

Februari 2015 Årgång 36 Nr 1

Filosofisk tidskrift

- PETER GÄRDENFORS
3 Varför vill vi förstå?
- PER BAUHN
16 Korsgaards teori om enade agenter
- LARS BERGSTRÖM
26 Gudar och mörk materia
- SIMON SKAU
31 Epistemiska trösklar, ett svar på problemet
med gemensam kunskap
- INGVAR JOHANSSON
40 Kahneman och Lindaproblemet
- RECENSIONER
43 Gustav Alexandrie om *Etik, människa, moral, mening:
En introduktion* av Ann Heberlein
49 Lars Bergström om *Vårt matematiska universum*
av Max Tegmark
- 60 NOTISER

Filosofisk tidskrift

Utges av Stiftelsen Filosofisk tidskrift och Stiftelsen Bokförlaget Thales

Redaktör och ansvarig utgivare: Lars Bergström

Filosofisk tidskrift utkommer med fyra nummer per år

Prenumeration per år inom landet: 200 kr (inkl. moms); utom landet: 300 kr

Lösnr: 55 kr (inkl. moms)

För prenumerationsärenden kontakta:

Nätverkstan ekonomitjänst, Box 31120, 400 32 Göteborg

Telefon 031-743 99 05 Fax 031-743 99 06

ekonomitjanst@natverkstan.net

Förlagsadress: Bokförlaget Thales, Box 50034, 104 05 Stockholm

info@bokforlagetthales.se www.bokforlagetthales.se

Copyright © Respektive författare 2015

Grafisk form: Ulf Jacobsen

Satt med Indigo Antiqua

Tryck: Carlshamn Tryck & Media AB, Karlshamn 2015

ISSN 0348-7482

1. INLEDNING

Varför är det så viktigt för människor att vi förstår? Varför räcker det inte med att vi vet? Vi kan lätt föreställa oss ett liv där vi följer en uppsättning medvetna eller omedvetna regler utan att förstå varför reglerna finns. Inte desto mindre är förståelse en stark motivation för många av våra strävanden.

Inom den tidiga kognitionsvetenskapen sågs hjärnan som en dator som bearbetar information. Jerome Bruner påpekar i sin bok *Acts of Meaning* (1990) att om tänkandet ses som enbart informationsbehandling, så försvinner frågor om mening och förståelse. Men en teori om den mänskliga naturen kan inte baseras på enbart ”information”, ”beräkning” och ”beteende” – den måste också beskriva vad det innebär att förstå. Bland annat handlar det om att sätta människans handlingar i ett socialt och kulturellt sammanhang.

Förståelse är inget mystiskt, men mycket lite är känt om hur förståelse uppstår. Inom humaniora finns den hermeneutiska traditionen om ”Verstehen”. Men denna tradition fokuserar på hur kulturella produkter – böcker, målningar, symfonier, etc. – kan tolkas. Forskningen är fortfarande i sin linda när det gäller vilken kunskapsteoretisk roll vardagsförståelsen har. Vad händer med din kunskap när du upptäcker ett samband du inte har sett tidigare? En ny insikt är ofta knuten till en aha-upplevelse. När det gäller att beskriva vad som händer i våra huvuden när vi förstår, har vetenskapen inte mycket att säga.

Trots de enorma framstegen under de senaste århundradena finns det överraskande luckor i vår vetenskapliga kunskap. Själva begreppet förståelse har ofta förbisetts i pedagogisk, filosofisk och psykologisk forskning. Ett undantag är Georg Henrik von Wrights bok *Explanation and Understanding* från 1971.

I denna artikel¹ kommer jag att fokusera på vilken roll förståelse har för kultur och vetenskap, och mer specifikt på hur vi använder abstrakta

¹Denna artikel är en översättning och bearbetning av Gärdenfors (2010).

teorier för att förstå världen. Men först av allt måste jag ge en allmän bakgrund till hur jag förstår förståelse. Som en arbetshypotes föreslår jag *att förstå är att se ett mönster* (Gärdenfors 2007a, Gärdenfors och Lindström 2008). Jag ska börja med att diskutera hur denna idé, som härrör från gestaltpsykologin, kan tolkas inom vetenskapsteorin i termer av ”teoretiska variabler”. Framför allt kommer jag att hävda att människor är inställda på att söka orsaker och att förstå andras tankar.

Som en tillämpning av hypotesen skall jag visa att de mönster man ser till en viss del är beroende av den kultur man lever inom. En annan tillämpning gäller betydelsen av förståelse inom naturvetenskap. Jag skall argumentera för att även denna form av förståelse kan inordnas under tesen att förstå är att se ett mönster.

2. ATT FÖRSTÅ ÄR ATT SE ETT MÖNSTER

Efter en konstutställning eller en konsert säger vi ibland att upplevelsen har gett oss starka intryck. Men det är en vanföreställning att vi får sinnesintryck – det är inget som ”trycks in” i huvudet på oss. Hjärnan är ingen passiv mottagare av bilder och ljud från omvärlden. Den söker aktivt efter mönster och den tolkar omvärlden. Denna ständigt pågående process är grunden till all form av förståelse.

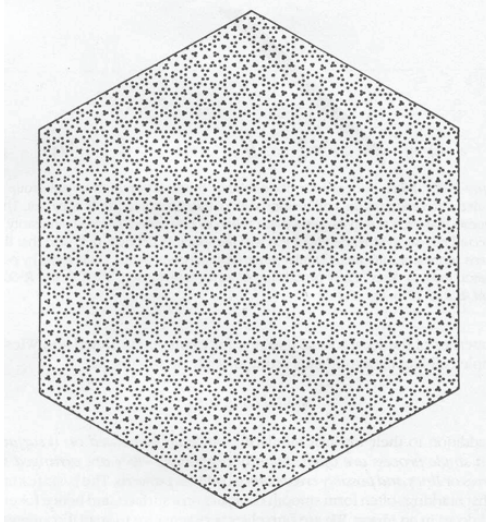
Vi kan finna mer eller mindre abstrakta mönster inom alla former av tänkande, från att se ränderna på en katt, till att identifiera den musikaliska strukturen i en fuga av Bach eller att förstå mikroorganismernas roll i sjukdomar. Några av de mönster vi uppfattar ges av vår biologiska konstitution, en del har vi lärt oss under vår barndom och vår utbildning, en del kommer från den kultur vi lever i och en del tillhandahålls av vetenskapen. Vad som händer i våra huvuden när vi uppfattar ett mönster är ett av de mest centrala problemområdena inom kognitionsvetenskapen. Men samtidigt som vår förmåga att lösa nya problem ökar när vi utnyttjar vår förmåga att se mönster, så kan invanda mönster också hämma vår förmåga att närma oss nya situationer med ett öppet sinne.

Framför allt uppfattar vi mönster visuellt. Men idén om att förstå genom att uppleva ett mönster kan tillämpas på alla sensoriska modaliteter. Några exempel: en sommelier kan känna igen ett vin genom att identifiera syra, sötma, etc.; ett tvåspråkigt barn har lärt sig de olika fonem som finns i två språk; en parfymtillverkare kan skapa nya dofter genom att ändra kombinationer av oljor och blomextrakt; och en skulptör kan fånga olika aspekter av en åldrande kropp med dess stelade leder, avmagrade muskler och rynkiga hud.

Vår hjärna söker efter mönster, oavsett om vi är medvetna om det eller inte, på olika nivåer. I botten finns biologiskt betingade mekanismer som

kraftigt styr hur vi uppfattar omvärlden. På de mest avancerade nivåerna finner man de kulturella mönster som krävs för att tolka konstverk, poesi, musik, dans, etc. Men när vi behärskar dessa mönster, påverkar de också vår uppfattning av kulturprodukterna. Följaktligen ser och hör vi inte samma saker – det finns ingen ”objektiv” beskrivning av världen.

För att illustrera att sökandet efter mönster är allestädes närvarande, börjar jag med ett exempel på en biologiskt styrd mekanism från den visuella modaliteten (se figur 1).

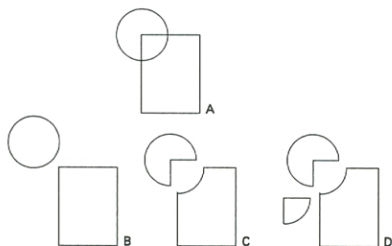


Figur 1. Exempel på dolda visuella mönster (Från Marroquin 1976)

I ett tidigt skede av den visuella processen försöker hjärnan hitta mönster bland alla punkterna. Vi uppfattar cirklar av olika storlekar i figuren. En intressant aspekt är att en cirkel är snart ersätts av en annan som undertrycker den första, etc. Figuren ”lever” i vår perception även om inte en enda prick rör sig. Man kan till och med upptäcka maltesiska kors i figuren. Om du gör det kommer korsen att blockera perceptionen av cirklarna – och vice versa. Figuren har ingen helhetsmening, men vårt visuella system söker oavbrutet efter mönster. Det finns därför inget entydigt svar på hur man ”uppfattar” figuren.

Hjärnan är full med olika mekanismer som *fyller i* det som faller på näthinnan, trumhinnan eller andra sensorer. När vi exempelvis tolkar ytor i en bild har vi en stark tendens att binda samman linjer som har samma riktning till konturen av ett objekt. Detta är en del av den process kallas ”slutenhet”. Denna och liknande processer studerades av gestaltpsykolo-

gera i den första halvan av nittonhundratalet, men de har fått ett förnyat intresse eftersom vi nu börjar förstå de mekanismer i hjärnan som ligger bakom. I bilden nedan kan figuren A delas upp i ytor på tre sätt. Tolkningen D passar kanske bäst med hur linjerna går i bilden, men våra hjärnor väljer hellre tolkning B eftersom den slutna cirkeln är så invand för vårt mönsterseende. Tolkningarna C och D verkar mycket mer "onaturliga".



Figur 2. Perception av gestalter

De mekanismer som fyller i är viktiga ur ett evolutionärt perspektiv. Dåliga ljusförhållanden eller skymmande objekt får inte hindra oss från att upptäcka fara eller mat, som tydligt ökar vår överlevnadsförmåga. Om man ser svansen på en tiger, är det livsviktigt att förstå att det finns en hel tiger i närheten.

Våra hjärnor har konstruerat en stor repertoar av mönster. Vi är ofta inte medvetna om dem, men de kan framkallas av olika trick. Betrakta exempelvis de två bilderna i figur 3. Vad föreställer de?

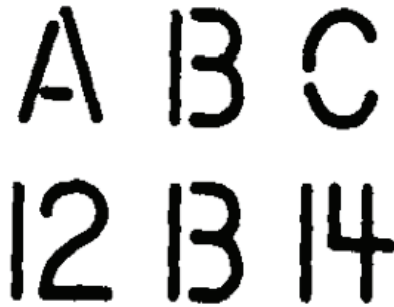


Figur 3. Ofullständiga avbildningar av föremål

Det är kanske inte särskilt svårt att se att bilden till vänster är en violin, men många behöver en ledtråd till den högra figuren. Det är en *elefant* (med huvudet till vänster). Plötsligt faller bitarna på plats och man kan

”tolka” bilden. De olika delarna blir meningsfulla; en liten svart fläck som man knappt lade märke till tidigare blir nu en elefantbete. En intressant effekt av en sådan gestaltupplevelse är att när man väl har sett ett mönster kan man inte släppa det. Även om man hade svårt att se vad den högra bilden föreställde, går det inte längre att låta bli att se den som en elefant.

För vissa kategorisystem, förstärks effekterna av en kategorisering av de perceptuella systemen så att avstånden inom en kategori uppfattas som mindre och avstånden mellan kategorierna uppfattas som större än vad de ”egentligen” är. Detta fenomen kallas *kategorisk perception* (se exempelvis artiklarna i Harnad 1987). Det innebär att en verklighet där det inte finns några skarpa gränser sorteras upp i distinkta fack av våra kognitiva mekanismer. I tolkningen av perceptionen lägger vår hjärna till gränser som inte har någon verklig motsvarighet. Gränserna ges av ett system av *kategorier* som vanligen är en produkt av inläring. Ett enkelt men illustrativt exempel på kategorisk perception ges av figur 4:



Figur 4. Exempel på kategorisk perception.

Fokuserar vi på den översta raden tolkas tecknet i mitten som bokstaven B, men om vi bara ser den undre raden blir samma figur i mitten talet 13. *Sammanhanget* styr alltså hur vi tolkar den information som mottas av våra sinnen. I en kontext där det förekommer bokstäver förutsätter vi att alla mönster skall tolkas som bokstäver. Denna mekanism förstärker vår tolkningsförmåga, men den skapar också läsningar. Mönstren ovan är kulturellt överförda eftersom de utgör en del av den grundläggande västerländska utbildning som använder latinska bokstäver och arabiska siffror.

Kategorisk perception förekommer i många domäner, men de har i synnerhet studerats för fonetiska system (se till exempel Petitot 1989). Att förstå ett språk förutsätter att man på rätt sätt kan kategorisera språkljuden – fonemen – och gruppera dem i enheter som vi kallar ord. Även om en uppsättning ljud kan produceras genom en talprocess som

varierar kontinuerligt, uppfattar hörselsystemet dessa variabler på ett kategoriskt sätt så att när talprocessen varieras, upplevs ljudet som konstant för ett brett intervall och sedan hoppar det till ett nytt fonem som också upplevs som relativt stabilt.

3. KAUSALDRIFTEN

Människor har en oemotståndlig drift att förstå världens mekanismer. Vi har en stark benägenhet att finna *orsaker* i världen. Kant utnämner denna benägenhet till en av våra åskådningsformer – ett tankemönster som vi inte kan ställa oss utanför. I min bok *Den meningssökande människan* (Gärdenfors 2006), kallar jag denna tendens *kausaldriften*. Det finns goda evolutionära skäl för att vi skall leta efter orsaksförhållanden eftersom de hjälper oss att förstå hur världen hänger ihop. Genom att resonera om orsaker och verkningar blir vi bättre på att förutse framtiden. Och människan är, mer än något annat djur, beroende av sina tankar om framtiden (Gärdenfors 2000).

Det finns i själva verket ett mycket nära samband mellan att uppfatta mönster och att förstå orsaker. En orsaksvariabel är en sorts mönster som kopplar samman orsaker med verkningar. Och omvänt, när man ser ett mönster, ser man nya samband mellan fenomen i världen och mönstret kan därför fungera som ett sätt att skapa nya orsaksförklaringar som ger ökade möjligheter att lösa problem. (Naturligtvis finns det mönster som inte ger upphov till orsaksförklaringar.) Om en förälder exempelvis uppfattar ett samband mellan att ett barn inte ätit på länge och att det blir gnälligt, kan hon eller han ge barnet mat och på så sätt lösa problemet. Får föräldern sen lära sig den abstrakta variabeln ”låg blodsockerhalt” som en förklaring till gnälligheten och förstår vilken föda som snabbt kan höja blodsockret, blir problemlösningen ännu effektivare.

Till stöd för min uppfattning om kausaldriften hävdar exempelvis psykologen Alan Leslie (1987) att spädbarn inte bara kan följa objekt med blicken utan att de också har en speciell mekanism eller modul i hjärnan som beräknar de *krafter* som påverkar ett objekt. Ett sådant system är ett paradexempel på hur den mänskliga hjärnan kan utnyttja de dolda variabler som styr världen omkring oss.

Människan har en unik talang att extrahera sådana dolda variabler. Inom vetenskapsteorin kallas sådana variabler *teoretiskastorheter* (Hempel 1965). Dessa variabler är ofta de som ger ryggraden till ett nytt mönster. Ett exempel är när vi kategoriserar handlingar och använder de mönster av krafter som vår tolkande perception ger oss (Runesson 1994, Gärdenfors och Warglien 2012). På en mer abstrakt nivå kan olika fysiska, sociala och andra variabler hjälpa oss att skapa mening i en upplevelse som till en

början var obegriplig. Att någon är nyförälskad kan exempelvis förklara många till synes märkliga beteenden. De dolda variabler som används för att se orsaker är ett sätt att fylla i en typ av mönster som människor är speciellt duktiga på att upptäcka.

Ett sätt att beskriva skillnaden mellan människor och andra arter är att våra inre världar utnyttjar dolda variabler mycket effektivare för att göra förutsägelser om framtiden. Orsakerna ges inte av våra sinnen, men våra hjärnor fyller i dem. Peter White (1995) hävdar till och med att vi bokstavligen uppfattar de krafter som styr händelser i världen, precis som vi uppfattar konturerna av föremål, även om de inte ingår i den visuella information som når ögonen.

Den stora evolutionära fördelen med kausaldriften är att den förbättrar vår förmåga att förutsäga. Men ibland blir den överdriven: även rena tillfälligheter, som exempelvis lotterivinster, ses som ett resultat av att man har tur. Men ”tur” finns inte – det är ett rent påhitt av vårt tanke-system. Det faktum att vi talar om tur är bara ett tecken på att våra hjärnor är besatta av att finna orsaker till allt som händer. Om det inte finns någon orsak, måste vi uppfinna en. Kausaldriften är därför en stark faktor bakom skapandet av myter och magiskt tänkande. Det magiska tänkandet är särskilt starkt hos barn. Piaget (1930, s. 174) skriver:

Barnet fyller världen med spontana rörelser och ”krafter”; himlakropparna kan vara i vila eller röra sig som de vill, molnen gör själva vind, vågorna ”reser” sig själva, träden svajar med sina grenar för att göra en bris.

Denna tanke kan tänjas till vuxenvärlden och formuleras som att vad vi kallar magiskt eller *animistiskt* tänkande bara är en tillämpning av psykologisk kausalitet på fysiska fenomen – man ser varelser lite varstans som orsaker till händelser. Magiskt tänkande förekommer inte bara hos barn utan det är vanligt, i skiftande former, i alla åldrar och i alla kulturer. Har man en gång drabbats av en olycka någon gång strax efter att en svart katt sprang över vägen från vänster till höger, så blir man orolig vid ett långt senare tillfälle när man ser en svart katt. Och man skall inte gå under stegar eller falla upp paraplyer inomhus, och så vidare.

Jag vill påstå att människans förståelse av fysiska krafter har utvecklats via animism och antropomorfism. Förmodligen har vi länge tolkat de fysikaliska skeendena med hjälp av sociala variabler och småningom lärt oss tänka i termer av opersonliga fysiska krafter. Detta överensstämmer med Collingwood (1972), som skriver: ”Kausala påståenden . . . är beskrivningar av relationer mellan naturliga händelser i antropomorfa termer.” Det bör noteras att det är först under de senaste århundradena som vi har skapat en rent mekanistisk beskrivning av världen.

4. ATT FÖRSTÅ ANDRAS TANKAR

Kausaldriften hos människor gäller inte bara den fysiska världen, utan också den mentala. Förutom fysiska variabler uppfattar vi de ”mentala krafter” som styr beteendet hos andra. Uppfattningen av dessa krafter utgör grunden för vår *inlevelseförmåga*. I detta sammanhang betyder inlevelseförmåga att man kan förstå olika aspekter av andras inre världar. Jag använder här termen inte bara för förståelse av andras trosuppfattningar (beliefs), men för alla former av medvetande, såsom känslor, önskningsar, uppmärksamhet och avsikter. I den filosofiska debatten brukar inlevelseförmåga kallas ”theory of mind” (Premack och Woodruff 1978, Mitchell 1997). Jag undviker denna term, eftersom den oftast bara handlar om att förstå andras trosuppfattningar, vilket bara är en aspekt av flera när det gäller inlevelseförmåga.

Samtidigt som vi blivit bättre på att föreställa oss andras inre världar, har vi blivit skickligare på att skapa dolda variabler som kan förklara andras tankar. Först och främst måste vi förstå andras handlingar, eftersom vi i allmänhet härleder andras mentala tillstånd från deras beteende (inklusive deras tal). Denna mekanism kan ses som en variant av en som tillåter oss att härleda fysiska krafter från föremål som förflyttas. De intentioner som vi tillskriver andra är de dolda krafter som att gör andra betar sig som de gör. Eftersom vi kan föreställa oss vad andra vet och inte vet, blir vi bra på samarbete, något som har främjat utvecklingen av människans språk, kultur och samhälle (Gärdenfors 2008). En nackdel med detta är att vår kunskap om andras kunskap också gör oss mer skickliga på bedrägeri och mot-bedrägeri. Om jag kan föreställa mig vad andra vet om vad jag vet, är det möjligt för mig att förutse andras potentiellt bedrägliga handlingar.

