

# FILOSOFISK TIDSKRIFT

2/93

DAN EGONSSON

*Om dödandet och utbyt-  
barheten av personer*

---

LARS BERGSTRÖM

*Att välja världsbild*

---

CLAUDIO M. TAMBURRINI *Ett försvar för spöstraff*

---



# INNEHÅLL

*Om dödandet och utbytbarheten av personer* 3

DAN EGONSSON

*Att välja världsbild* 16

LARS BERGSTRÖM

*Ett försvar för spöstraff* 30

CLAUDIO MARCELLO TAMBURRINI

*Recensioner* 47

PETER PAGIN om S Danielssons *Filosofiska Utredningar*

*Notiser* 59

Filosofisk tidskrift nr 2 1993 årg 14. Filosofisk tidskrift utges av Stiftelsen Filosofisk tidskrift och Bokförlaget Thales.

*Redaktör och ansvarig utgivare:* Lars Bergström

*Redaktionskommitté:* Aant Elzinga (Göteborg) Göran Hermerén (Lund), Göran Lantz (Uppsala), Sven-Eric Liedman (Göteborg), Åke Löfgren (Karlstad), Giuliano Pontara (Stockholm), Dag Prawitz (Stockholm) och Dag Westerstahl (Stockholm)

*Produktion och prenumerationer:*

Bokförlaget Thales, Box 50034, 104 05 Stockholm, redaktionssekreterare Bitte Malmnäs, tel 08 16 33 58 må, on, fr.

*Prenumeration:* Tidskriften kommer ut med 4 nummer/år, pris 120 kronor, pg 507991-8 Filosofisk tidskrift

*Lösnummer:* 35 kronor

Satt av Ord och Grafik, Stockholm

Tryckt hos Lagerblads tryckeri AB, Karlshamn, 1993

ISSN 0348-7482



## DAN EGONSSON

### *Om dödandet och utbytbarheten av personer*

---

Vi antar normalt att det i en acceptabel etisk teori skall finnas starka spärrar mot att döda personer. Detta har legat utilitarismen i fatet, eftersom man från olika håll har menat att utilitarismen saknar tillräckligt starka sådana spärrar. Därför har det varit viktigt för många företrädare för moderna varianter av utilitarismen att bemöta denna kritik. Jag tänkte se lite på hur Peter Singer går tillväga för att lösa det här problemet. Valet har fallit på honom, dels därför att han är en välkänd företrädare för utilitarismen också utanför fackfilosofin, dels därför att hans kanske mest kända arbete *Praktisk etik* sedan en tid finns i svensk översättning (av Jonas Josefsson, 1990).

#### *1. Dödandet av personer*

Singer hävdar att det i klassisk utilitarism inte har någon direkt betydelse om en varelse som dödas är en person eller bara en varelse med sensorisk förmåga. Klassisk utilitarism "bedömer handlingar på grundval av deras tendens att maximera välbehag eller lycka och minimera smärta eller olycka" (s 78). En person är en rationell och självmedveten varelse med begrepp om sitt eget förflutna och sin framtid (vilket bygger på John Lockes personbegrepp, s 75). Därför är det möjligt bara för en person att ha preferenser angående sin egen framtid, preferenser som alltså kan frustreras om hon dödas. Men antar vi att en person som vill fortsätta leva kan dödas smärtfritt och (för henne själv) omärkligt, så tycks vi i klassisk utilitarism inte kunna resa några invändningar mot att så sker, i varje fall inte med hänvisning till att något inträffar som personen i fråga inte vill skall inträffa. De unika förmågor som personen har i kraft av att hon är en person tycks alltså i detta sammanhang inte ha någon direkt relevans för klassisk utilitarism. Häri ligger en avgörande skillnad mellan

klassisk utilitarism och Singers preferensutilitarism, vilken inte värderar "handlingar utifrån deras tendens att maximera lycka och minimera smärta, utan utifrån i vilken utsträckning handlingarna är i överensstämmelse med de preferenser de hyser som påverkas av handlingarna eller deras konsekvenser" (s 79). Ett av de direkta skälen varför preferensutilitaristen inte godtar att personer dödas, även om det sker på ett "humant" sätt, är således att personen inte vill bli dödad.

Singer är medveten om att detta inte nödvändigtvis innebär en stark spärr mot att döda personer i preferensutilitarismen, men han tycks ändå hävda att denna teori reser större hinder mot sådant dödande än vad den klassiska utilitarismen gör (s 80).

Låt oss nu göra följande antagande: de känslor som har positivt värde i klassisk utilitarism är normalt sett föremål för en positiv preferens som står i proportion till dessa känslors styrka, och samma förhållande gäller normalt för negativa känslor och dito preferenser. Givet detta antagande, vilket jag tror är rimligt att göra för genomsnittsfallet, så ser vi att allt av värde i klassisk utilitarism också kommer att fångas upp, med bibehållet storleksförhållande mellan olika värden, av preferensutilitarismen, vilket betyder att alla de skäl som den klassiska utilitaristen kan anföra mot att döda personer också kan anföras av preferensutilitaristen.

För att dra slutsatsen att dödande av en person är allvarligare i preferensutilitarismen än i klassisk utilitarism, måste man således kunna visa att det omvända förhållandet inte råder, dvs att inte allt av värde i preferensutilitarismen med bibehållet storleksförhållande mellan olika värden också kommer att tas i beaktande av klassisk utilitarism. Och att visa detta tycks inte vålla något större problem. Visserligen tar klassisk utilitarism hänsyn till ett frustrerat intresse av att fortsätta leva, så till vida att det kommer att jämföras med en förlust av ett framtida värde, men det kommer inte i denna teori att dras något ifrån det totala värdet som personen hittills samlat i sin levnad. Det smärtfria och omärkliga dödandet av en person som vill leva vidare betyder enbart en avslutning av hennes personliga nytto-kalkyl, dvs den kalkyl där endast värden för henne själv figurerar, och inte, som för preferensutilitaristen, att något måste föras till minussidan innan en sådan avslutning kan ske.

Jag tror nog att det också finns fog för slutsatsen att det enligt

preferensutilitarismen i normala fall är värre att döda en person än att döda en annan varelse (s 80), men det är antagligen ett misstag att tro att personen är den enda typ av varelse som kan tillfogas direkt skada av att hon dödas, i betydelsen att hennes preferenser frusteras. Det är nämligen inte orimligt att hävda att även icke-personen har preferenser som riktar sig mot framtiden (somliga skulle kanske mena att alla preferenser rör framtiden, eftersom de rör objekt som innehavaren av preferensen ännu inte är i besittning av, och givet att icke-personer över huvud taget kan ha preferenser så är det också trivialt sant att de kan ha preferenser som rör framtida händelser). Det är fullt möjligt att tänka sig ett djur som har begrepp om skillnaden mellan en förestående och redan inträffad upplevelse utan att för den skull ha ett begrepp om sig själv som en existens med både ett förflutet och en framtid. Och redan denna logiska möjlighet skulle kunna vara användbar i ett argument. Men jag tror också att det verkligen förhåller sig på det här sättet, att t ex hönan som kämpar i trängseln för att komma åt fodret har ett begrepp om vad som väntar om hon lyckas ta sig fram till foderapparaten och att det är ganska vanligt att djuren i sitt dagliga liv, precis som vi, drivs av sådana förväntningar på framtiden, låt vara att dessa hos de djur som saknar personstatus antagligen är mera kortsiktiga. Detta skulle ju innebära att även ett sådant djur kan vederfaras något ont om det möter döden innan en preferens med ett objekt placerat i framtiden har tillfredsställts, och att denna omständighet måste medräknas i preferensutilitarismen. Men eftersom detta onda inte upplevs som något ont, så behöver det givetvis inte medräknas i den klassiska utilitarismen.

Dödandet av en varelse, av en person eller en icke-person, kan alltså i detta avseende sägas vara värre enligt preferensutilitarismen än enligt den klassiska utilitarismen. Hur skall man då kunna slå fast den ytterligare slutsatsen att dödandet av personer är värre än dödandet av icke-personer i preferensutilitarismen (även bortsett från de indirekta effekter som dödandet av en person kan ha på andra personer)? Åtminstone två argument står till buds.

(1) Man kan hävda att en person värderar normalt att leva i högre grad än vad icke-personen värderar positiva upplevelser i den närmaste framtiden. Detta skulle betyda att preferensfrustrationen blir jämförelsevis större när personen dödas. Bevisningen i samband med detta påstående måste bli bristfällig, eftersom vi inte med säkerhet

kan säga hur pass starka icke-personernas preferenser verkligen är. Men nog är det rimligt att tänka sig att denna skillnad finns mellan personer och icke-personer när den enligt alla vanliga begrepp finns inom klassen av personer. Nog värderar vi normalt sett att leva mera än vi värderar olika slag av upplevelser på kort sikt? Man kan således misstänka att (de psykologiska) möjligheterna för preferenser av en viss dignitet är begränsade hos icke-personerna på grund av deras bristande begreppsliga förmåga.

(2) Det är vanligare att det saknas skäl att avstå från att döda en icke-person än att döda en person. Jag tänker här på två olika situationer.

a) Det är inte realistiskt att tillvaron för t ex ett djur som saknar personstatus innebär att det lever under ständig utsikt att ha positiva upplevelser i den nära framtiden. Därför finns det troligen en mängd situationer där detta djur inte berövas ett nära förestående gott om det dödas och där vi därför inte heller kan tala om någon preferensfrustration som en följd av dödandet. Om en person har en önskan att leva vidare så kommer denna önskan att frustreras (mer eller mindre) oberoende av vad slags upplevelser som hon går miste om eller undviker genom att vi dödar henne. Vi kan därför inte på samma sätt vänta ut ett tillfälle där det inte vore fel att döda henne.

Emellertid utgör inte heller detta en absolut skillnad mellan icke-personen och personen. Om tillfället när det inte skulle vara fel att döda icke-personen är en situation där denna ändå inte har chans till kortsiktigt önskade upplevelser, så kommer det motsvarande tillfället i samband med personen att utgöras av en situation där hon (temporärt) inte längre har en önskan att leva. Men här är min gissning att dessa tillfällen, när personen saknar en preferens att leva, är inte lika frekventa som de tillfällen när icke-personerna av en eller annan anledning saknar en preferens riktad mot en nära förestående upplevelse.

b) Det kan antagligen inte finnas några direkta skäl att avstå från att döda den sovande icke-personen, eftersom det slag av preferenser riktade mot den nära framtiden som vi räknat med kan finnas hos icke-personer ju knappast kan existera hos en sovande varelse (vi behöver inte diskutera vilken status preferenser har i en varelses drömmar). Icke-personen har inte somnat in med en preferens om vad som skall ske när den vaknar. Därtill saknar den behövlig begreppsapparat. Däremot kan ju personen mycket väl somna med en prefe-



rens som handlar om vad som skall hända med henne när hon åter vaknar, och denna preferens kan då också bli frustrerad om någon dödar henne i sovande tillstånd.

Om dessa slutsatser, (1)–(2), stämmer så ser vi att vad som gör att dödandet av personer antagligen normalt är mera allvarligt än dödandet av icke-personer i preferensutilitarismen är en sammanvägning av omständigheter och förhållanden, där förmågan hos personen att ha ett begrepp om sin framtid, även om den antagligen spelar den viktigaste rollen, kan göra så också på ett mera indirekt sätt. Det är inte alltid vad preferensen är *riktad mot* som är avgörande, utan snarare vad den har för styrka.

## 2. Utbytbareheten av personer

De skäl som finns mot att döda personer i preferensutilitarismen kan dock knappast bilda underlag för en rätt till liv för personer. Man kan t o m hävda att vi inte på något sätt lyckats visa att det är (moraliskt) allvarligt att döda personer enligt preferensutilitarismen; vad vi visat är bara att det är *allvarligare* än i den klassiska utilitarismen. För antag att vi kompenserar för den preferensfrustration som enligt preferensutilitarismen blir följden om en person som vill leva vidare dödas. En sådan kompensation kan ske på två sätt. Antingen genom att man försäkras sig om att preferenstillfredsställelsen på ett optimalt sätt ökar hos andra existerande personer och varelser på så sätt att totalsumman blir större än den skulle ha blivit om personen i fråga hade fått leva vidare. Här är det fråga om att de andra existenserna i slutändan kommer att tjäna på att personen dödas. Kompensation kan också ske genom att man försäkras sig om existensen av en ny varelse eller person som kommer att åtnjuta (vad man nu vill att det skall betyda) ett sådant mått av preferenstillfredsställelse att det uppväger förlusten av tillfredsställelse hos den person som dödas. Här är det fråga om att personen kommer att bytas ut och få en (eller flera) ersättare. Vilket sätt vi än tänker på, så är det uppenbart att det är andra som kompenseras av att personen dödas (fast det är inte utslutet att också tänka sig en tredje metod som består i att personen kompenseras innan hon dödas).

Trots att vi konstaterat att det normalt är värre att döda personer i preferensutilitarismen än i klassisk utilitarism, så finns det inget i det som vi hittills sagt som hindrar oss från att betrakta personer som

utbytbara även i preferensutilitarismen. Det går an att döda personer om man bara ordnar med ersättare för dem, och det är fortfarande bara mängden preferenstillfredsställelse som räknas, inte på något sätt vems denna tillfredsställelse är. Det förefaller alltså som om kritiken mot utilitarismen kvarstår. Även preferensutilitarismen har en alltför liberal syn på dödandet och det hjälper inte mycket att dödandet av personer i denna teori moraliskt har försvårats jämfört med klassisk utilitarism, så länge personer betraktas som utbytbara i teorin.

Det finns emellertid ett sätt för Singer att förhindra (åtminstone delar av) denna kritik. Singer ber oss att jämföra två olika grunduppfattningar om nyttomaximering som en utilitarist kan ha. Enligt "det totala synsättet" så gäller det "för oss att öka det totala måttet av välbehag (och minska det totala måttet av smärta) oavsett om detta görs genom att välbehaget hos existerande varelser ökas, eller genom att antalet existerande varelser ökas" (1990, s 85; för att i stället knyta synsättet till preferensutilitarismen kan man ersätta "välbehag" och "smärta" med "preferenstillfredsställelse" och "preferensfrustration"). Den andra uppfattningen, som Singer kallar för "det existens-fokuserande synsättet", "består i att man endast tar med i beräkningen de varelser som redan existerar före det beslut vi fattar eller åtminstone existerar oberoende av beslutet" (s 85). Nu gäller det endast för "det totala synsättet" att en varelse är utbytbar, eftersom det bara är där som det är tillåtet att kalkylera med de framtida varelser vilkas existens är beroende av våra beslut.

Detta betyder att om Singer vill undanröja konsekvensen att personer är utbytbara i preferensutilitarismen, så måste han tillämpa "det existens-fokuserande synsättet" på dem. Detta är också vad han gör (s 100), men samtidigt vill han tillämpa "det totala synsättet" på varelser som inte är personer, som alltså inte är rationella och självmedvetna. Anledningen till att han gör denna åtskillnad är, som jag uppfattar det, att det endast är personerna som kan sägas ha ett intresse av ett förbud mot utbytbarhet. Det är endast de som har ett personligt förhållande till sitt eget liv, vilket betyder att deras livsinnehåll inte kan betraktas i termer av utbytbar lycka eller preferenstillfredsställelse. Singer skriver: "Vi kan hålla med om att 'det totala synsättet' är tillämpligt när det gäller varelser som inte kan existera som individer som lever sina egna liv. I dessa fall är det tillrådligt att man endast tar hänsyn till maximeringen av lycka. Men

när det i stället handlar om självmedvetna varelser finns det mer än opersonliga kvantiteter av lycka att ta hänsyn till” (s 100).

För att testa följderna av denna uppdelning på olika synsätt, skall jag presentera några möjliga konfliktsituationer.

(1) *Existerande personer mot existerande icke-personer*. Om vi definierar en person som en varelse som är rationell och självmedveten, så är det uppenbart att det finns människor som inte är personer, t ex nyfödda och mentalt gravt retarderade. Dessutom, menar Singer (ss 95–6), finns det personer som inte är människor, t ex schimpanser, valar, delfiner och kanske t o m grisar. Vi behöver därför inte, om detta stämmer, gå utanför klassen av (icke-mänskliga) djur för att ställa personerna och icke-personerna mot varandra.

Låt oss anta att vi ansvarar för en djurpark med så pass begränsade resurser att vi måste välja mellan att behålla schimpansburen och påfågelsgården och alltså mellan att döda antingen schimpanserna eller påfåglarna. Vi tänker oss nu att de djur som finns i schimpansburen är personer medan de som finns i påfågelsgården är icke-personer. Vidare antar vi, väl medvetna om det vanskliga i dylika uträkningar, att schimpansernas sammantagna preferenstillfredsställelse är något lägre jämfört med påfåglarnas. Låt oss anta att det finns många flera påfåglar än schimpanser och att detta är huvudförklaringen till att deras sammanlagda preferenstillfredsställelse uppväger schimpansernas.

Vad skall vi nu göra? Det framgår inte av vad vi sagt om uppdelningen på olika synsätt, eftersom vi där bara fastställde vilket synsätt som skulle gälla för de respektive varelserna. Uppdelningen löste endast frågan om en konflikt mellan olika synsätt vad gäller en och samma typ av varelse. Vad som nu målas upp är en situation där vi, givet uppdelningen, måste välja mellan att antingen göra vad ”det totala synsättet” uppmanar oss att göra med icke-personerna eller vad ”det existens-fokuserande synsättet” säger att vi skall göra med personerna. Konflikten finns alltså i någon mening mellan de olika synsätten, eftersom vi tar ställning till ett synsätt i samma ögonblick som vi tar ställning till vilken typ av varelse som skall få fortsätta att leva.

Antag då att vi bestämmer oss för att rädda schimpanserna. Detta beslut kommer att bli svårt att försvara, eftersom vi i så fall kommer att välja en handling som resulterar i en mindre mängd nytta, alltså preferenstillfredsställelse, än det alternativ som står till vårt för-

fogande. Varför föredra det mindre optimala av två möjliga alternativ? En förklaring kunde vara att man ansluter sig till en syn på nyttan där mera än de rent kvantitativa aspekterna har en inverkan vid värderingsfrågor. Skillnaden i värde mellan vad schimpanserna får om de räddas och vad påfåglarna får skulle alltså bestå av något annat än den skillnad i värde som dessa objekt har *för subjekten själva*. Eftersom jag inte kan se någon anledning att tolka Singer i en sådan objektivistisk riktning, så skall jag anta att ett sådant försvar för beslutet är ohållbart, åtminstone för Singer (jfr s 65).

Men antag då i stället att vi väljer att rädda påfåglarna, med hänvisning till att det övergripande utilitaristiska målet att maximera preferenstillfredsställelsen kräver detta. Det blir, enligt denna tanke, den kvantitativa skillnaden i preferenstillfredsställelse som avgör vilket av synsätten som skall tillämpas. Eftersom utsikten till sådan tillfredsställelse är större för påfåglarna, så blir det dessa som bör räddas och eftersom de är icke-personer, så skall "det totala synsättet" tillämpas på dem. Nu är det emellertid troligt att detta ställningstagande kommer att dra på sig samma slag av kritik som den som ursprungligen var en av anledningarna till att göra uppdelningen mellan olika synsätt.

