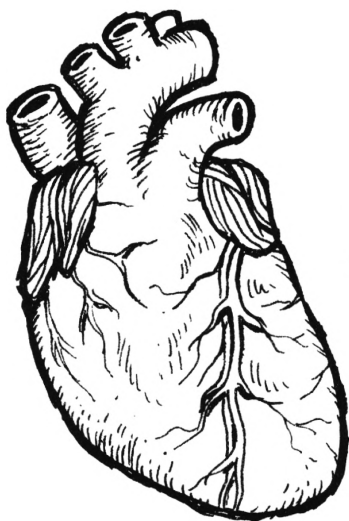

FILOSOFISK TIDSKRIFT

NR1
1987



*Mats Furberg
Ragnar Ohlsson
Ingmar Persson*

*Erik Ryding
Björn Eriksson*

Ingmar Persson

*Om dödsbegrepp och definitioner
De smutsiga händerna
Överkvalificerade handlingar ur ideal-
och realmoraliska perspektiv
Parfit och den personliga identiteten
Artisten och djuren — en filosofisk
cirkus
Tollands artistiska magplask*

FILOSOFISK TIDSKRIFT NR 1 • 1987 • ÅRGÅNG 8

<i>Mats Furberg</i>	Om dödsbegrepp och definitioner	1
<i>Ragnar Ohlsson</i>	De smutsiga händerna	11
<i>Ingmar Persson</i>	Överkvalificerade handlingar ur ideal- och realmoraliska perspektiv	20
<i>Erik Ryding</i>	Parfit och den personliga identiteten	29
<i>Björn Eriksson</i>	Artisten och djuren — en filosofisk cirkus	35
<i>Ingmar Persson</i>	Tollands artistiska magplask	39
	Recension	41
	Notiser	47

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 riktningar, men den vill undvika bidrag som man inte kan tillgodogöra sig utan speciella förkunskaper eller tekniska färdigheter. Utöver längre artiklar på omkring 10—25 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 (Göteborg)

Produktion, prenumerationer och annonser: Bokförlaget Thales, Box 50034, 104 05 Stockholm, redaktionssekreterare Astrid Thorsson, tel 08 162618 (arb), 08 7686512 (bost)

Prenumeration: Tidskriften kommer ut med 4 nummer/år, pris SEK 100:—, lösnummer SEK 30:—/ex, postgiro 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

Satt av EKEN fotosats, Täby

Tryckt hos Akademitryck AB, Edsbruk 1987

Mats Furberg

Om dödsbegrepp och definitioner

Men där bortom
spirar gräs där sjunger fåglar
någonstans någonting
ett Fäste

Karl Vennberg

Denna uppsats handlar om betänkandet *Dödsbegreppet* (SOU 1984:79). Åtskilligt finns att applådera i den statliga utredningen, framför allt polemiken mot tanken att sk hjärtdöd är vad vi normalt menar med någons död. Här vill jag dock främst redovisa några betänkligheter mot den ”grundläggande definitionen”, mot dess föregivna precisering och framför allt mot den bakomliggande förutsättningen att det överhuvud är möjligt att ge en realdefinition av ”händelsen död”.

1. Några huvuddrag i utredningen

Utredningen anser sig inte ha att ta ställning till ”det etiskt eller medicinskt berättigade i olika åtgärder som kan aktualiseras under döende-processen”, ej heller till ”frågor kring eutanasi (dödshjälp) och människans rätt att dö” (125). Dess uppgift är blott att ge ”en entydig och allmängiltig definition av *händelsen död*” (124). Denna definition ”skall [. . .] beskriva *övergången från liv till död*” och ”gälla enbart *människans död*” (125). Den ”måste ges ett sådant innehåll att den otvetydigt fastslår att döden är ett faktum som inträffar [!] när definitionens villkor är uppfyllda. Det är inte tillräckligt att man då kan säga att döden är en oundviklig konsekvens och att prognosen är hopplös” (126).

Efter att ha gått igenom och förkastat tidigare förslag framlägger utredningen sin ”grundläggande definition”:

En människa är död då hon totalt och oåterkalleligt har förlorat all förmåga att förena och samordna kroppens funktioner — fysiska och psykiska — till en fungerande enhet (147).

Med 'fysiska funktioner' förstås "människans rent kroppsliga funktioner — andning, matsmältning, blodtrycksreglering etc", medan uttrycket 'psykisk funktion' används för att "beteckna sådant som tankeverksamhet, minne, känslor m m" (140).

'Förmåga att fungera som enhet' bestäms aldrig. Tanken tycks vara att alla normala människor har egen tillgång till en repertoar av basala handlingar, varseblivningar o d. Den som inte har sådan tillgång annat än med hjälp utifrån saknar full förmåga att fungera som enhet.

Utredningen hävdar att det empiriskt "kan fastställas att det totala och oåterkalleliga bortfallet av samtliga hjärnfunktioner — i stora hjärnan, lilla hjärnan och hjärnstammen — innebär att människan oåterkalleligt har förlorat all förmåga att förena och samordna kroppens funktioner — fysiska och psykiska — till en fungerande enhet, dvs att hon är död" (150). Detta empiriska samband ger en precisering av dödsbegreppet:

En människa är död när hjärnans samtliga funktioner totalt och oåterkalleligt har fallit bort (152).

Med hjälp av den grundläggande dödsdefinitionen och dess "precisering" avvisas begreppen hjärndöd och hjärtdöd. En människas död är hennes, inte hennes hjärtas eller hjärnas, och hon kan inte vara både levande och död (154). Att hjärtverksamhet och andning har varaktigt upphört tyder dock på att samtliga hjärnfunktioner totalt och oåterkalleligt har fallit bort. Eftersom sådant bortfall obotligt gör det omöjligt för människan att samordna sin kropps funktioner till en fungerande helhet är varaktigt hjärt- och andningsstillstånd *kriterier* på att hon är död. Andra, mer direkta sådana kriterier är t ex djup medvetslöshet och reaktionslöshet utan tecken på högre hjärnfunktioner, frånvaro av alla hjärnstamsreflexer samt frånvaro av all elektrisk aktivitet i hjärnbarken (164). Varken hjärt- eller hjärnrelaterade döds-kriterier får emellertid ryckas in i dödsdefinitionen.

(Har jag rätt förstått detta, innebär det att den som inte längre har och inte heller längre kan återfå andning, matsmältning, blodtryck, tankar, minnen, känslor osv är död men att definitionen av död inte nämner hjärta, hjärna och deras "funktioner". Dock tillstår jag att jag är oklar över både vad utredningen menar med ett kriterium och huru-

vida den hävdar att kriterier under inga förhållanden får ingå i definitionen.)

En konsekvens av dessa begreppsövningar är enligt utredningen att den som har råkat ut för total och oåterkallelig hjärninfarkt inte är att betrakta som likställig med död och inte heller stipulativt bör räknas som död: han *är* död. Ty han har totalt och oåterkalleligt förlorat all förmåga osv. Att transplantera in hans slående hjärta i en annan människas kropp är alltså inte att beröva donatorn livet.

2. "Hjärtdöd" och "hjärndöd"

Utredningen fäster stor vikt vid definitioner. Den anser både att vardagens dödsbegrepp har en korrekt definition och att falska definitioner är i omlopp. I detta avsnitt stannar jag vid den senare åsikten.

Det har ibland sagts att vårt vardagsbegrepp är sådant att en mänskliga definitionsmässigt är död när hennes hjärta varaktigt och oåterkalleligt har slutat slå. Det kan inte gärna vara riktigt. Två illustrationer:

Pelle Lindberghs hjärta sätts in i en annan människas kropp vars utslitna hjärta därvid tas ut och slängs i vasken. Enligt definitionen är det då Pelle Lindbergh som överlever och hon som har fått hans hjärta är den som är död. Att hans kropp bevisligen är sänkt i graven och hennes bevisligen promenerar i solskenet har för anhängare av denna definitionstanke inte med saken att göra.

