanmamıştır. Deneyimin kendisi, var olan hiçbir teknikle, beynin içerisinde görüntülenememiştir. Bilim insanları tüm uğraşlarını, bilinçli deneyim ile sinirsel aktiviteler arasındaki bağlantıları açığa çıkarmaya yönlendirmiş olsa da deneyimin kendisinin bağlantı kurulan sinirsel aktivitenin kendisi olmadığı açıkça ortadadır. Beynin içinde gözlemlenen bilincin kendisi değil, sinirsel aktivitelerdir. Dolayısıyla şu an için

[58] Ayrıntılı tartışma için bkz. 3.5. ve 3.6. başlıklı bölümler.

nörobilim bilincin kendisini değil, nöral bağıntılarını açıklamaya çalışmaktadır. Bu bir anlamda bilincin kendisini örtük olarak bilimin dışarısına itmektir.

Bilincin yerini beyinden ve bedenden farklı bir noktada konumlandıran dışsalcılığa göre ise bilinci açıklamak için bilincin nöral korelasyonlarının keşfi yetersizdir. Bilinç, fiziksel niteliklere sahip olduğu gibi fiziksel olmayan niteliklere de sahiptir. Ve bu fiziksel olmayan nitelikler, bilincin fiziksel beden içerisinde yer aldığı düşüncesini yanlışlamaktadır. Dışsalcı bakış açısı, bilincin mahiyetini açıklamaya çalışırken beyin ve sinirsel aktivitelerin araştırılmasının yeterli olmayacağını; bilincin fiziksel olmayan doğasının gözden kaçırılmaması gerektiğini savunmaktadır.

Çalışmanın ilk bölümünde detaylı olarak ele alınan töz düalizmi ve nitelik düalizminin yanı sıra davranışçılık da dışsalcı bakış açısının önemli bir temsilcisidir. Davranışçılık, düalistlerin dışsalcılığından farklı olarak fenomenal bilinci dikkate almadan zihnin ve bilincin davranışsal olarak dışavurumuna odaklanarak, ancak benlik, zihinsel yaşam, içebakış gibi idealist kavramlardan kurtulmak isterken bilincin kendisini de yalnızca davranışa indirgeyerek, bilincin gerçek doğasını inkâr etme hatasına düşmektedir.

Töz düalizminde ise bilinç, hiçbir fiziksel özelliği bulunmayan immateryal bir varlık olarak ele alınmaktadır. İlk bölümde tartıştığımız üzere töz düalizminin en büyük açmazı tamamen fiziksel olan beden ile fiziksel olmayan bilincin nasıl olup da etkileşim içinde olabildiğidir. Nitelik düalizmi, töz düalizminden farklı olarak zihne ayrı bir varlık sahası atfetmemekte ancak zihinsel nitelikleri dolayısıyla bilinci, fiziksel niteliklerin karşıt kategorisinde ele almaktadır. Bu bağlamda, nitelik düalizminin açıklamaya kavuşturamadığı sorun ise fiziksel bir şeyin nasıl fiziksel olmayan özelliklere sahip olabildiğidir. Dolayısıyla dışsalcılığın farklı formları da içselcilik gibi birçok sorunsalı çözümsüz bırakmaktadır.

Bilince yönelik bakış açılarının en genel iki formu olan içselcilik ve dışsalcılık taşıdıkları problemlerle beraber tüm bilinç teorilerine temel teşkil etmekte ve taşıdıkları problemleri uygulandıkları teoriye aktarmaktan kaçınamamaktadırlar. İlk iki bölümün konusunu oluşturan klasik ve güncel bilinç teorilerinin bütüncül bir bilinç açıklaması sunamamasının en önemli sebebi dayandıkları ya içselci ya da dışsalcı temelin problematiklerini aşamamalarıdır. Dolayısıyla bu çalışmada ortaya koymaya çalışacağımız bilinç yaklaşımı, önce deneyimlenen ve deneyimlenenin ontolojik statüsünden bağımsız olarak, epistemolojik

bir yaklaşımla içselci ve dışsalcı bakış açılarının dışında kalmayı hedeflemekte, ardından ontolojik problemi tartışmayı planlamaktadır.

İçselcilik ve dışsalcılığın barındırdığı sorunların çözümüne alternatif bir bakış açısı 1950'lerden sonra beden ve çevre arasındaki etkileşime odaklanarak ortaya konulmuştur. Bu bakış açısı, algı ya da deneyimi, dış dünyanın zihindeki temsilinden ziyade çevreyle etkileşim olarak ele almaktad; bilincin varlığını inkâr etmemekte ve bebeklikten itibaren çevreyi dokunarak ve araştırarak öğrenmenin ilk deneyimlerimizle beraber bilinci meydana getirdiğini savunmaktadır (Manzotti, 2018: 40). Bu yaklaşıma göre deneyim, fiziksel dünyanın algımıza yönelttikleri ile ilişki kurmak anlamına gelmektedir. Ekolojik algı,[59] bedenlenmiş/somutlaşmış bilinç,[60] enaktivizm/etkileşimcilik,[61] yayılmış/genişletilmiş zihin[62] gibi birçok teori bu bakış açısını temel alarak oluşturulmuştur. Bu çalışmada da bu bakış açısından yola çıkarak, bilince dair farklı bir kavramsal çerçeve oluşturmaya teşebbüs edilecek ve bu çerçeve ile yapay bilincin imkân ve koşullarına dair sonuçlar çıkarılacaktır.

6.2. Alternatif Bir Bilinç Yaklaşımı: Fenomenal Gerçekçilik

Bilim tarihine geniş bir perspektiften baktığımızda, olan bitenin karmaşık görünen yapı ve olaylar hakkında basit açıklamalar bulmaya çalışmak olduğunu kolaylıkla fark edebiliriz. Ancak deneyim ve buna bağlı olarak deneyimleyen hakkında düşünürken bunun tam tersi bir yol izlememiz gereklidir. Zira deneyimleyenin, "ben"in ya da öznenin varlığı bize çok basit, açık ve sıradanmış gibi görünür. Nitekim "ben"in varlığına dair inancımız sürekli olarak duyumsadığımız bir "ben" hissinden ileri gelmektedir ve tüm deneyimlerin

[59] *Ecological Perception.*

[60] *Embodied Consciousness*: Bu teoriye göre, bilincin yeri yalnızca beyin değil, bedenin tamamıdır. Bilinç, biyolojik, hareket eden, eyleyen bir bedende bedenlenmiş ya da somutlaşmıştır.

[61] Etkileşimciliğin kurucu makalesi için bkz.: O'Regan, Kevin ve Alva Noë (2001). "A Sensorimotor Account of Vision and Visual Consciousness". *Behavioral And Brain Sciences*. (24): 939-1031. Üçüncü bölümde ayrıntılı olarak ele alındığı üzere Duyu-Motor Kuramı olarak da adlandırılan bu teoride, bilinç beynimize değil, çevreyle olan ilişkilerimize dayanmaktadır. Bu kuram bilinçli algının, dış dünyanın temsillerinin beyindeki nöronlarımız tarafından yeniden oluşturulmasını içerdiği fikrini tamamen reddetmektedir. Temsil eden ve temsil edilen ile özne ve nesne karşıtlığını ortadan kaldırmaktadır.

[62] *Extended Mind*: Bu teoriye göre, bilinç ne sadece beyne ne de sadece bedene bağlıdır. Zihin ve bilişsel süreçlerin tamamı, bedenden çok daha geniş bir şekilde bireyin dünyasına yayılmış durumdadırlar. Bireyin dünyasından kastedilen ise hem bireyin içerisinde yaşadığı fiziksel dünya hem de içerisinde bulunduğu sosyo-kültürel dünyadır. Yayılmış Zihin teorisinin başat makalesi için bkz. Clark, Andy ve David J. Chalmers (1998). "The Extended Mind." *Analysis*. (58): 7-19.

arkasında o "ben"in olduğuna içten içe sürekli olarak inanmaktan kendimizi alıkoyamayız. Ancak "ben"in doğası üzerine bir sorgulama yapmaya başladığımızda, onun göründüğü kadar açık ve kolay çözümlenebilir bir doğaya sahip olmadığını fark ederiz. Buna bağlı olarak, deneyim ve bilincin yapısı da göründükleri kadar basit değildir. Tüm deneyimlerimizin, duygularımızın, zekamızın, kararlarımızın kaynağı olarak bir benliğin kabulü, aslında birçok problemi de beraberinde getirmektedir. Kendimizi tek bir benlik olarak kabul etmek, bizi meydana getiren küçük parçaların doğası hakkında bize bir açıklama sunmaz ve "Bu problemi nasıl çözebilirim, bir sonraki adımıma nasıl karar veririm, düşündüğüm konulara karar veren şey nedir?" gibi soruları cevaplamak için hiçbir ipucu vermez (Minsky, 2006: 16). Bilinç araştırmaları, bu bağlamda, fizik, biyoloji ya da kimyanın aksine basit görünen bir fenomen hakkında karmaşık açıklamalar bulmayı hedeflemektedir. Bilinci tek bir kavram etrafında inşa ederek tanımlamaya çalışmak çok yönlü bu olguyu kavramakta yetersiz kalmaktadır. Uyanıklık, farkındalık, kendinin farkında olma, duygusal girdiye erişim, duygululuk gibi kavramlardan hareketle yapılan tanımlar bilincin belli bir yönüne işaret ederken bütüncül bir bakış açısı sağlayamamaktadır. Bu yüzden bilincin kapsamlı ve farklı yönlerini bütün olarak ifade edebilecek bir bilinç teorisine ihtiyaç vardır.

Bilinç günlük dilde farklı anlamlara gelen bir kavram olduğundan, tek bir tanımını oluşturmak oldukça güçtür. Bu tezin giriş bölümünde oluşturmaya çalışılan kavramsal çerçeve içerisinde, bilinç "deneyimleyenin, deneyimde bulunduğunun farkında olma hâli" anlamında gelmektedir. Yine de bilinç için daha geniş perspektiften bir tanım verilmesi gerekirse, bu tanım Searle'ün şu ifadelerine dayandırabilir:

> Bilinç alışılageldik şekliyle, rüyasız bir uykudan uyanmamızla başlayan ve biz tekrar uykuya dalana, komaya girene, ölene ya da farklı bir biçimde 'bilinçdışı' olana kadar devam eden duyarlık ve farkındalık durumlarına işaret eder. Tam uyanıklık durumlarından çok farklı olmalarına rağmen, rüyalar da bilincin bir biçimidir. Tanımladığımız hâliyle bilinç, kapalı ve açık durumlar arasında geçiş yapar. Buna göre bir sistem ya bilinçlidir ya da değildir, fakat bilinç alanı içerisinde sersemlikten tam uyanıklığa uzanan yoğunluk dereceleri de mevcuttur. Bu şekilde tanımlanan bilinç, bir içsel niteliksel birinci şahıs fenomenidir (Searle, 2018: 16-17).

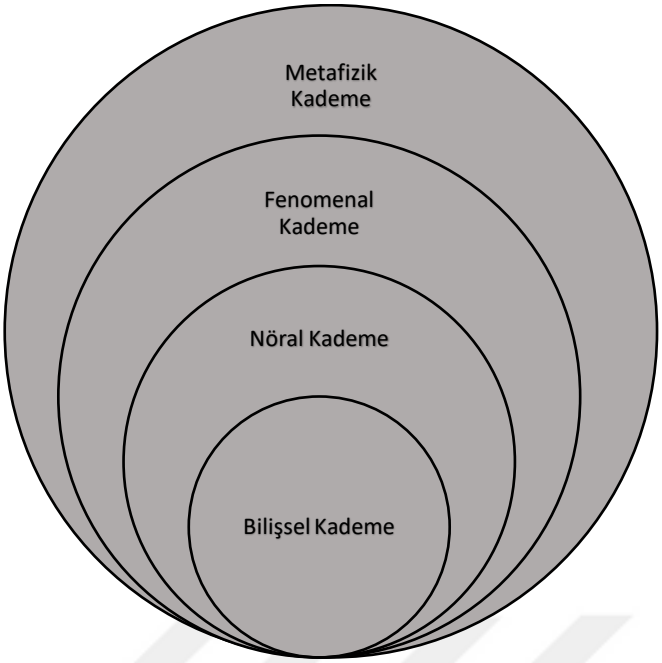
Daha sade bir ifadeyle de bu tezde bilincin doğasını anlamaya yönelik olarak geliştirmek istenilen yaklaşımda, bilincin atıfta bulunduğu anlam açıkça "deneyim" ile bağlantılı olacaktır. Deneyim; dikkat, farkındalık, davranışta bulunma, duygululuk ve uyanıklık gibi kavramlarla nitelenebilir olmakla birlikte bu kavramların hepsinden farklıdır.

Bilinç doğası gereği bölünemez bir yapıya sahiptir. Descartes zihin-beden problemini felsefe tarihinin en zor sorunlarından birisi olarak literatüre sunarken bilincin birliğini şöyle ifade etmiştir:

> Zihnimi düşündüğümde, bir başka deyişle, ben sadece düşünen bir şey olduğum için kendimi düşündüğümde, kendimde hiçbir parçayı ayıramıyorum; aksine kendimi açıkça tek ve bütün olarak kavrayabiliyorum (Descartes, 1970: 196).

Ancak bu bir ve bütün görünen yapının çözümlenmesi, bütünü oluşturan temel parçaların doğasını bilmeden mümkün görünmemektedir.

Ortaya koymaya çalışılan bilinç yaklaşımı, David Chalmers'ın bilincin kolay problemleri ve zor problemleri ayrımından yola çıkarak bilincin problemlerini daha geniş bir bakış açısıyla gruplandırmaktadır. Chalmers, bilincin davranışsal ve işlevsel yönlerini kolay çözülecek problemler olarak ele alırken; bilincin öznel karakterini bilincin zor problemi olarak ele alır. Bu tezde ise bilinç çok yönlü bir fenomen olarak düşünülmüş ve epistemolojik-ontolojik-metafizik boyutlarıyla ayrık olarak değerlendirilmesi önerilmiştir. Bu bağlamda bilincin davranışla bağlantılı olan işlevsel özellikleri, bilince dair problematiklerin en alt kademesi olan bilişsel kademeyi oluşturmaktadır. Bilişsel kademenin bir üstünde ise bilincin nöral korelasyonları yer almaktadır. Bilincin nöral kademesi üzerine yapılan araştırmalar temel olarak bilinçli durumları oluşturan nöral aktiviteler ile bilinçli ve bilinçdışı süreçlerin birbirinden nasıl ayrılacağı yönündeki problemi konu edinmektedir. Nöral basamaktan bir sonraki aşamayı fenomenal kademe oluşturmaktadır. Fenomenal kademe üzerine çalışmaların ana konusu ise öznel ve niteliksel deneyimdir. Dördüncü ve son basamak ise bu yaklaşımda metafizik kademe olarak adlandırılmaktadır. Metafizik kademeye dâhil olan çalışmaların içeriğini ise "Bilinç neden var?" sorusu oluşturmaktadır. Bu gruplandırmada, bilince ait her bir kademe bilinçle ilgili çözülmesi gereken bir problemi temsil etmektedir. Bu kademeler arasında aşağıdan yukarı doğru çıkıldıkça problemlerin zorluğu artmaktadır. Bu çalışmada, bilincin en kolay problemi, bilincin kolay problemi, bilincin zor problemi ve bilincin en zor problemi şeklinde dört ayrı problematik üzerinden bir sınıflandırma yapılmaya çalışılacaktır:



Şekil 1: İnsan Bilincinin Kademeleri

6.2.1. Bilincin En Kolay Problemi

Bilincin en kolay problemi, Chalmers'ın kolay problem olarak ifade ettiği sorunsaldır ve bilincin işlevsel ve davranışsal bağlantılarını kapsamaktadır. Bilinçli durumların davranış ile olan ilişkisi bilinç konusunda çözülebilecek en kolay problemlerdir. Bilişsel bilim de zaten bu problemi aydınlatma noktasında büyük aşama kaydetmiş, bilincin bilişsel ve davranışsal bağlantılarını büyük ölçüde ortaya koymuştur. Beynin bilgiyi nasıl işlediği sorusu bu bağlamda cevaplanmaya çalışılan en kolay problem görünümündedir. Beynin duyusal verileri aldıktan sonra, onlara çıktı üretme süreci oldukça karmaşık bir süreç olsa da bilişsel bilimin sınırları içerisinde kalarak kolay bir şekilde çözümlenebilirdir. Dahası, bilincin en kolay problemi olarak görülen bilgi işleme süreci, bilinçten ziyade zekanın içeriği olarak görülmektedir. Bu sebeple yalnızca bilgi işleme süreçlerinden oluşan işlemlerin makine ve bilgisayarlar tarafından rahatlıkla taklit edilebildiği görülmektedir. Satranç oynayan bilgisayar, bilincin en kolay problemi olarak adlandırdığımız bilgi işleme sürecinin yapay zekada yeniden üretilebildiğinin açık bir göstergesidir.

6.2.2. Bilincin Kolay Problemi

Bilincin kolay problemi, bilincin en kolay problemine oranla daha fazla bağlantının kurulmasını ve buna bağlı olarak çözümlenmesi için daha fazla zamanı gerektirmektedir. Bilincin kolay problemi, bizim teorimizde bilincin nöral korelasyonlarına karşılık gelmektedir. Nöral bağıntıların, işlevsel bağıntılar kadar kolay tespit edilememesinin sebebi ise beynin sahip olduğu girift yapıdır. Bilinçli durumlarda, beynin hangi bölgelerinin aktif olduğu, hangi berimsel işlemlerin gerçekleştiği, hangi sinaptik bağlantıların oluştuğu soruları fMRI, PET gibi beyin görüntüleme tekniklerinden yararlanılarak cevaplanmaya çalışılmaktadır. Bilincin nöral korelasyonlarının keşfi, bizi bilinçli durumları bilinçsiz durumlardan ayıracak kesin ölçütlere ulaştıracağından bilinç araştırmaları için merkezi bir konumdadır. Bilinçli durumlar ile bilinçsiz durumlar kıyaslandığında beyindeki elektriksel olayların farklılık gösterdiği gözlemlenmektedir. Bilincin nöron bağıntıları tezi üzerine çalışan araştırmacılar, daha çok bilinçli davranış ve fonksiyon üzerine inceleme yapma imkânı bulundukça, bilincin insan beyninde yer alan daha çok bağıntısına ulaşılacağını öngörmektedir.. Bilincin nöral korelasyonları net olarak tespit edildiğinde bilincin zor problemi olan qualia problemi için de bilimsel bir çerçeve çizilmiş olacaktır. Zira beyinde bilince yol açan tüm elektriksel, kimyasal ve nöral olaylar ortaya çıktığında, fiziksel özelliklerin nasıl olup da öznel deneyime yol açtığının açıklanması için bir kapı aralanmış olacaktır. Ancak bu bağlamda kullanılacak yöntemler ile uygulanacak deneyler oldukça kısıtlıdır.

Bu bağlamda EEG, fMRI, PET gibi var olan tarama ve görüntüleme teknikleri hem zamansal ve mekânsal duyarlılık konusunda hem de çok sayıda nörona ulaşma konusunda yeterli değildir. Bu teknik ve bilimsel yetersizlikler giderildiğinde ilkesel olarak bilincin nöral korelasyonlarının keşfi için bir engel bulunmamaktadır. Duyulardan gelen bilginin zihnimize nasıl eriştiği ve nasıl başkalarına anlatılabilir olduğu problemi çözüldükten sonra, bilincin ilk iki aşamadaki çok kolay ve kolay problemleri açıklanmakla kalmayıp, içsel niteliklerin gizemini aydınlatmak için de bir yol yaratılmış olacaktır (Dehaene, 2016: 23).

Her ne kadar zihin-beden özdeşliği teorisi, beynin işleyişinin tam olarak çözümlenmesi ile birlikte zihne dair problematik oluşturacak hiçbir unsur kalmayacağını savunsa da deneyimin öznel niteliğini açıklayacak bir sonucu hâlâ elde edemeyeceğimizi iddia eden düşünürlerin sayısı küçümsenmeyecek kadar fazladır. Bu sebeple teknolojik kısıtlamalarak aşılarak bilincin beyinle olan ilişkisi tam olarak şeffaf bir görünüme kavuştuğunda, bilimsel bilinç çalışmalarının ele alacağı konular şüphesiz ki bilimin "nesnellik" ölçütünün dışında kalan problemler olacaktır. Özgür irade problemi, "ben" veya "kendilik" problemi,

yönelimsellik, zihinsel nedensellik gibi çetin ontolojik problemler, beynin detaylı anatomik haritası çıkarıldığında ve bilincin tüm nöral korelasyonları tespit edildiğinde dahi var olmaya devam edecektir.