Den övergripande kritiken handlade om att utilitarismen tar alltför lätt på dödandet av personer. Men ligger det inte nära till hands att samma kritik även kommer att riktas mot idén att man i händelse av en konflikt bör föredra att tillgodose icke-personers intressen framför personers så snart de förras intressen har en kvantitativ övervikt (vars storlek inte har någon betydelse så länge som den bara är positiv). Det är sant att detta förhållande inte är det samma som att säga att personer är utbytbara, men väl, kan man mena, *eliminera*bara. Vi har inte sagt att personer kan ersättas av icke-personer i den meningen att vi dödar personer för att därmed bereda plats för framtida icke-personer, men vad vi har sagt är, att om det existerar en konflikt mellan samtida personer och icke-personer, så bör (under de förhållanden som beskrivits ovan) personerna elimineras till förmån för icke-personer. Och detta verkar inte vara en speciellt stark spärr mot att döda personer. Det här tycker jag visar att vårt exempel kan skapa ett dilemma för Singer: att rädda personerna framför icke-personerna förefaller vara partiskt, att göra tvärtom förefaller däremot att beröva personerna varje form av rätt till liv. Jag menar inte att Singer

någonsin har hävdad att personer har en rätt till liv i ordets fulla bemärkelse (jfr ss 80–3), men han menar att personer inte är ersättningsbara. Den retoriska frågan blir: varför skulle det vara värre att vara ersättningsbar än eliminerbar?

(2) *Existerande personer mot potentiella icke-personer*. Antag att vi modifierar exemplet i den föregående sektionen på så sätt att vi som ansvariga för en djurpark har att välja mellan att behålla schimpanserna eller att döda dem för att i stället använda de frigjorda resurserna till att skaffa ett fågelhus där man gör bedömningen att det kan "produceras" en större mängd preferenstillfredsställelse än i schimpansburen.

Låt oss även här först anta att vi väljer att behålla schimpanserna. Det kan då på samma sätt som i den förra sektionen tyckas vara ett utslag av partiskhet – vi är själva personer och vill därför gynna andra personer. Men låt oss i stället fråga hur stark denna partiskhet är, i betydelsen hur mycket nytta hos potentiella icke-personer som vi är beredda att offra för nyttan hos existerande personer. Om förtjänsten räknad i nytta blir tio gånger större om vi skaffar fågelhuset jämfört med om vi behåller schimpansburen, bör vi då fortfarande hålla fast vid beslutet att behålla schimpansburen? Och om så är fallet, hur ser den princip ut på vilken vi grundar ett sådant beslut? En tänkbar formulering i termer av "det totala synsättet" är:

(TS1) Maximera preferenstillfredsställelsen hos existerande och potentiella icke-personer endast i de fall där detta inte negativt påverkar preferenstillfredsställelsen hos existerande personer.

Vad denna formulering implicerar är att närhelst det uppkommer en konflikt mellan en existerande person och en existerande eller potentiell icke-person, så är det personens intressen som har företräde. Men TS1 måste sägas uttrycka en väldigt stark form av partiskhet, eftersom den ger vid handen att det skulle vara fel att frustrera varje slags intresse hos personen, alltså inte bara preferensen att leva utan också t ex en preferens för kroppslig njutning, till förmån för ett eller naturligtvis flera intressen hos icke-personer. Vad jag kan se blir det med TS1 inte bara rätt att döda existerande icke-personer om detta är enda sättet att förhindra den allra minsta olägenhet för existerande personer, det blir dessutom rätt att omöjliggöra den framtida existensen av alla icke-personer, t ex alla djurarter, om man samtidigt kan

visa att man därigenom gör den minsta lilla vinst för personerna.

Det är uppenbart att Singer inte ställer sig bakom en så pass kategorisk formulering som TS1. Han har själv sagt om sin uppdelning av synsätt som ett svar på problemet om dödandet i utilitarismen: "Det är inget riktigt utarbetat svar, eftersom det inte säger något om naturen hos det företräde vi ger till självmedvetna varelser. Skall det vara en lexikalisk prioritering, så att ingen förlust för en självmedveten varelse kan rättfärdigas genom skapandet av ett ytterligare antal lyckliga varelser? Eller finns det ett icke godtyckligt sätt att väga förluster för självmedvetna varelser mot skapandet av flera lyckliga varelser?" (1979, ss 152–3.) Eftersom Singer, även om han här tydligt visar att detta är viktiga frågor som måste besvaras, inte menar sig ha något svar på dem, så bör vi inte betrakta det som sägs här, om problemen i samband med uppdelningen av de två synsätten, som på något sätt ödesdigert. Men så länge som frågorna kvarstår obesvarade kan vi heller inte utan vidare acceptera Singers förslag.

Antag nu att vi väljer att skaffa ett fågelhus på bekostnad av de redan existerande personerna – man dödar schimpanserna för att i stället kunna skaffa sig ett stort antal icke-personer som totalt sett kommer att bidra med en större mängd preferenstillfredsställelse i världen än vad schimpanserna kan göra. I så fall kan man ansluta sig till följande formulering av "det totala synsättet":

(TS2) Maximera preferenstillfredsställelsen hos existerande och potentiella icke-personer i de fall där detta inte förhindrar realiserandet av en större mängd (jämfört med icke-personerna) preferenstillfredsställelse hos existerande personer.

Och, kunde man kanske tillägga, eftersom det sannolikt maximerar nyttan på det stora hela att vara så rättvis som möjligt, så bör vi nog också sträva efter jämn fördelning mellan de tillfällen då personerna och icke-personerna gynnas när mängderna av preferenstillfredsställelse är lika.

TS2 kommer först av allt, liksom fallet var ovan, att dra på sig den kritiken att den tar alltför lätt på värdet av en existerande person jämfört med värdet av en existerande icke-person. Redan här får formuleringen alltså kontroversiella följder. Men den är också kontroversiell i en kanske allvarligare bemärkelse. TS2 säger alltså att det är rätt att döda personer till förmån för potentiella icke-personer så

snart som vi har anledning att tro att det skulle maximera nyttan. Men detta skulle alltså betyda att de existerande personerna är *utbytbara* mot potentiella icke-personer. Hur skulle det vara möjligt att acceptera en sådan konsekvens när anledningen till att de två synsätten infördes var att omintetgöra personers utbytbarhet? Och inte borde det göra någon skillnad att det från början handlade om personer som ersattes av potentiella personer, medan det nu handlar om personer som ersätts av potentiella icke-personer. Om man har svårt att acceptera den första av dessa båda situationer, så borde man ha ännu svårare för att acceptera den andra.

Slutsatsen måste alltså bli, att ungefär samma sorts dilemma uppstår i samband med det här exemplet: risken är att antingen visa prov på partiskhet eller att avhända sig möjligheten att visa att utilitarismen trots allt kan ha förhållandevis starka spärrar mot att döda personer.

(3) *Existerande personer mot existerande personer.* Vad innebär "det existens-fokuserande synsättet" för personer? Antag att en grym diktator håller hundra personer i ett koncentrationsläger. Där inser vi att de kommer att svälta ihjäl om vi inte på något sätt oskadliggör diktatorn. Emellertid så är förhållandena sådana att enda sättet att göra detta tycks vara att döda honom. Vad skall vi göra? Att döda diktatorn för att rädda de hundra personerna och därmed se till att preferenstillfredsställelsen maximeras innebär att vi betraktar diktatorn såsom varande eliminerbar – han måste offras för att det större antalet människor skall kunna räddas. Och något annat alternativ tycks inte vara möjligt. Att avstå från att döda diktatorn skulle betyda att man lät de hundra personerna i koncentrationslägret dö. Och detta, menar Singer, är inte i sig bättre än om man själv hade dödat dem; han omfattar sålunda "slutsatsen att det inte finns någon *intrinsic* skillnad mellan att döda och att låta dö" (1990, s 145). Så om vi bortser från allt sådant som på indirekta grunder ger distinktionen en moralisk relevans (t ex den omständigheten att det normalt sett är svårare att avstå från att låta folk dö jämfört med att avstå från att döda dem) så tycks Singers åsikt vara att vi bör döda diktatorn.

Men om vi accepterar att personer under dessa omständigheter är eliminerbara, dvs när de har ett ont uppsåt som är livsfarligt för andra personer, vad skall vi då säga om ett fall där förhållandena för en grupp personer är sådana att den totala preferenstillfredsställelsen

skulle öka om gruppens storlek reducerades något? Detta beror varken på några onda avsikter eller på att någons beteende i den stora gruppen är livshotande för någon annan; ingen står direkt i vägen för någon annans lycka, men det från nyttomaximeringens synvinkel optimala antalet människor är alltså något lägre än antalet människor i den ursprungliga gruppen. Bör vi också här eliminera folk till förmån för den totala mängden preferenstillfredsställelse? Ja, det tycks i alla fall inte som om "det existens-fokuserande synsättet" hindrar något sådant, och det betyder, återigen, att uppdelningen mellan de båda synsätten tycks vara ett klenkt verktyg om vi vill förhindra moraliskt sanktionerat dödande av personer (vilket också R G Frey har påpekat, 1983, s 163). Och den retoriska frågan blir om det inte är något konstigt med en teori som vill förhindra att personer ersätts med framtida personer, men godkänner att de elimineras för att därmed åstadkomma förhöjd tillfredsställelse hos sina samtida. Det finns ett spänningsförhållande här, även om det inte rör sig om direkt inkonsekvens.

Låt mig nu avsluta med att diskutera villkoren för utbytbarhet och elimination i preferensutilitarismen, bortsett från uppdelningen på synsätt. Även dessa villkor måste förstås spegla det förhållande som vi redan tidigare noterat, nämligen att dödandet av en person som vill fortsätta leva är något direkt ont i preferensutilitarismen, men inte i den klassiska utilitarismen.

Vi kan se på ett förenklat räkneexempel som ändå illustrerar skillnaden. Antag att mängden framtida nytta i en persons liv motsvaras av siffran 10. Denna summa inbegriper då också den *tillfredsställda* preferensen att leva. Låt oss säga att halva totalsumman av nytta kommer från denna preferens. Om en sådan person dödas, så kommer 5 enheter nytta (från andra preferenser än livsintresset) att försvinna. Om nu livsintresset fungerat på samma sätt som dessa andra preferenser skulle räkneoperationen avslutas genom att vi drog av ytterligare 5 enheter nytta (för livsintresset) och kom fram till summan 0. Men i så fall skulle vi behandla dödandet uteslutande som ett berövande av något framtida gott (preferenstillfredsställelse) och inte dessutom som ett tillfogande av något ont (preferensfrustration). För preferensutilitaristen fullbordas räkneoperationen först när vi medräknat  $-5$  enheter från den frustrerade preferensen att leva vidare. Dödandet betyder alltså både en förlust av något framtida gott samt ett till-



fogande av något ont, även om vi enbart ser till livsintresset. Slutsumman framtida nytta för personen i fråga om hon dödas måste därför bli negativ, nämligen  $-5$ , eftersom den blir 0 bara om vi betraktar det frustrerade livsintresset som icke-realiserandet av något gott (vilket man alltså gör i klassisk utilitarism).

Det räcker således inte att ersätta en person med en annan person som har samma nyttopotential (10) som den första personen ( $-5+10$  blir bara 5, vilket alltså blir en lägre nivå än den ursprungliga), utan den framtida personen måste ha väsentligt *bättre* utsikter för att bytet skall gå jämnt ut.

Samma sak gäller för ett möjligt fall av elimination. Antag att den framtida nyttan för vardera existerande personerna *A*, *B* och *C* är 10 enheter. Deras sammanlagda nytta är således 30 enheter. Antag nu att om *A* dödades så skulle både *B*'s och *C*'s nytta höjas från 10 till 15, alltså till 30 sammanlagt. Detta skulle emellertid inte räcka för att eliminera *A*, eftersom dödandet av honom också innebär en negativ nyttomängd (om ett livsintresse frustreras) som också måste kompenseras för. För att uppväga dödandet av *A* således, så måste den sammanlagda extra nyttan för *B* och *C* vara väsentligt högre än *A*'s ursprungliga nyttopotential, dvs den nytta som skulle komma *A* till del om han inte hade dödats.

Slutsatsen av detta är att även om både utbytbarhet och eliminerbarhet skulle gälla för personer i preferensutilitarismen, så ställer ändå denna variant av utilitarism hårdare krav på de situationer där det skulle kunna vara moraliskt försvarbart (eller moralisk plikt) att praktisera dessa ting än vad klassisk utilitarism gör.

### Litteratur

FREY, R G: *Rights, Killing, and Suffering*, Blackwell, Oxford, 1983.

SINGER, PETER: *Praktisk etik*, Thales, Stockholm, 1990 (övers Jonas Josefsson).

SINGER, PETER: "Killing Humans and Killing Animals" i *Inquiry*, vol 22, nr 1–2, ss 145–56, 1979.

## LARS BERGSTRÖM

### *Att välja världsbild*

---

Livsåskådningsforskning – eller ”tros- och livsåskådningsvetenskap” – är ett förhållandevis nytt universitetsämne, som ligger inom den teologiska fakulteten. Livsåskådningar, och delar av livsåskådningar, studeras naturligtvis också på andra håll, inom samhällsvetenskapliga, humanistiska och mer traditionella teologiska ämnen. Förr brukade man för övrigt ofta säga att livsåskådningsproblem är vad man studerar inom den praktiska filosofin (medan däremot världsåskådningsproblem behandlas inom den teoretiska filosofin).

Vad är då skillnaden mellan livsåskådningsforskning och de studier av attityder, religioner, moralläror, ideologier, metafysiska system och så vidare som förekommer inom äldre ämnen som exempelvis religionshistoria, sociologi, antropologi, statsvetenskap, idéhistoria och filosofi? Ett sätt att närma sig ett svar på denna fråga är att studera en nyutkommen bok av Carl Reinhold Bråkenhielm, *Människan i världen. Om filosofi, teologi och etik i våra världsbilder*, Uppsala 1992.

I bokens inledning definieras *livsåskådningsforskning* som ”studiet av världsbild och värderingar i människors tänkande” (s 11). Detta antyder att en livsåskådning enligt Bråkenhielm består av två komponenter, en världsbild och vissa värderingar – eller möjligen att en livsåskådning kan vara antingen en världsbild eller en uppsättning värderingar. Att det finns samband mellan livsåskådningar och världsbilder skulle väl de flesta hålla med om. Däremot är det kanske lite ovanligt att se världsbilder som delar av livsåskådningar. Hur som helst världsbilder är uppenbarligen ett viktigt studieobjekt för livsåskådningsforskare, och det är framför allt världsbilder som Bråkenhielm intresserar sig för i boken.

## 1. Världsbilder

Vad är då en världsbild? Denna fråga behandlas i första kapitlet. Där ges emellertid två olika svar på frågan, nämligen (1) att en världsbild är en helhetstolkning av universum, samt (2) att en världsbild är en helhetstolkning som ger tillvaron mening och sammanhang (s 14). I slutet av boken sägs också (3) att en världsbild är ”en övergripande modell av tillvaron” (s 91). En ”modell” skall då förstås som ”en välbekant idé (t ex en mängd olika biljardbollar) som används för att tolka och åskådliggöra ett abstrakt fenomen (t ex en gas)”, och Bråkenhielm hävdar också att ”människan formar sin bild av verkligheten med hjälp av modeller som endast ofullständigt och indirekt ger oss tillgång till verkligheten själv” (s 53). En maskin – eller kanske idén om en maskin – kan exempelvis utgöra en modell av universum (jfr s 53). Bråkenhielm menar också att man bör skilja mellan en modell och en tolkning av denna modell. Modeller behöver inte tolkas ”bokstavigt”. De kan t ex också tolkas ”fiktionalistiskt”, och i så fall ”säger de egentligen ingenting om verkligheten” (s 53).

Detta är rätt förvirrande. Låt oss bortse från frågan om en gas, eller tillvaron, eller universum verkligen kan anses vara ett ”abstrakt fenomen” – och om en mängd biljardbollar utgör en idé. Låt oss också undvika att grubbla över om en modell som inte säger något om verkligheten kan vara en världsbild. Men kan Bråkenhielm verkligen mena att en maskin kan vara en världsbild? Jag tror inte det, och inte heller menar han väl att en tolkning av en maskin kan vara en världsbild. Vad han syftar på är väl snarare *påståendet* eller *uppfattningen* att universum är en maskin, eller att universum liknar en maskin. Detta stämmer också överens med att han tycks mena att världsbilder (åtminstone delvis) kan vara sanna eller falska (jfr s 87). (Här kanske någon skulle kunna få för sig att jag läser Bråkenhielm som fan läser bibeln. Men så är det inte alls. Bibelns författare uttrycker sig inte på långa vägar lika slarvigt som Bråkenhielm, och det krävs ingalunda någon djävulsk illvilja för att missförstå den senare.)

I själva verket menar Bråkenhielm säkerligen att en världsbild normalt är en uppsättning av påståenden eller uppfattningar om verkligheten. Detta låter ytterst rimligt, och det stämmer också bra överens med de exempel på världsbilder som han skisserar i kapitlen 2, 3 och 4. En fråga som kvarstår är förstås om man skall ta fasta på

(1) eller (2) ovan. Är det rent definitionsmässigt så, att varje världsbild ”ger tillvaron mening och sammanhang”? Förmodligen vill Bråkenhielm även räkna med världsbilder som *för honom själv* inte ger tillvaron mening. Men kanske kräver han att varje världsbild ger tillvaron mening för åtminstone *någon* person? Min gissning är att han inte ens kräver detta, men säker är jag inte.

Däremot verkar det klart att han är primärt intresserad av världsbilder som är relativt innehållsrika. Han bortser från ”fragmentariska och oartikulerade” världsbilder och koncentrerar sig på sådana som har ett visst ”djup och inre sammanhang”, som är ”etablerade” (s 15) och som är ”genomarbetade” (s 11). Han tänker sig att det finns olika ”verklighetsområden” – i första hand Gud, människan och naturen – och han kräver av en etablerad världsbild att den dels anger vilka verklighetsområden som finns och dels säger något om hur dessa områden förhåller sig till varandra, speciellt att den anger vilket område som är ”det mest centrala” (jfr s 16). I enlighet med detta skiljer han sedan mellan teocentriska, antropocentriska och biocentriska världsbilder.

Vidare tycks han mena att ”etablerade” världsbilder utöver metafysiska uppfattningar kan innehålla även kunskapsteoretiska och etiska komponenter (jfr s 17–20). Det innebär alltså att en världsbild kan vara ganska omfattande. En religion är däremot inte en världsbild (s 18–9), men jag antar att Bråkenhielm menar att t ex kristendomen innehåller en världsbild – eller kanske att olika varianter av kristendom kan innehålla olika världsbilder.

## 2. Varför studera världsbilder?

Detta med kristendomen är säkerligen en viktig punkt. Man får nämligen intrycket att livsåskådningsforskningen, liksom teologernas intresse för olika livsåskådningar och världsbilder, har något att göra med kristendomens försvagade ställning i samhället. Bråkenhielm påpekar att ”den traditionella kristendomen försvagas alltmer” (s 11), och han fortsätter:

Gamla världsbilder ersätts med nya. I en sådan situation kan det vara av särskilt intresse att klargöra tankealternativen, jämföra dem med varandra och bli på det klara med grundvillkoren för ett ställningstagande (s 12).

Med andra ord, så länge kristendomen, och därmed en kristen världsbild, är dominerande och föga ifrågasatt kan teologerna koncentrera sig på att studera och uttolka den, men när kristendomen tappar anhängare vill teologerna även studera "alternativen".

Detta kan låta vettigt, men jag menar ändå att det är något tvivelaktigt med hela projektet. Det tvivelaktiga har att göra just med det som är specifikt för teologisk livsåskådningsforskning, jämfört med dess motsvarigheter inom samhällsvetenskap och humaniora, nämligen *kombinationen* av, för det första, en inriktning på relativt genomarbetade, icke-fragmentariska och etablerade världsbilder och, för det andra, målsättningen att jämföra dessa för att ta ställning till vilken som är "den riktiga och sanna" (jfr s 87).

