

LAURI JÄRVILEHTO

KONE-

ONKO  
AJATTELU  
IHMISEN  
YKSI-  
OIKEUS?

MIELI

TAMMI

LAURI JÄRVILEHTO

# KONE- MIELI

ONKO  
AJATTELU  
IHMISEN  
YKSIN-  
OIKEUS?



TAMMI

HELSINKI



Ensimmäinen painos

© Lauri Järvilehto ja Tammi 2025

Tammi on osa Werner Söderström Osakeyhtiötä

ISBN 978-952-04-6357-1

Painettu EU:ssa

Tuoteturvallisuusasioihin liittyvät tiedustelut: [tuotevastuu@tammi.fi](mailto:tuotevastuu@tammi.fi)

# SISÄLLYS

<b>JOHDANTO</b> .....	7
<b>ENSIMMÄINEN OSA: IHMISMIELI</b> .....	11
<b>1. MIKÄ MIELI OIKEASTI ON?</b> .....	13
Mielen teorioista.....	15
Mitä mieli tarkoittaa? .....	20
Minuus: Missä mieli sijaitsee? .....	29
Tietoisuus ja tietoinen mieli .....	36
Tiedostamaton mieli: systeemi 1 ja systeemi 2 .....	50
Ympäristö mielen osana .....	58
<b>2. AJATTELU JA AIVOTOIMINTA</b> .....	66
Ajattelun ja aivojen rakenteelliset yhteydet.....	67
Hermoverkkojen synty ja neuroplastisuus.....	73
Ovatko aivot tietokone? .....	77
<b>3. MITÄ AJATTELU ON?</b> .....	87
Mielen ja maailman raja .....	88
Käsitteet ja uskomukset.....	93
Ajattelu ja meta-ajattelu.....	111

<b>TOINEN OSA: KONEMIELI</b> .....	115
<b>4. KONEMIELEN SYNTY</b> .....	117
<b>Tietokone ja informaation prosessointi</b> .....	120
<b>Tekoälyn aikakaudet</b> .....	127
<b>Uudet tekoälyratkaisut</b> .....	134
<b>5. KONEMIELI</b> .....	145
<b>Miten ChatGPT toimii</b> .....	148
<b>Tekoäly ja ajattelu</b> .....	157
<b>Todellisuus, malli ja simulaatio</b> .....	165
<b>Tekoäly tiedostamattoman simulaationa</b> .....	171
<b>6. TULEVAISUUDEN NÄKYMIÄ</b> .....	181
<b>Tekoäly lähitulevaisuudessa</b> .....	182
<b>Tekoäly mielen jatkeena</b> .....	189
<b>Miten tekoälyä käytetään?</b> .....	199
<b>Ajatteleva kone yhteiskunnassa</b> .....	207
<b>LOPUKSI</b> .....	211
<b>KIITOKSET</b> .....	215
<b>VIITTEET</b> .....	217
<b>LÄHTEET</b> .....	227

#### **Huomautus lukijalle**

Suurin osa kirjan loppuviitteistä on lähdeviitteitä. Osa viitteistä sisältää kuitenkin tarkennuksia ja teknisiä huomioita. Nämä on merkitty lihavoidulla viitenumerolla.

# JOHDANTO

Wikipedia määritteli muutamia tämän kirjan avaintermejä heinäkuussa 2024 seuraavasti:

**Ihminen** (*Homo sapiens*) on ainoa nykyisin elossa oleva ihmisten (*Homo*) sukuun kuuluva nisäkäs.<sup>1</sup>

**Kone** on laite, jonka avulla voidaan aikaansaada energian muodon, vaikutustavan, -suunnan tai -kohdan säännöllisiä muutoksia.<sup>2</sup>

**Mieli** ymmärretään yksilön sisäisenä, henkisenä olemuksena, jossa ajatukset, tunteet ja tietoisuus ilmenevät.<sup>3</sup>

Vuosituhansia olemme pohtineet, mikä ihmismieli oikeasti on ja miten se toimii. Vaikka ajattelu on saumaton osa arkipäiväämme, on se myös yhä yksi suurimmista ihmiskunnan mysteereistä. Miten on ylipäättään mahdollista, että tällaisesta solukimpusta voi kummuta kokonainen maailmankaikkeus muistoinen, arvoineen, käsityksineen ja tulevaisuudenhaaveineen?

