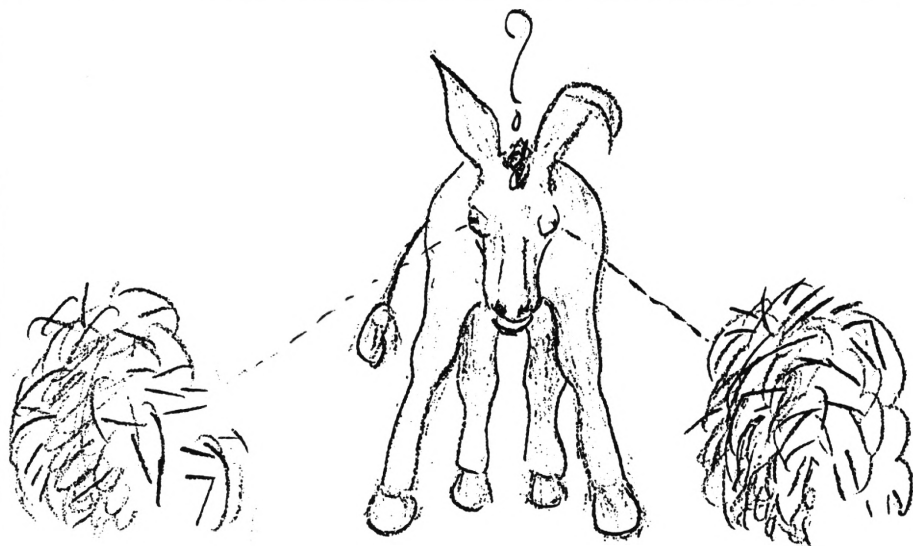


# Filosofisk tidskrift

ÅRGÅNG 11

NUMMER 4

1990



**Włodzimierz  
Rabinowicz**

*Rationella beslut under osäkerhet  
eller  
Hur skall man handla i blindo?*

**Thomas Brante**

*Rationalism, relativism och paradoxer*

**Ragnar Ohlsson**

*Tragiska valsituationer*

**Lars Bergström**

*Moraliska observationer*

**Torbjörn Tännsjö**

*Visst kan vi observera det goda!*

# FILOSOFISK TIDSKRIFT NR 4 \* 1990 \* ÅRG 11

<i>Wlodzimierz Rabinowicz</i>	Rationella beslut under osäkerhet eller Hur skall man handla i blindo?	1
<i>Thomas Brante</i>	Rationalism, relativism och paradoxer	13
<i>Ragnar Ohlsson</i>	Tragiska valsituationer	22
<i>Lars Bergström</i>	Moraliska observationer	32
<i>Torbjörn Tännsjö</i>	Visst kan vi observera det goda!	40
<i>Lars Bergström</i>	Replik till Tännsjö	46
Notiser		48

---

Filosofisk tidskrift utges av Stiftelsen Filosofisk tidskrift och Bokförlaget Thales.

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

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

---

*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 Westerståhl (Stockholm)

*Produktion, prenumerationer och annonser:* Bokförlaget Thales, Box 50034, 10405 Stockholm, redaktionssekreterare Bitte Malmnäs, tel 08 16 33 58 må, on, fr.

*Prenumeration:* Tidskriften kommer ut med 4 nummer/år (normalt i februari, april, september och november), pris SEK 120:-, lösnummer SEK 35:-, pg 507991-8 Filosofisk tidskrift

*Annonspriser:* 1/1-sida SEK 250:-, 1/2-sida SEK 150:-, 1/4-sida SEK 100:-

*Upplaga:* 1 000 ex/nummer

Ombrytning och sättning utförd av Ord och Grafik, Stockholm

Tryckt hos International Tryck AB, Uppsala

Włodzimierz Rabinowicz

***Rationella beslut under osäkerhet  
eller  
Hur skall man handla i blindo?***

Antag att en agent ställs inför ett något så när välstrukturerat beslutsproblem. Han har att välja mellan ett antal *handlingar*,  $A_1, \dots, A_m$ , vilkas utfall beror på hur världen är beskaffad (eller kommer att vara beskaffad). Världens möjliga alternativa *tillstånd* betecknar vi som  $S_1, \dots, S_n$ . Agenten vet att exakt ett av dessa tillstånd föreligger, men han vet inte vilket. En handling  $A_i$ 's *utfall* givet tillstånd  $S_j$  betecknar vi som  $o_{ij}$ . Vi får därigenom ett utfallsmatris:

	$S_1$	.....	$S_n$
$A_1$	$o_{11}$	.....	$o_{1n}$
$A_2$	$o_{21}$	.....	$o_{2n}$
.	.....	.....	
.	.....	.....	
$A_m$	$o_{m1}$	.....	$o_{mn}$

**Exempel 1.** En läkare måste välja mellan olika former av medicinering. Han är delvis osäker på vilken sjukdom patienten lider av (möjliga tillstånd). Beroende på patientens sjukdom leder en medicinering till olika utfall: patienten tillfrisknar utan bestående men, med bestående men, hans sjukdom påverkas ej, den förvärras, osv.

**Exempel 2.** En politiker väljer mellan olika alternativa åtgärder att plädera för. Hans osäkerhet gäller flera faktorer: åtgärdernas effektivitet med hänsyn till olika målsättningar, majoritetsförhållanden i riksdagen, opinioner i väljarkåren, etc. Alternativa tillstånd är alltså i detta fall relativt komplicerade. Utan denna komplikation vore det svårt att någorlunda entydigt specificera de utfall som hans olika val skulle kunna leda till.

Att identifiera alternativa handlingar, tillstånd och utfall innebär *redan* en långtgående strukturering (och förenkling) av beslutsproblemet. Men låt oss antaga att denna strukturering är genomförd. Beslutsfattarens situation är fortfarande svår.

Vilken handling skall han välja? Pondera att han tilldelar olika *sannolikheter* till de tänkbara tillstånden. Låt  $p$  vara hans sannolikhetsdistribution. Vi antar att varje tillstånd får ett  $p$ -värde mellan noll och ett, och att tillståndens  $p$ -värden summeras till ett. Pondera dessutom att han tilldelar vissa numeriska värden eller nyttor till de olika utfallen. Vi tänker oss att valet av numeriska värden till en del, men *endast* till en del, får vara godtyckligt – på samma sätt som valet av måttenheten och nollpunkten när man mäter temperatur är en fråga om en konvention, medan temperaturskillnadernas relativa storlek inte är konventionell.

Givet dessa antaganden ger den klassiska ("Bayesianska") beslutsteorin en entydig rekommendation: agenten bör välja den handling vars *förväntade värde* är *maximalt*. (Finns det flera sådana handlingar, spelar det ingen roll vilken som väljs.)

En handling  $A_i$ 's *förväntade värde* (förkortat,  $eu(A_i)$ ), där "eu" står för "expected utility") är en viktad summa av värden hos dess utfall under olika tillstånd. Som vikter används tillståndssannolikheterna.

$$eu(A_i) = p(S_1)u(o_{i1}) + \dots + p(S_n)u(o_{in}).$$

Att handlingens värde sammanfaller med dess väntevärde är en central tes inom den Bayesianska ansatsen. Vanligtvis argumenterar man för denna tes genom att man härleder den från någon uppsättning intuitiva antaganden som en rimlig handlingsvärdering bör uppfylla. Klassiska exempel på sådana "axiomatiska" argument förekommer hos von Neumann och Morgenstern (*Theory of Games and Economic Behaviour*, 1953) och hos Leonard J Savage (*The Foundations of Statistics*, 1954). I argument av denna sort antar man vanligtvis att agenten har en preferensordning över utfall och handlingar och försöker visa att denna preferensordning, om den uppfyller vissa axiom, kan ges en kvantitativ representation som uppfyller väntevärdeprincipen: den kan med andra ord representeras med hjälp av en numerisk värdefunktion över utfall och handlingar som låter handlingarnas värde vara lika med deras väntevärde.

Det är oklart hur mycket sådana argument visar. Att en preferensordning *kan* ges en representation som satisfierar väntevärde-

principen innebär ingalunda att varje kvantitativ representation av denna ordning måste satisfiera principen ifråga.<sup>1</sup>

Här skall jag i stället skissera en annan version av argumentet, som återfinns i en uppsats av Graham Oddie och Peter Milne.<sup>2</sup>

Oddie och Milnes utgångspunkt är från början kvantitativ: de antager från början en bestämd numerisk värdetilldelning till såväl utfallen som handlingarna. Vad de visar är att denna värdetilldelning, om den uppfyller vissa tillsynes plausibla antaganden, måste satisfiera väntevärdeprincipen.

Oddie och Milne utgår från fyra antaganden som här presenteras i en något modifierad form:

(i) Ett handlings värde – dess "instrumentella" värde, om man så vill – bestäms helt och hållet av utfallens värden och av tillståndssannolikheterna; det är, med andra ord, en funktion av  $p$  och av de värden som  $u$  tilldelar till handlingens olika tänkbara utfall.

(ii) Om alla utfall har samma värde, är handlingens värde identisk med värdet hos utfallen.

(iii) Låt  $A_i$  och  $A_k$  vara två olika handlingar vilkas utfall har samma värde för varje tillstånd med ett enda undantag: för exakt ett tillstånd,  $S_j$ , gäller det att  $u(o_{ij})$  skiljer sig från  $u(o_{ik})$ . Då är skillnaden mellan  $A_i$ :s och  $A_k$ :s värde helt och hållet bestämt av tre faktorer:  $u(o_{ij})$ ,  $u(o_{ik})$ , och sannolikheten  $p(S_j)$  för det särskiljande tillståndet.

(iv) Låt  $A_i$  och  $A_k$  vara som i antagandet (iii) ovan. Då är skillnaden mellan  $A_i$ :s och  $A_k$ :s värde som störst när sannolikheten för det särskiljande tillståndet är som högst, dvs när  $p(S_j)$  är lika med 1.

Givet dessa fyra till synes intuitiva antaganden bevisar Oddie och Milne att handlingens värde måste vara *lika* med dess väntevärde.

Antagandena i beviset kan säkert ifrågasättas. Antagandet (i), som ligger till grund för alla de andra, uppställer en kontroversiell

---

<sup>1</sup> En intressant diskussion av detta problem förekommer i Per-Erik Malmnäs' än så länge opublicerade uppsats "Expected Utility and the Herstein-Milnor Axioms".

<sup>2</sup> "Act and Value. Expectation and Representability of Moral Theories", opublicerat manuskript.

restriktion på vilka faktorer som kan vara handlingsvärderrelevanta. Men även om vi accepterar (i), får vi problem med antagandet (iii), som i vissa avseenden påminner om Savages berömda "sure-thing-principle".<sup>3</sup> Det kan angripas på ungefär samma sätt som man brukar angripa Savages princip.<sup>4</sup>

Följaktligen kan även det Bayesianska receptet – att maximera väntevärdet – kritiseras. (Antagandena (i) - (iv) är inte bara tillräckliga för en identifiering av handlingsvärdet med väntevärdet. De är också nödvändiga konsekvenser av en sådan identifiering. Det är lätt att visa att väntevärdet uppfyller alla dessa fyra antaganden. En invändning mot något av dessa blir därigenom *ipso facto* en invändning mot den Bayesianska identifieringen.) Men det är mera angeläget att ifrågasätta själva *förutsättningarna* för den aktuella ansatsen. Har beslutsfattaren alltid en bestämd sannolikhetsdistribution och en bestämd värdefunktion över utfallen att utgå ifrån? Tänk bara på vad som händer när man skall träffa ett beslut i en omdiskuterad politisk fråga. Valet av en förvaringsform för kärnavfallet kan vara ett lämpligt exempel. Olika vetenskapliga instanser och utredningar bombarderar beslutsfattaren med diver-

---

<sup>3</sup> "Sure-thing"principen kan formuleras så här: Om en handling  $A_i$  är att föredra framför en annan handling  $A_k$  och bägge handlingar leder till identiskt utfall under ett visst möjligt tillstånd, så är  $A_i$  fortfarande att föredra framför  $A_k$ , när vi tillskriver dem ett annat, men fortfarande identiskt, utfall under detta tillstånd och håller de andra utfallen oförändrade. Med andra ord: att en handling är att föredra framför en annan är något som helt och hållet bestäms av deras *divergerande* utfall och av sannolikheter för de tillstånd under vilka deras utfall divergerar. Likheten mellan denna princip och antagandet (iii) är uppenbar.

<sup>4</sup> Låt  $A_i$  och  $A_k$  vara som i antagandet (iii), och låt  $A_j$  och  $A_l$  vara ett annat par avhandlingar vilkas utfallsvärden endast skiljer sig åt när  $S_j$  är förhanden. Om nu  $u(o_{ij}) = u(o_{lj})$  och  $u(o_{kj}) = u(o_{lj})$ , så följer det ur (iii) att skillnaden mellan  $A_i$ 's och  $A_j$ 's värde måste vara lika stor som skillnaden mellan  $A_j$ 's och  $A_k$ 's värde.

Men om  $A_i$ 's utfallsvärden är genomgående höga (så att även  $u(o_{ij})$  är högt) medan  $u(o_{kj})$  är mycket lågt (så att  $A_k$ , som annars leder till höga utfallsvärden kan åstadkomma en katastrof om  $S_j$  blir fallet), kan en riskaversiv person uppfatta värdeskillnaden mellan  $A_i$  och  $A_k$  som ganska stor, hur liten sannolikhet han än tilldelar åt  $S_j$ . Å andra sidan, om utfallsvärden för  $A_k$  är genomgående mycket låga, så att även  $A_j$  får mycket låga utfallsvärden för alla tillstånd förutom  $S_j$ , kan den riskaversive komma fram till att värdeskillnaden mellan  $A_j$  och  $A_k$  inte är så stor. Sådana riskaversiva handlingsvärderingar strider alltså mot antagandet (iii). Det är emellertid en fråga för sig om en dylik form av riskaversion är en rimlig inställning.

gerande sannolikhetsberäkningar – riskuppskattningar varierar från en expert till en annan. För att ytterligare förvärra saken konfronteras den stackars beslutsfattaren med ett helt spektrum av konkurrerande utfallsvärderingar – tänkbara utfall evalueras olika och rankas t o m olika från varierande etiska perspektiv.

Beslutsfattaren tvingas därför att utgå från en *klass* P av alternativa värdefunktioner, snarare än från en bestämd p och en bestämd u. (Redan antagandet av avgränsade klasser P och U utgör en idealisering. Gränserna för P och U kan i det verkliga livet vara ganska obestämda och vaga.)

