

Recensioner

Att vara du: En ny vetenskap om medvetandet

Anil Seth. Översättning: Joachim Retzlaff

Fri tanke 2024. 378 s. ISBN 978-91-8973-370-1

Titeln på denna bok, "Att vara du", låter lite underlig. Men den kan nog uppfattas ungefär som "Att vara medveten". Hur som helst ska boken presentera en ny teori – en ny vetenskap – om medvetandet. Om teorin är ny eller inte spelar väl mindre roll. Det viktiga är förstås om den är korrekt eller åtminstone någorlunda rimlig. Och man kan ju undra om vi ens har några alternativa teorier om medvetandet. Det finns förstås en mängd olika påståenden om medvetandet, men är de alls *teorier* eller *vetenskaper* i någon mer avancerad betydelse?

En teori om medvetandet ska enligt Seth också vara en teori om jaget. Han säger att en aspekt av medvetandets mysterium är "*jagets natur*" (s. 14).

Till skillnad från en del andra tänkare tycks Seth alltså inte tro att medvetandet och jaget är illusioner. Han vill "förstå hur den subjektiva upplevelsens inre universum är kopplad till, och kan förstås i termer av biologiska processer i hjärnan och kroppen" (s. 15). Det låter som om han vill presentera en lösning – eller en del av lösningen – på det som brukar kallas "det svåra problemet" (*the hard problem*).

Seth säger sig också acceptera *fysikalismen*. Riktigt vad den går ut på är rätt oklart, men kanske innebär det att "universum består av fysisk materia" (s. 28). Han anser nog att medvetandet är något fysiskt, och han utgår från att upplevelser finns, så de är tydligen också i någon mening "fysiska".

En vanlig variant av fysikalism, inte minst bland neuroforskare, är *funktionalism*. Den går ofta ut på att medvetande är en sorts informationsbehandling, något som omvandlar fysisk input (nervimpulser) till fysisk output (beteende). Detta kallas *the computational theory of mind* (CTM). Seth är skeptisk till CTM, dels därför att det är oklart vad "information" är och dels för att man inte kan bortse från att medvetande kanske har något nödvändigt samband med biologi eller liv (och i så fall inte kan finnas i datorer). Han anser att vi inte kan utesluta att AI kan bli medveten, men att det är "ytterst osannolikt" (s. 277). Men han tänker sig att man på sikt

kanske kan skapa "entiteter", någon sorts robotar, som *känns* för oss som om de vore medvetna, även om vi tror att de inte är det (s. 289).

Själv lutar Seth åt en variant av funktionalism, som är agnostisk i fråga om hur informationsbehandlingen kan realiseras (s. 32, 273).

Men om medvetandet bara är en sorts informationsbehandling, så behöver det ju inte innefatta några upplevelser. Eller är tanken att output inte bara består av yttre beteenden, utan också av inre upplevelser? Seth skiljer mellan medvetandets *fenomenella* och dess *funktionella* och *beteendemässiga* egenskaper (s. 24). Själv utgår han från det fenomenella (i enlighet med Thomas Nagels uppfattning om medvetande). Men då är vi ju tillbaka till det svåra problemet: Hur kan fysiska fenomen (nervimpulser) ge upphov till fenomenella fenomen som subjektiva upplevelser?

Att fenomenella fenomen är "fysiska fenomen" betyder kanske bara att de *orsakas* av fysiska faktorer, nämligen elektro-kemiska händelser i hjärnan. Det kan låta lite konstigt, men tydligen menar Seth det.

Enligt det så kallade zombie-argumentet är det "tänkbart" att världen kunde vara precis som nu, utom i det att inget medvetande finns, det vill säga alla människor är zombier – vilket i så fall antas visa att medvetandet inte är ett fysiskt fenomen (s. 32f).

Vad ska man säga om det? I någon mening kan jag förstås "tänka mig" en helt annan värld, där inget medvetande finns trots att den är fysiskt identisk med vår värld (något som Seth på rätt oklara grunder *inte* kan tänka sig). Jag har ju just uttryckt denna tanke. Men det är fullt förenligt med att medvetandet i vår värld orsakas av fysiska faktorer. Att jag kan tänka mig en zombievärld – och att det kan finnas zombier i vår värld – hindrar inte att jag tror att mitt medvetande, inklusive mina tankar om vad som är tänkbart, orsakas av vad som händer i min hjärna. Men om medvetande är något mer än bara "informationsbehandling", något som även innefattar fenomenella fenomen (upplevelser), så är vi återigen tillbaka till det svåra problemet. Hur går det till? Och varför?