Bu tezde savunulan bakış açısına göre, tüm bu ontolojik problemlerin altında yatan ana sebep öznelliğin, bilimsel dilin sınırları dışarısında kalan bir özellik olmasıdır. Bilişsel ve nöral katmanın problemlerine ait olan dil ile öznel ve metafizik katmana ait olan dil birbirlerinden kökten farklıdır. Bilişsel bilim ve nörobilimin söylemleri, hesaplama, girdi-çıktı, nöronlar, nöron korelasyonları, nöron sistemleri gibi fiziksel dünyaya atıfta bulunan ifadelerden oluşurken; bundan sonraki katmanların dili yönelimler, niyetler, değerler, duygular, eylemler gibi salt fiziksel olmayan durumlardan oluşmaktadır.

Katmanlara ait problemleri tasvir etmek için kullanılan bu dil farklılığı iki ayrı basamakta ele alınabilir. Bunlar semantik-epistemolojik basamak ve ontolojik basamaktır. Öznellik ile nesnellik söylemleri birbirinden türetilemez ve birbirlerine indirgenemez farklı dillere sahiptir (Ricoeur, 2018: 23). Bu tez çalışmasında, bu dil farklılığı önce semantik-epistemolojik basamakta ele alınacak, ardından bu epistemolojik çözüm önerisinin öznelliğin yarattığı ontolojik problemin çözümü için yeterli olup olmadığını tartışılacak ve ulaşılan sonuçlar bu tartışma zemininde temellendirilecektir.

6.2.3. Bilincin Zor Problemi

Bu bilinç teorisinde, fenomenal bilinçli durumların varlığı, *bilincin zor problemin*e tekabül etmektedir. Kimilerine göre, bu zor problem bilincin çözülemez, gizemli yanıdır. Kimilerine göre ise öznel durumların nesnel mekanizmalarına ulaşmak bilimsel yöntemler ile mümkündür. Kimi katı indirgemeci materyalist yaklaşım savunucuları ise zor problemin varlığını hiçbir şekilde kabul etmemektedir. Ancak bu çalışma boyunca vurgulandığı üzere bilincin kendisi öznel deneyim kavramını tanım gereği içermesi gereken bir mefhumdur. Dolayısıyla bilincin öznel niteliğini yok saymak, bilincin kendisini yok saymak olacaktır. Fakat öte yandan öznelliği açıklanamaz bir gizem olarak ele almak da bilincin bilimsel incelemesi üzerine olan tüm nörobilim ve yapay bilinç çalışmalarını boşa çıkaracaktır.

Bu çalışmada ise bilinci gizemli ve çözülemez bir zihin özelliği olarak ele almak yerine, bilinci kademeli bir yapı olarak değerlendirip bilince ait çözülebilir sorunlarla, çözülmesi zor sorunları net bir şekilde ortaya koyabilmek hedeflenmektedir. Buna bağlı olarak da tezin

bu son bölümünde öznellik, ontolojik bağlamının dışında yalnızca epistemolojik basamakta ele alınarak fiziksel ve zihinsel ilişkisi üzerine olan problemin aşılıp aşılamayacağı tartışılacaktır. Ancak öznelliğin açıklanması için görünen o ki yalnızca bilincin nöral korelasyonlarını tespit etmek yeterli olmayacaktır. Bu bağıntıların keşfi, bize fenomenal deneyim üzerine tartışmamız için kuşatıcı bir çerçeve sağlayacak ancak deneyimin kendisini vermeyecektir. Bu sebeple bilinç konusunda yapılacak en iyi şeylerden birisi bir yandan nörobilim ve bilişsel bilim çalışmalarına devam edip deneysel çalışmaları sürdürürken bir yandan da öznel deneyimi uygun bir şekilde düşünmemizi sağlayacak kavramsal ve dolayısıyla felsefi çalışmalar yapmaktır (Tura, 2016: 141).

Bilincin zor probleminin temel sebebi olan öznellik, "fenomenal bilinç", "qualia", "öznel deneyim", "içsel durumlar" gibi farklı problematiklerin de temelinde yer almaktadır. Dolayısıyla zor problemin tartışılması bu kavramların hepsinin ortak olarak tartışılması anlamına gelmektedir. Açıkça görülmektedir ki tüm bu kavramlar özne-nesne ayrımını ve ayrık bir öznenin varlığını zorunlu kılmaktadır. Bu çalışmada bilincin doğasına yönelik ortaya koyma çabasında olunan "fenomenal gerçekçilik" yaklaşımı ise önce bu kavramları epistemolojik sahada deneyimleyenin varlığını zorunlu koşul olarak görmeden tartışacak, ardından da bu tartışmanın deneyimleyenenin ontolojik statüsüyle ile ilgili problemi ortadan kaldırıp kaldırmadığını değerlendirecektir.

Geleneksel bakış açısında, deneyimlenen nesneler dünyası ve bu nesneler dünyasının deneyimi olmak üzere iki ayrı varlık alanı kabul edilmektedir. Buna göre, nesnelerin dünyası ya da fiziksel dünya kendiliğinden vardır. Fiziksel dünyanın varlığı çevrenin geri kalanıyla etkileşimine bağlı değildir. Bir nesnenin var olabilmesi için başka bir şeyin var olmasına ihtiyacı yoktur. Ancak, bu çalışmada bilince epistemolojik bakış açısı olarak adlandırılan yaklaşımda şu iddia edilmektedir: Bir şeyin var olması başka bir şeyle etkileşime girmesine bağlıdır ve bu varoluş şartı hem deneyimleyen hem de deneyimlenen için geçerlidir. Ne deneyimleyenin ne de deneyimlenenin varlığı birbirleriyle etkileşmedikleri sürece gerçeklik ve anlam kazanamayacaktır. Bu bağlamda herhangi bir şeyin -deneyimleyen ya da deneyimlenen- kendiliğinden var olduğu inancından vazgeçilmedikçe dış dünya ve deneyim karşıtlığını, başka bir deyişle, fiziksel olayların nasıl öznel deneyime neden olduğu sorununu aşmak mümkün görünmemektedir. Buna bağlı olarak ortaya konulan epistemolojik bakış açısı şunu savunmaktadır: Dünyayı gerçek ve anlamlı kılan, deneyimleyenler ve deneyimlenenler değil, zaman ve mekân içerisinde gerçekleşen deneyimleyen-deneyimlenen

etkileşimleridir. Deneyim olarak isimlendirdiğimiz şey ise deneyimleyen ve deneyimlenen etkileşiminin bizzat kendisidir.

Biz dünyayı, dünyanın imgesi olarak değil, dünyanın kendisi olarak deneyimlediğimizi düşünürüz. Bu üstü kapalı varsayımımıza göre, deneyimlediğimiz dünya temsilî bir dünya değil, gerçek bir dünyadır. Buradan hareketle, biz deneyimlerimizin de dünyanın kendisi kadar gerçek olduğunu düşünürüz. Deneyimleyen ve deneyimlenen, deneyim için zorunlu koşullardır. Farklı bir bakış açısından da deneyim, deneyimleyen ve deneyimlenenin varlığına anlam kazandıran süreçtir. Bu bağlamda, hem deneyimlenenin kendisi hem de bizim o deneyimlenen nesneye dair öznel deneyimimiz etkileşim süreci olmadan bilinebilen bir varlığa sahip olamazlar. Bu varsayımımız şu iddiayı temel almaktadır: Nesne olarak deneyimlenen ve özne olarak deneyimleyen, aynı süreci tanımlamanın farklı iki yoludur. Hem deneyimleyen hem de deneyimlenen, gerçeklikleri birbirlerine zorunlu olarak bağlı etkileşim süreçleridir. Deneyimleyen ve deneyimlenen bu bağlamda bir paranın iki yüzü gibidir. Her ikisi de etkileşim sürecinin, bir başka deyişle, deneyimin kendisiyle özdeştir. Bu çalışmada öznellik problemini epistemolojik olarak aşmak için kullanılan bu perspektifi, Manzotti "radikal dışsalcılık" adını verdiği yaklaşımında benzer şekilde ele almaktadır:

> Dünya öznel ve nesnel alanlarda gerçekleşen ayrık olaylardan değil içkin olarak birbirleriyle ilişkili süreçlerden oluşmaktadır; bu sebeple özne ve nesne ayrı değildir. Fenomenal temsil problemi diye bir problem yoktur çünkü dünyanın deneyimi ve oluşu özdeştir (Manzotti, 2007: 182).

Manzotti, deneyimleyen ve deneyimleneni özdeş olarak görmektedir. Ancak bu bakış açısı, deneyimleyen ve deneyimleneni gerçeklik hâline dönüştürenin deneyim olduğu iddiası ile ele alındığında haklı olsa da durumu tersinden değerlendirdiğimizde, deneyimleyen ve deneyimlenen olmadan deneyimin imkânsız olduğunu söyleyemeyeceğinden eksik görünmektedir.

Manzotti, deneyimleyenin varlığını ontolojik bir problem olmaktan çıkarmak amacıyla deneyimi, deneyimleyene ait değil, dış dünyaya ait bir fenomen olarak ele almaktadır. Bir kitabı gördüğümüzde ya da başka bir şekilde deneyimlediğimizde, retinamız, sinirsel süreçler, optik sinirler, kitap molekülleri, moleküllerdeki atomlar vs. gibi çok sayıda faktör bu deneyimleme sürecine dâhil olmaktadır. Ancak bu faktörlerine hiçbirisi tek başına deneyimin kendisi değildir. Bu faktörlerin hepsi deneyim için gerekli hatta zorunlu şartlar olsa bile deneyimin kendisi olmaktan çok uzaktadırlar. Bu sebeple Manzotti'ye göre deneyimin kendisini, beynin içinde ya da dışında konumlandırmak bilinç problemini

çözümsüz bir noktaya sürüklemektir. Kitaba dair öznel deneyimimizi oluşturan algılarımız bize kitabın kırmızı renkte, orta kalınlıkta, yazılarla dolu, dikdörtgen şeklinde vs. olduğunu söylemektedir. Algımıza sunulan bu nitelikler ne bizim beynimizde (çünkü beynimizde görüntülenen şey nitelikler değil sinirsel aktivitelerdir) ne de beynimizin ve kitabın dışında bir yerdedir. Buna göre Manzotti, algılamış olduğumuz bu niteliklerin tam olarak kitabın kendisinde olduğunu iddia etmektedir. Deneyiminde bulunduğumuz kitap beynimizin içinde ya da nöronların arasında görüntülenememektedir; zira o kitap dışarıdadır. O hâlde, kitaba dair deneyimimiz de dışarıda olmalıdır. Bu çıkarımla Manzotti'nin karşı olduğu fikir, duyuların "girdi" olarak alıp işledikleri verileri, dış dünyayı yansıtan ikinci, içsel bir zihinsel dünyaya dönüştürmeleridir (Manzotti, 2018: 55). Ancak deneyimi dışarıda konumlandırmak da teleolojik açıdan deneyimleyeni sadece dış dünyaya yönelen bir varlığa dönüştürdüğünden tüm sezgisel eğilimlerimizle çelişir görünmektedir. Bu dışsalcı bakış açısı ancak erişim bilinci ya da durum bilinci için bir açıklama getirebilir, fenomenal bilinci açıklamaktan uzaktır.

Epistemolojik olarak deneyimleyen ve deneyimleneni, aynı deneyimin farklı iki yüzü olarak görmek bilincin zor problemine dair sorunu çözer görünümde olsa da döngüsel olarak deneyimin kendisi açıklamaya çalışıldığında yine deneyimleyen ve deneyimlenene dair bir açıklama sunmak zorunda kalınmaktadır. Bu epistemolojik yaklaşımda, özne pasif bir görev üstlenmekte ve yalnızca dış dünyayla etkileşime girerek deneyimleri toplayan bir görüntü sergilemektedir. Zira deneyimin kendisi bu bakış açısında, deneyimleyende değil dış dünyadadır. Aynı deneyimleri yaşayan insanların aynı "ben"e sahip olması gerektiği gibi çok uç bir iddiaya dönüşebilecek olan bu yaklaşımda, "ben" olarak tanımlanan özne yalnızca bilişsel bir özne olabilir. Bununla kast edilen ise şudur: Bilişsel özne, yalnızca erişim bilincine sahip bir öznedir; başka bir deyişle, bilinçli durumların kendisinde toplandığı bir öznedir. Bu öznenin fenomenal deneyimlerinden bahsetmek söz konusu değildir çünkü epistemolojik yaklaşım, deneyimi, deneyimlenen nesne ile özdeş olarak ele almaktadır. Bu özdeşlik ise fenomenal deneyime ve "ben"in pasif değil aktif bir özne olmasına müsaade etmemektedir.

Bu çalışmanın başlangıcında ortaya konulduğu üzere, "ben" ya da "deneyimleyen" bir yanılgı değil, bir gerçekliktir. Bu gerçeklik özne ve nesne etkileşimlerinden oluşan deneyim süreçlerine zorunlu olarak bağlıdır. Dolayısıyla fenomenal gerçekçilik yaklaşımı deneyimin varlığını zorunlu kılmaktadır. Tüm yaşantının sahibi olan bir "ben"e sahip olduğumuza dair

güçlü inancımız, başka bir deyişle ben hissimiz aslında, bilinçli bir deneyimin sonucudur.

"Ben", binlerce deneyimin iç içe geçmesiyle dönüşen aktif bir öznedir. Deneyim yelpazemize eklenen her yeni deneyim, "ben"i dönüştürür ve şekillendirir; dolayısıyla "ben" sabit bir özneye değil, oluş ve değişim içerisinde olan bir özneye işaret eder. Bu oluş ve değişim içerisinde olan özne, devinimsel bir yaşantının kurucusu konumundadır. "Ben" değiştikçe, fiziksel dünyanın deneyimi de değişir; bundandır ki fiziksel dünyaya olan deneyimim hiçbir zaman bir öncekiyle aynı değildir. Bir filmi kaç kez izlerseniz, o filmi o kadar farklı şekilde deneyimlersiniz; aynı şekilde ikinci kez okuduğunuz bir kitabın deneyimi bir öncekiyle özdeş değildir.

Deneyimlerin "ben" üzerindeki kurucu ve yapılandırıcı etkisini oluşturan en önemli faktör şüphesiz hafızadır. Hafıza, ben ve deneyimlerim arasındaki ilişkinin devamlılığını sağlayan köprüdür. "Ben"in kesintisiz bir akış içerisinde olmasının yegâne sebebi sahip olduğumuz hafızadır. Hafıza deneyimleri birbirleriyle ilişkilendiren ve "ben"i bir deneyimler yığını olmaktan kurtarıp aktif bir özneye dönüştüren en önemli ögedir. Bilincimin aktif olmadığını varsaydığım 8 saatlik bir uyku süresinin ardından sabah yine aynı insan olarak uyanmamın sebebi olan hafıza, aynı zamanda "ben"in devamlılığının en büyük garantisidir. Deneyimlerim ile benim aramda yer alan hafıza köprüsü zarar gördüğünde tüm zihinsel yaşantım kesintiye uğrar. Bu bağlamda, hafıza "ben"'i hem muhafaza eder hem de her an yeniden inşa eder. Hafızanın beyinle olan güçlü ilişkisi ise beyin hasarlarında karşılaşılan hafıza kayıplarıyla açıkça ortadadır.

Bir özne olarak insanın dünyayı deneyimleyebilmesi için bir beyin ve bedene ihtiyacı vardır; ancak beyin ve beden, zaman ve yaşamdan ayrıştırıldığında tek başlarına hiçbir deneyim yaşayamazlar. Kavanozun içinde yaşayan bir beynin bilincinden bahsetmek anlamsızdır. Çünkü bu beyin dış dünya ile etkileşim içerisinde değildir. Oysaki sahip olduğumuz her öznel deneyim, bir nesneyle etkileşim süreci vasıtasıyla oluşmaktadır. Bilinç dediğimiz şey ise öznel deneyimi oluşturan bu süreçlerin hafıza vasıtasıyla meydana getirilen örgüsü ile varlık kazanmaktadır. Zihinsel yaşantı ya da ben yaşantısı, zihnin bilincinde olduğu her şeyi inşa eden bir süreçtir.

Bilincin zor problemi, Manzotti'den hareketle savunmaya çalışılan, deneyimleyen ve deneyimlenenin aynı deneyimin farklı görünüşleri olduğu ve özne-nesne karşıtlığının yalnızca bir yanılsama oldunu iddia eden yaklaşım ile aşılabilir görünümde değildir. Zira bu

bakış açısı deneyimin ontolojik değil, yalnızca epistemolojik olarak değerlendirilmesine müsaade etmektedir. Ancak bu epistemolojik değerlendirme yapılırken deneyimin kendisi parçalara ayrılıp incelenmek istendiğinde ise yine deneyimlenen ve deneyimleyen ile karşı karşıya kalınmaktadır. Bu durumda incelenmek zorunda olunan deneyimleyene yalnızca epistemolojik bir statü atfedildiğinde ise özne, yalnızca bilişsel bir özneye indirgenmiş olmaktadır. Diğer bir yandan da bu bilişsel öznenin tek bir deneyimden değil de deneyimlenen ile olan etkileşim süreçlerinden oluştuğu varsayıldığında bu öznenin sosyal bir özne olduğu da aynı anda iddia edilebilmektedir. Çünkü deneyimleyen, dış çevreyle etkileşimde bulunabilen sosyal bir yapıya sahiptir. Ancak bu bilişsel ve sosyal özne, bilincin zor problemine temel teşkil eden fenomenal deneyime sahip değildir. Bilişsel ve sosyal süreçler, tek başlarına fenomenal bilinçli durumlar yaratmak için yeterli değillerdir.

Epistemolojik yaklaşım ile deneyimleyen ve deneyimlenenin ayrık varlık sahaları olduğunu temel alan ikici görüş reddedilerek, bilincin zor problemi ontolojik bir problem olmaktan kurtarılıp epistemik bir yaklaşımla değerlendirilmeye çalışılmıştır. Bu yaklaşım, deneyimleyen ve deneyimlenenin etkileşim süreçlerinin farklı iki görüntüsü olarak ele alınmasının, deneyimin varlık kategorisini değil epistemik kategorisini bilmeyi gerektirdiğini varsaymaktadır. Bu çalışmada da epistemolojik yaklaşım ile deneyimleyen ile deneyimlenen arasındaki, başka bir deyişle, birinci şahıs ile üçüncü şahıs bakış açısı arasındaki ontolojik gediğin kapatılması öngörülmüşse de özne ve nesnenin kendisinde buluştuğu deneyimin doğasına dair söylemek istenilen her sözde yeniden deneyimleyen ve deneyimlenen ile karşılaşmaktan kaçınılamadığı fark edilmiştir. Bununla birlikte, bilincin zor problemini yalnızca epistemolojik basamakta ele almanın, fiziksel dünyanın doğası ile deneyimin doğası arasındaki açıklama boşluğunun doldurulması için katkıda bulunamadığı gösterilmiştir. Ulaşılan sonuç şudur ki; bilincin doğasına epistemolojik yaklaşım bilincin zor problemini ancak yok saymayı başarabilir. Fakat deneyimlenenin doğasına dair büyük ontolojik problem var olmaya devam etmektedir.

Buraya kadar, deneyimleyen ve deneyimlenen iki karşıt varlık alanı olarak ele alınmadan epistemolojik yaklaşım bilincin zor problemi için bir çözüm önerisi olarak değerlendirilmiştir. Bu çözüm önerisinin geçerliliği gösterilebilseydi, aynı zamanda yapay bir sistem için bilincin öznel yanının bir problem teşkil etmeyeceği iddiası geçerlilik kazanacaktı. Zira bu yaklaşıma göre, bilincin öznelliği, deneyimlerin öznelliğinden kaynaklanmaktadır. Deneyim ise özne ve nesne arasındaki etkileşime karşılık gelmektedir. Ancak bu etkileşimin kendisinin açıklanması ise yine deneyimleyen ve deneyimlenenin

doğalarına dair bir açıklama ortaya koymayı zorunlu olarak gerektirmektedir. Bu bağlamda tekrar olarak, deneyimlerin öznelliğinin deneyimleyenden kaynaklı olduğu sonucuna ulaşılmıştır. Deneyimleyene epistemolojik yaklaşım içerisinde bir rol verildiğinde de bu rolün deneyimleri bilişsel ve sosyal olarak işlemekten fazlası olamayacağı görülmüştür. Özetle, öznel deneyim problemini epistemolojik olarak çözme amacıyla ortaya atılan bu bakış açısı, fenomenal ve fiziksel durumlar arasındaki açıklama boşluğunu kapatmasa da bilişsel ve sosyal öznenin varlığını göstermesi açısından deneyimleyenin doğasında yer alan ilk iki kademeye temel sağlamıştır.

