

# Hiiriä ja ihmisiä – vielä tietokoneiden uskonnollisuudesta

Ilkka Pyysiäinen

**Matti Kamppisen ja Sami Pihlströmin rakentavat kommentit kirjoitukseeni ”Uskonnollinen robotti kiinalaisessa huoneessa” (*Tieteessä tapahtuu* 7/2002) tuovat mielenkiintoisia ja hyödyllisiä tarkennuksia tärkeään aiheeseen. Kamppisen kirjoitukseen liittyy kuitenkin joitakin ongelmia, joita on syytä tarkastella. Pihlströmiä kommentoin hyvin lyhyesti.**

Pyrin taannoisessa artikkelissani vain osoittamaan, millä ehdoilla robotti voisi olla uskonnollinen ja että näiden ehtojen toteutuminen on *toistaiseksi* utopiaa (sanan alkuperäisessä etymologisessa merkityksessä). En mitenkään yksiselitteisesti sitoutunut Searlen argumenttiin, koska se voidaan ymmärtää eri tavoin, siihen on vastattu monin eri tavoin, ja koko problematiikka on paljon mutkikkaampaa kuin Kamppinen antaa ymmärtää (ks. *Preston & Bishop* 2002). Omat väitteeni ovat paljon

vaatimattomampia kuin Kamppisen, joka sitoutuu selkeästi ”komputationalismiin”. Uskonto on mielestäni kognitiotieteen kannalta kiinnostava esimerkki ihmisajattelusta, koska siihen liittyy uskomuksia asioista, joita pidetään absoluuttisen tosina tai ainakin mahdollisina; silti ”uskovat” rajoittavat niiden käytön vain tiettyihin erityiskonteksteihin. Ihminen voi esimerkiksi uskoa, että ehtoollisleipä on Kristuksen ruumis ja silti suhtautua siihen kognitiivisesti ja emotionaalisesti täysin eri tavoin kuin ruumiisiin yleensä (*Pyysiäinen* in press).

Kamppisen kirjoitus selventää varsin hyvin, mistä hänen edustamassaan komputationalismissa (funktionalismissa) on kyse. Siinä on kuitenkin kolme ongelmaa: pakonomainen vitsailun tarve hämärtää asioita, komputationalismia ei suhteuteta muihin vaihtoehtoihin, ja tunteiden käsittely on todella heikolla pohjalla.

Kamppinen ihmettelee, mitä tarkoitan 'luonnollisilla tunteilla'. Arvelisin hänen kuitenkin tietävän, että luonnollinen voidaan kontrastoida joko yliluonnolliseen tai kulttuuriseen. Koska uskontoa toisinaan väitetään yliluonnollisen ilmentymäksi, on tarpeen korostaa, että minä käsittelen teoriassani vain luonnollisia tunteita. Toiseksi perustunteet ovat luonnollisia siinä mielessä, että ne eivät ole kulttuurisesti vaan biologisesti määrittäytyneitä. Ei tässä ole mitään "eksoottista".

Kamppinen olettaa jostakin syystä minun väittäneen, että uskonnossa toimivat tunteet, joilla ei ole systemaattista suhdetta havaittaviin ilmiöihin. Tämä on todella käsittämätön tulkinta, koska puhuin tunteista nimenomaan biologisina ja siis periaatteessa mitattavina ilmiöinä. Kenties sana 'luonnollinen' olisi pitänyt vielä lihavoida ja kursivoida? Nyt Kamppinen tuhlaa ruutia torjuakseen näkemyksen, jota en ole koskaan esittänyt.

Kamppinen puhuu tunteista koko ajan mentaalisisina tiloina eikä mitenkään kommentoi näkemystäni tunteista nimenomaan biologisina ilmiöinä. Tämä oli kuitenkin koko argumenttini ydin. Perinteinen funktionalistinen vastaväite siihen on, että vaikka tunteet implementoidaan biologisessa ruumiissa, on ne määriteltävä loogisen muotonsa perusteella algoritmiksi, joka voi toteutua missä tahansa perustassa. Tämä on kuitenkin kuin Helsinki määriteltäisiin Helsingin kartaksi eikä fyysiseksi paikaksi. Vaikka jokin orgaaninen prosessi voidaan kuvata tiettyjen sääntöjen avulla, ei tästä kuitenkaan välttämättä seuraa, että ao. organismi olisi noudattanut sääntöjä prosessin toteuttamiseksi (ks. Panksepp 1998: 81-82). Itse asiassa Turingin konekaan ei seuraa sääntöjä, koska säännön seuraaminen edellyttää tietoisuutta niistä. En voi osata kiinaa ilman, että tiedän osaavani. Tietokone siis vain näyttää seuraavan sääntöjä (Preston 2002: 40-42).