Tiedämme ihmisielen toiminnasta jatkuvasti enemmän. Etenkin viime vuosikymmeninä yleistyneet ajattelun kaksois-prosessointiteoriat esittävät, että ihmismieli koostuu kahdesta erilaisesta järjestelmästä: tietoisesta systeemi 2:sta ja tiedostamattomasta systeemi 1:stä. Näihin ajattelua kuvaaviin teorioihin perustuvat mallit kykenevät ennakoimaan hämmästyttävän

hyvin ihmisen käyttäytymistä ja päätöksentekoa. Samaiset teoriat selittävät myös monia mielemme arkisia ilmiöitä, kuten vaikkapa sen, mistä ideat pulpahtavat päähän tai miksi saatamme herätä keskellä yötä pohtimaan jotain mitätöntä työasiaa.

Systeeminen psykologia puolestaan tarjoaa kiehtovan näkökulman ikiaikaisiin filosofisiin kysymyksiin siitä, mitä tietoisuus on tai miksi meillä ylipäätään on mieli. Kaksoisprosessointiteoriat ja systeeminen psykologia tarjoavat yhdessä systemaattisen näkökulman ihmismielen moninaisuuteen, josta tietoinen toiminta on vain yksi osa, kun taas mielen kokonaistoiminta on laajempaa, ja siihen kuuluu myös tiedostamattomia prosesseja ja ympäristöön ulottuvia vuorovaikutussilmukoita.

Olen tutkinut ihmismielen toimintaa yli 15 vuotta. Jos joku olisi kysynyt minulta keväällä 2022, milloin nähdään ihmisen kanssa kilpailukykyisiä tekoälyratkaisuja, jotka tuottavat tekstiä, kuvia ja musiikkia, olisi arvioni tuolloin ollut luultavasti, että niitä nähdään ehkä kolmenkymmenen vuoden päästä, ehkä ei koskaan. Kun sitten syksyllä 2022 pääsin ensimmäistä kertaa käyttämään Midjourneyta ja ChatGPT:tä, ensireaktioni oli, että tämä on mahdotonta. Oli selvää, että joko käsitykseni ihmismielestä tai käsitykseni teknologian mahdollisuuksista oli ollut puutteellinen. Mielessäni syttyi palava halu ymmärtää, mikä yhdistää ja mikä erottaa ihmismieltä ja nyt tarjolla olevia tekoälyratkaisuja.

Tekoälyn kehittäminen on ollut viime vuosikymmeninä kiistelty aihe. Jotkut ovat julistaneet jo 1950-luvulta alkaen, että meillä on pian ajattelevia koneita. Toiset taas esittävät, ettei todellista tekoälyä voida synnyttää koskaan. Jälkimmäiset väittävät, että ihmisen mielessä on jotain niin ainutlaatuista, ettei sitä voida koskaan teknologian keinoin jäljitellä. Tätä taustaa vasten syksyllä 2022 alkanut uutta luovan eli generatiivisen tekoälyn vallankumous onkin ollut suorastaan häkellyttävä.

Uudet tekoälyratkaisut pystyvät synnyttämään lähes samantasoista kieltä kuin ihminen. Ne voivat luoda kuvia, musiikkia

ja videoita. Ensimmäistä kertaa ihmiskunnan historiassa koneen kanssa on mahdollista keskustella kuin toisen ihmisen kanssa. Moni tekoälyn ja mielen tutkijoiden johtohahmo oli odottanut tällaisia läpimurtoja vasta kymmenien vuosien kuluttua – jos ne olisivat mahdollisia ylipäätään.

Nyt ne ovat totta.

Uusien tekoälyratkaisujen taustalla vaikuttavat teknologiat, erityisesti ChatGPT:n ja muiden uusien tekoälybottien taustalla toimivat suuret kielimallit, kertovat meille varmuudella ainakin sen, että tällaisella teknologialla voidaan synnyttää ihmismäistä kieltä – ja kenties siis myös ihmismäistä ajattelua. Samalla nämä teknologiat raottavat kantta aivan uuteen ymmärrykseen siitä, minkälaiset toimintaperiaatteet vallitsevat meidän oman mieleemme taustalla.