Leif Stenbergs hjärta togs bort och ersattes med ett konstgjort. För dem som tror på hjärtdödsdefinitionen måste han ha dött på operationsbordet. Det var alltså ett några veckor gammalt lik som efter operationen skrattade, pratade och gav teveintervjuer . . .

Beslut att inte tillåta hjärttransplantationer i Sverige har stundom motiverats med att hjärtats död är människans. Att motiveringen saknar tillstymmelse till stöd från sunda förnuftet torde de två illustrationerna ha gjort klart. Det är grovt osant att det vanliga dödsbegreppet sätter likhetstecken mellan hjärtats död och människans. Ett kriterium på att en människa är död är inte mer än ett kännetecken på att hon är död och får inte tas som en definition av död. Många liv kan räddas om utredningens insistens på denna elementära begreppssanning får politisk genomslagskraft.

En falsk dödsdefinition är alltså i omlopp. Att hjärtat strejkar kan inte likställas med att dess ägare dör. Hur är det då med hjärtdödsbegreppets konkurrent, hjärndödsbegreppet? Är vi alla eniga om att när

hjärnan är definitivt utslagen är ägaren död?

Bland dem som motsätter sig att respiratorn stängs av i fall av typ Pelle Lindbergh finns nog många som befarar att diagnosen total hjärninfarkt kan vara felställd. Dem kan vi bortse ifrån: deras invändning är ju att hjärnan trots allt kan vara igång. Verkligt motstånd mot en dödsdefinition i termer av total hjärninfarkt kommer från dem som genuint tänker sig att fastän hjärnan är totalt utslagen lever patienten så länge en apparat ännu mäktar hålla igång hans hjärta och syrsätta hans blod. Vet de vad 'död' betyder?

'Nej, det vet de inte!' Så måste svaret bli från dem som menar att det vanliga dödsbegreppet är definierat i termer av hjärnfunktioner. Motståndarna kan inte svenska! Stenåldersmannen *måste* ha ett annat dödsbegrepp än David Ingvar!

Absurditeten i svaret torde röja att också den definitionen är falsk. Om utredningen skulle tänka sig sin sk precisering som en precisering i vardagsbetydelsen (dvs som ett meningsförtydligande) måste den emellertid ge just det svaret. Såvitt jag kan se skulle den i så fall ramla i den grop den själv har varnat för: kriterier på död skulle ha förväxlats med en definition.

3. Dödsdefinition och dödsprecisering

Möjligen *kan* de flesta vardagsbegrepp definieras. Säkert är att många är odefinierade. Utredningen är dock övertygad om att just dödsbegreppet åtminstone för om den stund då betänkandet lades fram *har* en (real) definition, nämligen den i boken presenterade. Den definitionen är rent deskriptiv, utan "kvantitativa eller kvalitativa värderingar" (149, se även 146). Dessutom har den egenskapen att inte komma "i strid med [några som helst?] religiösa uppfattningar om vad döden innebär" (142).

Inget av dessa utredningens påståenden tycks mig trovärdigt.

Det kan vara gott att börja vid två krav som utredningens grundläggande definition sägs uppfylla: att både "vara enhetlig och ha samma innebörd i alla medicinska och rättsliga sammanhang". De kraven är *nödvändiga*. "Gränslinjen mellan liv och död *måste* vara klar och entydig" (127, min kursiv).

Varför måste den det? En skarp gräns mellan liv och död vore kanske önskvärd. Men måste den finnas? Är inte kontinuerliga övergångar vad man borde vänta sig? Utredningens såvitt jag kan finna enda icke-nor-

mativa skäl lyder: "En entydig dödsdefinition är en förutsättning för att läkaren skall kunna avgöra om hans patient är levande eller död" (129). Det förefaller mig inte stort mer övertygande än att en entydig skymningsdefinition är en förutsättning för att vi ska kunna avgöra om det är vackrast när det skymmer. Att läkaren skulle vilja slippa att i kinkiga fall lita till sitt eget omdöme är säkert sant. Men är en definition det som behövs? Och kan den i så fall undgå att vara normativ? Sådana frågor skjuter jag över till nästa sektion och betraktar här utredningens definition och dess föregivna precisering.

Ska definitionen ifråga vara brukbar måste det för det första finnas en ändlig och uppräknelig mängd "funktioner" hos människan eller hos hennes hjärna så att vi kan undersöka huruvida de samtliga är försvunna. Är det så? Kan vi säkert veta att de försvunna aldrig återkommer och aldrig *kan* återkomma? Kan vi veta det nu när hjärn- eller hjärtinfarkten just har inträffat eller först långt i efterhand?

För det andra: Vad är "funktionerna" ifråga? (Är också dysfunktioner funktioner? Om ja, var går en "fullständigt klar och entydig" gräns mellan sånt som inte överhuvud är funktioner och grava dysfunktioner? Om nej, hur går det med utredningens försäkran på s 149 att dess definition *inte* innefattar "kvantitativa eller kvalitativa värderingar"?) Som typexempel på fysiska funktioner nämns matsmältning och blodtrycksreglering. Ingendera faller klart under viljans kontroll. I vilken mån kan människan sägas av egen kraft förmå förena och samordna matsmältning och känsla, blodtryck och minne? Har en person med konstgjort hjärna någonsomhelst egen förmåga att samordna textilska och hjärtklappning? Om inte, är vederbörande då död?

Slutligen: antyder definitionens två sista ord att det finns icke-fungerande enheter? Vad vore i så fall exempel på dem? Är en klocka utan urverk en klocka? Är ett lik ett exempel på en icke-fungerande människa, en "enhet" som ohjälpligt har gått i stå?

Svärmen av frågor kan lätt förmeras. Det tycks mig rätt säkert att läkaren som står över trafikoffret eller juristen som grubblar över arvsfrågor i fallen Pelle Lindbergh och Leif Stenberg inte i utredningsdefinitionen får något "klart och entydigt" att jämföra den bångstyriga verkligheten med.

Bättre är han då betjänt av den skärpta preciseringen. Ty den verkar ge kriterier. Men hur dessa tänks förhålla sig till definitionen har sina problem:

I vardagslag brukar väl förhållandet mellan det som preciseras och

preciseringen tänkas vara av typen logisk inneslutning. Så kan det dock inte gärna vara i utredningen; då vore verkligen (hjärnrelaterade) kriterier förvandlade till definitionskonstituenten. Inte heller kan man rimligen ha menat (trots uttryckssättet på ss 163—165) att vardagens dödsbegrepp är bestämt i hjärnfysiologiska termer. Ty visst vore det en världsensation om någon totalt och oåterkalleligt förlorade samtliga hjärnfunktioner och ändå fortsatte andas, tänka, smälta maten, känna osv; men någon motsägelse eller annan semantisk absurditet går inte att skönja. Att hjärnan, den som fysiologen sysslar med, hör samman med tänkande, kännande m m är en empirisk upptäckt som ännu inte har blivit en begreppsanning. Det är därför svårt att förstå hur ”preciseringen” skulle kunna närma sig det krav eller barmhärtigare uttryckt det ideal som uppställs på s 157, nämligen att bestämningen under inga [!] förhållanden får ge utrymme för tveksamhet om huruvida någon lever eller är död.