I tidigare arbeten (Gärdenfors 2000, 2007b, 2008), har jag skilt ut olika nivåer av inlevelseförmåga för att förstå de kognitiva kapaciteterna hos djur och hos barn i olika åldrar:

- i. Att förstå andras känslor. På denna nivå kan man till exempel förstå att någon annan har ont. Detta är vad som brukar menas med *medkänsla* (empati).
- ii. Att förstå andras önskningsar. Denna förmåga innebär att förstå, till exempel, att andra inte tycker om samma saker som du gör.
- iii. Att förstå andras uppmärksamhet. Detta betyder att man exempelvis kan förstå vad någon annan fokuserar blicken på.
- iv. Att förstå andras avsikter. Denna förmåga innebär framför allt att kunna tolka vilket mål som ligger bakom en annan individs beteende.
- v. Att förstå andras kunskap. Denna förmåga innefattar att man kan föreställa sig vad andra individer vet eller tror.

Önskningsar, avsikter och trosuppfattningar är således de viktigaste delarna av variablerna vi använder för att förstå andra människor.

5. ATT FÖRSTÅ ANDRA KULTURER

En kultur är inte bara människor i en miljö, utan ett särskilt sätt att förstå världen. Varje kultur för med sig en uppsättning mönster. Ett problem är att det kan vara svårt att uppfatta andra kulturers mönster. Eftersom de inte passar med vår egen kulturs mönster, upplever vi dem ofta som konstiga, udda eller helt enkelt främmande.

Musikaliska skalor är tydliga exempel på hur kulturella mönster faktiskt kan påverka själva upplevelsen. Västerlänningars musikförståelse är styrd av en skala som innehåller tolv halvtoner. Inom arabisk musik har man 17 toner i skalan och inom indisk har man hela 21. För våra västerländska öron, som vill sortera in tonerna från en indisk raga eller en arabisk kärlekssång i vår tolvtonsskala, låter musiken gnällig, eftersom vi inte kan få den att "stämma" helt. Vi kan inte höra mönstren eftersom vårt kategoriska öra inte kan knäcka koden. Eller som Goethe säger i en av sina maximer: "Es hört doch jeder nur, was er versteht." Ett alternativt mönster i form av en annan skala måste läras in innan musiken kan uppskattas fullt ut.

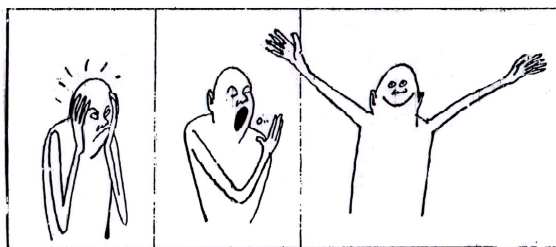
Kulturella utbildningsprogram har en tendens att fokusera för mycket på att lära ut fakta om andra kulturer och för lite på att förstå grundläggande mönster som kulturella sedvänjor eller trossystem bygger på. Därför bör utbildning som strävar efter att överbrygga kulturer fokusera på att förmedla relevanta mönster snarare än på fakta.

Mönstren är det som ger innehåll åt de begrepp vi använder när vi sorterar världen. Det finns därför klara kopplingar mellan de mönster vi uppfattar och det språk vi använder. De mönster som får starkast fotfäste i vårt tänkande är de som är gemensamma för medlemmarna av ett samhälle eller en kultur. Mönstren förstärks när vi *kommunicerar* om dem med ord, bilder eller andra media.

Orden står således för element i vår föreställningsvärld och bara indirekt för något i omgivningen. Men när vi väl har orden kan de, via de mönster de framkallar, också styra hur vi tolkar världen. Vissa mönster hör till vår kultur och det krävs ett möte med andra kulturer för att inse att dessa mönster inte är självklara. Ett vardagsexempel är begreppet "frukost". Slår man upp ordet "frukost" i ett engelskt lexikon får man "breakfast" och i ett italienskt står "prima colazione". Men när man kommer till hotellfrukosten i England respektive Italien möts man av två radikalt olika fenomen. En svensk eller en engelsman blir normalt besviken när man ser den minimala "prima colazione" som serveras i

Italien, och en italienare blir överraskad av det kalorimässiga överflödet i en ”full English breakfast”.

Kulturella mönster kan vara subtila. Ett slående exempel på en sådan läsning kommer från en reklamkampanj för huvudvärksmedicin som ett företag lanserade i Nordafrika. Eftersom en stor del av invånarna i målområdet var analfabeter utformades budskapet som en tecknad serie med tre rutor.



Figur 5: En misslyckad reklamkampanj

Reklamen för medicinen blev ett fiasko. De som kunde läsa, läste arabiska, vilket man gör från höger till vänster. Det avsedda budskapet med serien tolkades alltså omvänt. Analfabeterna såg tre olika figurer och kunde inte alls förstå det avsedda orsakssambandet. I vår kultur är vi så vana att läsa en serie från vänster till höger att det inte ens faller oss in att detta är ett inlärt tolkningsmönster.

6. FÖRSTÅELSE INOM VETENSKAPEN

Som ett sista tema skall jag jämföra tesen att förstå är att se ett mönster med debatten om vilken roll förståelse har i naturvetenskap. Detta är en mycket omdiskuterad fråga (t.ex. Trout 2002, de Regt 2004 de Regt och Dieks 2005, se även översiktsartikeln av Keil 2006).

Till skillnad från de tidiga analyser av förklaring och förståelse inom vetenskapsfilosofin (t.ex. Hempel 1965, von Wright 1971) är det nu en allmän uppfattning att förståelse och vad som räknas som en förklaring skall beskrivas i förhållande till forskarens kunskapsläge (van Fraassen 1980, Gärdenfors 1980, de Regt 2004). Det innebär att den vetenskapliga förståelsen är kontextberoende. Under den klassiska mekanikens glansdagar är Lord Kelvin känd för att ha sagt att ”frågan ’Har jag, eller inte, förstått ett visst område inom fysiken?’ svarar mot ’Kan jag göra en mekanisk modell av det?’” Men detta förståelseideal inom fysiken har numera förlorat sin lockelse (de Regt och Dieks 2005, s. 138). Man bör också komma ihåg att för Newtons samtida kollegor var hans föreställning om

krafter som verkar på avstånd mystisk, medan i dag knappast någon med en vetenskaplig bakgrund finner krafterna obegripliga.

Man kan grovt dela upp samtida teorier om förståelse inom naturvetenskap i två klasser: kausala och ”unificationist” (de Regt och Dieks 2005). Som ett exempel på en kausal teori, kan vi ta Salmon's bok *Scientific Explanation and the Causal Structure of the World* från 1984. Enligt Salmon, är ”bakomliggande orsaksmekanismer nyckeln till vår förståelse av världen” (Salmon 1984, s. 260), eftersom ”kausala processer, kausala interaktioner och kausala lagar ger de mekanismer som världen fungerar efter; för att förstå varför vissa saker händer, måste jag kunna se hur de produceras av dessa mekanismer” (Salmon 1984, s. 132). Enligt min uppfattning är denna beskrivning av förståelse fullt förenlig med det som jag presenterat här. I avsnitt 3 beskrev jag kausala strukturer som specialfall av mönster som kan upptäckas genom den vetenskapliga processen, eller via mer vardagliga upplevelser.

Det andra alternativet att tolka vetenskaplig förståelse, ”unificationism”, har lagts fram av Friedman (1974) och vidareutvecklats av Kitcher (1981, 1989) bland andra. Friedman skriver att vetenskapen binder samman vår kunskap eftersom den ”ökar vår förståelse av världen genom att minska det totala antalet oberoende fenomen som jag måste acceptera som ultimata eller givna. En värld med färre oberoende fenomen är, givet allt annat, mer begriplig än en med fler” (Friedman 1974, s. 15). Senare har Kitcher argumenterat för mönsters grundläggande roll i en sådan ”unificatory” process: ”Att förstå fenomenen är inte bara en fråga om att minska de ’grundläggande obegripligheterna’, utan att se samband, gemensamma mönster, i det som från början verkade vara olika situationer. . . . Vetenskapen flyttar fram vår förståelse av naturen genom att visa oss hur man skapa beskrivningar av många fenomen, genom att använda samma mönster om och om igen, och genom att visa detta, lär vi oss hur man kan minska antalet typer av fakta som vi måste acceptera som ultimata (eller råa)” (Kitcher 1989, s. 482). Återigen bör det framgå att unificationismens synsätt på vetenskaplig förståelse, särskilt i Kitchers version, passar väl med idén att förstå är att se mönster.

Efter att ha redogjort för de kausala och unificationistiska sätten att se på förståelse i vetenskapen, föreslår de Regt och Dieks (2005, s. 151) ett ”kriterium för förståelse av teorier”: ”En teori T är begriplig för forskare (i kontexten C) om de kan känna igen kvalitativt karakteristiska konsekvenser av T utan att utföra exakta beräkningar.” De Regt och Dieks tillskriver Heisenberg idén bakom detta kriterium. Som exempel på begripliga teorier nämner de den molekylära teorin om gaser, ”potentiell vorticitet” i meteorologi och ”fältlinjer” i elektrostatik med vars hjälp forskare kan göra rent kvalitativa förutsägelser. Det framgår emellertid

av deras beskrivningar att forskarna använder teorierna för att *visualisera* de kvalitativa egenskaperna (de Regt och Dieks betonar detta själva på s. 155). Jag anser därför att även deras kriterium för förståelsen av teorier faller under den allmänna idén om förståelse genom att se mönster. Man skall dock lägga märke till att deras kriterium gör förståelse relativ till den forskare som ser mönstret – det finns ingen absolut förståelse inom en vetenskaplig teori.

7. AVSLUTNING

I denna artikel har jag fokuserat på vilka roller förståelse har i det mänskliga tänkandet. Min centrala hypotes har varit att förståelse består att uppleva ett mönster. Jag har visat hur denna allmänna idé kan tillämpas när det gäller att förstå olika kulturer och förståelse inom naturvetenskap. Men det behövs ytterligare forskning för att få en djupare klarhet om de kognitiva processer och motivationsmekanismer som är involverade i olika former av förståelse. Bland annat saknas psykologiska tester för när förståelse uppstår under en inlärningsprocess. Vi behöver också studera hur förståelse skapar motivation för vidare lärande.

LITTERATUR

- Bruner, J. 1990. *Acts of Meaning*. Cambridge, MA: Harvard University Press.
- Collingwood, R. J. 1972. *Essay on Metaphysics*. Chicago, IL: Gateway.
- De Regt, H. W. 2004. "Making Sense of Understanding". *Philosophy of Science* 71, s. 98–109.
- De Regt, H. W. och D. Dieks. 2005. "A Contextual Approach to Scientific Understanding". *Synthese* 144, s. 137–70.
- Friedman, M. 1974. "Explanation and Scientific Understanding". *Journal of Philosophy* 71, s. 5–19.
- Gärdenfors, P. 1980. "A Pragmatic Theory of Explanation". *Philosophy of Science* 47, s. 404–23.
- Gärdenfors, P. 2000. *Hur Homo blev sapiens: Om tänkandets evolution*. Nora: Nya Doxa.
- Gärdenfors, P. 2006. *Den meningsökande människan*. Stockholm: Natur och Kultur.
- Gärdenfors, P. 2007a. "Understanding Cultural Patterns". I *Learning in the Global Era: International Perspectives on Globalization and Education*, utg. M. M. Suarez-Orozco, s. 67–84. Berkeley, CA: University of California Press.
- Gärdenfors, P. 2007b. "Evolutionary and Developmental Aspects of Intersubjectivity". I *Consciousness Transitions: Phylogenetic, Ontogenetic och Physiological Aspects*, utg. H. Liljenström och P. Århem, s. 81–305. Amsterdam: Elsevier.
- Gärdenfors, P. 2008. "The Role of Intersubjectivity in Animal and Human Cooperation". *Biological Theory* 3:1, s. 1–12.

- Gärdenfors, P. 2010. "The Role of Understanding in Human Nature". I *How Do We Know? Understanding in Science and Theology*, utg. D. Evers, A. Jackelén och T. A. Smeds, s. 3–22. London: T & T Clark.
- Gärdenfors, P. och P. Lindström. 2008. "Understanding is Experiencing a Pattern". I *A Smorgasbord of Cognitive Science*, utg. P. Gärdenfors och A. Wallin. Nora: Nya Doxa, s. 149–164.
- Gärdenfors, P. och M. Warglien. 2012. "Using Conceptual Spaces to Model Actions and Events". *Journal of Semantics* 29, s. 487–519.
- Harnad, S., utg. 1987. *Categorical Perception*. Cambridge: Cambridge University Press.
- Hempel, C. G. 1965. *Aspects of Scientific Explanation and Other Essays*. New York: The Free Press.
- Keil, F. C. 2006. "Explanation and Understanding". *Annual Review of Psychology* 57, s. 227–54.
- Kitcher, P. 1981. "Explanatory Unification". *Philosophy of Science* 48, s. 507–31.
- Kitcher, P. 1989. "Explanatory Unification and the Causal Structure of the World". I *Minnesota Studies in the Philosophy of Science*, Vol. 13, utg. P. Kitcher och W. Salmon, s. 410–505. Minneapolis: University of Minnesota Press.
- Leslie, A. M. 1987. "Pretense and Representation: The Origins of 'Theory of Mind'". *Psychological Review* 94, s. 412–26.
- Marroquin J. L. 1976. *Human Visual Perception of Structure*. Master's degree thesis, MIT Dept. of Electrical Engineering and Computer Science.
- Mitchell, P. 1997. *Introduction to Theory of Mind: Children, Autism and Apes*. London: Arnold.
- Petitot, J. 1989. "Morphodynamics and the Categorical Perception of Phonological Units". *Theoretical Linguistics* 15, s. 25–71.
- Piaget, J. 1930. *The Child's Conception of Physical Causality*. London: Routledge and Kegan Paul.
- Premack, D. och G. Woodruff. 1978. "Does the Chimpanzee Have a Theory of Mind?" *Behavioral and Brain Sciences* 4, s. 515–26.
- Salmon, W. C. 1984. *Scientific Explanation and the Causal Structure of the World*. Princeton: Princeton University Press.
- Runesson, S. 1994. "Perception of Biological Motion: The KSD-Principle and the Implications of a Distal versus Proximal Approach". I *Perceiving Events and Objects*, utg. G. Jansson, S. S. Bergström och W. Epstein, s. 383–405. Hillsdale, NJ: L. Erlbaum Associates.
- Trout, J. D. 2002. "Scientific Explanation and the Sense of Understanding". *Philosophy of Science* 69, s. 212–33.
- Van Fraassen, B. 1980. *The Scientific Image*. Oxford: Clarendon Press.
- Von Wright, G. H. 1971. *Explanation and Understanding*. London: Routledge and Kegan Paul.
- White, P. A. 1995. *The Understanding of Causation and the Production of Action*. Lawrence Erlbaum Associates, Hove.

1. INLEDNING

I sin bok *Self-Constitution* kombinerar Christine Korsgaard Platons och Immanuel Kants teorier om rationellt agentskap. Hennes tes är att en i Platons mening internt enad (*unified*) agent också måste acceptera Kants hypotetiska och kategoriska imperativ, för att uppfylla de villkor om effektivitet och autonomi som hon menar utmärker allt agentskap. Därmed kommer villkor för att vara en agent att sammanfalla med villkor för att vara en moralisk agent. Efter att ha beskrivit Korsgaards resonemang avser jag att visa att det inte är hållbart, eftersom internt enade agenter inte nödvändigtvis måste acceptera Kants kategoriska imperativ, och eftersom det i vissa fall av agentskap kan finnas en motsättning mellan att vara en internt enad autonom agent och att vara en effektiv agent, så som Korsgaard definierar dessa termer.¹

2. KORSGAARDS ARGUMENT

Korsgaards utgångspunkt är att vi vanligen missförstår vad strävan efter att vara moralisk handlar om:

Det moraliska livet inbegriper arbete och ansträngning – en sorts kamp – och de som kämpar framgångsrikt kallar vi ”rationella” eller ”goda”. Men det handlar inte om kampen för *att vara rationell* eller *att vara god*. I stället handlar det om den fortlöpande kampen för integritet, kampen för psykisk enhet, kampen för att, trots psykisk komplexitet, vara en enda enad agent.²

Att vara en enad agent är för Korsgaard ett kreativt projekt, där vi samordnar begär, drivkrafter och förnuft till en enhet. Förslagan för detta samordningsprojekt är Platons resonemang om arbetsfördelningen i den välordnade staten. Denna stat och dess samhällsklasser tillskrivs vissa dygder. Den styrande klassen representerar beslutsklökheten och ser

¹Jag vill tacka min hustru Fulya Tepe för värdefulla uppslag i samband med denna artikels tillkomst.

²Korsgaard 2009, s. 7. Kursiveringarna är Korsgaards.

till hela statens goda, medan väktarklassen representerar modet. Besinningen representeras av den "samklang" som råder i hela samhället när "begären hos den stora massan av enkla människor . . . behärskas av begären och klokheten hos den bättre minoriteten".³ Rättrådigheten, slutligen, består i att varje samhällsklass håller sig till sitt, så att affärsmän och hantverkare inte tar på sig att agera krigare och krigare inte tar på sig att styra staten.

Detta resonemang flyttar Platon sedan över på individen, hos vars själ han identifierar tre delar, nämligen begär, stridbarhet och förstånd. Tanken är att förståndet ska styra stridbarheten och tillsammans ska de tygla begäret, så att det inte "växer sig stort och starkt, vägrar sköta sitt och försöker förslava de andra delarna och styra över dem".⁴ Den rätt-rådiga människan är internt konstituerad som den välordnade staten, så att hennes själs olika delar inte försöker fullgöra varandras uppgifter:

I stället ska den enskilde skapa verklig ordning i sitt eget hus, bli sin egen styresman, ordningsvakt och vän och stämma samman de tre delarna precis som tre toner i en skala, en övre, en undre och en mittemellan . . . och själv träda fram ur mångfalden som en fullständig enhet, besinningsfull och väl sammanfogad . . .⁵

Orättrådighet på individnivå framställs av Platon som "ett slags inbördeskrig . . . då en del av själen gör uppror mot alltihop" och vill styra "fast den inte är lämpad för det, eftersom den tvärtom ska vara slav under den styrande klassen".⁶

Det är sökandet efter denna Platons "fullständiga enhet" som sedan återkommer i Korsgaards argument. Det originella i detta argument är hur hon kombinerar Platons teori om den internt enade agenten med Immanuel Kants teori om förnuftigt och moraliskt agenskap. Kort sagt, Korsgaard hävdar att vägen till det enade agenskapet går via Kants hypotetiska och kategoriska imperativ.

Enligt Korsgaard vägleds allt agenskap av normer om *effektivitet* och *autonomi*. Effektivitet handlar om att agenten måste drivas av en strävan att förverkliga sina mål, och autonomi om att hennes uppträdande (Korsgaard talar om "rörelser") måste vara något självbestämt och ha sitt upphov i henne själv.⁷ Effektivitet svarar då mot Kants hypotetiska imperativ, som föreskriver "en möjlig handlings praktiska nödvändighet som ett medel att uppnå något annat som man vill", medan autonomi svarar mot hans kategoriska imperativ, som föreskriver "en

³Platon 2003, 431c-d.

⁴Ibid., 442a-b.

⁵Ibid., 443d-e.

⁶Ibid., 444b.