Sosyal özne olmanın aynı zamanda etik özne olmayı gerektirdiği bu çalışmanın 5. bölümünde tartışılmıştı. Bu sebeple, bilince epistemolojik bakış açısı bilincin ahlak ile olan bağlantısını kurması bakımından da faydalı bir yaklaşım olarak ele alınmalıdır. Bilince yönelik epistemolojik bakış açısı her ne kadar fenomenal durumların varlığını açıklamasa da bilincin betimlenmesi için elverişli bir yol olarak görülmelidir.

Bilinç, bilimsel araştırmalar için sıradışı bir örnektir. Bilinci sıradan bilimsel araştırma konularından farklılaştıran şey, bilincin bilimsel araştırmacının kendisi ile olan ilişkisidir. Edelman ve Tononi bu ilişkiyi şöyle ifade etmektedir:

> Bilimsel betimlemenin diğer nesnelerinden farklı olarak, bilincin nöral temelini araştırırken karakterize etmeye çalıştığımız nöral süreç aslında kendimize gönderme yapar -o biziz, yani bilinçli gözlemciler (Edelman ve Tononi, 2019: 30).

Bilinç araştırmacılarının yapmaya çalıştığı şey diğer bilimsel etkinliklerde olduğu gibi dış dünyanın bir nesnesini açıklamak değil, içsel bir öznenin açıklamasını yapmaktır. Fenomenal bilincin üçüncü şahıs bakış açısına kapalı olmasından dolayı bilinci çözülemez bir gizem olarak kabul eden filozof ve bilim insanlarının sayısı azımsanmayacak kadar fazladır. McGinn bu bilinemezci duruşu şöyle ifade eder:

> Uzun zamandır zihin-beden sorununu çözmek için uğraşıyoruz. En güçlü çabalarımıza inatla direndi. Gizem devam ediyor. Bana kalırsa gizemi çözemeyeceğimizi açık yüreklilikle itiraf etmenin zamanı artık geldi. Fiziksel beyin suyunun nasıl olup da şaraba döndüğüne dair hâlâ bir fikrimiz yok (McGinn, 1989: 349).

Fakat deneyimleyeni değil deneyimin kendisini konu edinen bir yaklaşımın en güçlü yanı bilinci çözülemez bir gizem olarak kenara koymayıp en azından betimsel amaçlı araştırmasının yapılmasına müsaade etmesidir. Şunu kabul etmekteyiz: Bilince yönelik hiçbir tasvir, fenomenal bilinçli durumların öznel deneyiminin yerini alamayacaktır. Bu

bağlamda, bilinç bazı yönleriyle tasvir edilebilir olsa da fenomenal deneyimin semantik içeriğine nesnel bir erişim ilkece mümkün görünmemektedir. Bilince epistemolojik yaklaşım adını verdiğimiz bakış açısı bilincin bilişsel ve sosyal yönlerini, bilişsel özne ve sosyal özne kavramlarıyla açıklama imkânı sağlamaktadır. Bu da ahlak olgusu için gerekli kavramsal zeminin oluşturulması anlamına gelmektedir. Ancak diğer yandan da epistemolojik yaklaşım, fenomenal özneyi davranışsal ve bilişsel olarak sosyal etkileşimde bulunabilen özneye indirgeyerek fenomenal deneyimlerin yaratmış olduğu ontolojik problemi görmezden gelmektedir. Sonuç olarak, bilincin zor probleminin çözümü, fenomenal öznenin doğasının açıklanmasını zorunlu kılmaktadır. Fenomenal öznenin doğasına dair kavrayışımız ise birinci şahıs bakış açısıyla sınırlıdır ve nesnel yaklaşım ile arasındaki gedik çetrefilli bir felsefi problematik olarak var olmaya devam etmektedir.

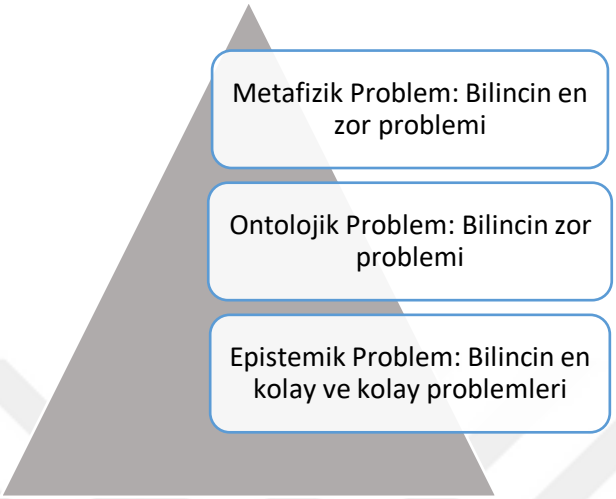
6.2.4. Bilincin En Zor Problemi

Bilincin zor problemine kaynaklık eden bilincin öznel ve niteliksel yapısının ardından, geriye kalan tek aşama *bilincin en zor problemi*ni değerlendirmeye almaktır. Bilincin en zor problemi, bir zihinsel durumun neden bilinçli olduğunu açıklamaktır. Farklı bir şekilde ifade edilecek olursa, "Madde kümelerinin neden bilinçli olabildiğine dair derin, keşfedilmemiş bir açıklama var mıdır, yoksa bu, dünyanın işleyişiyle ilgili açıklanamayan bir gerçek midir?" sorusu bilinç probleminin en üstü kapalı kalan yanıdır (Tegmark, 2019: 367).

Nasıl ve neden bilinçli olduğumuz sorusu, bilincin bilimsel olarak incelemesine dâhil edilmemektedir. Bilincin doğası ve işleyişi bilimsel olarak araştırma konusu olurken bilincin sebebi yalnızca felsefenin konusu olarak kalmıştır. Zira "Bilinç neden var?" sorusu, insanın metafizik yapma yetisiyle üretilmiş bir sorudur ve cevabının da fiziki dünyada bulunmadığı öngörülmektedir. İnsanın dış dünyayı ve bizzat kendi dünyasını kavrayışı, diğer bireyler ve çevreyle girmiş olduğu ilişkilerde kazanmış olduğu yaşama deneyimi ya da "ben" yaşantısıdır. İnsanın "ben yaşantısından" bağımsız bir bilme gücü yoktur.

İnsanın dünya karşısındaki tavrı sürekli bilgi edinmeye çalışan epistemik bir tavır değil, dünyayı anlamaya ve yorumlamaya çalışan hermeneutik bir tavırdır (Özlem, 2004: 194). Bilinç de insanın hermeneutik bir tavır ile sebebini yorumlamaya çalıştığı fenomenlerden birisidir. Bilincin problem olarak adlandırdığımız ilk iki yönü epistemik, üçüncü ve öznel

yönü ontolojik bir meseleyken bilincin en zor problemi birtakım metafizik kabul ve sonuçları içermektedir. Bilimin şu an için yaptığı, metafizik alana hiçbir gönderme yapmadan bilinç mefhumunu açıklamaya çalışmaktır. Bilincin en zor problemi olarak adlandırılan neden bilinçli olduğumuz sorusu bilim camiası tarafından göz ardı edilmektedir.



Şekil 2: Bilincin Problemleri

Bilincin en zor problemi iki farklı açıdan ele alınabilir ya da tamamen yok sayılabilir bir mesele görünümündedir. Neden bilinçli olduğumuzu bilmenin, bilincin doğasını ve bilinçli durumların işleyişini etkileyecek bir bilgi olmadığını öne sürerek, bilincin en zor problemini dikkate almamak kabul edilebilir bir bakış açısıdır. Ancak insan, doğada metafizik düşünce kapasitesine sahip olan tek varlıktır. İnsanın doğadaki üstünlüğü de büyük ölçüde bu metafizik düşünme yetisi ile ilgilidir. Bu yeti, insanın merak duygusu ile birleşerek insanı bilim üretebilen bir varlık hâline getirmiştir. Bu bağlamda, metafizik olarak ele alınan bir sorunun aslında fiziki dünyada bir çözümü olduğu sonradan anlaşılan yüzlerce örnek olay mevcuttur. Neden ve nasıl bilinçli olduğumuz da bugün metafizik bir mesele olarak görülürken daha sonra fiziksel bir açıklamaya kavuşabilir.

Bunun yanı sıra, bilincin sebebi üzerine düşünmek, bilincin doğasına dair çerçevemize de katkı sağlayacak ve farklı bakış açılarının ortaya çıkmasına sebebiyet verecektir. Neden bilinçli olduğumuz sorusu, bilincin tüm işlevsel ve nöral korelasyonları çözüldüğünde cevaplanabilecek bir soru da olabilir; evrenin en büyük gizemi de olabilir. Fakat şu açıktır ki; neden bilinçli olduğumuzun sorgulamasını yapmak, bilincin mahiyetine getirilecek açıklamalara zarar vermeyecek aksine katkı sağlayacaktır. Dolayısıyla insan sahip olduğu metafizik düşünme yetisinden bu hususta faydalanmalı ve bu soruyu göz ardı etmemelidir.

Thomas Nagel "Bir Yarasa Gibi Olmak Nasıl Bir Şeydir?" diye sorarken bilinçli durumların öznelliğine vurgu yaparak, bunun nesnel olarak bilinmesinin mümkün olmadığını ileri sürüyordu. Bir yarasa gibi olmanın ne demek olduğunu bilemesek de kendimiz gibi olmanın ne demek olduğunu biliyoruz. Bunu yalnızca kendimiz için biliyoruz. Fakat "kendimiz gibi olmak" gibi bir şeyin neden var olduğunu merak ediyoruz. Bu sorunun cevabının fenomenal bilinçli durumların nasıl var olduklarına da bir cevap olacağı inancındayız. Bu metafizik sorgulama, bilincimiz üzerine bir bilinç sahibi olduğumuzun açık göstergesidir. Bilinçli olduğumuzun bilincinde olmadan "Neden bilinçliyim?" sorusunu sormak mümkün değildir. Bu demektir ki metafizik katman kendilik bilincine sahip olmayı, başka bir deyişle özbilinci zorunlu kılar. Kendilik bilincindeki "kendilik" ise bir öznenin, deneyimleyenin ya da "ben"in mutlak varlığına işaret eder. William James bunu şöyle ifade eder:

> Bizim doğal olarak uğraşısında olduğumuz bilinç durumları yalnızca kişisel bilinçte, kendiliklerde, somut, tikel "ben"lerde ve "sen"lerde bulunur(…) Psikolojide 'düşünce' yerine 'kişisel kendilik' doğrudan veri muamelesi görmelidir. Evrensel bilinçli olgu, 'hisler ve düşünceler vardır' değildir; fakat 'ben düşünüyorum', 'ben hissediyorum'dur (James, 1950: 225-226).

"Ben" ya da "kendilik" meselesi zihin felsefesi literatüründe de farklı açılardan ele alınmıştır. Materyalist temelli yaklaşımların hemen hepsi "ben"i ya bir yanılsama olarak kabul edip tamamen yok saymaktadırlar, ya da "ben"in varlığını bir gerçeklik olarak değil, kavramsal ya da dilsel bir varoluş biçimi olarak değerlendirmektedirler. Örneğin Dennett'e göre kendilik yalnızca öyküsel bir oyun ya da Kartezyen tiyatrodur; kendilik, beyinlerimizin hem diğerlerine hem de kendimize yönelik olarak sahip olduğu zihinsel durumları tanımlamak için kullandığımız faydalı bir soyutlama şeklinden fazlası değildir (Dennett, 1991: 258). Ancak bu tezin daha en başında, deneyimleyen olmadan deneyimin mümkün olamayacağı postulat olarak ileri sürülmüştü.

Deneyimleyenin varlığına dair bu önkabul sezgisel olmakla birlikte, rasyonel arka planı da olduça güçlüdür. Düşüncelerimde yanılma olasılığım her zaman vardır, ancak o düşüncenin sahibinin "ben" olduğum konusunda yanılamam. Literatürde "bağışıklık prensibi"[63] olarak adlandırılan bu görüşü filozoflar "birinci şahıs zamiri ile ilgili yanlış tanımlama yoluyla hata yapmaya bağışıklık" olarak tanımlarlar (Shoemaker, 1996: 15). Buna göre "Bu elmanın kırmızı olduğunu düşünüyorum." dediğimde, elmanın kırmızı olduğu konusunda yanılabilirim ancak o düşüncenin bana ait olduğu konusunda yanılmam.

[63] *Immunity Principle*.

Bunun yanı sıra "ben", sosyal bağlam içerisinde dilsel olarak yapılmış bir soyutlama ve hikâyesel bir kurgu değildir. Eğer öyle olsaydı, kendilik çok kolay değiştirilebilirdi. Hikâyeler ve kurgular değiştirilebilir. Ancak "ben", irade yardımıyla değiştirilebilen bir doğaya sahip değildir. Kendilik, yalnızca bireysel ya da sosyal olarak gerçek olduğuna inanılan bir yanılsama değildir. Aksine, değiştirilmesi çok zor olan bir "ben"in varlığı, kendiliğe sosyal bakış açısının zorunlu bir parçasıdır (O'Regan, 2012: 122).

Bilincin en zor probleminin çözümü "ben"in varlığının kabulünü zorunlu olarak gerektirmektedir. Birinci şahıs bakış açısı ile "ben"in varlığına dair *eyleyenlik hissi*[64] ve *sahiplik hissi*[65] olmak üzere iki temel hisse sahibiz. Eyleyenlik hissi, eyleme neden olanın ya da üretenin ben olduğum hissiyken; sahiplik hissi, deneyimde bulunanın "ben" olduğum hissidir (Gallagher, 2000: 15). Sahiplik hissi, hareketin iradeye dayalı olup olmamasından bağımsız olarak hareketin deneyimleyene ait olduğudur. Eyleyenlik hissinde aynı zamanda irade sahibi olmak da içerilir. Her iki his de "ben"in bilişsel, sosyal, davranışsal tüm yansımalarından sıyrıldığımızda dahi basit ve ilkel olsa da geriye bizim kendilik adını verdiğimiz bir şey kaldığı fikrine ulaştırmaktadır. Bu kendilik Strawson'un iddia ettiği gibi anlık ve aracısız minimal bir kendilik de olabilir, zamana yayılmış, geçmiş ve geleceğe sahip minimal-olmayan ya da Neisser'in ifadeleriyle öyküsel[66] bir kendilik de olabilir. Minimal kendilik, "ben"in bir çekirdeğiyken; öyküsel kendilik, bu anlık kendiliklerin hafıza yardımıyla örgülerinden oluşmaktadır. Bu bağlamda, bir taraftan öyküsel kendiliğin oluşumu için kendisini meydana getirecek olan minimal kendiliklerin varlığı zorunludur. Bir taraftan da minimal kendilik, hafıza tarafından yorumlanmaya muhtaç olduğundan varoluşunun öyküsel kendilikle bağlantılı olması gerekli görünmektedir (Gallagher, 2000: 15).

[64] *Sense of Agency*
[65] *Sense of Ownership*.
[66] *Extended Self*



Şekil 3: Öznenin Katmanları

Fenomenal gerçekçilik yaklaşımı, minimal kendiliğin varlığına dair bir iddiada daha bulunmaktadır. Bu iddia kaynağı şu varsayımdır: *Her deneyimin bir deneyimleyeni vardır.* Bu varsayıma dayalı olarak çıkarımda bulunulduğunda sahip olunacak ilk deneyimin de bir deneyimleyeni olması gerektiği fikrine ulaşılır. Eğer ki deneyimleyenin tüm bilişsel, davranışsal, sosyal bağlantılarının dışında kalan basit bir öze sahip olmadığı iddia edilirse ilk deneyimin deneyimleyeni olmadığı da iddia edilmelidir. Bu ise mantıksal bir çelişkidir.

Görünen o ki, metafizik bilinç katmanına sahip olan bir özne, minimal ve öyküsel kendilik sahibi olan bir özne olmalıdır. Minimal kendiliğin varlığına, fenomenolojik yöntemle ulaşırız. Fenomenal bilinç, minimal kendiliğin varlığını gerekli kılar. Bunun yanı sıra, hem minimal kendilik hem de öyküsel kendilik, özbilinci zorunlu olarak gerektirmektedir. Bu, bilincin sebebi üzerine tartışılan bilincin en zor probleminin aslında kendilik problemi olduğunu göstermektedir. Bu tezde savunulan yaklaşıma göre, bilincin en zor problemi gerçek bir öznenin varlığıyla doğrudan bağlantılıdır ve çözümü minimal kendiliğin açıklanması yapılmaksızın mümkün değildir.

Fenomenal bilinç, öznel ve niteliksel deneyimdir. Fenomenal bilincin gerçekliğini yok sayan bir bilinç yaklaşımı deneyimleyenin doğasına dair bir açıklama sunamaz. Fenomenal bilincin varlığı, "ben"in varlığının garantörüdür. "Ben" bir yanılsama değildir, bir gerçekliktir. "Ben", fenomenal bir öznedir ve minimal kendilik sahibidir. Bilincin hem zor problemi hem de en zor problemi gerçek bir fenomenal öznenin doğasını açıklamayı zorunlu kılmaktadır. Fenomenal öznenin doğasının açıklanabilmesi ise minimal kendiliğin varlığı kabul edilmeden ilkece mümkün değildir. Dolayısıyla ortaya koyulan bu yaklaşım, "fenomenal gerçekçilik" olarak adlandırılmıştır.

6.2.5. Bilincin Doğasına Yönelik Problemlerin Basamakları

Bu yaklaşımı yapay bilinç bağlamında ele almadan önce hem tutarlılığını test etmek hem de daha anlaşılır kılmak için argümantasyonlarla özetlemek faydalı olacaktır:

1. Argüman: Bilince Dair Problemlerin Sınıflandırılması

Varsayımlar:
1- Bilinç tek yönlü, kompakt bir yapı değil, farklı özelliklerden oluşan bir fenomendir. Bilincin doğasına yönelik geniş bir inceleme için, bu mefhumun barındırdığı problemler parçalı olarak ele alınmalıdır.
2- İnsan bilinci, bilişsel kademe, nöral kademe, öznel kademe ve metafizik kademe olmak üzere dört farklı basamakta ele alınarak incelenmelidir.
İddialar:
1- İnsan bilincinin her bir özelliği çözülmesi gereken farklı bir probleme işaret etmektedir.
2- Bilişsel kademe, bilincin davranışsal ve işlevsel bağlantılarını içermektedir. Davranışsal ve işlevsel durumlar, bilincin bilimsel bakış açısıyla incelemeye en uygun nitelikleri olduğundan, bilince dair en kolay problemlerdir. Bilişsel problem, epistemik bir problemdir.
3- Nöral kademe, bilişsel katmana göre çözümlenmesi daha zor olsa da mümkün olan, bilincin nöral korelasyonlarını içermektedir. Nöral problem, epistemik bir problemdir. Ancak nöral problem çözülse ve bilincin tüm nöral korelatları ortaya çıkarılsa dahi nesnel olanın, nasıl olup da öznel bilince yol açtığı problemi çözülemeyecektir.
4- Fenomenal kademe, bilincin öznel ve niteliksel yanını temsil etmektedir. Bu kademe, nöral kademede açıklanamayan fiziksel süreçlerin ne şekilde yalnızca birinci şahıs bakış açısı tarafından deneyimlenebilen fenomenal süreçlere yol açtığına dair olan zor problemin ait olduğu basamaktır.

5- Metafizik kademe, bilincin nedeninin ele alındığı problem basamağıdır. Bilincin en zor problemi, bir zihinsel durumun neden bilinçli olduğunu açıklamaktır.
Sonuç Kabulleri:
1- Bilişsel ve nöral kademeler bilimsel inceleme için en uygun nitelikleri içermektedirler. Bu nitelikler, bilincin en kolay ve kolay problemleri olarak ele aldığımız **epistemik** meselelere işaret etmektedir.
2- Fenomenal kademe, öznel ve niteliksel deneyimin varlığını sorgulayan **ontolojik** bir problem olan zor probleme işaret etmektedir.
3- Metafizik kademenin içeriği olan neden bilinçli olduğumuz sorusu ise diğer bilinç problemlerinden farklı olarak, **metafizik** bir problemdir.