Ett problem med detta är att det på ett orimligt och försåtligt sätt favoriserar vad vi kunde kalla "stora paketlösningar". Mer begränsade uppfattningar om t ex metafysiska och moraliska problem kan då lätt komma att uppfattas som "fragmentariska" (för att använda Bråkenhielms ord) och därmed bristfälliga, om de inte ingår som en del av en större och i någon mening "fullständig" världsbild.

Men bör man överhuvud taget ha en "fullständig" världsbild? Är det något fel på att endast ha en "fragmentarisk och oartikulerad" världsbild? Är inte det egentligen det enda vettiga i en situation där man inte tycker sig ha några tillräckliga grunder för att tro på någon av de mer "genomarbetade" och "etablerade" världsbilder som livsåskådningsforskningen räknar med?

Dessutom kan man väl inte bortse från att olika människor har olika intressen och att vissa människor kanske är helt ointresserade av en del av de frågor som Bråkenhielm anser att en världsbild skall besvara. Hit hör t ex de som är ointresserade av religiösa ting. De tror inte att Gud finns, men de tror kanske inte heller att Gud inte finns. Likaså finns det säkert personer som inte bryr sig ett dugg om vilket "verklighetsområde" som är "det mest centrala". Sådana personer har ju inget behov av en världsbild i Bråkenhielms mening, men man kan väl inte säga att det är något fel på dem för det. Eller kan man det?

Ett annat och viktigare problem är den utgångspunkt som Bråkenhielm antyder i följande formulering när han har gått igenom en rad olika världsbilder: "Efter denna inventering av olika världsbildsalternativ uppkommer naturligtvis frågan om vilken världsbild som

är den riktiga och sanna världsbilden” (s 87). Detta antyder att en av de genomarbetade och etablerade världsbilder som Bråkenhielm har uppmärksammat är riktig och sann. Såvitt jag förstår finns det ingen som helst anledning att anta detta, och just genom att världsbilder i Bråkenhielms mening är så innehållsrika (om de inte är ”fragmentariska”), så är det väl också ytterst osannolikt. För att man skall kunna utgå från att ett av flera alternativ skall vara sant, så måste alternativen vara uttömmande. Och ju innehållsrikare alternativen är, desto fler måste de vara för att vara uttömmande. Om alternativen skall vara så innehållsrika och etablerade som Bråkenhielm tänker sig, kan man vara rätt säker på att man aldrig har en uttömmande uppsättning. Alltså finns det inte heller någon anledning att utgå från att någon av de världsbilder man har identifierat är sann.

Jag föreställer mig att det är bland annat av detta skäl som man inom filosofin – där man försöker fastställa vad som är sant i fråga om metafysik, kunskapsteori och moralfilosofi, områden som ju spelar stor roll inom världsbilder i Bråkenhielms mening – *inte* brukar ställa sig frågan om vilken världsbild som är sann, utan i stället inriktar sig på att diskutera en sak i sänder. Såvitt jag förstår är detta det enda rimliga tillvägagångssättet om man är ute efter att hitta sanningen. Livsåskådningsforskningens metod – att leta efter sanningen genom att jämföra hela livsåskådningar eller världsbilder – verkar bakvänd och oframkomlig.

### 3. Hur skall världsbilder testas?

Men däremot kan man ju, i fråga om varje enskild världsbild som man av någon anledning intresserar sig för, fråga sig om just den världsbilden är ”riktig och sann”, eller om den överhuvud taget är acceptabel. Såvitt jag förstår skulle Bråkenhielm hålla med om detta.

Bråkenhielm tycks på det hela taget anse att världsbilder bör betraktas i analogi med vetenskapliga hypoteser, och att de därmed också skall prövas på i princip samma sätt som dessa (s 87). Närmare bestämt tänker han sig att världsbilder måste vara motsägelsefria, sammanhängande, och förenliga med väletablerade vetenskapliga resultat och teorier (s 90–2), samt att man i valet mellan olika världsbilder som uppfyller dessa krav bör välja den som i störst grad uppfyller vissa ytterligare prioriteringsprinciper (som ibland kan komma i konflikt) med krav på enkelhet, räckvidd, samt ”förmågan

att på ett nyanserat sätt göra olika typer av erfarenheter och kunskaper rättvisa” (s 94). De flesta av dessa kriterier känns ju igen från vetenskapsteoretisk litteratur.

Det enda som förefaller nytt är det sista kriteriet, ”förmågan att på ett nyanserat sätt göra olika typer av erfarenheter och kunskaper rättvisa”. Bråkenhielm kallar det ”varsamhetsprincipen”. Han tycks mena att detta är en princip som brukar förekomma i vetenskapsteoretiska sammanhang, men såvitt jag förstår är det ett misstag – eller kanske snarare ett utslag av önsketänkande. Det beror visserligen på hur principen närmare bestämt skall tolkas, men här måste det ju rimligen vara fråga om något annat än de tidigare kraven på empirisk provbarhet och förenlighet med vetenskapliga resultat. Bråkenhielm förklarar att det skall innebära ”respekt för den vetenskapliga kunskapens komplexitet och trohet mot erfarenheten, dvs det som människor själva uppfattar som innehållet i sin upplevelse av tillvaron, sig själva, andra människor och naturen” (s 94–5). Det är inte lätt att veta vad han menar med det första, ”respekt för den vetenskapliga kunskapens komplexitet”, speciellt som det ju inte gärna kan vara detsamma som det tidigare kravet på förenlighet med etablerade vetenskapliga resultat och teorier. Men den andra halvan av varsamhetsprincipen, kravet på trohet mot människors egna upplevelser av tillvaron, torde innebära att en världsbild skall vara förenlig med, och på något vis inkorporera, människors mer eller mindre spontana uppfattningar om verkligheten.

Här tycks Bråkenhielm helt ha glömt bort vad han var på jakt efter, nämligen ett sätt att avgöra om en världsbild är ”riktig och sann”. Ty det finns väl ingen som helst anledning att tro att människors spontana uppfattningar om ”sig själva, andra människor och naturen” är sanna. Vore det så, behövdes ju ingen livsåskådningsforskning. Man måste komma ihåg att det här rimligen är fråga om uppfattningar som inte kan prövas med vanliga empiriska metoder, ty i så fall vore vi ju tillbaka till kravet på empirisk provbarhet. En sann världsbild skall alltså uttrycka, eller vara förenlig med, en *riktig* uppfattning om verkligheten, inte en uppfattning som vissa människor *upplever* som riktig – oavsett hur ofrånkomlig, djup eller central upplevelsen känns.

Dessutom har ju olika människor ofta helt olika och oförenliga upplevelser av tillvaron. Somliga upplever t ex att det finns en Gud,

medan andra lika starkt upplever att det inte finns någon Gud. Innebär Bråkenhielms varsamhetsprincip att en sann världsbild skall vara trogen mot alla sådana upplevelser? Det tror jag inte han menar. Det verkar omöjligt. I alla händelser skulle en världsbild som innebär både att det finns en Gud och att det inte finns en Gud knappast uppfylla kravet på motsägelsefrihet. Snarare menar väl Bråkenhielm (a) att han själv föredrar en världsbild som passar hans egna djupaste upplevelser, eller kanske (b) att var och en bör välja den världsbild som bäst överensstämmer med vederbörandes egna upplevelser. Men däremot kan han knappast mena (c) att en riktig och sann världsbild måste överensstämma med Carl Reinhold Bråkenhielms upplevelser, eller med allas upplevelser.

#### *4. Tro och vetande*

Jag gissar att Bråkenhielm själv har kristna upplevelser som gör att han för egen del föredrar en kristen världsbild. Låt oss nu anta att en kristen världsbild är motsägelsefri, sammanhängande, och förenlig med väletablerade vetenskapliga resultat och teorier (även om detta nog kan diskuteras). Hur uppfyller en kristen världsbild kraven på enkelhet och räckvidd? Bråkenhielm skulle nog säga att den är mindre enkel, men har större räckvidd, än en icke-religiös världsbild som inte antar existensen av en Gud (jfr s 94).

Att en religiös världsbild är mindre enkel än en icke-religiös verkar rimligt. Men på vilket sätt har en religiös världsbild större räckvidd? Man kunde säga att en teori har större räckvidd än en annan om den förklarar fler fakta. Det säger också Bråkenhielm (s 94). Men en religiös världsbild förklarar knappast fler fakta än en icke-religiös. Emellertid tolkar Bråkenhielm detta krav på ett lite speciellt sätt. Han skriver:

Om en världsbild lämnar stora fält i vår erfarenhetsvärld utan någon tolkning, är detta en grund för kritik. Räckviddsprincipen kan för världsbildens vidkommande formuleras som ett krav på maximal räckvidd och har inte sällan varit en grund för kritik mot t ex den antropocentriska världsbilden (s 94).

Enkelt uttryckt är väl tanken denna. Somliga människor har religiösa erfarenheter, t ex av att Gud existerar, och en världsbild som inte innehåller antagandet att Gud existerar uppfyller därför inte



räckviddsprincipen. Men detta är uppenbarligen ett flagrant missbruk av den vetenskapsteoretiska räckviddsprincip som Bråkenhielm åberopar. Den vetenskapsteoretiska principen säger, grovt uttryckt, att man bör förklara så många fakta som möjligt, men som fakta räknas då inte religiösa uppfattningar. Till fakta hör naturligtvis det att somliga människor *har* religiösa upplevelser, men däremot kan man inte räkna *innehållet* i sådana upplevelser (alltså det som upplevs) som fakta. Innehållet i religiösa upplevelser – eller för den delen i vilka upplevelser som helst – skulle räknas som fakta endast om de i sin tur förklarar några mer neutrala och okontroversiella fakta bättre än alternativa hypoteser. Och såvitt jag vet har ingen, åtminstone ingen nutida tänkare, hävdad att religiösa antaganden uppfyller detta krav.

Med andra ord, det kan knappast råda någon tvekan om att de principer som brukar användas för att välja mellan vetenskapliga teorier skulle utesluta världsbilder av den typ som brukar förknippas med kristendomen och andra religioner. Detta är knappast någon nyhet. Det är ju bara en lite annorlunda formulering av den gamla vanliga uppfattningen om att det råder någon sorts konflikt mellan tro och vetande, mellan religion och vetenskap. Detta behöver ju inte uppfattas så att en person inte samtidigt skulle kunna tro på såväl vetenskapliga teorier som på någon religiös världsbild. Sådant förekommer säkerligen ofta. Men religiösa världsbilder tävlar så att säga i en helt annan division än vetenskapliga teorier. Det är fel att som Bråkenhielm antyda att religiösa världsbilder kan uppfylla de krav man brukar ställa på vetenskapen.

Somliga religiösa personer anser att deras tro strider mot vissa vetenskapliga teorier, t ex Darwins utvecklingslära, och att de därför måste förkasta dessa teorier. Men vanligare är väl – åtminstone i dagens Sverige – att religiösa personer är beredda att godta vilka teorier som helst, som uppfyller gängse vetenskapliga krav, utan att detta alls berör, eller antas strida mot, deras religiösa övertygelser. De lyckas med andra ord, åtminstone i egna ögon, förverkliga en sorts fredlig samexistens mellan olika teorier. Deras vetenskapliga och religiösa uppfattningar har så att säga olika funktion, och de ser inget fel i att hålla sig med bägge. Framför allt har de inga illusioner – som Bråkenhielm tycks ha – om att deras religiösa uppfattningar uppfyller de krav som brukar ställas på vetenskapliga (eller filoso-

fiska) teorier. För egen del tycker jag nog att denna inställning är betydligt rimligare än Bråkenhielms.

### 5. Val av kriterier

Om det jag hittills har sagt är riktigt, så kan religiösa världsbilder inte hävda sig mot mer "vetenskapliga" världsbilder, när de jämförs med hjälp av de kriterier som brukar användas för att jämföra vetenskapliga teorier. Detta är väl egentligen självklart och borde inte förvåna någon. Men man kunde kanske fråga sig varför just dessa kriterier skall användas när man vill jämföra världsbilder (som ju inte är vetenskapliga teorier). Hur vet vi vilka kriterier som är relevanta vid bedömningen av världsbilder? Kanske är detta överhuvud taget inte något som man kan *veta*. Kanske är det snarare en fråga om att *välja* en uppsättning kriterier, med vars hjälp man sedan kan bedöma olika världsbilder?

Lika väl som vetenskapliga (icke-religiösa) världsbilder favoriseras av de vanliga kriterierna för teorival i vetenskap, så kan förmodligen religiösa världsbilder favoriseras av den varsamhetsprincip och den speciella form av räckviddsprincip som Bråkenhielm vill använda sig av. Lite förenklat kan vi säga att valet av världs-åskådning till sist beror på om man lägger störst vikt vid enkelhetsprincipen eller varsamhetsprincipen. Så framställs faktiskt också saken av Bråkenhielm själv på sista sidan i boken (s 96). Han tycks mena att man inte kan ge något rationellt svar på frågan om vilken av dessa principer som skall ges företräde, men för sin egen del menar han ändå att "trohet mot erfarenheten är viktigare än enkelheten" (s 96). En religiös person kan på denna grund sedan välja en religiös världsbild. Ett sådant val strider visserligen mot enkelhetsprincipen – vilket även Bråkenhielm tycks anse – men det stöds av varsamhetsprincipen, som är "viktigare".

(Låt oss inom parentes notera det persuasiva språkbruket. Bråkenhielms "varsamhetsprincip" har ett namn med en positiv klang. Varsamhet är naturligtvis något bra. Men om principen innebär att man skall inkorporera andra människors djupt kända åsikter i sin egen världsbild, så kunde den ju också kallas "godtrogenhetsprincipen". Om den enbart innebär att man inte skall överge sina egna djupt kända åsikter, så kan den kallas "fördomsprincipen". Sådana namn skulle göra principen mindre lockande.)

När det gäller frågan om vilka kriterier som skall tillämpas, så tycks alltså Bråkenhielm till sist anse att man endast kan "ta ett språng, göra en personlig bekännelse" (s 96). Ett rationellt eller objektivt svar kan inte ges. "Här har vi nalkats en slags slutpunkt för den rationella analysen" (s 96). Om han har rätt i detta kan tydligen inte heller valet av världsbild vara något annat än en rent privat fråga, en smaksak. I så fall är det väl inte heller något man bör syssla med inom en akademisk disciplin som livsåskådningsforskning.

Men såvitt jag förstår har Bråkenhielm fel. Man kan föra en rationell diskussion om vilka kriterier som skall tillämpas, och hur de skall tillämpas. Det gör man ju faktiskt sedan länge inom filosofi och vetenskapsteori. En rationell diskussion kan exempelvis utgå från en viss målsättning, för att sedan motivera olika kriterier för teorival med hänvisning till denna målsättning. En rimlig och intressant målsättning i sammanhanget är t ex att ta reda på "vilken världsbild som är den riktiga och sanna världsbilden" (s 87), för att nu använda Bråkenhielms egen formulering. Givet denna målsättning förefaller det rationellt att utgå från de gängse kriterierna för teorival i vetenskap, ty dessa kriterier kan antas ha utvecklats för just detta syfte. Man kan till och med hävda att dessa kriterier utgör en precisering av vad vi menar med "rationalitet" i fråga om teorival. Dessa kriterier innefattar bl a enkelhetsprincipen – inte därför att sanningen med nödvändighet är enkel, utan därför att vi på sikt lättare kan närma oss sanningen om vi i allmänhet föredrar enkla teorier (eller möjligen därför att enkelhet har ett självständigt värde som ibland till och med kan väga tyngre än sanning i varje liten detalj). Däremot förefaller Bråkenhielms varsamhets- och räckviddsprinciper fullständigt omotiverade om målet är att hitta sanningen. Detta tycks han också själv inse när han antyder att dessa principer kan bereda plats för "verklighetsfrämmande ideologier" (s 96).

## 6. *Naturalism*

Det finns emellertid en komplikation här. Man kunde hävda att en "fullständig" världsbild bl a kan innehålla kriterier eller principer för teorival. Bråkenhielm är också inne på detta (s 87). Vilka kriterier skall man i så fall tillämpa i valet mellan världsbilder som själva innehåller olika kriterier? Ligger det inte något orättfärdigt i att bedöma en världsbild med kriterier som tillhör en annan världsbild?

En kristen världsbild kan exempelvis innehålla teorin att Gud uppenbarar sig för vissa människor. Om man tror på detta, så tror man väl också att dessa människors gudsupplevelser är viktiga källor till kunskap – t ex så att de utgör evidens för Guds existens – och för att garantera att de inte avfärdas som ”verklighetsfrämmande ideologi” kan man därför finna det motiverat att acceptera varsamhetsprincipen. Är det då rimligt att en sådan världsbild bedöms med vanliga vetenskapliga kriterier för teorival, där varsamhetsprincipen inte ingår?

Ja, det är enligt min mening helt rimligt. Överhuvud taget måste varje problem lösas inom ramen för vad man i andra avseenden tror. Det innebär bl a att även valet mellan världsbilder – i den mån sådana val någonsin är aktuella – måste ske inom ramen för den, eventuellt fragmentariska, världsbild och de andra teorier som man är anhängare av. Inom den analytiska filosofin har denna tes särskilt drivits av Otto Neurath och W V Quine. I sammanhanget brukar man ofta hänvisa till Neuraths liknelse om att vi i kunskapsteoretiskt avseende är som sjömän i en båt, som måste repareras mitt ute på havet: vi kan inte gå in i en torrdocka och bygga om hela båten på en gång, utan vi måste begränsa oss till mindre reparationer i flera omgångar, och ingen enskild reparation får bli mer genomgripande än att vi hela tiden kan hålla oss flytande. Neuraths båt motsvarar alltså totaliteten av det vi tror vid en viss tidpunkt. Denna ”totala teori” kan inte som helhet förkastas. (Ett skäl till detta är att den inte ens kan artikuleras i sin helhet; ingen av oss kan ge en uttömmande beskrivning av vad vi tror vid en viss tidpunkt.) Men den kan naturligtvis ha mer eller mindre allvarliga brister. Dessa brister måste identifieras utifrån teorin själv, inklusive dess principer om teorival. Teorin kan exempelvis ur sin egen synpunkt framstå som ofullständig eller osammanhängande eller onödigt tillkrånglad. Sådana brister bör då åtgärdas, så att den totala teorin gradvis förändras till det bättre – och vad som skall räknas som ”förbättringar” måste avgöras på grundval av teorin själv. Denna sk naturalistiska uppfattning innebär med andra ord att det inte finns någon fristående och neutral ”första filosofi”, någon sorts högsta instans utifrån vilken vi kan bedöma olika teorier om verkligheten. Filosofiska teorier befinner sig så att säga på samma plan som alla andra teorier. Våra bedömningar måste ske utifrån de teorier som vi godtar, och vi kan därför bara ifrågasätta vissa mindre delar av dessa teorier åt gången.

Om någon nu skulle föreslå oss att acceptera en alternativ världsbild, en som innehåller varsamhetsprincipen, så har vi ingen annan möjlighet än att bedöma den med utgångspunkt från vår egen totala teori, inklusive våra egna principer för teorival. I själva verket måste vi utgå från vår egen teori redan när det gäller att förstå vad den alternativa världsbilden innebär (dvs hur de språkliga satser som används för att uttrycka den skall översättas till vårt eget språk). Detta har jag t ex gjort när jag har försökt göra klart för mig vad Bråkenhielms varsamhetsprincip går ut på. Bedömningen av den alternativa världsbilden måste sedan, eller i samband med att tolkningen pågår, ske med avseende på enskilda delar i tur och ordning. För varje del får vi fråga oss om den kan inkorporeras i vår egen teori, utan att teorin försämras (när nödvändiga modifieringar gjorts). Vi kan t ex fråga oss om varsamhetsprincipen och antagandet om Guds existens kan infogas i vår totala teori, utan att priset blir för högt.