Utredningen kan inte ha lagt nämnvärd vikt vid ordet ’precisering’. Oftast verkar det utbytbart mot ’extensionell ekvivalens’ e d; avsnittet 8.2 bär den både syntaktiskt och semantiskt förbryllande titeln ”Dödsbegreppets precisering (empiriskt samband mellan den grundläggande definitionen och funktionerna hos organ i människans kropp)”. Den senare dryga hälften av boken handlar om sådana ”empiriska samband”. Dessa är säkerligen både intressantare och viktigare än några semantiska. Men boken vill även här framtona som en begreppsutredning. Hjärnfysiologiska rön ges därigenom semantiskt status; de framträder som nya insikter om dödens begrepp, inte som nya fakta relevanta vid en begreppsrevision.

Därigenom förutsätter utredningen det som bör bevisas. Det är en sak att göra det normativa hävdandet att nya kunskaper gör det rimligt att ompröva härdvunna föreställningar om döden. Det är en vida annorlunda sak att göra deskriptiva anspråk på att vardagsbegreppet redan är sådant som det borde göras efter en normativ prövning.

4. Dödsbegreppets porositet

Vanligen yrkar man på definition av besvärliga begrepp bl a för att få hjälp att i förväg avgöra hur kinkiga fall ska behandlas. Ideligen anför utredningen just det motivet för sin definition och dess precisering: beslutets vända inför tex en respiratorstängning kommer tydligen att förflyktigas när sjukvårdspersonal och anhöriga vid fall av typen Pelle

Lindbergh har tillgång till en definition.

Det förefaller mig vara en besynnerlig människosyn. Jag betvivlar dessutom att en definition som ger den önskade botten mot beslutsångest skulle kunna åstadkommas ens om den gjordes stipulativ och starkt normativ. Men låt gå! Här nöjer jag mig med att ge skäl mot att någon definition skulle kunna vara sådan som utredningen drömmar om: *både* en pålitlig ledstång i brydsamma fall *och* icke-normativ *och* en trogen beskrivning av vardagens dödsbegrepp.

Vetenskapens begrepp är tämligen ofta tämligen välbestämda. Så är det sällan med vardagsbegreppen. De är t ex vaga så att man i många fall är osäker på deras tillämplighet; och av och till är de dessutom porösa, dvs saknar avgränsningar åt vissa håll.

Vardagens dödsbegrepp är inte vetenskapligt. Så länge vi håller oss till det menar vi nog alla att övergången mellan liv och död inte är knivskarp. Den sjuke svävar ofta mellan liv och död, vilket inte betyder att han än lever och än är död utan att han är i gränsområdet, närmare död än liv eller omvänt. Att med utredningen påstå att han antingen lever eller är död förefaller knappast vettigare än att stirra ut i gryningen och med djupsinnig min slå fast att utanför fönstret är det antingen fullt dagsljus eller kolsvart. Liv och död är liksom ljus och mörker vaga begrepp. Övergången från det ena ledet i begreppsparet till det andra är bara på helger något drastiskt, ett kvalitativt språng. Till vardags är den trippande.

'Huruvida vi är döda eller blott djupt medvetlösa avhänger alltså av hur våra medmänniskor beslutar sig för att dra en gräns i en skymningszon! Så skrämmande! Så får det inte vara! Granskar vi verkligheten djupare, *måste* den uppvisa ett skarpt, ett otvetydigt brott mellan liv och död!'

Så larmar det i oss, av förklarliga skäl. Men om verkligheten inte fogar sig efter våra önskningar, vad gör vi?

Med blott och bart en skärpning av vardagens dödsbegrepp kommer vi inte långt. Ty det begreppet är inte blott vagt utan därtill poröst. Från urminnes tider har vi kunnat nödtorftigt skilja mellan levande människor och lik, mellan medvetlöshet och död. Den levande människan drar andan, hennes puls slår och hennes hjärta tickar. I döden upphör andning och hjärtverksamhet, ögonen brister, kroppen svalnar och stelnar och blir fläckig. Men om medvetandet för alltid är borta och andning, syrsättning och hjärtverksamhet pågår endast genom att hon är inkopplad i en maskin? Vad säger vi då? På en sådan tidigare oanad typ

av situation är vi sannolikt begreppsligt oförberedda. En Pelle Lindbergh är ju inte vanligt medvetlös. Får vi tro läkardiagnosen är han bokstavligen utan varje form av medvetande; hans hjärna är numera blott en grå massa utan varje funktion. Att säga att han är eller inte är död är inte att krympa ett vaghetsområde. Det är att skapa en gräns där förut allt var öppet.

En sådan gräns erfordras. Det vanligaste sättet att staka ut den är att fastställa kriterier, tydliga kännetecken. Utredningen försöker sig på det. Den påpekar att vardagskriterier på död som att hjärtat upphör att slå och att andhämtningen slutar kräver tillägget att verksamheterna inte börjar igen. Det tillägget förutsätter att långvarigt hjärt- och andningsuppehåll obotligt fördärvar den drabbades hjärna och därmed hans möjlighet att åter spontanandas osv. Det är därför, menar utredningen, också i vardagen den totala hjärninfarkten man sig själv ovetande spanar efter när man som döds-kriterier tar hjärtats sista slag och lungornas sista pust.

Detta förefaller bestickande. Eftersom därtill starka normativa skäl talar för att fall av typ Pelle Lindbergh inte bör räknas som levande, så bör kanske de väsentligaste döds-kriterierna förklaras vara just de som är relaterade till hjärnans totala kollaps. Vi kan då stänga dödsbegreppets porer och avlägsna en del av dess vagheter. Men från att vi kanske bör så göra följer inte att vi har gjort det, utan fastmer att vi *inte* har gjort det. Utredningen medger indirekt detta när den inte ens vågar förordna att dödsögonblicket fastställs till 15 à 20 minuter efter den sista pusten, dvs till den tid då hjärnan är fullständigt förstörd av syrebrist (s 225).

Även om vi skulle godta total hjärninfarkt som nödvändigt och tillräckligt villkor för död följer ingen dödsdefinition. Det är ingen intellektuell katastrof, kanske tvärtom. Definitioner är stundom förträffliga men gör ibland större skada än nytta — vore det t ex önskvärt att ha en skarpt avgränsande definition som en gång för alla sade ifrån vad konst eller kultur eller för den delen vetenskap är?

'Liv och död är inte kulturella begrepp! De är därmed inte i viss mån artificiella som konst och religion!' Inte det? Är du säker på att de inte är sammanvävda med bl a religiösa och moraliska föreställningar? Utan sådan sammanvävning vore det godtyckligt och inte blott orealistiskt att kräva en dödsdefinition som inte kommer i strid med religiösa uppfattningar (142). Och antingen de är kulturella begrepp eller ej, är det inte tämligen sannolikt att en *definition* av dem inte har någon nytta

med sig?

Tydligare sagt: Är inte en dödsdefinitions huvuduppgift besvärjelsens? Med kaos för våra ögon och kaos i vårt inre stämmer vi in i bönen om Fästet: Hur bjärt och klar är egentligen inte världen! Hur fri är den inte i verkligheten från suddighet och tvetydighet! Hur återkommer inte alla situationer, djupast sett, så vi en gång för alla kan bestämma eller av någon auktoritet få besked om hur vi ska tänka, känna, handla! Sådana föreskrifter är i sanning inget vi människor har funnit på; dikterade är de av verkligheten själv, verkligheten som omutligt befast ett svalg mellan ljus och mörker, liv och död men fördolt det för fåkunniga blickar! O trygghet, du som vi skumögda inte har men som finns, som *måste* finnas — kom! Fräls oss ur villrådigheten, ur oron, ur beslutsångesten! Herre Definition, du som avslöjar vad Liv och Död innerst *är*, ur djupen ropar vi till dig!