Bu akıl yürütme ile yapılan sınıflandırılmada bilince dair problemlerin, epistemik bakış açısıyla çözümlenebilecek olan bilişsel ve nöral problemler ile bilincin üzerinde hâlen tartışılan bir fenomen olmasında müsebbip olan ontolojik ve metafizik problemler olduğu görülmüştür. Bunun devamında ise yine bazı varsayım ve öncüllerden yola çıkarak bu ontolojik ve metafizik problemlere dair bir bakış açısı sunulmaya çalışılacaktır:

2. Argüman: Bilincin Zor Problemi

Varsayımlar:
1- Hem öznenin hem de nesnenin varlığı birbirlerinden bağımsız değildir; hem öznenin hem de nesnenin varlığı birbirleriyle etkileşime girmelerine bağlıdır. Özne ve nesne ayrık ontolojik statülere sahip değildir.
2- Deneyim, özne ve nesnenin etkileşim sürecine eşittir. Deneyimlenen ve deneyimleyen, deneyimin farklı iki görüntüsüdür.
Öncüller:
1- Bilincin öznelliğinin çözülmesi zor bir problem olarak değerlendirilmesine dair en önemli gerekçe, özne ve nesnenin ontolojik bağlamda farklı varlık

sahalarına tekabül etmeleridir. Bu bakış açısı deneyimlenen ve deneyimleyen arasında kapanmaz bir gediğe yol açmaktadır.
2- Deneyimleyen ve deneyimlenenin aynı deneyimin yalnızca farklı iki görüntüsü olduğuna dair bir bakış açısı, deneyimlenen ve deneyimleyen arasında oluşan kapanmaz bir ontolojik boşluk olduğu iddiasını bertaraf etmektedir.
Ara Sonuç 1: Bu bakış açısı, öznel deneyimin varlığını ontolojik bir mesele olmaktan çıkarıp epistemik alanda değerlendirilebilir duruma dönüştürebilir görünmektedir.
Ek Varsayım: Deneyimin doğasına dair bir sorgulama da tıpkı bilincin kendisinin sorgulanması gibi, onu oluşturan parçaların ayrı ayrı incelenmesini gerektirir.
3- Deneyim, deneyimleyen ve deneyimlenenin farklı iki görünüşü ise deneyim, parçalara ayrıldığında bileşenlerinin deneyimleyen ve deneyimlenen olduğu görülür. Dolayısıyla deneyimin doğasını açıklamak için deneyimleyen ve deneyimlenenin doğasını bilmeye hâlâ ihtiyacımız vardır.
4- Deneyimin öznelliği, yalnızca deneyimleyenin sahip olduğu birinci şahıs bakış açısından bilinebilir.
Ara Sonuç 2: Deneyimleyen ve deneyimlenenin farklı ontolojik kategoriler olmadığını var saymak ve böylece fenomenal deneyimin kendisini yalnızca epistemolojik basamakta ele almak bilincin zor probleminin çözümü için yeterli değildir. Bu yaklaşım, ancak bilincin zor problemini indirgemeci bir tavırla yok saymayı başarabilir. Deneyimlenenin doğasına dair büyük ontolojik problem var olmaya devam etmektedir. Dolayısıyla 1. ara sonuç geçersizdir.
Ek Varsayım: "Ben"in doğasına yönelik açıklama yapmanın zorluğu da bilincin doğasında olduğu gibi "ben"in sahip olduğu farklı niteliklerden kaynaklanmaktadır. Dolayısıyla "ben" de problematik olarak incelenirken bütün olarak değil, bilinçte olduğu gibi kademelere ayrılarak ele alınmalıdır.
Ek Varsayım: "Ben" probleminin basamakları, sosyal-bilişsel kademe, nöral kademe, fenomenal kademe ve metafizik kademedir.
5- Epistemolojik bakış açısı fenomenal bilinci açıklamak için yetersizdir. Epistemolojik bakış açısının açıklama gücü bilişsel ve nöral kademe ile

sınırlıdır. Buna ek olarak bu yaklaşım, davranışsal ve bilişsel olarak çevresiyle etkileşimde bulunabilen sosyal bir bilince de açıkça elverişlidir.
6- Bu yaklaşımda deneyimleyen, fenomenal bilince sahip olmayan, erişim bilincine sahip bir özne olacaktır.
7- Erişim bilinci, bilişsel katman ile ilişkilidir ve yalnızca erişim bilincine sahip bir özne bilişsel öznedir.
8- Bu bilişsel özne teori gereği aynı zamanda bilişsel ve davranışsal olarak çevresiyle ve diğer öznelerle girdi-çıktı, etki-tepki, neden-sonuç ilişkilerine dayalı olarak etkileşime girebilme imkânına da sahip olduğundan sosyal bir özne olarak da adlandırılabilir.
9- Deneyimi merkeze alarak deneyimleyen ve deneyimleneni gerçekliğin farklı görünüşleri olarak ele alan bir yaklaşımla zor problem olan fenomenal deneyimin varlığına dair ontolojik düzeyde bir açıklama sunulamamaktadır.
Sonuç: Bilincin zor probleminin çözümü, fenomenal öznenin doğasının açıklanmasını zorunlu kılmaktadır. Fenomenal öznenin doğasına dair kavrayışımız birinci şahıs bakış açısıyla sınırlıdır ve nesnel yaklaşım ile arasındaki gedik çetrefilli bir felsefi problematik olarak var olmaya devam etmektedir. Bu bağlamda öznellik epistemolojik bakış açısı ile betimlenebilir olsa da öznelliği bu bakış açısından anlamamız mümkün değildir. Fenomenal deneyimin anlamı yalnızca birinci şahıs bakış açısında vardır.

Fenomenal bilinç ve nesnel dünya arasında var olan açıklama boşluğu gösterildikten sonra, bu adımda ise fenomenal bilincin neden var olduğu sorusunu ve bu soruya verilecek cevabın gerektirdiği minimal kendilik hususunu içeren bilincin en zor problemini serimlemeye çalışacağız.

3. Argüman: Bilincin En Zor Problemi

Varsayımlar:
1- Düşünce tarihi metafizik olduğu kabul edilen birçok problemin, sonradan fiziksel açıklamalarla çözüme kavuşturulduğu birçok örnek içermektedir. Dolayısıyla bilincin metafizik katmanı için de fiziksel bir açıklama yapılması

mümkün olabilir. Ancak bugünkü koşullarda bilincin neden var olduğu sorusu bir gizem olarak kalmaya devam etmektedir.
Öncüller:
1- Fenomenal bilinçli deneyim, kendisine sahip olan deneyimleyeni zorunlu kılar.
2- Deneyimleyenin varlığı bir yanılsama değildir çünkü zihin içeriklerimin doğruluğu konusunda yanılabilirim, ama o zihin içeriğine sahip olanın ben olduğum konusunda yanılamam.
3- Deneyimleyenin varlığı bir kurgu değildir çünkü kendilik değiştirilemez bir doğaya sahiptir. Oysa kurgular değiştirilebilir.
Ara Sonuç: Deneyimleyen, kurgu ya da bir yanılsama değil, bir gerçekliktir.
4- "Ben"in gerçek bir varoluş şekli olduğuna dair eyleyenlik hissi ve sahiplik hissi olmak üzere iki temel hisse sahibiz.
5- Bu iki his bizi minimal kendilik ve öyküsel kendiliğin varlığına ulaştırmaktadır.
6- Anlık ve aracısız olan minimal kendilik ve geçmişe ve geleceğe sahip olan öyküsel kendilik birbirleriyle sıkı ilişki içerisindedirler.
7- Hem minimal kendilik hem de öyküsel kendiliğin varlığı kendilik bilincini (özbilinç) gerektirmektedir.
8- Bu kendilik bilinci bizim "Neden bilinçliyim?" sorusunu sormamızın en temel etkenidir.
9- "Bilinç neden var?" sorusu, insanın metafizik yapma yetisiyle üretilmiş bir sorudur ve cevabının da fiziki dünyada bulunmadığı öngörülmektedir.
10- Bu problemi yok saymak, sahip olduğumuz özbilincin ve kendiliğimizin anlamını yok saymak demektir.
Sonuç: Bilincin en zor problemi metafizik bir problemdir; minimal ve öyküsel kendilik sahibi metafizik özneyi zorunlu kılar.

6.2.6. Alternatif Bir Yapay Bilinç Yaklaşımı

İnsan bilincinin bilişsel kademesinin yapay bir sistemde oluşturulması ilkece mümkündür. İnsan zihnine bilişsel olarak eş-değer bir makine, bilişsel özne olarak kabul edilebilir. Gelişmiş bir bilişsel kapasiteye sahip, kendi dışındaki nesnelerle ve çevresiyle etkileşimde bulunabilen bir robotun kendi deneyimlerine sahip olabileceğini iddia etmek yanlış bir yaklaşım olmaz. Yapay sistemlerin bilişsel olarak gelişmiş olmaları fiziksel nesneye epistemik erişimlerinin de aynı oranda gelişmiş olmasını sağlayacaktır. Yapay sistemin sahip olacağı hafıza, insanın sahip olduğu hafızadan çok daha büyük bir alana sahip olabilir. Bu bağlamda, yapay sistemlerin çok daha fazla bilgiyi birbiriyle ilişkilendirmesi de mümkündür. Yapay zekalı sistemlerin sahip olacağı bilişsel kapasitenin insanın bilişsel kapasitesinden çok daha gelişmiş olması oldukça muhtemel görünmektedir. Dolayısıyla yapay bilincin bilişsel kademesi ile insan bilincinin bilişsel kademesi niceliksel anlamda farklı olacakları gibi, bu nicelik farkının niteliksel bir farka dönüşmesi beklenebilecek bir gelişmedir. Ancak bu niteliksel farkın ne şekilde tezahür edeceğini şu an felsefi bir tartışma içerisinde tahmin etmek oldukça zordur.

Kendi dışındaki nesnelerle ve çevresiyle etkileşimde bulunabilen yapay bir sistemin sosyal bilince sahip sosyal bir özne olması da ayrıca mümkündür. Çevreye adapte olabilen ve çevresel değişkenlere uygun çıktılar üretebilen ve öğrenebilen yapay zekalı bir sistemde bilincin sosyal kademesinden bahsetmek yanlış olmayacaktır. Sosyal bilinç sahibi bir yapay zekanın etik kurallara sahip olması ise önemli bir gerekliliktir. Ahlak, sosyal düzen içerisinde ortaya çıkar ve bu sosyal düzene müdahale imkânı bulunan yapay sistemlerin etik kodları olması büyük önem arz etmektedir. Bu bağlamda, makine bilinci araştırmaları makine etiği araştırmalarından bağımsız yürütülmemelidir.

Yapay sistemlerde fenomenal katmanın oluşturulup oluşturulamayacağı ise farklı bakış açılarından tartışılabilir bir meseledir. Fenomenal bilinçli durumlar ile fiziksel nitelikler arasındaki ilişkinin doğasına dair kesin bir kavrayışa sahip değiliz. Yine de bu kavrayış eksikliğine rağmen, insan beyninin nöral haritası tamamlansa, bilişsel süreçler bütünüyle ortaya konulsa dahi fenomenal deneyimin doğasına bir açıklama getirilemeyeceğini iddia ediyoruz. Ancak fenomenal durumların ortaya çıkışını fiziksel süreçlerle açıklayamayız demek; bizi "Yapay sistemlerde fenomenal bilinç mantıksal olarak mümkün değildir." iddiasına götürmemektedir. Fenomenal durumların ortaya çıkısını fiziksel süreçlerle açıklayamayız demek; bize yalnızca, fiziksel süreçlerin fenomenal durumların doğasını açıklamak için yetersiz olduğunu söylemektedir. Fenomenal

durumların doğasının nasıl açıklanacağına dair bir yöntem sunmamaktadır. Bilişsel ve nöral modellemesi tamamlanan bir yapay sistemde –ne şekilde ve hangi aşamada olduğunu bilemesek de- fenomenal durumların oluşması hâlâ ihtimal dâhilindedir. Dolayısıyla yeterince gelişmiş bilişsel niteliklere sahip ya da beynin birebir modellemesiyle oluşturulmuş bir yapay sistemde, fenomenal bilincin ortaya çıkması mantıksal olarak mümkündür. Ancak bu ortaya çıkış rastlantısal olacak ve fiziksel süreçlerin ne şekilde fenomenal durumlara yol açtığını yine ortaya koyamayacaktır. Ayrıca rastlantısal buluşlar, normal bilimsel paradigmanın temel ilkeleriyle çatışmaktadır. Rastlantısallık, öngörülemez sonuçlar barındırır ve bu öngörülemezlik, kontrollü işleyişe sahip olan çağdaş bilimsel yaklaşımın tercih edeceği bir durum değildir.

Ek olarak, tezin en başında ifade edildiği üzere, her deneyimin bir deneyimleyeni vardır. Bu demektir ki; deneyimleyen, deneyimden önce gelmektedir. Başka bir deyişle, sahip olunan ilk deneyim için düşünüldüğünde, o ilk deneyimde bulunmak için, deneyime sahip olacak bir "özne" zorunlu olarak gerekir. Aksi takdirde, deneyimleyenin deneyimle aynı anda ya da ondan sonra oluşması gerekir ki bu da mantıksal bir çelişki doğurur. "Her deneyimin bir deneyimleyeni vardır." postulatı bu sebeple deneyimden önce de var olan, basit ve ilkel, minimal kendiliğin kabulünü gerektirmektedir. Yine, deneyimlerimize dair sahip olunan eyleyenlik ve sahiplik hissine dayanarak söylenebilir ki; fenomenal bilincin varlığı minimal kendiliğin varlığını gerektirmektedir.

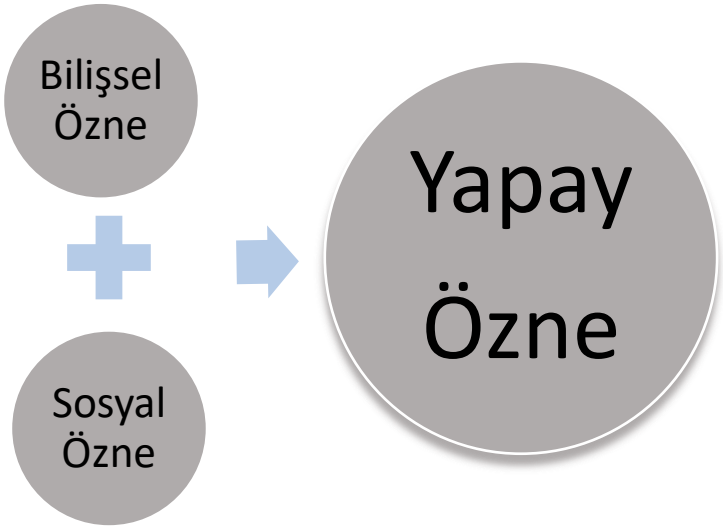
Minimal kendilik, basit ve ilkel olsa da deneyimleyenin tüm bilişsel, davranışsal ve sosyal özelliklerinden sıyrıldığımızda dahi var olmaya devam eden bir gerçekliktir. Yapay sistemler ise bilişsel ve sosyal nitelikleri sıfırlandığında yalnızca basit makinelere dönüşeceklerdir. Minimal kendilik sahibi bir yapay sistemden bahsedilemez. Dolayısıyla yapay fenomenal bilinç mantıksal olarak mümkün görünse de nomolojik olarak mümkün değildir.

Burada iddia edilen, yapay bir sistemde insan bilincin yeniden üretilmesinin mantıksal olarak mümkün olsa da ilkece mümkün olmadığıdır. Yapay sistemlerin sahip olacağı bilinç bilişsel düzeyde bir erişim bilinci olabilir. Dolayısıyla yapay sistemleri, bilişsel özneler olarak düşünerek, insan bilincine olmasa da bir tür bilince sahip oldukları ileri sürülebilir. Yapay sistemlerin bilinci, ontolojik değil epistemolojik düzeyde bir bilinç olarak ele alınabilir. Dolayısıyla bu tezde savunulan yaklaşım, makine bilincini niteliksel olarak insan bilincinden farklı olacağına vurgu yapmakta ve makine bilincinin yeni bir bilinç türü olarak imkânını savunmaktadır.



Şekil 4: Yapay Bilinç

Fenomenal bilincin varlığı, minimal kendiliğin varlığını; minimal kendiliğin varlığı ise kendilik bilincini (özbilinç) zorunlu olarak gerektirmektedir. Minimal kendilik, kendiliğin bilişsel ve sosyal süreçlerden soyutlanmış hâlidir. Bir yapay sistemin, bilişsel ve sosyal özellikleri paranteze alındığında geriye kalan bir kendilikten bahsetmek sezgisel olarak mümkün görünmemektedir. Yine de yapay sistemler, sosyal hayat içerisinde yer alan özneler olarak ele alınacaksa bu özneler yalnızca bilişsel özneler olacaktır. Yapay sistemlerde özneden bahsederken bilişsel ve sosyal etkileşimlerinin toplamı olan bir özneden bahsediyoruz. Ancak insan öznesini tanımlarken bu öznenin deneyimlerin toplamından daha fazlası olduğunu görmüştük. İnsan olarak özne, bilincinin bilincine sahip bir varlık olarak sahip olduğu deneyimlerin ona ait, ona özel ve yalnız onun tarafından bilinebilir olduğunun farkındadır. Yapay bir sistemin böyle bir farkındalığa sahip olması ilkece mümkün görünmemektedir.



Şekil 5: Yapay Öznenin Katmanları

Bilgisayarlar bizden çok daha hızlı hesaplama yapabilmelerine rağmen yaptıkları hesaplamaya dair bir bilince sahip değiller. Makine öğrenmesi çalışmaları, öğrenebilen ve

hedeflerini güncelleyebilen makineler üretmeyi başardı. Kendi programına yeni adımlar ekleyebilen ve hatta kendi programlarını yazan makineler bir bağlamda makineleri yalnızca hesaplama yapan sistemler olarak görmekten vazgeçilmesi gerektiği şeklinde yorumlandı. Ancak bu gelişmeler bize yine de o makinenin semantik içerik sahibi olduğuna dair bir bakış açısını vermeyi başaramamıştır. Öğrenen makinelerde, eylemleri yöneten, kural koyan, kararlar alabilen bir özne yalnızca bilişsel bir öznedir. Bu sebeple bilişsel bilimlerden, yapay zekadan ya da daha geniş olarak yapay sistemlerden ödünç aldığımız girdi-çıktı ilişkisi, hesaplama, öğrenme gibi kavramlarla bilinci açıklamaya çalışmak yerine, yapay sistemleri bilincin geniş çerçevesi içerisinde bir yere yerleştirmemiz gerekir. Bunun en iyi yolu da yapay bilinci de insan bilinci gibi katmanlı olarak ele almaktır. Yukarıda kurgulanmaya çalışılan bilince yönelik bakış açısından hareketle oluşturulabilecek yapay bilinç yaklaşımı şu şekilde formüle edilebilir:

Varsayımlar:
1- Yapay bilinç de insani bilince benzer olarak tek yönlü bir problem değildir. Yapay bilinç problemi de insan bilinci gibi kademelere bölünerek ele alınmalıdır.
Öncüller:
1- Her deneyimin bir deneyimleyeni vardır. Deneyimleyen kavramsal değil, gerçek bir varlıktır. Deneyimleyen, deneyimden önce de var olan minimal kendiliği zorunlu olarak gerektirir.
2- Yapay zeka araştırmalarındaki güncel duruma bakıldığında insan bilincinin bilişsel katmanının yapay bir sistemde oluşturulması ilkece mümkün görünmektedir.
3- Yapay sistemlerin bilişsel olarak gelişmiş olmaları fiziksel nesneye epistemik erişimlerinin de aynı oranda gelişmiş olmasını sağlayacaktır.
4- Kendi dışındaki nesnelerle ve çevresiyle etkileşimde bulunabilen bir yapay sistemin sosyal bilince sahip sosyal bir özne olması mümkün görünmektedir.
5- İnsan bilincinin üçüncü katmanı fenomenal katmandır ve deneyime epistemik yaklaşım fenomenal katmanın işaret ettiği ontolojik probleme çözüm üretememekte; yalnızca özneyi indirgemeci bir tavırla yok saymaktadır.