En undersökning av detta slag kan naturligtvis komma att mynna ut i att vi förkastar den alternativa världsbilden, med motiveringen att den inte kan inkorporeras i vår egen utan att helheten försämras. Men det bör betonas att detta på intet sätt är en nödvändig konsekvens. Det är t ex inte självklart att varsamhetsprincipen måste förkastas. Kanske skulle det innebära en förbättring av vår totala teori om varsamhetsprincipen inkorporerades i den. För att avgöra hur det förhåller sig med detta krävs argument.

För egen del tror jag som sagt att tillräckliga argument talar för att varsamhetsprincipen bör förkastas. Ett av argumenten antyds ju redan av Bråkenhielm: varsamhetsprincipen kan legitimera "verklighetsfrämmande ideologier". Ett annat argument tar fasta på att varsamhetsprincipen inte brukar tillämpas i vetenskapliga sammanhang, där det ju också är fråga om att avgöra vilken teori som är "den riktiga och sanna". Att tillämpa olika principer för vetenskapliga teorier och för världsbilder verkar omotiverat och därmed onödigt komplicerat. Ett tredje argument pekar på svårigheten att dra en gräns mellan de upplevelser som varsamhetsprincipen skall respektera och de som den inte skall respektera. En sådan gräns behövs om principen inte skall bli orimligt tolerant och leda till motsägelser.

## 7. *Relativism?*

Enkelt uttryckt kan alltså frågan ställas så här: vilken världsbild passar bäst ihop med vad vi för övrigt tror. Men nu kunde kanske någon invända att det vi för övrigt tror kan variera väldigt mycket från en individ (eller grupp) till en annan. Är det då inte troligt att man kan komma till helt olika svar på frågan? En anhängare av en kristen världsbild (som innehåller varsamhetsprincipen) skulle kanske kunna komma fram till att det skulle innebära en försämring att t ex förkasta varsamhetsprincipen, beroende på att vederbörandes totala teori även i många andra avseenden skiljer sig starkt från vår. Innebär detta att vi måste godta någon form av relativism?

Nej, det följer lyckligtvis inte. För det första tror jag inte att folks totala teorier skiljer sig så dramatiskt att en relativism av detta slag någonsin blir aktuell – åtminstone inte så länge vi håller oss till personer som är någorlunda samtida och någorlunda lika i kulturellt avseende. Snarare är det nog så att vi inte vill eller orkar driva diskussionen så långt att vi blir överens om vilken världsbild som är bäst. Men låt oss ändå anta att vi stöter på en verklig avvikare, en medlem av en helt annan kultur än vår, som hävdar att hans världsbild passar bättre ihop med resten av hans totala teori än vad vår världsbild skulle göra. Låt oss också anta att han har rätt i detta – annars blir ju tankeexperimentet ointressant. Vi måste då ta ställning till resten av hans totala teori. Om den är acceptabel ur vår synpunkt kan vi inkorporera den med vår, och om detta inte rubbar vår världsbild kan vi hävda att denna är överlägsen, eftersom den passar ihop med en mer omfattande teori. (Om vår världsbild inte längre passar ihop med resten av teorin bör vi förstås förkasta den.) Om däremot den främmande teorin är oacceptabel ur vår synpunkt, så måste vi hävda att den är felaktig, och då kan vi sedan förkasta den främmande världsbilden av detta skäl. En tredje möjlighet är att vi (ännu) inte kan avgöra vilken av de två världsbilderna som passar bäst ihop med vad vi för övrigt tror. I så fall är det mest rationella att hålla frågan öppen, och alltså inte tro på någon av de bägge världsbilderna. I inget av dessa fall behöver vi dra någon relativistisk slutsats.

## 8. *Sammanfattning.*

Etablerade och icke-fragmentariska världsbilder och livsåskådningar

är av intresse som exempel på vad folk faktiskt tror på. Men om man vill veta vad som är sant på detta område bör man snarare intressera sig för mindre enheter, och diskutera en sak i sänder. Detta låter sig göra. Vi behöver inte hamna i någon relativism på detta område. Vi behöver inte heller tro att världs- och livsåskådningsfrågor måste lösas genom att man helt enkelt bestämmer sig, "tar ett språng, gör en personlig bekännelse", utan att ha någon rationell grund för sitt beslut. Vi *kan* naturligtvis lösa frågor på detta sätt, men för den som är ute efter sanningen förefaller detta tillvägagångssätt missriktat och helt irrationellt. I valet mellan två världsbilder – om vi nu någonsin står inför ett sådant val – är det rationellt att välja den vars ingredienser lättast kan infogas i vår totala teori om verkligheten. Om vi finner att ingen av dem kan inkorporeras i vår totala teori, utan att det innebär en försämring av helheten, är det mest rationella att hålla frågan öppen, och alltså inte tro på någon av dem.

# CLAUDIO MARCELLO TAMBURRINI

## *Ett försvar för spöstraff*

---

Vad, om något, kan rättfärdiga att myndigheterna straffar brottslingar? Olika svar ges på denna fråga, beroende på vilken straffrättslig ansats man bekänner sig till.

Avskräckningsteorin är ett av dessa svar. Kortfattat går teorin ut på att straff är moraliskt rättfärdigt i och med att det avskräcker både aktuella och presumptiva brottslingar från att begå brott.

I denna artikel skall jag diskutera en traditionell moralisk invändning mot avskräckning: att ett avskräckningsgrundat straffsystem kräver alltför stränga straff. Jag kommer att argumentera för att invändningen inte är ödesdiger för avskräckningsidén, och för att en typ av lindriga kroppsbesträffningar – spöstraff – faktiskt är att föredra framför nuvarande fängelsestraff. Men först behöver både avskräckningsidén och invändningen preciseras. Låt mig då börja med det.

### *1. Avskräckningsidén*

Avskräckningstanken har traditionellt blivit formulerad i ekonomiska termer. Brott påstås medföra vissa kostnader. Det behövs pengar (eller ett tidigare brott) för att anskaffa de verktyg som är nödvändiga för att bryta mot lagen (ett vapen, antilarmsystem etc); det krävs tid för att begå brottet (åka till brottsplatsen, kanske till och med kartlägga offrets rörelser eller brottsplatsens rutiner under en lång tids övervakning, etc). Alla dessa är *naturliga* kostnader av brottslig verksamhet. När de förväntade vinsterna av brott överstiger dess naturliga kostnader, resonerar avskräckningsanhängarna, så kommer agenten förmodligen att känna sig motiverad att begå brott. Men om det nu till brottets naturliga läggs på extra kostnader, så skulle man enligt avskräckningsanhängarna kunna minska den förväntade vinsten



av brott, och därmed agentens motivation att begå det. När man genom lagstiftningen straffbelägger en viss handlingstyp blir den förväntade vinsten av den, i varje fall om straffhotet aktualiseras, mindre än vad den skulle varit, om handlingen hade lämnats ostraffad. På så sätt förvandlas brott enligt avskräckningsanhängarna till en *irrationell* verksamhet, eftersom dess kostnader (i fall gärningsmannen blir upptäckt och dömd) överstiger den vinst som gärningsmannen förväntar sig att få av det. Jeremy Bentham har formulerat avskräckningsidén på följande vis:

Brottets vinst är den kraft som tvingar en människa till brottslighet: straffets lidande är den kraft som används för att få henne att avstå från den. Om den förra kraften är större kommer brottet att äga rum: om den senare är det, så kommer brottet inte att begås.<sup>1</sup>

Genom straffhotet görs då brottet irrationellt; och genom att aktualisera straffet görs straffhotet trovärdigt. Det som enligt avskräckningsanhängarna rättfärdigar straff är då det faktum att det bidrar till att minska antalet brott i samhället. Detta sägs vara resultatet av att presumptiva brottslingar avskräcks från brott genom straffhotet.

## 2. *Invändningen*

Motståndare till avskräckning har påpekat att, om brottsreduktion genom avskräckning är det som rättfärdigar brottspåföljder, så kräver en konsekvent tillämpad avskräckningspolitik att man inför bestraffningar som klart framstår som alltför stränga i förhållande till brottets svårhetsgrad. Proportionalitet mellan brottets svårhetsgrad och påföljdens stränghet är ett grundläggande inslag i moderna straffsystem. Den motiveras både med hänvisning till rättviseskäl och av humanitära hänsyn. En väletablerad juridisk tradition anser det vara både orättvist och inhumant att straffa en lagöverträdare hårdare än vad hon gjort sig förtjänt av genom sin illgärning. Och hur man nu än avgör vad välförtjänt och humant straff är, så har denna tanke-

<sup>1</sup> Citerat i Cook, Philip, J: "The Economics of Criminal Sanctions", in *Sanctions and Rewards in the Legal System: a Multidisciplinary Approach*; M L Friedland (utg); Univ of Toronto Press, 1989, s 58-70 (Benthams citat på s 51-2, min översättning).

riktning resulterat i att de flesta moderna straffsystem bannlyser alltför stränga straff.

Den aktuella kritiken mot avskräckning säger då att, för att åstadkomma effektiv brottsreduktion, så verkar det vara nödvändigt att åsidosätta ovannämnda restriktioner i påföljdsvalet. Framför allt i tider av ständigt ökande brottslighet, har dessa kritiker hävdad, bör avskräckningsanhängarna gå med på att införa alltför stränga straff. I varje fall om de inte är villiga att allvarligt äventyra straffsystemets avskräckande verkan. K G Armstrong har formulerat denna invändning i följande ordalag:

Om vi bestämmer straffet på basis av avskräckningsprincipen, så varför stanna vid ett minimum, varför inte vara på den säkra sidan och straffa honom (lagöverträdaren) på ett något spektakulärt sätt – skulle inte det förmodligen avskräcka andra mera? Piska honom till döds, offentligt naturligtvis, för en parkeringsförseelse, sådant skulle säkerligen avskräcka mig från att parkera på den plats reserverad för Vice-Kanslern!<sup>2</sup>

Invändningen är relevant för frågan om avskräckningsidéns rimlighet eftersom ett straffsystem som alltför mycket avvek från vad allmänt anses vara rättvisa och humana påföljder förmodligen skulle förlora acceptans hos allmänheten. Och låg acceptansnivå för lagssystemet kan rimligen förväntas minska lagundersåtarnas benägenhet att lyda lagen i situationer där – som fallet är i stora urbana centra – lagens synbara arm inte kan garantera heltäckande övervakning.

Detta resonemang leder till ett annat problem som ofta citeras i samband med avskräckningsteorin, och som stödjer hållbarheten hos den ovannämnda invändningen. Straffets avskräckande verkan beror inte enbart på dess stränghet, utan i hög grad också på dess sannolikhet. Det är allmänt vedertaget inom kriminologisk forskning att väldigt lindriga straff inte kan garantera en någorlunda acceptabel nivå av laglydnad; men ett sådant mål skulle inte heller kunna uppnås med aldrig så stränga men osannolika påföljder. Och faktum är att i stora, komplexa samhällen är upptäcktsrisken ganska låg; resursknapphet omöjliggör en effektiv poliskontroll och, även i de fall där

2 I "The Retributivist Hits Back"; i Acton, HB *The Philosophy of Punishment*; MacMillan, St Martin's Press. 1969, s 152 (min översättning).

brott uppklaras, saknas det ofta resurser för att ställa de ansvariga inför rätta. Till detta bör det också tilläggas att i många länder (t ex USA) är trångboddheten på anstalterna så stor att brottslingar ofta får villkorliga domar eller sätts på fri fot långt innan de hunnit sitta av den ådömda strafftiden, för att bereda plats för nya interner. Alla dessa faktorer sammantagna gör då att den objektiva sannolikheten för att faktiskt bli utsatt för straff är ganska låg.

Ett annat problem som brukar förknippas med avskräckningsidén är vad som i den tekniska litteraturen kallas för "tendensen att diskontera värdet av framtida händelser". Kort sagt går detta ut på att omedelbara, i tiden närliggande faktorer har en mycket större motivationskraft på vårt beteende än mer avlägsna, dock allvarliga, konsekvenser av vårt handlande.

Inom ramen för straffsystemet, och trots att ovannämnda teori är empiriskt svårprövad, ger detta upphov till två olika svårigheter.

Den första svårigheten drabbar i synnerhet frihetsberövande påföljder: om det nu är sant att vi faktiskt har en tendens att diskontera värdet av framtida händelser, så borde fängelsestraffets avskräckande värde avta med tiden. Ta t ex ett tioårigt fängelsestraff. Den motivationskraft som, låt oss säga, de första tre åren kan förväntas ha på vårt beteende måste vara betydligt starkare än den som följer av de tre sista fängelseåren.<sup>3</sup> Skälet därtill är helt enkelt att våra föreställningar om det liv som vi kommer att leva då, och därmed om hur straffet kommer att påverka oss, är mycket mer vaga och diffusa än våra föreställningar om de närmast följande tre åren.

Den andra svårigheten drabbar brottspåföljder överhuvudtaget: frestelsen av att på brottslig väg skaffa sig en omedelbar välfärdsökning väger många gånger tyngre än den diffusa och avlägsna välfärdslust som straffet för med sig. Det följer ur detta att vad som krävs för att övervinna denna svårighet är påföljder som gör ett så starkt intryck på potentiella brottslingar att de, trots påföljdens medelbara karaktär, ändå känner sig starkt motiverade till att avstå från brott.

Tendensen att diskontera värdet av framtida händelser kan rimligen förväntas minska straffhotets avskräckande verkan på vårt beteende.

3 Argumentet har framförts av Philip J Cook i hans redan citerade uppsats, s 64.

Visserligen kunde man försöka motarbeta dessa problem genom att inrätta fler och effektivare säkerhetsanordningar (larmsystem, utökad vaktpersonal, etc), så att åtminstone upptäcktsrisken blev något större. Eller också kunde myndigheterna sprida felaktig information om den objektiva risken att bli upptäckt och dömd för brott. Ja, de kunde t o m iscensätta fingerade rättegångar, där skådespelarna fick långa fängelsestraff eller blev skenavrättade. Så länge inga läckor inifrån själva systemet förekom, så kunde en sådan strategi kanske hålla nere brottsstatistiken genom att öka, eller åtminstone hålla på en acceptabel nivå, straffets subjektiva sannolikhet.

Problemet med denna sistnämnda strategi är emellertid att läckor förmodligen inte går att undvika. Och vetskapen om att straffsystemet var grundat på verklighetsförvrängning och lögn skulle troligen medföra stora prestigeförluster för lagsystemet. Medborgarnas benägenhet att respektera lagar och förordningar kunde då hamna under den nivå som *samhället kan tåla*.

Beträffande den första strategin (att utöka och effektivisera säkerhetsanordningarna), så skulle detta förmodligen kunna hålla brottslingarna i schack, men till priset av att vi övergick till ett i det närmaste polisiärt samhälle.

Det verkar då som om den enda utvägen för avskräckningsanhängarna att tackla den ökade brottsfrekvensen skulle vara att införa strängare straff. Låg faktisk sannolikhet för straff samt straffhotets begränsade intryck på kriminella agenter driver upp straffsatserna.

Invändningen mot avskräckningsidén verkar då vid en första analys vara både hållbar och relevant och förtjänar därför att tas på allvar. Men hur pass ödesdiger är invändningen för avskräckningsidealet egentligen? Inte mycket. I resten av denna artikel kommer jag att påstå att invändningen blir rimlig endast om den tolkas som om den refererade till vissa lindriga former av kroppslig bestraffning som vi idag ser som förlegade. Men då klarar sig avskräckningsprincipen inte bara helskinnad ur kritiken, utan blir också betydligt förstärkt. I synnerhet kommer jag att försöka visa att en variant av kroppsbestraffning – spöstraff – är effektivare som avskräckningsmedel och mer humant än nuvarande frihetsberövande påföljder. Men, innan jag går över till att presentera argumenten för min tes, låt mig först precisera innebörden i uttrycket "för stränga straff".

Kräver avskräckning egentligen för stränga straff? Det beror naturligtvis på vad man menar med detta uttryck. Ett straff kan vara för strängt på så vitt jag kan se två olika sätt: (i) i förhållande till dess *varaktighet*: och (ii) i förhållande till dess *intensitet*.

Exempel på för stränga straff enligt (i) är för långa frihetsberövande påföljder (dvs sådana som överskrider vad brottets svårighetsgrad kräver). Exempel på för stränga straff enligt (ii) är påföljder vars intensitet i termer av fysiskt lidande gör dem till grymma och inhumana bestraffningar (tänk t ex på sträckbänken och amputering av lemmar).

Nu är vi i stånd att granska den aktuella invändningen mot avskräckningstanken. Kräver avskräckning verkligen för stränga straff? För att svara på det, behöver vi fråga oss om det vore rationellt att i ett avskräckningsgrundat straffsystem införa a) för långa fängelsestraff och b) grymma och inhumana påföljder.

### *3. Kräver avskräckning för långa fängelsestraff?*

Många avskräckningsanhängare skulle förmodligen se denna invändning som missriktad. De skulle säkert hänvisa till det faktum att det var just avskräckningsanhängare (i synnerhet Bentham men också den italienske kriminologen Cesare Beccaria) som först pläderade för införandet av proportionalitetsprincipen i straffbestämningen. Detta såg de som en motvikt mot de alltför hårda och godtyckliga påföljder som var brukliga på deras tid. Och deras argument för proportionalitet var avskräckningsgrundad: tar vi inte hänsyn till kravet på proportionalitet mellan brottets svårighetsgrad och straffets stränghet, så skapar vi en missriktad avskräckningsstruktur. Idén är den följande: för att minimera brottsskadorna är det rationellt att införa en straffpolitik som inte bara framgångsrikt avskräcker presumptiva brottslingar, utan som också leder dem som inte blir avskräckta till att välja mindre skadliga brott. Ett straffsystem som tillmäter långa odifferentierade frihetsberövande påföljder för brott av varierande skadlighetsgrad skulle förmodligen misslyckas med detta senare mål. Riskerade man lika långa, eller i varje fall väldigt långa fängelsestraff både för t ex väpnat rån och mord, så skulle man inte kunna sig tillräckligt motiverad att välja det mindre allvarliga brottet (mord kunde t ex vara ett bra sätt att skydda sig från rånoffrets vittnesmål).