Det är inte lätt att fatta ett beslut under såpass obehagliga omständigheter. Låt oss därför, för att förenkla problemet, först fundera på vad som händer om vi bortser från de "axiologiska" komplikationerna – från osäkerheten på värdeplanet. Vi antar alltså att beslutsfattaren utgår från en klass P av sannolikhetsdistributioner – han är osäker ifråga om sannolikheter – men vi låter honom förfoga över en bestämd värdefunktion u över utfallen. Vi skall titta på några beslutsprinciper som har föreslagits för ett sådant fall. Först därefter kommer vi att ta ställning till frågan om dessa principer kan överleva i någon form när själva utfallsvärdena blir osäkra.

Den första principen vi skall betrakta har diskuterats av flera beslutsteoretiker (Robbins, Wald, Hurwicz). Under senare år har den försvarats av två Lundafilosofer, Peter Gärdenfors och Nils-Eric Sahlin.<sup>5</sup> Här presenteras deras ansats i en något förenklad form.

Låt  $A_i$  vara en möjlig handling.  $A_i$  får olika väntevärden  $e_{ip}$  ( $A_i$ ) för de olika sannolikhetsdistributioner p som tillhör klassen p. Vi definierar  $A_i$ :s minimala väntevärde som det lägsta väntevärde som  $A_i$  kan få för någon p i P.

Gärdenfors & Sahlin föreslår nu en beslutsregel som de benämner MMEU (Maximize the Minimal Expected Utility):

---

<sup>5</sup> Se deras uppsats "Unreliable probabilities, Risk-taking, and Decision-Making", *Synthese* 53, 1982, s 361-386; omtryckt i deras antologi av beslutsteoretiska texter *Decision, Probability, and Utility – Selected Readings*, Cambridge University Press 1988, s 313-334. Se också Nils-Eric Sahlins "Three decision rules for generalized probability representations", *The Behavioral and Brain Sciences* 84, 1985, s 751-753.

Välj den handling vars minimala väntevärde är jämförelsevis maximalt.<sup>6</sup>

Deras ansats kan alltså ses som en kombination av två idéer: å ena sidan är det viktigt att *maximera väntevärde* – precis som Bayesianer alltid har hävdad, å andra sidan får man gardera sig mot katastrofer – det gäller att *maximera minimum*.

Dessa två huvudidéer återkommer hos en annan filosof – Isaac Levi.<sup>7</sup> Levi börjar med att introducera begreppen "*e-tillåtenhet*" (e-admissibility) och "säkerhetsnivå". En handling  $A_i$  är e-tillåten om det finns någon sannolikhetsdistribution i  $P$  med avseende på vilken  $A_i$  maximerar väntevärde. Dvs om det finns  $p$  i  $P$  sådan att  $eu_p(A_i) \geq eu_p(A_j)$  för varje handlingsalternativ  $A_j$ . Med andra ord, en handling är e-tillåten om det finns någon sannolikhetsdistribution i  $P$  med avseende på vilken handlingen ifråga framstår som det bästa alternativet.

En handlings *säkerhetsnivå* är lika med värdet hos det värsta utfall som kan inträffa om handlingen utförs.  $A_i$  kan leda till något av utfallen  $o_{i1}, o_{i2}, \dots, o_{in}$ , med varierande  $u$ -värden.  $A_i$ 's säkerhetsnivå är lika med det lägsta av dessa  $u$ -värden.

Levis rekommendation är nu som följer:

---

<sup>6</sup> Egentligen är deras ansats lite mer komplicerad. De antar att sannolikhetsdistributionerna kan ha olika grad av *epistemisk opålitlighet*. Beslutsfattaren bör endast ta hänsyn till de distributioner som är "tillräckligt pålitliga" – som framstår som seriösa alternativ. Vilken grad av pålitlighet som får anses vara tillräcklig beror på beslutsfattarens beredskap att ta "epistemiska risker". Ju djärvare han är, desto högre kan han sätta pålitlighetströskeln och desto fler alternativa distributioner kan han därigenom bortse ifrån.

Begreppet "epistemisk risk" diskuteras närmare av Nils-Eric Sahlin i "On Second Order Probabilities and the Notion of Epistemic Risk", *The Foundations of Utility and Risk Theory and Applications*, utg av BP Stigum and F Wenstop, Reidel Publ Comp 1983, s 95-104, och i "On Epistemic Risk and Outcome Risk in Criminal Cases", *In So Many Words – Philosophical Essays Dedicated to Sven Danielsson*, utg av S Lindström och W Rabinowicz, Department of Philosophy 1989, s 176-186.

<sup>7</sup> Se hans böcker *The Enterprise of Knowledge*, The MIT Press 1980, och *Hard Choices* Cambridge University Press 1986. Se också hans uppsatser "On Indeterminate Probabilities", *Journal of Philosophy* 71, 1974, s 391-418, och "Ignorance, Probability and Rational Choice", 53, 1982, s 387-414. Den förra uppsatsen omtrycktes med vissa ändringar i Gärdenfors & Sahlins antologi *Decision, Probability, and Utility*, Cambridge UP 1988, s 287-312.



Välj den e-tillåtna handling vars säkerhetsnivå, i jämförelse med andra e-tillåtna handlingar, är maximal.<sup>8</sup>

Som vi märker, är det igen fråga om en kombination av de två huvudidéerna: maximering av väntevärde kommer in via kravet på e-tillåtenhet, medan säkerhetsnivån utgör det minimum som skall maximeras.

Men Levis sätt att kombinera de två idéerna skiljer sig från Gärdenfors & Sahlins. För att illustrera skillnaden låt oss betrakta ett exempel.<sup>9</sup>

*Biltillverkningens dilemma.* En biltillverkare är osäker om framtida bensinpriser. Låt HÖGT stå för *Bensinpriset kommer att gå upp* och LÅGT för *Bensinpriset kommer att förbli oförändrat eller minska*. Biltillverkaren har att välja mellan tre handlingsalternativ: antingen *investera i småbilstillverkning* (SMÅ) eller *investera i stora bilar* (STORA), som slukar mycket bensin, eller, slutligen, att *avstå från nya investeringar* (PASS). Alternativens utfall blir givetvis olika beroende på bensinprisets utveckling. Utfallsvärden ges i nedanstående matris:

	Bensinpriset	
	HÖGT	LÅGT
SMÅ	9	-6
STORA	-9	12
PASS	0	0

PASS-alternativet ger varken vinst eller förlust. Att satsa på små bilar är mindre farligt än att satsa på de stora: den eventuella förlusten om man satsar fel blir mindre. Å andra sidan blir även den eventuella vinsten – om man satsar rätt – mindre: storbilstillverkning ger större vinstmarginaler. Biltillverkaren är osäker på den framtida bensinprisutvecklingen. Olika ekonomiska experter har framlagt divergerande prognoser. Låt oss anta att deras sannolik-

---

<sup>8</sup> Levi anser inte att maximering av säkerhetsnivån är ett rationalitetskrav. Snarare uppfattar han den som en rimlig policy som är värd att rekommendera. Dessutom utesluter han inte användning av extra kriterier för att göra ett ytligare urval bland e-tillåtna handlingar med maximal säkerhetsnivå. Det kan ju finnas flera sådana handlingar i ett givet beslutsproblem.

<sup>9</sup> Gärdenfors & Sahlin använder sig av ett liknande exempel när de visar att MMEU och Levis beslutsprincip ger olika rekommendationer.

hetstilldelningar till HÖGT varierar mellan  $1/3$  och  $2/3$ . (Sannolikheten för LÅGT varierar på ett motsvarande sätt mellan  $2/3$  och  $1/3$ .) Med andra ord: Biltillverkarens klass P av sannolikhetsdistributioner omfattar flera olika  $p$  där  $p$  (HÖGT) varierar mellan  $1/3$  och  $2/3$ .

Följande sammanställning ger de minimala väntevärdena för de olika handlingsalternativen:

SMÅ	-	-1	(tänk på det fall när HÖGT får sin lägsta sannolikhet)
STORA	-	-2	(tänk på det fall när HÖGT får sin högsta sannolikhet)
PASS	-	0.	

Gärdenfors & Sahlin skulle därför rekommendera PASS som det rationella valet.

Levi skulle däremot rekommendera alternativet SMÅ. Ty PASS är inte e-tillåtet! Det är lätt att kontrollera att det saknas en sannolikhetsdistribution för vilken PASS maximerar väntevärde. (Om  $p$  (HÖGT)  $\geq 1/2$ , så är  $eu_p$  (SMÅ)  $>$   $eu_p$  (PASS); om  $p$  (HÖGT)  $\leq 1/2$ , så är  $eu_p$  (STORA)  $>$   $eu_p$  (PASS).) För varje sannolikhetsdistribution är antingen SMÅ eller STORA, eller bägge, att föredra framför PASS. Samtidigt, bland de e-tillåtna handlingarna (SMÅ och STORA), är det SMP som maximerar säkerhetsnivå: det värsta som kan inträffa när SMÅ väljs är inte så dåligt som det värsta utfallet för STORA.

Det kan här noteras att Levis princip är ytterst känslig för vilka handlingar som är tillgängliga i beslutssituationen. Anta att agenten får reda på att alternativet STORA – det alternativ som *inte* rekommenderas av Levi – i själva verket inte är möjligt att utföra. (Regeringen har av miljöskäl förbjudit nya satsningar på bensinlukande stora bilar.) Stryker vi STORA från listan, blir plötsligt PASS e-tillåtet. (När  $p$  (HÖGT) =  $2/3$ , blir PASS' väntevärde maximalt – större än hos alternativet SMÅ.) Men när både SMÅ och PASS är e-tillåtna, blir det PASS som maximerar säkerhetsnivå bland de e-tillåtna handlingarna. Efter att STORA har stukits från listan, upphör SMÅ att vara den rekommenderade handlingen! Levi är helt klar över denna instabilitetseffekt, men han är ändå beredd att hålla fast vid sin beslutsprincip.

Gärdenfors & Sahlin råkar ut för en annan malör: deras rekommendation i exemplet skulle ändras om vi som extra handlingsalter-

nativ tillät *lotterier* över bashandlingar. Det är t ex lätt att räkna ur vad som kommer att hända om vi introducerar som ett möjligt alternativ ett lotteri i vilket SMÅ och STORA får lika stora chanser: vi låter alltså ett myntkast fälla avgörandet mellan dessa två satsningar. Detta lotteri har 1 som sitt minimala väntevärde. (Tänk på det fall när HÖGT tilldelas den maximala sannolikheten  $2/3$ . Lotteriets väntevärde blir då  $2/3(1/2(9) + 1/2(-9) + 1/3(1/2(-6) + 1/2(12))) = 1$ . Väntevärdet ökar när sannolikheten för HÖGT minskar.) Om man vill höja minimiväntevärdet ytterligare – till 1,5 – kan man välja ett lotteri som ger SMÅ chansen  $7/12$  och STORA chansen  $5/12$ . I vilket fall som helst blir det inte längre PASS som är det rekommenderade alternativet enligt Gärdenfors & Sahlin. Att deras beslutsprincip föreskriver att man skall förlita sig på slumpen är både oroande och smått absurt.<sup>10</sup>

Själva det faktum att förnuftiga personer som Gärdenfors & Sahlin och Levi kommer med så olika rekommendationer i ett så pass enkelt fall gör en fundersam. Det är kanske rimligt att lägga vikt vid maximering av väntevärde och vid maximering av minimum. Men vad hjälper det en, om man kan lägga vikt vid dessa faktorer på *olika sätt? Hur kan man avgöra vilket sätt som är rimligast?*

Dessutom: Finns det inte andra relevanta faktorer att ta hänsyn till? Varför skall man stirra sig blind på minimum och strunta i *maximum*? Det är viktigt att gardera sig mot det värsta, men det är

---

<sup>10</sup> Även Levi skulle rekommendera ett lotteri mellan SMÅ och STORA (med chanserna  $7/12$  resp  $5/12$ ). Ett sådant lotteri skulle maximera säkerhetsnivå bland de e-tillåtna alterantiven. (Notera att utfallsvärden för ett handlingslotteri är lika med viktade summor av utfallsvärden för de handlingar man lottar emellan. Som vikter används de olika handlingarnas lotterichanser. Ett lotteri mellan SMÅ och STORA som fördelar chanserna i förhållande sju mot fem är e-tillåtet ty det maximerar väntevärde när  $p(\text{HÖGT}) = p(\text{LÅGT}) = 1/2$ . Dess säkerhetsnivå är samtidigt så högt som 1,5.)

Levi kan emellertid kunna undgå denna glorifiering av slumplösningar genom att *partitionera* om beslutsproblemet. Om man i stället för att laborera med två tillstånd – HÖGT och LÅGT, väljer att arbeta med en finare indelning kommer lotteriets säkerhetsnivå att falla. Om vi tillåter sådana tillstånd som *Bensinpriset kommer att gå upp och lotteriet, om det spelades, skulle leda till en satsning på stora bilar*, så skulle lotteriets säkerhetsnivå minska till -9. Säkerhetsnivån beror på hur fin tillståndspartition vi väljer att laborera med – något som Levi själv framhåller. Han vill emellertid inte påstå att ju finare partition desto bättre. I själva verket ger han oss inga tips om hur man skall förfara när man strukturerar ett beslutsproblem.

också viktigt att bereda möjligheter för det bästa. Så skulle man kunna tycka. Men ju fler faktorer man tar hänsyn till, desto svårare blir det att uppställa en entydig, men inte godtycklig, beslutsregel.

Situationen förvärras ytterligare när vi introducerar värdeosäkerhet i beslutssituationen. Vi antar alltså att agenten inte längre har någon bestämd  $u$ -funktion att utgå ifrån. I en sådan situation kan det mycket väl hända att rekommendationen att maximera minimum inte längre är *meningsfull*.

Antag t ex att världen kan vara i ett tillstånd  $S_1$  eller  $S_2$  och att agenten har att välja mellan två olika handlingar,  $A_1$  och  $A_2$ . Dessa antas leda till samma utfall givet  $S_1$ , men till olika utfall givet  $S_2$ .

	$S_1$	$S_2$
$A_1$	$o$	$o'$
$A_2$	$o$	$o''$

Ponera att agenten konfronteras med två olika värdeperspektiv som han inte kan välja emellan. Enligt bägge dessa perspektiv är  $o$  det bästa utfallet, men de skiljer sig ifråga om inbördes ranking av  $o'$  och  $o''$ .

Rankning 1	Rankning 2
$o$	$o$
$o'$	$o''$
$o''$	$o'$

Givet att agenten inte kan avgöra mellan perspektiven, kan han inte längre identifiera det värsta möjliga utfallet. Minimum blir inte längre väldefinierat.

Givetvis går det att variera exemplet så att inte heller maximum blir väldefinierat. Går man över från en bestämd  $u$ -funktion till en klass  $U$  av konkurrerande värdefunktioner, förlorar begreppen minimum och maximum sin innebörd.