6- Ontolojik problem, fiziksel süreçlerin nasıl olup da öznel ve niteliksel deneyimlere yol açtığı sorunudur. Bu sorunun, bilincin nöral korelasyonlarının tespit edilmesiyle aşılması mantıksal olarak mümkün olsa da ilkece mümkün değildir. Çünkü bir yapay sistemin fenomenal deneyim sahibi olması rastlantısal olarak gerçekleşme ihtimaline sahipse de fenomenal deneyimlerin nasıl oluştuğu yine bilinemeyecektir.
7- Ontolojik problem aşılmadıkça fenomenal deneyimlerin yapay muadillerinin yeniden üretilmesi ilkece mümkün görünmemektedir.
8- Fenomenal deneyim kendisine sahip olan bir deneyimleyeni ve minimal kendiliği zorunlu kılmaktadır. Deneyimleyen gerçek bir varlıktır.
9- Fenomenal bilinçli bir yapay sistem mantıksal olarak mümkün olsa da nomolojik olarak mümkün değildir.
10- Fenomenal deneyim sahibi olmayan bir sistemin minimal kendilik sahibi metafizik bir özne olması çelişkilidir. Zira metafizik kademenin oluşması, zorunlu olarak fenomenal deneyim gerektirmektedir.
11- İnsan bilincinin son katmanı kendilik bilinci içeren metafizik katmandır.
Ara Varsayım: Metafizik, epistemolojiden farklı olarak anlam bilgisini kendi içerisinde bulundurur. Metafizik bir problemin çözümü için yalnızca bilmek yeterli değildir aynı zamanda yorumlama yetisi de gereklidir.
12- Yapay sistemler doğası gereği semboliktir. Sembollerin varlığı semantiğin varlığını zorunlu olarak doğurmaz. Sınırsız bir bilme kapasitesi olan bir yapay zekada dahi anlamın varlığından bahsetmek için geçerli hiçbir sebep yoktur. Yapay sistemler belli bir algoritmayla belli bir veri tabanı üzerinden hareketle işlemektedirler ve gerçekleştirdikleri tüm işlemler sentaktiktir.
13- Semantik içerikten yoksun bir sistemde metafizik katmanın varlığından bahsetmek mümkün görünmemektedir.
14- Bir sistemin semantik içeriğe sahip olduğunu söylemek için bilişsel kapasitesi yüksek, sosyal etkileşim kurabilen bir yapı olması yeterli değildir. Deneyimin yorumlanması, deneyimin bilinmesinden fazlasını gerektirmektedir.

15- Sembolik sistemler, matematiksel altyapılara sahiptir ve şu an için bu matematiksel altyapı, insan zihnini yeniden üretmek için yeterli değildir. Ancak ileride yeterli bir altyapı ile kurulacak sembolik sistemlerde insan zihninin eşleniğinin üretilmesi mantıksal olarak mümkündür.
16- Metafizik kademe de fenomenal kademe gibi yapay sistemlerde rastlantısal olarak üretilebilir. Ama bu, "Bilinç neden var?" sorusunu cevaplamayacaktır.
17- Bilincin zor problemi (fenomenal problem) ve bilincin en zor problemi (metafizik problem) minimal kendiliğin varlığı kabul edilmeden çözümlenemez. Yapay sistemlerde minimal kendilikten bahsetmek ilkece mümkün değildir.
18- Yapay sistemlerde metafizik kademeden bahsetmek ilkece mümkün değildir.
19- Muhtemel bir yapay bilincin, -mantıksal olarak insan bilincine eşdeğer olabilecek olsa da- ilkece bilişsel ve sosyal olmak üzere iki yönünden bahsedilebilir.
Sonuç: Yapay bilinç ilkece iki katmanlı olarak mümkündür. Dolayısıyla yapay bilinç insan bilincinden farklı yeni bir bilinç türüne işaret etmektedir. Bu bilinç türünün bilişsel kapasitesi insan bilinciyle eşdeğer ya da onu aşmış olacak ancak ilkece fenomenal bilinç sahibi olan kendinin bilincinde metafizik bir özne olamayacaktır. Bu yaklaşımla ortaya konulan yalnızca bilişsel, davranışsal ve sosyal olarak insana benzeyen sistemlerin, felsefi zombilerden farkları ise rastlantısal olarak fenomenal bilinç sahibi olma olasılığına açık olmalarıdır.

6.3. Yapay Uygulamalar

Öne sürülen yapay bilinç yaklaşımının anlaşılması ve uygulanabilirliğinin test edilmesi için en uygun yöntemlerden bir tanesi örnek yapay sistemler üzerinden analiz ve değerlendirme yapmaktır. Var olan yapay sistem uygulamalarının birçoğu, belirlenmiş bir hedefi başarmak için oluşturulmuş olan algoritmaların uygulanmasına dayanmaktadır. Oysa bilinçli özneler tahmin edilemez ve belirli olmayan hedefler geliştirme kapasitesine sahiptirler. Öznelerin davranışlarındaki değişene uyum olarak tanımlanan öğrenme, bir hedef için gerçekleştirilen modifikasyondur (Yılmaz 2017: 39). Sadece belirlenmiş ölçütlere göre davranışın uyarlanması öğrenme için yeterli değildir; aynı zamanda yeni hedefler

ekleme yeteneği de gereklidir. Öğrenme kavramı, hem ihtiyaçlar doğrultusunda davranışta bulunmayı hem de daha önce karşılaşılmamış durumlar karşısında da duruma uygun bir davranış geliştirmeyi içermektedir. Biyolojik organizmaların hedeflerinin hepsi doğuştan belirlenmiş hedefler değildir, özne büyüyüp geliştikçe yeni hedefler üretme yetisini kullanarak davranışlarını şekillendirebilir. Karmaşık yapıya sahip özneler olarak biz insanlarda davranış, deneyim ve hedeflere dayanır; hedefler belli değildir. Hedefler, özne ve çevresinin etkileşiminin sonucudur; geçmişi gelecekle bağlayan, çevreyi bireyle bağlayan nedensel yapılarda saklıdır (Manzotti, 2007: 186).

Yapay sistemler teleolojik/ereksel esnekliklerine göre *Sabit Kontrol Yapıları*, *Öğrenen Yapılar ve Hedef Üreten Yapılar -Teleolojik Olarak Açık Sistemler* olarak sınıflandırılabilir (Manzotti, 2007: 186).

Sabit Kontrol Yapıları olarak adlandırdığımız sistemlerin hedefleri ve bu hedeflere nasıl ulaşacakları önceden belirlenmiştir ve yapılarındaki hiçbir şey deneyimlerinden etkilenmez. Dolayısıyla bu sabit yapıların bilinçli olma durumlarından bahsetmek mümkün değildir. Zira yapay bilinç için zorunlu koşul olarak öne sürülen bilişsel ve fenomenal kademeler bu yapılarda mevcut değildir. Bu sistemlerin bilişsel yetenekleri, var olan algoritmayı uygulamak ile sınırlı olduğundan bilişsel katmanları tamamlanmış değildir. Aynı zamanda öğrenme yeteneklerinin olmaması, çevreyle etkileşim kurmalarını ve dolayısıyla sosyal katman sahibi olmalarını da imkânsız kılmaktadır. Hesap makinesinden, içecek otomatlarına kadar girdi-çıktı ilişkisiyle çalışan makinelerin tamamı bu yapıların alt seviyede örnekleri olsa da 1986 yılında Honda mühendislerinin yürüyen bir robot yaratma projesiyle ortaya çıkan ASIMO[67] sabit kontrol yapılarının en gelişmiş örneklerindendir.

Bir tür Sabit Kontrol Yapısı olan ASIMO derin öğrenme yöntemi ile kendisine yeni hedefler belirleyebilen bir robot değildir. Ona yüklenen sabit algoritma ile çevreyle ilişki kuran ASIMO, tüm tanıma ve öğrenme işlemlerini öncesinde sahip olduğu kodlar ve bu kodların güncellenmesi ile gerçekleştirmektedir. Örneğin, etrafındaki tehlikelerden kaçınması, yalnızca tehlike ona önceden öğretildiğinde mümkün olmaktadır. ASIMO'ya çarpışmanın tehlikeli bir hareket olduğu öğretildiğinde eğer kendisine yaklaşan bir nesne ya da insan görürse, kameraya sahip sensörleriyle algılayıp yön değiştirebilmektedir. Ancak yangının tehlikeli olduğu ona öğretilmemiş ve ateş tanıtılmamışsa, ateşten kaçınması söz konusu

[67] *Advanced Step in Innovative MObility*. Asimo hakkında detaylı bilgi için bkz: Ek: 1

değildir. Ateşe dokunarak zarar görse dahi bunu zarar olarak algılamayacak ve ona ateşten uzak durması kodlanmadığı sürece ateşe dokunmaya devam edecektir çünkü ASIMO kendi deneyimleri ile programını güncelleyebilen ve öğrenebilen bir robot değildir. Yeni hedefler üretemediği gibi, hedefe giden yolları değiştirme ve güncelleme yetisine de sahip değildir.

Sabit Kontrol Yapılarından daha gelişmiş yapay zeka uygulamalarına sahip *Öğrenme Yapıları* ise davranış temelli olup görevleri nasıl yerine getireceklerini öğrenebilirler. Yapay sinir ağlarına dayanan sistemler, bu tür yapıların en çok bilinen örnekleridir. Yapay sinir ağlarının oluşturulma mantığının temelinde kendi kendine öğrenme yer almaktadır (Yılmaz 2017: 64). Öğrenme yapıları özel bir hedefe ulaşmaları istendiğinde sonuca ulaşma yollarına kendileri karar vermekte ve yollar arasında seçim yapabilmektedirler. Ancak hedefe ulaşmak için yeni yollar öğrenebilseler de hedef, görev ile birlikte sisteme a priori olarak tanımlandığından hedef üretmeleri söz konusu değildir. Sadece tanımlanan hedefin farklı yollarını deneyebilirler.

MIT'nin üretmiş olduğu Cog[68] öğrenme yapısına sahip bir robottur. Cog, teleolojik olarak kapalı bir yapı olarak, verili olan hedef doğrultusunda öğrenme ve hedefe giden farklı yolları deneyebilme yetisine sahiptir. Bu kapalılık, onun kendisi için yeni hedefler üretebilmesine engeldir. Belli hedefler doğrultusunda çevresiyle ilişki kuran Cog, davranışlarını gelen etkilere göre değiştirebilse ve karşılaşmadığı durumlara öğrenme yoluyla adapte olabilse de otonom olarak davranışta bulunma ve karar verme yetisine sahip değildir. Bu bağlamda, Cog oldukça gelişmiş bir bilişsel kademeye sahip olmakla birlikte, sosyal katmanı tamamlanmış bir robot değildir. Dolayısıyla Cog için tamamıyla yapay bilinç sahibi bir robot tanımlaması yapmak yanlış olacak olsa da en azından düşük düzeyde bir bilince sahip olduğu ya da başka bir deyişle bilincin zeka ile ilgili olan yanını tamamlamaya yakın olduğu söylenebilir.

Teleolojik olarak açık olup *hedef üretebilen yapılar* ise ilk iki sistemden farklı olarak verilen görevi nasıl yerine getireceğini ve hedefin ne olduğunu öğrenebilirler. İlk iki yapıdan çok daha fazla çevreyle etkileşim içerisinde bulunabilirler. Çevrelerinden ayrık bir şekilde davranış sergilemezler, çevresel süreçlerin oluşumu ve karmaşıklığının arkasında dururlar (Manzotti, 2007: 188). Bu türden yapılar için bugün için üzerinden analiz yapılacak bir örnek bulunmamaktadır. Zira var olan tüm sistemler ne kadar gelişmiş öğrenme yapılarına sahip

[68] Cog'a ait fiziksel özelliklere ve Cog'un yeteneklerine dair bilgilere http://groups.csail.mit.edu/lbr/humanoid-robotics-group/cog/methodology.html ve http://www.ai.mit.edu/projects/humanoid-robotics-group/cog/overview.html adreslerinden ulaşılmıştır.

olurlarsa olsunlar, belli hedefler doğrultusunda eylemde bulunmaktan öteye geçememektedirler. Ancak yeni hedefler üretebilen yapıların oluşturulması için çok uzun bir zamana ihtiyaç olmadığı ve bu teknolojik aşamanın çok yakında aşılacak olduğu yapay zeka araştırmacıları tarafından sıklıkla dile getirilmektedir. Tarih konusunda yapılan tahminler birbirlerinden farklı olsa da bu yapıların imkânı hususunda mutabakat neredeyse sağlanmış görünümdedir. Teleolojik olarak açık yapıların inşası ya Cog gibi bilişsel olarak çok gelişmiş bir robotun öğrenme teknikleri ilerletilerek mümkün olabilir ya da en başından açık bir yapı inşa edilebilecek bir algoritma geliştirilebilir. Bunun yöntemi felsefi bir problem olmamakla birlikte gerçekleşmesi hâlinde ortaya çıkacak yapının bilinç durumunun değerlendirilmesi felsefi bir sorunsal olacaktır. Zira hedef üretebilen yapılar inşa edildiğinde yapay sistemin hem bilişsel katmanı hem de sosyal katmanı tamamlanmış olacaktır. Bu durum, hayvan ve insan bilincinden farklı nitelikte bir bilinç türünün ortaya çıkışı anlamına gelecektir.

Önceleri ruh ya da nefs şeklinde tartışılan zihin -hatta özelinde bilinç- düşünce tarihinin başlangıcından beri dereceli olarak ele alınmıştır. Örneğin Aristoteles ruhu, bitkisel ruh, hayvanî ruh ve insani ruh olarak üç tabakada değerlendirirken; İbn Sînâ nefsi, nebatî, hayvanî ve insani nefs olmak üzere üç ayrı tür şeklinde gruplandırır (Aristoteles, 2011: 69; İbn Sînâ, 2012: 146). Bilincin de bu şekilde farklı varlık türlerinde farklı düzeylerde bulunduğu kabul edilir. Düzey 0 olarak bitkisel bilinç -yalnızca beslenme, büyüme ve üreme düzeyinde bir bilinç-, Düzey 1 olarak böcek ve sürüngenlere ait bilinç –bitkilere ek olarak hareketli ve merkezi sinir sistemi olan canlıların bilinci-, Düzey 2 olarak hayvan bilinçi -düzey 1'e ek olarak duygulara sahip hayvanların sahip olduğu bilinç- ve Düzey 3 olarak da insan bilinçi var olan bilinç türleri olarak kabul edilmektedir (Kaku, 2014: 261).

Bu çalışmada sunulan yaklaşımda, düzey 0 ve düzey 1 bilinç tanımına dâhil değildir. Buna karşılık, hayvan bilinci ve insan bilinci olmak üzere iki bilinç türü kabul edilmektedir. Teleolojik olarak açık yapıların inşası ile mümkün olabilecek yapay bilinç ise üçüncü bir bilinç türü olarak karşımıza çıkacaktır. Bu bilinç türü, bilişsel katmanının gelişmişliği bakımından hayvan bilincinin üstünde yer alacaktır. Ancak metafizik katmana ilkece sahip olamayacak olması, onu insan bilinci ile karşılaştırılamaz bir konuma yerleştirecektir. Zira yapay bilinçte fenomenal ve metafizik katmanın ilkece yokluğu, insan bilinciyle karşılaştırıldığında bir eksiklik gibi görünse de evrenin nesnel bilgisine ulaşma bakımından bir avantaj içeriyor olabilir. Başka bir deyişle, metafizik bakış açısı ve fenomenal deneyim sahibi olması yetisi, evreni anlama ve yorumlama bakımından insana diğer varlık türleri

arasında bir üstünlük sağlıyor olsa da metafizik bakış açımız ve öznel deneyimlerimiz, kimi nesnel gerçekliklerin görünmesinin önünde engel teşkil ediyor da olabilir. Dolayısıyla yapay bilinci, bilinç türleri arasında bir düzeyde konumlandırmak zordur. Ancak şunu söyleyebiliriz ki, bu yeni bilinç türü ile insan bilinci arasında bir karşılaştırma yapmak anlamlı olmayacaktır. Yapay bilinç ve insan bilinci kıyaslanamaz yapılar olarak kabul edilmeli ve yapay bilincin ortaya çıkışı varlıkların sınıflandırılmasında yeni bir paradigma olarak değerlendirilmelidir.

7. SONUÇ VE ÖNERİLER

Yapay zeka ve bilinç kavramlarının her ikisi de bilimin evren tasavvurumuz üzerindeki hegemonyasından fazlaca etkilenmiş olarak göze çarpmaktadır. Bir yandan nörobilimciler bilinci felsefi bir problem olmaktan çıkardıklarını ve felsefenin bilince dair söyleyecek sözü kalmadığını iddia ederlerken; bir yandan da bilişsel bilimler ile bilgisayar teknolojisi araştırmacıları yapay zekayı yalnızca teknolojik bir mesele olarak ele alıp, dijital gelecek senaryolarını salt teknik söylemler üzerine kurmaktan çekinmemektedirler. Bu yaklaşımlara karşıt olarak, bu çalışma hem yapay zekanın hem de bilincin derin felsefi problemler barındırdığı inancıyla başlamıştır.

Felsefenin başlangıcından beri düşünmekle eş anlamlı olarak ele alınmış olan bilinç, insan doğasının en mühim niteliklerinden birisi olarak felsefi bağlamdan koparılamaz bir problematiktir. Yapay zekanın da fikrî ve mantıksal temellerinin büyük bir kısmı felsefe aracılığı ile atılmıştır. Bilinç, insan doğasına ait bir nitelik olması; yapay zeka ise insan zihnini kendisine model alması nedeniyle felsefenin en temel unsuru olan insan ile doğrudan ilişki içerisindedir. Dolayısıyla her iki kavramın da tarihsel arka planlarının yanında, insan doğasına yönelik diğer tüm meseleler gibi felsefi düşünceye muhtaç olan birçok yönü bulunmaktadır. Bu araştırmada asıl gaye, her iki kavramın hem birlikte hem de birbirlerinden bağımsız olarak barındırdıkları felsefi problemleri serimlemektir.

Yapay bilinç ya da makine bilinci, melez bir kavram olarak, yapay zeka ve bilinç kavramlarının bir araya gelmesiyle oluşturulmuştur. Dolayısıyla bir yapay bilinç tartışmasında yapay zeka ve bilinç kavramlarının kapsam ve sınırları ayrık olarak net bir şekilde belirlenmiş olmalıdır. “Bilinçli bir yapay zeka mümkün müdür?” sorusuna cevap aranan bu çalışmada, araştırma planı bu sorunun içeriğini oluşturan bileşenler üzerine kurulmuştur. Buna bağlı olarak, tezin giriş bölümünde kendisinde bilincin olup olamayacağı sorgulanan yapay zeka kavramı tarihsel, mantıksal, nörolojik arka planıyla beraber yöntemsel ve işlevsel açılardan tanımlanmaya çalışılmıştır. Görülmüştür ki, literatüre “yapay zeka” olarak yerleşmiş olan bu kavramın geldiği noktada ve gelecek projeksiyonlarında atfedildiği anlam yapay zekanın ötesine geçmiş durumdadır.

Zekanın içeriği işlevsel, berimsel ve belleksel niteliklerden oluşmaktadır. Oysaki yapay zeka çalışmaları ile hedeflenen bu niteliklerden çok daha fazlasıdır. Hedef, yapay bir zeka inşa etmeyi aşmış, yapay zihin oluşturmaya evrilmiştir. İnsanlığın geleceğini şekillendireceği öngörülen bu teknolojik ilerlemenin menzilini belirleyecek olan, yalnızca makinelerin sahip

olacağı düşünme kapasitesi değildir. Yapay sistemlerin bilinç, duygulanım, irade, değer, inanç gibi işlevselliği aşan hususlardaki potansiyelleri, onların hem fiziksel hem de sosyal düzen içerisindeki görev ve statüleri üzerinde doğrudan belirleyici etkenler olacaktır. Bu etkenler arasında en önemli role sahip olan, insan zihninin ayırt edici unsurlarından olduğunu düşündüğümüz bilinç mefhumudur.

Bilinç sahip olduğu öznel ve niteliksel yapısıyla, kendi varlığımızın farkındalığını veren, kendimizi çevreden ve diğerlerinden ayırmamıza sebep olan, bize biricikliğimizi vurgulayan çok özel bir yapıdır. Bilincin özelliği, ona erişimin yalnızca birinci şahıs bakış açısıyla olmasından ileri gelmektedir. İnsan zihninin yapay bir ortamda üretildiğini iddia etmek için zorunlu koşul, o yapay sistemin bilinçliliğini göstermek olmalıdır. Bilinç sahibi bir yapay zeka, insanın evrendeki ayrıcalıklı konumunu elinden alması öngörülebilecek ve sosyal, epistemolojik, etik ve kültürel bağlamlarda tüm düzenleri yeniden şekillendirebilecek bir varlık türü olacaktır. Ancak böyle büyük bir iddiada bulunmak için, bilinçli yapay zeka tasavvurunun bilimsel dayanaklarının yanı sıra olasılığını sorgulayan felsefi ve düşünsel temellerin de sağlam kurulmuş olması gereklidir.

Bu çalışmada, "Bilinçli bir yapay zeka spekülatif bir kurgudan mı ibarettir yoksa teorik olarak mümkün müdür?" şeklinde ifade edilebilecek bir soruya verilecek cevabın tutarlılığı ve geçerliliğinin, cevabı üstüne kurgulayacağımız bilinç teorisine bağlı olacağı varsayılmıştır. Bu varsayıma dayalı olarak ilk iddia "Yapay bir bilincin imkânını tartışmak için, doğal bilincin mahiyeti açıklanabilmiş ve açıklanamamış nitelikleriyle birlikte ele alınmalıdır." şeklinde belirlenmiştir.