Att avsluta i denna gycklande ton vore fel; men sannolikt ligger det mycket i besvärjelsemotivet. Ty det är genuint svårt att se varför döden annars behöver definieras. Tänk t ex på att medicinsk teknologi nog kommer att förverkliga vilda fantasier! Om hjärntransplantationer bleve möjliga, så skulle utredningens ”precisering” leda till absurditeter ifall den togs som ett definitionsförslag. Och är det inte nästan säkert att *varje* definition som vi idag försöker kan komma att i morgondagens värld ställas inför fall där den ger intuitivt oacceptabla resultat?

Naturligtvis kan vi göra rullande definitionsrevisioner; men poängen med en definition av döden är väl inte att den i efterhand klaffar med vad vi uppfattar som det rimligaste sättet att i situationen ifråga dra en begreppsgräns? Ska vi ha den, så ska den hjälpa oss *i förväg*. Men kan en sådan definition göras? Minns — (i) den måste bli stipulativ eftersom inga gränser är dragna; (ii) den bör grundas på moraliska överväganden, så mänskligt viktig som den är; och (iii) det är tvivelaktigt om dessa kan göras en gång för alla, oberoende av den faktiska situationen och vår kunskap om den . . .

5. Sammanfattning

Försök att definiera den sk händelsen död har hittills misslyckats. Därmed är inte sagt att företaget är fåfängt; men inte mycket talar för att en definition kan åstadkommas som är till stor hjälp i brydsamma situationer.

Om vardagens dödsbegrepp är både vagt och poröst måste det up-

penbart ändras i varje försök att göra det klart och tydligt, det må ske i definitionsform eller på annat sätt. Ändringarna bör därtill vara av normativ och värdemässig natur. Det vore överraskande om de inte då och då kommer i kraftig konflikt med t ex religiösa föreställningar.

Men är det som krävs för att komma till rätta med problem kring händelsen död verkligen en definition? Betänkandet har givit starka skäl för att hjärnbaserade kännetecken på död bör tas allvarligare än de gängse hjärt- och andningsbaserade, eftersom det oåterkalleliga i hjärt- och andningsstillståndet beror på att hjärnan är fördärvad. Accepterar man denna prioritering av de hjärnbaserade dödsriterierna, så har sjukvårdspersonal, anhöriga och jurister all vägledning som rimligen kan begäras i de typer av kinkiga fall som hittills har dykt upp. En befordran av hjärnbaserade kriterier till *de* viktiga skulle dessutom inte ändra på det hävdvunna dödsbegreppet; det skulle förbli lika vagt och öppet som det alltid har varit. Om nu *den* saken är något att stå efter. □

Ragnar Ohlsson

De smutsiga händerna

Jean-Paul Sartre använde sitt skönlitterära författarskap för att kasta en sidobelysning över sina filosofiska idéer. I släpljuset framstår deras knaggligheter i nya mönster. Pjäsen *De smutsiga händerna* från 1948 är ett exempel på detta.

Hugo Barin är en ung intellektuell med högborgerlig bakgrund. Han är medlem i Proletärpartiet — en underjordisk koalition av ett socialdemokratiskt och ett kommunistiskt parti — i Illyrien, beläget någonstans i Centraleuropa. Landet styrs av en fascistisk diktator och har varit i allians med Hitler. Andra världskriget lider mot sitt slut. Röda armén närmar sig, och regenten, den liberala oppositionen och Proletärpartiet förhandlar om att bilda en samlingsregering. Den kommunistiska delen av partiet vill mörda partiledaren Hoederer för att förhindra detta. Hugo, som dittills arbetat i den underjordiska kommunistpressen, utber sig och får uppdraget att verkställa mordet. Väl installerad som sekreterare hos Hoederer vacklar han, inser att Hoederer har goda skäl för sin kompromissvilja, men skjuter honom slutligen, när han överraskar sin fru Jessica i hans armar. När Hugo släpps efter två år i fängelse, upptäcker han att den kommunistiska fraktionen av partiet ändrat ståndpunkt och numera betraktar mordet på Hoederer som ett misstag. Hugo kan inte acceptera att hans mord var meningslöst och låter sig avrättas av kamraterna.

I Hugo ger Sartre ett nidporträtt av sig själv, genom hans mun flyter anspelningar på centrala teman i Sartres filosofi. Genom att framställa Hugo som löjlig och omogen, distanserar sig Sartre från sådant han för fram på fullt allvar på andra håll. Genom att de problem och idéer som Hugo sysslar med, i pjäsen kopplas till hans speciella bakgrund och psykologi, får man anledning att ifrågasätta den allmänna giltigheten hos teser som annars doceras.

Exempel på idéer från Sartres filosofi är det gott om: (a) Att människan saknar väsen, eller rättare att människans väsen är summan av

hennes handlingar, dvs att existensen föregår essensen, antyds flera gånger av Hugo: ”. . . en familjefar är aldrig en verklig familjefar. En mördare är aldrig helt och hållet mördare. De leker, begriper ni. Men en död, det är en verklig, definitiv död. / . . . / Det finns ingenting jag skulle kunna *vara* — annat än död med sex fot jord över skallen” (Tablå I, scen 4). I *Varat och intet* är detta ett metafysiskt faktum som Människan förgäves strävar att komma ifrån: denna väsenslöshet är den frihet vi är dömda till. Medvetandets sätt att existera är just detta att aldrig ha något väsen — utom denna tomhet. Ständigt strävar vi efter att fylla ut vår tomhet, att bli något bestämt, solitt, avgränsat och tydligt. Men förgäves — så länge vi är medvetna kan vi inte uppnå tingens entydighet. Och när vi en gång gör det, i döden, är vi inte längre medvetna. Enligt slutorden i *Varat och intet* är människan därför ”une passion inutile”, ett fåfängt begär.

I *De smutsiga händerna* är detta en speciell egenskap hos Hugo, han längtar efter en handling som skall ge honom en bestämd egenskap, ett väsen, därför att han är en onyttig utväxt på samhället och familjen. Hugo vill hänga ett brott om halsen ”som en sten” (VII). ”Mördare borde bära något igenkänningstecken. En röd blomma, en vallmo i knapphålet” (VII). En mördare eller en familjefar ”borde man känna igen. På någonting. En viss min. En smak i munnen. En tagg i hjärtat” (IV:6). Hugo upptäcker förstås att han inte förändras genom sin handling. Varken genom att gå in i partiet, genom att bestämma sig för mordet på Hoederer, eller genom att utföra det, skaffar han sig ett väsen i den mening han hoppas.

Men det som i *Varat och intet* är ett slags ontologisk tragik blir i *De smutsiga händerna* ett komiskt inslag av Hugos omogenhet, och samtidigt den psykologiska förklaringen till Hugos handlingar.

(b) Hugo uppfattar sig inte bara som tom, väsenslös, utan också som överflödig. Också detta har tidigare förts fram som en ontologisk sanning: allt som finns är överflödigt, dvs saknar nödvändigt samband med resten av världen; varje ting taget för sig är ”överflödigt”. I *De smutsiga händerna* är det tydligt att denna upplevelse hänger samman med Hugos klassbakgrund och familjeförhållanden. Upplevelsen av att vara väsenslös och överflödig får Hugo att gå in i Proletärpartiet, att underkasta sig sträng disciplin och att drömma om handling. Den hierarkiska ordningen och ordergivningen inom partiet skulle ge Hugo en bestämd plats och en uppgift i livet — han skulle inte längre vara överflödig, och han skulle vara någonting bestämt, nämligen ett instrument för världs-

revolutionen. När han finner att partimedlemskapet inte räcker för att ge honom en identitet, vill han handla: handlingen skall ge honom ett väsen. (Notera följande hur "handling" för Hugo betyder "att döda".)