Är det något i de tidigare skisserade ansatserna som fortfarande kan användas? Ja, begreppet *e-tillåtenhet* överlever. En handling är *e-tillåten* om den maximerar väntevärde för *något* val av  $p$  ur  $P$  och  $u$  ur  $U$ .

Levi kan alltså hävda att vi fortfarande får en viss vägledning: det är irrationellt att välja en handling som inte är *e-tillåten*. Men vägledningen är av tveksamt värde. Det kan ibland vara rimligt att

välja en "evig tvåa", fast eviga tvåor inte är e-tillåtna. För att vara e-tillåten måste en handling komma som en etta för något val av  $p$  och  $u$ .

Levi bemöter detta problem genom att kräva att såväl  $P$  som  $U$  skall vara konvexa klasser: om distributioner  $p$  och  $p'$  tillhör  $P$ , så skall även varje "kompromiss" mellan de två – varje mellanliggande sannolikhetsdistribution – tillhöra  $P$ . Ett analogt krav gäller  $U$ . Antag nu att en handling är en god tvåa med avseende på par  $(p, u)$  och  $(p', u')$ . Den är ett bra kompromissförslag beroende på att de andra handlingarnas väntevärde våldsamt varierar. I en sådan situation, menar Levi, kommer handlingen ifråga att maximera väntevärde (och därigenom vara e-tillåten) med avseende på något par  $(p^*, u^*)$ , där  $p^*$  är en kompromiss mellan  $p$  och  $p'$ , medan  $u^*$  är en kompromiss mellan  $u$  och  $u'$ . Med andra ord: Givet konvexitetsantagandet kan inte en god tvåa vara en "evig" tvåa.

Drar man sig för att anta konvexitet, bör man ersätta kravet på e-tillåtenhet med ett svagare krav: för att kunna komma i fråga, får en handling  $A_i$  inte vara *dominerad* av någon annan handling som agenten kan utföra. Dvs det får inte finnas ett alternativ till  $A_i$  vars väntevärde överstiger väntevärdet hos  $A_i$  för varje val av  $p$  ur  $P$  och  $u$  ur  $U$ . Notera att en dominerad handling inte kan vara e-tillåten, men inte omvänt: en handling som är e-otillåten behöver inte vara dominerad.

Även om man vore beredd att anta konvexitet, kvarstår svårigheten att Levis vägledning inte är särskilt detaljerad. Andelen e-tillåtna handlingar kommer normalt att vara relativt stor. Seriösa handlingsalternativ är ofta alla e-tillåtna.

Det finns tre möjliga sätt att se på beslutsteori. Den kan betraktas som

- en deskriptiv disciplin (en underavdelning till psykologi),
- en preskriptiv disciplin (en uppsättning rationalitetsprinciper och beslutsregler,
- en verktygslåda (en uppsättning begreppsliga verktyg som kan komma till användning när vi strukturerar beslutsproblem och letar efter lösningar).

När agenten tvingas välja en epistemisk *och* axiologisk osäkerhet, lyser beslutsreglerna med sin frånvaro. Det mesta han kan hoppas på är att beslutsteoretikern förser honom med begreppsliga

instrument som han sedan får använda på eget ansvar.<sup>11</sup> Men även detta är kanske bara en from förhoppning. Den axiologiska osäkerheten gör att många beslutsteoretiska standardverktyg kan visa sig oanvändbara. Vi sitter där med en hammare utan spikar.<sup>12</sup>

□

---

<sup>11</sup> Det var Jan Odelstad som hade fått mig att inse att verktygs-lådetolkningen på ett intressant sätt skiljer sig från den preskriptiva synen på beslutsteori.

<sup>12</sup> Denna uppsats presenterades vid ett seminarium om osäkerhet och beslut i Hässelby Slott 4-6 april 1990. Seminariet anordnades av Statens Kärnbränsle-nämnd och Samrådsnämnden för Kärnavfallsfrågor. Jag vill gärna tacka seminariets organisatörer, Christian Munthe och Folke Tersman, samt deltagarna, särskilt Sven Ove Hansson, Jan Odelstad, Per-Erik Malmnäs, Torbjörn Tännsjö och Lennart Sjöberg.

**Thomas Brante**

## ***Rationalism, relativism och paradoxer***

Ett av de mest svårlösta problem kring vilket den klassiska filosofin har utvecklats är frågan om rationalistiska versus relativistiska uppfattningar om kunskapens natur. Idag diskuteras denna motsättning framför allt i termer av skillnaden mellan rationalistiska/empiricistiska och relativistiska *vetenskapsteorier*. Båda dessa framstår som mycket rimliga och plausibla när de tillåts utveckla sina egna epistemologier på egna villkor, både vad gäller principiell filosofisk argumentering och illustrativa fallstudier av vetenskapliga förlopp.

Samtidigt som rationalismen och relativismen framstår som rimliga tagna var för sig så är de uppenbarligen enkla att attackera och underminera med utgångspunkt i det motsatta perspektivet. En rationalist kan med lätthet peka ut ett flertal avgörande svagheter i ett relativistiskt filosofiskt system, och vice versa.

En grundläggande orsak till detta förhållande är alla filosofiska systems natur. Alla filosofier tycks nämligen bygga på obevisbara antaganden, eller axiom, som ger paradoxliknande konsekvenser. Och när dessa angripes så rasar systemet samman och vad som återstår kan endast upprätthållas medelst ett visst mått av tro eller ogrundad övertygelse, dvs ett visst mått av dogmatism. Det finns mängder av exempel på detta, såsom att det empiriska verifikationskriteriet inte självt kan verifieras empiriskt, att falsifikationskriteriet inte kan falsifieras, att paradigmteorin bygger på ett paradigm, och så vidare.

Men det är också så att själva förhållandet mellan rationalism och relativism är paradoxliknande. I denna korta uppsats skall jag först illustrera vad jag menar med ovanstående påståenden genom exempel. Exemplet visar vilka versioner av relativism och rationalism jag åsyftar, kort sagt Feyerabend's relativism, och en standard-tolkning av Poppers rationalism. När paradoxen är frilagd övergår jag till att kommentera den och kortfattat ge förslag till hur den möjligen kan förstås, för att därmed kunna passeras.

## 1. Tes

Enligt den popperska kritiska rationalismen finns det en grundläggande form för vetenskaplig härledning, som utgör vetenskapernas gemensamma bas. Sociologen Arthur Stinchcombe har försökt konkretisera denna universiella metodologi genom att rekonstruera en berömd sociologisk teori: Emile Durkheims teori om självmordets sociala orsaker. Durkheims grundläggande tes är att en ökning av antalet (egoistiska) självmord i en population är orsakad av en minskning av gruppsolidariteten i samma population, och vice versa. Från detta teoretiska påståenden kan empiriskt observerbara påståenden härledas, t ex att "Protestanter i Frankrike kommer att ha högre självmordsfrekvens än katoliker i Frankrike" (med premissen att protestanter har lägre gruppsolidaritet än katoliker).

Detta empiriska påstående kan testas genom observation, t ex statistik över antal självmord i olika grupper. När observationerna är utförda kan vi jämföra dem med den empiriska härledningen och avgöra om påståendet är sant eller inte. Låt oss kalla teorin för A, och den empiriska konsekvensen för B.

Deduktionen har följande logiska form:

$A \rightarrow B$

B är sann

A är sannolik

För att testa och eventuellt stärka teorin kan man nu låta den undergå multipla test genom att härleda och pröva flera empiriska konsekvenser, t ex att protestantiska länder uppvisar högre självmordsfrekvens än katolska länder (B1), att protestantiska regioner i Frankrike har en högre självmordsfrekvens än katolska regioner (B2), att protestanter i Tyskland har en högre självmordsfrekvens än katoliker i Tyskland (B3), och så vidare:

$A \rightarrow B1, B2, B3, \dots Bn$

B1, B2, B3, ... Bn är sanna

A är mer sannolik

För att ytterligare stärka teorin letar vi nu efter empiriska konsekvenser av ett annat slag, som är härledbara från den allmänna teorin. Durkheims teori implicerar också att gifta män i Frankrike kommer att ha lägre självmordsfrekvens än ensamma



män, och att de kommer att ha ännu lägre självmordsfrekvens om de har barn; professionella och välutbildade personer kommer att ha högre självmordsfrekvens än arbetare och lägre utbildade, och så vidare. Låt oss kalla dessa nya empiriska påståenden för D, E och F, och vi antar att de konfirmerats genom observation, varvid vi erhåller:

$$\frac{A \rightarrow D, E, F}{D, E, F \text{ är sanna}}$$

A är trovärdig

Ytterligare ett sätt att pröva och eventuellt stärka teorin är att vederlägga de mest sannolika *alternativa* teorierna, t ex att mental-sjukdomar är orsaker till självmord. Det vill säga; vi utför ett *avgörande experiment* genom att explicera de konsekvenser av den ursprungliga teorin A vars *negation* impliceras av de viktigaste alternativa teorierna. Låt oss kalla det huvudsakliga alternativet C:

$$\frac{A \text{ eller } C}{A \rightarrow B_i}{C \rightarrow \text{icke } B_i}{B_i \text{ är sann}}$$

C är falsk

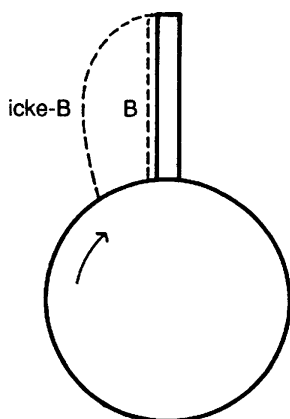
A är mycket trovärdig

Alternativa teorier falsifieras systematiskt, varvid den ursprungliga teorin ytterligare ökar i trovärdighet. Detta tillvägagångssätt är själva grundbulten i den kritiska rationalismens metodologi, givet villkoret att man verkligen försöker finna alternativ vars konsekvenser är "farliga" för den egna teorin.

## 2. Antites

Den kritiska rationalismens främste vedersakare, Paul Feyerabend hävdar att den ovan beskrivna metoden *de facto* aldrig använts inom vetenskapen (historiskt argument), att den inte *kan* tillämpas (logiskt argument), och att den inte *bör* användas (normativt argument). Feyerabends viktigaste historiska bevis för dessa påståenden härstammar från hans analys av Galileis bidrag till den heliocentriska världsbilden. Så för att förstå Feyerabends tankegång återvänder vi till 1500-talet och situationen inom astronomi vid

tiden för den kopernikanska revolutionen. Vi förkortar de pedagogiska stegen ovan och går direkt på det avgörande experimentet. Låt oss kalla den ptolemaiska teorin A och den kopernikanska teorin C. Ett av de viktigaste och ovedersägligaste bevisen för den ptolemaiska teorin är faktumet att "tung kroppar ... som faller ner från en höjd rör sig längs en rät och vertikal linje mot jordens yta" (Galilei). Det vill säga, jorden måste vara orörlig eftersom om den rör sig "borde en sten träffa jorden ett stycke ifrån tornets bas" (Galilei). Låt B vara det observerbara påståendet om rakt fall, som testas genom tornexperimentet:



Således:

A eller C

A → B

C → icke B

B är sann

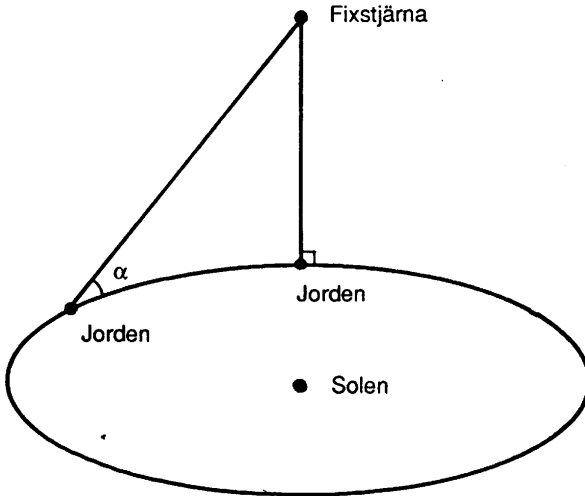
C är falsk

A är trovärdig

A kan stärkas ytterligare genom andra experiment med fallande kroppar: B1, B2, ... Bn, vilket gör A mer trovärdig.

Empiriska påståenden *av andra slag* som är härledbara från A och som är bekräftade genom observation gör A "ännu mer trovärdig". En konsekvens som talar för att jorden inte är i rörelse, och som Tyko Brahe ansåg vara avgörande, var följande: Om jorden snurrar

kring solen så bör den vinkel från vilken en fixstjärna observeras ändras med tidens gång, allteftersom jorden rör sig från en sida av solen till en annan<sup>1</sup>:



Men Brahe kunde inte iaktta några förändringar vad gäller vinkeln till fixstjärnorna. Låt oss kalla Brahes empiriska konsekvens för D, och komplettera den med andra konsekvenser av A av andra slag, kallade E och F (t ex himlarnas rörelser):

A eller C

A → B, D, E, F

C → icke (B, D, E, F)

B, D, E, F är sanna

C är falsk

A är mycket mer trovärdig

Således är den kopernikanska teorin vederlagd, enligt den kritiska rationalismens egna principer. Om den kritiska rationalismen verkligen är det vetenskapliga tänkandets essens, så borde den kopernikanska revolutionen inte ha ägt rum vid denna tid.

### 3. Inte en syntes men ... en paradox

Låt oss nu undersöka Feyerabendts tankekedja. Med stöd av resultat som de ovanstående drar Feyerabend slutsatsen att den kritiska rationalismens metodologi måste vara felaktig, *givet dess egna*

*principer*. Feyerabend hävdar att det inte var "rationella" faktorer och empiriska observationer utan snarare *övertalning* som utgjorde den viktigaste drivkraften bakom det kopernikanska paradigmskiftet, och att detta är fallet för alla vetenskapliga framsteg. Han menar vidare att det är orsaker som lekfullhet, makt och propaganda som ligger till grund för förändringar vad gäller vetenskapens riktning och innehåll i allmänhet. Låt oss denna gång kalla Feyerabends tes för A, och den kritiska rationalismens tes för C. Hur går Feuerabend tillväga för att stärka sina egna påståenden? Först tar han fram negativ evidens mot C, som sin egen Galileistudie, tillsammans med resultat inom den moderna vetenskapen som den browniska rörelsen, övergången från Newton till Einstein, och annat. B, D, E och F är den kritiska rationalismens version av hur teorier etableras:

$$C \rightarrow B, D, E, F$$

$$A \rightarrow \text{icke } (B, D, E, F)$$

$$\hline B, D, E, F \text{ är falska}$$

C är falsk  
A är trovärdig

Därefter förstärker Feyerabend A genom olika slag av positiv evidens såsom Piagets teori om barnets utveckling, lekens betydelse för den intellektuella utvecklingen, perspektivets betydelse för observationen (G, H, I) – faktorer som samtliga betraktas om irrelevanta eller metafysiska och som befinner sig utanför vetenskapens logik, enligt den kritiska rationalismen. Således:

$$A \text{ eller } C$$

$$A \rightarrow G, H, I$$

$$C \rightarrow \text{icke } (G, H, I)$$

$$\hline G, H, I \text{ är sanna}$$

C är falsk  
A är mer trovärdig

#### 4. Diskussion

Feyerabend använder sig således av den kritiska rationalismens metod för att visa att den kritiska rationalismen är felaktig. Så vad har egentligen skett i de tre stegen ovan? Feyerabend själv skulle svara att det enda sättet att övertyga eller övertala rationalister är att använda just deras egna metoder – något annat förstår de inte. Men

så enkelt är det naturligtvis inte. Ty om Feyerabend visat att rationalismens metod är felaktig med hjälp av en felaktig metod, så har han inte visat någonting. Och hur kan han då hävda att metoden är felaktig? Om han å andra sidan verkligen har visat att rationalismens metod är felaktig, så verkar den ju fungera – fast i så fall är den förstås falsk. Och de metoder Feyerabend själv anser vara paradigmskiftande – makt, till exempel – tycks inte ha använts av honom i detta försök till filosofiskt paradigmskifte. Möjligen har han använt sig av övertalning.