Andrew von Hirsch har avfärdat detta avskräckningsgrundade

försvar för proportionalitetsprincipen genom att ifrågasätta en av argumentets förutsättningar: att det faktiskt finns (och att det går att reliabelt mäta) en relevant påverkan ("tvärelasticitet") mellan straffvärden och brottsfrekvensen för olika brottstyper. Här är hans argument:

Benthams och Beccarias argument för proportionalitet förutsätter t ex en hög grad av (för att använda ekonomernas språkbruk) tvärelasticitet mellan frekvenserna för olika brott. Brotsstatistiken för en typ av brott påstås vara påverkad inte bara av dess straff men av straffet för alternativa brott. Lagöverträdare skulle därför övergå till allvarliga brott om inte straffen vore rangordnade enligt brottets allvarsgrad. Men så vitt vi vet kan tvärelasticiteten vara svag eller obefintlig. Vi kan helt enkelt inte avgöra med någon nivå av pålitlighet hur inbrottsfrekvensen påverkas av öknings eller sänkningar i motsvarande straff för andra allvarligare brott som rån.<sup>4</sup>

Har von Hirsch rätt i det han säger? Inte helt och hållet. För det första måste man konstatera att tvärelasticitet faktiskt förekommer mellan vissa brottstyper som är artliknande<sup>5</sup>. För det andra räcker det inte med att konstatera att artbaserad tvärelasticitet många gånger inte föreligger för att kunna avfärda ett avskräckningsgrundat försvar av proportionalitetsprincipen: tvärelasticitet kan också förekomma mellan artskilda brottsformer. Dels kan tvärelasticitet förekomma mellan artskilda men *vanligen relaterade* brottstyper. Tänk t ex på väpnat rån och mord. Om straffet för mord inte görs avsevärt högre än straffet för väpnat rån, eller om båda sätts för högt, skulle många rånare känna sig motiverade att döda rånoffret för att bli av med ett besvärligt vittne. Dels kan tvärelasticitet också förekomma mellan artskilda och *vanligen icke-relaterade* brottstyper. Föreställ er ett avskräckningssystem vars huvuduppgift är att minimera välfärdsförluster i samhället. Och föreställ er vidare att i det här samhället vållar felparkeringarna, som är mycket utbredda, avsevärda sociala

4 *Past or Future Crimes – Deservedness and Dangerousness in the Sentencing of Criminals*: Manchester University Press, 1985, s 32, min översättning.

5 (dvs, brott som vållar samma typ av skada). Exempel på artliknande brott är olika sorters egendomsbrott, sexuellt ofredande och våldtäkt, etc.

kostnader. I enlighet med straffsystemets huvuduppgift, borde man då straffa felparkerare väldigt hårt (med t ex 10 års fängelse). Å andra sidan kunde det också i detta samhälle hända att mord väldigt sällan förekom. Avskräckningsprincipen skulle då rekommendera att döma mördare till ett fängelsestraff av högst 20 år.

Fastän jag är beredd att medge att det i normala fall inte förekommer någon artgrundad tvärelasticitet mellan felparkering och mord (och att de inte heller brukar inträffa tillsammans), tycker jag ändå inte att i mitt exempel mordstatistiken skulle förbli helt opåverkad av att införa så stränga straff för felparkerare. Blev jag upptäckt när jag parkerade min bil i ett förbjudet område, så skulle jag förmodligen känna mig mindre motiverad att avstå från att mörda lappvisan och smita undan från brottsplatsen. Så, även om artbaserad tvärelasticitet mellan felparkering och mord inte är tillräckligt belagd, förekommer ändå ibland det jag kallar för *situationell tvärelasticitet* mellan (vanligen) orelaterade brottstyper.

Min slutsats är då att, även i de fall där det inte verkar vara befogat att påstå att artgrundad tvärelasticitet mellan olika brottstyper faktiskt förekommer, kan situationell tvärelasticitet utgöra ett tillräckligt skäl för att införa ett proportionellt, differentierat påföljdssystem. I den bemärkelsen kunde man säga att det tycks finnas ett tak för straffvärdena som det ur avskräckningssynpunkt inte lönar sig att överskrida. Och det är mycket som talar för att nuvarande fängelsestraff, med sina inbyggda restriktioner på påföljdslängden, sammanfaller med det taket. Proportionalitetsprincipen kan med andra ord, som Bentham och Beccaria hävdade, rättfärdigas på rena avskräckningsgrunder.

#### *4. Kräver avskräckningsprincipen alltför grymma och inhumana bestraffningar?*

Om kritiken mot för långa fängelsestraff betonar betydelsen av straffets varaktighet, så har kritiken mot för grymma och inhumana bestraffningar med straffets intensitet att göra. Må påföljden vara aldrig så kort, den får inte heller vara för intensiv, annars klassas den som grym och inhuman.

Om mina argument ovan är rimliga, så verkar det som om att, åtminstone enligt en möjlig tolkning av uttrycket "alltför strängt straff", avskräckningsprincipen inte kan fällas av den aktuella invändningen.

Men kritikerna kunde säga att ett avskräckningsystem inte behöver begränsa sig till fängelsestraff. Brottslingar kunde bli dömda till olika former av kroppsstraff i stället för att bli inspärrade. Kroppsliga bestraffningar, i synnerhet om de är riktigt intensiva, skulle enligt dessa kritiker vara ytterst rationella ur avskräckningssynpunkt: de skulle skrämra ett stort antal presumptiva lagbrytare från att begå brott. Tänk t ex på ett straffsystem som bestraffade tjuvar genom att klippa av deras fingrar med en tång, utan bedövningsmedel. Eller, för den delen, på det exempel som Armstrong tar upp i ovannämnda invändningen: spöa felparkerare till döds! Skulle vi inte säga att dessa bestraffningar är alltför grymma och inhumana på det oerhörda fysiska lidande som de som blir utsatta för dem får utstå?

Invändningen pekar på det faktum att, givet att man av avskräckningsskäl bestämt sig för att införa kroppsbestraffningar, så borde man inte stanna vid lindriga sådana. Om det skulle visa sig att några få spöslag avskräcker felparkerare mer än, säg, en vecka i fängelse, då borde enligt samma resonemang spöstraff till döds avskräcka felparkerare mer effektivt än begränsat spöande. Om skälet för att föredra spöstraff framför en veckas fängelse är att det förra reducerar felparkeringar mer än det senare, har dessa kritiker sagt, så varför inte välja riktigt grova kroppsbestraffningar i stället för lindriga sådana?<sup>6</sup> De skulle säkerligen minska antalet parkeringsförseelser ännu mer! Låt oss kalla denna invändning för "våld-trappar-upp-våld" argumentet.

Jag tycker att denna invändning kan bemötas med hjälp av samma argument som anfördes i det föregående avsnittet för att bemöta invändningen om för långa fängelsestraff. Där argumenterade jag att avskräckningsprincipen tycks ha någon sorts inbyggd dynamik som

6 Godtycklighet kan inte undvikas när man försöker dra en gräns mellan grova och lindriga kroppsbestraffningar. Jag begränsar mig här till att stipulera följande vart för sig tillräckliga villkor för användningen av termen "grovt kroppsligt straff": ett kroppsligt straff är grovt när det medför antingen i) stympling och andra bestående fysiska men för brottslingen, eller ii) psykiska men av en sådan omfattning att ett någorlunda välfungerande socialt liv i framtiden omöjliggörs för den bestraffade. Lindrigt kroppsligt straff är alla kroppsliga påföljder som inte vållar någon av dessa effekter. Gränsdragningen är som sagt svår och godtycklig. Men på basis av den stipulativa definitionen ovan kan vi ändå fastslå att kortvarigt spöstraff, men inte stympling av lemmar, är ett exempel på lindrig kroppsbestraffning.



utesluter för stränga oproportionella fängelsestraff. Annars kan straffsystemet förvandlas till en missriktad avskräckningsstruktur i och med att det skulle misslyckas med att leda potentiella lagbrytare till att välja mindre skadliga brottstyper. Samma resonemang kan tillämpas på ett straffsystem som införde grova kroppsbestraffningar.

Men skulle inte en liknande invändning kunna riktas mot lindrigare former av kroppsliga straff? Avskräckningsmotståndarna kan i och för sig acceptera att grova kroppsbestraffningar inte vore rationella ur avskräckningssynpunkt. Men de skulle insistera på att lindriga kroppsliga straff är det, eftersom de förmodligen skulle vara mer effektiva som avskräckning än nuvarande frihetsberövande påföljder. Men icke desto mindre är dessa kroppsbestraffningar alltför grymma och inhumana. Eller så skulle åtminstone dessa kritiker säga.

För att bemöta den här invändningen kommer jag att koncentrera mig på en typ av lindrigt kroppsligt straff: spöstraff. Den ståndpunkt jag kommer att försvara här är att spöstraff är både effektivare som avskräckningsmedel och mer humant än nuvarande frihetsberövande påföljder. Argumenten för denna tes är de följande:

a) Genom att koncentrera tillfogandet av lidande på ett enda, intensivt ögonblick av kroppssaga gör spöstraff mer intryck på presumtiva brottslingar. Detta kan bidra till att öka straffets avskräckande verkan. Straffhotets effektivitet som avskräckning reduceras av individernas tendens till att diskontera värdet av framtida händelser. Incitamentet från en omedelbar och överskådlig välfärdsvinst är många gånger starkare än aversionen mot en lika stor (eller t o m större) medelbar och diffus välfärdsförlust. Spöstraff är då särskilt lämpligt för att rikta presumtiva brottslingars uppmärksamhet till det framtida obehagsmomentet. Till skillnad från frihetsberövande påföljder (där straffobehaget är mer abstrakt och utsträckt över tiden), så ter sig spöstraff som ytterst kännbart och påtagligt för potentiella lagbrytare. Därför kan det vara ett effektivare avskräckningsinstrument än nuvarande fängelsestraff.

b) Genom att koncentrera tillfogandet av lidande på ett enda, intensivt ögonblick (när straffet verkställs), för att sedan omedelbart sätta brottslingen på fri fot, kan spöstraff minska kostnaderna för straffsystemet. Stora summor som idag går till att bygga upp och försörja fängelseanstalter skulle kunna sparas om man införde ett spöstraffsystem.

c) Kostnaderna för samhället skulle också kunna reduceras i en annan, mer indirekt bemärkelse. Spöstraff medför faktiskt färre negativa effekter för brottslingen och hennes anhöriga än vad nuvarande frihetsberövande påföljder gör. Eftersom hon blir släppt strax efter straffverkställigheten, så drabbas inte lagbrytarens sociala liv, familjeförhållanden och arbetssituation på samma sätt som de skulle drabbas om hon hölls inspärrad under månader eller år. Brottslingen förblir med andra ord en producerande samhällsmedborgare och aktiv familjemedlem också under strafftiden.

Men spöstraff är inte bara mer humant för den bestraffade lagbrytaren. Om det skulle visa sig att det också är ett mer effektivt avskräckande, så är det samtidigt mer humant för framtida brottsoffer och för framtida brottsförövare. För de förra, eftersom deras lidande förebyggs; för de senare, eftersom man genom ökad avskräckningseffekt skulle kunna bespara dem det straff de skulle fått, om de inte hade blivit avskräckta av spöstraffet.

Min slutsats är då att i jämförelse med nuvarande frihetsberövande påföljder, så är ett spöstraffsystem: a) effektivare som avskräckning; b) billigare för samhället; och c) mer humant för aktuella brottslingar såväl som för framtida brottsoffer och lagförbrytare.

Men är det verkligen bara fördelar med spöstraff? Kan inga invändningar riktas mot det? Låt mig granska en annan möjlig invändning mot ett spöstraffsystem. Det kunde sägas att århundranden av reformistiskt arbete inom straffområdet kulminerade i att man införde ett förbud mot kroppsbestraffningar. Att återinföra spöstraff ter sig då som en återgång till en strafftradition som vi redan lämnat bakom oss. Enklare sagt, och trots mina ovan anförda argument som stödjer det motsatta, skulle dessa kritiker insistera i att det finns *något* med spöstraff som gör att det bör klassas som grym och inhuman bestraffning. Låt mig avslutningsvis granska vad det här något kan vara.

### *5. Spöstraff är grymt och inhumant*

I min diskussion av denna invändning kommer jag att använda mig av den amerikanska domaren Brennans test på grymma och inhumana bestraffningar ("cruel and inhuman punishments"). Med utgångspunkt i tanken att brottspåföljder inte får kränka brottslingarnas människovärde, har Brennan lagt fram fyra olika kriterier för att avgöra när en påföljd är grym och inhuman och därför kränker människovärdet.

Brennans fyra kriterier ser ut på följande sätt:

Om ett straff är ovanligt strängt, om det finns en stor sannolikhet för att det utmäts på ett godtyckligt sätt, om det förkastas av det samtida samhället, och om det inte finns något skäl att tro att det realiserar något straffrättsligt syfte på ett effektivare sätt än vad andra, mildare straff gör, så kränker detta straffs fortsatta tillämpning kravet på att staten inte skall tillfoga inhumana och ociviliserade bestraffningar på de som blivit dömda för brott.<sup>7</sup>

Nu måste vi ställa oss två frågor: 1) är Brennans kriterier rimliga?; och 2) om de nu är det, uppfylls de av spöstraff? Jag skall börja med att diskutera det kriterium som jag bedömer som uppenbart orimligt: att ett straff är inhumant och ociviliserat om det förkastas av samhället. Detta måste rimligen tolkas som en beskrivning av hur vanligt folk känner och tycker beträffande vissa straffpåföljder, samt hur nuvarande lagstiftning faktiskt ser ut. Nu är det onekligen sant att kroppslig bestraffning, även mildare sådan som t ex spöstraff, faktiskt är bannlyst i de flesta moderna straffbalkar, och förmodligen avfärdas den också av den stora majoriteten människor. Men jag har svårt att se hur detta faktum kan tala mot spöstraffets rimlighet utan att man förutsätter det som skulle bevisas. Den tes som nu är under diskussion är den att spöstraff borde accepteras av det stora flertalet. Att det faktiskt inte är så kan därför inte vara en konklusiv, inte ens en någorlunda relevant, invändning mot denna tes.

Frågan om huruvida spöstraff är inhumant och ociviliserat måste avgöras på basis av en uppskattning av hur mycket skada brottslingen utsätts för, snarare än på basis av hur pass accepterad denna straffpraxis är.

<sup>7</sup> Citerat i Tom Sorell, *Moral Theory and Capital Punishment*, Basil Blackwell, 1987, s 110, min översättning. Brennan utarbetade detta test för att avgöra om dödsstraffet är en grym och ovanlig, och därför kränkande, påföljd. Jag skall använda mig av hans test för att undersöka om spöstraffet är det. Vidare klargör domaren Brennan inte om han tänker sig de här fyra kriterierna som nödvändiga eller som tillräckliga för att avgöra frågan om en påföljd är grym och ovanlig. Jag skall tolka honom som om han menade att de var för sig är tillräckliga; det räcker med att ett av dem är uppfyllt för att straffet i fråga skall kunna klassas som grymt och inhumant.

Om man i stället tolkar Brennan som om han menade att samhällets avfärdande *i sig* gör att en påföljd blir inhuman och kränkande av människovärdet, så är argumentet helt enkelt inte hållbart. Att döma brottslingar till att arbeta för hjälporganisationer i tredje världen skulle förmodligen förkastas av dagens allmänna opinion; ännu mindre vore det realistiskt att förvänta sig att sådana påföljder skulle bli accepterade av rättsystemets etablissemang. Men detta i och för sig gör inte dessa bestraffningar till grymma och kränkande påföljder.

Det andra kriteriet, att ett straff är ovanligt strängt, är relevant för att avgöra om straffet i frågan är inhumant och kränker människovärdet. Men jag tycker inte att spöstraff uppfyller det kriteriet. Jag har tidigare hävdad att det är långtifrån uppenbart att spöstraff tillfogar brottslingen mer lidande än vad fängelsestraff gör. Detta är självfallet en svårprovbar empirisk fråga, och meningarna kan därför gå isär här. Men även den som vill påstå att fängelsestraff tillfogar mindre lidande hos brottslingen än spöstraff måste gå med på att detta inte räcker för att avgöra frågan till förmån för fängelsestraff. Andra skäl kunde ändå tala för att spöstraff är mer humant till slut. Spöstraff kan fortfarande vara mer humant för potentiella brottsoffer, vars välfärdsförluster kunde förebyggas genom spöstraffets ökade avskräckningseffekt. Eller för potentiella lagförbrytare som, i och med att de blir effektivt avskräckta från att begå brott, undgår det lidande som de skulle blivit utsatta för om de hade begått brott och straffats. När man skall ta ställning till om en viss påföljdsform är human eller ej, så verkar det vara orimligt att begränsa sig till att enbart uppskatta dess effekter på den dömda. Brottsoffers och potentiella brottslingars lidande är minst lika relevant här.

Dessutom verkar det också förutsättas i Brennans kriterium att ett nödvändigt samband föreligger mellan ett straffs stränghet och dess karaktär som kränkande av människovärdet. Men så är emellertid inte fallet: sambandet är endast kontingent. Tom Sorell har försökt visa detta med hjälp av följande exempel:

En bestraffning som onekligen är sträng, t ex sträckbänken, kunde tänkas tillfogas därför att det var hedrande att frivilligt eller utan motstånd bli utsatt för det. Man kunde tänka sig ett militärt straff av denna sort, relaterat till en militär hederskod. En sådan bestraffning kunde rimligen sägas vara grym och ovanlig utan att

den var kränkande. Å andra sidan kunde ett mildt straff anses vara kränkande just därför att det inte också var fysiskt smärtsamt. Att raka skallen på kvinnor som hade hjälpt och tröstat Nazisoldater i det ockuperade Frankrike var mildare än den långsamma avrättningen av andra kollaboratörer.<sup>8</sup>

Detta argument leder Sorell till slutsatsen att straffstränghet är varken ett nödvändigt eller ett tillräckligt villkor för att en påföljd skall kränka människovärdet. Jag finner hans ståndpunkt rimlig.

Brennans tredje kriterium säger att om ett straff tillämpas på ett godtyckligt sätt, så blir detta straff kränkande. Men även om det vore så att spöstraff blev summariskt och hastigt verkställt (vilket naturligtvis återstår att bevisa), så kunde denna godtycklighet åtgärdas genom strängare procedurella regler och utökad rättssäkerhet.

Slutligen, beträffande Brennans fjärde kriterium har jag tidigare hävdat att spöstraff förmodligen är effektivare som avskräckning än frihetsberövande påföljder. Jag skall därför inte upprepa min argumentation här. Men en sak kan ändå poängteras. När det gäller frågan om straffstränghet, så är det den verkliga skada som brottslingen lider, och inte den som till synes tillfogas henne, som är avgörande. Men när det gäller straffeffektiviteten som avskräckningsmedel, så är det precis tvärtom: där är det den skada hon ser ut att lida, och inte den verkliga, som är relevant. Spöstraff är effektivare som avskräckning och samtidigt mer human än fängelsestraff därför att det ger intryck av att tillfoga brottslingen en större skada än vad det faktiskt gör. Däri ligger spöstraffets stora fördel, eftersom "(d)et är den starkaste rekommendation ett straff kan få att det skulle se ut att vara strängare än vad det är; för dess praktiska styrka beror mycket mindre på vad det är än på vad det verkar vara".<sup>9</sup>

Min slutsats är att inte heller Brennans test har lyckats identifiera det som skulle kunna göra att ett spöstraff rimligen kunde klassas som grymt, inhumant straff som kränker människovärdet.

Avslutningsvis tänkte jag ta upp en annan invändning som kunde riktas mot ett spöstraffssystem, och som inte ingår som kriterium i

<sup>8</sup> Op cit, s 113, min översättning

<sup>9</sup> John Stuart Mill, *Parliamentary Debates*, citerat i Sorell, op cit, s 87, min översättning. (Mills uttalande gäller dödsstraff).

Brennans test. Att utsätta brottslingar för spöstraff kommer förmodligen att öka snarare än att hjälpa dem att komma över deras agg och antisociala tendenser. Kortare uttryckt: *våld kan bara föda våld*:

Mitt svar på denna invändning är enkelt och direkt: fängelselivet tycks mig vara allt annat än en försonande och läkande erfarenhet. Jag har svårt att se hur det att tillbringa månader eller t o m år mellan fyra väggar, utan praktiskt taget någon kontakt med yttervärlden, med ens familje-, sex- och arbetsliv kraftigt eller helt nedskurna, hur allt detta kan förväntas föda mindre våld och agg än alternativet att spöa den dömda för att sedan sätta henne på fri fot. Det är för mig uppenbart att nuvarande straffpraxis innehåller mycket mer våld än den lindriga formen av kroppslig bestraffning som jag har försvarat här.