(c) En tredje tes som Sartre utförligt diskuterat i *Varat och intet* säger att människors handlingar inte bestäms av det förflutna utan är uttryck för människans totala frihet. I *De smutsiga händerna* visar Sartre att indeterminismen inte är något smakligt alternativ till determinismen. Bekymret med determinismen är ju att om den är sann tycks den beröva oss vårt ansvar, ty är varje handling bestämd av det förflutna, kan vi aldrig göra något annat än det vi faktiskt gör. Men om indeterminismen är sann — gör vi då någonsin något som vi har ansvar för: blir inte våra handlingar då slumpmässiga, de bara händer oss? Efter sin långa redogörelse för mordet, säger Hugo: " 'Om' jag hade stannat lite längre ute vid kastanjeträden, 'om' jag hade fortsatt längre in i trädgården, 'om' jag hade gått in i flygelbyggnaden, där vi bodde . . . /hade jag inte skjutit/. Vad blir det av mig i allt det där, vart tar jag vägen? Det är mord utan mördare" (VII).

(d) En fjärde tes som pjäsen anspelar på är att "alla har ansvar för allt". Hugo säger att Hoederer har ansvar för sin egen död (VII). Och Jessica säger: ". . . jag har ingenting gjort, jag är oskyldig till allt" (V:2). Jessica förnekar alltså tesen från *Varat och intet*. Det är oklart hur Sartre ställer sig till hennes uttalande: antagligen kan regissör och skådespelare ge henne antingen rätt eller fel. Men man kan lägga märke till att det är Jessica som möjliggör mordet — genom att gömma revolvern och genom att rädda Hugo ur flera brydsamma situationer; och det är hon som utlöser mordet — genom att flirta med Hoederer. Om någon är ansvarig för hans död, är det hon.

I *Varat och intet* förfäktade Sartre tesen att vi är skyldiga till allt som ett ontologiskt faktum: varat-för-sig väljer den värld det lever i. Om man tar fasta på denna tes skulle styckets titel få en speciell innebörd: om vi är skyldiga till allt, är våra händer förvisso smutsiga. Men en sådan utsaga verkar ointressant som grund för ett ställningstagande till vad vi har ansvar för, bl a därför att det verkar omöjligt att välja en radikalt annorlunda värld än den där vi faktiskt befinner oss. Att vi kan föreställa oss alternativ till den värld vi ser omkring oss, och bitvis faktiskt också kan förändra den, är inte tillräckligt för att påstå att vi i någon djupare mening är moraliskt ansvariga för att den ser ut som den gör.

Detta är fyra exempel på filosofiska idéer som Sartre anspelar på och

leker med i pjäsen. Men jag tänkte framför allt uppehålla mig vid två andra idéer. 1) Det finns inga giltiga order. 2) Man kan inte handla politiskt utan att ådra sig moralisk skuld.

1. Ogiltiga order

Hugo framhäver flera gånger, att lydnad och sträng disciplin är en hederssak för honom (III:1). Men det är också ett behov för honom. "Jag har behov av att lyda. Lyda och inget annat. Ät; sov; lyd" (III:4). Han har alltså behov av order för att ge sitt liv en mening, för att inte känna sig överflödig, för att motivera sin existens. Man kan tycka att också högt ställda krav på lydnad och ordergivning borde bli tillfredsställda i ett leninistiskt parti. Men se, i dramat blir Hugo besviken. Orderna räcker inte för att tysta hans pladdrande skalle, de räcker inte för att hindra honom från att tänka.

Hugo: "Jag behöver disciplin."

Hoederer: "Varför det?"

Hugo: "Det finns för mycket tankar i skallen på mig. Jag måste bli kvitt dem" (III:4).

Men det visar sig alltså att det inte räcker med att lyda order. "Från ett visst ögonblick så lämnar alla order dig så ensam, överger dig. Ordern hade stannat kvar bakom mig och jag gick framåt alldeles ensam och jag dödade ensam och . . ." (I:1).

Jag tror att man bör uppfatta Hugos upptäckt som en genuin existentialistisk insikt — och att det är Sartres avsikt. Det är inte det faktum att det finns motstridiga order, som kommer från olika fraktioner av partiet, som är Hugos egentliga dilemma. Bekymret är inte heller, att det är en minoritetsfraktion som ger honom ordern att skjuta Hoederer, så att allt skulle vara gott och väl bara det vore centralkommittén i sin helhet som hade gett ordern. Det skulle vara att göra det för lätt för sig. Vi tvingas *alltid* välja vilka order vi skall utföra — oavsett vem som utfärdat dem. Det spelar ingen roll vem som utfärdar en order, eller vilken process som lett fram till ett påbud: det finns *ingen* beslutsordning som garanterar att de order som utfärdas, bör följas. Det är det som är Sartres poäng.

Det kan vara lätt att förbise den poängen i *De smutsiga händerna*, där Hugos order utfärdas av en liten grupp som själv utnämnt sig till representant för partiet. Det skulle inte vara mycket svårare att se att ordern inte har någon giltighet om den utfärdats av ett litet elitparti, som

själv utnämnt sig till arbetarklassens sanna företrädare. Men naturligtvis skulle inte orderns giltighet ha varit garanterad ens om den utfärdats av en i god ordning utsedd demokratisk församling. I alla situationer måste den handlande själv ta ställning till befallningens giltighet. Det är en idé om moraliska agents autonomi som är väl förankrad i västerländsk moralfilosofi, framhävd av Kant, driven till sin spets hos Sartre.

För både Hugo och Hoederer framstår det som självklart att politiskt handlande förutsätter disciplin och hierarkisk ordergivning. Det är idén att militär ordning är det effektivaste sättet att ta sig an världshistorien. Effektivt kanske det är. Särskilt effektivt verkar det vara om man vill åstadkomma koncentrationsläger och masslakt; sådana saker låter sig svårigen genomföras utan en hierarkisk ordning, där en (grupp) fattar besluten och en annan verkställer dem. Sådana beslut fattas bara av dem som känner sig ha ansvar för helheten, för nationen, rasen, klassen eller socialismen. Sådana beslut verkställs bara av dem som håller sig till givna order, som litar på auktoriteter, som lyder.

Alternativet är alltså att själv ta ansvar för sina egna handlingar, att inte tro att moraliskt ansvar låter sig överlåtas, att vägra lyda. Det var också Sartres uppmaning till de franska soldater som beordrades till Algeriet.

Detta är den första och viktiga innebörden i idén om orders ogiltighet. Tolkad på detta sätt, tror jag att den är sann och har vittgående och inte alltid beaktade konsekvenser. Men denna enkla och fundamentala sanning kombineras hos Sartre ofta med en tes om moraliska utsagors giltighet. Sartre tänker sig moraliska utsagor i orderform. I detta avseende liknar Sartre en del moderna anglosachsiska moralfilosofier — framför allt Hare: i metaetiken är Sartre preskriptivist. I denna tolkning säger alltså tesen att det inte finns några giltiga normer över huvud taget.

I *Existentialismen är en humanism* lekte Sartre en smula med tanken att rättfärdiga moraliska utsagor på ett kantianskt sätt, genom att kräva universaliserbarhet; men det är något som blev en parentes i Sartres författarskap: varken förr eller senare försökte Sartre rättfärdiga moraliska utsagor med hjälp av universaliserbarhetskriteriet. För Sartre återstår alltså ståndpunkten att inga moraliska utsagor kan ha objektiv giltighet.