Seth vill dock inte ta sig an det svåra problemet. I stället fokuserar han på något han kallar "det *verkliga* problemet", det vill säga "att *förklara, förutsäga och kontrollera* den medvetna upplevelsens fenomenella egenskaper". Vilket innebär att förklara varför en viss fysisk process är förbunden med ett visst slag av medveten upplevelse (s. 35).

Man ska till exempel förklara varför vissa specifika hjärnaktiviteter är korrelerade med en viss upplevelse av rött, snarare än av blått eller tandvärk eller svartsjuka (s. 37). Det är alltså fråga om själva upplevelsen, snarare än ett visst beteende, som att uttala orden "Det ser rött ut" eller

att peka på något rött i en färgkarta. Seth föreställer sig att om detta program genomförs, så kommer det svåra problemet att förblekna och gå upp i "metafysisk rök" (s. 41).

Men hur kan man då veta vilken fenomenell kvalitet som en upplevelse har? Hur vet man exempelvis att det som en person kallar rött inte upplevs som det som en annan person kallar blått? Det är ett klassiskt problem. Kan man verkligen komma åt fenomenella kvaliteter? Kommer man inte bara att fastna i beteenden? Speciellt språkliga?

Seth anser att det finns olika *nivåer* av medvetande och att det vore bra att kunna *mäta* detta (ungefär som för temperatur eller liv). Först då kan man kanske hitta förklaringar (analogt med att värme är molekylrörelser). Det man mäter är graden av vissa typer av hjärnaktivitet, vilket jämförs med grad av medvetande så som det bedöms av neurologer (från narkos via REM-sömn till full medveten vakenhet). Men såvitt jag förstår mäter man då inte *fenomenella* egenskaper – annat än möjligen indirekt, vilket i så fall förutsätter någon teori.

En rätt populär teori om medvetandet bland neurologer är *teorin om integrerad information* (IIT). Enligt IIT är medvetande *detsamma som* integrerad information (s. 74) och "ett system är medvetet i den mån helheten genererar mer information än dess delar" (s. 75). Men den teorin har enligt Seth brister, bland annat kan den inte testas. Hans kritik av IIT (s. 73–86) är intressant och verkar rätt övertygande.

Vilken är då Seths teori? Jag är inte säker på att jag har lyckats förstå det, men kanske kunde den kortfattat formuleras som att medvetandet består av *kontrollerade hallucinationer*, vilka bygger på hjärnbaserad "minimering av förutsägelsefel" (s. 211).

Vi tror normalt att en varseblivning uppstår genom att något i verkligheten påverkar våra sinnesorgan, vilka i sin tur tolkar dessa nervimpulser och därmed levererar en upplevelse av verkligheten. Processen är här utifrån-och-in. Men Seth anser att processen i stället är inifrån-och-ut. Varseblivning är enligt honom en sorts slutledning, nämligen "abduktion" eller "slutledning till bästa förklaring" (s. 113). Hjärnan skapar förväntningar som sedan testas gång på gång mot sensorisk input. Den minimerar förutsägelsefel och därigenom uppstår "kontrollerade hallucinationer". Detta är vad vi upplever, våra upplevelsers innehåll (s. 99f).

Det här kan ju leda tankarna till liknande idéer hos exempelvis Platon och Kant. Men hos dem fanns det ju något verkligt bakom fenomenvärlden, något som vi kanske inte kan veta något om. Men vad är verkligt hos Seth?

Talet om "hallucinationer" antyder ju att det vi upplever inte är verkligt, men samtidigt menar Seth att när vi är överens om våra hallucinationer, så *kallar* vi det verkligt (s. 103). Men *är* det verkligt?

Hjärnan och processerna i den hör ju till det vi upplever (varseblir), inte minst i de laboratorier där Seth är verksam. Enligt hans teori är de alltså hallucinationer, även om Seth kallar dem verkliga när han är överens med sina kolleger. Han säger också att "våra perceptuella bästa gissningar måste upplevas som *verkligen existerande i världen* för att vara nyttiga för organismen" (s. 215). *Upplevas* som verkliga, alltså, men tydligen inte vara verkliga!

Jag har svårt att begripa detta, och jag är också tveksam till om det kan kallas fysikalism.