Kamppisen mukaan "Pyysiäistä vievät 'aito uskonnollisuus' ja 'luonnolliset tunteet'". Minua ei vie väkisin mihinkään edes Kamppisen kokoinen mies. Aidosta uskonnollisuudesta en ole puhunut lainkaan, ja luonnollisia tunteita vastaan Kamppisella itsellään ei näytä olevan mitään sanottavaa.

Kamppisen mukaan Searlen "kiinalaisen huoneen" argumenttia on kritisoitu hyvin perusteellisesti. Vakuudeksi hän viittaa "tarunhoitoiseen" konferenssiin, jota oli järjestämässä kahdeksan vuotta sitten, sekä omiin "oiviin" teksteihinsä. Uudempi keskustelu näyttää olevan hänelle tuntematonta. Ongelmana on se, että Kamppinen ei erittele kunnolla sen enempää Searlen argumenttia kuin sen vasta-argumenttejakaan.

Koska ulkopuolisen on jokseenkin mahdollista saada tarkkaa kuvaa Searlen argumentista minun ja Kamppisen kirjoitusten pohjalta, on syytä selvittää asiaa. Searle väittää, että pelkkä tietokoneohjelma on syntaktinen rakenne vailla semantiikkaa ja intentionaalisuutta. Siksi se ei voi koskaan tuottaa aitoa ymmärrystä tai tietoista mieltä. Kuitenkin "vahva tekoäly" väittää Searlen mukaan, että ohjelmat eivät vain simuloi mielen prosesseja, vaan ne ovat samanlaisia mielellisiä prosesseja kuin ihmismieli. Siispä mielen kognitiivisten prosessien täytyy selittyä samalla lailla kuin tietokoneohjelmien. Kognitio on aivoprosessi, jossa "mielen kielen" symboleita manipuloidaan vain niiden syntaktisten ominaisuuksien perusteella (vaikka niillä on myös semanttisia ominaisuuksia). Aivoissa ei ole mitään erityistä; sama syntaktinen ohjelma voitaisiin yhtä hyvin toteuttaa missä tahansa perustassa.

Searlen mukaan intentionaalisuus ja tietoinen ymmärrys sen sijaan ovat riittävän monimutkaisten aivoprosessien biologisia ominaisuuksia. Ne ovat emergenttejä ominaisuuksia aivan kuten veden nestemäisyys on veden molekyyliarakenteen emergentti ominaisuus. Aivoilla on kausaalinen voima tuottaa tietoisuus. Millään tietokoneohjelmalla ei tällaista voimaa ole. Tietokone toimii kuin suljetussa huoneessa oleva kiinaa ymmärtämätön mies, joka antaa kiinankielisiin kysymyslappuihin kiinankielisiä vastauslappuja ohjekirjan mukaan: tämännäköiseen lappuun annetaan vastaukseksi tämännäköinen lappu. Mies ei ymmärrä kiinaa lainkaan, mutta ulkona oleva kyselijä luulee hänen osaavan, koska saa asianmukaisia vastauksia. Searle ei kiistä ajattelevan koneen mahdollisuutta; ihminenkin on sellainen. Hän kiistää vain sen, että puhtaasti syntaktinen prosessi voisi taata mentaalisen tai semanttisen sisällön ilmaantumisen. Tällainen sisältö voidaan ehkä löytää tietokoneesta, mutta se on silloin peräisin jostakin erityisestä implementaatiosta, ei ohjelmasta sinänsä.

Searlelle on vastattu mm., että kiinalainen huone kokonaisuutena osaa kiinaa. Tällöin on kuitenkin jo luovuttu siitä premissistä, että pelkkä ohjelma riittää ja hyväksyty, että implementaatiolla on merkitystä. Sama pätee ns. robottivastaukseen, jonka mukaan asianmukaiseen ympäristöön sijoitettu liikkuva robotti voisi olla tietoinen ja intentionaalinen toimija. Tällöinkin hyväksymme, että semantiikka ei voi syntyä yksinomaan syntaksista.