Men Hugo, han tror på objektivitet i normativa sammanhang. När han försöker motivera inför Jessica varför han bör mörda Hoederer, gör han gällande att han *vet* att de som gett honom ordern har rätt: "Louis, han tar aldrig fel" (V:5). Hugo *vet* att ordern är riktig: "Politiken är en vetenskap. Du kan bevisa att du har rätt och att de andra har

fel" (V:2). Det är svårt att inte se den här repliken som exempel på "mauvaise foi", det självbedrägeri man enligt Sartre kastar sig in i för att undslippa den ångest som följer medvetandet att man är fri. Sartre anger i *Varat och intet* som en av de former "mauvaise foi" kan ta sig, just att tro på objektiva värden och normer. Att tro på objektiva värden innebär att man ser sig själv som ett redskap för deras förverkligande, dvs att man betraktar sig som ett ting med en given funktion och förnekar sin frihet.

Men är den motsatta ståndpunkten — den som verkar vara Sartres egen — mer acceptabel? Jessica antyder den, när hon frågar varför Hugo valt just Louis' och Olgas åsikter (V:2). Som Hugo mycket riktigt påpekar, låter det som om "alla åsikter har lika värde och att man ådrar sig dem som man ådrar sig sjukdomar" (V:2). Det är knappast en acceptabel ståndpunkt, och framställs inte heller så av Sartre i *De smutsiga händerna*.

2. Politiskt handlande och moralisk skuld

Tesen att moralen saknar objektiv giltighet — särskilt i Sartres extrema version — tycks också underminera tesen att politik medför smutsiga händer. Vad innebär tesen, bl a framförd av Hoederer, att man inte kan handla politiskt utan att få smutsiga händer, om inte att den politiskt aktive ådrar sig moralisk skuld? Om ingen objektivt giltig moral existerar tycks begreppet moralisk skuld inte få någon särskilt intressant innebörd. I en värld där normer väljs av den handlande själv tycks det vara svårt att ådra sig moralisk skuld, utom i någon psykologisk mening: man känner att man brutit mot någon norm man själv accepterar. Å andra sidan visar man ju vilka normer man accepterar genom sitt handlande, hur kan man då ådra sig skuld enligt Sartre?

Men låt oss för en stund bortse från problemet att förena preskriptivism med idén om moralisk skuld; det finns åtskilliga sätt att göra de två idéerna förenliga, men jag skall inte närmare gå in på det i det här sammanhanget. Låt oss i stället se närmare på vad tesen om de smutsiga händerna egentligen säger.

Man kunde för det första tänka sig att det är en varning för att ägna sig åt politik. Politiskt handlande betyder att man tvingar på andra sin vilja. Problemet är att alla goda ting inte kan realiseras samtidigt: något måste offras för att något annat skall segra. Vad värre är: troligen företräds de värden jag är beredd att offra av personer för vilka de är

viktiga. Jag måste därför hindra andra människor från att få sina önskingar tillfredsställda, tvingas hindra dem från att försöka — jag måste till slut hota dem med fängelse eller andra straff. Vilka goda saker jag än genomför, kommer jag alltid att ha offrat någon annan sak som också var bra och ha utfört handlingar som har dåliga konsekvenser.

Anarkismen är den lära som förnekar det politiska tvångets legitimitet. En tolkning av *De smutsiga händerna*, förenlig både med många av Sartres politiska ställningstaganden och med tesen om orders ogiltighet skulle vara att pjäsen innebär en förkastelsedom över politiken. En sådan ståndpunkt rymmer givetvis sina problem. Genom att avstå från politiskt handlande lämnar man vägen öppen för dem som är beredda att gripa den politiska makten och använda den för sina syften. Anarkismens svar skulle vara att rekommendera allmän olydnad. Om ingen lyder kan ingen regera. Var och en skulle tvingas ta ansvar för sina egna handlingar. Detta vore en utopi i överensstämmelse med Sartres existentialistiska ståndpunkter i *Varat och intet*.

Men denna idé problematiseras i dramat genom att Hoederers politiska ståndpunkt framstår som så förnuftig och realistisk. I en värld präglad av kamp om den politiska makten, vad kan man göra annat än försöka att själv gripa makten, tillsammans med dem som omfattar liknande värderingar, för att genomföra det som tycks en önskvärt? Tesen att politiskt handlande medför smutsiga händer kan ges en macchiavelisk innebörd: endast de som är beredda att tillgripa vilka medel som helst förmår gripa och behålla den politiska makten; gör dig därför av med dina moraliska fördomar och använd de mest effektiva medlen för att genomföra det som är viktigast!

Hoederer framstår i pjäsen som utilitarist. Den sanne utilitaristen ser inte något annat problem än svårigheten att räkna rätt. Väger så här mycket framtida lycka faktiskt upp så här mycket lidande i det närvarande? Det är klart att man kan räkna fel i så komplicerade kalkyler. Men i princip finns det inga problem annat än de rent praktiska att utföra de nödvändiga beräkningarna på det bräckliga underlag vi har. Ett problem utöver det aritmetiska är naturligtvis att avgöra vilka värden som är viktiga: är det vissa principer, är det jämlikhet eller är det människors lycka, som är det viktiga? Hoederer, den gamle cynikern, tar här ett skamgrepp på idealisten Hugo: "För mig betyder det något om det finns en människa mer eller mindre på jorden. Det är något dyrbart" (V:3).

Utilitaristen Hoederer hämmas inte heller av några föreställningar om

att någon viss kategori av handlingar skulle vara förbjuden. ”Jag har inga principiella invändningar mot det politiska mordet” (IV:3). Allt är enligt Hoederer tillåtet för att avskaffa klassamhället, den enda fråga som återstår är den om vilka medel som är effektivast. ”Alla medel som är effektiva är goda” (V:3).

Skulle Sartre helt enkelt bara vilja illustrera en utilitaristisk tes: avgörande för handlingars riktighet är deras konsekvenser och därvid är inga handlingar i princip förbjudna? Riktigt så enkelt är det nog inte. Det är sannolikt att den cynism som utmärker Hoederer och många andra realpolitiker i längden är skadlig. Den som så beredvilligt som Hoederer accepterar lögn och mord, kompromisser och svek bara de leder mot målet har alltför lätt att acceptera medel som faktiskt inte är nödvändiga, men som tycks effektiva för stunden. På sikt tenderar medlen att överta målets karaktär: man vet att man måste gripa makten, men har glömt till vad ändamål. Hugo spekulerar en smula över effekterna på partiet av Hoederers kompromissvilja (V:3) på ett sätt som verkar välgrundat: ett parti som tillåter sig vilka metoder som helst kommer att korrumpiera sina medlemmar, dra till sig fel personer och förlora förmågan att genomföra det mål för vars skull man var beredd att offra allt — också sin moral. Men faktum kvarstår att Hoederer är den som framstår som mest förnuftig och sympatisk.

Om det alltså är rimligt att läsa pjäsen som en kritisk kommentar till ståndpunkter som Sartre själv förfäktat i andra sammanhang, kanske också tolkningen av tesen om politikens tendens att smutsa händerna på sina utövare kan ses som ett försök att formulera en kritik av Sartres egen hållning. Som kritisk intellektuell stod Sartre i praktiken utanför det politiska skeendet. Vad han skrev och sade, hade på sin höjd marginella effekter på den politiska verkligheten. Då är det också möjligt att behålla sina händer rena. Ser man *De smutsiga händerna* ur det perspektivet blir pjäsen ett försvar för den praktiske realpolitikern, som faktiskt tar på sig ansvaret att försöka förändra verkligheten och som därvid tvingas till utilitaristiska avvägningar.

Läser man pjäsen på detta sätt, ger den alltså stoff att ifrågasätta några av Sartres viktigaste teser. Tesen att existensen föregår essensen framstår som ett psykologiskt uttryck för en bestämd klassposition och psykologisk läggning. Tesen att vårt förflutna inte bestämmer våra handlingar framstår i ny dager, när det slumpmässiga i Hugos mord inte heller tycks ge honom ett större ansvar. Tesen att det inte finns några giltiga normer blir till den rimligare uppmaningen att vi själva måste ta

ställning till vilka normer som tycks oss mest acceptabla, vilket innebär att inga normer är giltiga i kraft av de auktoriteter som utfärdat dem. Därvid spelar överväganden om vilka konsekvenser våra handlingar får en viktig roll: det är kanske inte alldeles likgiltigt vilka normer vi väljer.