Seth säger att hjärnan "översköljs av en flod av sensoriska signaler från världen och kroppen" (s. 133). Gör den verkligen det, eller är det bara något vi hallucinerar? Vi upplever inte en objektiv verklighet, utan en projektion, säger han (s. 137). Och "organismer samlar och modellerar sensorisk information för att maximera de sensoriska beläggen för sin egen existens" (s. 225). Verkligheten är en projektion på samma sätt som Hume menade att kausalitet är en projektion (s. 156). Men Seth anser att det är just detta, att vi tycker oss uppleva en objektiv verklighet, som gör att det svåra problemet uppstår (s. 158). Anser han alltså att vi borde sluta att tro på en objektiv verklighet och nöja oss med hallucinationer? Och finns "vi" över huvud taget, när även jaget är en kontrollerad hallucination (s. 167)?

Är det vi upplever som verkligheten detsamma för oss alla? Seth säger att "våra perceptuella upplevelser av världen är inre konstruktioner, formade av vår biologis och vår levnadshistorias idiosynkrasier" (s. 105). Det talar väl för att vi inte upplever verkligheten på samma sätt. Att vi ändå är "överens om våra hallucinationer" innebär kanske bara att vi *beskriver* dem språkligt på samma sätt.

Seth kan uppenbarligen en hel del om människors (och andra djurs) hjärnor och hans bok innehåller en stor mängd detaljer om hur de fungerar. Jag har här inte gått närmare in på detta, utan mest hållit mig till det som är av filosofiskt intresse. Så låt mig avslutningsvis beröra det han säger om viljans frihet.

Viljefrihet är enligt Seth en fråga om *känslor*: vi *känner* att vi gör det vi vill göra, att vi kunde ha gjort något annat och att våra handlingar tycks komma inifrån (s. 240f). Men att man har dessa känslor innebär inte att man faktiskt kunde ha handlat annorlunda (s. 245). De har inga direkta

motsvarigheter i "en bestämd egenskap hos världen". Men *upplevelser* av viljefrihet är verkliga och de är "oundgängliga för vår överlevnad" (s. 248). De är nyttiga i framtida liknande men inte identiska situationer. Det tycker jag låter rimligt.

Lars Bergström

Världen vaknar: Känslornas plats i människans liv

Fredrik Svenaeus

Gidlunds förlag 2023. 210 s. ISBN 978-91-7844-507-3

Boken handlar om känslornas centrala betydelse i våra liv, och den består av en kortare inledning, sju kapitel och en kort avslutning. Kapitel 1 handlar om existenskänslan och om skillnaden mellan människor och maskiner. Kapitel 2 har titeln "Kroppens vibrationer: grundstämningar och sociala emotioner" och kapitel 3 handlar om människoblivandets olika stadier. Därefter upphör kopplingen till människolivets olika faser, för att sedan upptas igen i kapitel 7, som handlar om döendet och döden. Detta är bokens svagaste och mest ostrukturerade kapitel, och den som intresserar sig för något av de ämnen som behandlas där gör bäst i att söka information på annat håll. Kapitel 4 och 5 kretsar kring teman som etik och politik samt kunskap och vetenskap, och här utgår Svenaeus från att känslor spelar en viktig roll i alla mänskliga aktiviteter och praktiker, även i dem som ibland betraktas som känslobefriade. Jag har ibland undrat över vem Svenaeus skriver för, men dessa båda kapitel tror jag är riktade till andra filosofer. Kapitel 6 handlar om sjukdom, och dess kärna är en fenomenologisk analys av sjukdomskänslan.

Svenaeus beskriver själv sin bok som en filosofiskt argumenterande essä: han lägger ut tanketrådar åt olika håll som han sedan försöker knyta ihop till ett övergripande resonemang (s. 11). En traditionell filosofisk text är lättare att granska än en essä. Man kan då fråga sig vad författaren vill säga, vilka de centrala frågeställningarna är och vilka metoder hon har valt för att besvara sina frågor. Man kan undersöka om det som påstås är tillräckligt precist formulerat och om det är tillräckligt välgrundat. Man kan även påpeka att vissa tankegångar lyser med sin frånvaro eller att delar av texten är irrelevanta och överflödiga. Men är det tillåtet att använda ett sådant angreppssätt när man recenserar en essä? Jag tar mig friheten att göra detta, men är samtidigt medveten om att det inte är problemfritt.