Searle huomauttaa, että syntaktisuus ei ole aineen sisäinen ominaisuus. Se on aina havait-sijasta riippuvainen: me näemme jonkin syntaksina. Emme esimerkiksi voi selvittää, onko jokin kapine tietokone pelkästään avaamalla sen. Sisältä ei löydy ykkösiä ja nollia, vaan jotakin, minkä me voimme tulkita toimivan ikään kuin ykkösinä ja nollina. Hilary Putnam huomautti jo 1988, että mikä tahansa aineellinen rakenne voidaan nähdä minä tahansa funktionaalisen järjestelmänä; tietokoneohjel-mia on siis kaikkialla ja vahva tekoäly johtaa panpsykismiin (vrt. Block 2002: 77).

## Kamppisen kritiikki

Kamppisen mukaan ”on selvää, ettei kiinalai-sen huoneen argumenttiin vetoaminen ky-seenalaista uskonnollisen robotin mahdolli-suutta”. Ymmärrän tämän tarkoittavan, että uskonnollisen robotin periaatteellista mahdottomuutta ei voida osoittaa kiinalaisen huoneen argumentilla.

Kamppinen perustelee kantaansa sillä, että mentaaliset tilat perustuvat rakenteisiin, joissa objekteja prosessoidaan niiden syntaktisten ominaisuuksien perusteella ”eli koko joukkoon kiinalaisia huoneita”. Esimerkiksi kieltä voidaan oppia ”sopivassa ympäristössä, lukuisten kiinalaisten huoneiden ympäröimänä”. Huoneet muodostavat systeemisen kokonaisuuden, joka prosessoii merkityksiä. Syntaktiset rakenteet saavat intentionaalisen sisäl-lön käyttötilanteista.

Mutta tämä on jo eri näkemys kuin se, mitä Searle kritisoi. Hän sanoo vain, että mieli, semantiikka ja intentionaalisuus eivät voi perustua pelkkään syntaktiseen ohjelmaan. Kamp-pinen näyttää hyväksyvän tämän, mutta jättää avoimeksi, miten intentionaalisuus syntyy käyttötilanteista. Olemmeko siirtymässä kohti konnektionistisia malleja? Mikä on oppimisen rooli (ohjelman suorittaminen), mikä geneetti-

sen koodin (ohjelmointi)? Fodorin (2000) mukaan komputationalismille aiheutuu ongelmia juuri siitä, että mentaalisten prosessien kau-saaliseen luonteeseen vaikuttavat myös kontekstuaaliset tekijät. Arkipäiväisten rutiinien mallintaminen on vaikeaa koska emme voi etukäteen määrittää kaikkia relevantteja konteksteja. Mikä merkitys tällä on uskonnollisen robotin kannalta, sitä Kamppinen ei käsittele lainkaan.

Yksi vastaus on, että vaikka ohjelma sinän-sä ei synnytä tietoista mieltä, niin ohjelman suorittaminen synnyttää (Hauser 2002). Sear-len ratkaisu on kuitenkin toinen, ja sitä Kamp-pinen kritisoi. Hän syyttää Searlea neuronis-mista ja aivomystiikasta, koska tämä sanoo tietoisuuden olevan neuronijärjestelmien kau-saalisesti emergoituva systeemi. Onkin totta, että Searle jättää tässä turhan paljon avoimek-si. Hänen väitteensä voidaan kuitenkin tulkita niin, että tietoisuus syntyy dynaamisista neu-ronien itseorganisoituvista systeemeistä. Täl-löin aivojen kausaaliset voimat voidaan ym-märtää kyvyksi toteuttaa sisällöstä riippumat-tomia vuorovaikutuksen sääntöjä. Pahaksi on-neksi vain tämä tuhoaa Searlen kritiikin, jonka mukaan tietoisuus on biologinen eikä formaali ilmiö! Tietoisuus olisi sittenkin formaalin sys-teemin ominaisuus (Wheeler 2002).

Kamppinen jatkaa toteamalla, että mikäli aivojen ”kausaalista osista” (?) yksikin voi-daan korvata ei-biologisella osalla, on avattu polku keinoaivojen rakentamiselle. Tässä on kolme ongelmallista sanaa: ”mikäli”, ”avattu” ja ”keinoaivojen”. Tarvitaan lisäargumentteja osoittamaan, että keinoaivot tuottaisivat myös tietoisesta kokemuksesta (esim. Bringsjord & Noel 2002), että avattu polku johtaa päämäärään, ja että tällainen ajatuskoe toimisi käytännössä. Nyt Kamppinen lähinnä olettaa sen, mikä hä-nen piti todistaa. Toisaalta totesin omassa kir-joituksessani ihmisen ainutlaatuisuuden kon-tingentiksi ilmiöksi. En siis väitä tietäväni, että virtuaalinen kokemus on periaatteellisesti mahdotonta.