Monissa mielen nykyteorioissa on tunnistettu, että ajattelu nojaa erilaisiin käsitteellisiin assosiaatorakenteisiin. Suuret kielimallit perustuvat puolestaan erittäin laajoihin ja monisyisiin sanojen keskinäisiin tilastollisiin yhteyksiin. Kun kielimallien toimintaa tarkastellaan yhtäältä viime vuosien mielen tutkimuksen, nykyaikaisen aivotutkimuksen ja myös tiettyjen perinteisempien mielenfilosofisten teorioiden valossa, syntyy hämmästyttävä kuva siitä, miten mieli oikeasti toimii.

Ihmismielen tutkimus on auttanut keksimään menetelmät, joilla on synnytetty ihmismäiset kielimallit. Niiden tutkimus voi puolestaan auttaa ymmärtämään paremmin, miten ihmisen kieli ja mieli toimivat. Jos nyt keksityt teknologiat mahdollistavat lähes ihmismäisen kielen synnyn, voivatko ne mahdollistaa myös keinotekoisen ajattelun ja sitä kautta keinotekoisien mielen – todellisen tekoälyn, konemielen – synnyn?

Mitä tämä kertoo meille meistä itsestämme? Mitä se kertoo meille keinotekoisien ajattelun mahdollisuudesta? Olemmeko astumassa aikakauten, jossa ihmisen kanssa planeetta Maata asuttaa pian toinen älykäs elämänmuoto?

Vai onko tämä todellisuutta jo nyt?



ENSIMMÄINEN OSA



# IHMISMIELI

# 1. MIKÄ MIELI OIKEASTI ON?

Olipa kerran ilkeä tiedemies Jank Frackson, joka lukitsi vastasyntyneen tyttärensä Maryn suljettuun huoneeseen. Marysta piti huolta laaja huoltajajoukko, ja hän eli innostavan ja antoisan lapsuuden. Huomionarvoista oli kolme asiaa.

Ensinnäkin Marya koulutettiin systemaattisesti tieteentekijäksi, ja hänestä varttuikin jo teini-ikäisenä näköhavainnon neurofysiologiaan erikoistunut huipputieteilijä. Toiseksi 20-vuotispäiväänsä asti Mary ei saanut poistua huoneestaan. Hänelle tuotiin huoneeseen kaikki tarvittava – ruoka, viihdykkeet, tieteelliset artikkelit ja niin edelleen. Kolmanneksi kaikki, mitä huoneessa oli, oli mustavalkoista. Seinät olivat valkoiset, huonekalut harmaan sävyissä. Maryn iho käsiteltiin aina öisin harmaanvalkoiseksi. Kaikki huoneessa käyneet avustajat olivat pukeutuneet tyystin mustavalkoisiin pukuihin. Huoneessa oli mustavalkoinen videonäyttö, jonka kautta Mary pystyi tarkastelemaan maailmaa. Sinne oli tuotu laaja valikoima tieteellisiä kirjoja, jotka oli kaikki käsitelty mustavalkoisiksi. Ja niin edelleen.

Kasvaessaan Mary oppi kaiken, mitä havaintopsykologiassa ja havainnon filosofiassa tiedetään näkemisestä. Hän oppi ymmärtämään, miten silmät ja aivot käsittelevät aistihavainnot. Hän ymmärsi, mikä valon aallonpituus vastaa sinistä ja mikä punaista ja minkälainen prosessi johtaa siitä, kun joku katsoo taivasta, siihen, että hän sanoo ”se on sininen”. Mutta vaikka hän sai huoneessa olevista videoruuduista ja loppumattoman paksuista kirjoista kaiken tarvitsemansa tiedon väreistä, hän ei ollut

koskaan nähnyt itse muita värejä kuin harmaan sävyjä. Maryn maailma oli mustavalkoinen.