Av detta kan vi kanske lära att det vore nyttigt om fler filosofer ägnade sig också åt skönlitterärt författarskap, där de konkretiserade, prövade och omprövade sina filosofiska idéer. □

Litteratur

Jean-Paul Sartre: *L'être et le néant* (1943).

L'existentialisme est un humanisme (1946).

Les mains sales (1948). Citaten från Eyvind Johnsons svenska översättning.

De smutsiga händerna, publicerad i Jean-Paul Sartre, *Fem dramer* (1964).

Hänvisning till tablå, romerska siffror, och scen, arabiska siffror.

Ingmar Persson

Överkvalificerade handlingar ur ideal- och realmoraliska perspektiv

I en artikel i *Filosofisk tidskrift* (1985:1) behandlar Erik Ryding kritiskt några försök att komma tillrätta med s k överkvalificera(n)de handlingar, d v s handlingar som tycks bestå i att man gör "mer" än vad moralen kräver eller i berömvärda handlingar, vars uraktlåtanden inte hade varit klandervärda. Bl a skärskådar han mina bemödanden i ämnet i uppsatsen "Moraliskt hänsyntagande och dess gränser" (*Filosofisk tidskrift*, 1982:4). Min uppfattning är att Ryding misslyckas med att exakt lokalisera min position, med påföljd att han fyrar av sina invändningar i fel riktning. Jag känner mig emellertid inte speciellt hågad att utförligare påvisa detta, eftersom jag numera tror att det finns andra argument som torpederar den ståndpunkt jag intog i nämnda alster.

Hellre än att av solidaritet med mitt jag i det förgångna dra fram avlagda åsikter för att demonstrera att de inte är så malättna, vill jag skissera hur jag i dag är böjd att tolka överkvalificerade handlingar i termer av en moralsyn som i grund — om än inte i formulering — förblivit densamma sedan nyssnämnda artikel. Detta företag nödvändiggör att jag parerar några reflektioner av mer allmän räckvidd, som Ryding avslutningsvis låter undslippa sig, men det ger mig också möjlighet att ge en vink om varför jag menar att hans kritik av min tidigare ståndpunkt missar sitt mål.

Först skall jag i nya ordalag karakterisera vad som, enligt mitt förmenande, gör en viljeinställning — och därmed den handling som verkställer attityden — moraliskt riktig. Antag att en agent befinner sig i en situation där han har att avgöra vilket ställningstagande som är moraliskt rätt: kanske måste han besluta om han bör rädda en eller flera personers liv till priset av ett större eller mindre offer för egen del. Ett moraliskt beslut måste träffas på grundval av en beskrivning av situationen i rent *universella* termer eller åtminstone i termer som bara på ett *eliminerbart*

sätt refererar till individer och företeelser i den faktiska världen. Ett exempel på en eliminerbar referens till något faktiskt existerande är om jag beskriver en persons näsa genom att säga att den ser ut "som *min*". I princip — men kanhända inte i praktiken p g a ordförrådets torftighet — kan "som min" bytas ut mot en fras som karakteriserar min näsa i rent universella termer, som en näsa av den och den storleken och formen. Ifrågavarande universella beskrivning av situationen måste nämna alla situationens egenskaper — såväl egenskaper som den besitter i sig själv som relationer till andra situationer — vilka i någon riktning influerar agentens viljeattityder. I övrigt behöver ingen fullständighet eftersträvas.

En sådan beskrivning lämnar det helt öppet vilka, om några, faktiska individer de uppmålade individerna är identiska med. En attityd formad på basis av enbart denna typ av beskrivning kan följaktligen inte vara partisk till ens egen eller någon annan faktisk individs fördel. Naturligtvis har en moralagent som regel till sitt förfogande en oerhörd mängd information både om sig själv och om andra, men denna information får han inte begagna sig av om han vill nå en moraliskt oantastlig hållning. Han måste alltså bortse från den kunskap han äger om den faktiska världen, vilken gör det möjligt för honom att konkludera att han är identisk med agenten snarare än med en av de personer, vars liv agenten kan rädda, och vilken möjliggör bedömningar av sannolikheten för att han hamnar i den ena positionen snarare än i den andra (eller ingen alls). Om agenten vet t ex att han själv är man, kan han inte tillåta sig att förespråka en policy där kvinnor behandlas illa av det skälet att han är medveten om att sannolikheten för att han skall bli kvinna är praktiskt taget obefintlig. Däremot får han ta hänsyn till sannolikhetsuppskattningar vilka helt grundar sig på förhållanden som specificeras i den universella beskrivningen, som exempelvis numrära förhållanden. Om situationen är den att 5 personer kan räddas genom att agenten offrar sitt eget liv, måste han ta med i beräkningen att det är 5 gånger så sannolikt att han får åtnjuta förmånen att bli räddad som att han får underkasta sig ett offer av sitt liv, givet att han intar viljeattityden att agenten skall begå denna offergärning.

Självfallet är det ingen enkel sak att försäkra sig om att denna kunskap om den faktiska världen inte infiltrerar den universella beskrivningen och förvränger ens ställningstagande. Man kan lätt bedra sig själv till att tro att man hellre vill att en person, som i universella hänseenden är mycket lik en själv, skall leva än flera andra, som har färre

egenskaper gemensamma med en själv, alldeles oberoende av det faktum att de universella egenskaperna är utmärkande för just en själv, medan det i själva verket är så att dessa egenskaper tilltalar en just precis därför att de är framträdande hos en själv. För att motverka denna form av självbedrägeri kan man försöka att livligt åskådliggöra för sig den logiska möjligheten att man själv hade utvecklat (eller kommer att utveckla) radikalt skilda drag, förslagsvis dem som andra varelser som är inblandade i situationen är utrustade med. Man kan med andra ord leva sig in i deras omständigheter och begrunda hur man då skulle ha velat att agenten uppförde sig. Denna imaginära partiskhet tjänar som en motvikt mot den verkliga favorisering av en själv som man vanligtvis tenderar till.

Mitt förslag är således att en persons vilja att göra en viss handling i en viss situation är moraliskt riktig om och endast om han, på basis av enbart en i alla relevanta (= viljeinfluerande) avseenden komplett, universell beskrivning av situationen, vill att agenten i situationen utför denna handling. Kommer viljeinställningar bildade under dessa betingelser att uppvisa individuella differenser? Det är frestande att, i likhet med t ex John Rawls i *A Theory of Justice* (Harvard U P, 1971), besvara frågan nekande. Rawls förmodar att om beslutsfattare tvingas ignorera sin kännedom om — nödgas dra ner "en okunnighetens slöja" framför — den egna personligheten, sociala rollen etc, så möjliggörs enhällighet i fråga om normativa ställningstaganden: "Okunnighetens slöja möjliggör ett enhälligt val av en speciell konception av rättvisa" (s 140). Grunden för detta antagande blottläggs när Rawls diskuterar risktagande:

Den väsentliga saken är att inte tillåta de valda principerna bero av speciella attityder till risk. Av detta skäl utesluter okunnighetens slöja också kunskapen om dessa böjelser: parterna vet inte om de har en karakteristisk motvilja mot att ta chanser eller ej (s 172).