Lika litet som en principers giltighet kan visas genom hänvisning till just denna principers giltighet kan en principers ogiltighet visas genom användning av densamma, dvs det går inte att lita på en felaktig principers följd. Man måste stå utanför principens domän i båda fallen. Detta insåg också Popper när han hävdade att själva valet av rationalismen i grunden är irrationellt.

Förhållandet mellan rationalism och relativism tycks vara paradoxalt i den meningen att den relativistiska kritiken av rationalismen förutsätter ett visst mått av rationalistisk ordning och reda, samtidigt som den rationalistiska kritiken av relativismen förutsätter åtminstone ett irrationellt val. Båda förutsätter varandra, samtidigt som de förnekar varandras giltighet, vilket borde vara litet pinsamt, för båda.

(Med "paradox" menar jag helt enkelt en självmotsägande utsaga, t ex "Det finns inga sanningar", eller "Denna utsaga är falsk". Det tycks inte vara något formellt fel på dessa utsagor, ändå är deras konsekvenser, dvs effekten av att de uttalas, paradoxala. Rationalismen och relativismen innehåller, som jag försökt visa ovan, just paradoxala konsekvenser.)

Den lösning eller "upplösning" jag skulle vilja föreslå är att tänka förhållandet mellan rationalismen och relativismen som inkommensurabelt. Inkommensurabilitet anger förhållandet att två storheter inte kan mätas med samma mått. Inom vetenskapsteorin betecknar det förhållandet mellan teorier som tillhör olika paradig, och som därför bedömes med olika kriterier. "Intressanta" inkommensurabiliteter uppstår när sådana teorier är rivaliserande, dvs de försöker besvara samma fråga men gör det utifrån olika kriterier på vad som är ett godtagbart svar, vad som räknas som den korrekta vetenskapliga metoden, vad som ses som relevanta argument och fakta för frågans besvarande.

Ur det perspektiv jag använder här, som kan kallas "metavetenskapsteoretiskt", är rationalismen och relativismen inkommensu-

rabla i samtliga dessa avseenden. De förespråkar som vi har sett olika kriterier och olika bevisföringsregler, och har olika åsikter om vad som är relevanta fakta – för rationalismen utgör t ex så kallade "externa faktorer" irrelevanta eller till och med illegitima byggestenar i en bevisföring. Samtidigt tycks det problem båda syftar till att analysera och förklara vara detsamma, dvs fenomenet vetenskap och den vetenskapliga kunskapens grunder.

Om rationalismen och relativismen är inkommensurabla, så borde deras gemensamma problem vara "både lika och olika", dvs det finns avseenden i vilka de definierar och försöker lösa olika problem. Och så är det ju; rationalismens yttersta syfte är att blottlägga vetenskapens *logik*, och därför arbetar man med rationella rekonstruktioner av den empiriska verkligheten, dvs vetenskapens historia och samhällliga vara. Rationalismen skapar idealtyper av vetenskapens essens, som kan överensstämna mer eller mindre väl med den "faktiska" verkligheten, något som till viss del måste vara sekundärt för rationalismen – historiens många tillfälligheter får inte störa den genuina epistemologin alltför mycket. Relativismens yttersta syfte är å andra sidan att just spåra faktiska *orsaker* till den vetenskapliga kunskapens utveckling och innehåll, tillfälliga eller ej. Rationalistisk metodologi kan därvidlag utgöra *en* orsak bland flera – det är en empirisk fråga som inte kan besvaras *a priori*.

Om förhållandet mellan rationalism och relativism är inkommensurabelt, så följer att de i själva verket utgör två disparata forskningstraditioner som – idealt sett – skulle kunna ses som komplementära, dvs de konkurrerar inte utan arbetar med olika aspekter på ungefär samma fenomen, eller snarare olika delar av ett brett forskningsområde. Varför då denna historiska, synbarligen eviga animositet, varför dessa angrepp? Svaret kan ges av vetenskapssociologin, även om jag därmed tycks lämna den neutrala positionen och hemfalla åt empirisk spekulatoin. Även vetenskapsteoretiker och filosofer lever i ett komplext samhällsligt sammanhang, fyllt av kognitiva, professionella och gruppsspecifika intressen och motsättningar. Därav följer revirtänkande och försök att monopolisera forskningsområden, inom filosofin såväl som vid alla andra institutioner med begränsat antal tjänster och ansträngd budget. Det finns alltså irrationella grunder till den ofta mycket hätska debatten.

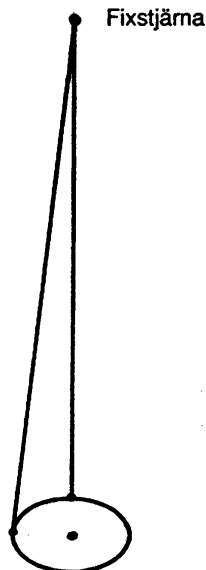
De båda perspektiven har också olika konsekvenser vad gäller frågor om vetenskapens/akademins autonomi, legitimitet och auktoritet. Kort sagt: rationalismen fungerar bra som försvarare av

vetenskapen, medan relativismen tenderar att profanera densamma.

Men rivaliteten mellan dessa båda perspektiv bör också ses som funktionell och fruktbar. Intellektuell utveckling kräver motstånd, ifrågasättande av grunder, klargörande av anomalier. Konflikt och inte harmoni är den bästa grunden för kreativitet och mot försoffning och dogmatik.

Så, för att avsluta: rationalismen och relativismen har gjort varandra goda tjänster i det att de har tvingat rivalen att finslipa sina argument. Men, som exemplifieringen ovan antyder skall vi inte vänta oss att någon av dem någonsin på kognitiva grunder kommer att kunna "övertyga" den andre och få dess utövare att byta sida – det är nämligen omöjligt, givet deras inkommensurabla relation. En förening kan endast uppnås på basis av en genuin syntes, eller troligare, en helt ny epistemologi. □

<sup>1</sup> Orsaken till Brahes felaktiga (vet vi nu) slutsats var att även om Brahe använde de för den tiden bästa instrumenten, så var hans beräkning av avståndet mellan jorden och fixstjärnorna alldeles för kort. Förhållandet är inte som i bilden ovan, utan snarare på detta sätt:



Ragnar Ohlsson

## *Tragiska valsituationer*

En tragisk valsituation är en situation, där två eller flera moraliska krav befinner sig i olöslig konflikt. Hur agenten än handlar, kommer hon att bryta mot någon giltig norm. Sådana konflikter är vanliga i skönlitteraturen, från Kreon i *Antigone* till Sophie i Styrons roman *Sophie's Choice*, som av lägerkommendanten får följande alternativ: antingen väljer hon ut ett av sina två barn, och det hon väljer får leva, eller också mördas båda genast.

Den traditionella uppfattningen bland moralfilosofier har länge varit att tragiska valsituationer är omöjliga. Visserligen kan det ibland verka som om någon befann sig i ett olösligt moraliskt dilemma, men sådana tragiska val är alltid skenbara. Såväl kantianer som utilitarister har i allmänhet varit övertygade om att olösliga moraliska konflikter måste vara omöjliga.

Sedan 1940-talet har emellertid denna ståndpunkt ifrågasatts av Jean-Paul Sartre, E J Lemmon, Bernard Williams, Thomas Nagel och många andra. Detta har utlöst en livlig debatt under 70- och 80-talen, där den traditionella ståndpunkten förfäktats med nya och nyslipade argument av Terrance McConnell, Alan Donagan, Earl Conee, o a, och möjligheten av tragiska valsituationer har försvarats av Ruth Barcan Marcus, Bas van Fraassen, Walter Sinnott-Armstrong, m fl.

Jag vill här introducera debatten genom att presentera två relativt nyutkomna böcker, som ger en god översikt över debatten, och diskutera var de viktiga motsättningarna mellan de två lägren finns. Den ena boken heter *Moral Dilemmas* (Oxford University Press 1987) och är en antologi, som innehåller de viktigaste texterna från Kant till T M McConnell. Den innehåller också en utmärkt, sammanfattande och klagörande inledning av Christopher W Gowans, som sammanställt boken. Den andra boken är en monografi av Walter Sinnott-Armstrong, med samma titel som Gowans' men utgiven av Basil Blackwell 1988. Den presenterar de



viktigaste argumenten för och emot möjligheten av moraliska dilemman och pläderar för möjligheten av deras existens.

### 1. Definition

Hur bör man definiera "tragisk valsituation" eller "moraliskt dilemma", som en del författare föredrar att benämna fenomenet? "Standarddefinitionen" ger vid handen att en person P befinner sig i en tragisk valsituation T, omm P i T bör utföra vardera av (minst) två handlingar men inte kan utföra bägge (alla). Om "förbud" respektive "förpliktelse" är definierbara i termer av varandra – så att t.ex. en handling är förbjuden omm det är obligatoriskt att underlåta den, spelar formuleringen ingen roll. Om man inte tror att så är fallet, är det lätt att i det följande företa nödvändiga förändringar för att resonemanget skall bli tillämpligt på en formulering i termer av förbud.

### 2. "Prima facie-bör" och "resultat-bör"

För många framstår idén om tragiska valsituationer som själv-motsägande. Att *prima facie*-bör eller moraliska skäl i största allmänhet kan komma i konflikt, finns det knappast någon som ifrågasätter. Det är klart att det i många situationer finns skäl, som talar för att en handling bör utföras och samtidigt skäl, som talar för att någon alternativ handling bör utföras. Men det kan ändå vara så att skälen som talar för den ena handlingen är starkare än skälen som talar för den andra, och att följaktligen de förra "väger upp" eller "slår ut" de senare. En sådan konflikt är lösbar och anses i allmänhet inte vara problematisk. De som har förnekat möjligheten av tragiska valsituationer har förutsatt att det "bör" som uttrycker de moraliska förpliktelser som kommer i konflikt i en tragisk val-situationsituation, skall tolkas på något starkare sätt. Man kan t ex uttrycka det så att det är "resultat-bör", resultat av en samman-vägning av alla moraliskt relevanta skäl som talar för respektive emot de olika alternativen i situationen. "Resultat-bör" uttrycker ett totalomdöme om situationen, som talar om vilken handling som är *den bästa* i situationen, vilken handling som är *överlägsen* de andra eller som uttrycker den norm som *tar överhand* över de andra, det är m a o *dominerande*. Alternativt kan det vara ett kategoriskt, konklusivt "bör", som inte förutsätter jämförelser mellan alternativen. Om "bör" uttrycker resultatet av en jämförelse mellan alternativen, kan det bara finnas en handling i varje situation som är *den bästa*. A kan inte vara bättre än B, när allt tagits med i

beräkningen, samtidigt som B är bättre än A, när allt tagits med i beräkningen.

Sinnott-Armstrong, som alltså hör till dem som försvarar möjligheten av tragiska valsituationer, instämmer i att två handlingar inte kan vara bättre än varandra. Han inför därför en distinktion mellan 'dominerande' ("overriding") och 'icke-dominerade' ("non-overridden") moraliska krav, och menar att ett moraliskt dilemma är en situation, där ett icke-dominerat moraliskt krav står mot ett annat icke-dominerat moraliskt krav. Earl Conee påpekar i en nypublicerad artikel (Conee 1989) att två icke-dominerade krav ställs t ex i situationer där två alternativ är lika bra. Men hittills har ingen tyckt att det är något tragiskt med sådana situationer – här har man bara att välja vilket som helst av de två alternativen! Vad som måste komma till är att kraven är kategoriska eller ovillkorliga, att man faktiskt *måste* göra både den ena och den andra handlingen (ehuru det är omöjligt att göra bägge). Ett "kategoriskt bör" eller ett "resultat-bör" anger en moralisk nödvändighet – och det är först om sådana "bör" kolliderar, som vi står i en tragisk valsituation. Men en sådan kollision är omöjlig enligt Conee, eftersom omdömen om moralisk nödvändighet förutsätter jämförelse mellan alternativen. Emellertid är det nog möjligt att kombinera Sinnott-Armstrongs ståndpunkt med den insikt Conees invändning uttrycker: man måste lägga till att de moraliska kraven är ovillkorliga, kategoriska eller dylikt, men *inte* att de är dominerande, dvs det kan vara moraliskt nödvändigt att utföra en handling, trots att detsamma gäller för flera alternativ.

### 3. Deontisk logiska argument mot tragiska valsituationer

Ett sätt att visa att tragiska valsituationer är omöjliga är att visa att det går att härleda motsägelser från antagandet att det finns tragiska valsituationer. Det går inte att göra direkt, men med hjälp av några standardantaganden i deontisk logik är det lätt att göra det. De som förnekar möjligheten av tragiska valsituationer ser detta som ett avgörande argument mot möjligheten av moraliska dilemman. Försvarearna menar i stället att detta visar att det är något fel på deontisk logik. Den är konstruerad för att undvika tragiska valsituationer: att hänvisa till den är alltså att förutsätta vad som skall bevisas.

Det finns flera sätt att visa oförenligheten mellan antagandet att tragiska valsituationer är möjliga och vissa standardantaganden i deontisk logik. Om vi tex accepterar att det är tillåtet att sluta från

Bör A och Bör B till Bör (A&B) – Bernard Williams kallar det för "agglomerations-principen" – och vidare att 'bör' implicerar 'kan', dvs att ingen kan ha en förpliktelse att göra något som han inte kan utföra, är det lätt att härleda motsägelsen :

P1 Bör A	Antagandena uttrycker att
P2 Bör B	A och B är alternativ i ett
P3 - Möjligt (A&B)	moraliskt dilemma
4 Bör (A&B) 1, 2,	Agglomerationsprincipen
P5 Bör $X \supset M X$ ,	'Bör' implicerar 'kan'
6 - $M(A\&B) \supset -Bör(A\&B)$	5, kontraposition
7 - Bör (A&B)	3, 6, modus ponens

4 och 7 motsäger varandra.