Lähestyessään kahdenkymmenen vuoden ikää Mary oli noussut maailman johtavien havaintopsykologien joukkoon. Hänen tutkimuksiaan ja artikkeleitaan siteerattiin laajasti. Hänelle myönnettiin etänä tohtorintutkinto Arvovaltaisesta Yliopistosta. Pian Marylle tarjottiin myös professuuria Hienosta ja Tärkeästä Yliopistosta. Mary oli erityisesti erikoistunut värien havainnon tulkintaan, ja monia hänen teorioistaan pidettiin urauurtavina. Marya nimitettiin tiedepiireissä jopa ”väri-Maryksi”. Koko tiedeyhteisöstä ei löytynyt yhtään tutkijaa, jonka ymmärrys ja tieto koskien värejä olisi ollut lähellekään niin kaikenkattava kuin Maryn.

Lopulta Mary täytti 20, ja Jank Fracksonin kokeen viimeisen vaiheen oli määrä alkaa. Professori Frackson saapui huoneelle ja kertoi Marylle, että oli vihdoinkin tullut aika päästää hänet mustavalkoisesta huoneestaan. Pelonsekaisella innostuksella Mary seurasi professori Fracksonia ulos mustavalkoisesta huoneesta, pitkin valkoista käytävää, kohti mustia umpioivia, jotka olivat ylöspäin viettävän käytävän päässä.

Frackson pysähtyi ovien eteen.

”Mary”, hän sanoi, ”sinähän tiedät kaiken väreistä, eikö totta?”

”Kyllä, isä”, Mary sanoi. ”Ainakin enemmän kuin kukaan muu koko tieteellisessä yhteisössä.”

”Hieno juttu”, Frackson sanoi. ”Kerro minulle nyt, mitä näet, kun avaan ovet.”

Jank Frackson avasi oven ulkomaailmaan. Sisäänkäynti Maryn huoneeseen sijaitsi keskellä vehreää metsää, jossa maa tulvi kukkien kirjoja ja puissa lennähteli monenkirjavia lintuja. Kevätaurinko paistoi kirkkaana pilvettömältä sinitaivaalta. Mary astui ulos maailmaan silmiään siristäen.

Kahdenkymmenen elinvuotensa aikana Mary oli oppinut tietämään kaiken mahdollisen, mitä väreistä voi tietää.

Oppiko hän nyt jotain uutta?

Tarina on mukailtu filosofi Frank Jacksonin vuonna 1982 julkaiseman klassikkoartikkelin ”Epiphenomenal Qualia” pohjalta. Toisin kuin tarinan fiktiivinen ilkeä tiedemies, Frank Jackson on tiettävästi oikein mukava ihminen ja lukitsee ihmisiä huoneisiin vain teoreettisesti.

## MIELEN TEORIOISTA

Ymmärrämme yhä ihmismielen toimintaa hämmästyttävän vähän. Monet ihmismielen ilmiöt, kuten päätöksenteon epärationaalisuus, mieleen pulpahtavat oivallukset ja mietemadot sekä tietoisien ajattelun rajallisuus ovat saaneet selityksen viime vuosikymmeninä läpi lyöneissä teorioissa. Samaan aikaan monet perustavanlaatuiset kysymykset mielen luonteesta odottavat yhä selitystään. Mitä tietoisuus on? Miksi meillä on ylipäätään mieli? Miten mieli ja muu maailma liittyvät toisiinsa? Miksi asiat tuntuvat ylipäätään miltään?

Ymmärrämme yhä yksityiskohtaisemmin, mitä aivoissa tapahtuu kun ihminen ajattelee – mutta minkäänlaista kokeelliseen tutkimukseen perustuvaa yksimielisyyttä ei vallitse sen suhteen, miten ajattelu ja aivotoiminta viime kädessä suhteutuvat toisiinsa. Yhtäältä mittausmenetelmät eivät edelleenkaan riitä kovinkaan tarkkaan elävän olion aivotoiminnan tarkasteluun. Toisaalta yksikään teoreettinen malli ei selitä tunnettuja ilmiöitä niin tarkkaan, että voisimme varmuudella sanoa selitystavan olevan ainoa oikea.

Mitä se tarkoittaa, että ihmisellä on mieli? Onko mieli vain ihmisen ominaisuus, vai onko myös eläimillä mieli? Entäpä kasveilla? Jotkut filosofit ovat esittäneet, että termostaatilla on eräänlainen mieli, koska se pystyy muuttamaan käytöstään suhteessa ympäristöönsä.<sup>4</sup> Entäs muut koneet? Onko tietokoneella mieli? Miten on uusien tekoälyratkaisujen laita?