## 6. Slutsatser

Vilka slutsatser kan man dra av den föregående diskussionen? Den slutsats som först kommer till tankarna är att, om mina argument ovan är rimliga, så borde vi byta ut nuvarande fängelsesystem mot spöstraff. Hittills har jag medvetet inte yttrat mig i den frågan. I stället har jag begränsat mig till att ge argument som stödjer att spöstraff är rimligare än frihetsberövande påföljder, utan att säga något om huruvida det bör införas eller inte.

Men nu kan man undra: om spöstraff är både effektivare och mer humant, så varför inte införa det?

Ett argument för att inte göra det vore att hänvisa till det faktum att det redan finns tillräckligt våld i världen. Om man nu dessutom öppet sanktionerade dess tillämpning, så skulle man implicit legitimera våldets existens.

Problemet med detta argument är att det likväl talar för att införa ett spöstraffssystem. Givet att det redan finns mycket våld i världen, både innanför och utanför fängelsesystemet, så är det bättre att det tillämpas under reglerade, offentliga former.

Den andra slutsats man kan dra är att, om vi nu ger vika för vår fördöm mot kroppsbestraffningar och bestämmer oss för att inte införa spöstraff, så gör vi en påtaglig välfärdslust. Det är möjligt att vi många gånger agerar på basis av våra förutfattade meningar, även när vi skulle kunna tjäna mycket mer på att slå in på nya, eller nygamla vägar. Syftet med denna artikel har varit att dra uppmärksamheten till kostnaderna av att så göra inom straffområdet.

Den tredje slutsats man kan dra är att fängelsestraff faktiskt medför stora sociala kostnader. Om det verkligen är så, att fängelsestraff är mindre effektivt och mer inhumant än påföljder som vår kultur så bestämt (om än felaktigt) tar avstånd ifrån, så borde vi få oss en funderare om fängelsestraffets fortsatta existens. Ett outtalat syfte med denna artikel har då också varit att uppmana till att hitta rimligare alternativ till frihetsberövande påföljder.

Låt detta vara min viktigaste slutsats. Jag lämnar frågan om spöstraffets införande öppen.

*Dela med dig  
av ditt  
överskott.*

 *Amnesty* / HUMANFONDEN

*Din allemansfond  
020-23 53 00*

## THALES 1993

Filosofi från

*FilosofiAntologin*

red JAN WILLNER

*Filosofin genom tiderna – 1800-talet*

red KONRAD MARC WOGAU

*Filosofin genom tiderna – 1900-talet*

red KONRAD MARC WOGAU

*Argumentationsteori, språk och vetenskapsfilosofi*

DAGFINN FØLLESDAL, LARS WALLØE, JON ELSTER

*Effektivitet i ekonomisk analys*

ROLAND GRANQVIST

*Moraliskt tänkande*

RICHARD HARE

*En Ingemar Hedenius bibliografi*

ANN-MARI HENSCHEN-DAHLQUIST

*Manipulerat liv*

TORBJÖRN TÄNNSJÖ

*Vårdetik*

TORBJÖRN TÄNNSJÖ

*Samhällsvetenskapens idé*

PETER WINCH

THALES 1993

Filosofi från



## RECENSION

Sven Danielsson: *Filosofiska Utredningar. Om modus, tempus och moral*, Thales förlag, Stockholm 1988.

---

*Filosofiska Utredningar. Om modus, tempus och moral* rymmer tre uppsatser av Sven Danielsson, vid tiden för publicering docent, nu professor, i praktisk filosofi i Uppsala.

I. Den första uppsatsen, 'Normlogik och svensk språklära', är till formatet närmast en mini-avhandling, omfattande tre delar och sammanlagt femton avsnitt. Den är en välskriven, skarpsinning och till stora delar övertygande genomgång av förslag till logisk form och logiska lagar för moraliska utsagor med hänsyn till intuitioner om mening hos moraliska utsagor formulerade på vanlig svenska. Titeln på uppsatsen är delvis rättvisande, delvis missvisande. Danielsson är väl inläst på traditionell, klassifikatorisk, grammatik, "språklära", och tillämpar denna flitigt på sina exempel-satser, framför allt i diskussioner av eventuella relationer mellan morfologi och syntax å ena sidan och språklig mening å den andra. Även om denna kunskap är respektgivande, så fyller användandet av den ofta inte något syfte för det huvudsakliga resonemanget. Det

som är relevant för Danielssons syften i uppsatsen är hans intuitiva tolkningar av ett antal exempel-satser, eftersom dessa är steg på vägen mot hans förslag till logisk analys, medan de grammatiska karakteriseringarna oftast inte spelar någon roll i detta sammanhang. Det hade varit annorlunda om han verkligen arbetat sig fram till en bestämd semantisk teori om ett relevant fragment av svenska språket, men detta mål ligger en bra bit bortom uppsatsens syften. Det är dock möjligt att Danielsson, vid sidan av den centrala diskussionen, haft ett bredare pedagogiskt syfte med uppsatsen.

Det som Danielsson i första hand pläderar för är att vi i den logiska formen för moraliska utsagor, dvs den form som normlogiken appliceras på, skall skilja mellan *alternativ* och *omständigheter*. Alternativ är det som räknas som öppet, det som t ex rekommenderas eller påbjuds, medan omständigheter är det som i sammanhanget tas som givet. Distinktionen är såväl enkel och generell som kraftfull, i den meningen att ett stort antal intuitioner om skillnader i mening mellan snarlika *bör*-utsagor på ett i mitt tycke träffande sätt kan karakteriseras med hjälp av den.

En typ av fall representeras av satserna 'Per bör vara framför Pål' och 'Pål bör vara bakom Per', vilka intuitivt har olika mening, trots att det är samma relation mellan Per och Pål som enligt satserna bör råda, eftersom 'Per är framför Pål' och 'Pål är bakom Per' är ekvivalenta. Med hjälp av Danielssons operator 'Alt', som anger att ett visst sakförhållande, med avseende på situationer av en viss typ, är ett alternativ, kan vi representera den intuitiva skillnaden genom att ange att, vad gäller den första satsen, Påls placering inte är ett alternativ, och, vad gäller den andra, att Pers placering inte är ett alternativ (precis hur det ska formuleras i detta fall anges inte av Danielsson; det förefaller som om vi behöver använda kvantifikatorer, och gör vi det, så kan vi faktiskt uttrycka skillnaden utan Alt-operatören).

En annan typ av fall gäller modifikationer. Satsen 'Du bör köra på höger sida' används normalt inte såsom implicerande 'Du bör köra'. Snarare antas körandet vara en omständighet i vilken alternativet att göra det på höger sida anbefalls. Detta representeras av Danielsson som 'O(Du kör på höger sida) och icke Alt (Du kör)'. Körandet är inte i denna läsning ett alternativ, utan en omständighet.

En tredje typ av fall gäller tidsordning. Allmänt gäller att något som kan föreskrivas vid en viss tidpunkt inte kan vara en omständighet vid denna tidpunkt. Det 'bör' som

är relevant för föreskrifter är alltså tidsrelaterat. Danielsson skiljer noga mellan att föreskriva och att klassificera en handling med hänsyn till normer. Man kan klassificera en handling (eller en underlåtenhet) i efterhand (varvid vi typiskt använder 'borde'), men föreskriva endast så länge det är en öppen fråga om handlingen ifråga kommer att utföras eller inte. Dessa distinktioner tillåter Danielsson att skilja mellan 'löften som ges bör hållas' och 'löften som har givits bör hållas'. Den förra men inte den senare kan vara giltig utan inskränkning. Problemet är att jag kan lova sådant som är omöjligt för mig att hålla, t ex lova Per att gå bort och Pål att stanna hemma. I detta fall skall jag enligt den senare normen, med hänsyn till vilken givna löften utgör *omständigheter*, både gå bort och stanna hemma. Men man kan inte föreskrivas att göra något som är logiskt omöjligt. Enligt den förra normen är emellertid såväl löftesgivandet som löfteshållandet *alternativ*. Situationer i vilka givna och oförenliga löften utgör omständigheter är alltså inte relevanta för denna norm. Vad normen säger är att man inte både skall ge och sedan bryta ett löfte. Det som följer ur den är därför att oförenliga löften inte skall *ges*. Detta är en utsökt analys. Å andra sidan, vad gäller intuitioner om vanlig svenska, så förefaller det inte särskilt övertygande att med hänsyn till normformuleringen 'löften som ges bör

hållas' påstå att jag efter att ha givit ett löfte *inte* är förpliktad att hålla det, men det är precis vad Danielssons tolkning går ut på.

Det som här refererats härrör från uppsatsens två första delar. Danielsson har där i mitt tycke övertygande visat på behovet att skilja mellan alternativ och omständigheter när det gäller föreskrivande normer, och i samband med detta också på att normer bör uppfattas som tidsrelaterade. I det normlogiska formalismen skall distinktionen mellan alternativ och omständigheter fångas med hjälp av Alt-operatorm, vilket bl a skall tillåta Danielsson, till skillnad från andra, att behålla 'bör', eller snarare dess formella motsvarighet 'O', som en satsoperator, i enlighet med traditionen.

Förslaget är knappast oproblemiskt. Frågan är nu vad en sats som 'O(du kör på höger sida)' skall anses betyda ('du' antas här riktad till alla normen att köra på höger sida gäller). Det traditionella förslaget att karakterisera dess mening genom att säga att denna sats är sann i en viss värld  $w$  om, och endast om, det gäller i samtliga norm-perfekta (deontiskt perfekta) världar (tillgängliga från  $w$ ) att alla kör på höger sida, duger inte. 'O(du kör på höger sida)' är inte en fullständig normutsaga utan kompletteras i Danielssons framställningar med andra delsatser innehållande Alt-operatorm. Endast om ingenting relevant som impliceras av 'du kör

på höger sida' utgör omständigheter, dvs inte är alternativ (såsom 'du kör', 'du är på höger sida', 'du är på en väg' och så vidare) är karakteriseringen i termer av möjliga världar korrekt för normen som helhet. Vi kan naturligtvis inte tolka satsen 'O(du kör på höger sida)' såsom implicerande att körande på höger sida utgör i sin helhet utgör alternativ, eftersom vi då skulle erhålla motsägelser när vi lägger till klausuler om att vissa förhållanden utgör omständigheter och därmed inte är alternativ (dvs, om 'du kör' relevant impliceras av 'du kör på höger sida' och enligt förslaget därför måste uttrycka något som är ett alternativ, så får vi en motsägelse genom att lägga till 'och icke Alt(du kör)'). Inte bara har vi problem med att förstå vad en sats som 'O(du kör på höger sida)' betyder. Vi kan heller inte veta vad för norm som tänks vara uttryckt i denna formalism förrän vi har erhållit fullständig information, genom Alt-klausuler, om vad som är respektive inte är alternativ. Och när är sådan information fullständig?

I uppsatsens tredje del diskuterar Danielsson några kända normlogiska paradoxer, samt föreslår lösningar grundade just på distinktionen mellan alternativ och omständigheter. Paradoxerna grundar sig på vissa principer för normlogik, i synnerhet på principen att om  $B$  är en logisk konsekvens av  $A$ , så är  $O(B)$  en logisk konsekvens av  $O(A)$ . Detta verkar för vissa

exempel fullt rimligt, men det ger oss också t ex den barmhärtige samaritens paradox:

Den rånade bör få hjälp.

Att det finns en rånad är en logisk konsekvens av att den rånade får hjälp.

Alltså: det bör finnas en rånad.

Att denna paradox kan undvikas genom Danielssons distinktion mellan alternativ och omständigheter förefaller troligt. Frågan är hur denna skall införas i de logiska principerna. Danielsson utgår från en alternativ formulering: för det första principen om substituerbarhet av logiska ekvivalenter under deontiska operatorer. Det vill säga, om A är logiskt ekvivalent med B, så är  $O(A)$  logiskt ekvivalent med  $O(B)$ . Och för det andra slutledningsregeln:  $O(A \& B)$ ; alltså  $O(A)$  (respektive  $O(B)$ ). Den första av dessa principer vill Danielsson behålla på grund av dess logiska rimlighet. Den andra däremot tänker han sig ska ersättas med:  $O(A \& B)$ ; Alt (A); Alltså  $O(A)$ . Detta blockerar paradoxen. Med hjälp av substitutionsprincipen sluter vi oss från  $O(\text{den rånade får hjälp})$  till  $O(\text{det finns en rånad och den rånade får hjälp})$ , men vi kan inte därifrån sluta oss till  $O(\text{det finns en rånad})$ , eftersom Alt (det finns en rånad) är falskt.

Danielsson fullföljer inte uppsatsen genom att ge en deontisk logik innehållande Alt-operatoren, men det förefaller mig som om detta skulle vara möjligt. För vissa

typer av slutledningar räcker det emellertid inte med att veta att något inte utgör ett alternativ; man måste också veta om det är sant eller falskt.

Det tycks emellertid också, i slutet av avsnitt 3.3, som om Danielsson med hjälp av sin föreslagna logiska begränsning menar sig undvika Ross' paradox.

Du bör posta brevet

Ur 'du postar brevet' följer 'du postar brevet eller bränner det'

Alltså: du bör posta eller bränna brevet.

Trots ett flertal genomläsningar har jag inte lyckats bli klar över hur denna paradox kan undvikas genom Danielssons modifiering av normlogiken. Danielsson påpekar själv, som ett skäl mot att införa en begränsning direkt på konsekvensprincipen, att vi knappast har något skäl att förneka att 'du postar brevet eller bränner det' är ett alternativ. Men om det är ett alternativ, så blockeras slutledningen så vitt jag kan se inte heller genom Danielssons metod. Riktigt hur Danielsson tänker här förstår jag inte.

II. Den andra uppsatsen, 'Om konsekvensetikens gränser' är av mer traditionellt diskuterande slag. Av uppsatsens fem avsnitt är de tre första ägnade åt en diskussion om de principiella skillnaderna mellan konsekvensetik och pliktetik, och de två sista åt en diskussion av olika idéer om moraliskt relevanta conse-

kvenser av handlingar.

Intuitivt tycks det föreligga en tydlig skillnad mellan konsekvensetik och pliktetik, väl illustrerad av ett exempel där en person kan uppfylla en plikt genom att hålla ett löfte samtidigt som konsekvenserna av ett eventuellt löftesbrott skulle vara bättre. Konsekvensetiken skulle då förespråka det senare alternativet och pliktetiken det förra.

Så är det emellertid inte enligt de mer abstrakta och generella begrepp som moralfilosofier använder för att karakterisera skillnaden. Exemplet ovan illustrerar då närmast en skillnad mellan en mer och en mindre stelbent moral. Vad som utgör de principiella skillnaderna mellan konsekvensetik och pliktetik är svårare att säga. En pliktmoral kan enligt Danielsson mycket väl formuleras i termer av konsekvenser ("Handla så att ...!"), samtidigt som en konsekvensmoral också fastställer plikter, dvs. hur man ur moralisk synpunkt bör handla. Utan att själv närmare argumentera för det antar Danielsson också uppfattningen att varje pliktmoral kan skrivas om till en "extensionellt ekvivalent" konsekvensmoral, och *vice versa*. Med "extensionellt ekvivalenta" moraler (uppsättningar av moraliska föreskrifter eller villkor) menas moraler som påbjuder, tillåter respektive förbjuder precis samma handlingar. Man kan naturligtvis tänka sig att det som kännetecknar en konsekvensmoral just är att vara formulerad i termer av konsekvenser av

handlingar, men dels finns en principiell svårighet att skilja handlingar från konsekvenser (ett förbud mot att mörda kan formuleras som ett förbud mot handlingar vilkas konsekvens är andra människors död (och som planerats i detta syfte och dessutom utförts på grundval av planen)), dels är det också en öppen fråga vad som ska räknas som moraliskt relevanta konsekvenser (vilket diskuteras i de avslutande avsnitten).

Det som ofta tänks skilja pliktetik från konsekvensetik, och som är skälet till att en omskrivning av moralföreskrifter från alls kan behövas, är att den senare men inte den förra utgår från *värden*. Värdeförhållanden, som t ex att en värld med mer lycka är mer värdefull än en med mindre, tänks utgöra grunden för en konsekvensmoral, vilken formuleras som föreskrifter för handling i termer av resulterande värde. Danielsson invänder mot detta. Även i en pliktmoral ligger värderingar *implicit*. I ett förbud mot stöld ligger redan en implicit negativ värdering av resultatet av stölder. Dessutom finns det ingen klar mening i talet om att ett värdeförhållande utgör *grunden* för en handlingsföreskrift annat än vad som kan anges genom tillräckliga och nödvändiga villkor, och sådana räcker inte för att karakterisera skillnaden mellan de bägge moralfilosofiska uppfattningarna. Det finns en formell skillnad i formulering, eftersom konsekvensetikerns föreskrifter relaterar handlingar till

empiriska förhållanden *via* uttalade värdeomdömen, medan pliktetikern relaterar dessa direkt, men det är en skillnad utan substans.

Emellertid finns det enligt Danielsson en viktig skillnad, en som är till konsekvensetikens nackdel och visar dess begränsning. För att karakterisera våra moraliska intuitioner krävs såväl begreppet *plikt* som begreppet *värde*, men pliktetikern kan relatera dessa på ett friare sätt än konsekvensetikern. Pliktetikern kan t ex säga att det inte är en plikt att åstadkomma de bästa konsekvenserna, t ex om det skulle kräva för mycket av personen, medan konsekvensetikern har svårt att redogöra för denna uppfattning i termer av konsekvens och värde.

Var och en av dessa uppfattningar kan tyckas ha ett visst fog för sig. Det är emellertid inte uppenbart att de går väl ihop. Om värderingar, som Danielsson säger, ligger *implicit* i en uppsättning moraliska föreskrifter, så kan vi, förmodar jag, anta existensen av en viss funktion  $F$  från moralförhållanden till värdeförhållanden, sådan att  $F(P) = V$ , där  $P$  är en pliktmoral och  $V$  ett värdesystem. Omvänt bör vi, utifrån Danielssons beskrivning av konsekvensetik, kunna anta existensen av en viss funktion  $G$  från värdeförhållanden till moralförhållanden, sådan att  $G(V) = K$ , där  $V$  är ett värdesystem och  $K$  en konsekvensmoral. Här kan man tänka sig att det finns fler än en sådan funktion. Moralfilosofen,

även Danielsson i föreliggande uppsats, diskuterar olika förslag.

En fråga är nu om det finns en viss funktion  $G_i$  sådan att, för varje pliktmoral  $P$ ,  $G_i(F(P))$  är en konsekvensmoral som är extensionellt ekvivalent med  $P$ ? Danielssons tal om att vi kan gå mellan pliktmoral och konsekvensmoral genom rena omskrivningar ger stöd för tanken. Antag därför att en sådan finns. Vad innebär då talet om att pliktetiken kan relatera plikter och värden på ett friare sätt? Om något så förefaller talet om implicita värderingar tyda på motsatsen. Jag förmodar att det Danielsson har i tankarna är våra intuitiva sätt att *resonera* om moral. Vi har vissa intuitioner om pliktförhållanden och vissa intuitioner om värdeförhållanden, ingendera i formen av en fullständig teori, och använder bägge, återigen intuitivt, när vi resonerar oss fram till pliktförhållanden i olika typer av situationer. Poängen är emellertid att under rådande antaganden så gäller att om konsekvensetikern inte kan redogöra för vårt intuitiva sätt att resonera om moral, så kan heller inte pliktetiken göra det. Om de intuitiva uppfattningarna kan systematiseras pliketiskt, så att den systematiserade intuitiva värdeteorin faller ut ur den systematiserade intuitiva pliktmoralen, så har vi också en extensionellt ekvivalent konsekvensmoral som faller ut ur samma värdeteori, och därmed så kan även konsekvensetikern redogöra för de intuitiva resonemangen.