Ett annat sätt att härleda en motsägelse är via principen att vadhelst som är nödvändigt för att utföra en förpliktelse också är en förpliktelse.

P1 (Bör A & (A $\supset$ B)) $\supset$ Bör B	Deontisk princip
P2 Bör A	
P3 Bör B	Antagandena definierar en
P4 - $M(A\&B)$	tragisk valsituation
5 -(A & B)	Modal logik, 4
6 A $\supset$ -B	5, predikat logik
7 Bör -B	1, 2, 6, Modus ponens

7 motsäger inte direkt 3, men det verkar ändå egendomligt att man både bör utföra och underlåta B. Antar vi dessutom den deontiska logiska princip som säger att Bör  $X \supset -Bör -X$ , så kan vi från 3 härleda:

8 -Bör -B, vilket motsäger 7.

Accepterar man att "bör" definieras som "inte tillåtet att underlåta" och vidare att det som är förbjudet inte kan vara obligatoriskt går det också att härleda en motsägelse. (Den som är intresserad av härledningen hänvisas till Gowans bok, från vilken också de två tidigare härledningarna har hämtats i lätt förändrad och försvenskad form.) Men gentemot alla dessa argument kan man alltså invända att de egentligen förutsätter det som skall bevisas – de deontiska

principer de vilar på är konstruerade bl a för att omöjliggöra tragiska valsituationer.

Gowans menar att de deontiska principer som anförts i det föregående är giltiga för det dominerande 'bör', det "bör" som uttrycker moralisk nödvändighet (a a s 22ff). Det är också något som förklarar varför deontisk logik kunnat konstrueras med modallogiska förebilder: det uttrycker vad som *måste* göras.

#### 4. Argument för tragiska valsituationer

Vad kan man då ge för skäl *för* att tragiska valsituationer är möjliga? Bernard Williams hävdar att om man inte vet vilket av två motsägande påståenden som är sant, men till slut bestämmer sig för det ena, så förkastar man det andra som osant; om man däremot befinner sig i en moralisk konflikt och till slut väljer den ena handlingen, så känner man ändå ofta ånger över att man inte utförde den andra, man försöker kompensera den förfördelade parten, eller framför åtminstone ursäkter. Om man t ex har lovat någon att besöka honom vid en viss tid, men finner det viktigare att hjälpa en person efter en olyckshändelse, så förklarar man sitt beteende och ber om ursäkt i efterhand för att man inte kunnat hålla sitt löfte. Det visar, menar Williams, att även om ett moraliskt krav får vika i en viss situation, så behåller det sin bjudande kraft. Det finns en konflikt, även om det ena kravet tar överhand.

Philippa Foot (1983 omtryckt i Gowans 1987) menar att det är ett föga övertygande argument: avgörande är inte om en person känner ånger efter att ha handlat i en moralisk konfliktsituation, frågan är i stället om det är på sin plats att känna ånger. Sinnott-Armstrong skärper Williams' argument på just det sättet: han hävdar att det är passande att känna ånger efter att ha handlat i en tragisk valsituation. Men det är väl å andra sidan att förutsätta vad som skulle bevisas, som Foot påpekar när det gäller Williams.

Den andra typen av argument hävdar att moraliska krav uppstår på olika sätt, att det finns olika normer och förpliktelser. En del har att göra med våra roller – Agamemnon har en plikt som far och en annan som överbefälhavare för de akajiska trupperna – en del har att göra med speciella åtaganden (t ex löften) andra med människors rättigheter och intressen. Detta är Lemmons, Nagels och van Fraassens argument. Eftersom dessa förpliktelser är mycket olika varandra är de inte jämförbara.

Sinnott-Armstrong däremot menar att det inte är rimligt att hävda att moraliska skäl måste vara ojämförbara bara därför att de här-

stammar ur olika källor. Det är ofta klart att ett löfte som gäller något trivialt är mindre viktigt än förpliktelsen att rädda en människas liv, eller att plikter, som hänger samman med ens yrkesroll, får vika, om de kommer i konflikt med en rättighet. Och det är inte så att vissa typer av moraliska skäl alltid dominerar över andra; i så fall funnes en prioriteringsordning, som skulle lösa moraliska konflikter.

Däremot, menar Sinnott-Armstrong, kan det i bestämda situationer vara så att inget av de moraliska kraven är avgjort starkast. I sådana situationer råder begränsad ojämförbarhet, det finns ingen klar, objektiv prioritering mellan de olika kraven. Om olika ideala väljare skulle välja olika, utan att man kan visa på några defekter hos någon av dem, kan man dra slutsatsen att det faktiskt inte finns något moraliskt skäl att föredra det ena alternativet framför det andra, åtminstone inte något med universell giltighet. Att då införa en prioritering vore godtyckligt och skulle förutsätta vad som skulle bevisas, nämligen att tragiska valsituationer inte är acceptabla. Att det i vissa situationer finns ojämförbara krav beror enligt Sinnott-Armstrong bl a på bristande exakthet i "mätningen" av olika kravs styrka. Varje försök att införa en exakthet, som skulle utesluta moraliska dilemman, skulle bli godtycklig och föga plausibel, menar han.

Härtill kommer, enligt Sinnott-Armstrong, situationer där de moraliska kraven är av samma slag och lika starka, där kraven är *symmetriska* men oförenliga: det finns lika starka skäl för Sophie att välja det ena barnet som att välja det andra. Gentemot detta kan man invända att det knappast är detta som är problemet: har vi möjlighet att rädda högst en av två tvillingar som håller på att drunkna, tycks vi visserligen stå inför en tragisk situation, men inte en tragisk valsituation. Det är tragiskt att bara en tvilling kan räddas, men vilken tvilling vi än räddar, så handlar vi rätt. Fel gör vi bara om vi låter båda drunkna. Men Sophies val verkar värre: hon bör inte välja ut något av sina barn till att dödas, och inte heller bör hon låta båda dö. Alla alternativ verkar förbjudna. Det är därför hon står i en tragisk valsituation.

## **5. Ingen konflikt mellan dilemman och realism**

Nu tror jag vi kan börja se var den grundläggande motsättningen finns mellan försvarare och förnekare av tragiska valsituationer, den gäller olika syn på vad moraliska omdömen handlar om och vad ett moraliskt system är. Motsättningen gäller närmast vad slags

egenskaper moraliska egenskaper är, vad som utgör "moraliska fakta" s a s.

Bernard Williams har i flera uppsatser hävdad att möjligheten av tragiska valsituationer är oförenlig med moralisk realism. Argumentet är det som antydde ovan: om en moralisk utsaga är sann eller falsk, blir en moralisk konflikt analog med en logisk konflikt mellan motstridiga påståenden. En sådan konflikt måste alltid i princip vara lösbar – minst ett av påståendena måste vara falskt – medan en moralisk konflikt inte alltid är lösbar på samma sätt – det finns alltid en rest av ånger eller krav på kompensation eller ursäkter sedan man valt i en moralisk konflikt, även om man inte väljer "fel". Någon sådan rest finns inte sedan man bestämt sig för vilket av två påståenden som är sant. Alltså är moraliska konflikter inte logiska konflikter, problemet är inte att veta vilket av två påståenden som är *sant*. Därför är den moraliska realismen falsk, enligt Bernard Williams.

Men nu är det ju inte inkonsistent att påstå att A bör utföras och att B bör utföras, men att det är omöjligt att utföra A och B, om vi godtar Sinnott-Armstrongs distinktion mellan "dominerande" och "icke-dominerade moraliska krav" och definierar "moraliskt dilemma" med hjälp av det senare begreppet. Alla tre leden kan vara sanna. Ja, i själva verket kan man tycka att alla tre *måste* vara sanna för att det *verkligen* skall föreligga en tragisk valsituation. Om moraliska utsagor "bara" är rapporter om subjektiva attityder eller uttrycker attityder hos den som yttrar dem, eller om någon annan icke-objektivistisk analys av moraliska utsagor är riktig, förefaller det närmast oegentligt att säga att det *verkligen* föreligger en tragisk valsituation. För att man skall kunna skilja mellan skenbara moraliska konflikter och verkliga måste man förutsätta moralisk realism!

Vore verkligen påståendet att en person befann sig i en tragisk valsituation inkonsistent, måste man uppge tanken att tragiska valsituationer är möjliga. Konjunktionen av påståenden som beskriver en tragisk valsituation får inte vara logiskt inkonsistent. Däremot är normerna i situationen handlingsinkonsistenta, dvs de kan inte samtidigt uppfyllas. Det är just det som gör situationen till ett tragiskt moraliskt dilemma.

## 6. Den grundläggande skillnaden

Nej, skillnaden mellan dem som förnekar och dem som förfäktar möjligheten av olösliga moraliska konflikter, ligger inte i att den

förra gruppen skulle bestå av moraliska realister som hävdar att moraliska normer och värderingar är sanna eller falska, oberoende av attityder och trosföreställningar hos dem som faller omdömet, medan den senare gruppen skulle förneka det. Sinnott-Armstrong t ex hävdar att moraliska dilemman i hans mening mycket väl kan vara moraliska dilemman i realistisk mening. Jag vill t o m gå ett steg längre och hävda att det finns skäl som talar för att det är troligare att det finns tragiska valsituationer om vissa former av moralisk realism är sanna än om alla former av moralisk realism är falska.

Först och främst tycks det, som jag redan antytt, som om själva idén förutsätter moralisk realism. Frågan formuleras ibland så här: det tycks stundom som om vi befann oss i ett moraliskt dilemma. Frågan är nu, är detta dilemma (i likhet med alla andra) bara skenbart eller kan det verkligen finnas olösliga moraliska konflikter? Den frågan kan ges en klar innebörd om vi förutsätter moralisk realism, medan det blir mycket oklarare hur frågan skall förstås, om vi *inte* antar någon form av moralisk realism.

Den grundläggande skillnaden mellan dem som omfattar och dem som förnekar möjligheten av tragiska valsituationer är kanske den att de senare ser moralen som ett system i den meningen att det finns en bestämd funktion från de moraliskt relevanta omständigheterna i varje situation till en norm som säger vilken (vilka) handlingar som bör utföras. *Ett exempel på ett system är t ex en beslutsteori.* T o m en beslutsteori av Isaac Levis typ (Levi 1986), som erkänner värdekonflikter och moraliska dilemman, försöker ange vägar att lösa sådana problem genom att ange hur man skall gå tillväga för att fatta rationella beslut också i sådana situationer. Men uppenbarligen kan en beslutsteori inte ge till resultat att man absolut och ovillkorligen måste utföra en viss handling och också absolut och ovillkorligen underlåta samma handling.

Ett moraliskt *system* – som skall ge vägledning i valsituationer – bör vara konsistent, både logiskt konsistent och handlingskonsistent. Det måste därför var konstruerat så att tragiska valsituationer inte kan uppstå. Flera alternativ i en situation kan vara lika bra, så att det är hugget som stucket vilket alternativ man väljer, men det får inte kunna uppstå olösliga moraliska dilemman, ty i så fall pekar systemet i vissa situationer inte ut vilken handling som bör utföras, det lämnar agenten i sticket. Ett sådant system kan antingen innehålla en enda grundläggande norm, t ex utilitarismens: välj alltid den handling som ger det största överskottet av gott över

ont – eller det kan innehålla flera olika normer med prioriteringsnormer. Det är rimligt att se föreslagna deontisk logiska system som *desiderata* hos ett sådant system, vilket också förklarar svårigheten att konstruera en etiskt neutral deontisk logik av någon kraft. (Jfr Andréén 1970 och Sayre-McCord 1986)

Försvare av tragiska valsituationer, som E J Lemmon och Thomas Nagel uppfattar i allmänhet moralen på ett annat sätt. De ser moraliska krav som konkreta krav, som uppstår ur situationer, roller, relationer. Är jag far till ett litet barn, ställer detta barn direkt och uppenbart krav på mig. Det krävs ingen speciell kunskapsteoretisk förmåga för att identifiera ett sådant krav. (Givetvis är andras medvetanden ett problem både kunskapsteoretiskt och ontologiskt; men det krävs ingen speciell *moralisk* observation för att fastställa att barnet kräver något av mig.) Har jag rollen av far (socialt) ställs det krav på mig. Andra människor och andra roller ställer mig inför andra krav. Själv åtar jag mig att göra saker och ting: jag lovar att vattna grannens blommor och ge hennes katt mat. Nu finns det en rimlig förväntan hos grannen att katten är mätt och blommorna frodas, när hon kommer hem. På detta sätt uppstår krav: jag bör ge mitt barn mat, och trösta honom när han haft en mardröm; blommorna skall vattnas och katten få mat av mig. En del av dessa krav utgör helt enkelt moraliska krav på mig.

Det är uppenbart att sådana krav kan komma i konflikt. Och är moralen identisk med dessa krav, finns det inga garantier för att alla konflikter kan lösas. Det är svårt att se att någon naturlig prioriterings-ordning skulle *finnas* som gjorde alla mina plikter som far viktigare än *alla* mina plikter som son (äkta man, granne, arbetare, medborgare, osv).

Man skulle kunna tro att det är här den grundläggande skillnaden mellan anhängare och förnekare av möjligheten av moraliska dilemman ligger. Förnekarna vill hitta ett mönster i den kaotiska väv av plikter och krav vi ser oss invävda i, antingen genom att se alla dessa krav och plikter som uttryck för ett (eller ett fåtal) mera grundläggande krav – att maximera överskottet av lust över olust, t ex – eller genom att hitta prioriteringsprinciper, som möjliggör en entydig rangordning av alla normer och krav. De som tror på möjligheten av tragiska val, förnekar att det det finns någon given struktur i våra plikter och åtaganden. En del är givetvis viktigare än andra, men olika människor kan ge olika plikter olika vikt; en del krav har helt enkelt inte i sig själva någon inbyggd rangordning.

Men jag tror att den avgörande skillnaden ändå är en annan.



Också inom ramen för ett till synes enkelt normativt system kan normkonflikter uppträda. T o m om systemet bara innehåller en enda norm, t ex "Du skall aldrig bryta ett givet löfte," kan olösliga normkonflikter uppkomma, som Ruth Barcan Marcus påpekat (Marcus 1980 omtryckt i Gowans 1987). Det är den grupp av tragiska konflikter som innebär att två symmetriska krav ställs på oss, som är de mest intressanta, de där två två förpliktelse av samma slag och styrka står mot varandra.