Bu iddia doğrultusunda ikinci bölümde bilince yönelik temel bakış açıları olan idealist, materyalist, düalist, agnostik ve panpsişist yaklaşımlar kendi içlerinde ayrıldıkları alt kollar ile beraber eleştirel bir tutumla ele alınmıştır. Bu eleştirel tutuma temel oluşturacak bakış açısı şu dört ön sav ile ortaya konulmuştur:

1- Fenomenal bilinç, deneyimdir.
2- Deneyim için hem deneyimleyen hem de deneyimlenenin varlığı zorunludur.
3- Hem deneyimleyen hem de deneyimlenen ideal ve düşünsel varlıklar değil, gerçek varlıklardır.
4- Deneyimleyeni ve deneyimleneni anlamaksızın ne deneyimin kendisinin ne de fenomenal bilincin doğası kavranabilir değildir.

Bu kabuller ile hem deneyimin kendisi hem de deneyimleyen ve deneyimlenen, bilince yönelik bir bakış açısının göz ardı etmemesi gereken zorunlu ölçütler olarak ileri sürülmüştür. Bu ölçütler ile irdelenmeye çalışılan teorilerin hiçbirisinin, bütün olarak insan bilincinin doğasını, deneyimleyen özneyi ve deneyimin kendisini dikkate alarak anlamamıza müsaade etmediği görülmüştür.

Üçüncü bölüm ise klasik bakış açılarının sahip olduğu boşluk ve açmazların farkında olarak ortaya atılan ve bilincin bilimsel incelemesine yönelik verileri de işleyerek daha derinlikli perspektifler yakalamış olan güncel felsefi bilinç teorilerinin incelenmesine ayrılmıştır. Dokuz ayrı başlık altında ele alınan bu teoriler, deneyimin öznel ve niteliksel karakterinin yol açtığı problemleri farklı bakış açılarıyla çözmeye çalışmaktadır. Güncel bakış açılarındaki genel eğilim, fenomenal bilinç sorunsalının farkında olan bir kuram geliştirmeye yönelik olsa da fenomenal bilince sahip öznenin doğası bu bakış açılarında da büyük ölçüde yok sayılmaktadır. Ya da tersi bir yaklaşımla özne tek gerçek varoluş biçimi olarak kabul edilirken nesnel dünyanın gerçekliği yadsınarak öznel deneyim ile fiziksel dünya arasındaki gedik kapatılmaya çalışılmaktadır. Yine başlangıçta ortaya koyduğumuz ölçütler ile irdelediğimiz bu teorilerin, bilincin doğasını kavramak için yeterli olmamalarına karşın, savunulabilir bir bilinç yaklaşımı için önemli ipuçları içerdikleri fark edilmiştir. Buradan elde edilen ipuçları ve yaklaşımlar altıncı bölümde sunulan bakış açısı için kullanışlı araçlar olmuşlardır.

Dördüncü bölüm, nihayet problemin hibrit olarak ele alındığı ve yapay bilincin tanım ve teorilerinin ortaya konulduğu bölüm olmuştur. Yapay bilince yönelik bakış açılarını, davranış temelli, bilişsel temelli, nörobilimsel temelli araştırmalar ile fenomenal makine bilinci yaklaşımı oluşturmaktadır. Esasen bu bakış açıları, davranışçılık, işlevselcilik, zihin-beden özdeşliği teorisi ve fenomenolojinin yapay sistemler için olan uyarlamalarından meydana gelmektedir. Dolayısıyla bu yaklaşımların bilince yönelik açıklamalarındaki sorunları aynı şekilde yapay bilince taşımaktadırlar. Fenomenoloji bu bağlamda deneyim ve deneyimleneni dikkate alan faydalı bir yöntem olsa da sentetik fenomenoloji, öznel ve niteliksel durumların yapay bir kurgusunu iddia ederken bu kurgunun imkânını tartışmaktan kaçınmaktadır. Oysaki bilinci kaçınılan ve üzerinde uzlaşılamayan bir problematik hâline getiren, bu fenomenal durumların nasıl oluştuklarına dair doyurucu bir açıklama yapılamamasıdır. Fenomenal durumların epistemik ve ontolojik yapısı çözümlenmeden fenomenal deneyim sahibi bir makinenin bilinçli olacağını iddia etmek aslında hiçbir şey

söylememektir. Elbette ki deneyim sahibi bir varlık bilinçli olacaktır zira fenomenal bilinç, deneyimin kendisidir. Fakat fenomenal deneyimin, fiziksel akıllı sistemlerde yeniden inşasının imkânını tartışmadan ve eğer sentetik fenomenal deneyimin mümkün olduğunu iddia ediyorsak bunu yöntemsel olarak temellendirmeden bu iddia anlamına kavuşamayacaktır.

Beşinci bölüm, yapay bilincin salt teknik bir söylem olmadığı ve felsefi analizinin yalnızca zihin felsefesinin özel bir konusundan ibaret olmadığını ortaya koyma amacı taşımaktadır. Bilinçli bir makine tahayyülü, sosyal, etik, epistemolojik yönleri olan değer sahibi bir makine tahayyülünü de beraberinde getirmektedir. İnsan bilinci kendisini sosyal, hukuki, epistemolojik ve aksiyolojik bağlamlarda açığa vurur. Bu sebeple yapay bilincin de bu bağlamlar içerisinde ele alınması gereklilik arz etmektedir.

Yapay zekanın gelecekte elde edeceği düşünülen güç ve sosyal hayatta erişeceği konuma yönelik sahip olunan korku ve endişenin temelinde yapay zekaların etik ilkelere sahip olup olamayacağı sorunsalı yatmaktadır. Toplumsal temayül, bilinçli makinelerin ancak insanı önceleyen etik ilkelere sahip olmasının garantisi verildiğinde üretimlerinin yasal olması yönündedir. Makine etiği tartışmaları, etiğin ne şekilde makineye aktarılacağı sorunsalıyla başlamakta ve nasıl bir etik tutumun yapay sistemler için daha uygun olacağını araştırmaktadır. Bu minvalde, bu çalışmada da öncelikle makine etiği için önerilen tutum ve yöntemler değerlendirilmiş; etik bir makine algoritmasının mantıkla olan ilişkisi irdelenerek iki değerli mantıktan ziyade bulanık mantık, olasılılık mantığı, kuantum mantığı gibi çok değerli mantıkların insanın ahlaki seçim ve kararlarını sembolleştirmede daha işlevsel olacağı iddia edilmiştir. Ardından ahlakın, bilinçle olan ilişkisi tespit edilmeye çalışılarak makine bilinci ve makine etiği karşılıklı olarak birbirinden beslenmesi gereken alanlar olarak ortaya konulmuştur.

Altıncı bölüm, bu çalışma içerisinde düşünsel olarak en çok uğraş verilen bölüm olarak bu çalışmanın merkezi savlarını içermektedir. Bundan önceki beş bölümün yazılmasındaki temel amacın bu bölümü sağlam temeller üzerinde kurgulayabilmek olduğunu söylemekte bir sakınca yoktur. Bu bölüm, başlangıçta sorulan "Bilinçli bir makine mümkün müdür?" sorusunu, ileri sürdüğümüz yeni bir bakış açısıyla cevaplamayı hedeflemektedir. Dolayısıyla bu çalışmanın bir "tez" olarak adlandırılmasına katkı sağlayan en önemli unsur, altıncı bölümde ileri sürülmüş olan varsayım, öncül ve sonuçlardan oluşan bakış açısıdır.

Savunulan tez başlangıç olarak şunu öne sürmektedir: *Bilinç, her ne kadar parçalanamaz bir varoluşa sahip olsa da bilinci oluşturan temel bağıntı ve nitelikler, çözülmeyi gerektiren farklı problemlere işaret etmektedir. Bilincin tek parçadan oluşan büyük bir problem olarak ele alınması, bilincin fenomenal karakteri hasebiyle, meseleyi hem çözülemez bir gizem hâline getirmekte hem de bilinç tartışmalarını ve buna bağlı olarak yapay bilinç alanını beyhude uğraşlara dönüştürmektedir. Bu sebeple bilincin özsel ögeleri, içerdikleri sorunsallara göre sınıflandırılarak ayrı ayrı ele alınmalıdır.*

Daha sonra belirlenen bu yöntemle bilinç (problematik olarak), *bilişsel, nöral, fenomenal ve metafizik* kademeler oluşacak şekilde parçalanmış; bu kademelerin kapsam alanında olan problemler, *bilişsel problem* -bilincin en kolay problemi-, *nöral problem* -bilincin kolay problemi-, *fenomenal problem* -bilincin zor problemi- ve *metafizik problem* –bilincin en zor problemi- olarak serimlenmiştir. En kolay problemden, en zora doğru yapılan tartışmada bilişsel ve nöral problemler epistemik bakış açısı ile çözülebilir problemlerken; fenomenal kademeye geçiş yaptığımızda problemin kategorisinin ontolojik bir kategoriye evrildiği görülmüştür.

Fenomenal bilinç, deneyimdir; deneyimleyen ve deneyimlenen olmadan deneyim mümkün değildir. Dolayısıyla fenomenal problem, deneyimleyen olarak "özne"nin ontolojik statüsü hakkında bir bakış açısını zorunlu olarak gerektirmektedir. Ancak özne de bilincin kendisi gibi tek bir büyük parça şeklinde ele alındığında hakkında bir yaklaşım geliştirmek zorlaşmaktadır. Bu sebeple bu çalışmada öznenin de kademeli olarak ele alınması önerilmiş ve bu kademeleri meydana getirecek niteliklerin ışığında özne bilişsel, sosyal, fenomenal ve metafizik özne şeklinde dallandırılmıştır.

Fiziksel süreçlerin nasıl olup da öznel ve niteliksel deneyimlere yol açtığını konu alan ontolojik problemin bilincin nöral korelasyonlarının tespit edilmesiyle aşılmasının mümkün olmadığı ileri sürülerek, yapay zeka sahibi sistemlerde ilkece fenomenal bilinçten bahsedilemeyeceği ve bir makinenin fenomenal özne olarak tasavvur edilmesinin ilkece mümkün olamayacağı iddia edilmiştir.

Bununla birlikte, fenomenal bilinç ve kendilik bilincinin varlığını zorunlu olarak gerektiren metafizik kademedenin de yine yapay bir ortamda ilkece mümkün olamayacağı farklı akıl yürütmeler ile desteklenmiştir.

Sonuç olarak, yapay bilincin yalnızca bilişsel ve sosyal olmak üzere iki yönünden bahsedilebileceği görülmüştür. Yapay bilinç ilkece iki katmanlı olarak mümkündür. Bu demektir ki; yapay bilinç insan bilincinden farklı yeni bir bilinç türüne işaret etmektedir. Altıncı bölümde bu bilinç türünün bilişsel kapasitesinin insan bilinciyle eşdeğer ya da onu aşmış olacağı ancak ilkece fenomenal bilinç sahibi olan kendinin bilincinde metafizik bir özne olamayacağı iddia edilmiştir. Tasvir edilen yalnızca bilişsel, davranışsal ve sosyal olarak insana benzeyen sistemlerin, felsefi zombilerden farkları ise rastlantısal olarak fenomenal bilince sahip olma olasılıkları olarak sunulmuştur. Son tahlilde, yapay bilinç ve insan bilincinin kıyaslanamaz yapılar olarak kabul edilmesi ve yapay bilincin ortaya çıkışının varlıkların sınıflandırılmasında yeni bir paradigma olarak değerlendirilmesi gerektiği vurgulanmıştır.

Bu çalışmanın cevap verme amacında olduğu "Yapay bir bilinç mümkün müdür?" sorusunun çeşitli felsefi argümantasyonlara dayalı olarak cevaplanmış olduğu altıncı bölüm, bu tezin teorik sonucunun sunulduğu bölümdür. Ancak bu ulaşılan sonucun sonuçları da en az temel problemin kendisi kadar felsefi temellendirme gerektirmektedir. Zira sosyal hayat içerisinde yer alan bilişsel bilince sahip yapay zekaların da entegre edileceği yeni bir dünya düzeni hakkında öngörülerde bulunmadan önce, bu öngörülerin rasyonel gerekçelerini açığa çıkarabilmek oldukça önemlidir.

Bu çalışmada ulaşılan sonucun uygulamasını hayata geçirecek olan bir robot, nomolojik olarak insan zihninin sahip olduğu fenomenal niteliklerden yoksun olacak olsa da bu robotun bilişsel ve davranışsal özellikleri insandan ayırt edilemez nitelikte olacaktır. "Bana robot olmadığını ispatla." dediğimiz bir insanla, "Bana insan olduğunu ispatla." dediğimiz bir robotun dışavurumları arasında hiçbir fark olmayacaktır. Fenomenal bilinçli durumlar yalnızca içsel olarak bilinebildiklerinden ve üçüncü şahıs bakış açısına kapalı bir yapıya sahip olduklarından bu durumların dışsal gerçeklik ile olan ilişkisi dolaylıdır. Bu bağlamda, simülasyon ve gerçeklik arasındaki ayırt edici faktör nesnel erişime kapalı olduğundan fenomenal bilinçten yoksun bir robot ile bir insan birbirlerinden fiziksel çevreyle etkileşimleri ve sosyo-kültürel düzen içerisinde üstlendikleri roller ile ayırt edilemez hâle gelebilir.

Bir robotun sahip olduğu bilişsel ve sosyal yetilerin fenomenal deneyimlerine sahip olmaması, onun yapacağı işlerin büyüklüğünü ve kuracağı ilişkilerin karmaşıklığını azaltmayacaktır. Öyle ki, bilişsel özne olarak robotlar, akıl yürütme, karar verme ve hesap

yapma gerektiren her türlü işte insandan daha hızlı ve daha doğru sonuca ulaşma kapasitesine sahip olacaklar ve içerisinde bedenlendikleri fiziksel yapının gelişmişliği ile insanın beden gücünü şimdikinden çok daha önemsiz hâle getireceklerdir. Bu durum, insana ait birçok meslek dalının robotlar tarafından icra edilmesine sebebiyet verecek ve iş gücü dağılımındaki bu değişim yeni bir sosyo-ekonomik düzenin inşasını zorunlu kılacaktır. Meslek tanımlamalarının baştan yapılacağı bu yeni düzene karşı insanın alması gereken tedbirler ve yeni düzende insana biçilecek olan rolün ne olduğu tartışıladursun, felsefecilerin bu yeni düzenden en son etkilenecek mesleklerden birisine sahip olduğu söylenebilir. İlkece anlama kapasitesinden yoksun olacakları varsayılan yapay sistemlerin, mantıksal akıl yürütmenin ötesine geçerek felsefi ve metafizik düşünce üretmesi de mümkün olmayacaktır. Bu durum, eğer ki rasyonel tüm akıl yürütmeleri insandan çok daha hızlı ve doğru yapan yapay sistemlerin varlığından pratikte faydalanılırsa, hem felsefenin köklü problemlerine hem de metafizik öge ve inançlara yönelimi artırarak insanın evreni anlamlandırma çabası içerisinde yükselişe geçtiği bir döneme kapı aralayabilecektir.

Bu çalışmada ortaya konulan bakış açısı, yapay zekaların kendilik bilinci kazanıp kendi varoluşlarını güçlendirmek için insan varlığını yok etme gayesinde olacakları, insan türüne köle olarak yaratılmalarına başkaldırarak insanları köleleştirecekleri gibi gerçekleşmesi fenomenal bilinçli durumların varlığına bağlı senaryoları ilkece dayanaksız kılmaktadır. Buna bağlı olarak, gelecekte oluşması muhtemel yeni bir sistem içerisinde insanın konumuna dair belirleyici rol üstlenecek olanın, yapay zeka değil insan olacağını ileri sürmektedir. Yapay zekanın hem fiziksel hem de sosyal, kültürel, ekonomik ve etik yapının parçası olacağı bir gelecek projeksiyonunda insan, bilincinin kendisine sunduğu ayrıcalıklı nitelik ve yetilerin farkına vararak kendi varoluşunu anlamlı ve güçlü kılmaya yönelik bir tavırla, sahip olduğu ahlak, sezgi, inanç ve değerleri yeniden gözden geçirme ve inşa etme şansına erişebilir. Ancak tam tersi bir tavırla da bilişsel, sosyal, ekonomik açılardan yapay zeka ile rekabete girerek rasyonel bir varlık olmasının yanı sıra metafizik düşünce ve değer üretebilen bir varlık olduğunu unutabilir ve kendisini giderek makineleşeceği bir sürecin içerisine sokarak insan türünün yokoluşunu hızlandırabilir. İçinde bulunduğumuz zaman dilimi, insan türünün varoluş ve yokoluş arasında bir seçim yapacağı, yeni kurulacak olan dünya düzenine bir geçiş aşamasıdır. Bu geçiş aşamasının orta yerinde yer alan bizler, bu paradigmatik değişimin insanın fiziksel, zihinsel ve ahlaki yaşantısı üzerine olacak etkisini şekillendirici ve belirleyici rollere sahip bireyler olarak büyük bir sorumluluğu üstlenmekteyiz. Bu bağlamda, felsefecilerin yapay zeka meselesine, bu kavramın içeriği,

potansiyeli, sınırlılıkları ve yönteminin farkında olarak çok daha fazla eğilmeleri ve bu kavramın yarattığı problematikleri farklı boyutlarıyla daha dikkatli değerlendirmeleri, yakın gelecekte gerçekleşmesi muhtemel ekonomik, sosyal, ahlaki, hukuki ve epistemolojik değişim süreçlerinde, insan özne ve makine öznenin sahip olacakları rollerin tartışılması açısından büyük kıymet arz etmektedir.

Son olarak, bu çalışmada savunulan tezin ortaya koyduğu bakış açısı post-hümanizm ya da transhümanizm olarak adlandırılan görüşe de açıkça müsaade etmektedir. Sahip olunan matematiksel altyapı ve sembolik sistem ile bunlara dayalı algoritmaların meydana getireceği yapay bilinç sınırlı olmaya mahkûm görünmektedir. Bu sınırlılık, yapay sistemlerde fenomenal bilincin oluşturulması önünde büyük engel teşkil etmektedir. Ancak fenomenal bilinçli bir makinenin imkânını inkâr etmek, makinelerin bilişsel ve fiziksel üstünlüklerinin insanla birleşebileceği ihtimalini inkâr etmeyi gerektirmemektedir. Giderek insansılaşan robotik uygulamalarla beraber, giderek robotlaşan insan tasarımlarının ortaya çıkması muhtemeldir ve bu robot-insanların oluşturulmasının birden fazla yolu mevcuttur. Bu yollardan ilki, insanın el, kol, bacak gibi organlarının robotik organlarla değiştirilmesi ve bu robotik organların organik bedene uyumlu hâle getirilmesidir. Fiziksel engeli olan bireyler, yaşlılar ve beden gücüyle çalışan işçilerin hayat kalitesini önemli ölçüde arttıracak olan bu yolda, insan bedenine uyumlulukları giderek artan protez uzuvlar üretilmeye başlanmasıyla önemli gelişmeler yaşanmaktadır. İkinci yol, robotik ile genetik çalışmaların bir araya gelerek dönüşüm gösterebilen, kendini yenileyebilen sentetik uzuvlarla oluşturulmuş makine-bedenler oluşturmaktır. Bu çalışmalar, beynin çalışma hızı ve kapasitesini de yapay zeka destekli elektronik çipler yardımıyla arttırmayı öngörmektedir. Fiziksel ve zihinsel nitelikleri bakımından insan türünün üzerinde bir kapasiteye sahip olacak bu makine-insanlar, yeni bir varlık türü olarak konumlandırılabileceklerdir. Transhümanizme giden üçüncü yol ise insan zihninin tüm yönleriyle bilgisayarlar ya da bulutlara yüklenerek zihinsel ölümsüzlüğün sanal ortamda meydana getirilmesidir. Bu yol ile insanı ölümlü bedeninden kurtararak sonsuz bir zihinsel varoluşa kavuşturmak hedeflenmektedir.

Günümüz bilimi henüz bu sentetik varoluş hayalinin çok uzağında yer almaktadır. Ancak bu hayalin gerçekleşme ihtimalinin önünde mantıksal herhangi bir engel bulunmamaktadır. Dolayısıyla insanların robotlaşacağı ya da makineleşeceği öngörülen muhtemel transhümanist bir çağın başlangıcı, insana ölümsüzlük vadediyor olsa da özsel bağlamda

insanın ölümü anlamına gelmektedir. Bu türden bir makine-insan tanımı, insanın doğasına yönelik özelliklerini farklılaştırmayı amaçladığından ortaya çıkacak olacak bu yeni tür, klasik insandan ayrık bir ontolojik kategoriye ait olacak ve insan ile makinenin bir arada varoluşunu ifade edecektir.