Detta förefaller riktigt åtminstone om vi bortser från att vi formulerar (axiomatiserar) en moral på olika sätt, genom att välja vissa principer som grundläggande och andra som härledda. Om jag förstått Danielsson riktigt framhåller han inte någon skillnad mellan plikt-moral och konsekvensmoral på denna punkt, utan en skillnad ifråga om vad som måste och vad som inte måste vara plikt, givet vissa värden.

Kanske bör slutsatsen därför vara att ge upp antagandet om existensen av funktionen  $G_i$ . Vi kan tänka oss åtminstone två logiska försvagningar av relationen mellan pliktmoraler och konsekvensmoraler. Dels kan vi anta att det visserligen inte finns någon unik funktion  $G$  sådan att, för varje pliktmoral  $P$ ,  $G(F(P))$  är en med  $P$  extensionellt ekvivalent konsekvensmoral, men att det ändå, för varje pliktmoral  $P$  finns åtminstone någon funktion  $G$  sådan att  $G(F(P))$  är en konsekvensmoral som är extensionellt ekvivalent med  $P$ . Vi kan också gå ett steg längre. Kanske är det bara så att det för varje pliktmoral  $P$  finns någon funktion  $G$  och något värdesystem  $V$  sådan att  $G(V)$  är en konsekvensmoral som är extensionellt ekvivalent med  $P$ .  $V$  behöver inte vara identisk med  $F(P)$ . Skälet för att anta sådana inskränkningar skulle vara att de funktioner  $G$  som är tillgängliga enligt konsekvensetiska principer är alltför schematiska för att kunna likna pliktetikens  $F$ . Åtminstone den senare försvagningen

passar väl ihop med Danielssons uppfattning att pliktetiken kan relatera plikter och värden på ett friare sätt. Men med dessa begränsningar torde det vara svårt att visa att det till varje pliktmoral finns en extensionellt ekvivalent konsekvensmoral och omvänt. I varje fall kan vi knappast erhålla de respektive moralerna ur varandra genom någon enkel omskrivning.

Uppsatsens två avslutande avsnitt utgör en utmärkt diskussion om vad som är rimligt att konsekvens-etiskt räkna som moraliskt relevanta konsekvenser av en handling. Speciellt framhåller Danielsson att vi inte dit behöver räkna vad som faktiskt senare sker (vilket i sin extrema form ger vad Danielsson kallar *domedagsmoral*: det är vad som till sist visar sig ske som avgör en handlingens moraliska status). Som det moraliskt relevanta kan vi i stället räkna de sannolikheter (chanser, risker) som resulterar ur en handling. Detta ligger väsentligt närmare vår intuitiva moraluppfattning.

III. Den tredje uppsatsen är betitlad 'Om om' och handlar om villkorssatser. Den har jag svårt för. Den är snårig i sin uppbyggnad och företer i mitt tycke egendomligheter både vad gäller resonemang och språkliga intuitioner. I inledningen vill Danielsson göra gällande att man kan acceptera satserna

- (1) Om Oswald inte sköt Kennedy, så sköt någon annan Kennedy.

(2) Om Oswald inte sköt Kennedy, så sköt han någon annan.

samtidigt som man inte accepterar

(3) Om Oswald inte sköt Kennedy, så sköt någon annan Kennedy, medan Oswald sköt någon annan än Kennedy.

trots att (3) förefaller följa ur (1) och (2) tillsammans.

Redan detta förefaller mig besynnerligt. Emellertid händer det ofta att man överraskas av möjligheter till oförutsedda tolkningar av satser. Särskilt lingvister har utvecklat mästerskap i att finna sådana. För att ro hem en tes som Danielssons skall man då visa på godtagbara tolkningar, det som lingvister kallar "läsningar", av (1), (2) och (3), sådana att under dessa tolkningar (1) och (2) kan vara sanna och (3) samtidigt falsk.

Danielsson resonerar så här. Tolkningarna av (1) och (2) kan erhållas genom att lägga tonvikten på 'Oswald' i första fallet och på 'Kennedy' i det andra. Det som sägs i försatsen till de bägge satserna, 'Oswald sköt Kennedy', är det samma till ordalydelsen, men uttrycker inte samma *egentliga hypotes*. Termen är Danielssons. Skillnaden kan göras explicit genom parafraseringarna

(1.1) Om den som sköt Kennedy inte var Oswald, så sköt någon annan Kennedy.

(2.1) Om den som Oswald sköt inte var Kennedy, så sköt Oswald någon annan.

Enligt den semantiska teori Danielsson föredrar är (1.1) meningslös om inte exakt en person sköt Kennedy, medan (2.1) är meningslös om inte Oswald sköt exakt en person. För bägge satserna gäller då att om de inte är meningslösa, så är de sanna (vilket också kastar tvivel över tolkningarna, eftersom (1) och (2) skulle kunna vara falska).

Enligt resonemanget skall vi nu acceptera (1.1) och (2.1), och därmed såväl att Kennedy sköts av exakt en person som att Oswald sköt exakt en person. Sista steget i argumentet skall nu vara att visa att det finns en rimlig tolkning av (3) som under dessa förutsättningar är falsk. Men just detta försöker Danielsson inte visa. I stället hävdar han att konjunktionen av förleden i (1.1) och (2.1), dvs 'Om den som sköt Kennedy inte var Oswald och den som Oswald sköt inte var Kennedy', inte är en rimlig tolkning av 'om Oswald inte sköt Kennedy' såsom förled i (3). På den punkten kan jag inte annat än hålla med, men varför detta skulle visa att man kan acceptera (1) och (2) men inte (3) förstår jag helt enkelt inte. Jag är lika frågande inför flera andra passager.

I ljuset av uppsatsens senare delar tycks syftet med detta exempel vara att introducera begreppet *egentlig hypotes*. Begreppet



används för att skissera och exemplifiera en teori om tolkningar av villkorssatser i termer av tidsreferens. Teorin används i sin tur polemiskt, för att vederlägga ett argument för att 'om, så' är väsentligt tvetydigt däri att det har en "indikativ" och en "konjunktiv" läsning. Argumentet utgår från att en sats som (1) kan vara sann, samtidigt som en sats som

(4) Om Oswald inte hade skjutit Kennedy, så skulle Kennedy ha vunnit valet 1964.

också är sann. Tanken är nu att eftersom satserna innehåller samma antagande i försatsen, nämligen att Oswald inte sköt Kennedy, och ut-säger oförenliga saker i eftersatserna (att någon annan sköt Kennedy respektive att Kennedy vann – skulle vinna – valet 1964), så skulle de vara oförenliga om 'om, så' betydde detsamma i (1) och (4). Eftersom satserna är förenliga finns därför en tvetydighet: en indikativ läsning i (1) och en konjunktiv läsning i (4).

Danielssons reaktion på detta argument är att först påpeka att slutledningen är riktig endast om det finns läsningar av (1) och (4) med samma egentliga hypotes, och därefter att denna läsning fortfarande skall ge (1) och (2) olika egentliga hypoteser (jag antog att läsningar som gav (1) och (2) olika egentliga hypoteser endast skulle vara möjliga; här förefaller det som om skillnaden är obligatorisk).

Detta kallar Danielsson en svår beställning.

Man skulle vid det här laget tro att argumentationen går ut på att visa att dessa uppställda villkor inte kan uppfyllas, och att argumentet för tvetydighet därför faller. Emellertid har jag inte lyckats bli klar över om Danielsson verkligen resonerar så. Det som närmast följer i texten är en diskussion om distinktionen mellan indikativa och konjunktiva villkorssatser, vari skillnaden kan tänkas bestå och vilka grammatiska kriterier vi har för den. Detta vävs samman med diskussioner om när försatser i villkorssatser framställer förhållanden som överkliga eller endast okända, om när villkorssatser uttrycker relationer mellan förhållanden respektive relationer mellan kunskap om förhållanden, om relationen mellan 'om, så' och 'endast om', om tidsreferenser och om annat.

Om jag förstått det hela något så när riktigt, så utmynnar resonemanget i följande uppfattning. För det första, så görs om-så-satser sanna eller falska av verkliga om-så-relationer mellan förhållanden (om t ex du har bättre kort än jag, så råder en om-så-relation mellan förhållandet att jag synar och förhållandet att jag förlorar). För det andra, så kan sådana om-så-relationer vara tidsbundna; de kan inträda som fakta vid vissa tidpunkter. För det tredje, så finns det en skillnad mellan det som klassas som indika-

tiva och det som klassas som konjunktiva villkorssatser. Skillnaden är temporal. Konjunktiva, men inte indikativa, villkorssatser görs sanna eller falska av om-så-relationer som består *före* de händelser som omtalas i satsen. I denna mening är konjunktiva, men inte indikativa, villkorssatser prediktiva, dvs de uttrycker förutsägelser eller möjliga förutsägelser.

Denna teori tillämpad på (4) och (1) ger följande resultat. (4) är en konjunktiv, prediktiv villkorssats. Dess tempusformer är pluskvamperfekt, och den förlägger därmed något till det förflutnas förflutna, nämligen till någon tidpunkt före tidpunkten för mordet på Kennedy. Det som förläggs dit är en om-så-relation. I kraft av denna om-så-relation skall den till (4) relaterade satsen

(5) Om Oswald inte kommer att skjuta Kennedy, så kommer Kennedy att vinna valet 1964

ha varit sann vid tidpunkten i fråga, dvs en tidpunkt före mordet (vilket alltså visar den prediktiva karaktären hos (4)). Om nu (5) var sann vid denna tidpunkt, så kan inte

(6) Om Oswald inte kommer att skjuta Kennedy, så kommer någon annan att skjuta Kennedy

också ha varit sann vid samma tidpunkt. Men satsen (1) är av indikativt slag, och därmed inte relaterad till (6), dvs den implicerar inte att (6) var sann vid någon tidpunkt före mordet.

Slutpoängen stavas inte ut av Danielsson, men jag förmodar att läsaren skall förstå att (1) och (4) inte har samma egentliga hypotes. Den egentliga hypotesen i (4) är, via relationen till (5), framtidsrelaterad: "Oswald kommer inte att skjuta Kennedy". Den egentliga hypotesen i (1), vilken är densamma som i (1.1), är relaterad till förfluten tid.

Har Danielsson därmed en sammanhängande argumentation mot det inledande argumentet för att 'om, så' är tvetydigt? Nej. För att ha det måste också visas att ingen tvetydighet hos 'om, så' framtvings av hans egen teori. Det kan tyckas att så faktiskt är fallet, just på grund av de semantiska olikheterna mellan (1) och (4), och på grund av att bägge är om-så-satser. Om jag förstått Danielssons intentioner riktigt, så undviker han emellertid tvånget att postulera en tvetydighet hos 'om, så' av det skälet att (4) *inte* betraktas som en om-så-sats. Dvs, på ytan är (4) en om-så-sats, men i dess djupare grammatiska struktur är (4) bildad ur en om-så-sats genom att applicera en tempus-operator på hela om-så-satsen, en operator för förfluten tid i förhållande till en annan tidpunkt, nämligen tidpunkten för den händelse som omtalas i villkorssatsens försats, dvs mordet på Kennedy. Denna djupare struktur visar sig på ytan endast i verbformerna.

Lägg nu märke till att detta är det egentliga skälet till att vi inte

behöver postulera en tvetydighet. (4) är, djupare sett, inte en om-så-sats. Men till detta resultat kunde vi ha kommit utan tal om tidsreferenser och om-så-relationer, helt och hållet inom det vanliga semantiken för satser som (4), dvs kontrafaktiska villkorssatser. Vi kan nämligen säga att (4), djupare sett, är bildad genom att applicera en nödvändighetsoperator på en om-så-sats. Denna djupare struktur visar sig på ytan, återigen, genom verbformerna. Genom analysen i termer av en yttre nödvändighetsoperator erhåller vi sådana möjliga-världars-semantiska sanningsvillkor som i vissa gängse teorier (vilka inte diskuteras av Danielsson) associeras till kontrafaktiska villkorssatser.

Det förefaller alltså som om Danielssons teori i stor utsträckning är irrelevant för slutsatsen om tvetydighet. Å andra sidan förefaller diskussionen om tvetydighet egentligen vara ett retoriskt grepp, en infallsvinkel till hans egen teori, vilken är det verkliga ämnet för uppsatsen.

Mig veterligen är teorin originell, och värd att ta i beaktande. Samtidigt finns en hel del som talar emot den. Den är problematisk ur metafysisk synvinkel, eftersom den postulerar existensen av verkliga om-så-relationer som en del av vår ontologi, en typ av verkliga fakta. Logiskt komplexa fakta av detta och liknande slag, t ex negativa, är notoriskt svåra att redogöra för. Som del av en semantik för natur-

liga språk är den metodologiskt problematisk. Något som talar för teorier som postulerar tvetydighet hos 'om, så' är att de behandlar det som förefaller vara om-så-satser såsom verkliga om-så-satser, dvs bildade ur två satser genom en om-så-sammansättning, vilket är en fördel om vi vill bevara naturliga principer för s k kompositionalitet.

Dessutom förefaller mig Danielssons teori vara direkt felaktig. Jag kan mycket väl hävda (4) på grundval av omständigheter som inträdde först *efter* mordet på Kennedy. Antag t ex att det en tid före 1964 års val skedde en kraftig uppgång i amerikansk ekonomi och att orsakerna till denna uppgång var oberoende av mordet på Kennedy. Jag kan mycket väl anse att just dessa omständigheter är vad som gör satsen (4) sann.

Det bör också påpekas att det som är tillgängligt i uppsatsen i form av *skäl* för teorin är väl svagt, men jag skall inte här gå närmare in på det.

Sammantaget kan sägas att Danielsson i denna bok, som behandlar flera för modern analytisk filosofi centrala problem, visar prov såväl på originalitet som på ett stort mått av skarpsinne. Å andra sidan är framställningen ofta snårig. Läsaren har svårt att bli klar såväl över argumentationens gång som över dess slutsats. I vissa fall har jag själv misslyckats helt, men möjligheten finns att detta säger mer om mig än om Danielsson. *Peter Pagin*

Ludwig Wittgenstein

---

*Tractatus logico-philosophicus*

*Filosofiska undersökningar*

*Om visshet*

*Särskilda anmärkningar*

BÖCKER OM WITTGENSTEIN

*Essäer om Wittgenstein*

red. LARS HERTZBERG

*Wittgensteins Wien*

ALLAN JANIK och  
STEPHEN TOULMIN

*Wittgenstein  
om kunskap och visshet*

GUNNAR SVENSSON

---

Filosofi från THALES

## NOTISER

### ¶ *Mer om vetbarhetsparadoxen.*

Man kan skilja mellan *ovetade* sanningar och *ovetbara* sanningar. De förra är sanningar som man de facto är okunnig om, medan de senare är sådana som man i någon mening inte kan veta. Den så kallade vetbarhetsparadoxen, som Lars Bergström diskuterar i en notis i FT 4/92, utsäger att om alla sanningar är vetbara så är alla sanningar vetade.

Att det finns ovetade sanningar är en truism så god som någon. Men om man vill ha ett bevis för denna självklarhet råkar man i svårigheter. Det tycks nämligen vara omöjligt att uppvisa en sanning utan att den blir vetad.

Bergström söker ta udden av denna paradox genom att visa att alla sanningar inte är vetbara. Det förefaller dock som Bergströms resonemang rör sig i cirkel. Han antar att en viss sats är sann och dessutom ovetbar och drar därav slutsatsen att det finns ovetbara sanningar.

Ett bevis för att det finns åt-

minstone ovetade sanningar kan formuleras så här:

Det finns en utsaga  $p$  sådan att jag varken vet  $p$  eller icke- $p$  (ex: det finns fysiker i Andromedagalaxen);

Jag vet dock:  $p$  är sann eller icke- $p$  är sann;

Av detta följer att det finns en utsaga som är sann men ovetad.

Det förefaller rimligt att den första premissen kan förstärkas så att man kan visa att det också finns sanna utsagor som är *ovetbara*. Det bör påpekas att Kurt Gödel har ett bevis för att det finns *matematiska sanningar* som är ovetbara i den meningen att de inte kan bevisas.

Bergström menar att vetbarhetsparadoxen är oförarglig om man kan visa att inte alla sanningar är vetbara. Men är det inte paradoxalt att man ur det icke helt orimliga antagandet att alla sanningar är vetbara kan härleda att alla sanningar är vetade? Den paradoxen berörs dock inte av Bergströms resonemang. *Per Ericson*

### ¶ Svar till Ericson

Per Ericson menar att man "råkar i svårigheter" om man vill ha ett bevis för att det finns ovetade sanningar. Men svårigheterna verkar överkomliga, ty han ger sedan själv ett bevis för just detta. En annan sak är att beviset kanske inte skulle övertyga den som är skeptisk till lagen om det uteslutna tredje.

Ericson anser också att Kurt Gödel har ett bevis för att det finns sanningar som inte kan bevisas. Men det är väl inte riktigt sant. Såvitt jag vet säger Gödels teorem att det inom varje system (av en viss karaktär) finns någon sats sådan att varken satsen själv eller dess negation kan bevisas *inom detta system* – de kan med andra ord inte kan härledas från systemets egna axiom med systemets egna härledningsregler. Detta utesluter inte att en sådan sats eller dess negation kan bevisas på annat sätt.

Men hur är det då med själva vetbarhetsparadoxen. Ericson menar att mitt resonemang rör sig i cirkel. Det tror jag inte, men kanske har jag uttryckt mig oklart. Frågan är om det ligger något paradoxalt i att man från antagandet att alla san-

ningar är vetbara kan härleda att alla sanningar är vetade. Vid första anblicken kan det nog tyckas så, men vad jag hävdade var att det *inte* är paradoxalt. Sålänge det finns ovetade sanningar är det nämligen *uppenbart* att det också finns sanningar som är ovetbara, nämligen de sanningar som består av konjunktionen av (1) en ovetad sanning, och (2) det sanna påståendet att första ledet i konjunktionen inte är vetad. Denna konjunktion är ju sann, och den är naturligtvis också ovetad (eftersom första konjunkten är ovetad). Men den är dessutom ovetbar. Den *kan* nämligen inte vara vetad, ty antagandet att den är vetad leder till en motsägelse, närmare bestämt till motsägelsen att någonting är både vetat och inte vetat.

Alltså: om det finns ovetade sanningar, så finns det ovetbara sanningar. Detta är uppenbart, och alltså inte paradoxalt. Och härav följer att om alla sanningar är vetbara, så är alla sanningar vetade. Eftersom detta är ekvivalent med något som inte är paradoxalt, kan det inte gärna självt vara paradoxalt.

Lars Bergström

¶ Platons dialoger *Gästabudet*, *Försvartalet* och *Faidon* har under den gemensamma titeln *Om kärleken och döden* utgivits i nyöversättning av Jan Stolpe och med inledning av Staffan Bergsten. Förlag: Studentlitteratur.

¶ Ludwig Wittgensteins *Om visshet* i svensk översättning av Lars Hertzberg har nyutgivits av bokförlaget Thales.