Kamppinen sanoo, että uskomukset yms. on *voitava* kuvata alempien tasojen avulla il-man samantasoista intentionaalisuutta ja että mentaalisten tilojen systeemisestä perustana *täytyy* olla rakenteita, joissa objekteja proses-soidaan (vain?) niiden syntaktisten ominai-suuksien perusteella. Näin on kuitenkin vain, mikäli komputationalismi/funktionalismi on tosi kuvaus mielen toiminnasta. Mutta juuri tästähän on kiistelty.

Pihlströmin kanssa olemme yhteistuumin todenneet, että polveileva keskusteluamme on nyt tullut määränpäähensä. Em. artikkelissani olin reagoinut lyhyesti Pihlströmin kahteen eri kirjoitukseen; laajemmin käsittelen niitä *Tiede & edistys* -lehden numerossa 3/2002, joka sisältää myös lehden Pihlströmiltä pyytämän vastineen vastineeseeni. Tässä totean vain sen, että en edes hulluimmista houreissani ole väittänyt kognitiivisen uskontotieteen olevan "kaikkivoipainen uskonnollisten ilmiöiden äärellä" (pace Pihlström).

\*

Searlen kriitikkojen kesken ei ole yksimielisyyttä siitä, *miten* argumentti epäonnistuu tai miten siihen pitäisi vastata. Joidenkin mukaan se ei edes ansaitse tulla keskustelluksi (Preston & Bishop 2002). Tästä seuraa, että ei ole olemassa selkeää näkemystä siitä, miten uskonnollisen robotin rakentaminen olisi periaatteessa mahdollista. Miten esimerkiksi erotamme uskonnollisen toiminnan ja uskonnollisen toiminnan jäljittelyn? Tätä olen yrittänyt sanoa. Lisäksi yritin sanoa, että ihmisten uskonnollisuudessa on merkittävä osuus evoluutiossa kehittyneillä perustunteilla. On olemassa valmiuksia (*preparedness*), joiden pohjalta jotkut kulttuuriset representaatiot voivat laukaista tiettyjä tunteita, erityisesti pelon tunteen (*Pyysiäinen* in press). Tämä olisi otettava huomioon robotin rakennuksessa. En kuitenkaan näe uskontotieteen päämääräksi robotin rakennusta edes periaatteessa. Mallit ja teorit eivät ole päämäärä sinänsä, vaan vain välineitä ilmiöiden selittämiseksi.

- Block, Ned (2002): "Searle's arguments against cognitive science". Teoksessa *Views into the Chinese room: new essays on Searle and artificial intelligence*, edited by John Preston & Mark Bishop, pp. 70-79. Oxford: Clarendon Press.
- Bringsjord, Selmer & Ron Noel (2002): "Real robots and the missing thought-experiment in the Chinese room dialectic". Teoksessa *Views into the Chinese room*, pp. 144-166.
- Fodor, Jerry (2000): *The mind does not work that way*. Cambridge, MA: MIT Press.
- Hauser, Larry (2002): "Nixin' goes to China". Teoksessa *Views into the Chinese room*, pp. 123-143. Oxford: Clarendon Press.
- Panksepp, Jaak (1998): *Affective neuroscience: the foundations of human and animal emotions*. New York: Oxford University Press.
- Preston, John (2002): "Introduction". Teoksessa *Views into the Chinese room*, pp. 1-50. Oxford: Clarendon Press.
- Preston, John & Mark Bishop (eds.) (2002): *Views into the Chinese Room*.
- Pyysiäinen, Ilkka (In press): "True fiction: philosophy and psychology of religious belief". *Philosophical Psychology*.
- Searle, John R. (1980): "Minds, brains, and programs". *Behavioral and Brain Sciences* 3, 417-457.
- Searle, John R. (1992/1995): *The rediscovery of the mind*. Cambridge, MA: MIT Press.
- Wheeler, Michael (2002): "Change in the rules: computers, dynamical systems, and Searle". Teoksessa *Views into the Chinese room*, pp. 338-359. Oxford: Clarendon Press.

*Kirjoittaja on Helsingin yliopiston tutkijakollegiumin tutkija sekä Helsingin ja Turun yliopistojen uskontotieteen dosentti.*