Jotta näissä kysymyksissä olisi ensinkään järkeä, pitää ensin luoda ymmärrys siitä, mitä mieli ja ajattelu ylipäätään

tarkoittavat. Kuten yleensäkin filosofian ja tieteen historiassa, tähän aika perustavanlaatuisen kysymykseen on olemassa lukematon määrä erilaisia vastauksia.

Usein filosofian ja psykologian kirjallisuudessa mielestä puhuttaessa tarkoitetaan tietoista mieltä – siis kielellistä ajattelua, havaintoja, tuntemista ja niin edelleen. Jo Platon oli vihjannut mielen kaksijakoisuudesta *Faidros*-dialogissaan, mutta vasta 1800-luvulla alkoi kunnolla kehittyä ajatus siitä, että mieli ei olisikaan yksinomaan tietoinen.<sup>5</sup> Ajatus ei-tietoisesta mielestä kypsyi muun muassa Hegelin, Schopenhauerin, Goethen ja Wordsworthin pohdinnoissa.<sup>6</sup>

1900-luvulla mielen tutkimuksessa tehtiin valtava määrä läpimurtoja. Logiikan ja matematiikan tutkimus piirsi rajat käsitteelliselle ajattelulle. Kokeellinen psykologia synnytti alati syvenevän ymmärryksen aivojen ja hermoston toiminnasta. Tietokoneen kehitys synnytti mielen tutkimuksen kanssa rinnakkaisia teorioita, jotka mahdollistivat myös monien mielen ominaisuuksien simuloinnin. Ja lopulta 1900-luvun lopulle tullessa erityisesti ajattelun kaksoisprosessointiteoriat tarjosivat ensimmäistä kertaa merkittävän selitysvoimaisen teoriajoukon, jonka valossa voitiin ymmärtää entistä paljon tarkemmin monia inhimillisiä mysteereitä, kuten päätöksentekoa, intuitiivista ajattelua, ajatusvinoumia ja tunnesäätelyä.

Karkeasti ottaen mielen teoriat voidaan jakaa kuuteen luokkaan:

- 1. Dualistiset mielen teoriat:** Näissä teorioissa mieli katsotaan jonkinlaiseksi aineettomaksi substanssiksi, joka vaikuttaa aineellisen substanssin – ruumiin, ympäristön – kanssa. Nämä teoriat vastaavat arkikokemukseen siitä, että mieli ja ruumis ovat jotenkin erilliset. Ne ovat yhä vahvasti läsnä kulttuurissamme, mutta tieteellisesti dualismi on osoittautunut kestäättömäksi. Jos mieli ei

siis ole henkiolento, avainkysymys on, missä se mieli oikein luuhaa.

- 2. Materialistiset mielen teorit:** Nämä teorit olivat 2000-luvun alkuun saakka tieteellisen mielentutkimuksen valtavirtateorioita, ja niillä on yhä paljon kannattajia. Materialismin ytimessä on ajatus siitä, että mieli on yhtä kuin aivot. Nykytiede on osoittanut kiistatta, että aivoissa tapahtuvilla muutoksilla on välitön vaikutus ajatteluun. Tietävästi kukaan ei myöskään ole toistaiseksi ajatellut ilman aivoja. Materialismin ongelmana on kuitenkin, että nämä teorit sivuuttavat kokonaan peruskysymyksen siitä, mitä itse mieli ja ajattelu ovat. Siis se, että jokin tuntuu vaikkapa punaiselta tai kivalta tai että päässä pyörii huomisen ostoslista kun pitäisi nukkua. Se, että aivot ovat ajatellessa aktiiviset, ei ole kovin kiistanalaista. Aivotoiminnan mittaukset voivat ennakoida erilaisia psyykkisiä tapahtumia. Mutta aivotoiminnan analyysit eivät vielä kerro mitään siitä, mitä itse ajattelu on.
- 3. Laskennalliset mielen teorit:** Materialismin ohella myös laskennalliset mielen teorit ovat saaneet valtavasti kannatusta etenkin matemaattisen eleganssinsa ja intuitiivisuutensa vuoksi. Näillä teorioilla on ollut valtava vaikutus tekoälyratkaisujen kehittämiseen. Ne ovat saaneet tukea sitä mukaa kuin aivotutkimusmenetelmät ovat kehittyneet ja tekoälyratkaisut tulleet yhä vaikuttavammiksi. Ja myönnettävä se on: onhan se aika intuitiivinen ajatus, että kun silmään osuu valonsäde, se laukaisee aivoissa prosessointiketjun, jonka perusteella aivot ”päättelevät”, että edessä tönöttää kissa. Laskennallisia mielen teorioita pidetään kuitenkin nykyään kahdella tapaa ongelmallisina. Ensinnäkin