Förutsättningen för att tragiska valsituationer skall vara möjliga är att riktighet och moralisk nödvändighet är egenskaper som tillkommer handlingar utan att dessa jämförs med sina alternativ, deras riktighet eller moraliska nödvändighet är inte en funktion av alternativens normativa status. Är det så att vissa typer av handlingar är absolut förbjudna, eller påbjudna, oavsett vilka normativa egenskaper alternativen har, då är tragiska valsituationer möjliga. Då kan en handling vara absolut påbjuden och dess alternativ likaså. □

### Referenser

Andrén G, "Deontisk logik som normativ etik", *Ditt & Datt, festskrift till Harald Ofstad*, Stockholm 1970

Conee E, "Why Moral Dilemmas Are Impossible", *American Philosophical Quarterly*, Vol 26, No 2, 1989

Gowans, C W (ed), *Moral Dilemmas*, Oxford University Press 1987 (Innehåller uppsatser av bl a E J Lemmon, B Williams, B van Fraassen, T C MacConnell, T Nagel, E Conee, P Foot, R M Hare)

Levi, I, *Hard Choices*, Cambridge University Press, 1986.

Sayre-McCord G, "Deontic Logic and The Priority of Moral Theory", *Nous*, XX, No 2, 1986

Sinnott-Armstrong W, *Moral Dilemmas*, Basil Blackwell 1988

Ytterligare hänvisningar finns i de nämnda arbetena.

Lars Bergström

## *Moraliska observationer*

Empiriska vetenskaper kan nog ge oss en hel del kunskap, men de kan inte lösa alla våra problem. Bland annat kan de inte lösa våra grundläggande moraliska och politiska problem. Detta är åtminstone en mycket utbredd uppfattning. Empiriska vetenskaper kan lära oss något om vilka medel vi kan använda för att nå våra mål. De kan kanske också säga något om vilka moraliska mål som överhuvud taget är möjliga att förverkliga. Men de kan inte – tänker man sig normalt – fastställa vilka mål som är riktiga eller önskvärda eller som *bör* eftersträvas. Skälet till detta är i sista hand att moraliska teorier och moraliska principer inte kan uppfylla empirismens krav. De kan inte grundas på empiriska observationer.

Men denna syn på saken – som ju är rätt allmänt accepterad – kan kanske ifrågasättas. En som ifrågasätter den är Torbjörn Tännsjö i sin bok *Moral Realism* (Rowman & Littlefield, Savage, Maryland, 1990). Tännsjö menar att moraliska teorier kan grundas på observationer. Man kan fråga sig om han också menar att dessa observationer är empiriska (i den meningen att vi gör dem med hjälp av något av våra fem sinnen), eller om han menar att de sker med hjälp av något sjätte eller sjunde sinne. Jag tror han menar att det är fråga om observationer (inklusive introspektiva iakttagelser) av samma slag som i vanliga empiriska vetenskaper, även om det man observerar är moraliska fakta. Några andra sinnen än våra vanliga fem förutsätts inte. I så fall tycks hans tes kunna formuleras så, att moralen är en empirisk vetenskap. Det låter ju rätt uppseendeväckande. Kan han verkligen ha rätt?

### **1. Observationer.**

Till att börja med bör vi ha klart för oss vad Tännsjö menar med att någonting är en "observation". Någon klar definition av denna centrala term ger han inte, men han säger att han skall använda den på ett rätt tekniskt ("slightly technical") sätt, och han diskuterar en rad krav som observationer har ansetts böra uppfylla och som han

delvis ansluter sig till (s 54-7, 64, 71-5). Jag har fått intrycket att han menar att observationer definitionsmässigt uppfyller följande krav.

(1) De implicerar *tro*, dvs om jag observerar att p – där "p" står för något sakförhållande, t ex att lampan framför mig är tänd – så tror jag också att p. (2) De implicerar *sanning*, dvs om någon observerar att p, så är det sant att p. (3) De implicerar dessutom *rättfärdigad tro*, dvs om jag observerar att p, så är därmed min tro att p kunskapsteoretiskt berättigad. De tre villkoren (1) - (3) innebär alltså tillsammans att om någon observerar att p, så *vet* vederbörande att p.

(4) Kanske kan man säga att en observation är en sorts *omedelbar händelse*, en händelse som består i att en person börjar tro att något förhåller sig på ett visst sätt, utan att denna tro är resultatet av en slutledning eller ett resonemang. (5) Dessutom har observationer *predikativ form*, dvs de innebär att man tillskriver en viss egenskap till ett individuellt föremål. Härav tycks följa att man i Tännsjös bemärkelse inte kan "observera" att en viss relation råder mellan två föremål, t ex att det ena är större än det andra. Men så bokstavligen skall han nog inte tolkas. (6) Observationer handlar vidare om föremål eller objekt som är *konkreta* i den meningen att de existerar på en viss plats i rum och tid. Sålunda kan man t ex inte "observera" att talet tre är ett primtal, ty talet tre är inte ett konkret föremål. (7) Observationer är också *orsakade* av verkliga förhållanden. Tännsjö tycks mena att om jag observerar att p, så är det faktum att p orsaken till att jag tror att p. (8) Han kräver också att om jag observerar att p, så utgör det faktum att p (en del av) den *bästa förklaringen* till att jag då tror att p. (9) Slutligen kräver han att observationer skall vara *stabila* ("indefeasible") i ungefär den meningen att de inte skulle kunna undermineras av ny information.

Ibland uttrycker sig Tännsjö som om en observation skulle vara en sorts respons som orsakas av yttre stimuli (s 54-5). Men det menar han nog inte, ty han anser att även introspektion är (eller kan vara) en variant av observation (s 77). Å andra sidan hävdar han också att introspektion sker med hjälp av våra vanliga sinnen (s 68). Enligt min mening är detta felaktigt. Om jag med hjälp av introspektion lägger märke till, och alltså får den åsikten (Å), att jag har en viss upplevelse (U), så antar jag visserligen att U är en orsak till Å, men jag kan inte föreställa mig att detta orsakssamband går via något av mina fem sinnen (syn, hörsel, känsel, lukt eller smak).

## 2. Moraliska observationer.

Nu är alltså frågan: kan vi göra några observationer i Tännsjös mening av moraliska förhållanden och därmed skaffa oss en fast grund för en moralisk vetenskap, en vetenskap om vad som är rätt och orätt, gott och ont, osv?

Tännsjö anser att vi kan det. Men han menar att de enda moraliska observationer som kan göras gäller det intrinsikala värdet hos upplevelser. Upplevelser skall då uppfattas som konkreta föremål eller händelser, som har ett bestämt läge i rum och tid. Jag tror han menar att en upplevelse är en del av en persons kropp, ungefär som ett finger eller en blindtarm, under en viss tidsperiod.

Det kan förefalla egendomligt att på detta sätt uppfatta upplevelser som konkreta föremål. (För egen del vill jag inte göra det. Jag tänker mig – grovt uttryckt – hellre att upplevelser är *egenskaper* hos kroppar, än att de skulle vara *delar* av kroppar. Att jag t ex upplever glädje är detsamma som att jag är glad, dvs att jag har en viss egenskap, nämligen att vara glad. Det betyder inte att en viss bit av min kropp är en glädjeupplevelse! Vidare menar jag att det som här har värde är det *faktum* att jag är glad, inte en viss *kroppsdel*).

Men jag tror att det är en omistlig del av Tännsjös program att uppfatta upplevelser som konkreta föremål. Ty det ligger nära till hands att hävda att om man observerar att X har en viss egenskap, så måste X vara ett konkret föremål. Skälet är att det faktum att X har en viss egenskap inte gärna kan vara en orsak till att jag observerar detta om inte X är ett föremål med bestämt läge i tid och rum. Orsaker skall ju ligga nära sina verkningar i rum och tid. En förutsättning för att man skall kunna observera att en upplevelse har en viss egenskap är alltså att denna upplevelse är ett konkret föremål – jfr även villkoren (5) och (6) ovan.

Låt oss nu begrunda ett exempel på en "moralisk observation" av det slag som Tännsjö har i tankarna. Antag att Tännsjö blev glad när han för första gången fick se sin nyfödda yngsta dotter. Vi kan kanske – med en viss tvekan – uttrycka det så, att en lustupplevelse då befann sig i hans kropp. Ungefär samtidigt konstaterade Tännsjö tyst för sig själv att denna upplevelse var intrinsikalt god, dvs värdefull för sin egen skull. Och nu vill han alltså försöka få oss att gå med på att detta konstaterande var en observation, och att denna observation har samma kunskapsteoretiska kvalitet som observationer i vanliga empiriska vetenskaper. Vad skall man säga om det?

Till att börja med kan man lägga märke till att om det är det faktum att Tännsjö är glad, snarare är den konkreta glädjeupp-

levelsen, som har värde, så kan detta inte fastställas med observation – av de skäl som ovan antytts.

Men låt oss nu ändå (försöka) anta att Tännsjö verkligen hade en konkret glädjeupplevelse och att han observerade att denna upplevelse hade egenskapen att vara god. Det är då tveksamt om någon annan än Tännsjö själv någonsin kan veta att han gjort denna observation. Ty för att veta det måste vi ju veta åtminstone tre saker som förefaller rätt oåtkomliga.

För det första måste vi veta att Tännsjö verkligen hade en konkret upplevelse i hans mening. Jag förstår inte hur det skall gå till. Andra närvarande kan nog veta att han blev glad, men de har nog inte observerat någon konkret lustupplevelse. De kan möjligen ha observerat att Tännsjö blev glad (även om det kanske vore rimligare att påstå att det de observerade var hans ansiktsuttryck och kroppsspråk, och att de därav drog slutsatsen att han blev glad). Men denna observation kan inte tas till intäkt för existensen av den upplevelse Tännsjö talar om. Det de observerade var ju Tännsjö själv och inte en upplevelse. Tännsjö är inte en upplevelse.

För det andra måste vi veta att det Tännsjö trodde om sin upplevelse var sant, dvs att den verkligen var intrinsikalt värdefull, jfr villkor (2) ovan. Hur skall vi veta det? Om vi kunde observera det Tännsjö påstår sig observera, så vore ju saken klar, åtminstone i praktiken. Men såvitt jag förstår kan vi inte det. Tännsjö hävdar dock motsatsen (s 86). Hans argument har att göra med hans tes att det intrinsikala värdet hos upplevelser kan orsaka psykologiska förändringar, som i sin tur kan observeras (s 87). Jag skall strax återkomma till detta.

Men även om ingen utomstående kan observera att Tännsjös upplevelse är intrinsikalt värdefull, så kan det kanske ändå vara sant att den är det. Tännsjö menar att sådant kan vara sant (s 68), men han har såvitt jag kan se inget argument för detta. Han säger egentligen bara att om vi måste anta detta i samband med vissa vetenskapliga förklaringar, vilket han för sin del tror att vi måste, så är det också rimligt att göra detta antagande. Som argument betraktat är detta mer än lovligt svagt.

För det tredje måste vi veta att orsaken till Tännsjös observation (som alltså gick ut på att upplevelsen var intrinsikalt god) var just det faktum att upplevelsen var intrinsikalt god – jfr villkor (7) och kanske också (8) ovan. Här stöter vi åter på det problem som jag nyss nämnde, nämligen huruvida det faktum att något har intrinsikalt värde – låt oss kalla det ett värdefaktum, givet att det nu alls

finns sådana – kan orsaka andra händelser eller fakta som i sin tur kan observeras, antingen direkt eller indirekt. För enkelhetens skull kan vi här begränsa oss till psykologiska fakta. Jag tror att detta är själva kärnpunkten: *kan värdefakta orsaka psykologiska fakta?* Tännsjö svarar ja på denna fråga, under det att de flesta andra filosofer nog skulle svara nej. Huvudtesen i Tännsjös bok tycks stå och falla med svaret på denna fråga. Den "moraliska realism" han förespråkar bygger på att vi kan observera värdefakta, vilket i sin tur förutsätter att värdefakta kan orsaka psykologiska fakta.

### 3. Värde och kausalitet.

Vad skall man då säga om detta problem? Kan värdefakta orsaka psykologiska fakta? Tännsjös argument ser ut ungefär så här: Den bästa förklaringen till att en person vill ha en upplevelse för dess egen skull är att denna upplevelse är intrinsikalt god. Det senare, som är ett värdefaktum, är alltså orsaken till det förra, som är ett psykologiskt faktum! (s 83, 87).

Men jag kan inte se att detta är en särskilt bra förklaring. Den tycks förutsätta att folk är konstruerade på ett sådant sätt att de eftersträvar det goda. Men vad har vi för anledning att tro det? Det är möjligt att personer som Tännsjö genom introspektion kan konstatera att de har en tendens att eftersträva det som de *tror* är bra. Med andra ord: de *tror* sig kunna konstatera att de eftersträvar det som är bra. Men vad visar det? Det visar väl inte ens att det föreligger en hög korrelation mellan att något eftersträvas och att det är bra, och ännu mindre visar det att det föreligger ett orsakssamband här.

För att det skall vara rimligt att tro att värdefakta kan orsaka psykologiska fakta, så bör man kanske ha åtminstone någon uppfattning om den kausala mekanism som då är inblandad. Detta har ofta påpekats, och Tännsjö tar också upp denna fråga. Han tycks gå med på att vi inte vet något om den kausala mekanism som förbinder oss med värdefakta, men han menar att vi vet lika lite om de kausala mekanismer som är inblandade i vanlig varseblivning (s 85).

Detta är inte övertygande. När jag tittar på den tända lampan framför mig, så uppstår i mitt medvetande den uppfattningen att lampan framför mig är tänd. De flesta av oss skulle väl gå med på att det faktum att lampan är tänd (T) är orsaken till att jag tror att den är tänd (T\*). Det finns mycket vi inte vet om orsaksprocessen mellan T och T\*. Men något vet vi ändå. Vi vet en del om hur lampan reflekterar och utsänder ljusstrålar, hur ljuset träffar mina ögon, hur synnerven aktiveras, och hur detta ger upphov till

processer i min hjärna. Vad vi inte vet är framför allt hur hjärnprocesserna ger upphov till T\*. Men vi är övertygade om att T\* på något sätt är kopplat till dessa hjärnprocesser. Detta stöds bl a av att vi kan framkalla olika synintryck genom att påverka hjärnprocesserna.

Däremot har vi ingen som helst aning om hur det faktum att Tännsjös lustupplevelse är värdefull (V) skulle orsaka hans uppfattning att den är värdefull (V\*). Vi vet inget om V:s kausala förmåga överhuvud taget, och därmed inte heller något om hur V skulle kunna framkalla några speciella hjärnprocesser hos Tännsjö. Att V\* är förknippat med vissa hjärnprocesser tror vi däremot. Om vi skulle förstöra Tännsjös hjärna, så skulle han t ex inte tänka alls.