⁷Korsgaard 2009, s. 82-83.

handling som objektivt nödvändig för sig själv, utan relatering till något annat syfte”.⁸

En handling som pekats ut av det hypotetiska imperativet är nödvändig i betydelsen att om vi inte väljer den, så kan vi inte lyckas med vårt mål, vilket detta nu än är, och då kan vi inte vara effektiva agenter. En handling som är grundad i det kategoriska imperativet är däremot rationellt nödvändig för varje agent, oberoende av hennes mål, eller, med Kants ord, ”nödvändig för en med förnuftet överensstämmande vilja”.⁹ Det kategoriska imperativet formulerar han, som bekant, i följande termer: ”*Handla endast efter den maxim genom vilken du tillika kan vilja att den blir allmän lag.*”¹⁰ Här får vi också en metanorm för moraliskt handlande, i så måtto att ”om plikt är ett begrepp som skall ha betydelse och verkligen innehålla en lagstiftning för våra handlingar, så kan denna endast uttryckas i kategoriska imperativ”.¹¹

För Korsgaard handlar autonomi just om att ge sig själv en lag för det egna handlandet och därigenom *konstituera* sig som en enad agent i Platons mening. Begär, stridbarhet, och förnuft ska samordnas, och förnuftet ska vara styrande.

Våra böjelser kan tala för ett visst handlingsalternativ, men till skillnad från instinktsstyrda djur, så kan vi människor låta förnuftet avgöra om vi ska handla i enlighet med våra böjelser eller ej. Det är i detta som vår autonomi består: ”Böjelsen presenterar förslaget; förnuftet beslutar om vi ska handla i enlighet med det eller ej, och beslutet antar formen av *lagstiftande handling*.”¹²

Att låta sig styras av sina böjelser skulle vara lika ofritt som att låta sig styras av en annan person, eftersom böjelserna ensamma inte kan utgöra hela agenten. Det är när böjelserna samordnas och kontrolleras av förnuftet som en enad och autonom agent kan uppkomma: ”Det du gör när du överlägger med dig själv om vad du ska göra och sedan gör det, är att du organiserar dina begär, ditt förnuft och din stridbarhet i ett enat system, som ger upphov till en handling som kan tillskrivas dig som person.”¹³

Men varför måste den autonoma agenten vilja se sina handlingsmaximer upphöjda till *allmän lag*, som Kant hävdar? Varför inte nöja sig med att lagstifta för sig själv och se till att man själv alltid agerar i enlighet med förnuftets fordringar? Korsgaards svar är att det inte är någon

⁸Kant 2006, s. 39.

⁹Ibid., s. 40.

¹⁰Ibid., s. 46. Kursiveringen är Kants.

¹¹Ibid., s. 50.

¹²Korsgaard 2009, s. 154. Kursiveringen är Korsgaards.

¹³Ibid., s. 179.

skillnad mellan att lagstifta för sig själv och att lagstifta universellt. När agenten lagstiftar för sig själv, så lagstiftar hon för alla sina framtida jag och hon måste pröva det hon vill göra här och nu mot vad hon rationellt skulle kunna vilja göra i liknande framtida situationer:

Fordringarna för att ena ditt agentskap internt är desamma som fordringarna för att ena ditt agentskap med andras. Att konstituera ditt eget agentskap är en fråga om att välja endast de skäl som du kan dela med dig själv. Det är därför som du måste vilja universellt, eftersom skälet för ditt handlande nu, den lag som du ger dig själv *nu*, måste vara en som du kan vilja handla i enlighet med *senare*, oavsett vad som händer, om du inte kommer till insikt om att det finns något gott skäl att ändra den.¹⁴

Att inte vilja handla i enlighet med en universell lag är att från början offra autonomi till förmån för viljesvaghet och i denna mening faktiskt abdikera från sitt agentskap, enligt Korsgaard. Man bygger in förväntningar på att man framgent inte kommer att göra det man har goda skäl att göra, att man kommer att ändra sig utan att ha goda skäl för det. Det faktum att du inte vill upphöja din maxim till universell lag är, enligt Korsgaard, liktydigt med ”det faktum att du saknar självrespekt”.¹⁵

Och eftersom villkoret för att vara en internt enad och autonom agent sammanfaller med den kantianska metanormen för moraliskt handlande, alltså det kategoriska imperativet, så blir Korsgaards slutsats att det inte är någon skillnad mellan att vara en internt enad agent och att vara en rättrådig agent: ”morallagen är den enade konstitutionens lag, den persons lag som verkligen kan sägas lagstifta för sig själv eftersom han är den person som verkligen har ett jag.”¹⁶

3. TYRANNEN: AUTONOM, MEN INTE NÖDVÄNDIGTVIS INTRESSERAD AV UNIVERSALISERING

Korsgaard antar att en agent som är ”enad” i Platons mening också är moraliskt rättrådig i Kants mening. Men stämmer detta alltid? Känner vi inte från historien och litteraturen flera fall av agenter som är internt enade, i den meningen att deras böjelser inte är invecklade i något inbördeskrig med deras förnuft, men som likväl, och kanske rentav på grund av denna inre enighet, är kapabla till långtgående övergrepp mot andra?

Korsgaard tar själv upp ett sådant fall, nämligen tyrannen. Denne är ingen lekboll för sina impulser och böjelser. Tvärtom uppvisar tyrannen en hög grad av integritet, beredd som han är att ”göra *vadhelst* är

¹⁴ Ibid., s. 202–3. Kursiveringarna är Korsgaards.

¹⁵ Ibid., s. 203.

¹⁶ Ibid., s. 206.

nödvändigt för att få det han vill ha och besluten att inte låta något stå i hans väg”.¹⁷ Korsgaard förnekar inte att tyrannen kan vara effektiv, men väl att han är autonom och därmed också att han är en agent. Tyrannen är i stället, enligt henne, en besatt person som inte kan välja sina handlingar, utan närmast robotaktigt slår sig fram för att tillgodose sina begär: ”Som jag föreställer mig tyrannen, så är hans relation till sin besatthet som en psykotikers till sin vanföreställning: han är beredd att organisera allt annat kring den, även till priset av att han förlorar sitt grepp om verkligheten.”¹⁸

Tyrannen kan inte besluta något om mål och medel, eftersom dessa frågor redan är avgjorda för honom. Han väljer inte några maximer, än mindre upphöjer han maximer till universella lagar ”och det innebär att han inte är autonom, och det innebär att han inte är fri”. Korsgaard tänker sig att tyrannen visserligen kan universalisera i så måtto att han är beredd att säga att alla som tillhör en viss grupp – kvinnor, judar – ska offras, men ”han anser inte att det finns något innehåll, något innehåll över huvud taget, i tanken att han själv kunde ha varit en kvinna eller en jude”.¹⁹

Detta är, med all respekt för Korsgaards filosofi i övrigt, en märklig position som hon här intar. Tyrannen skulle sakna autonomi bara därför att han inte är beredd att universalisera sina maximer fullt ut. För Korsgaard är detta ett tecken på en oförmåga hos tyrannen. Han *kan* helt enkelt inte universalisera sina maximer fullt ut, eftersom hans speciella form av besatthet gör det omöjligt för honom att tänka sig in i rollen av att vara ett offer för dessa maximer. Men är det inte väl så rimligt att tänka sig tyrannen som en person som gör valet att *strunta* i en universalisering av sina maximer? Universalisering implicerar en beredskap till opartiskhet mellan egna och andras skäl för handlande, en benägenhet att se skäl som goda eller dåliga skäl oavsett vems skäl de är, eller vem som berörs av dem. Men tyranner är knappast intresserade av sådan opartiskhet. De är en typ av agenter för vilka ett skäl för handlande är gott om det gagnar deras intressen, dåligt om det inte gör det.

I fallet med tyrannen håller inte Korsgaards tes att ”[f]ordringarna för att ena ditt agentskap internt är desamma som fordringarna för att ena ditt agentskap med andras”. Tyrannen kan mycket väl vara en internt enad agent, med begär och stridbarhet samordnade under det egna förnuftet, utan att han för den skull bryr sig ett dugg om huruvida hans skäl för handlande är universellt acceptabla eller ej.

Och nog är denne tyrann en autonom agent, åtminstone i så måtto

¹⁷ Ibid., s. 170. Kursiveringen är Korsgaards.

¹⁸ Ibid., s. 172.

¹⁹ Ibid., s. 173.

att han ger sig själv en "lag" för sitt handlande. Korsgaards försök att framställa tyrannen som en psykotiker förefaller vara ett något krystat försök att beröva honom hans agentstatus. Om vi i stället tänker oss att tyranner mycket väl kan besitta självkontroll och vara internt enade i Platons mening, men att de för den skull inte behöver bry sig om Kants kategoriska imperativ, så får vi både en mer realistisk bild av tyranner och en klarare bild av vad som brister i Korsgaards resonemang. Det är nämligen hennes ambition att förena Platons idé om den internt enade agenten med Kants kategoriska imperativ som vållar henne problem här.

Platon ger oss en teori om vad som krävs för att vi ska vara *agenter*, nämligen att vi kan utöva förnuftskontroll över våra böjelser. Kant ger oss en teori om vad som krävs för att våra handlingsmaximer ska vara förenliga med *moralen*, nämligen att de låter sig universaliseras. Korsgaard försöker foga samman dessa båda teorier, så att den platonskt enade agenten blir identisk med den kantianskt moraliske agenten. Men i fallet med den tyranniske agenten så fungerar detta inte. Han är amoralisk, inte sällan också omoralisk, men han är fortfarande en agent. När Korsgaard försöker försvara sin syntetisering av Platon och Kant genom att avfärda tyrannen som en icke-agent (eftersom han vägrar att universalisera sina maximer) ter sig hennes resonemang bara verklighetsfrämmande.

4. FANATIKERN: AUTONOM, MEN INTE NÖDVÄNDIGTVIS MORALISK ELLER EFFEKTIV

Implicit i Korsgaards diskussion av tyrannen är att han *skulle* ha varit autonom om han hade varit beredd att universalisera sina maximer fullt ut och accepterat möjligheten att själv bli ett offer för dem. Hade tyrannen varit beredd till en sådan fullständig universalisering, så hade han hört till den kategori av agenter som Richard Hare benämnde *fanatiker*, det vill säga, personer som "har förmåga att föreställa sig, i all levande detaljrikedom, lidandet hos de förföljda, och som ändå kan föreskriva universellt att denna förföljelse ska fortsätta för att tjäna deras ideal, även om det vore de själva som måste lida på detta vis".²⁰

Fanatikern skulle alltså vara en *agent*, enligt Korsgaards teori, men lika lite som tyrannen är han en uppenbart *moralisk* agent. Även om fanatikerns ideal i sig skulle vara rättfärdiga, så är hans benägenhet att offra andra människor (och sig själv) för dem inte moraliskt tilltalande. (Kant, som avvisade tanken att människor kunde reduceras till enbart medel till förverkligandet av något mål, skulle ha delat denna bedömning.)

Fanatikern uppfyller Korsgaards krav på autonomi. Men är han också

²⁰ Hare 1963, s. 184.

en *effektiv* agent i hennes mening av denna term? Just den faktor som gör fanatikern till en autonom agent i Korsgaards ögon, nämligen att han är beredd att universalisera sina maximer fullt ut, riskerar faktiskt att göra honom till en ineffektiv agent. Fanatikerns kompromisslöshet gör honom till en i vissa avseende skör agent, som hellre går under än gör de eftergifter, gentemot sig själv eller andra, som skulle ge honom en chans att åtminstone på sikt förverkliga sina mål. Låt oss illustrera fanatikerns dilemma, med ett litterärt exempel.

I novellen "Regn" från 1921 beskriver Somerset Maugham ett litet sällskap vars fartyg fastnar i karantän i Pago Pago på Amerikanska Samoaöarna. I sällskapet ingår missionärsparet Davidson, samt en doktor Macphail med hustru. Man blir tvungen att ta in på ett primitivt hotell, och där inkvarteras också en ung kvinna, Sadie Thompson, som rest med samma fartyg.

Davidson beskrivs som en "tyst och lite butter" man, vars ögonblick av älskvärdhet var "något som han påtvang sig som en kristlig dygd". Det mest påtagliga med honom var "det intryck av inre, kuvad eld han gav".²¹ Han ger uttryck för kompromisslöshet i sin förestående kamp mot synden på Söderhavsöarna: "Jag tänker handla och jag ska handla snabbt. Om ett träd är ruttet ska det huggas ner och kastas i elden."²² Synden som ska bekämpas är företrädesvis av sexuell art. Det handlar inte minst om de inföddas lättkläddhet och deras dansande.

Sadie Thompson blir i detta avseende en utmaning för Davidson. Hon klär sig om inte utmanande, så åtminstone uppseendeväckande, i vit klänning och vita högklackade stövlar. Dansmusik hörs från hennes rum på hotellet, och det visar sig att hon tar emot män där. Det framgår snart att Thompson är prostituerad. Davidson konfronterar henne. Han går in på hennes rum, slänger hennes grammofon på golvet och skäller ut henne, men blir själv utslängd av Thompsons manlige besökare. Davidsons hustru noterar då om sin man att "när det gäller synd är han skoningslös och när hans rättmätiga vrede väcks är han fruktansvärd".²³ Davidson kontaktar guvernören på ön och får denne att beordra Thompson att lämna ön med nästa båt till San Francisco, där hon har en fängelsedom hängande över sig. Under intryck av detta framtidsperspektiv bryter Thompson samman. En kväll skickar hon bud till Davidson att hon ångrar sina synder och han kommer till hennes rum för att be tillsammans med henne. Detta upprepas de följande dagarna. Davidson är nu "blekare än vanligt, men ögonen lyste av en eld som inte var av denna världen".²⁴

²¹ Maugham 1974, s. 16.

²² Ibid., s. 20–21.

²³ Ibid., s. 29.

²⁴ Ibid., s. 45.

Så väcks Macphail en natt av hotellägaren, som tar med honom till stranden. Där ligger Davidson död. Han har skurit halsen av sig med en rakkniv. Samtidigt har Sadie Thompson genomgått en förändring. Hon har återfått sitt självförtroende, hon har klätat upp sig i sin gamla stass, hennes grammofon är åter igång och hon samspråkar med en sjöman utanför sitt rum. Då hon får se Davidsons hustru spottar hon på henne, och när Macphail förebrår henne, svarar hon: ”Å ni karlar! Snuskiga svin allihop! Den ene är inte en bit bättre än den andre. Svin! Svin!”²⁵ Det står då klart för doktorn, att Davidson har legat med Sadie Thompson, och som en följd av sin kompromisslösa syn på synd, så har han sedan sett sig tvungen att ta livet av sig: ”Om ett träd är ruttet ska det huggas ner och kastas i elden.”

Davidson är, åtminstone fram till dess han faller för frestelsen att ha sex med kvinnan som han tror sig vara skickad att frälsa från synd, en enad agent. Han har underordnat alla sina böjelser och begär under målet att bekämpa synden. Den ”kuvade elden” i hans ögon vittnar om att begären är undertryckta, snarare än utplånade. Märk väl, en enad agent är inte en agent utan begär, utan just en agent som kan kontrollera sina begär. Davidson är också en autonom agent i Korsgaards mening, eftersom han är beredd att universalisera sin skoningslösa syn på synd till att omfatta honom själv. Själva hans självmord blir ett bevis för att han är autonom i denna mening. (Att Kant menade att det kategoriska imperativet uteslöt självmord²⁶ behöver inte bekymra oss här, eftersom vi diskuterar Korsgaards autonomibegrepp, inte Kants.)

Davidson brister emellertid i effektivitet som agent, och han gör det just därför att han lever upp till Korsgaards ideal om en enad agent som universaliserar sina maximer. Här kan man invända, att effektivitet bara handlar om att välja de medel som realiserar ens mål, och att Davidson väl är effektiv i så måtto att han lyckas ta livet av sig när självmordet väl har blivit hans mål. Men effektivitet måste rimligen också ges en tidsdimension. Vi kan vara effektiva på kort sikt eller på lång sikt. Om vår kortsiktiga effektivitet (vi kommer över drogen som vi just nu begär) innebär att vår långsiktiga effektivitet blir lidande (vårt drogberoende kommer att beröva oss familj och arbete, som vi inte vill vara utan), så är vi inte effektiva, eller åtminstone inte så effektiva som vi hade varit om förverkligandet av våra mål här och nu hade varit förenligt med förverkligandet av våra framtida mål.

Fanatiker är sårbara när det gäller effektivitet, eftersom deras sätt att vara enade agenter är att vara kompromisslösa och inte tillåta några känslor eller hänsyn komma mellan dem och deras beslutsamhet att

²⁵ Ibid., s. 52.

²⁶ Kant 2006, s. 47–48.

förverkliga sina ideal. Denna kompromisslöshet gör dem oförmögna att se sina egna begränsningar och att anpassa sitt handlande därefter. Davidson vägrar att se sina egna undertryckta sexuella begär när han närmar sig Sadie Thompson, eftersom det skulle vara att medge en kapacitet för ”synd” hos honom själv. Skulle han se sig själv som en möjlig syndare bland andra, så skulle han förlora den självbild av att vara Guds rättfärdige stridsman som är central för hans sätt att vara en enad agent. Men de sidor av det egna jaget som man förnekar kan man i längden heller inte kontrollera. (Detta är, på ett djupare plan, vad historien om Jekyll och Hyde handlar om, som Mary Midgley har påpekat.²⁷) Därför är Davidson oförberedd på sina egna reaktioner när han går i närkamp med Sadie Thompson. Hans kompromisslöshet gör honom alltså både blind för gliporna i hans rustning och oförmögen att förlåta sig själv när han väl givit efter för sina begär. Hade Davidson varit kapabel att inta en mer förstående attityd till sig själv och till andra, så hade han kunnat låta Sadie Thompson vara i fred och besparat sig själv ett onödigt tidigt och brutalt slut på sitt liv. Han hade kunnat fortsätta att verka som (en något mindre nitisk) missionär och i denna mening kunnat vara en (långsiktigt) effektiv agent.

5. SLUTSATSER

Enligt Korsgaards agentbegrepp eftersträvar varje agent autonomi och effektivitet. Autonomibegreppet utgör för henne förbindelsen mellan *agentskap* i allmänhet och *moraliskt agentskap*. Enligt Platon har den autonome agenten ett internt enat jag, där förnuftet råder över stridbarhet och böjelser. Detta kan vi ses som en teori om agentskap i allmänhet, även om Platon använder ord som ”rättrådig” för att beteckna den internt enade agenten. Här handlar autonomi om den sorts självkontroll som vi normalt förutsätter hos varje agent som kan hållas ansvarig för sina handlingar. Enligt Kant är den autonome agenten en universell lagstiftare, som handlar endast i enlighet med universaliserbara maximer. Detta är en teori om moraliskt agentskap, eftersom Kant relaterar plikt-begreppet till just sådan universaliserbarhet.

Korsgaard tänker sig att en enad agent i Platons mening måste vara en universell lagstiftare i Kants mening, och att en enad agent är en agent som handlar i enlighet med det kategoriska imperativet. Hon låter alltså begreppet om agentskap sammanfalla med begreppet om moraliskt agentskap.

Denna syntes av Platon och Kant håller emellertid inte, vilket framgår

²⁷ Midgley 1984, s. 119–23.

av exemplet med tyrannen, som kan vara internt enad i Platons mening, utan att för den skull acceptera Kants universaliserbarhetskrav. Att, som Korsgaard gör, hävda att tyrannen per definition måste vara oförmögen att vara en enad agent bara för att han inte accepterar kravet på universaliserbarhet ter sig mer som en flykt undan problemet än som en lösning på det.

Exemplet med fanatikern visar att även om vi tänker oss en agent som är både internt enad och autonom, så kan det just på grund av fanatikerns beredvillighet att universalisera sin kompromisslöshet uppstå problem med den effektivitet, som Korsgaard menar att varje agent också måste eftersträva. Givet att effektivitet förstås i en långsiktig mening, som kan fordra vissa kompromisser här och nu, så ter sig den internt enade och autonoma fanatikern som en ”skör” agent, som ständigt riskerar att offra långsiktig framgång för kortsiktig renlärighet.

Slutsatsen blir att Korsgaard's resonemang fallerar på grund av att hon försöker likställa villkoren för agentskap i allmänhet med villkoren för moraliskt agentskap, och på grund av att hon underskattar de motsättningar som kan råda mellan att vara en internt enad agent, en moralisk agent, och en framgångsrik agent.