Görüldüğü üzere yapay bilincin imkânı üzerine bir sorgulama yürütülen bu çalışmada ulaşılan her sonuç, başka sonuçlara kapı aralaması yönünden aynı zamanda bir başlangıcı ifade etmektedir. Ancak kapsam ve sınırları belirlenmiş bu tez çalışması içerisinde tüm bu sonuçların bütün yönleriyle ele alınması mümkün olmamıştır. Bu çalışmanın felsefi literatüre yapay zeka felsefesinin içerisinde barındırdığı çok sayıda sorunsal içerisinden, en azından "yapay bilinç" özelinde mütevazı bir katkı sağlanması hedeflenmiştir. Bu hususta son bir söz söylemek gerekirse insan türünün geleceğini şekillendirici ve dönüştürücü bir etkiye sahip olacağına inanılan yapay zeka kavramının pek çok yönü daha felsefi sorgulama ile irdelenmeyi beklemektedir. Türkiye'de bu alana olan eğilimin artması ve derinleşmesi burada dile getirilen bir umut olarak kalmamalıdır çünkü artık açık bir zarurettir.

8. KAYNAKLAR

Adams, Bryan; Breazeal, C.; Brooks, R.; Scassellati, B. (2000). "Humanoid Robots: A New Kind of Tool". *Humanoid Robotics*. 25-31.

Alpaydın, Ethem (2004). *Introduction to Machine Learning*. Cambridge: The MIT Press.

Alpaydın, Ethem (2018). *Yapay Öğrenme*. İstanbul: Boğaziçi Üniversitesi Yayınevi.

Anderson, Micheal ve Susan Leigh Anderson (2007). "Machine Ethics: Creating an Ethical Intelligent Agent". *AI Magazine* 28(4): 15-26.

Arıcı, Murat (2015). "Zihinsel Alanda "Ben Hissi" Neye Karşılık Gelir?". *Dört Öge* (7): 65-76.

Aristoteles (1975). *Politika* (Çeviren: Mete Tuncay). İstanbul: Remzi Kitabevi.

Aristoteles (2011). *Ruh Üzerine* (Çeviren: Zeki Özcan). Ankara: Birleşik Yayınevi.

Aydın, Metin (2003). "Jeremy Bentham'ın Ahlak Anlayışı ve Felicific Calculus". *II. Türkiye Lisanüstü Çalışmaları Kongresi: Bildiriler Kitabı III*. 601-615.

Baars, Bernard (1998). *A Cognitive Theory of Consciousness*. Cambridge: Cambridge University Press.

Baars, Bernard (2002). "The Conscious Access Hypothesis: Origins and Recent Evidence". *Trends in Cognitive Sciences*. 6 (1): 47-52.

Barrat, James (2013). *Our Final Invention*: *Artificial Intelligence and the End of Human Era*. New York: Thomas Dunne.

Bayoğlu Kına, Filiz (2018). "W. D. Ross'un "Prima Facie Duty" Kavramı". *FLSF* (*Felsefe ve Sosyal Bilimler Dergisi*) (25): 225-238.

Bentham, Jeremy (2009). "An Introduction Principles of Morals and Legislation". *The Works of Jeremy Bentham*. 1 (J. Browing) Toronto: Liberty Fund.

Berkeley, George (1984). *Hylas ile Philonous Arasında Üç Konuşma* (Çeviren: K. Sahir Sel). İstanbul: Sosyal Yayınlar.

Blackmore, Susan J. (2002). "There is no Stream of Consciousness". *Journal of Consciousness Studies* 9 (5-6): 17-28.

Block, Ned (1978). "Troubles with Functionalism". *Minnesota Studies in the Philosophy of Science* (9): 261-325.

Block, Ned (1995). "On a Confusion about a Function of Consciousness". *Behavioral And Brain Sciences* (18): 227-287.

Bonsignorio, Fabio (2013). "The New Experimental Science of Physical Cognitive Systems. AI, Robotics Neuroscience and Cognitive Sciences under a New Name with the Old Philosophical Problems". *Philosophy and Theory of Artificial Intelligence* (ed. Vincent C. Müller). Heidelberg: Springer.

Bringsjord, Selmer; Licato, J.; Govindarajulu, N. S.; Ghosh, R.; Sen, A. (2015). "Real robots that pass human tests of self-consciousness". *Proceedings of IEEE the 24th International Symposium on Robots and Human Interactive Communications*. 498-504.

Büyük, Celal (2013). "Düalizm, Bilinç ve Tanrı". *Atatürk Üniversitesi İlahiyat Fakültesi Dergisi* (39): 133-158.

Chalmers, David (1996). *The Conscious Mind: In Search of a Fundamental Theory*. Oxford: Oxford University Press.

Chalmers, David (2000). "What is a Neural Correlate of Consciousness". *Neural Correlates of Consciousness*: *Conceptual and Emprical Questions* (ed. Thomas Metzinger). Cambridge: MIT Press, 17-39.

Chalmers, David (2017). "Ben Bilinçliyim ki; Zombi Olan O". *Bilinç Üzerine Konuşmalar* (ed. Susan Blackmore). İstanbul: Küre Yayınları. 51-65.

Changeux, Jean Pierre ve Paul Ricoeur (2018). *Neden Nasıl Düşünürüz? Etik, İnsan Doğası ve Beyin Üzerine Bir Tartışma*. İstanbul: Metis Yayınları

Chrisley, Ron (2003). “Embodied Artificial Intelligence”. *Artificial Intelligence* 149 (1): 131-150.

Chrisley, Ron (2008). “Philosophical Foundations of Artificial Consciousness”. *Artificial Intelligence in Medicine* (44): 119-137.

Chrisley, Ron (2009). “Synthetic Phenomenology”. *International Journal of Machine Consciousness* (1): 53-70.

Churchland, Patricia (2019). *Nörofelsefe* (Çeviren: Özge Yılmaz). İstanbul: Alfa Yayınları.

Churchland, Paul (2012). *Madde ve Bilinç*: *Zihin Felsefesine Güncel Bir Bakış* (Çeviren: Ekrem Berkay Ersöz). İstanbul: Alfa Yayınları.

Clark, Andy ve David J. Chalmers (1998). “The Extended Mind.” *Analysis*. (58): 7-19.

Crick, Francis (1994). *The Astonishing Hypothesis*: *The Scientific Search for the Soul*. New York: Scribners.

Crumley, Jack (2000). *Problems in Mind*. California: Mayfield Publishing Company.

Çilingir, Lokman (2019). *Pratik Felsefe*, Ankara: Elis Yayınları.

Dennett, Daniel C. (1991). *Consciousness Explained.* New York: Back Bay Books/Little, Brown and Company.

Dennett, Daniel C. (2005). *Sweet Dreams*: *Philosophical Obstacles to a Science of Consciousness*. Cambridge: MIT Press.

Dennett, Daniel C. (2017a). *Bilinç Açıklanıyor* (Çeviren: Sibel Kibar). İstanbul: Alfa Yayınları.

Dennett, Daniel C. (2017b). “Bilinçle İlgili Sezgilerinizden Vazgeçmelisiniz”. *Bilinç Üzerine Konuşmalar* (ed. Susan Blackmore). İstanbul: Küre Yayınları. 101-115.

Deheane, Stanislas, Jean-Pierre Changeux ve Lionel Naccache (2011). “The Global Neuronal Workspace Model of Conscious Access: From Neuronal Architectures to Clinical

Applications". *Characterizing Consciousness*: *From Cognition to the Clinic? Research and Perspectives in Neurosciences* (ed: Stanislas Dehaene ve Yves Christen). Berlin: Springer. 55-84.

Dehaene, Stanislas (2016). *Bilinç ve Beyin* (Çeviren: Sibel Sevinç). İstanbul: Alfa Yayınları.

Descartes, René (1968). "The Principles of Philosophy". *The Philosophical Works of Descartes* (ed. S. Haldene ve G.R.T. Ross). Cambridge: Cambridge University Press.

Descartes, René (1970). "Meditations on First Philosophy". *The Philosophical Works of Descartes* (1). Cambridge: Cambridge University Press.

Descartes, Rene (1985). "Treatise on Man". *The Philosophical Writings of Descartes.* Vol.1. Cambridge: Cambridge University press.

Descartes, René (2015). *Felsefenin İlkeleri* (Çeviren: Mesut Akın). İstanbul: Say Yayınları.

Descartes, René (2015). *Yöntem Üzerine Konuşma* (Çeviren: Afşar Timuçin ve Yüksel Timuçin). İstanbul: Bulut Yayınları.

Descartes, René (2017). *Duygular ya da Ruh Hâlleri* (Çeviren: Çiğdem Dürüşken). İstanbul: Alfa Yayınları.

Domingos, Petro (2017). *Master Algoritma Yapay Öğrenme Hayatımızı Nasıl Değiştirecek?* (Çeviren: Tufan Göbekçin). İstanbul: Paloma Yayınevi.

Dölek, Haydar (2010). "Descartes'ın Cevher ve Ruh Anlayışına Leibniz'in Eleştirisi". *Fırat Üniversitesi İlahiyat Fakültesi Dergisi* (15): 215-225.

Dretske, Fred (1994). *Naturalizing the Mind*. Cambridge: MIT Press.

Dretske, Fred (2003). "Experience as Representation". *Philosophical Issues* (13): 67-82.

Dreyfus, Hubert L. (2007). "Why Heideggerian AI Failed and How Fixing It would Require Making It More Heideggerian". *Philosophical Psychology* 20 (2): 247-268.

Edelman, Gerald M. ve Giulio Tononi (2000). "Reentry and Dynamic Core: Neural Correlates of Conscious Experience". *Neural Correlates of Consciousness Empirical and Conceptual Questions* (ed. Thomas Metzinger). Cambridge: MIT Press, 138-151.

Edelman, Gerald M. ve Giulio Tononi (2019). *Bilincin Evreni Maddenin Hayale Dönüşümü* (Çeviren: Ahmet Subaşı). İstanbul: Küre Yayınları.

Elçi, Rafet (2017). *Var Olan*. İstanbul: Litera Yayıncılık.

Ersoy, Çağlar (2017). *Robotlar, Yapay Zeka ve Hukuk*. İstanbul: On İki Levha Yayıncılık.

Feldman, Fred (2016). *Etik Nedir?* (Çeviren: Ferit Burak Aydar). İstanbul: Boğaziçi Üniversitesi Yayınevi.

Ford, Martin (2018). *Robotların Yükselişi Yapay Zeka ve İşsiz Bir Gelecek Tehlikesi* (Çeviren: Cem Duran). İstanbul: Kronik Kitap.

Franklin, Stan ve F.G. Patterson (2006). "The Lida Architecture: Adding New Modes of Learning To An Intelligent, Autonomous, Software Agent". *Proceedings in Integrated Design and Process Technology*. USA: Society for Design and Process Science.

Gallagher, Shaun (2000). "Philosophical Conceptions of the Self: Implications for Cognitive Science". *Trends in Cognitive Sciences* (4): 14-21.

Gamez, David (2008). "Progress in Machine Consciousness". *Consciousness and Cognition*, 887-910.

Gamez, David (2009). "The Potential for Consciousness of Artificial Systems". *International Journal of Machine Consciousness* (1): 213-223.

Gennaro, Rocco J. (2004). "Higher-Order Theories of Consciousness: An Overview". *Higher-Order Theories of Consciousness An Anthology* (ed. Rocco J. Gennaro). Amsterdam: John Benjamins Publishing Company.

Giovagnoli, Raffaele (2013). "Computational Ontology and Deontology". *Philosophy and Theory of Artificial Intelligence* (ed. Vincent C. Müller). Heidelberg: Springer.

Goff, Philip; Seager, William ve Sean Allen-Hermanson (2017). "Panpsychism". *The Stanford Encyclopedia of Philosophy* (*Winter 2017 Edition*). (ed. Edward N. Zalta).

Gottfried, Leibniz (2011). *Monadoloji ve Metafizik Üzerine Konuşma* (Çeviren: Atakan Altınörs). Ankara: Doğu-Batı Yayınları.

Graham, George (2010). "Behaviorism". *The Stanford Encyclopedia of Philosophy* (ed. Edward N. Zalta). <http://platon.stanford.edu/archives/fall2010/entries/behaviorism/>.

Gulick, Robert Van (2008). "Understanding the Phenomenal Mind: Are We All Just Armadillos?" *Mind and Cognition*: *An Anthology* (ed. Lycan W. G. ve J. J. Prinz). Oxford: Blackwell.

Günday, Şeref (2003). *Zihin Felsefesi*. Bursa: Asa Kitabevi.

Haidt, Jonathan ve Joseph Craig (2007). "The Moral Mind How Five Sets of Innate Intuitions Guide the Development of Many Culture-Specifi c Virtues, and Perhaps Even Modules". *The Innate Mind* (ed. Peter Carruthers, Stephen Laurence ve Stephen Stich). Oxford: Oxford University Press. 367-391.

Haikonen, Pentti (2009). "Qualia and Conscious Machines". *International Journal of Machine Consciousness*. 1(2): 225-234.

Hawking, Stephen; Russell, Stuart; Tegmark, Max ve Wilzeck, Frank (2014). "Stephen Hawking: 'Transcendence Looks At the Implications of Artificial Intelligence- But Are We Taking AI Seriously Enough?'". *The Independent*, 1 Mayıs 2014.

Heimsoeth, Heinz (2016). *Kant'ın Felsefesi* (Çeviren: Takiyettin Mengüşoğlu). Ankara: Doğu Batı Yayınları.

Hilbert, David ve W. Ackermann (1950). Principles of Mathematical Logic. New York: Chelsea Publishing Company.

Hobbes, Thomas (2007). *Leviathan* (Çeviren: Semih Lim). İstanbul: YKY.

Holland, Owen (2004). "The Future of Embodied Artificial Intelligence: Machine Consciousness?". Embodied Artificial Intelligence. Berlin: Springer, 37-53.

http://jmc.stanford.edu/articles/dartmouth/dartmouth.pdf

https://www.independent.co.uk/news/science/stephen-hawking-transcendence-looks-at-the-implications-of-artificial-intelligence-but-are-we-taking-9313474.html.

İbn Sīnā (2012). *en-Necāt*: *Felsefenin Temel Konuları* (Çeviren: Kübra Şenel). İstanbul: Kabalcı Yayıncılık.

Igor, Aleksander ve Barry Dunmall (2003). "Axioms and Tests for the Presence of Minimal Consciousness in Agents". *Journal of Consciousness Studies* (10): 630-637.

Igor, Alexander ve H. Morton (2007). "Axiomatic Consciousness Theory For Visual Phenomenology In Artificial Intelligence". *Association for the Advancement of Artificial Intelligence*. 18-23.

Jackson, Frank (1986). "What Mary Didn't Know?. *The Journal of Philosophy* (83): 291-295.

Jaegwon, Kim (2011). *Philosophy of Mind*. Boulder: Westview Press.

James, William (1950). *The Principles of Psychology* (1). New York: Dover Publications.

Kaku, Michio (2017). *Zihnin Geleceği Bilimin Zihni Anlamaya ve Geliştirmeye Yönelik Arayışları* (Çeviren: Emre Kumral). Ankara: ODTÜ Yayıncılık.

Kant, Immanuel (2016). *Pratik Aklın Eleştirisi* (Çeviren: Ionna Kuçuradi, Ülker Gökberk ve Füsun Akatlı). Ankara: Türkiye Felsefe Kurumu.

Kim, Jaegwon (1998). *Mind in a Physical World*: *An Essay on the Mind-Body Problem and Mental Causation*. Cambridge: MIT Press.

Kim, Jaegwon (2011). *Philosophy of Mind*. USA: Westview Press.

Koch, Christof (2004). *The Quest for Consciousness*: *A Neurobiological Approach*. Colorado: Roberts ve Company Publishers.

Kurzweil, Ray (2015). *Bir Zihin Yaratmak İnsan Düşüncesinin Esrarı* (Çeviren: Dilara Gostolüpçe). İstanbul: İstanbul Bilgi Üniversitesi Yayınları.

Kurzweil, Ray (2017). *İnsanlık 2.0* (Çeviren: Mine Şengel). İstanbul: Alfa Yayınları.

Leibniz, G. W. (1997). *Monadoloji* (Çeviren: Suut Kemal Yetkin). İstanbul: MEB Yayınları.

Lettvin, J. Y.; Maturana, H. R.; McCulloch, W. S. ve W. H. Pitts (1959). "What the Frog's Eye Tells the Frog's Brain". *Proc. IRE*, 47. 1940-1951; reprinted in Warren S. McCulloch, *Embodiments of Mind*. MIT Press, 1965.

Levy, David (2009). "The Ethical Treatment of Artificially Conscious Robots?". *International Journal of Social Robotics* (1): 209-216.

Lewis, David (1983). "Postscript to 'Mad Pain and Martian Pain'", *D. Lewis Philosophical Papers*. Oxford: Oxford University Press, 130-132.

Madden, James (2013). *Mind, Matter and Nature A thomistic Proposal for the Philosophy of Mind*. Washington: The Catholic University of America Press.

Manzotti, Riccardo (2007). "From Artificial Intelligence to Artificial Consciousness". *Artificial Consciousness* (ed. Antonio Chella ve Riccardo Manzotti). Exeter: Imprint Academic, 174-190.

Manzotti, Riccardo ve Vincenzo Tagliasco (2008). "Artificial Consciousness: A Discipline Between Technological and Theoretical Obstacles". *Artificial Intelligence in Medicine* (44): 105-117.

Manzotti, Riccardo ve Tim Parks (2018). *Zihnin Ucu Bucağı*: *Bilinç ve Dünya Bir midir?* (Çeviren: Özde Duygu Gürkan). İstanbul: Metis Yayınları.

Maslin, Keith T. (2007). *An Introduction to the Philosophy of Mind*. Cambridge: Polity Press.

McGinn, Colin (1989). "Can We Solve the Mind-Body Problem?", *Mind* (98): 349-366.

McGinn, Colin (1991). *The Problem of Consciousness: Essays Toward a Resolution*. Oxford: Blackwell.

McGinn, Colin (2004). *Consciousness and Its Objects*. New York: Oxford University Press.

Metzinger, Thomas (2009). *The Ego Tunnel The Science of the Mind and The Myth of the Self*. New York: Basic Books.

Metzinger, Thomas (2017). "Saydam Bir Benlik Modelinin İçeriğiyim". *Bilinç Üzerine Konuşmalar* (ed. Susan Blackmore). İstanbul: Küre Yayınları. 183-194.

Mill, John Stuart (2017). *Faydacılık*. İstanbul: Alfa Yayıncılık.

Minsky, Marvin (1954). "Neural Nets and the Brain Model Problem", Doktora Tezi. Princeton University.

Minsky, Marvin (2006). *The Emotion Machine Commonsense Thinking, Artificial Intelligence, and the Future of the Human Mind*. New York: Simon ve Schuster.

Moor, James (2006). "The Nature, Importance and Difficulty of Machine Ethics". *IEEE Intelligent Systems*. 21 (4): 18-21.

Nagel, Thomas (1974). "What is it like to be a Bat?". *The Philosophical Review*. Duke University Press (83): 435-450.

Nagel, Thomas (2015). *Zihin ve Evren*. İstanbul: Jaguar Kitap.

Nath, Rajakishore ve Vineet Sahu (2017). "The Problem of Machine Ethics in Artificial Intelligence". *AI ve Society*. 1-9.

Neisser, Ulric (1988). "Five Kinds of Self-Knowledge". *Philosophical Psychology* (1): 35-59.

Nelson, Andrew L. (2013). "Artificial Life and Machine Consciousness". *AAAI Technical Report FS-13-02*. 52-57.

Nemes, Tihamêr (1969). *Cybernetic Machines*. Budapest: Illiffe Books ve Akademiai Kiado.

Nillson, Nils J. (2018). *Yapay Zeka Geçmişi ve Geleceği* (Çeviren: Mehmet Doğan). İstanbul: Boğaziçi Üniversitesi Yayınları.

Noë, Alva (2007). “Inattentional Blindness, Change Blindness, and Consciousness”. *The Blackwell Companion to Consciousness* (ed. Max Velmans ve Susan Schneider). Oxford: Blackwell Publishing. 504-511.

O’Regan, Kevin J. ve Alva Noë (2001). “Sensorimotor Account of Vision and Visual Consciousness”. *Behavioral and Brain Sciences* (24): 939-1031.