kokeellinen tutkimus osoittaa, ettei aivoilla ole aikaa suorittaa tällaisia laskutoimituksia, vaan ajattelu tapahtuu usein salamannopeasti. Toiseksi vaikka laskennalliset mielen teorit ovatkin intuitiivisia ja kiehtovia, kukaan ei ole toistaiseksi kyennyt varsinaisesti osoittamaan, että aivot laskisivat yhtikäs mitään – muutoin kuin silloin kun ihminen tekee niiden avulla itse laskutoimituksia.

- 4. Mielen emergenssiteorit:** Monet nokkelikot ovat kiiruhtaneet ratkaisemaan materialismin perusongelmaa esittämällä, että psyykkiset ominaisuudet syntyvät jonkinlaisesta fyysisten ominaisuuksien vuorovaikutuksesta, mutta niin ettei psyykkistä voi palauttaa jäännöksettä fyysiseen. Nämä teorit ratkaisevat näppärästi materialismin ongelman. Seurauksena on kuitenkin eräänlainen kahden maailman ongelma. Emergenssiteorit flirttailevatkin dualismin kanssa siihen malliin, että on syytä kysyä, onko tämä ”emergoitunut” mieli sittenkin jonkinlainen ”rinnakkaissielu”? Se, että mieli ja aivot tanssivat samaan tahtiin, ei kuitenkaan varsinaisesti vielä selitä sitä, mihin koko mieltä ylipäättään tarvitaan. Eikö homma sitten hoituisi vain ihan puhtaasti aivotoiminnalla?
  
- 5. Modulaariset mielen teorit:** Modulaariset mielen teorit voidaan jakaa karkeasti kahteen joukkoon. Perinteisemmät mielen moduuliteorit, kuten Noam Chomskyn nativismi tai Jerry Fodorin mentaalisten moduulien teoria, eivät ole kestäneet tarkempaa empiiristä tarkastelua. Mieli ei ainakaan koostu selkeistä ja tarkkarajaisista moduuleista, jotka vastaisivat esimerkiksi kielellisestä tai visuaalisesta informaatiosta. Nykyaikaisiin modulaarisiin teorioihin kuuluu myös

kognitiivisen psykologian ylivoimaisesti menestyksellä mielen teoriakokonaisuus: ajattelun kaksoisprosessointiteoriat. Näiden teorioiden mukaan ihmisellä ei ole vain yhtä mieltä, vaan itse asiassa kaksi. Ajattelun kaksoisprosessointiteorioiden vahvuus on ennen kaikkea niiden laaja empiirinen ennustusvoima. Ne muun muassa selittävät sen, miksi ihminen käyttäytyy ajoittain tyhmästi vaikka ei haluaisi, sekä sen, mistä luovat oivallukset tulevat. Ajattelun kaksoisprosessointiteoriat pääsevät jo pitkälle mielen *toiminnan* kuvaamisessa. Mutta itse kysymys mielen luonteesta luiskahtaa lopulta niiltäkin pakoon.