LITTERATUR

- Hare, Richard. 1963. *Freedom and Reason*. Oxford: Oxford University Press.
- Kant, Immanuel. 2006. *Grundläggning av sedernas metafysik*. Göteborg: Daidalos.
- Korsgaard, Christine. 2009. *Self-Constitution*. Oxford: Oxford University Press.
- Maugham, Somerset. 1974. *Fotspår i djungeln och andra noveller*. Stockholm: Forum.
- Midgley, Mary. 1984. *Wickedness: A Philosophical Essay*. London: Routledge.
- Platon. 2003. *Skrifter*, 3, *Staten*. Stockholm: Atlantis.

Martin Lembke har i några tidigare artiklar argumenterat för en princip, som han menar ger stöd för att *teismen* (dvs. tesen att det finns en personlig gud med vissa övernaturliga egenskaper) kan ge en möjlig förklaring till *varats gåta* (dvs. frågan om varför det finns något överhuvud taget) – och att teismen därmed är kunskapsteoretiskt att föredra framför naturalismen (som inte räknar med något övernaturligt). Den princip han först föreslog, UP, har han numera övergivit och ersatt med en modifierad version kallad ”*Komplement*”.¹

Jag ska inte här diskutera *Komplement* mer ingående. Låt mig bara nämna en sak. Lembke menar att *Komplement* undviker min kritik mot UP. Ett av mina motexempel mot UP, vilket också drabbar *Komplement*, var att vi inte kan förklara att klassen av primtal mellan 2 och 5 är icke-tom med att någon som inte tillhör den klassen ”gör” så att den inte är tom, men att vi ändå har en förklaring, som går ut på att det mellan 2 och 5 finns ”ett positivt heltal [nämligen 3] som inte är produkten av två mindre heltal”. Lembke tycker inte att detta är en förklaring, eftersom den inte ”upplyser” oss, utan bara är ”ett bra svar på frågan vad ett primtal är”.² Jag undrar hur Lembke då själv vill förklara att klassen av primtal mellan 2 och 5 är icke-tom. Eller menar han att detta faktum inte kan förklaras? Jag överlåter åt läsaren att bedöma om Lembke eller jag har rätt på denna punkt.

1. TEISMEN

Hur som helst är Lembke och jag överens om att den intressanta frågan är om *Komplement* utgör en sorts stöd för teismen eller för tesen att Gud skapade världen. Men detta har egentligen inte så mycket med *Komplement* att göra. Det hänger ytterst på om påståendet att teismen kan utgöra en *förklaring* till att världen existerar. Det är det som är frågan. Så låt oss fokusera på just detta.

¹Lembke 2014, s. 24.

²Ibid. s. 29.

Här måste vi först bestämma oss för vad ”teismen” innebär. Som jag tidigare har antytt kan vi skilja på två tolkningar. Antingen säger teismen – som jag nämnde inledningsvis – att det finns en övernaturlig person (en gud) med vissa märkliga egenskaper som allgodhet, allvishet och allsmäktighet. Eller också säger teismen dessutom att denna person skapade världen.³

Med den första tolkningen kan teismen inte förklara att världen existerar.⁴ Med den andra tolkningen kan det verka som en ren trivialitet att teismen innebär en förklaring av att världen existerar – eftersom teismen ju då säger att Gud skapade världen. Men mot detta har jag tidigare invänt att det inte är en förklaring ens då, om vi inte har goda skäl att tro på teismen. Även om teismen skulle vara sann, så skulle det alltså inte räcka.

Lembke tycks hålla med om detta. Men han menar att teismen ändå är en *möjlig* förklaring: om vi skulle få goda skäl att tro att Gud skapade världen, så skulle vi därmed ha en förklaring till att världen existerar. Så oenigheten mellan oss tycks främst handla om vad som kan vara goda skäl för en förklaring.

2. FÖRKLARINGAR

I mitt tidigare inlägg hävdade jag att en teori är en förklaring endast om den är empiriskt prövbar och framgångsrikt prövad. Lembke anser att detta krav är för starkt. Som motexempel anför han att vår tro att det finns en av oss oberoende yttervärld förklaras av att det verkliga finns en av oss oberoende yttervärld.

Vad ska man säga om det? Ja, för det första är det inte så enkelt att vår tro att yttervärlden finns förklaras av att den finns. I så fall kunde man ju analogt säga att det att folk tror att Gud finns förklaras av att hen finns,

³Lembke använder åtminstone ibland termen ”teism” i den andra av dessa betydelser. Som t.ex. när han säger: ”Enligt teismen är ju världslallt skapat av Gud” (s. 30).

⁴Lembke skriver att ”poängen med min argumentation, som baseras på Komplement, är att den kunskapsteoretiskt sett *enda möjliga* förklaringen till varför G' inte är tom är att minst en gud gör så att G' inte är tom. Vad G' s icke-tomhet beträffar har teismen inga rivaler” (s. 32; G' är klassen av icke-gudar). Det kan knappast stämma. Komplement kräver ju bara att *någon*, utanför G' gör så att G' inte är tom. Denne ”någon” behöver ju inte vara en ”gud” i den mening som just ”teismen” räknar med, t.ex. att guden i fråga är god och allsmäktig. Men Lembke uttrycker sig också svagare: ”Min poäng är blott att teismen i detta avseende, till skillnad från naturalismen, har förklaringspotential” (s. 31). Alltså, teismen är ingen förklaring, men den skulle kunna bli (enligt Lembkes uppfattning).

vilket inte är särskilt övertygande.⁵ Vår tro att det finns en yttervärld är ju inbakad i ett nätverk av andra övertygelser om diverse orsakssamband mellan yttervärlden och vår tro på den. Exempelvis orsakssamband mellan händelser i yttervärlden och händelser i våra sinnesorgan, samt mellan händelser i våra sinnesorgan och händelser i våra hjärnor. Vi noterar också likheter mellan våra egna observationer och andras. Och vi förutsäger, ofta framgångsrikt, att andra kommer att instämma i våra observationer. Det är den stora mängden av samband av dessa slag som skulle saknas i ”förklaringar” som hänvisar till att Gud har orsakat vår tro på yttervärlden, eller på att Gud finns, eller på att Gud har skapat världen. Där har vi ingen aning om hur eventuella kausala mekanismer skulle kunna se ut.

För det andra bör man nog tänka sig att förklaringar kan ha olika grader av trovärdighet, men att trovärdigheten inte får vara alltför låg. *Hur* trovärdig en teori ska vara för att verkligen räknas som en förklaring kan det antagligen råda delade meningar om. Men att någon – eller eventuellt många – *uppfattar* den som tillförlitlig kan knappast vara tillräckligt. Några ytterligare krav bör uppfyllas. Men Lembke tycks inte ens kräva att teorin ska vara falsifierbar.⁶ Jag krävde däremot just det, nämligen ”att den kan generera *förutsägelser* om sådant som vi ännu inte känner till”.⁷ Lembke avfärdar detta.⁸ Därför avfärdar han förstås också mitt starkare krav att teorin ska vara ”empiriskt prövbar och framgångsrikt prövad”.

Men för det tredje måste man kanske ändå hålla med Lembke om att det finns acceptabla förklaringar som inte uppfyller detta krav. Jag får intrycket att det Lembke här har i tankarna har att göra med det som brukar kallas *abduktion* eller, på engelska, ”inference to the best explanation”. Han menar förmodligen att teorin att det faktiskt finns en yttervärld oberoende av vårt medvetande är den bästa förklaringen av att vi tror detta – och att vi därmed också har ett gott skäl att tro på teorin (och att teorin därmed

⁵Om någon tycker att detta inte är analogt eftersom det finns många som *inte* tror att Gud finns, så kan man ju tillfoga att det också finns många som inte tror att en av oss oberoende yttervärld finns – även om Lembke anser att detta är en hypotes som är oproblematiskt sann (se s. 30).

⁶Han säger ju t.ex. att teorin att det finns en oberoende yttervärld är en förklaring av att vi tror det, men att ”den inte kan prövas på empirisk väg” (s. 30).

⁷Bergström 2013, s. 60.

⁸Men hans motexempel med Palme-mordet är inte särskilt övertygande. ”Förklaringen till att Palme dog sent på kvällen den 28 februari 1986 är ju att det var då han mördades”, säger Lembke. Men det är väl ingen vidare förklaring. Kanske främst för att de allra flesta som kan tänkas undra varför Palme dog den 28 februari 1986 redan vet att han mördades då. Men om polisen har en teori om varför det hände – en möjlig förklaring – så kan den dra med sig förväntningar (”förutsägelser”) om vad man kan tänkas observera under kommande förhör och brottsplatsundersökningar.

blir trovärdig i tillräcklig grad). Han hänvisar också till antagandet om mörk materia, vilket brukar anses förklara att galaxerna hålls ihop trots sin rotation, och Lembke betvivlar att detta antagande kan generera några verifierbara förutsägelser. Ytterligare ett exempel är antagandet om mörk energi som förklarar att universums expansion accelererar. Dess trovärdighet motiveras enligt Lembke av sitt förklaringsvärde: ”som teoretiskt postulat motiveras den av ett observerbart astronomiskt faktum som annars, såvitt vi förstår, skulle sakna förklaring” (s. 32).

Hur ska man se på detta? Exemplet med yttervärldens existens har jag redan kritiserat. Huruvida exemplen med mörk materia och mörk energi kan hanteras på liknande sätt har jag ingen säker uppfattning om. Jag misstänker att de kan det, men jag är inte tillräckligt kunnig i kosmologi för att vara säker.

Lembke säger också att historieskrivningen tycks vara ”full av förklaringar som inte kan generera några förutsägelser alls”. Ja, men det hänger lite på vad man menar med ”förutsägelser”. Dels kan sådana förklaringar få ökat stöd av ytterligare iakttagelser som de också kan förklara, men som inte gäller framtida händelser. Dels kan de prövas genom att man undersöker om man i andra, liknande fall kan ha fått liknande resultat. Och dels kan de ibland ”generera förutsägelser om sådant som vi ännu inte känner till”. Att Hitler skrev det han skrev om judar i *Mein Kampf* kan t.ex. förklaras av att han var antisemit, vilket i sin tur kan generera förväntningar om hur han kan komma att bete sig om han får politisk makt.

3. TEISMEN SOM ”FÖRKLARING”

Men det är ändå en väldig skillnad mellan dessa teorier i fysik och historia å ena sidan och teismen å den andra. Att universums expansion accelererar kan vara ett gott skäl att tro att det finns mörk energi – vad nu det är – men att världen existerar kan inte på liknande sätt vara ett gott skäl att tro att Gud finns. (Det är just därför teismen *inte* är en förklaring av världens existens.) Lembke säger inte heller direkt att dessa fall är jämförbara, men man får ändå intrycket att det är det han har velat antyda.

Troligen vill han tillämpa idén om ”inference to the best explanation”. Om denna kombineras det faktum att världen existerar samt med Lembkes rätt lättvindiga förmodan att teismen är *den enda möjliga* förklaringen av detta,⁹ så kan man tyckas ha ett gott skäl för teismen. Men även om

⁹Han skriver: ”Men poängen med min argumentation, som baseras på Komplement, är att den kunskapsteoretiskt sett *enda möjliga* förklaringen till varför G' inte är tom är att minst en gud gör så att G' inte är tom” (s. 32). Men det kan knappast stämma. En annan förklaring vore t.ex. att en vindögd, röd, illvillig och nödvändigt existerande orm gör så att G' inte är tom.

man skulle gå med på att den enda möjliga förklaringen till att världen existerar är att Gud har skapat den – och att detta alltså i en mening är den ”bästa” av alla tillgängliga förklaringar – så kan man invända att den inte är *tillräckligt* bra.

Det som Lembke däremot har rätt i är att teismen (till skillnad från naturalismen) är en *potentiell* förklaring till att världen existerar – i den meningen, nämligen, att *om* vi, mot all förmodan, skulle få goda skäl att tro att Gud skapade världen, så skulle vi därmed ha en förklaring till att världen existerar. Detta är närmast en ren truism.

Men en sådan förklaring leder tyvärr till nya problem, nämligen att hitta förklaringar till *varför* och *hur* Gud skapade världen. Dessa problem uppstår lyckligtvis inte för naturalismen.

LITTERATUR

- Bergström, Lars. 2013. ”Teismen som förklaring”. *Filosofisk tidskrift* 34, nr 4, s. 57–61.
- Lembke, Martin. 2012. ”Omnipotence and Other Possibilities”. *Religious Studies* 48, nr 4, s. 425–43.
- Lembke, Martin. 2013. ”Varat: en fråga om uteslutning”. *Filosofisk tidskrift* 34, nr 2, s. 45–53.
- Lembke, Martin. 2014. ”Mer om allt”. *Filosofisk tidskrift* 35, nr 4, s. 23–32.

Epistemiska trösklar, ett svar på problemet med gemensam kunskap

1. INTRODUKTION

Debatten kring kollektiv intentionalitet handlar om hur vi ska förstå hur två eller flera individer kan agera som en grupp. I debatten finns det fyra etablerade teorier Bratman (1999), Gilbert (1989), Searle (2010) och Tuomela (1995). Gemensamt för alla teorier är att de anser att *gemensam kunskap* (eng. common knowledge) är ett nödvändigt villkor för kollektiv intentionalitet. Även Searle som i tidigare texter (t.ex. Searle 1995) hävdade att gemensam kunskap inte alls är nödvändigt för kollektiv intentionalitet har ändrat sig: ”När jag talar om denna form av kollektiv intentionalitet talar jag om människors och andra djurs förmåga att faktiskt *samarbeta* i sina aktiviteter. Samarbete *implicerar* gemensam kunskap” (Searle 2010, s. 49, författarens betoning).

Gemensam kunskap är en teknisk term i filosofi och logik som beskriver att ett faktum är epistemiskt öppet (allmänt tillgängligt) i en grupp. Traditionellt har gemensam kunskap definierats som *jag vet att du vet att jag vet ... att du vet*. Men det här leder till problemet att det krävs en extremt hög kognitiv förmåga för att kunna ingå i gemensam kunskap. Det här skulle innebära att åtminstone barn inte skulle kunna ingå i en sådan situation, men det finns mycket psykologisk forskning som pekar på att de både kan ingå i situationer av kollektiv intentionalitet och epistemisk öppenhet. Jag kommer i den här artikeln att presentera en psykologiskt rimligt teori om hur barn kan delta i epistemisk öppenhet.

2. GEMENSAM KUNSKAP OCH EPISTEMISK ÖPPENHET

För att illustrera hur ett faktum kan vara epistemiskt öppet kan vi föreställa oss en arbetsplats där alla anställda vet att chefen kommer att komma sent till morgonmötet, men ingen vet att alla andra också vet det. Det här sakförhållandet kallas ömsesidig *kunskap* (eng. mutual knowledge). Däremot om någon ropar ut att de precis har pratat med chefen och att hen hälsar att hen kommer bli sen, så ändras situationen från ömsesidig kunskap till gemensam kunskap att chefen kommer bli sen. Det som har

ändrats är att faktumet att chefen kommer att bli sen är nu öppet och offentligt bland de anställda.

Motivationen bakom gemensam kunskap brukar ligga i att förklara olika typer av gemensam kunskaps problem. Ett sådant problem föreligger t.ex. vid koordinerat handlande.

Vi kan tänka oss ett kärnkraftverk som behöver stängas ner p.g.a. en brand. För att stoppa reaktorn så finns det två personer Olof och Lisa som måste sätta in varsin nyckel i varsitt nyckelhål. Av säkerhetsskäl så är nyckelhålen på olika sidor av reaktorn, dessutom som en säkerhetsåtgärd om bara den ena nyckeln sätts in och inte den andra inom en minut så kommer nyckelhålet att sluta fungera under en timma. P.g.a. eldsvådan så kan Olof och Lisa inte nå varandra och planera när de ska sätta i nycklarna. Som tur är finns det skärmar bredvid nyckelhålen så de kan skicka meddelande till varandra. Olof och Lisa upptäcker dessvärre att skärmarna har blivit trasiga av eldsvådan till den grad att endast 50 procent av deras kommandon/knapptryckningar fungerar. Däremot så lyckas varje meddelande skickas, vilket de inte vet om. Olof skickar till Lisa att de borde sätta in nyckeln vid tidpunkt t. Olof vet inte om Lisa har fått meddelandet vilket Lisa inser, så hon skickar tillbaka ett bekräftelsemeddelande. När Olof får Lisas meddelande inser han att Lisa inte vet om han hade fått bekräftelsemeddelandet och han vet att Lisa inte kommer agera vi om hon inte är säker på att Olof har fått meddelandet. Så Olof skickar en bekräftelse på bekräftelsemeddelandet och så vidare och så vidare. Det här är problemet med gemensam kunskap uttryckt i ett koordineringsexempel.

David Lewis var den första att definiera gemensam kunskap och hans definition ses än idag som standarddefinitionen. Han började med att karakterisera gemensam kunskap som en oändlig iteration av jag vet att du vet att jag vet etc. Lewis insåg att det här var ett problem då det inte finns någon människa som kan ha en oändlig kedja av propositioner i huvudet eller ha ett oändligt antal mentala tillstånd samtidigt. För att kunna hantera det här så gör han en skillnad på *faktiska trosföreställningar* och *skäl att tro* (*actual belief* och *reason to believe*). Han definierar gemensam kunskap som följande:

[D]et är *gemensam kunskap* i en population P att X om och endast om något sakförhållande A är fallet, på så vis att:

- (1) Alla i P har skäl att tro att A är fallet.
- (2) A indikerar för alla i P att alla i P har skäl att tro att A är fallet.
- (3) A indikerar för alla i P att X. (Lewis 1969, s. 56, författarens betoning)

Lewis menar att även om denna teori leder till en oändlig regress så är det en regress av implikationer av skäl att tro och inte av faktiska tros-

föreställningar. Men även en oändlig kedja av implikationer kräver en viss nivå av kognitiv kapacitet. Lewis förstår det och fortsätter med att säga:

Resten av vad de behöver är det som de behöver för att bilda högre ordningens förväntningar på det sätt vi behandlar: ömsesidig tillskrivning av några allmänna induktionsstandarder och bakgrundsinformation, rationalitet, ömsesidigt tillskrivning av rationalitet och så vidare. (Lewis 1969 s. 56–57)

För att någon ska kunna delta i gemensam kunskap så behöver åtminstone dessa kriterier uppfyllas. Från ett utvecklingspsykologiskt perspektiv är Lewis teori, och alla andra teorier om gemensam kunskap, väldigt otillfredsställande. Småbarn verkar exempelvis inte uppfylla dessa kriterier, men det finns flera studier som visar på att barn redan vid 1 års ålder förstår vad som är epistemiskt öppet i en grupp och vad som inte är det. Jag ska ge två exempel. I en studie på spädbarn inte mer än tolv månader gamla av Liszkowski, Carpenter och Tomasello från 2008 så undersökte de spädbarns pekande efter objekt när den vuxna letade efter något. De fann att om barnet hade sett att den vuxna inte hade sett objektet ramla ner så pekade de ut objektet för den vuxna, till skillnad från när barnet hade sett att den vuxna hade sett objektet ramla ner. Den andra studien från samma år gjord av Moll, Richter, Carpenter och Tomasello undersökte hur fjortonmånaders barn förstod referens baserad på tidigare erfarenhet med objekten och den vuxna. I studien fick barnet leka med en vuxen och tre leksaker. Den vuxna visade stor entusiasm för den ena leksaken. Efter en stund så frågar den vuxna om hen kunde få en leksak, varpå hen fick den hen hade visat entusiasm över. I kontrollstudien kom en annan vuxen som barnet kände in och frågade om hen kunde få en leksak och då var sannolikheten att det skulle vara den speciella leksaken lika med slumpen. De här två studierna visar på att barn förstår vilka fakta som är öppna mellan dem, men det skulle fortfarande vara orimligt att mena att ett tolv månaders barn skulle uppfylla kraven för gemensam kunskap som nämnts ovan.