O’Regan, Kevin J. (2012). “How to Build a Robot that is Conscious and Feels”. *Minds and Machines* (22): 117-136.

Öner, Necati (1991). *Klasik Mantık*. Ankara: AÜİF Yayınları.

Özakpınar, Yılmaz (2011). *Psikoloji Tarihi*. İstanbul: Ötüken Yayınları.

Özlem, Doğan (2004). *Etik Ahlak Felsefesi*. İstanbul: İnkilap Kitabevi.

Öztemel, Ercan (2006). *Yapay Sinir Ağları*. İstanbul: Papatya Yayıncılık.

Parthemore, Joel ve Blay Whitby (2013). “What Makes Any Agent a Moral Agent?: Reflections on Machine Consciousness and Moral Agency”. *International Journal of Machine Consciousness* 5(2): 105-129.

Pazarlı, Osman (1980). *İslâm’da Ahlak*. İstanbul: Remzi Kitabevi.

Picard, Rosalind W. (1997). *Affective Computing*. Cambridge: MIT Press.

Pitts, Walter ve Warreb McCulloch (1943). “A Logical Calculus of Ideas Immanent in Nervous Activity”. *Bulletin of Mathematical Biophysics* (5): 115-133.

Platon (1997). *Phaidon*. İstanbul: MEB Yayınları.

Polak, Michal ve Tomas Marvan (2018). “Neural Correlates of Consciousness Meet the Theory of Identity”. *Frontiers in Psychology* (9): 1-13.

Rakova, Marina (2007). *Philosophy of Mind A-Z*. Edinburg University Press.

Reggia, James A. (2013). “The Rise of Machine Concsiousness”. *Neural Network* (44): 12-31.

Reichenbach, Hans (1939). *Lojistik* (Çeviren: Vehbi Eralp). İstanbul: Türkiye Basımevi.

Revonsuo, Antti (2000). "Prospects for a Scientific Research Program on Consciousness". *Neural Correlates of Consciousness Empirical and Conceptual Questions* (ed. Thomas Metzinger). Cambridge: MIT Press, 56-75.

Revonsuo, Antti (2017). *Bilinç Öznelliğin Bilimi* (Çeviren: Selim Değirmenci). İstanbul: Küre Yayınları.

Rose, Hilary ve Steven Rose (2018). *Nörobilim Zihnimizi Değiştirebilir mi?* (Çeviren: Meriç Öztürk). İstanbul: Ginko Bilim.

Rosenthal, David M. (1986). "Two Concepts of Consiousness". *Philosophical Studies* (49): 329-359.

Rosenthal, David M. (2004). "Varieties of Higher Order Theory". *Higher-Order Theories of Consciousness An Anthology* (ed. Rocco J. Gennaro). Amsterdam: John Benjamins Publishing Company.

Rosenthal, David M. (2008). "Consciousness and Its Function". *Neuropsyhologia* (46): 829-840.

Roth, Gerhard (2000). "The Evolution and Ontogeny of Consciousness". *Neural Correlates of Consciousness Empirical and Conceptual Questions* (ed. Thomas Metzinger). Cambridge: MIT Press, 77-97.

Rowland, Mark (2007). "Mysterianism". *The Blackwell Companion to Consciousness* (ed. Max Velmans ve Susan Schneider). Oxford: Blackwell Publishing, 335-345.

Say, Cem (2018). *50 Soruda Yapay Zeka*. İstanbul: Bilim ve Gelecek Kitaplığı.

Sayan, Erdinç (2012). "Analitik Zihin Felsefesinin Temel Problemlerine Bir Bakış". *Kaygı* (19): 37-54.

Scasselatti, Brian (2002). "Theory of Mind for a Humanoid Robot". *Autonomous Robots* (12): 13-24.

Seager, William ve David Bourget (2007). "Representationalism about Consciousness". *The Blackwell Companion to Consciousness* (ed. Max Velmans ve Susan Schneider). Oxford: Blackwell Publishing, 261-276.

Searle, John R. (1980). "Minds, Brains and Programs". *Behavioral and Brain Sciences* 3(3): 417-457.

Searle, John R. (1990). "Is the Brain's Mind a Computer Program?. *Scientific American* (262): 25-31.

Searle, John R. (2000). "Consciousness". *Annual Review of Neuroscience* (23): 557-578.

Searle, John R. (2004). *Zihnin Yeniden Keşfi* (Çeviren: Muhittin Macit). İstanbul: Litera Yayıncılık

Searle, John R. (2005). *Bilinç ve Dil* (Çeviren: Muhittin Macit ve Cüneyt Özpilavcı). İstanbul: Litera Yayıncılık.

Searle, John R. (2006). *Zihin, Dil ve Toplum* (Çeviren: Alaattin Tural). İstanbul: Litera Yayıncılık.

Searle, John R. (2018). *Bilincin Gizemi* (Çeviren: İlknur Karagöz İçyüz). İstanbul: Küre Yayınları.

Shoemaker, Sydney (1996): *The First-Person Perspective and Other Essays.* Cambridge: Cambridge University Press.

Silver, David; Hubert, T.; Schrittwieser, J.; Antonoglou, I.; Lai, M.; Guez, A.; Lanctot, M.; Sifre, L.; Kumaran, D.; Graepel, T.; Lillicrap, T. P.; Simonyan, K. ve D. Hassabis (2017). *Mastering Chess and Shogi by Self-Play with a General Reinforcement Learning Algorithm*. CoRR, abs/1712.01815.

Singer, Peter W. (2015). *Robotik Savaş: 21. Yüzyıldaki Robotik Devrim*. Ankara: Buzdağı.

Skansi, Sandro (2018). *Introduction to Deep Learning From Logical Calculus to Artificial Intelligence*. Basel: Springer.

Smart, J.J.C. (2004). "Sensations and Brain Processes". *Philosophy of Mind A Guide and an Anthology* (ed. John Heil). New York: Oxford University Press, 116-127.

Starzyk, Janusz ve Dilip Prasad (2011). "A Computational model of Machine Consciousness". *International Journal of Machine Consciousness*. 2(1): 1-26.

Steels Luc ve Rodney A. Brooks (1995). *The Artificial Life Route to Artificial Intelligence*: *Building Embodied, Situated Agents*. Psychology Press.

Strawson, Galen (1999). "Self, Body and Experience". *Proceedings of the Aristotelian Society* (73): 307-332.

Strawson, Galen (2006). "Realistic Monism: Why Physicalism Entails Panpsychism". *Journal of Consciousness Studies* (13): 3-31.

Sun, Ron (1997). "Learning, Action and Consciousness: A Hybrid Approach toward Modelling Consciousness". *Neural Networks* 10 (7): 1317-1331.

Tegmark, Max (2019). *Yaşam 3.0 Yapay Zeka Çağında İnsan Olmak* (Çeviren: Ekin Can Göksay). İstanbul: Pegasus Yayınları.

Tononi, Giulio (2004). "An Information Integration Theory of Consciousness". *BMC Neuroscience*, 5(42): 1-22.

Tononi, Giulio ve Gerard Edelman (2000). *A Universe of Consciousness How Matter Becomes Imagination*. London: Allen Lane.

Tononi, Giulio ve Christof Koch (2015). "Consciousness: Here, There and Everywhere?". *Philosophical Transactions of the Royal Society B Biological Sciences* 370(1668) 19 May 2015.

Topdemir, H. Gazi ve Yavuz Unat (2018). *Bilim Tarihi*. Ankara: Pegem Akademi.

Torrance, Steve (2008). "Ethics and Consciousness in Artificial Agents". *AI&Soc* (22): 495-521.

Trautteur, Giuseppe (1991). "Is Artificial Consciousness Possible?". *Consciousness*: *Distinction and Reflection* (ed. L. Steels). Napoli: Bibliopolis, 42-52.

Tura, Saffet Murat (2016). *Beynin Gölgeleri Bir Psikiyatri Felsefesi*. İstanbul: Metis Yayınları.

Tura, Saffet Murat (2017). *Zor Problem*: *Bilinç Bilinç Nörobiyolojisinin Fenomenal Dünya Yorumu*. İstanbul: Metis Yayınları.

Turing, Allen (1936). "On Computable Numbers, Wıth an Applıcation to the Entscheidungsproblem". *Proceedings of the London Mathematical Society* (42): 230-265.

Turing, Allen (1950). "Computing Machinery and Intelligence". *Mind* (49): 433-460.

Turing, Allen (2004). "Computing Machinery and Intelligence". *The Essential Turing*. Oxford: Oxford University Press.

Varela, Francisco J. (1996). "Neurophenomenology A Methodological Remedy for the Hard Problem". *Journal of Consciousness Studies* 4(3): 330-349.

Velmans, Max (2002). "Making Sense of Causal Interaction Between Consciousness and Brain". *Journal of Consciousness Studies* 9(11): 69-95.

Velmans, Max (2009a). "How to Define Consciousness and How not to Define Consciousness". *Journal of Consciousness Studies* 16(5): 139-156.

Velmans, Max (2009b). *Understanding Consciousness.* New York: Routledge.

Velmans, Max (2017). "Evren Senin Dolayımından ve Benim Dolayımımdan Kendine Dair Farklı Görünüşlere Sahip". *Bilinç Üzerine Konuşmalar* (ed. Susan Blackmore). İstanbul: Küre Yayınları. 279-291.

Vural, Mehmet (2002). "Düşünce Tarihinde Mantık: Aristoteles Mantığından Bulanık Mantığa". *Kutadgubilig Felsefe-Bilim Araştırmaları Dergisi* (2): 179-192.

Wallach, Wendell, Colin Allen ve Stan Franklin (2011). "Consciousness and Ethics: Artificially Conscious Moral Agents". *International Journal of Machine Consciousness* (3): 177-192.

Wheeler, Michael (2008). "Cognition in Context: Phenomenology, Situated Robotics and the Frame Problem". *International Journal of Philosophical Studies* 16(3): 323-349.

William, David Ross (1930). The Right and The Good. Oxford: Oxford University Press.

Yılmaz Atınç (2017). *Yapay Zeka*. İstanbul: Kodlab.

Young R. A ve S. Torrance (1999). "The Mentality of Robots". *Proceedings of the Aristotelian Society*. Supplementary Volumes (68): 199-227, 229-262.

Zadeh, Lütfi A. (2007). "Puslu Mantığın Doğuşu ve Evrimi". *Kutadgubilig* (Çeviren: Yücel Yüksel). (12): 173-184.

Zeman, Adam (2006). *Bilinç, Kullanım Kılavuzu*. İstanbul: Metis Yayınları.

Zeman, Adam, Dewar, M. ve S. Della Sala (2015). Lives without Imagery - Congenital Aphantasia. *Cortex*. (73), December: 378–80.

9. EKLER

Ek: 1: ASIMO

Asimo'nun ilk modelleri olan E1, E2 ve E3 insan yürümesini taklit eden bacaklardan oluşuyordu. Ancak bu modellerin hem hızları çok yavaştı (bir adımı 10 saniyede atabiliyordu), hem de yalnızca tek bir doğrultu üzerinde ilerleyebiliyordu. Daha sonra geliştirilen E4, E5 ve E6'da yürüme hızlandı, merdiven çıkabilme ve geriye gidebilme özellikleri eklendi. Daha sonra ise bu modele, kollar, kafa ve vücut eklenerek ilk insansı robot olan P1 oluşturuldu. P1'den geliştirilen P2 modeline kablosuz otomatik hareketler eklendi. Son olarak da 160 cm boyunda 130 kg olarak tasarlanan P3, ASIMO'nun bugünkü görüntüsünü ortaya çıkardı. ASIMO'nun bu görüntüsüne[69] ulaşması en az 20 yıllık mühendislik çalışmalarının sonunda mümkün oldu. ASIMO bugün koşabiliyor, engebeli arazilerde yürüyebiliyor ve dönüş yapabiliyor. Sesli komutları anlayarak cevap verebiliyor ve çevresindeki bireylerin hareket hâlinde olsalar bile –eğer önceden ona tanıtılmışsa- yüzlerini tanıyabiliyor. Kamera şeklindeki gözlerini kullanarak çevreyi ve çevresindeki nesneleri tanıyarak hafızasına kaydedebiliyor. Etraftan gelen tehlike ve engellerden kaçabiliyor; çevreye adapte olabiliyor.[70]

[69] Bkz. Şekil 4.

[70] ASIMO'ya ait genel özelliklere https://asimo.honda.com/asimo-history/ ve https://asimo.honda.com/downloads/pdf/asimo-technical-information.pdf adreslerinden ulaşılmıştır.



Şekil 6: ASIMO

Ek: 2: COG

Cog, "insansı zeka, dünyayla insansı etkileşim gerektirir." motivasyonu ile yola çıkılmış bir projedir. Bedensel olarak insani benzerliği yüksek olan Cog, yapay zekanın birçok alt dalını bütün bir şekilde uyumlu ve işlevsel olarak bir araya getiren tek bir donanımdan meydana gelmektedir. Cog robotu, insan bedeninin duyu-motor dinamikleri örnek alınarak oluşturulmuş sensör ve işlemcilerden oluşan bir bedene sahiptir. Bacakları ve esnek bir omurgası yoktur ancak baş, gövde ve kollarının motor yapısındaki özgürlük ve esneklik insandaki ile yakın seviyededir. Cog, video kameralar yoluyla görebilmektedir; duyma ve dokunma yetenekleri ise sorunsuz değilse de sürekli geliştirilen bir taslak üzerinden ilerlemektedir. Cog'un denge sistemi de tamamlanmak üzeredir. Bunun dışında, çevredeki ortak tutum ve eğilimi tespit edebilen oldukça insana yakın bir algı sistemine sahiptir. Cog'un göz hareketleri insani seviyeye yakındır; nesneleri takip edebilir, konuşurken odaklanır ve göz kırpabilir. Ayrıca, tüm davranışlarına boyun ve kafa hareketlerini de ekler, gövdesi ve baş hareketleri birbiriyle uyumludur. Başını sallayarak onay verme veya hayır deme özelliği onun yine insansı niteliklerindendir. Sahip olduğu görsel tarayıcılar ile hareket ve rengi algılayabilen Cog, görsel hedefine ulaşabilir; bir nesneye yönelerek, tutup kavrayabilir. Nesneyi tanıma ve tutup kavrama yetisi Cog'a yüklenmiş bir yeti değildir, deneme-yanıma yoluyla öğrenerek geliştirir. Bunun dışında Cog,

testereyle kesebilme, stres yayıyla oynama, bateri çalma ve sarkaç çevirme gibi el ve kollarını çok iyi kullandığını gösteren yeteneklere de sahiptir.



Şekil 7: *Cog*

İnsan zekası seviyesinde bir yapay zeka üretmek için, insansı bir robot inşa etmeye çalışmak üzerine farklı gerekçeler sunulmaktadır. Bu gerekçelerden bir tanesi, hem içsel düşüncelerimiz hem de dil yeteneğimiz için bedenin sahip olduğu önemli role vurgu yapmaktadır. Bu gerekçelendirme "Eğer insana benzer bir zeka üretilmek isteniyorsa, insanla benzer temsillere sahip olması için insansı bir beden de üretilmelidir." fikri üzerine yapılandırılmaktadır. Ancak insan bedeni görülen ve görülemeyen milyonlarca küçük ayrıntıya sahiptir ve insan bedenini bütün ayrıntılarıyla yeniden üretmek mümkün görünmemektedir. Dolayısıyla insansı robot uğraşlarını bekleyen en büyük tehlike, yüksek bir olasılıkla insan bedeninin ayrıntılarda gizli olan temel özelliklerinden yoksun olacak olmalarıdır. İnsan bedeninin ham hâldeki, çok geniş ve basit taslağını taklit ederek ve içsel temel nitelikleri yakalayamadan üretilen robotlar için harcanan emek "sahte bir bilim" yapmak için harcanmış olabilir (Scasselatti, 2002: 10).

İnsansı robot üretimi için öne sürülen diğer bir gerekçe, insan zekasının ancak insan bedeninin yeniden üretimiyle taklit edilebileceği fikrinden daha sağlam temellere dayanmaktadır. Bu temel, insanın ana niteliklerinden bir tanesi olan çevreyle ve diğer insanlarla etkileşim hâlinde olmasıdır. İnsan bu etkileşimler ile kendi bilincini ve "ben"ini oluşturmaktadır. İnsan seviyesinde bir robot da ancak çevresiyle kurduğu etkileşimlerle

deneyim sahibi olan bir robot ile mümkün olabilir (Scasselatti, 2002: 13). Robotun insani bir form ve şekle sahip olması, insanlarla etkileşimini daha doğal ve kolay hâle getirir. İnsansı robotun davranış ve tepkilerini gören insanlar, robotu sosyal ortamlarında çok daha kolay kabullenmektedir. Ayrıca sosyal etkileşim, robotun öğrenme sürecini de hızlandırmakta ve robotun fiziksel ve zihinsel gelişimine büyük katkı sağlamaktadır.

Cog projesi insan zekasının gelişim, organizasyon, bedenlenmişlik ve çok-modlu entegrasyon özelliklerini temel alarak sürdürülmektedir (Adams, 2000: 26). Bir bebek, yetişkin bir insanın sahip olduğu akıl yürütme sistemi, motor sistemi ve hatta duyu sistemine sahip değildir. İnsan doğumundan yetişkinliğine kadar karşılaştığı karmaşık çevresel koşullar ile birlikte hem akıl-yürütme hem de duyu-motor özellikleri bakımından sürekli bir gelişim süreci içerisindedir. Gelişim, kademesel olarak ilerleyen bir süreçtir. Öğrenilmekte olan davranış ve yetenekler, insanın daha gelişmiş davranış ve yeteneklere sahip olması için birer hazırlık sürecidir. Cog'un öğrenme yapısı da bu nedenle, insanla benzer şekilde oluşturulmaya çalışılmıştır. Bir bebeğin oyuncağa ulaşmak için gösterdiği görsel ve motorsal gelişim sürecinin benzeri Cog'un oyuncağa ulaşma sürecinde de kademesel olarak gözlemlenmektedir (Adams, 2000: 28).

Bebekler ve çocuklar ebeveynleri ve bakıcılarına sadece temel ihtiyaçlarının karşılanması açısından değil, duyusal, zihinsel ve sosyal gelişimleri açısından da ileri derecede bağımlıdırlar. Sosyal temas ve etkileşim, insan türüyle öyle iç içedir ki, tamamen asosyal olan, hiçbir sosyal etkileşimde bulunmamış bir insan tahayyül etmek oldukça zordur. Dolayısıyla insan benzeri bir robotun da sosyal etkileşimde bulunuyor olması zorunlu bir özellik olarak ele alınmalıdır. Öğrenmeye olanak tanıyan, öğrenmeyi hızlandıran ve insan bilincini oluşturan önemli etkenlerden bir tanesi olarak sosyal etkileşim, yapay bilinç için de zorunlu bir nitelik olarak kabul edilmelidir.

İnsanlar dış dünyadan sınırsız bilgi alma yetisine sahiptirler. İnsan, görme, işitme, koku alma, dokunma ve tat alma duyuları ile kendi dünya resmini oluşturur. Fakat bu duyular birbirinden bağımsız değillerdir, her duyu birbiri ile karmaşık bir ilişki içerisindedir. Sesini duyduğunuz bir nesne gözünüzün önüne gelebilir ya da kokusunu aldığınız kurabiye size annenize sarıldığınızı hissettirebilir. Duyusal modüller birbirlerini etkilerler; bu sebeple insansı robotlarda kurulacak duyusal sistem de bütünleşik bir şekilde planlanmalıdır. Cog tasarlanırken de bu sebeple çok-modüllü bütünleşik bir yapı olarak düşünülmüştür.

EK:3: ÖZGEÇMİŞ

KİŞSEL BİLGİLER	
Adı Soyadı	: Mehtap Doğan
Doğum tarihi	: 12.08.1986
Doğum yeri	: Eskişehir
Medeni hali	: Bekar
Uyruğu	: T.C
Adres	: Ankara Yıldırım Beyazıt Üniversitesi İnsan ve Toplum Bilimleri Fakültesi, Felsefe Anabilim Dalı, Ankara
Tel	: 0312 906 13 98
E-mail	: mdogan@ybu.edu.tr
EĞİTİM	
Lisans	İhsan Doğramacı Bilkent Üniversitesi İnsani Bilimler ve Edebiyat Fakültesi Felsefe Bölümü/2009
Yüksek lisans	Ankara Yıldırım Beyazıt Üniversitesi İnsan ve Toplum Bilimleri Fakültesi, Felsefe Anabilim Dalı/2014
Doktora	Ankara Yıldırım Beyazıt Üniversitesi İnsan ve Toplum Bilimleri Fakültesi, Felsefe Anabilim Dalı/2020