- 6. Systemiset mielen teoriat:** Etenkin itämaisessä filosofiassa on jo vuosituhansia ollut vallalla ajatus siitä, ettei mieli ole vain yksilön yksityisomaisuutta, vaan itse asiassa mieli on jollakin tavalla jonkinlainen koko maailmankaikkeuden ilmentymä. Nämä teoriat eivät ole aiemmin olleet kovinkaan suosittuja tieteellisessä valtavirrassa, koska ne esittävät sellaisia perustavanlaatuisia kysymyksiä, joihin ei tieteen keinoin ole kerta kaikkiaan löydettävissä vastauksia. Jos sinun ajattelusi on universumin ajattelua, miten sitä edes päästäisiin tutkimaan tieteellisesti? Vaikka jonnekin Alfa Kentaurin tienoille saataisiinkin vietyä magneettikuvauslaite tai kuussa voitaisiin ampua magneettipulssia TMS-laitteella, ei se kertoisi meille vielä paljoakaan juuri sinun mielesi toiminnasta. On kuitenkin selvää, että ympäristöllä on välitön vaikutus siihen, mitä ajattele. Nykytutkimuksen valossa on myös selvää, että ympäristö muokkaa aivojen hermoratoja myös fysiologisesti – eli aivosi on piuhoitettu eri tavoin riippuen esimerkiksi siitä, elätkö hälyisässä kaupungissa vai metsän siimeksessä. Systemiset mielen teoriat tarjoavat myös



kiehtovan teoreettisen viitekehyksen sille, mikä mielen funktio on äkkiseltään fyysiseltä vaikuttavassa maailmassa.

Tämä jaottelu on osin keinotekoinen, ja moni mielen teoria kuuluu useampaan kuin yhteen kategoriaan. Tuonnempana esittelemäni malli ihmismielen toiminnasta hyödyntää erityisesti ajattelun kaksoisprosessointiteorioita ja systeemistä psykologiaa. Vastaavasti esittelemäni näkökulma on osin kriittinen materialistisia ja laskennallisia mielen teorioita kohtaan, dualismista puhumattakaan.

Kuitenkin konemielen pariin siirtyessämme näemme, että sekä materialismista että laskennallisista mielen teorioista on lähes paradoksaalisesti valtavasti iloa silloin kun yritämme jäljitellä ihmismielen toimintaa – siis myös ne kertovat kuitenkin jotain olennaista mielen toiminnasta. Siksi näitäkään teorioita ei voida ykskantaan sivuuttaa, jos haluamme luoda kattavan ja uskottavan näkemyksen ihmismielen toiminnasta. Koska yksikään yllä kuvatuista teorioista ei selitä ihmismielen toimintaa kokonaan ja jäännöksettä, on tärkeää pyrkiä muodostamaan yhtenäinen kuva siitä, mitä eri teorat lopulta kertovat meille mielen toiminnasta.

On hämmästyttävää, miten tämä meille ehkä kaikkein intiimein olemassaolomme osa on yhä niin valtava mysteeri. Sitä merkityksellisemmäksi kysymys käy nyt, kun olemme ensimmäistä kertaa todellisuudessa tekemisissä teknologian kanssa, jonka kohdalla on syytä pohtia kysymystä siitä, voiko myös koneella olla mieli.

## **MITÄ MIELI TARKOITTA?**

Jotta voimme ymmärtää, voiko ihmisellä, eläimellä tai koneella olla mieltä, meidän pitää ensin muodostaa käsitys siitä, mitä ylipäätään tarkoitamme mielellä. Mikä mieli oikeastaan on?

Mitä se tarkoittaa, että jollakin on mieli? Mistä tunnistamme, että jollakin oliolla on mieli – ja mistä tiedämme, että sitä ei ole?

Mielen luonnetta voidaan tarkastella monin eri tavoin. Kiinnostavimpia ovat mielen rakenteellinen ja toiminnallinen tarkastelu. Mistä mieli koostuu? Ja miten se toimii?

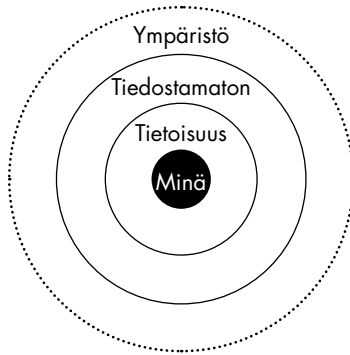
## **MIELEN OSAT**

Dualistiset teoriat ja monet uskonnot käsittelevät minuutta ruumiista erillisenä, jonkinlaisena sieluoliona. Tällaista oliota ei kuitenkaan ole toistaiseksi tieteen keinoin löydetty. Siksi dualismi tarjoaa heikot eväät konkreettisen tieteellisen tutkimuksen tekemiseen.