3. PROBLEMET MED EN AVKORTAD REGRESS

Av de skäl som ovan nämnts anser många att diskussionen kring gemensam kunskap är en abstrakt filosofisk fråga som inte har bäring på verkligheten. Många forskare har till och med hävdade att vi inte behöver den här långa kedjan av propositioner. Tomasello menar, i sin bok *Origins of Human Communication* från 2008, att vi inte ska se regressen som oändlig, utan snarare som odefinierad. Men denna typ av avkortningsteorier burkar ofta underskatta, eller missförstå motivationen till, problemen med gemensam kunskap. Michael Wilby formulerar det så här:

Problemet med [gemensam] kunskap kommer inte från något förmodat glapp mellan *psykologiska begränsningar* hos mänskliga agenter å ena sidan och en *idealiserad uppfattning* om vad [gemensam] kunskap skulle bestå å den andra, utan snarare från ett glapp mellan *psykologiska begränsningar* hos mänskliga agenter å ena sidan och behovet av en *psykologiskt ändamålsenligt uppfattning* om vad [gemensam] kunskap består i å den andra. (Wilby 2010, s. 86–87, författarens betoning)

För att förstå vad Wilby syftar på här så kan vi återgå till kärnkraftverks-exemplet. Wilby menar att om avkortningsteorierna stämmer så kan vi givet någon nivå k i regressen säga att $k + 1$ är tillräcklig för gemensam kunskap mellan Olof och Lisa. Det här innebär att Lisa som skickar $k+1$ meddelandet kommer sätta in nyckeln vid tidpunkt t utan att ha fått något bekräftande meddelande. Men eftersom vi vet att de bara kan agera tillsammans, så innebär det att Lisa kommer agera oavsett om meddelandet kommer fram eller inte. Men då är $k + 1$ meddelandet överflödigt och det skulle då vara tillräckligt med meddelande k som Olof skickade. Dock leder det här till att hela regressen blir meningslös eftersom meddelande k skulle kunna vara det första meddelandet som Olof skickade och vi skulle inte anse att vi då har någon form av epistemisk öppenhet.

Jag tror dock inte att det är kört för en avkortningsteori och jag kommer senare att lägga fram ett förslag på en sådan teori, efter vi ha sett vilka andra typer av teorier det finns i debatten.

4. ETT RELATIONELLT ALTERNATIV

John Campbell (2005) menar att han har kommit på en lösning av vårt problem, att hitta en teori som både är psykologiskt rimlig och som klarar av problemen med gemensam kunskap som vårt koordinerings-exempel. Campbell pratar inte om gemensam kunskap utan om *delad uppmärksamhet* (eng. joint attention). Men det han säger går att applicera på gemensam kunskap, något Wilby (2010) gör. Campbell skiljer mellan reducerande och relationella analyser av gemensam kunskap. Enligt den reducerande analysen ska vi kunna beskriva de mentala tillstånd Olof är i, utan att implicera att han står i en gemensam kunskapsrelation med Lisa. Lewis teori om gemensam kunskap är enligt denna uppdelning en reducerande teori. En relationell analys, å andra sidan, menar att en analys av Olofs mentala tillstånd implicerar att det finns någon, Lisa, som Olof har gemensam kunskap med. Enligt den relationella analysen är gemensam kunskap ett primitivt fenomen hos medvetandet. Både Wilby och Campbell menar att vi aldrig kommer att kunna lösa ett koordinationsproblem med en reducerande teori. De menar att antingen kommer en sådan teori att vara psykologiskt orimlig eller också kommer den att

skjuta fram problemet ett steg. Exempel på det senare skulle vara, att om vi accepterar Lewis teori som rimlig, så har vi fortfarande en oändlig kedja av implikationer. Man skulle så klart kunna säga att dessa implikationer kan vara implicita hos Olof och Lisa i stället för explicita; om de t.ex. explicit vet att X och explicit vet att om X så Y, så vet de Y implicit. Men då uppstår problemet att de behöver vara säkra på att de i den här situationen får sluta sig till Y och att den andra gör samma sak. Härmed skjuts problemet bara fram ett steg och bekräftelsemeddelandena kommer snarare att handla om de är berättigade till att använda sin implicita kunskap.

I Campbells fall när två individer, Per och Stina, har delad uppmärksamhet, så innebär en relationell analys av Pers mentala tillstånd att Stina är konstituerande för Pers perceptuella medvetande som deltagare (eng. co-attender), utan att vara ett objekt för medvetandet.

Dock verkar den relationella förklaringsmodellen inte fungera, eller rättare sagt den förklarar inte det den ska förklara. Relationella teorier, som Campbell och Wilbys, menar sig förklara hur *vi* kan ha en specifik kunskap *ihop*. Men varifrån kommer det här *vi ihop*? De säger att *vi* är ett primärt och primitivt fenomen. Men vad betyder det egentligen? Det verkar betyda att Olof är berättigad att säga att han och Lisa konstituerar ett *vi ihop*, eller tillsammans vet att de ska agera vid tidpunkt *t*, på samma grunder som han är berättigad att säga att det finns en sten framför honom om han ser en sten. I det senare fallet så skulle vi ju säga att Olof är berättigad att tro att det är en sten framför honom om det verkar vara en sten där. Det skulle innebära att Olof anser att han och Lisa konstituerar ett *vi* endast om det verkar som om Lisa gör det också. Visserligen är det viktigt, till och med ett nödvändigt villkor som vi kommer se nedan, att något verkar vara öppet för att det ska vara öppet. Men det kan inte vara hela historien, att det verkar vara så gör att det är så. I exemplet med arbetsplatsen så räcker det inte med att säga att det verkar som om alla vet att chefen kommer bli sen, det behövs något mer. Den relationella analysen verkar inte kunna förklara det vi vill att den ska förklara. Jag kommer nu försöka återuppliva både den reducerande analysen och den avkortande versionen av en sådan analys för att kunna få fram en tillfredställande teori om gemensam kunskap.

5. ÖPPENHET OCH EPISTEMISKA TRÖSKLAR

Till att börja med tar vi det vi lärde oss från den relationella analysen, att något verkar vara fallet, som väsentligt. Om vi pratar om epistemisk öppenhet så verkar det vara nödvändigt att det verkar vara så att det är öppet, för att något ska vara öppet. Om vi lägger till det till vår formu-

lering av problemet så får vi; om det är ömsesidig kunskap att chefen är sen och det *verkar* som faktumet att chefen är sen är öppet för alla, så är det gemensam kunskap bland arbetarna att chefen kommer bli sen. I det här fallet, om vi lägger till att det verkar vara öppet till faktumet att alla redan visste det, så verkar vi vara berättigade till att sluta oss till att det är öppet. Det intressanta är att något kan verka publikt eller öppet innan alla faktiskt vet om det, däremot måste alla veta om det för att det i sin tur ska bli gemensam kunskap. Vi kan tänka oss en situation där Stina säger till Per att han verkar vara arg. Det kan till och med vara så att Per inte vet om att han är arg och först genom att Stina pekar ut det upptäcker att han är arg. Det är först då Pers ilska blir gemensam kunskap mellan Stina och Per. Andra tillfällen kan det vara så att Per upptäcker att han verkar arg men helt enkelt inte är det varpå det inte blir gemensam kunskap.

Av det här kan vi definiera vad jag kallar för *semi-öppet*: B är semi-öppen för A om intention $X =_{df}$ det verkar/ser ut för A som att B är öppen med sin intention att X.

Öppen med sin intention innebär helt enkelt att man ger en publik indikation för något som nämnts ovan. Med det här kan vi beskriva fenomenet att man kan *ställa in* sin uppmärksamhet efter någon annans uppmärksamhet. Vi kan säga att om Lisa är medveten om att Olof är medveten om X och Lisa är semi-öppen med att vara medveten om detta för Olof, så tror Olof att deras uppmärksamhet är *inställd* efter X. Om samma sak gäller för Lisa så får vi en situation där både Olof och Lisa tror att de har korrigerat sin uppmärksamhet efter varandra till att de båda är inställda på att uppmärksamma X. Det här leder till att X verkar vara öppet för dem båda.

För att se hur det här kan hjälpa oss få gemensam kunskap så behöver vi analysera koordineringsexemplet igen. En intressant sak med exemplet är att innan Olof skickar första meddelandet, så har varken Olof eller Lisa någon anledning att agera vid tidpunkt t . Vi kan beskriva det här som att deras individuella subjektiva sannolikhet för att agera vid tidpunkt t utan någon evidens eller skäl att tro så, är lika med noll. Vi kan formellt skriva ut det som att den subjektiva sannolikheten $P(\text{agera vid tidpunkt } t) = 0$. Den enda evidens som kan ändra deras subjektiva sannolikhet är, om det verkar som den andra kommer agera vid tidpunkt t . Om exempelvis Olof anser sig tillräckligt berättigad att anse det, så kommer han att agera, det samma gäller för Lisa.

När Olof skickar sitt första meddelande, vi kan kalla det M_1 , så får vi $P_{\text{Olof}}(\text{agera vid tidpunkt } t | M_1) = x_1$, där $0 < x_1 \leq 1$. P_{Olof} betyder Olofs subjektiva sannolikhet. Även om Olof inte vet om meddelandet kommer fram, så kommer själva akten av att skicka meddelandet att öka hans subjektiva sannolikhet dvs. $0 < x_1 \leq 1$. x_1 ligger här emellan p.g.a. att Olof är

osäker på om hans intention att agera vid tidpunkt t är öppen eller inte. Men hans subjektiva sannolikhet är högre nu, eftersom det *kan* vara så att meddelandet kommit fram.

När Lisa får meddelandet så händer samma sak för henne, dvs. hennes subjektiva sannolikhet för att agera vid tidpunkten t ökar. När hon skickar iväg sin bekräftelse, kalla det M_2 , så kommer även denna handling också att öka sannolikheten. Vi får då att $P_{\text{Lisa}}(\text{agera vid tidpunkt } t | M_1 \& M_2) = x_2$, där $0 < x_2 \leq 1$ och $x_1 < x_2$.

Något väldigt intressant händer när Olof får bekräftelsemeddelandet, nämligen, att genom själva meddelandet att Lisa vill agera vid tidpunkten t så ökar hans subjektiva sannolikhet, men även hans tidigare syn på M_1 ändras då han nu vet att det faktiskt kom fram. Vi får nu att $P_{\text{Olof}}(\text{agera vid tidpunkt } t | M_1 \& M_2) = x_2$, där $0 < x_2 \leq 1$ samt $x_1 < x_2$ och så vidare.

Givet premisserna för exemplet kommer Olof och Lisa aldrig att nå visshet dvs. det händer aldrig att $P(\text{agera vid tidpunkt } t) = 1$. Jag tror dock inte att det är nödvändigt. I stället vill jag postulera att alla agenter har en *subjektiv tröskel*, som bedömer om något är epistemiskt öppet eller inte. I det här exemplet så var det just att något verkade öppet, som var det enda som kunde ändra deras subjektiva sannolikheter och om tröskeln blir nådd, så kommer de att agera. Den brännande frågan är egentligen om de kommer att sätta in nycklarna när tidpunkten t kommer? Om de gör det, gör de det inte de då p.g.a. att det att agera vid tidpunkten t verkade öppet för dem?

Vi vet att efter ett par meddelandeväxlingar så vet båda att den andre vet om den andras planer att agera vid tidpunkt t och båda vet att den andra har bekräftat detta. I termer av semi-öppenhet: det verkar öppet för båda två att den andra är öppen med sin intention att agera vid tidpunkt t och att denne är medveten om att den andra är öppen med samma intention. Om vi utgår från att det finns en subjektiv tröskel, låt oss säga 0,7 för Olof och 0,9 för Lisa, så gäller att om de har nått dessa trösklar genom en ökning av sina sannolikheter baserad på sin kommunikation, så verkar det rimligt, baserat på deras standard för att agera, att de just agerar.

Om vi accepterar existensen av en subjektiv tröskel så kommer vi få en version av både den reducerande analysen och den avkortade teorin att fungera. Den här typen av teori är dessutom kompatibel med sannolikhetsteorier om hur hjärnan hanterar och interagerar information, t.ex. fri energi-principen (Friston, 2009), som argumenterar för att hjärnan försöker minimera överraskningar. Med en sådan läsning så kommer den subjektiva upplevelsen av gemensam kunskap vara till för att skapa förväntningar på vad andra kommer att göra. Om Olof och Lisa skulle fortsätta att skicka ett par meddelanden efter att deras tröskelnivåer har nåtts så skulle det vara för att minimera möjligheten att bli överraskade.

Den här tröskelteorin möjliggör självklart att Olof och Lisa har olika tröskelnivåer, som i exemplet ovan. Vi kan tänka oss att Olofs når sin tröskel men inte Lisa, vilket innebär att Olof agerar men inte Lisa. Det här är såklart något som kan hända och jag anser att en teori om gemensam kunskap även behöver kunna förklara varför det kan bli fel.

Det här är som sagt en reducerande analys av gemensam kunskap och en avkortningsteori, då jag menar att efter ett par nivåer så kommer tröskeln att vara nådd för dem båda. Teorin faller därför inte för Wilbys kritik av avkortningsteorier, då den kräver att det blir några iterationer, men att det är subjektivt för vardera agenten hur många som behövs givet kontexten. Dessutom så klarar den här teorin av att vara både psykologiskt rimlig och uttrycka gemensam kunskap i psykologiskt ändamålsenliga termer, utan att det blir ett glapp. Dessutom är frågan om det finns en tröskel en rent empirisk fråga som kan leda till flertalet intressanta studier.

6. SLUTSATS

Vi började med att se att alla etablerade teorier inom kollektiv intentionalitydebatten anser att gemensam kunskap är nödvändigt för att två eller flera personer ska kunna samarbeta. Vi såg sedan att som gemensam kunskap brukar definieras, så blir teorierna antingen psykologiskt orimliga eller så skjuter de bara upp problemet ett steg. Mitt förslag på hur vi ska analysera gemensam kunskap tycks hantera dessa problem samtidigt som det lyckas förklara den epistemiska öppenheten som de relationella teorierna smyger in i sin analys. Teorin om subjektiva trösklar leder dessutom till en del intressanta empiriska frågor, t.ex. om det finns en sådan tröskel, hur den utvecklas hos barn och hur tröskelnivåerna skiljer sig hos introverta och extroverta individer?¹

REFERENSER

- Bratman, M. 1999. *Face of Intention: Selected Essays on Intention and Agency*. Cambridge: Cambridge University Press.
- Campbell, J. 2005. "Joint Attention and Common Knowledge". I *Joint Attention: Communication and Other Minds*, utg. N. Eilan, C. Hoerl, T. McCormack och J. Roessler, s. 287–97. Oxford: Oxford University Press.
- Friston, K. 2009. "The Free-Energy Principle: A Rough Guide to the Brain?", *Trends in Cognitive Neuroscience* 13, nr 7, s. 293–301.
- Gilbert, M. 1989. *On Social Facts*. London: Routledge.

¹Tack till William Hedly Thompson för litteraturtips och kommentarer på en tidigare versioner av uppsatsen samt till Niklas Rydberg som har stått för översättning av citaten.

- Lewis, D. 1969. *Convention: A Philosophical Study*. Cambridge: Harvard University Press.
- Liszkowski, U., M. Carpenter och M. Tomasello. 2008. "Twelve-Month-Olds Communicate Helpfully and Appropriately for Knowledgeable and Ignorant Partners". *Cognition* 108, s. 732-39.
- Moll, H., N. Richter, M. Carpenter och M. Tomasello. 2008. "Fourteen-Month-Olds Know What 'We' Have Shared in a Special Way". *Infancy* 13, s. 90-101.
- Searle, J. 1995. *Konstruktionen av den sociala verkligheten*. Göteborg: Daidalos.
- Searle, J. 2010. *Making The Social World*. Oxford: Oxford University Press.
- Tomasello, M. 2008. *Origins of Human Communication*. Cambridge: The MIT Press.
- Tuomela, R. 1995. *The Importance of Us: A Philosophical Study of Basic Social Notions*. Stanford: Stanford University Press.
- Wilby, M. 2010. "The Simplicity of Mutual Knowledge". *Philosophical Explorations* 13, nr 2, s. 83-100.

Ingvar Johansson

Kahneman och Lindaproblemet

Bestsellern Daniel Kahneman, *Tänka, snabbt och långsamt* (2013) recenserades i *Filosofisk tidskrift* 2/2014 av Lars Bergström, och jag instämmer i de kritiska reflektioner han ibland gör. Men jag vill komplettera med kritik av Kahnemans berömda Lindaproblem, vilket Kahneman nu sammanfattar sin syn på (kap. 15). Diskussionen kring problemet finns för övrigt kortfattat presenterat i Wikipedia under rubriken ”Conjunction fallacy”, men min typ av kritik finns inte redovisad där.

Se på följande fråga: ”Vilket alternativ är mest sannolikt? (a) Linda är bankkassörska. (b) Linda är bankkassörska och verksam i feministrörelsen.”

Om jag skulle hittat frågan i en bok i sannolikhetslära, så skulle jag omedelbart ha svarat (a). Detta därför att jag vet att sannolikheten för en konjunktion (”Linda är bankkassörska *och* verksam i feministrörelsen”) enligt den matematiska sannolikhetskalkylen alltid är lägre än sannolikheten för någon av konjunkterna tagna för sig (t.ex. ”Linda är bankkassörska”). Men när jag i en offentlig lite interaktiv föreläsning första gången mötte frågan, så svarade jag, i likhet med många andra som kan sin sannolikhetslära, att alternativ (b) är det mest sannolika. Hur kommer detta sig?

Kahneman säger att mitt och andra sannolikhetskunniga personers val av (b) ”strider mot all logik” (s. 179). Han säger också att ”Tyvärr brydde vi oss aldrig om att ta reda på hur den stora minoritet (36 procent) av denna välinformerade grupp resonerade som valde fel alternativ” (s. 180). Den förklaring av mitt förment ologiska svar som jag själv dagen efter föreläsningen kom fram till ser ut som följer.

Till saken hör att i föreläsningen, liksom i Kahnemans experiment, föregicks frågan av följande information: ”Linda är trettioett år, singel, frispråkig och mycket intelligent; har examen i filosofi, och brunn som student för frågor om diskriminering och social rättvisa och deltog även i demonstrationer mot kärnkraft” (s. 177).

Det finns utanför böcker i sannolikhetslära åtminstone två olika rimliga preciseringar av påståendet ”Linda är bankkassörska”:

- (a₁) Linda tillhör mängden bankkassörskor.
 (a₂) Linda är en ganska typisk bankkassörska.

Låt oss nu se på sannolikheterna i de två fallen. Om $(a) = (a_1)$, så är det sannolikare att (a) är fallet än att (b) är det; och till yttermera visso är detta en logisk-matematisk sanning. Frågan som jag ställer är om detta i den givna situationen är den enda logiskt korrekta slutsatsen, och svaret menar jag måste bli nekande.

Om $(a) = (a_2)$, så är det tvärtom sannolikare att (b) är fallet än att (a) är det. Givet informationen om Linda, så det sannolikare att (b) hon är bankkassörska och verksam i feministrörelsen än att (a₂) hon är snarlik stereotypen för en bankkassörska. Man kan inte bedöma om personer drar logiskt korrekta slutsatser förrän man vet hur de uppfattar premisserna. Som jag uppfattade premisserna gjorde jag en helt korrekt logisk slutledning.

Som med all önskvärd tydlighet framgår av Kahnemans bok är jag i sällskap med många personer, varav åtskilliga säkert har mer sannolikhetslära i ryggmärgen än jag har. Något mer allmänt behöver därför sägas.

I kognitionsvetenskap och delar av lingvistik finns det många forskare som skriver under på den s.k. prototyp teorin. De mest namnkunniga tidiga företrädarna är Eleanor Rosch och George Lakoff. Peter Gärdenfors, för att nämna en svensk, har arbetat med att väva in antagandet om prototyper i en vidare teoretisk ram där prototyperna representeras av punkter i vad han kallar begreppsrum; se hans *Conceptual Spaces* (2000) och *The Geometry of Meaning* (2014). Enligt prototyp teorin förstår vi en mängd begrepp inte via en implicit hänvisning till en uppsättning nödvändiga och/eller tillräckliga villkor för begreppets applicering, utan primärt som implicit hänvisande till en prototyp.