Tännsjö förnekar emellertid detta. Han menar att vi inte vet hur T orsakar hjärnprocesser, eller att vi endast vet mycket lite om detta, och att vi vet lika lite om detta som vi vet om hur V orsakar hjärnprocesser. Men här tror jag att han helt enkelt har fel. Han hänvisar till att den bästa vetenskapliga förklaringen av hjärnprocesser inte innehåller några vardagliga uttryck av typ "lampa" och "tänd" (s 85), men jag kan inte se att det har ett dugg med saken att göra. (Det är väl klart att orsaken till att min lampa är tänd är att jag vred på strömbrytaren, trots att den bästa fysikaliska teori som förklarar sambandet inte innehåller termer som "vrida" och "strömbrytare".)

Kort sagt: Tännsjö har inte givit några goda skäl för sin huvudtes, nämligen att vi kan observera värdefakta, och själva tesen verkar dessutom felaktig.

#### **4. Mer om introspektion.**

Emellertid skulle man kanske kunna utnyttja Tännsjös förarbete och ta ytterligare ett steg i argumentationen och därmed trots allt komma i närheten av hans huvudtes. Ty Tännsjö menar ju att värdeobservation ofta (men inte alltid) är introspektion – och i fråga om introspektion vet vi väl egentligen inte heller någonting om hur kausalmekanismerna ser ut!

Låt oss återgå till Tännsjös upplevelse av glädje, eller närmare bestämt till det faktum (som vi kan kalla G) att han blev glad när han fick se sin dotter. Detta faktum skall inte förväxlas med det faktum (som vi kan kalla G\*) att han introspektivt observerade att han var glad. Många skulle kanske här säga att G är orsak till G\*. Men vi vet knappast något om hur G orsakar G\*. Är det inte då för mycket begärt att Tännsjö skall kunna säga något om hur V orsakar V\*?

Jo, det är nog för mycket begärt. Men dessa två fall är inte riktigt analoga. För det första är det väl inte fråga om introspektion i bägge fallen. Med "introspektion" menar man normalt en sorts observation av, eller åtminstone tänkande på, *mentala* tillstånd hos en själv. I denna mening kan G\* sägas vara ett introspektivt tillstånd eftersom det handlar om G som är ett mentalt tillstånd. Men V\* är inte i denna mening ett introspektivt tillstånd, ty det handlar om V som *inte* är ett mentalt tillstånd, utan ett värdefaktum. Att uppfatta V som ett mentalt tillstånd skulle vara en form av naturalism, som Tännsjö knappast skulle acceptera (jfr s 30-1).

För det andra kunde man kanske hävda att orsakprocessen från G till G\* går via processer i hjärnan som vi åtminstone vet någonting om (eftersom de är fysikaliska och vi vet en del om fysikaliska processer). Men om det finns en orsakprocess från V till V\* så måste den åtminstone delvis vara av annat slag, just därför att V varken är ett mentalt eller ett fysiskt tillstånd och därför inte kan identifieras med eller direkt motsvaras av något tillstånd hos hjärnan. Hur en sådan orsakprocess skulle se ut är såvitt jag vet ett fullständigt mysterium på vetenskapens nuvarande standpunkt (och det är ett *annat* mysterium än det som gäller förhållandet mellan mentala tillstånd och tillstånd hos hjärnan).

Dessutom kan man ifrågasätta om introspektion verkligen innefattar ett orsakssamband. Det är inte självklart att introspektion uppfyller Tännsjö's villkor (7). Om det inte gör det så är det inte alls observation i hans mening. Introspektion är kanske snarare en sorts teoribildning. Och introspektiva teorier (dvs introspektioner) underdetermineras kanske av de eventuella data som de skulle kunna förklara (om det nu finns några sådana).

Att det skulle råda ett orsakssamband mellan V och V\* är kanske ännu orimligare – inte minst därför att vi inte tycks ha något sinnesorgan som kan förmedla eventuella stimuli från värdefakta till processer i våra hjärnor (som sedan i sin tur på något okänt sätt kan framkalla värderingar i vårt medvetande). Dessutom strider det väl faktiskt mot våra mest väletablerade teorier om verkligheten att värdefakta som V skulle kunna orsaka några fysikaliska förändringar överhuvud taget. Redan av detta skäl verkar det uteslutet att man skulle kunna observera värdefakta.

## 5. Tännsjö's projekt.

Men även om Tännsjö skulle ha rätt i att vi kan observera värdefakta, så kan jag inte se att det skulle räcka särskilt långt som



grundval för en moralisk teori. Det skulle knappast räcka för att lösa några allvarliga moraliska problem.

För det första kvarstår frågan om hur man skall kunna veta att en upplevelse är *bättre* än en annan? Enligt varje rimlig moralisk teori är detta av avgörande betydelse, men sådant kan man inte observera enligt Tännsjö, jfr villkor (5) ovan. Kanske har jag missförstått honom på denna punkt. Men det klassiska problemet om "interpersonella jämförelser" (dvs hur man skall jämföra värdet av upplevelser hos *olika* personer) kvarstår i alla händelser. Tännsjö menar kanske att detta problem helt enkelt försvinner i och med att man kan observera värdet av andras upplevelser. Men ett sådant påstående skulle knappast övertyga den som oroas av problemet om interpersonella jämförelser.

För det andra kan man fråga sig hur man skall kunna veta att *endast* upplevelser kan ha intrinsikalt värde. Många skulle kanske föreställa sig att även t ex rättvisa har intrinsikalt värde? Sådant kan enligt Tännsjö inte observeras, men av detta kan man väl inte utan vidare dra slutsatsen att det är falskt.

För det tredje säger Tännsjöns teori egentligen ingenting om hur man *bör handla*. Det stör i och för sig inte honom, ty han menar att det inte finns några genuina normativa problem (s 91). På denna punkt har han i stort sett samma åsikt som de logiska positivisterna. Däremot anser han att vi bör använda normativa ord som "bör", "rätt" och "orätt" för att uttrycka hypoteser om det intrinsikala värdet av de totala konsekvenserna av givna handlingar jämfört med deras alternativ. Han menar t ex (ungefär som G E Moore 1903) att påståendet att en viss handling bör utföras bör tolkas till att betyda att universum som helhet skulle bli intrinsikalt bättre om handlingen utförs än om den inte utförs (s 95). När Tännsjö säger att vi bör använda "bör" på detta sätt får han väl antas mena att universum som helhet blir intrinsikalt bättre om vi gör det än om vi låter bli. Såvitt jag kan förstå är detta helt gripet ur luften. Även om vi skulle anta att Tännsjö har rätt i att man kan observera vissa värdefakta, så är det minst sagt tveksamt om man ens i princip testa en sådan hypotes.

**Not:** Jag vill tacka Hans Mathlein och Torbjörn Tännsjö för värdefulla synpunkter på en tidigare version av ovanstående.

**Torbjörn Tännsjö**

## ***Visst kan vi observera det goda!***

Lars Bergström betvivlar min tes (framförd i kapitel tre av min bok *Moral Realism*) att vi kan observera intrinsikalt värde och han för fram argument mot den. Bergströms invändningar har inte fått mig att ändra uppfattning.

Utgångspunkten för mitt resonemang är ett argument mot moralisk realism framfört av Gilbert Harman i första kapitlet av boken *Introduction to Ethics*. Enkelt uttryckt består Harmans argument i ett kriterium på när det är rationellt att anta att något har objektiv existens och ett påstående om att moraliska "fakta" inte uppfyller kriteriet. Kriteriet går ut på att vi ska anta att något har objektiv existens om och endast om vi (på ett essentiellt sätt) hänvisar till det, direkt eller indirekt, i våra bästa förklaringar av någon observation. Flera författare har hävdat att detta kriterium är orimligt strängt. Moralisk realism är en rimlig hållning *trots* att moraliska fakta inte uppfyller kriteriet. Det är t ex Thomas Nagels uppfattning. Själv har jag emellertid valt en "tuffare" strategi. Jag har försökt visa att (vissa) moraliska fakta uppfyller Harmans kriterium. Och de uppfyller kriteriet på ett mycket direkt sätt, har jag hävdat. Det är just vissa moraliska observationer (av intrinsikalt värde) som bäst förklaras av moraliska fakta. Kl 5 i morse observerade jag t ex att den upplevelse av intensivt lidande jag är utsatt för när min dotter Marja väckte mig genom att hoppa på min mage var intrinsikalt dålig. Den observationen förklaras bäst, tror jag, av att denna upplevelse *var* intrinsikalt dålig. Kanske var det också det faktum att upplevelsen var en upplevelse av lidande, som orsakade att den var dålig? Det är en moralisk hypotes jag vill tillmäta ett visst mått av rimlighet. Det är till den jag vänder mig då jag i ett sökande efter moralisk förklaring vill göra en "inference to the best explanation" av vad jag erfarit.

Detta tycker Bergström är en uppseendeväckande tes och han ifrågasätter dess rimlighet. Till detta gör jag kommentaren att den nog var uppseendeväckande, då jag först framförde den (i boken

*Betrakta det goda*, 1979), men att den inte är så uppseendeväckande idag. Liknande åsikter har sedan dess framförts t ex av Robert Nozick i *Philosophical Explanations* och utförligt försvarats av Nicholas Sturgeon. Jag tror också att min tes är rimlig.

## 1. Ontologi

Bergström noterar att han och jag är oense i vissa ontologiska frågor. Han bygger mycket av sitt resonemang på dessa skillnader och hävdar att min moraliska realism, står och faller med min (orimliga) ontologi.

Bergström tror inte att upplevelser är konkreta föremål utan egenskaper hos kroppar. Han tror att det är en omistlig del av mitt "program" att uppfatta upplevelser som konkreta föremål. Det är ett missförstånd. Jag omfattar visserligen sedan filosofiska barnben en ontologi där *händelser* är en grundläggande kategori (jag tror jag fick det från Bertrand Russell), men jag förutsätter den inte i mitt resonemang. Allt jag vill ha sagt om moralisk realism kan jag uttrycka också i Bergströms ontologi (där *kroppar* tycks ha fundamental status). Om olusten jag känner när Marja väcker mig klockan 5 är en egenskap hos *mig* så gäller, skulle jag säga, detsamma om det intrinsikala värdet. Det är sant om mig både att jag lider och att jag uppvisar eller är bärare av negativt intrinsikalt värde (det "är illa ställt" med mig, kunde vi kanske säga). Både lidandet och dåligheten är i så fall egenskaper hos mig. Såväl lidandet som dåligheten observerar jag, introspektivt. Och en hypotes jag har är att det är därför att jag lider som det är illa ställt med mig. Eftersom både lidandet och dåligheten ger upphov till olika psykiska reaktioner hos mig, så kanske man också kan tala om att de i någon mening är observerbara för andra (tränade iakttagare). Låt mig emellertid för korthets skull hålla mig till det mest fundamentala, introspektiva fallet. Låt mig också i förbigående notera att jag är nöjd med Bergströms referat när det gäller min syn på vad en observation är för något – med kommentaren att jag naturligtvis, som han förmodar, godtar att det predikat man tillskriver individuella föremål i observationsutsagor kan vara flerställt.

## 2. Värdeobservation

Vad min tes går ut på är alltså, uttryckt i Bergströms ontologiska vokabulär, att dåligheten hos mig (det negativa intrinsikala värde som realiserar hos mig just nu, då en fullkomligt omutlig Marja

väcker mig kl 5 på morgonen genom att hoppa på min mage) är något jag introspektivt observerar. Det är själva dåligheten som orsakar min observation av dess dålighet. Den bästa förklaringen till att jag gör min observation är att det verkligen *är* illa ställt med mig. Varför skulle den ståndpunkten vara så orimlig?

Bergström går med på att det faktum att Marja hoppar på min mage orsakar (på det sätt som är relevant i sammanhanget) min observation att hon gör det. Han går också (med tvekan?) med på att obehaget jag känner orsakar mitt introspektiva omdöme om att jag känner obehag. Men han vill inte gå med på min tes att det dåliga i min upplevelse orsakar mitt omdöme om att den är dålig. Varför inte?

Tesen syns mig plausibel, den stämmer överens med olika ting vi brukar hålla för sanna om moralisk argumentation och den rimmar väl med mina vardagliga erfarenheter. Och det finns, menar jag, inga riktigt starka argument *mot* den. Detta ifrågasätter Bergström.

Harman har hävdad att det finns ett starkt argument mot tanken om moralisk observation. Bergström gör det till sitt. Då jag observerar att Marja hoppar på min mage har vi åtminstone ett hum om hur hennes hoppande orsakar min uppfattning att hon hoppar. Då jag observerar att något har intrinsikalt värde har vi ingen aning om hur detta skulle kunna orsaka något som helst omdöme hos mig. I det förra fallet är vi vissa kausala mekanismer på spåren, i det senare fallet är vår okunnighet total.

Mot detta har jag invänt att vi i *båda* fallen är okunnigare än vi vanligtvis är benägna att tro. När det gäller kausala mekanismer (så som vi känner dem i våra mest avancerade – fysikaliska – teorier) görs inga hänvisningar till hoppande treåringar. Och, viktigare, även om vi grovt kunde göra reda för något sådant som att en treåring hoppar på min mage i den mest avancerade fysikens termer, och säga något om hur hennes hoppande reflekterar ljustrålar, som träffar mitt öga och aktiverar min synnerv o s v, (för att inte tala om hur hennes fötter aktiverar mina känsloorgan i magtrakten), så står vi oss ändå slätt då vi ska förklara "andra ändan" av den kausala kedjan. Hur producerar dessa fysikaliska händelse ett visst *omdöme* hos mig? Dessutom: vid den mest grundläggande formen av moralisk observation, som jag tänker mig saken, är det vi observerar inte något yttre, fysikaliskt föremål, utan ett inre tillstånd (eller, i Bergströms ontologi, att vi *är* i ett visst mentalt tillstånd). Därmed ligger också de första länkarna i den kausala kedjan, ånyo, i dunkel.

Bergström medger att vi vet föga om den sista länken i den kausala kedjan då vi observerar fysikaliska föremål och föga om båda ändarna av den då vi gör introspektiva omdömen, men han menar att vi ändå vet mer i båda dessa fall än då vi faller värdeomdömen. Varför? Hans argument är att mentala tillstånd är "förknippade" med hjärnprocesser (om min hjärna utplånas så kan jag inte tänka). Något liknande skulle inte gälla för intrinsikalt värde. Här tror jag han misstar sig.

Utplånar man min hjärna kan jag inte tänka mer men samtidigt, det känner jag mig förvissad om, så kommer jag då inte heller att ha några dåliga upplevelser – intet ont kommer då att vederfaras mig (och det vore ju en tröst i eländet). Jag tror vi har ungefär samma (vaga) hum om hur intrinsikalt värde hänger samman med hjärnprocesser som vi har om hur lidande gör det.