Materialistiset teoriat ja niitä venyttävät emergenssiteoriat puolestaan esittävät, että tietoisuus on joko jonkinlainen aivotoiminnan ilmentymä tai se ainakin seuraa aivotoiminnasta. Radikaaleimmat tutkijat ovat peräti esittäneet, että kyseessä on illuusio – ajattelu on aivotoimintaa ja sillä sipuli. Luulet vain olevasi tietoinen.<sup>7</sup>

Ajattelun kaksoisprosessointiteoriat jakavat mielen karkeasti tietoiseen ja tiedostamattomaan mieleen. Systemisen psykologian teoriat puolestaan esittävät, että myös ympäristöllä ja sosiaalisilla suhteilla on välitön vaikutus mielen syntyyn.

Mielen toiminnan kannalta keskeistä on myös selittää, mistä syntyy käsitys *itsestä* tai *minuudesta*. Lukuun ottamatta dualismia, jonka mukaan mieli on yhtä kuin sielu, muiden teorioiden täytyy pystyä selittämään, miksi aivoista tai niistä emergoituvista prosesseista tai kahden systeemin mallista, puhumattakaan holistisemmista systeemisistä malleista, seuraa käsitys siitä, että sinä olet juuri sinä.



*Mielen ytimessä on minuus, joka ohjaa tietoista tarkkaavaisuutta.  
Kaikki mielen toiminta nojaa tiedostamattomiin ja ympäristöön  
kytkeytyviin ajatusprosesseihin.*

Mielen keskeiset osat ovat siis minuus, tietoisuus ja ei-tietoinen. Neljäntenä mielen osana voidaan myös ainakin osin pitää myös ympäristöä. Ympäristö vaikuttaa suoraan kaikkeen siihen, mitä ajattelet. Kuten William James kirjoitti vuonna 1890 teoksessaan *Principles of Psychology*, nämä painetut kirjaimet muuttavat konkreettisesti sitä, miten ajattelet, ja voivat myöhemmin myös vaikuttaa esimerkiksi sanavalintoihisi keskustelutilanteessa.<sup>8</sup>

Tässä kirjassa käytän sekä ihmismielen että konemielen tarkastelussa yllä hahmoteltua mielen rakennetta. Keskiössä on minuus ja sen määrittävä tietoisuus. Tätä tukee puolestaan tiedostamaton käsitteellinen ja tiedollinen rakenne, joka perustuu käsitteiden ja uskomusten keskinäisiin suhteisiin ja joka muovautuu jatkuvasti kokemuksen ja ympäristön kanssa käytävän vuorovaikutuksen myötä. Mielen osat ovat siis minuus, tietoisuus, tiedostamaton ja ympäristö.

Tätä mallia ja sen ennustusvoimaa voidaan havainnollistaa esimerkiksi klassisella *Zeigarnik-efektillä*.<sup>9</sup> Latvialaisen psykologi Bluma Zeigarnikin mukaan nimetty efekti kuvaa sitä, miten ihminen muistaa hoitamattomat asiat hoidettuja paremmin.

# LENNOKAS AJATUSMATKA IHMISMIELLEN JA TEKOÄLYN YDINKYSYMYKSIIN

**K**onemieli vertaa ihmismieltä ja tekoälyä toisiinsa ja päätyy perimmäisten kysymysten äärelle. Kuinka tekoälyn matemaattinen prosessointi eroaa ihmisaivojen hermo-verkoista? Missä kaikessa tekoäly ja ihmismieli taas ovat samantaisia? Voiko kone ajatella?

Kirja avaa ajattelun rakenteita sekä ihmismielen että tekoälyratkaisujen näkökulmasta. Ihmismielen toimintaa valotetaan uusimman kognitiivisen psykologian, neurotieteen ja mielenfilosofian kautta. Tekoälyä puolestaan tarkastellaan tietojenkäsittelytieteen, koneoppimismenetelmien ja neuroverkkoarkkitehtuurien avulla.

Kirjan luettuasi ymmärrät, miten mieli toimii, mitä ajattelu on ja mistä tekoälytyökaluissa on oikeasti kyse.



**FT Lauri Järvilehto** on tietokirjailija, filosofi ja muusikko, joka tutkii ihmismielen ja uusien tekoälyratkaisujen toimintaa.