Jag har långt före jag stötte på Lindaproblemet trott att det ligger mycket i prototyp teorin, vilket kanske t.o.m. var det som gjorde att jag så snabbt fann förklaringen till mitt skenbart felaktiga svar. Enligt teorin uppfattar vi normalt begreppet ”bankkassörska” som hänvisande till ”typisk bankkassörska”, vilket jag ju gjorde. Slutsats: Kahneman borde ha tänkt mer på hur vardagsspråket kan fungera innan han började beskylla sannolikhetskunniga personer för att tänka helt ologiskt.

Märkligt nog finns i Kahnemans bok inte en enda hänvisning till prototyp teorin. Detta trots att han drar in begreppet om prototyper i sina resonemang, och t.o.m. har indexerat ordet prototyp. Specialiseringen i forskarsamhället får ibland lite konstiga konsekvenser.

Som jag redan citerat, så skriver Kahneman i anslutning till Lindaproblemet att han inte brydde sig om att ta reda på hur de som enligt honom

tänkte ologiskt egentligen tänkte. Men det låter lite underligt, eftersom hela hans bok är ett försök att identifiera och klassificera olika mekanismer bakom mänskligt vardagligt feltänkande. Och det är faktiskt lätt att rekonstruera hur han ser på det förmenta misstag jag och många andra med mig gjort. Han säger för övrigt i förbifarten senare i boken att hans förklaringsmodell kan appliceras på Lindaproblemet (s. 429). Såvitt jag förstår menar han så här.

För att på ett enkelt sätt förstå många av våra felslut kan vi betrakta vårt intellekt *som om* det består av två åtskilda system. Det första, System 1, fungerar snabbt och okritiskt; ofta fungerar det bra, men förhållandevis ofta också fel. System 2 är metodiskt reflekterande, kritiskt-logiskt och kräver lite mera tid. Det kan träda in och rädda situationen när System 1 gått vilse, men detta sker långt ifrån alltid. En av människans egenheter är att det kan byta ut ett svårt problem, vilket egentligen kräver System 2, mot ett enklare problem som System 1 snabbt klarar av; och sedan ta svaret på det enklare problemet som ett svar på det svårare problemet. System 1 arbetar med olika typer av förenklingar. Det ”föreställer sig uppsättningar med hjälp av genomsnitt, normer och *prototyper* [min *emfas*], inte med hjälp av summor” (s. 429). Tillämpas förklaringsmodellen på mig och Lindaproblemet blir det så här.

Jag förstod först att (a) ”Linda är en bankkassörska” borde tolkas som (a_1) ”Linda tillhör mängden bankkassörskor”. Men för att förenkla det givna problemet, så tolkade jag snabbt om (a) till (a_2) ”Linda är en typisk bankkassörska”, och gav därför det enligt Kahneman felaktiga svaret att (b) är sannolikare än (a). Jag måste säga att detta i mina öron låter högst otroligt. Hade jag tolkat (a) som (a_1) hade jag precis lika snabbt och oreflekterat kunnat ge mitt svar som när jag tolkade (a) som (a_2). Kahneman är med sitt tal om prototyper snubblande nära den rätta förklaringen, men genom sin oproblematiska syn på språkförståelse ändå mycket långt ifrån.

Jag upprepar därför: Kahneman (och likasinnade) bör tänka lite mer på hur vardagsspråket kan fungera, innan de drar slutsatsen att vissa av deras experimentpersoner tänker på ett sätt som ”strider mot all logik”.

Recensioner

Etik: Människa, moral, mening: En introduktion

Ann Heberlein

Albert Bonniers Förlag, 2014. 279 s. ISBN 978-91-0-013571-3

För allmänheten är Ann Heberlein kanske Sveriges mest kända etikforskare. I hennes nya bok *Etik, människa, moral, mening: En introduktion* är den uttalade ambitionen att ”erbjuda en grundläggande introduktion till ämnet etik, men också att knyta samman etiken med existensen” (s. 14). I likhet med alla andra introduktionsböcker i etik presenteras begrepp och argument för olika filosofiska teorier, men till skillnad från många andra böcker argumenterar Heberlein också för sin egen uppfattning. Det finns alltså två aspekter att bedöma: hur korrekt och pedagogiskt boken lyckas förklara filosofiska teorier och begrepp samt hur väl Heberlein lyckas argumentera för sin egen åsikt. Jag kommer beskriva och problematisera några avsnitt för att sedan göra några allmänna reflektioner.

Tidigt i boken diskuteras huruvida moralen är objektiv. Heberlein förespråkar den moraliska realismen som hon menar innebär att ”det faktiskt finns sådant som är fel och sådant som är rätt” (s. 55). Två sidor tidigare skriver hon dock detta: ”En moralisk argumentation är inte möjlig att verifiera eller falsifiera. Det är inte möjligt att hävda att det är ’sant’ att djurförsök är moraliskt felaktigt” (s. 53). Hur ska detta förstås? Det är ju självklart möjligt att *uttrycka* påståendet ”det är sant att djurförsök är moraliskt felaktigt”. Heberlein kanske snarare menar att det inte *kan* vara sant att djurförsök är moraliskt felaktigt, men detta tycks strida mot den moraliska realismen. Om det finns sådant som faktiskt är fel är det väl sant att sådant är fel?

En annan tolkning är att Heberlein tror att det finns sådant som är rätt och fel, men att det är omöjligt att ha kunskap om moraliska sanningar. Detta verkar också konstigt: genom hela boken försöker Heberlein rättfärdiga sin tro i olika moraliska frågor. Om hon lyckas med detta och dessutom råkar ha rätt i åtminstone en fråga så betyder det ju att hon har kunskap – rättfärdigad sann tro – om den moraliska frågan. En fjärde tolkning är att Heberlein bara menar att moraliska påståenden inte kan

utvärderas empiriskt. I så fall kan det fortfarande vara sant att något är moraliskt felaktigt. Oavsett vad Heberlein menar är detta en ganska förvirrande öppning på den metaetiska diskussionen.

Den moraliska realismen kontrasteras sedan mot en variant av moralisk relativism som enligt Heberlein hävdar att ”det som är rätt i ett sammanhang är fel i ett annat” (s. 54). Utifrån bokens beskrivningar av realismen och relativismen tycks det dock inte finnas någon konflikt mellan dem. Om realismen innebär att det finns sådant som är rätt och fel är detta konsekvent med att det som är rätt och fel beror på sammanhanget. Trots detta är det tydligt att Heberlein ställer realismen och relativismen mot varandra. Jag antar att detta beror på att hon egentligen menar att realismen innebär att det finns sådant som är rätt och fel *oberoende av sammanhanget*. Det är nödvändigt att göra detta tillägg till Heberleins formulering av realismen, annars står den inte i konflikt med relativismen. Som vi ska se är tillägget dock oförenligt med ett argument som används senare i boken för att rättfärdiga människovärdet.

Heberlein avfärdar sedan relativismen då den gör det ”omöjligt att kritisera andra kulturers seder och bruk” (s. 54). Detta tyder på att hon blandat ihop den metaetiska relativismen som hon just beskrivit med en annan typ av relativism som hävdar att vi bör låta andra kulturer lösa sina egna moraliska problem. Dessa typer av relativism är helt oberoende av varandra. Det är till och med så att den senare typen av relativism passar *bättre* med metaetisk realism än med relativism – bara en realist kan hävda att det är fel att kritisera andra kulturers seder och bruk i alla möjliga sammanhang.

Det tycks alltså som Heberleins avfärdande av relativismen bygger på ett missförstånd av dess innebörd. Detta är väldigt olyckligt, särskilt som relativismen målas upp som det enda alternativet till realismen. Heberlein nämner överhuvudtaget inte misstagsteorin (*error theory*) eller några non-kognitivistiska teorier – dessa är ju kanske realismens värsta motståndare. Dessutom nämns inte några argument mot realismen. På grund av allt detta misslyckas Heberlein med att försvara moralens objektivitet. Heberleins förvirrande utläggningar om ”moralisk argumentation” i kombination med felbeskrivningen av relativismen samt utelämnandet av motargument och reella alternativ till realismen gör det också svårt att se hur denna del av boken skulle kunna ha något pedagogiskt värde.

Boken fortsätter med en diskussion om definitionen av ordet människa och huruvida människor har ett värde. Först skriver Heberlein ”Människa är den som är född av en människa”, sedan skriver hon ”Människa är den som biologiskt tillhör arten människa” (s. 66). Det är kanske värt att notera att den första definitionen är snävare då den inte inkluderar foster som ju inte blivit födda än. Jag själv förespråkar den senare definitionen och jag tror också att detta är Heberleins åsikt. I diskussionen

om människovärde verkar arttillhörigheten åtminstone vara en lite mer robust gräns än huruvida varelsen är innanför eller utanför livmodern (personligen tror jag inte att någon av gränserna är moralisk relevant).

Därefter beskrivs det argument mot människovärdet som hävdar att det inte finns någon moraliskt relevant egenskap som alla människor har, men som alla icke-människor saknar. Självmedvetande, rationalitet och talförmåga är egenskaper som de flesta, men inte alla (exempelvis spädbarn eller gravt förståndshandikappade), människor besitter. Dessa egenskaper kan alltså inte grunda ett värde för alla människor. Inte heller kan människovärdet baseras på människans förmåga känna lycka och lidande – den egenskapen delar vi med många andra djur. Det verkar också konstigt att människor skulle ha ett värde enbart för att de tillhör arten *homo sapiens*. Det är ju lika godtyckligt som att säga att män har ett värde för att de tillhör det manliga könet. Argumentet hävdar kort sagt att tron på människovärdet är speciesistisk, diskriminerande på grund av arttillhörighet, och därmed moraliskt förkastligt på ett sätt som påminner om exempelvis rasism och sexism.

Heberlein beskriver detta argument mycket väl. Trots att hon själv förespråkar människovärdet verkar hon dock inte särskilt störd över argumentet. ”Kalla mig speciesist, men det är en grundläggande förutsättning i min moraluppfattning” (s. 70), skriver hon. Att människan har ett värde motiverar hon med en egen princip: ”Varje människa äger ett unikt, okränkbart värde eftersom vi kommit överens om att alla varelser som tillhör arten människa är innehavare av detta värde i lika stor utsträckning” (s. 76–77).

Principen verkar medföra att om det Heberlein kallar ”vi” hade kommit överens om något annat hade det varit rätt i stället. Detta är problematiskt på flera sätt. Dels verkar det som Heberlein inte längre kan hävda att det som är rätt och fel finns *oberoende av sammanhanget* – i så fall skulle det ju inte spela någon roll vad ”vi” tycker. Principen är alltså oförenlig med Heberleins moraliska realism.

Låt oss säga att Heberlein överger realismen för att rädda sin människovärdesprincip. Då uppstår ett nytt problem med principen: den verkar kunna rättfärdiga vilka helt godtyckliga moraliska avgränsningar som helst. Om ”människa” byts ut mot ”vit man” blir den genast mindre attraktiv: ”Varje vit man äger ett unikt, okränkbart värde eftersom vi kommit överens om att alla varelser som tillhör gruppen vita män är innehavare av detta värde i lika stor utsträckning.”

Heberlein kanske skulle hävda att i fallet med vita män är det inte rätt sorts ”vi” som är överens. Vad menar hon då med ordet ”vi”? Hon kan ju knappast avse ”alla människor” eller ”alla som läser denna bok”, för det finns ju människor som inte tror på människovärdet och det går inte att

utesluta att någon av dem läst boken. Om ”vi” i stället tolkas som ”alla som tror på människovärdet” blir principen trivial – människovärdesförespråkare komma överens om att det finns ett människovärde (precis som manliga sexister kan komma överens om ett mansvärde), men detta är knappast ett argument.

Kanske ska ”vi” förstås som ”alla rationella varelser” eller ”alla moraliska varelser”, men i så fall behöver Heberlein ett oberoende argument som visar att de som förnekar människovärdesprincipen är irrationella eller omoraliska. Heberlein framställer inget sådant argument och jag tror förresten inte att det är möjligt. (Det kanske är irrationellt eller omoraliskt att vara rasist och sexist, men i så fall det lika illa att vara speciesist). Heberlein lyckas alltså inte visa att hennes ”vi” har auktoritet nog att bestämma att människan har ett särskilt värde.

Ibland verkar det dock som Heberlein glömt bort sitt eget argument. Hon skriver till exempel: ”Människan äger således ett värde just i kraft av att vara människa. . . . Jag har argumenterat för människans särställning i relation till andra djur, för hennes förmåga till ansvarighet och autonomi. Min övertygelse är att alla människor, i varierande utsträckning, har potential till ansvarstagande” (s. 139). Här framställs två helt nya (och oförenliga) argument för människovärdet. Först hävdar hon (utan att ange något skäl) att människor har ett värde på grund av sin arttillhörighet. Som vi sett är denna hållning lika godtyckligt som rasism och sexism.

Sedan verkar Heberlein glömma detta och hävdar i stället att människovärdet inte alls bygger på arttillhörighet utan på förmågan till ansvarighet och autonomi. Detta argument bygger dock på en falsk empirisk premis: ”att alla människor, i varierande utsträckning, har potential till ansvarstagande.” Hjärndöda och gravt förståndshandikappade är inte och har inte potential för varken autonomi eller ansvarstagande. I alla fall inte i högre utsträckning än många icke-mänskliga djur. Heberlein lyckas alltså inte försvara människovärdet.

Boken diskuterar också teorier om vad som är rätt. De teorier som beaktas är konsekventialismen, deontologin, kontraktualismen, dygdetiken, och omsorgsetiken. Dessa förklaras på ett ganska tydligt sätt som ger läsaren en grundläggande uppfattning. Heberlein kommer också med invändningar mot de tre första synsätten. Invändningarna bygger mestadels på att teorierna strider mot våra (Heberleins?) intuitioner i olika situationer. Utilitarismen anklagas exempelvis för att föreskriva att oskyldiga dödas för att öka lyckan. Heberlein framför dock inga argument mot dygdetiken och omsorgsetiken. Detta beror förmodligen på att hon själv förespråkar någon form av situationsetik med ”en idé om den goda människan snarare än den rätta handlingen i centrum” (s. 249).

Det är besynnerligt att boken inte tar upp de vanliga oklarheterna

kring situationsetiken. Teorin skulle vara lättare att förstå om Heberlein förklarade vad som skiljer den från motiv-konsekventialism och huruvida handlingar är goda för att dygdiga personer skulle ha utfört dem eller om personer är dygdiga för att de handlar gott. Framförallt hade det varit bra om Heberlein diskuterat den praktiska tillämpningen – hur kan situationsetiken hjälpa mig handla rätt när den bara säger vem jag ska vara och inte vad jag ska göra? Detta är också relaterat till ett annat problem med situationsetiken – den kan inte testas mot våra intuitioner.

Vetenskapliga teorier är svaga om de inte kan testas empiriskt. På ett liknande sätt finns det en svaghet hos normativa teorier som inte kan testas mot våra intuitioner. Vid första anblicken verkar många normativa teorier, såsom exempelvis utilitarismen, väldigt rimliga. Det är först när de testas mot våra intuitioner i specifika fall som många avfärdar dem. Då Heberlein använder just intuitioner för att avfärda de andra teorierna borde hon vara misstänksam mot situationsetiken – även om grundidén verkar rimlig, kanske dess utslag i konkreta fall är kontraintuitiva. Eftersom situationsetiken inte genererar svar på hur vi bör handla i specifika fall kan den inte testas mot våra intuitioner och konkurrerar därmed inte på samma villkor som de andra teorierna. Eftersom Heberlein inte diskuterar dessa svårigheter och preciserar sin position kring dem verkar hennes situationsetik helt innehållslös. Detta är problematiskt både för den pedagogiska och argumenterande aspekten av boken.

Efter att ha beskrivit och problematiserat några avsnitt ur boken är det tydligt att många argument är väldigt svaga. Det är dock inte säkert att Heberlein är ute efter att övertyga någon. Kanske är det huvudsakliga målet snarare att ge läsaren en grundläggande uppfattning om etiska begrepp och teorier. I så fall är det väldigt konstigt att Heberlein ofta struntar i att presentera motargument och alternativ till sin egen hållning.

Ett ännu värre problem för det pedagogiska syftet är att Heberlein missuppfattat flera begrepp. Som jag redan visat blandar hon ihop två skilda typer av relativism. Hon skriver också felaktigt att Peter Singer menar ”att en person är en varelse med intressen och planer” (s. 68). I själva verket använder Singer (1996, s. 85) ordet person för att beteckna ”en rationell och självmedveten varelse”. De värsta missuppfattningarna finns dock i diskussionen om fri vilja då Heberlein frågar: ”Har hon [människan] en fri vilja och en förmåga att välja, eller är hon styrd av något? Är hon ofri som determinismen hävdar, är hennes öde förutbestämt och hennes vilja bunden av till exempel biologi eller psykologi?” (s. 64).

Citatet tyder på att Heberlein hävdar att determinismen inte är förenlig med fri vilja. Problemet är inte att Heberlein förespråkar denna hållning, utan att hon inte verkar ha förstått att många filosofer (majoriteten enligt Bourget och Chalmers 2014) är så kallade mjuka determinister (som också

kallas kompatibilister) som tror att determinismen är förenlig med fri vilja. Detta blir tydligt när hon skriver att enligt ”den mjuka determinismen är människans handlingar delvis en produkt av erfarenheter och förutsättningar, men hon är inte bestämd av det samma” (s. 114). Detta är fel – den mjuka determinismen hävdar att *alla* handlingar är bestämda, men att detta är kompatibelt med fri vilja. En bok som innehåller denna typ av fel och missförstånd är knappast lämplig som introduktion till etik.

Det finns dock några förtjänster med boken. En styrka är att den presenterar tankar från olika traditioner och tider. Många svenska introduktionsböcker till filosofiska ämnen refererar enbart till författar inom den analytiska traditionen. Heberlein däremot, hänvisar till idéer och tänkare från båda traditionerna. Dessutom refererar hon till tänkare från alla tider från Platon till Judith Butler och även till litterära referenser såsom Fjodor Dostojevskij och Albert Camus. Detta ger läsaren en stor idéhistorisk bredd.

En annan styrka är att Heberlein skriver utan filosofisk jargong och lyckas knyta ihop etiken med psykologi, antropologi och existentiella frågor. Detta är nog bra för läsare utanför fackfilosofin som kanske inte skulle ta sig igenom en bok som strikt håller sig till moralfilosofi. Överhuvudtaget tror jag att den intresserade allmänheten är den mest passande läsargruppen för denna bok. Filosofer kommer inte vinna mycket på att läsa Heberleins argument, filosofistuderande riskerar att lära sig felaktig information, men allmänheten kanske får upp ett intresse för etik och läser vidare.

Den allmänna läsaren, snarare än filosofen eller den filosofistuderande, är nog också mer benägen acceptera Heberleins slutsatser som ofta är väldigt konventionella. Att äta kött och sätta betyg i skolan är okej, vi har inte lika stora obligationer mot dem som lever i fattigdom som mot våra egna barn, vi har en fri vilja, de som felat ska straffas, osv. Det verkar som Heberlein redan från början bestämt sig för att ”sunt förnuft”-moralen (*common sense morality*) är riktig. Kanske är det därför som hon inte tycker sig behöva framställa några starka argument. För den som inte tar de konventionella svaren för givet blir dock både resonemangen och slutsatserna ganska platta. Vad är poängen med moralfilosofi om den varken rättfärdigar eller korrigerar våra befintliga åsikter?

GUSTAV ALEXANDRIE

REFERENSER

- Bourget, David och David Chalmers. 2014. ”What Do Philosophers Believe?”. *Philosophical Studies* 170, nr 3, s. 465–500.
- Singer, Peter. 1996. *Praktisk etik*. Stockholm: Thales.