Bergström medger att det är något av ett mysterium på vetenskapens nuvarande ståndpunkt *hur* fysiska tillstånd kan producera mentala tillstånd; vi tror ändå att sådana orsakssamband föreligger. Det är emellertid ett annat (och värre?) mysterium, hur värdetillstånd skulle kunna producera mentala tillstånd. Varför? Jo, skriver Bergström, eftersom värdeförhållanden (enligt mig) varken är fysikaliska eller mentala.

Detta är ett missförstånd. Jag har ingen bestämd uppfattning på den punkten. Kanske är mentala tillstånd fysikaliska? Kanske är värdetillstånd mentala? Förvisso vore detta i så fall exempel på en form av naturalism. Denna naturalism har jag emellertid varken tagit ställning för eller emot! Då jag diskuterat den (i kapitel 2 av *Moral Realism*) har jag nöjt mig med att konstatera att sådana teorier om identitet mellan egenskaper *förutsätter* moralisk observation. Skulle en sådan form av naturalism vara korrekt så är det inget mysterium alls *att* fysikaliska tillstånd producerar mentala tillstånd som producerar värdetillstånd – alla dessa tillstånd är i så fall fysikaliska. *Hur* går det till är vi emellertid, på vetenskapens nuvarande ståndpunkt, okunniga om. Jag är i själva verket, liksom f ö Sturgeon, benägen att acceptera en sådan naturalism.

### 3. Introspektion

Den mest grundläggande formen av moralisk observation är alltså, om jag har rätt, av introspektiv karaktär. Kan verkligen introspektion godtas som en form av observation?

Åsikterna om introspektion har historiskt sett gått isär. Några har hävdats att introspektiva uppfattningar är fullkomligt säkra. Berg-

ström tycks luta åt en annan extrem, att introspektion är en ren konstruktion, eller åtminstone en sorts teoribildning, som underdetermineras av eventuella data. Han ifrågasätter också att vi når introspektiv kunskap via våra fem sinnen.

Kanske har jag uttryckt mig oklart på den sista punkten. Jag har argumenterat mot tanken att introspektion förutsätter att vi vid sidan av våra fem sinnen har något extra, "inre öga", med vars hjälp vi gör introspektiva omdömen. Med det har jag inte velat hävda att vi inte har något sådant "inre öga", även om jag starkt betvivlar existensen av det. Jag har inte heller velat hävda att vi kan göra introspektiva omdömen uteslutande med stöd av något visst av våra fem sinnen. Hela idén om *fem* sinnen är mycket primitiv. Information "processas" på olika mycket komplicerade sätt i vårt nervsystem. Jag tror att det är samma typ av processer som föreligger då vi gör introspektiva omdömen som då vi gör perceptuella omdömen, även om i olika fall olika delar av nervsystemet är aktiva. Det var allt jag ville ha sagt om introspektion och våra sinnen.

I all perception finns säkerligen ett teoretiskt moment. Jag tror inte heller att introspektiva omdömen är ofelbara eller sådana att vi aldrig kan ha anledning att ompröva dem. Vi kan komma att vilja revidera ett introspektivt omdöme i ljuset av ytterligare information. Men jag tror att de är ungefär lika stabila mot revision som vissa andra (observations-)omdömen. Omdömet att jag kände obehag (då Marja hoppade på min mage) i morse var ungefär lika "centralt" och "immunt" mot revision, som omdömet att någon hoppade på mig, och, för den delen, som omdömet att det jag då kände var intrinsikalt dåligt. *Alla* dessa omdömen var för övrigt säkrare, mer immuna mot revision, än att det var just *Marja* som hoppade på mig.

#### 4. Mitt projekt

Jag har en gång lärt mig genom att läsa Lenin att ett enda steg av verklig rörelse är viktigare än hundra program och vidare att man ska akta sig för "projektmakeri". Det är därför med visst bekymmer jag konstaterar att jag enligt Bergström håller mig både med ett filosofiskt program (i vilket ingår tron på existensen av händelser) och med ett projekt. Men kanske har han rätt i det. Programmet har jag redan kommenterat. I ett avslutande avsnitt kritiserar Bergström mitt projekt. Det är inte storstatligt nog, enligt Bergström. Även om jag skulle ha rätt i att vi kan observera intrinsikalt värde, så räcker det inte långt då det gäller att lösa "allvarliga moraliska

problem". Åtminstone tre problem återstår. (1) Hur ska vi kunna göra interpersonella värdejämförelser? (2) Hur kan vi veta att inte annat än det vi kan observera som värdefullt har värde? (3) Hur kommer vi från en teori om vad som är värdefullt till *hur vi bör handla*?

När det gäller den tredje frågan klagar Bergström på att jag i stort sett är av samma åsikt som de logiska positivisterna. Hans karaktäristik är riktig, men jag vantrivs inte med sällskapet. Jag tror verkligen inte att det finns några viktiga *moraliska* problem kvar att lösa, om vi i en situation vet vilket bland olika alternativ öppna för oss, som realiserar mest intrinsikalt värde. (Däremot finns det givetvis ett praktiskt problem kvar – något måste göras!). Själv är jag som Bergström noterar benägen att låta ordet "bör", tillämpat på konkreta handlingar, betyda detsamma som "har bättre konsekvens än varje alternativ". När jag hävdar att också andra *bör* använda ordet "bör" på samma sätt och alltså implicerar att universum som helhet blir intrinsikalt bättre om det följer mitt råd, så är det inte gripet ur luften men väl menat som ett skämt. Min pessimism i fråga om vi kan veta huruvida vi med våra handlingar gör världen som helhet bättre eller sämre överträffar Bergströms – men den diskuterar jag i andra sammanhang.

När det gäller den andra frågan, om annat än upplevelser kan ha värde, så håller jag den öppen. Jag har emellertid svårt att tro att vi, om så är fallet, någonsin kommer att kunna ha *kunskap* om dessa värden.

När det gäller den första frågan, om interpersonella värdejämförelser, så medger jag att jag inte har löst (eller ens försökt lösa) den. Men Bergström har ju själv i *Filosofisk tidskrift* (nr 3 1982) ett förslag till lösning – det ligger väl i linje med hur jag ser saken! □

**Lars Bergström**

## ***Replik till Tännsjö.***

Tännsjö menar att vår oenighet i ontologi inte spelar någon roll, att allt han vill ha sagt kan sägas i min ontologi. Men när han skall tala om att det är intrinsikalt dåligt att han lider när Marja hoppar på hans mage, att det "är illa ställt" med honom, så får han det – i min ontologi – till att dåligheten är en egenskap hos honom. Det är väl ändå orimligt. Det är ju det faktum att Tännsjö lider som är dåligt. Inte Tännsjö själv!

Frågan är om man kan observera att något är dåligt. Tännsjö menar (krav 6) att om man kan observera att X har egenskapen E, så är X ett konkret objekt. Om dålighet är en egenskap hos sakförhållanden, så kan man alltså inte observera att något är dåligt. Enligt min ontologi är (intrinsikal) dålighet en egenskap hos sakförhållanden. Jag har svårt att tro att Tännsjö kan acceptera det.

Min huvudinvändning är emellertid att värdetillstånd knappast kan orsaka mentala tillstånd eftersom värdetillstånd varken är mentala eller fysikaliska tillstånd. Jag trodde Tännsjö höll med om att värdetillstånd varken är mentala eller fysikaliska, och att han alltså förkastade den form av naturalism som hävdar motsatsen. Jag hänvisade här till s 30-1 i hans bok, där han nämner en del invändningar mot en sådan naturalism. Han tycks fästa vikt vid dessa invändningar och han nämner inga argument åt andra hållet. (På s 7 säger han också: "I will argue that moral statements cannot be reduced to empirical statements".)

Nu visar det sig emellertid att han vill hålla frågan öppen, men att han lutar åt att naturalismen är riktig. Det gör enligt min mening hans position ännu svagare. Ty antingen är naturalismen riktig, eller också är den inte riktig. Om den är riktig, så blir ju Tännsjös tes rätt trivial. Att man kan observera eller introspektera fysikaliska och mentala tillstånd är ingen nyhet. Och om naturalismen inte är riktig, så har Tännsjö fortfarande inget motargument mot min invändning.

Tännsjö håller frågan öppen om annat än upplevelser kan ha värde, och han menar att vi nog aldrig kan veta något om detta.



Såvitt jag förstår innebär detta att vi nog aldrig kan veta om en given handling bör utföras eller underlåtas (i Tännsjös mening). Och det tycks vara oförenligt med moralisk realism (i Tännsjös mening), även om han skulle ha rätt i att vi kan observera vissa värdefakta. □

## Notiser

På förlaget Dualis har nyligen utkommit tre klassiska verk av Martin Buber i svensk översättning, *Människans väg*, *Skuld och medkänsla* och, den kanske mest kända, *Jag och du*. Dualis förlag har adressen Box 136, 771 01 Ludvika.

Tre nya antologier med filosofiska texter i översättning har utkommit. Per Bauhn har sammanställt en volym med titeln *Teorier om mänskliga rättigheter* (Thales 1990). Här ingår viktiga avsnitt ur böcker av Locke, Rousseau, Alan Gewirth, Ronald Dworkin och Robert Nozick. Utgivningen har möjliggjorts genom bidrag från Erik och Gurli Hultengrens fond för filosofi.

Vidare har volymerna 2 och 3 i Tidens idéserie kommit. *Idéer om välfärd* (1990) innehåller uppsatser av N Furniss, T Tilton, R M Titmuss, J Tobin och M Walzer. *Idéer om frihet* (1990) innehåller uppsatser av G MacCallum, S Benn, W Weinstein, D Miller, G A Cohen, J Reiman, I Young och E Loevinsohn. Introduktioner har författats av Sven Ove Hansson.

På RFSUs förlag Sesam har Thorsten Sjövall utgivit *Från sexualism till humanism* (1990).

För någon tid sedan pågick en debatt om Guds existens i radions "Klarspråk". Ett inlägg utmynnade i den intressanta tesen att "om jag tror på Gud, så är det *sant för mig* att Gud existerar". *Filosofisk tidskrifts* läsare inbjuds härmed att försöka hitta en tolkning av detta uttalande som gör det till något annat än en ren truism utan innehåll.

### **Svar till Lars Bergström**

I förra numret av *Filosofisk tidskrift* (sid 47-48) uppmanar Lars Bergström läsarna att syna den logiska hållbarheten i den slutledning, som jag antagit vara Hägerströms. Jag är därför angelägen att få formulera min ståndpunkt tydligare. Det är riktigt att jag i dialogen med Sven Danielsson säger att slutledningen är logiskt giltig. Det är en olycklig och slarvig formulering. Min ståndpunkt är den jag uttryckt i min avhandling, nämligen att vi får en logiskt hållbar slutledning, endast om vi förutsätter en viss tolkning av premisserna. Jag antar nu att Hägerström i sin tankegång uppfyller detta krav och därför är det kanske lätt att utan restriktion tala om giltighet hos resonemanget. Men att strukturen i sig skulle ge upphov till en logiskt giltig slutledning alldeles oberoende av

tolkningen av premisserna, har jag således inte menat, även om det kan tyckas så i min uppsats. För att göra min ståndpunkt klar vill jag därför återge den tankegång i avhandlingen som berör just detta.

När man läser Hägerströms framställning av argumentet i installationsföreläsningen, har man ingen känsla av att han skulle göra ett logiskt misstag. Därför utgår jag i mitt tolkningsförsök från att det borde gå att precisera hans resonemang så att det blir logiskt hållbart (även om det inte är sakligt riktigt). De premisser och den slutsats, som jag finner överensstämma väl med hans formuleringar är den anförda:

- 1) För att en värdering skall föreligga hos en person är det nödvändigt att en viss känsla föreligger hos honom.
- 2) För att ett omdöme skall föreligga hos en person är det inte nödvändigt att en känsla föreligger hos honom.
- 3) En värdering är inte något omdöme.

Mitt problem blir nu hur premisserna skall tolkas. Slutledningen blir nämligen inte logiskt korrekt, om man tolkar premisserna helt efter ordalydelsen. I avhandlingen framhåller jag att möjligheten, att värderingen ändå skulle kunna vara ett slags omdöme, ett som alltid åtföljs av vissa känslor, inte utesluts, och att slutsatsen 3) inte följer. Jag antar då att det är premissen 2) som orsakar problemet. Kan den ges en sådan innebörd att vi får slutsatsen att följa?

Hägerström använder sig i sina formuleringar av uttrycks sätt som att omdömet är "oberoende" av känslan, att omdömet gäller "utan hänsyn till" känslor. Kanhända kan 2) ges en "starkare" tolkning. I avhandlingen säger jag följande:

"För att Hägerströms slutledning skall vara korrekt måste 2) åtminstone utsäga, att det för varje omdöme gäller att det kan föreligga i ett psyke utan att därför samtidigt en känsla måste föreligga. Det får inte finnas någon klass av omdömen, som alltid föreligger tillsammans med känslor. I ett sådant fall kunde nämligen värderingen vara ett dylikt omdöme" (sid 118).

Om vi antar att Hägerström gör denna förutsättning – dvs att 2) skall innebära att *varje* förekomst av ett omdöme skulle kunna föreligga utan att en känsla behöver vara närvarande – så blir argumentet logiskt hållbart. Riktigheten och rimligheten i förutsättningen kan naturligtvis ifrågasättas, men den förefaller Hägerströmsk, så jag antog att Hägerström nog tänkt så. □

*Bo Petersson*

Medarbetare i detta nummer av *Filosofisk tidskrift*: *Włodzimirz Rabinowicz* är docent i praktisk filosofi i Uppsala, *Thomas Brante* är docent i sociologi i Göteborg, *Ragnar Ohlsson* är högskoleadjunkt i praktisk filosofi i Stockholm, *Lars Bergström* är professor i praktisk filosofi i Stockholm och *Torbjörn Tännsjö* är docent i praktisk filosofi i Stockholm.

Omslagsbilden är gjord av *Lisa Rossholm*.

### *Manuskript 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 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

### *Honorar*

- införda bidrag honoreras ej för närvarande

### *Särtryck*

- I stället för särtryck erhåller författaren, gratis, 10 fullständiga exemplar av det nummer av tidskriften i vilket bidraget varit infört.

**GLÖM INTE ATT PRENUMERERA  
PÅ  
FILOSOFISK TIDSKRIFT 1991  
120:- postgiro 507991-8**

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

För årg 1/1980—7/1986 är priset 25 kr/årg, lösa ex 10 kr.

För årg 8/1987—10/1989 är priset 100 kr/årg, lösa ex 30 kr.

Porto tillkommer.

**BOKFÖRLAGET THALES**

ISSN 0348-7482