¶ I en recension i SvD den 22/1 av Georg Henrik von Wrights nyutkomna bok *Myten om framsteget* (Bonniers) hävdar författaren Lars Gustafsson att "det är inkonsekvent att resa till fina internationella symposier med jetplan för att hävda teknikens generella skadlighet för mänskligheten". Har Gustafsson rätt? Är det verkligen inkonsekvent? Man kan tycka att det beror på vad som närmare bestämt skall menas med termen "inkonsekvent". I *någon* mening är det förstås inkonsekvent att inte leva som man lär. Men är det inkonsekvent i en mening som innebär att man gör något *fel* när man betar sig inkonsekvent? Är det med andra ord irrationellt eller omoraliskt att inte leva som man lär? Detta kan kanske tjäna som ett enkelt övnings-exempel i logik och/eller moralfilosofi.

¶ Under tjugo år har forskare från Sverige kom två deltagare, och hela världen, inte minst filosofer, övriga länder som representerades samlats till seminarier vid *Inter-University Center of postgraduate studies* i den vackra staden Dubrovnik i södra Kroatien. På grund av kriget i det forna Jugoslavien har naturligtvis verksamheten varit ganska reducerad på senare tid, men den har hela tiden upprätthållits. Våren 1992 hölls t ex som vanligt ett möte med den grupp för vetenskapsteori som startades redan 1974 av IUC:s grundare Ivan Supek. Från Sverige kom två deltagare, och övriga länder som representerades var England, Polen, Kanada, USA, Grekland och Tjeckoslovakien. IUC:s egen byggnad, alldeles utanför den medeltida stadskärnan, var helt söderskjuten av serbiskt artilleri, så de vetenskapsteoretiska mötena fick äga rum i stadens musikskola. Även i år kommer ett möte med den vetenskapsteoretiska gruppen att äga rum i april (om inget oförutsett inträffar).

¶ På Daidalos förlag har utgivits *Emmanuel Lévinas. En introduktion* av Peter Kemp, i översättning av Rikard Hedenblad, 1992.

¶ Alfred Tarskis berömda utredningar om sanningsbegreppet har bland annat efterlämnat den allmänt spridda uppfattningen att satsen "snö är vit" är sann om och endast om snö är vit. Eller är inte detta en spridd uppfattning? En visserligen ovetenskaplig, men dock undersökning, vid filosofiska institutionens i Stockholm lunchbord ger vid handen att teoretiska filosofer säger "snö är vitt" medan praktiska filosofer säger "snö är vit". Korrelationen är inte perfekt, men anmärkningsvärt hög. Hur skall detta förklaras? Vem har rätt? Om bägge uttrycken eventuellt är korrekta, har de i så fall något olika betydelser? Särskilt den sista frågan vore intressant att få ett svar på. (Ett korrekt svar, alltså.)

¶ *Något att tänka på.*

Den unge adelsmannen Johan Ludvig Morgonstierna går och väntar på ett arv efter en gammal onkel, som ligger för döden. Han är på rätt goda grunder förvissad om att han skall ärva; såvitt han vet är han ende kvarlevande släkting, och onkeln har tidigare sagt att Johan Ludvig skulle få ärva allt. Av en tillfällighet råkar han emellertid få syn på en kopia av onkelns nyupprättade testamente, där alla pengar testamenteras bort till en "Jean Louis Morgenstern".

Johan Ludvig ser framför sig ett liv i fattigdom, men kommer snart på tanken att mörda denne för honom okände arvtagare. Med Morgenstern ur vägen måste pengarna gå till Johan Ludvig. Han gör vissa ansträngningar att spåra upp Morgenstern, men spåren leder bara till advokatbyrån Mill, Frege, Kripke & Partners, där man säger sig vara förhindrad att avslöja något om Morgenstern. Däremot kan man tänka sig att vidarebefordra brev, paket eller meddelanden.

Alltså måste Johan Ludvig skicka en brevbomb för att nå sitt mål. Han bygger en sådan och skickar iväg den till advokatbyrån. Där fattar man misstankar och larmar polisen. Polisen kommer och desarmerar bomben. Spåren leder polisen till Johan Ludvig, som erkänner.

Men en faktor komplicerar utredningen. Efter en kontroll med advokatbyrån visar det sig att den

mystiske Jean Louis Morgenstern helt enkelt är Johan Ludvig Morgonstierna själv.

Åklagaren har då att ta ställning till frågan om Johan Ludvig Morgonstierna skall anses ha försökt begå självmord – när allt kom omkring skulle ju ingen annan ha dött. Det är inte olagligt att försöka ta livet av sig. Däremot borde Morgonstierna då kanske skickas till psykolog, för att tala ut om sina självmordstankar.

Åklagarens andra alternativ är att förbereda ett åtal för mordförsök – mordförsök är ju avgjort något olagligt. Det står ju alldeles klart att Morgonstierna hade "ont i sinnet" när han planerade sitt dåd: hade Morgenstern varit en annan person, så hade ju saken varit klar. Morgonstiernas beteende påminner ju inte alls om en presumtiv självmördare, utan han har på alla sätt och vis betett sig som en som försöker mörda någon annan, och sådant beteende måste ju beivras.

Skall åklagaren lägga ner målet, eller skall han fortsätta att förbereda ett åtal för mordförsök? Här finns flera frågor, som tidskriftens läsare inbjuds att kommentera.

En grupp frågor som juridiskt kunniga läsare kan kommentera är frågor om det juridiska läget. Vad skulle åklagaren faktiskt göra, och enligt vilka principer skulle detta beslut fattas? Skulle man fatta olika beslut i olika länder? Finns det till och med någonstans ett autentiskt rättsfall med samma struktur?



De filosofiska frågorna rör främst frågor kring vad Morgonstierna hade för avsikter, och vad det var han försökte göra. Dessa

frågor går utöver de juridiska, och därför skulle kommentarer också till dessa frågor uppskattas.

Fredrik Stjernberg

¶ Tore Nordenstam, som är professor i filosofi vid universitetet i Bergen sedan 1968, har utgivit en bok om humanvetenskapernas utveckling från forntiden till de modernas humanvetenskapernas genombrott för 100 år sedan och fram till våra dagar. Bokens titel är *Från konst till vetenskap*, Carlssons Bokförlag 1993.

¶ Kenneth Holmqvist har disputerat i kognitionsvetenskap på en avhandling med titeln *Implementing cognitive semantics*, Lund University Cognitive Studies nr 17, Lund 1993.

¶ *De gröna smaragdernas gåta*  
Nelson Goodmans problem (från "The New Riddle of Induction", 1953, i *Fact, Fiction, and Forecast*, 1955): Varför är bara vissa egenskaper "projicerbara" i induktiva sammanhang? Alla hittills observerade smaragder har varit gröna, men de har också alla varit gröda, där egenskapen *gröd* definieras som *antingen observerad före år 2000 och grön eller ej observerad före år 2000 och röd*. Varför drar vi då den induktiva slutsatsen att alla smaragder är gröna och inte att de alla är gröda? (Observera att vi inte kan dra bägge slutsatserna samtidigt om vi vill undvika oförenliga förutsägelser om de smaragder som inte observeras före år 2000.) Är det därför att *grön* är en enklare och mera naturlig egenskap än *gröd*? Men i vilken mening? Beror det på att den senare egenskapen är definierad i termer av den förra och att definitionen dessutom involverar en

godtycklig tidsreferens? Men definitionsordningen skulle ha kunnat vara annorlunda. I vårt språk är sådana predikat som "grön" och "röd" primitiva. Vi använder dem för att definiera komplexa predikat "gröd" och "rön" (= *antingen observerad före år 2000 och röd eller ej observerad före år 2000 och grön*). Men i ett språk i vilket "gröd" och "rön" är primitiva predikat är definitionsordningen den motsatta. I ett sådant språk är "grön" ett komplext predikat som definieras i termer av "gröd" och "rön": *antingen observerad före år 2000 och gröd eller ej observerad före år 2000 och rön*. (Definitionen av "röd" blir analog. Att den ovanstående definitionen av "grön" är korrekt kan man lätt övertyga sig om, om man betänker att "grön" och "gröd", resp "grön" och "rön", sammanfaller ifråga om objekt som har resp inte har observerats före år 2000.) Är då vårt induktiva beteen-

de beroende av hur vårt språk råkar vara utformat? Är induktiv projicerbarhet en produkt av språkliga tillfälligheter?

Lars Bergström ("Gröna smaragder", FT 1/93) avvisar detta defaitistiska alternativ och föreslår en (partiell) lösning på Goodmans problem: Det finns en viktig skillnad mellan *grön* och *gröd* som inte är av språklig karaktär – den förra egenskapen, men inte den senare, är direkt observerbar. Och direkt observerbara egenskaper är *cæteris paribus* enklare och därmed lämpligare för induktiv projektion än andra egenskaper.

Låt oss säga att en egenskap E är *direkt observerbar* om det för varje föremål x och varje tidpunkt t gäller, att om man vid t kan observera x i alla observerbara avseenden, så kan man vid t genom observation avgöra om x har E (vid t). I denna mening är grönhet en direkt observerbar egenskap. Men grödhet är inte en direkt observerbar egenskap. Ty antag att vi vid någon tidpunkt *efter* år 2000 observerar en sak som är [grön eller röd] ... vi kan normalt inte genom observation vid en given tidpunkt *efter* år 2000 avgöra om den sak vi då observerar har observerats före år 2000.

Observationer lämnar inga observerbara märken, och ännu mindre lämnar de några observerbart daterade märken. Därför

kan vi då inte heller avgöra om den sak vi observerar är gröd eller ej. ... Ur empiristisk synpunkt verkar det vidare rimligt att hävda att en direkt observerbar egenskap är *enklare* än en egenskap som inte är direkt observerbar (allt annat lika). Den är enklare i den meningen att det, i åtminstone vissa fall, är enklare att avgöra om den föreligger eller inte.

Nu menar Bergström inte att lösningen ifråga besvarar alla problem som uppkommer i samband med induktiv projicerbarhet. Dels är det inte sant att alla direkt observerbara egenskaper kan projiceras. Projicerbarhet kan upphävas av bakgrundsevidens: "Om t ex alla åsnor jag har sett är grå, så drar jag ändå inte slutsatsen att alla åsnor är grå, för jag vet redan att djur av samma sort kan ha olika färger". Dels finns det fall där vi är beredda att projicera egenskaper som inte är direkt observerbara: det kan ju finnas fall i vilka en icke-observerbar egenskap saknar direkt observerbara konkurrenter som är förenliga med vår induktiva evidens: "...om vi någon gång *efter* år 2000 märker att alla smaragder som vi observerat både före och efter år 2000 har ändrat färg från grönt till rött, och att detta tycks ha skett just på nyårsdagen år 2000, så ligger det nära till hands att dra den induktiva slutsatsen att alla smaragder faktiskt är [grönaföre-år-2000-och-röda-därefter]"

även om denna egenskap (som emellertid inte får sammanblandas med grödhet) är allt annat än direkt observerbar.

Bergström påstår emellertid att han åtminstone har lyckats ta hand om Goodmans exempel: han har lyckats förklara varför grönhet ur projektionssynpunkt är att föredra framför grödhet och diverse andra grödlika konkurrenter.

Egenskapen grönhet är enligt ovanstående resonemang enklare än egenskapen grödhet, och den är också enklare än alla andra egenskaper som kan konstrueras i analogi med grödhet. Därför är också hypotesen att alla smaragder är gröna enklare än alla konkurrerande hypoteser av Goodman-typ. Därför bör den också föredras. Vilket skulle bevisas.

Har Bergström verkligen lyckats bevisa grönhetens överlägsenhet i jämförelse med *alla* grödlika konkurrenter? Svaret måste bero på vad vi anser vara utmärkande för grödlika egenskaper. Det är uppenbart att Bergström med sin lösning klarar av alla egenskaper av formen *antingen observerad före t och grön eller ej observerad före t och F*, där *t* står för en godtyckligt vald tidpunkt och *F* är en godtyckligt vald egenskap (som inte sammanfaller med grönhet). Inga sådana egenskaper är direkt observerbara.

Vi kan därför gå längre och säga att Bergström klarar av alla grödlika

egenskaper av formen *antingen (O och grön) eller (icke-O och F)*, förutsatt bara att den särskiljande O-egenskapen inte är direkt observerbar.

Men vad händer om vi väljer en egenskap *O* som är direkt observerbar? Låt oss fundera en stund på detta alternativ. I Goodmans exempel har varje hittills observerad smaragd den icke direkt observerbara egenskapen att vara observerad före år 2000 – en egenskap som vi inte tror tillkommer alla smaragder. (Detta är viktigt, ty annars skulle man ju inte kunna hävda att hypotesen att alla smaragder är gröda är oförenlig med hypotesen att alla smaragder är gröna.) Men kan vi inte hitta en *direkt observerbar* egenskap som vi inte heller tror tillkommer alla smaragder men som råkar tillkomma varje smaragd som hittills observerats? Detta borde inte vara svårt. (Tänk på Bergströms exempel med gråa åsnor!) Antag *tex* att varje hittills observerad smaragd har befunnits väga under fyra kilo (en direkt observerbar egenskap), men att vi ändå är övertygande om att det borde finnas smaragder som väger betydligt mer än så. Men om *väger under fyra kilo* är en direkt observerbar egenskap, så måste ju detsamma gälla egenskapen

gröd\* = antingen väger under fyra kilo och är grön eller väger åtminstone fyra kilo och är röd.

För varje objekt *x* och vid varje tidpunkt då objektet ifråga är "ob-

serverbart i alla avseenden” kan jag avgöra om  $x$  är grött\* genom observation. (Jag antar här att vägning också räknas som observation. Annars får man ta ett annat exempel.)

*Gröd\** tycks vara en värdig grödliknande konkurrent till *grön*, ty (1) alla hittills observerade smaragder är såväl gröna som gröda\*; (2) vi tror att det finns smaragder (de som väger åtminstone fyra kilo)

#### ¶ Svar till Rabinowicz

I min ursprungliga artikel påpekade jag att min lösning av Goodmans problem gällde just hans eget exempel, och att man inte kan dra slutsatsen att alla direkt observerbara egenskaper kan projiceras. Rabinowicz exempel illustrerar detta. För övrigt utgick jag från vanliga idéer om hypotesval, och det jag närmare bestämt tog fasta på var tanken att man bör föredra enkla hypoteser (allt annat lika). Det mer allmänna problem som Rabinowicz – och om man så vill redan Goodmans – exempel aktualiserar, är om *alla* konkurrenter till grönhet *med samma satslogiska struktur* som grödhets kan avfärdas med hjälp av de vanliga idéerna om hypotesval, alltså de som faktiskt brukar användas i vetenskap.

Beträffande just Rabinowicz egenskap grödhets\*, är jag benägen att säga att den *inte* är lika enkel som grönhet. Intuitivt skulle väl många vara beredda att hålla med om detta, men går denna intuition att förklara mer i detalj? Det är inte

som inte kan vara gröna och gröda\* på en gång; men (3) till skillnad från *grön* är *gröd\** en egenskap som vi inte är beredda att projicera. Samtidigt är denna konkurrent till grönhet en direkt observerbar egenskap! Bergströms lösning förefaller inte räkna särskilt långt: han har inte givit oss några goda skäl att tro att alla smaragder är gröna.

*Wlodek Rabinowicz*

alls säkert att det går med hjälp av begreppet direkt observerbarhet, men låt mig ändå försöka.

Rabinowicz säger att grödhets\* är direkt observerbar. Jag är inte så säker på det. För resonemangets skull skall jag dock gå med på att egenskapen att väga under fyra kilo är direkt observerbar (även om det kan diskuteras). Men även om var och en av en uppsättning egenskaper är direkt observerbar, så följer inte därav att varje satslogisk kombination av dem också är direkt observerbar. I det aktuella exemplet går jag med på att gröd\* är en satslogisk kombination av observerbara egenskaper, men jag skulle trots det vilja hävda att grödhets\* inte är direkt observerbar. Detta framgår inte, åtminstone inte på något klart sätt, av min definition av ”direkt observerbar”, men denna utnyttjar ju termer som ”observera”, ”observerbar” och ”observation”, och frågan är då hur dessa skall tolkas. Jag är benägen att säga att en egenskap är observerbar, om den kan tillskrivas ett objekt med hjälp av

en "observationssats" i den amerikanske filosofen W V Quines mening, vilket bland annat innebär att satsen i fråga skall kunna läras ut genom ostension. Man skall kunna lära ett barn att använda satsen korrekt genom att peka ut saker och samtidigt uttala satsen. Jag tror inte att detta är möjligt med en sats som "Denna sak är gröd\*".

Antag att jag har fel på denna punkt, och att det faktiskt är möjligt att lära ut "gröd\*" genom ostension. I så fall skulle jag ändå vilja hävda att det är *svårare* att lära ut "gröd\*" än "grön" genom ostension, och att grönhet i denna mening är en *enk-lare* egenskap än grödhhet\*. Detta talar då i sin tur för att hypotesen att alla smaragder är gröna bör före-

dras framför hypotesen att alla smaragder är gröda\*.

Härmed är ämnet ingalunda ut-tömt, men eftersom detta nummer av *Filosofisk tidskrift* just nu skall skickas till tryckning kan jag för närvarande inte göra något mer åt det. Tills vidare överlåter jag i stället åt tidskriftens många be-gåvade läsare att komma med nya uppslag. En skillnad mellan Rabino-wicz och Goodmans exempel som kanske kan vara värd att fundera över är denna: vi har hittills aldrig observerat någon sak som är gröd och inte grön, men vi har däremot ofta observerat saker som är gröda\* och inte gröna (nämligen röda saker som väger fyra kilo eller mer).

*Lars Bergström*

Medarbetare i detta nummer av *Filosofisk tidskrift*: *Dan Egonsson* har doktorerat i praktisk filosofi i Lund, *Lars Bergström* är professor i praktisk filosofi i Stockholm, *Claudio Tamburrini* har doktorerat i praktisk filosofi i Stockholm, och *Peter Pagin* är doc i teoretisk filosofi i Stockholm.

**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 fackfilosofier, 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.

#### MANUSKRIFT TILL FILOSOFISK TIDSKRIFT

- sändes till redaktören, Lars Bergström, Stora Ängby allé 24, 161 54 Bromma
- skall vara försedda med namn och adress
- skall som regel vara skrivna på svenska, men bidrag på norska och danska accepteras också
- skall vara maskinskrivna på A4-papper med skrift endast på arkets ena sida; bör ej innehålla mer än 1 500 nedslag per sida och skrivs med bred vänstermarginal (räkna med 50 nedslag/rad och med 30 rader/sida = 1 500 nedslag)
- skicka gärna med manus även på diskett för Macintosh (disketten kommer naturligtvis att returneras)
- noter och litteraturhänvisningar inarbetas i texten
- särskild litteraturförteckning upprättas i alfabetisk ordning och placeras sist i manus
- för icke beställt material ansvaras ej
- korrektur läses i regel endast av redaktören
- ändringar mot manuskript bekostas icke av förlaget; författaren debiteras kostnad för sådana ändringar
- införda bidrag honoreras ej för närvarande
- i stället för särtryck erhåller författaren gratis 10 exemplar av det nummer av tidskriften i vilket bidraget varit infört

Tidigare årgångar av Filosofisk tidskrift kan köpas direkt från förlaget. Adressen är Box 50034, 104 05 Stockholm.

För årg 1/1980 — 10/1989 är priset 25 kr/årg, lösa ex 10 kr

För årg 11/1990 — 13/1992 är priset 100 kr/årg, lösa ex 30 kr

Porto tillkommer