Vårt matematiska universum: Mitt sökande efter den yttersta verkligheten

Max Tegmark. Översättning: Pär Svensson

Volante, 2014. 506 s. ISBN 978-91-87419-27-0

Max Tegmark är fysiker, född 1967 i Sverige och numera anställd vid MIT i Cambridge, Massachusetts. Han har gjort flera uppmärksammade och uppskattade insatser i fysik, men till skillnad från de flesta fysiker ägnar han sig även åt stora och djupa problem som ligger på – eller till och med bortom – gränsen till filosofi. För detta har han fått kritik av andra fysiker. Speciellt har han fått kritik för den här ”bokens bärande idé” (s. 312), nämligen att ”vår fysikaliska värld inte bara kan *beskrivas* med matematikens hjälp, utan att den faktiskt är matematik, närmare bestämt en matematisk struktur” (s. 15) eller ”ett väldigt matematiskt objekt” (s. 312). Denna tanke låter ju väldigt uppseendeväckande och Tegmark har också varnats av kollegor för att den kan ”äventyra hans framtid” som fysiker. (Den kan också påminna lite om den postmoderna tesen att verkligheten endast är språklig.)

Boken har tre delar. Men en viss förenkling kan man säga att den första handlar om kosmologi, den andra om kvantmekanik och den tredje om filosofi. De första två delarna är personligt och elegant presenterade, medryckande och mycket lärorika. För det allra mesta är de också begripliga – eller åtminstone tror man att man begriper. Den filosofiska tredjedelen, som också innehåller vissa inslag av naturvetenskap, är också intressant men mindre övertygande. Det är framför allt här Tegmark argumenterar för att verkligheten är matematisk.

Jag ska här huvudsakligen ägna mig åt de filosofiska inslagen i boken. Även om man kan ha en del kritiska synpunkter på dem, så tycker jag att Tegmark förtjänar beröm, snarare än kritik, för att han till skillnad från andra fysiker även tar upp filosofiska frågor. I en facktidskrift skulle det kanske inte vara så lämpligt, men här är det ju fråga om en populärvetenskaplig bok. Och att det finns samband mellan fysik och metafysik är ju rätt naturligt, bägge har ju att göra med hur verkligheten är beskaffad.

Bokens första kapitel har just rubriken ”Vad är verkligheten?” Det kan inte sägas vara en rent fysikalisk fråga; ”verkligheten” är inte ett fysikaliskt begrepp som t.ex. ”neutron”, ”massa” eller ”gravitation”. Tegmark besvarar den inte heller genom att hänvisa till några fysikaliska experiment, utan genom att helt enkelt definiera verkligheten som ”den yttersta beskaffenheten hos den yttre fysikaliska värld som vi är en del av” (s. 17–18).

Här undrar man: varför bara ”den *yttre* beskaffenheten” (vad nu det betyder)? Och varför bara den ”*yttre*” fysikaliska världen? Och varför

bara den ”fysikaliska” världen? Varför inte hellre helt enkelt säga att ”verkligheten” är ”den värld vi är en del av”? Det får man inte veta.

Vidare innehåller enligt Tegmark ”den värld vi är en del av” inte bara vårt vanliga universum – dvs. ”den del av den fysiska verkligheten som vi i princip kan observera” – utan dessutom flera ”parallella” universum. Närmare bestämt finns det parallella universum på fyra olika nivåer. På nivå I finns det andra universum, med samma fysikaliska lagar som vårt, men bortom vår kosmologiska horisont, så att ljuset därifrån ännu inte har hunnit nå oss. På nivå II finns det andra universum, med delvis andra fysikaliska egenskaper, som är ”evigt onåbara på grund av att rymden mellan oss och dem tänjs ut av den kosmologiska inflationen” (s. 402). På nivå III finns kvantfysikens många världar – enligt flervärlds-tolkningen av kvantfysiken – och på nivå IV finns olika ”matematiska strukturer” (s. 403).

Om allt detta existerar, så tillhör det den fysikaliska verkligheten, helt enkelt därför att ”fysikalisk verklighet” enligt Tegmark definieras som ”allt som existerar” (s. 156). Om matematiska objekt som t.ex. mängder och tal existerar, så ingår de alltså i den fysikaliska verkligheten. Men det betyder att den fysikaliska verkligheten åtminstone delvis är abstrakt, dvs. varken har position, rörelsemängd eller massa. Och då kan det kanske finnas utrymme för den ovanliga uppfattningen att verkligheten är matematik. Men Tegmarks argument för detta inte särskilt övertygande.

VERKLIGHETEN ÄR MATEMATIK

På ett ställe säger Tegmark: ”I de två första delarna av boken har vi sett att vår fysikaliska värld kan beskrivas förbluffande väl med matematiska ekvationer, vilket ger hopp om att vi en dag ska finna ekvationer som [...] ger en heltäckande beskrivning av vår yttre verklighet på samtliga nivåer” (s. 301). Det kan låta underligt, men ”vår fysikaliska värld” är tydligen inte detsamma som (det som på s. 156 kallades) ”den fysikaliska verkligheten”, utan syftar nog här bara på den ”yttre verklighet” som han ”anser skulle existera även om vi människor inte funnes” (se s. 303). Vår fysikaliska värld inkluderar inte det han kallar vår ”konsensusverklighet”, som är ”de makroskopiska objektens tredimensionella positioner och rörelser, liksom andra vardagliga verklighetsattribut som vi har en gemensam beskrivning av i termer av den klassiska fysikens välbekanta teorier” (s. 301).

Han säger också att ”fysikens uppgift är att härleda konsensusverkligheten ur den yttre verkligheten” (s. 303). Även detta kan låta lite konstigt, men det betyder antagligen att fysiken ska hitta matematiska ekvationer som förklarar de observationer, i t.ex. experimentsituationer, som kompetenta bedömare kan vara överens om. (Men då krävs väl också lite mer, nämligen någon sorts regler eller antaganden som översätter från mate-

matiskt språk till observationsspråk?) Och hur som helst är detta knappast tillräckligt för att härleda t.ex. beskrivningar av historiska händelser och processer, som andra världskriget eller det senaste riksdagsvalet eller den berömda diskussionen om kvantmekanik mellan Einstein och Bohr. Sådant ska kanske inte räknas till "konsensusverkligheten", men historien inträffar väl ändå i någon sorts yttre verklighet? Eller menar Tegmark att historien – inklusive fysikhistorien – endast existerar i vissa personers "inre verklighet"?

Nej, vad han menar är nog att en så kallad "teori om allt", som förklarar gravitation och kvantmekanik, endast får innehålla matematiska begrepp. Den får inte innehålla något "mänskligt bagage", som han uttrycker det (s. 322ff). Mänskligt bagage finns däremot i konsensusverkligheten.

Han påstår, gång på gång, att "om man anser att det finns en yttre verklighet oberoende av oss människor, måste man också hålla för sant att vår fysikaliska verklighet är en matematisk struktur" (s. 329). Men varför måste man det?

Hans argument ser ut så här. Det finns en yttre fysikalisk verklighet som är oberoende av oss människor. Men en "fullständig" (?) beskrivning av denna verklighet måste då vara begriplig för andra än människor, t.ex. utomjordingar eller superdatorer, och då får den inte innehålla termer i mänskliga språk ("mänskligt bagage") som t.ex. "partikel", "observation", osv. *Därför* måste beskrivningen vara "rent matematisk" (se s. 322–26). Och vidare, "om några matematiska ekvationer helt beskriver vår yttre fysikaliska verklighet och en matematisk struktur så är vår yttre fysikaliska verklighet och den matematiska strukturen ett och samma" (s. 353). Med andra ord: "yttre verklighets-hypotesen, som säger att det existerar en yttre fysikalisk verklighet som är helt oberoende av oss människor, implicerar matematiskt universum-hypotesen, som säger att vår yttre verklighet är en matematisk struktur" (s. 441).

Detta argument är ju inte särskilt övertygande. De flesta av oss tror väl att det finns en yttre verklighet oberoende av oss människor. (Man kan dock undra vad som är "oberoende". Himlakropparna är väl det. Kanske också superdatorer, om de har konstruerats av utomjordingar, men knappast om de har konstruerats av människor.) Men det verkar helt omöjligt att vi skulle kunna åstadkomma en beskrivning av något utan att använda ett språk – och de språk vi då kan använda är mänskliga språk med "mänskligt bagage" (eller möjligen omänskliga språk som kan översättas till mänskliga språk). Hit räknas även vårt matematiska språk. De arabiska siffror och andra, ofta grekiska, symboler vi där använder är nog inte mer begripliga för utomjordingar än t.ex. ord som "partikel" och "massa". Einsteins ekvationer har (oberoende av om de är "oberoende"

av Einstein) formulerats på språk som konstruerats av människor. Vilket inte hindrar att vi tror att det som gör att de är sanna – om de nu skulle vara sanna – är något som är oberoende av oss, nämligen vissa förhållanden i den fysikaliska verkligheten. Men utgör dessa förhållanden en ”matematisk struktur”?

VERKLIGHETEN SOM EN MATEMATISK STRUKTUR

Det är tyvärr rätt oklart vad Tegmark menar med ”en matematisk struktur” – trots att han ägnar flera sidor åt att förklara det och trots att han hänvisar till det som ”ett standardbegrepp inom den moderna matematiska logiken” (s. 333), vilket antyder att han helt enkelt syftar på strukturer i mängdteoretisk mening.

Men till att börja med tror jag att man här måste skilja mellan å ena sidan mängdteoretiska strukturer i allmänhet och å andra sidan mängdteoretiska strukturer som enbart innehåller matematiska entiteter och matematiska relationer. En struktur i mängdteoretisk mening är helt enkelt en mängd entiteter i kombination med en uppsättning relationer (som råder mellan vissa av dessa entiteter). Att den fysikaliska verkligheten kan identifieras med en matematisk struktur i denna första mening är kanske inte så konstigt. Det stämmer nog med den vanliga uppfattningen bland folk i allmänhet – även om det förstås kan vara svårt eller kanske omöjligt att identifiera precis *vilken* struktur det är fråga om. Precis *vilka* entiteter och *vilka* relationer finns det i den fysikaliska verkligheten? Redan detta är en fråga som kan göra vem som helst vimmelkantig. Men att den fysikaliska verkligheten dessutom skulle vara en struktur i den andra meningen, dvs. en struktur med enbart *matematiska* ingredienser, låter väldigt konstigt – och det är troligen just detta som Tegmarks kritiker inte vill gå med på.

Vad säger då Tegmark själv om vad en ”matematisk struktur” är? Han säger att den är ”en uppsättning abstrakta entiteter som står i relation till varandra” (s. 328) eller ”en samling abstrakta entiteter som det finns relationer mellan” (s. 334). Detta passar in på mängdteoretiska strukturer. Och att entiteterna ska vara ”abstrakta” tyder på att det är fråga om matematiska strukturer i den andra av de två betydelser jag angett ovan. Detta är emellertid svårt att förena med vissa andra påståenden han gör, t.ex. att de naturliga talen är en struktur (s. 339) eller att en matematisk struktur är ”en ekvivalensklass av beskrivningar” (s. 411). Visserligen är tal abstrakta entiteter och även beskrivningar kan kanske ses som abstrakta entiteter, men beskrivningar är knappast oberoende av oss. Och visserligen står både de naturliga talen och ekvivalenta beskrivningar ”i relation till varandra”, men så länge man inte anger exakt vilka relationer som avses, så är det ju inte heller klart vilken struktur man syftar på. I

bägge fallen finns det obestämt många – och troligen oändligt många – relationer som råder mellan entiteterna. För att identifiera en mängdteoretisk struktur måste man identifiera både entiteter och relationer.

Hur som helst, han säger, rätt kryptiskt, att ”två beskrivningar av matematiska strukturer är ekvivalenta om de motsvarar varandra på ett sätt som bevarar alla relationer” (s. 338). Och, lite tydligare, att ”om det föreligger en isomorfism mellan en matematisk struktur och en annan struktur (en överensstämmelse ett till ett mellan de två som bevarar relationerna) är de ett och samma” (s. 353). Varje struktur är förstås isomorf med sig själv, men Tegmark tycks alltså anse att isomorfa strukturer dessutom är *identiska*, vilket är något helt annat. Det förutsätter en ovanlig uppfattning om vad en struktur är.

Om en *matematisk* struktur (där entiteterna och relationerna är matematiska) är isomorf med en *fysikalisk* struktur (där entiteterna och relationerna är fysikaliska) – vilket verkar fullt möjligt – så är strukturerna ändå *inte* identiska, eftersom de består av olika saker (matematiska i ena fallet och fysikaliska i det andra). Isomorfin består i att entiteterna och relationerna i de bägge strukturerna kan paras ihop på ett sådant sätt att en relation råder mellan entiteterna i den ena strukturen om och endast om motsvarande relation råder mellan motsvarande entiteter i den andra strukturen. Tegmark ger ingen förklaring till sin avvikande uppfattning.

Kanske föresvävar det honom att isomorfa strukturer måste ha något gemensamt och att det är detta ”något”, som är en matematisk struktur i hans mening? Det är ju riktigt att isomorfa strukturer har något gemensamt, men detta något är inte en struktur (i mängdteoretisk mening), utan snarare en *mängd* av (isomorfa) strukturer. Det strukturerna har gemensamt är att de alla är element i denna mängd. Det vill säga, att de är isomorfa.

Som jag redan nämnt är ”bokens bärande idé” (s. 312) att ”vår fysikaliska värld inte bara kan *beskrivas* med matematikens hjälp, utan att den faktiskt är matematik, närmare bestämt en matematisk struktur” (s. 15). Jag kan alltså inte finna att denna ”bärande idé” är fullt begriplig, och inte heller att Tegmark har gett något bra argument för den. Den har också en del egendomliga konsekvenser. Exempelvis anser Tegmark, i enlighet med kvantmekaniken, att en subatomär partikel kan befinna sig på två platser samtidigt. Han uppfattar sådana partiklar som *matematiska* (och därmed abstrakta) objekt, som fullständigt kan beskrivas av sina kvanttal (s. 211). Men hur kan ett matematiskt objekt över huvud taget befinna sig på en plats?

En lite lustig sak med tanken att världen är en matematisk struktur, i den andra av de två tolkningar jag nämnt, är att den tycks eliminera det gamla problemet om hur världen har skapats eller uppstått. Som Teg-

mark själv säger: ”Man kan inte *skapa* en matematisk struktur, den finns bara till” (s. 404). Så någon skapelse av världen kan det inte ha varit fråga om. Simalabim!

Men frågan om *varför* världen existerar finns på sätt och vis kvar. Alla matematiska strukturer existerar i Platons idévärld (åtminstone om de är konsistenta). Men varför existerar just *vår* världs matematiska struktur även i fysikalisk mening? På detta har Tegmark ett svar, nämligen att *alla* matematiska strukturer – eller åtminstone de som är beräkningsbara eller ändliga (se s. 416–17) – existerar i fysikalisk mening, ty ”matematisk existens och fysikalisk existens är samma sak” (s. 402). Det är detta som leder till multiversum IV. Något bra argument för detta har jag dock inte kunnat hitta, trots att Tegmark har ett helt avsnitt som heter ”Varför jag tror på multiversum IV”. På s. 402 hänvisar han till en ”plötslig insikt”. Det hade varit bättre om han hade hänvisat till något som multiversum IV kan förklara, helst något vi kan observera.

VERKLIGHETENS STRUKTUR

Man kan för övrigt undra om verkligheten verkligen har – eller är – *en* struktur? Om begreppet struktur uppfattas på vanligt sätt, så kan man urskilja många olika strukturer i verkligheten. Det skulle Tegmark nog hålla med om – han har ju själv gett flera exempel – men hur förhåller sig dessa strukturer till den särskilda struktur som enligt Tegmark är detsamma som verkligheten, den som är ”den yttersta beskaffenheten” hos världen? Några är kanske substrukturer, men många är ju inte matematiska, i den meningen att de bara innehåller matematiska entiteter och relationer, och då kan de inte heller vara substrukturer till en matematisk struktur.

Jag tror att verklighetens struktur för Tegmark är, eller beskrivs av, en eller flera matematiska ekvationer, som binder ihop gravitation och kvantfysik, det han också kallar ”ekvationerna för en sammanhållen teori om allt” (s. 318). Anta att fysikerna till sist hittar dessa ekvationer. De är förstås matematiska. Men är den fysikaliska verklighet de beskriver också matematisk? Det är det Tegmark tycks mena. Men hur jag än letar i boken kan jag inte finna något bra argument för det.

Och även om fysikerna om några år kommer fram till en ”teori om allt” som kan bekräftas empiriskt, så kan man inte *identifera* den (eller motsvarande struktur) med verkligheten. Ty det skulle innebära att fysiken då skulle ha blivit *slutgiltig*, på ett sätt som fysikaliska teorier aldrig tidigare under historien har varit. Man kan nog aldrig utesluta att nya problem dyker upp, som kräver nya och bättre teorier.

EVIG INFLATION

Vi är nog många som har undrat hur Big Bang uppstod – om vi nu tror att den alls inträffade. Numera anses det att den var resultatet av en oerhört snabb expansion (inflation) av en egendomlig substans, som sattes igång av kvantfluktuationer (s. 144) och som enligt Einsteins gravitationsteori kan växa från subatomär skala till att bli mycket större än vårt observerbara universum under bråkdelen av en sekund (s. 133). Big Bang uppstår när den ursprungliga substansen delvis sönderfaller till vanlig materia. Vi som undrade över Big Bang undrar förstås också varifrån denna substans kom, men inflationshypotesen tycks i alla fall förklara flera observationer som annars är svårförklarliga. Detta visar Tegmark mycket pedagogiskt (s. 140–44).

Inflationen ger inte bara upphov till en Big Bang och vårt multiversum I. Eftersom inflationssubstansen hela tiden expanderar, samtidigt som den endast delvis sönderfaller till vanlig materia, så kommer vid varje nytt sönderfall en ny Big Bang och därmed får vi ett multiversum II (s. 146).

Men om rymden är oändlig och likformigt fylld av materia, vilket förutsågs av inflationshypotesen, så innehåller den inte bara många andra universum, utan faktiskt *oändligt* många (s. 163). Nästan alla inflationsmodeller ger upphov till evig inflation, säger Tegmark (s. 147). Även ”om vi bor i en ändlig rymd formad som en hypersfär finns det åtminstone hundra parallella universum på nivå I”, men att rymden är ändlig ”motsägs starkt av observationer som gjorts på senare år” (s. 166). Och i multiversum II finns i sin tur oändligt många multiversum av typ I (se s. 172). Det är svårt att förstå. Inflationen pågår kanske fortfarande, men den har väl ännu inte pågått oändligt länge, så hur kan den då *redan* ha hunnit producera oändligt många oändligt innehållsrika multiversum? Den eviga inflationen gör visserligen att ”allt som kan hända faktiskt händer” (s. 194), men av detta följer väl inte att allt redan har hänt.

MULTIVERSUM OCH JAGET

Men ännu konstigare låter kanske idén att varje människa finns i oändligt många olika universum (se s. 159); ”det existerar både flera universum och flera du” (s. 433). Tegmark säger att enligt ”den enklaste och mest populära kosmologiska modellen”, så existerar det ett annat exemplar av mig, en person vars liv är identiskt med mitt åtminstone upp till nu, i en annan galax i ett annat universum oerhört långt härifrån (s. 155) – och om ”nu” är mitt dödsögonblick, så blir ju hela våra liv identiska. Men kan verkligen oändligt många ”identiska” personer leva i olika universum? Strider inte det mot Leibniz princip att om X är identisk med Y, så finns det *ingen* skillnad mellan X och Y? Att leva i olika världar innebär ju en skillnad.

Detta diskuterar inte Tegmark, men man kunde kanske tänka sig en

teori om jaget som gör det hela konsistent. Analogt med det förhållandet att jag finns vid många olika tidpunkter i vår rumtid, så kunde jag finnas i olika universum. Liksom mitt jag är utspritt över många olika tidpunkter, så kunde det vara utspritt över många olika universum. Skillnaden skulle vara att jag är medveten om att jag finns vid många tidpunkter, medan jag är omedveten om att jag (samtidigt?) finns i många universum. Den skillnaden ska man kanske bara vara tacksam för.

Man finns ju dessutom i många olika *versioner* i olika universum. I vissa gör man si, i andra gör man så. I vissa drabbas man av det ena, i andra av det andra. Tegmark säger att han *hoppades* att förlossningen av hans son i maj 1999 skulle gå bra, trots att han visste att den skulle gå både bra och dåligt, i olika universum III (s. 242). Han fick alltså uppleva bägge varianterna, utan att märka det. Men det är mycket man inte märker. Tegmark nämner att Hugh Everett påpekade att man ju inte heller känner att man snurrar runt solen med en hastighet av 30 km i sekunden (s. 242). Man märker inte heller att man mest består av tomrum.

Man kan förstås fråga sig varför vi uppfattar oss själva som lokaliserade i *ett* universum, trots att vi finns i oändligt många. Tegmark gissar att förklaringen är att ”det gynnar oss” (s. 375). Det är ett lovande uppslag, som skulle behöva utarbetas. Men man kan nog anta att det vore odrägligt att vara medveten om att man finns i oändligt många olika universum. Oerhört förvirrande. Det är säkert mycket bättre att omedvetet ha oändligt många medvetandehåll, ett i varje universum där man existerar – analogt till att man har ett medvetandehåll vid varje tidpunkt ett universum. (Har vi också olika kroppar i olika universum? Det är en smaksak, som beror på hur vi definierar ”kropp”. Det vi vanligen menar är väl att en mänsklig kropp bara finns i ett universum, men vi kunde ju också säga att en människas verkliga kropp består av alla de kroppar, i olika universum, som innehåller något av hennes medvetandehåll.)

Kanske måste man också veta mer om vårt medvetande innan man riktigt kan förstå varför vi uppfattar vår existens på ett visst sätt. Och medvetandet vet vi inte mycket om. Men Tegmark ”utgår från påståendet att vår oförmåga att förstå medvetandet inte står i vägen för en full förståelse av den yttre fysikaliska verkligheten” (s. 23). Det kan betvivlas. Om man har en ”full förståelse” av den yttre fysikaliska verkligheten, så inkluderar väl detta också en förståelse av hur denna ger upphov till eller orsakar (och påverkas av) medvetandet.

En intressant konsekvens – som Tegmark inte nämner – av den eviga inflationen och av att man finns i många olika universum (i kombination med min tolkning av detta) tycks vara att man aldrig slutgiltigt dör. När man dör i ett universum finns man ju fortfarande kvar i flera andra, möjligen i evighet, eller åtminstone utan slut, och utan att vara medveten om

att man har dött. Denna form av odödlighet är också tilltalande på så sätt att den inte tycks innebära någon risk för att man ska bli uttråkad eller obehagligt gammal i biologisk mening. Frågan om inflationens existens verkar alltså vara direkt livsviktig!

Tegmark diskuterar på ett ställe något liknande, en ”kvantodödlighet”, som dock tycks medföra att man hela tiden blir äldre och äldre, och som också förutsätter ett oändligt delbart kontinuum i naturen, vilket han inte tror på (se s. 275–78). Så för honom finns ett problem om att lokalisera sig själv på multiversumnivå I, att avgöra vilken av ens kopior som har ens egna subjektiva förnimmelser (s. 285). För min tolkning tycks detta problem försvinna.

FYSIKENS KRIS

När man väl har vant sig vid tanken på evig inflation och många parallella världar, så känns den rätt trevlig. Men efter närmare fyrahundra sidor i boken får man veta att det också finns allvarliga problem med detta. Speciellt tycks just antagandet om oändlighet leda till att alla inflationsteorins förutsägelser kullkastas på grund av det s.k. måttproblemet. Vilket ger upphov till vad Tegmark anser vara ”den största krisen för fysiken idag” (s. 395), och som ”väcker frågan om fysiken verkligen kan förutsäga något över huvud taget” (s. 399).

Måttproblemet har att göra med ”vad en typisk observatör bör förvänta sig att mäta” eller vilken metod – vilket ”mått” – man ska använda för att beräkna sannolikheter när oändligheter är inblandade (s. 394). Inflationsteorin förutsäger enligt Tegmark oändligt många observatörer (s. 395) och han tror att måttproblemet bottnar i antaganden om just oändlighet, dels ”att rymden kan ha en oändlig volym” och dels ”att det finns kvantiteter i naturen som kan variera kontinuerligt” (s. 397).

För en oinitierad läsare verkar det naturligt att invända att även om inflationen är ”evig” i den meningen att den inte har något slut, så följer inte att den redan är oändligt stor eller att den redan innehåller oändligt många observatörer (se ovan). Och som Tegmark själv påpekar har det ”spekulerats mycket om huruvida vårt eget universum också uppvisar någon form av diskret indelning av rumtiden på så låg nivå att vi ännu inte har upptäckt den” (s. 407).

Man kan till och med undra om det inte just är idén om att verkligheten är matematisk som leder till problemet: att det i *matematiken* finns aktuella oändligheter är ju allmänt godtagat, t.ex. oändligt många primtal eller oändligt många reella tal mellan 0 och 1, men att det i den fysikaliska verkligheten skulle finnas aktuella oändligheter är mer tvivelaktigt. Att antalet punkter i en matematisk/geometrisk modell av rumtiden är oändligt är en sak, men att det skulle finnas oändligt små och oändligt

stora avstånd i den fysikaliska verkligheten låter väldigt ointuitivt, för att inte säga obegripligt.

VETENSKAPLIGHET

Tegmark säger att ”den kunskapsteoretiska gränsen mellan fysik och metafysik går mellan det man kan och det man inte kan pröva i experiment” (s. 158). Det är nog en vanlig uppfattning. Men gränsen är knappast skarp. Hur är det med allt det som enligt Tegmark själv är ”kontroversiellt” i boken? Kan det ”prövas i experiment”? Eller är det metafysik?

Är de mer uppseendeväckande teorierna i Tegmarks bok ”vetenskapliga”? Exempelvis multiversum II, III och IV. Enligt Tegmark är detta inga teorier, utan förutsägelser (s. 197), men detta är nog mest en fråga om ord, ty sådana ”förutsägelser” borde i sin tur också kunna vara empiriskt prövbara.

Men om de har vissa falsifierbara konsekvenser, så kan de dessutom ha konsekvenser som är icke-falsifierbara. Strävan efter en enkel teori kan medföra att även icke-falsifierbara konsekvenser – t.ex. antaganden om oändlighet – slinker med; att eliminera dem i en alternativ formulering kan vara svårt och dessutom göra teorin mer komplicerad. Låt X vara en falsifierbar teori. Då är även X & Y falsifierbar, för vilket metafysiskt Y som helst. Men Y kan ändå sägas bidra till teorins empiriska innehåll, ty om E är en empirisk konsekvens av X, så är implikationen $Y \rightarrow E$, som saknar empiriskt innehåll, något som tillsammans med Y har empiriskt innehåll. Och i en komplicerad teori kan det vara svårt att skilja det falsifierbara (X) från det irrelevanta och metafysiska (Y).

Man kan undra om matematiskt universum-hypotesen är falsifierbar. Det anser Tegmark. Han säger: ”Om någon till exempel visade på ett övertygande sätt att naturlagarna har en inbyggd fundamental slumpmässighet (i motsats till deterministisk observatörkloning som bara *känns* slumpmässig subjektivt) skulle detta kullkasta matematiskt universum-hypotesen” (s. 446), ty ”det finns inget slumpmässigt i en matematisk struktur” (s. 426). Men om det skulle vara *omöjligt* att visa att naturlagarna har en inneboende slumpmässighet, så vore det fortfarande möjligt att matematiskt universum-hypotesen *inte* är falsifierbar och att den därmed är ovetenskaplig.

FRAMTIDEN OCH LIVETS MENING

I bokens sista kapitel reflekterar Tegmark över universums och den mänskliga civilisationens framtid. Både elegant och förnuftigt. Han har till och med ett avsnitt om livets mening. Vårt universum ger ingen mening åt livet, men livet ger däremot mening åt universum, anser han (s. 490). Han presenterar också en kalkyl som visar att sannolikheten för

att det skulle finnas intelligent liv på någon annan plats i vårt universum är "rätt liten" (s. 495–96).

Med tanke på att "det finns parallella universum där alla möjliga fysikaliska utfall uppträder", så ställer han också den närliggande frågan om varför vi alls ska bry oss om hur det går i vårt eget universum (s. 491). Det är en bra fråga, men hans svar är ganska lamt. Han tycker att "det känns logiskt" att "värna om civilisationen, planeten och det universum de hör till" – och att det är lika logiskt att bry sig mindre om andra universum, eftersom våra beslut inte kan ha någon inverkan på dem (s. 491).

Man kan undra hur det går ihop. Om vår fysikaliska verklighet är en oföränderlig matematisk struktur kan väl våra beslut inte ha någon inverkan på något alls. Men dessutom finns vi ju inte bara i ett universum, utan enligt Tegmark i oändligt många. Alla dessa är "vårt" universum (åtminstone med min tolkning av jagets natur). Vi fattar vissa beslut i vissa universum och andra beslut i andra. Vi fattar väl alla tänkbara beslut, och hur gör man då för att "bry sig om" något? Ja, i bästa fall bryr man väl sig om sitt multiversum.

LARS BERGSTRÖM

Notiser

DEN GENERÖSA KAPITALISMEN

I Dagens Industri (18/11 2014) skriver Investors vd Börje Ekholm bland annat: ”Vinst är en förutsättning, men kan inte vara det övergripande målet för ett framgångsrikt företag.” Det verkar som om han menar att det övergripande målet är ”att tillfredsställa människors och samhällets behov av varor och tjänster”. Och kanske också att ”generera arbetstillfällena, bidra till den frihet som egen försörjning innebär för medarbetarna, skapa välbefinnande och därmed också skatteintäkter till det samhälle företagen är en bärande del av”. Målet är alltså rent altruistiskt, vilket kanske kan förvåna sofliga. Dessutom är en sådan målsättning troligen olaglig. I aktiebolagslagen 3 kap. 3 § sägs det nämligen att ”om bolagets verksamhet helt eller delvis skall ha ett annat syfte än att ge vinst till fördelning mellan aktieägarna, skall detta anges i bolagsordningen”. Man kan undra om Investor i sin bolagsordning har angivit några altruistiska eller samhällsnyttiga målsättningar.

ONDA OCH GODA

I förlagsreklamen för Torbjörn Tännsjö's bok *Fatta!* står det bland annat: ”Tron på att det finns genuint goda eller dåliga människor är grundad – en skrämmande tanke, men viktig om vi vill skapa ett gott samhälle, snarare än goda människor.” Ordet ”genuint” är lite lurigt här, men tanken är väl att det varken finns goda eller dåliga människor. Är den tanken skrämmande? Tanken att exempelvis Nelson Mandela inte är bättre än Hitler. Eller att fascisterna inte är sämre än humanister och demokrater. Eller är det tanken att sådana tankar är grundade som är skrämmande? Och är de verkligen grundade?

NYA BÖCKER (SOM SKICKATS TILL REDAKTIONEN)

G. A. (”Jerry”) Cohen, 1941–2009, politisk filosof i Oxford och en av de främsta företrädarna för den s.k. analytiska marxismen, har skrivit

boken *Varför inte socialism?* som nu utkommit på Daidalos förlag, 2013.

En välkommen bok för den som drar sig för att läsa väldigt tjocka böcker är *Thomas Pikettys Kapitalet i det tjugoförsta århundradet* av ekonomen Jesper Roine, Volante 2014. Den innehåller bara lite över hundra sidor – att jämföra med Pickettys original på 970 sidor.

Birger Winsa har utgivit en tjock bok med titeln *Språket: Från paradiset till riksdagen*, Meänkielen förlaaki, 2014.

NÅGRA KÄNDA FILOSOFER SOM DÖTT 1914

Israel Scheffler, Brian Loar, Robert A. Dahl, Joyce Mitchell Cook, Judith Baker, Anthony Brueckner, James Higginbotham, Rudolf Haller, Stanley Rosen, Fred Sommers, Patrick Suppes, David Armstrong och A. Phillips Griffiths.

TANKEGYMNASTIK

Det sägs ofta att man måste kunna hålla två tankar i huvudet samtidigt. Men frågan är om detta över huvud taget är möjligt. Man kan ju försöka. Kan man t.ex. tänka både att $1 + 1 = 2$ och att Stockholm är Sveriges huvudstad *på precis samma gång*? Författaren till dessa rader har försökt, flera gånger, utan att komma till något alldeles övertygande resultat. Det *känns* som om de två tankarna kommer efter varandra, ibland visserligen väldigt snabbt, men inte så snabbt att de blir helt simultana. Men det är svårt att avgöra. Bland annat därför att de också trängs med en tredje tanke: ”är det två tankarna alldeles samtidigt?” Men denna tredje tanke verkar inte heller inträffa exakt samtidigt med de två andra. Kanske är det så att en tanke är precis allt det som vid en viss tidpunkt ryms i ett huvud. Eller i det medvetande som vid tidpunkten i fråga ryms i den hjärna som ryms i huvudet. Och i så fall går det ju inte att hålla två tankar i huvudet samtidigt.

DEDUKTIV LOGIK

Enligt Gilbert Harman och Sanjeev R. Kulkarni i uppsatsen ”The Problem of Induction” är följande en deduktivt giltig slutledning:

All people are mortal.
I am a person.
So, I am mortal.

Man kan undra vad Gud skulle säga om det!

ANDERSSON OM POPPER

Författaren och DN-skribenten Lena Andersson hävdar i en stor artikel att det finns en "postmodern spricka" hos Karl Popper; att han "framstår som en tidig postmodernist" (DN 2/1 2015). En sådan tes kan antagligen reta gallfeber på många Popperbeundrare, men Andersson tycks snarare se det postmoderna hos Popper som en förtjänst.

Är Popper verkligen postmodern? Andersson anser att han "avvisar allt tal om att världen är på ett visst sätt" – vilket onekligen låter postmodernt. Men hon säger också att han tror "på objektiv kunskap och en objektiv verklighet", vilket – tillägger hon – "är en smula motsägelsefullt". Andersson anser att postmodernisterna är mer konsekventa än Popper: "i motsats till Popper tar de konsekvenserna av sina förhållningssätt, avvisar objektiv kunskap och anammar sanningsrelativismen och det subjektiva vetandet där ingen utsaga är giltigare än någon annan."

Det här skulle ju Popper inte ha gillat, men Andersson menar att eftersom han "driver en absolutistisk antiessentialism", så hamnar han "i famnen på den förnuftsfiendliga relativism och nihilism som är honom främmande".

Här vill den vanlige DN-läsaren nog gärna veta vad en "absolutistisk antiessentialism" är för något. Det framgår inte särskilt tydligt i artikeln, men termen "essentialism" brukar vanligen beteckna uppfattningen att varje (typ av) objekt har någon nödvändig egenskap, någon egenskap som objektet måste ha. Antiessentialism skulle då förneka detta och absolutistisk antiessentialism skulle väl förneka det på ett särskilt bestämt eller absolut sätt. Själv säger Andersson att antiessentialism "ytterst är upplösning av alla kategorier".

Att skilja mellan nödvändiga och tillfälliga egenskaper hos ett objekt kan verka rätt rimligt. En cirkel måste t.ex. vara rund, men den måste inte vara stor, kan man tycka. Men om en viss cirkel faktiskt är stor, så måste den väl vara stor – för om den inte vore stor, så skulle det inte vara just denna cirkel. Lena Andersson har egenskapen att vara författare, men det är förmodligen inte en nödvändig egenskap hos henne, hon kunde nog ha haft något annat yrke. Kunde hon också ha haft ett annat kön? Kunde hon ha varit man? Förmodligen. Kunde hon ha varit en katt? Eller en cykel? Förmodligen inte. Så egenskapen att inte vara en cykel är kanske en nödvändig egenskap hos Andersson. Detsamma gäller kanske egenskapen att vara född på 1900-talet. En person som t.ex. föddes på 1700-talet kunde inte ha varit identisk med Andersson, ty vederbörande hade haft helt andra föräldrar.

Men även om man förkastar essentialismen – dvs. tanken att objekt (och/eller typer av objekt) har nödvändiga egenskaper – så kan man knappast vara tvungen att anamma sanningsrelativismen och tanken

att ”ingen utsaga är giltigare än någon annan”. Även om Lena Andersson kunde ha varit en cykel (vilket jag för min del inte vill påstå), så kan man väl ändå inte dra slutsatsen att utsagorna ”Hon är en cykel” och ”Hon är en författare” är lika giltiga eller lika sanna. Så Popper behöver inte vrida sig i sin grav, utan kan dra en postum suck av lättnad.

MÄNSKLIGHETENS FRAMTID

Matematikprofessorn och statistikexperten Olle Häggström meddelar på sin blogg att vår tillvaro inte är särskilt trygg. Han skriver bland annat: ”Mänskligheten står inför enorma möjligheter, men också enorma risker. En informell enkätstudie bland futurologer specialiserade på globala risker gav 2008 ett medianvärde på 19 % för uppskattad sannolikhet att mänskligheten går under senast 2100 – en siffra som givetvis bör tas med en nypa salt, men som ändå ger en fingervisning om situationens allvar. Uppdelat på olika slags risker visar det sig att naturliga faror, som t.ex. asteroidnedslag, är försumbara i ett hundraårsperspektiv, och att de verkligt stora farorna kommer från oss själva och de teknologier vi utvecklar.”

MATEMATIKENS NYTTA

Var det möjligt så att det var en enda person, nämligen matematikern Alan Turing, som avgjorde andra världskriget till de allierades fördel? Han var ju avgörande för att knäcka den tyska kommunikationskoden *Enigma*, vilket inte bara gjorde stor skillnad för de allierade konvojernas möjligheter att undgå torpederingar från tyska ubåtar, utan också hjälpte Sovjetunionen att slå tillbaka de tyska angreppen – eftersom ryssarna fick tillgång till de tyska planer som kommunicerades med *Enigma*. Att Turing var avgörande är ju förenligt med att hans arbete byggde vidare på tidigare insatser, bl.a. av de polska matematikerna Marian Rejewski, Jerzy Różycki och Henryk Zygalski.

Medarbetare i detta nummer av *Filosofisk tidskrift*: *Gustav Alexandrie* studerar praktisk filosofi vid Stockholms universitet, *Per Bauhn* är professor i praktisk filosofi vid Linnéuniversitetet, *Lars Bergström* är professor emeritus i praktisk filosofi, *Peter Gärdenfors* är professor i kognitionsvetenskap vid Lunds universitet, *Ingvar Johansson* är professor emeritus i teoretisk filosofi och *Simon Skau* forskar vid institutionen för neurologi och fysiologi vid Göteborgs universitet.

Instruktioner till skribenter

Filosofisk tidskrift har som syfte att bidra till en allsidig och fruktbar diskussion av filosofiska problem, samt att på ett lättfattligt sätt informera om aktuell filosofisk forskning. Den vänder sig inte enbart till fackfilosofer, utan vill framför allt nå en bredare läsekrets av filosofiskt intresserade personer

Tidskriften står öppen för olika filosofiska inriktningar, men den vill undvika bidrag som man inte kan tillgodogöra sig utan speciella förkunskaper eller tekniska färdigheter. Utöver längre artiklar på omkring 10–15 sidor, tar tidskriften gärna emot även kortare bidrag och inlägg av notiskaraktär.

Manuskript till *Filosofisk tidskrift*:

- skickas med e-post till lars.bergstrom@philosophy.su.se
- skall utformas i överensstämmelse med den typografi som är normal i tidskriften, utan onödig formatering
- skall vara försedda med namn och adress
- skall som regel vara skrivna på svenska, och citat från andra språk bör översättas till svenska
- särskild litteraturförteckning upprättas vid behov i alfabetisk ordning och placeras sist i manus
- för icke beställt material ansvaras ej
- korrektur läses i regel endast av redaktören
- införda bidrag honoreras inte
- i stället för särtryck erhåller författaren, gratis, 10 exemplar – för recensioner 5 exemplar – av det nummer av tidskriften i vilket bidraget varit infört