Tydligen gör Rawls här det bisarra antagandet att en viljeattityd man hyser påverkar ens val endast om man är *medveten om* att man hyser attityden. Men självfallet är det så, att en spelarnatur tar risker därför att han *är* benägen och villig att ta risker och inte därför att han *inser* att denna disposition hårbärgeras i hans inre; som regel är denna insikt *en följd* av att man observerar risktagande beteende hos sig själv.

De kognitiva restriktioner jag har introducerat garanterar sålunda ingen motivationell likformighet. Mitt moralkriterium utmynnar följ-

sätt refererar till individer och företeelser i den faktiska världen. Ett exempel på en eliminerbar referens till något faktiskt existerande är om jag beskriver en persons näsa genom att säga att den ser ut "som *min*". I princip — men kanhända inte i praktiken p g a ordförrådets torftighet — kan "som min" bytas ut mot en fras som karakteriserar min näsa i rent universella termer, som en näsa av den och den storleken och formen. Ifrågavarande universella beskrivning av situationen måste nämna alla situationens egenskaper — såväl egenskaper som den besitter i sig själv som relationer till andra situationer — vilka i någon riktning influerar agentens viljeattityder. I övrigt behöver ingen fullständighet eftersträvas.

En sådan beskrivning lämnar det helt öppet vilka, om några, faktiska individer de uppmålade individerna är identiska med. En attityd formad på basis av enbart denna typ av beskrivning kan följaktligen inte vara partisk till ens egen eller någon annan faktisk individs fördel. Naturligtvis har en moralagent som regel till sitt förfogande en oerhörd mängd information både om sig själv och om andra, men denna information får han inte begagna sig av om han vill nå en moraliskt oantastlig hållning. Han måste alltså bortse från den kunskap han äger om den faktiska världen, vilken gör det möjligt för honom att konkludera att han är identisk med agenten snarare än med en av de personer, vars liv agenten kan rädda, och vilken möjliggör bedömningar av sannolikheten för att han hamnar i den ena positionen snarare än i den andra (eller ingen alls). Om agenten vet t ex att han själv är man, kan han inte tillåta sig att förespråka en policy där kvinnor behandlas illa av det skälet att han är medveten om att sannolikheten för att han skall bli kvinna är praktiskt taget obefintlig. Däremot får han ta hänsyn till sannolikhetsuppskattningar vilka helt grundar sig på förhållanden som specificeras i den universella beskrivningen, som exempelvis numrära förhållanden. Om situationen är den att 5 personer kan räddas genom att agenten offrar sitt eget liv, måste han ta med i beräkningen att det är 5 gånger så sannolikt att han får åtnjuta förmånen att bli räddad som att han får underkasta sig ett offer av sitt liv, givet att han intar viljeattityden att agenten skall begå denna offergärning.

Självfallet är det ingen enkel sak att försäkra sig om att denna kunskap om den faktiska världen inte infiltrerar den universella beskrivningen och förvränger ens ställningstagande. Man kan lätt bedra sig själv till att tro att man hellre vill att en person, som i universella hänseenden är mycket lik en själv, skall leva än flera andra, som har färre

egenskaper gemensamma med en själv, alldeles oberoende av det faktum att de universella egenskaperna är utmärkande för just en själv, medan det i själva verket är så att dessa egenskaper tilltalar en just precis därför att de är framträdande hos en själv. För att motverka denna form av självbedrägeri kan man försöka att livligt åskådliggöra för sig den logiska möjligheten att man själv hade utvecklat (eller kommer att utveckla) radikalt skilda drag, förslagsvis dem som andra varelser som är inblandade i situationen är utrustade med. Man kan med andra ord leva sig in i deras omständigheter och begrunda hur man då skulle ha velat att agenten uppförde sig. Denna imaginära partiskhet tjänar som en motvikt mot den verkliga favorisering av en själv som man vanligtvis tenderar till.

Mitt förslag är således att en persons vilja att göra en viss handling i en viss situation är moraliskt riktig om och endast om han, på basis av enbart en i alla relevanta (= viljeinfluerande) avseenden komplett, universell beskrivning av situationen, vill att agenten i situationen utför denna handling. Kommer viljeinställningar bildade under dessa betingelser att uppvisa individuella differenser? Det är frestande att, i likhet med t ex John Rawls i *A Theory of Justice* (Harvard UP, 1971), besvara frågan nekande. Rawls förmodar att om beslutsfattare tvingas ignorera sin kännedom om — nödgas dra ner "en okunnighetens slöja" framför — den egna personligheten, sociala rollen etc, så möjliggörs enhällighet i fråga om normativa ställningstaganden: "Okunnighetens slöja möjliggör ett enhälligt val av en speciell konception av rättvisa" (s 140). Grunden för detta antagande blottläggs när Rawls diskuterar risktagande:

Den väsentliga saken är att inte tillåta de valda principerna bero av speciella attityder till risk. Av detta skäl utesluter okunnighetens slöja också kunskapen om dessa böjelser: parterna vet inte om de har en karakteristisk motvilja mot att ta chanser eller ej (s 172).

Tydligen gör Rawls här det bisarra antagandet att en viljeattityd man hyser påverkar ens val endast om man är *medveten om* att man hyser attityden. Men självfallet är det så, att en spelarnatur tar risker därför att han *är* benägen och villig att ta risker och inte därför att han *inser* att denna disposition hårbärgeras i hans inre; som regel är denna insikt *en följd* av att man observerar risktagande beteende hos sig själv.

De kognitiva restriktioner jag har introducerat garanterar sålunda ingen motivationell likformighet. Mitt moralkriterium utmynnar följ-

aktligen i en *individualisering* eller *relativisering* av moralbegreppet: en persons attityder eller handlingar kan vara moraliskt rätta blott *givet* fundamentala fakta rörande den personens typ av viljeliv. (Om det empiriskt förhåller sig så, vilket jag betvivlar, att alla moralagenter ytterst är harmoniskt stämda, behöver denna relativisering aldrig göras explicit.) Det exempel som Ryding med många andra har använt för att illustrera fenomenet överkvalificerande handling kan konkretisera saken: om en soldat kastar sig över en exploderande handgranat, räddar han livet på flera av sina kamrater på bekostnad av sitt eget, medan dessa går döden till mötes om han i stället räddar sitt eget skinn genom att söka skydd. Det är tänkbart att det är moraliskt korrekt för honom att inta attityden *A*, att föredra att rädda sitt eget liv. Han kanske föredrar att leva i en värld där djungelns lag råder och där ingen ger avkall på något för någon annan, trots att han inser att hans chanser att överleva i en sådan värld är sämre än i en värld där agenter är beredda att offra det egna livet för att rädda flera medvarelsers. (Denna kalkyl gäller utifrån den universella information moralagenten är hänvisad till.) Eller också kanske han ser någon universell kvalitet, vilken han själv, till skillnad från hans kamrater, besitter, som så värdefull att dess bevarande gör världen bättre än vidmakthållandet av kamraternas liv. Tänkbart är också att den moraliska attityden för soldaten är *B*, att föredra att kasta sig över granaten och rädda kamraternas liv. Anledningen till detta kan vara att, med utgångspunkt från den universella informationen, överlevnadschanserna är störst för var och en i en tillvaro där denna policy efterlevs. Alternativt kan anledningen vara den att han vill se förverkligad en värld i vilken agenter iscensätter självupppoffrande handlingar.

Det var denna relativisering som utgjorde bakgrunden till min tolkning av s k överkvalificerade handlingar i den av Ryding kritiserade uppsatsen. Jag summerade min uppfattning som så:

en agents handlingar och attityder är inte övermeriterande relativt hans egen moralkodex utan relativt andra mindre hjältemodiga eller helgonlika individers moralkodex vilka bildar genomsnittet (s 13).

Jag tänkte mig att man på grundval av vittskilda, konsekvent universellt baserade viljeinställningar kunde konstruera divergerande moraliska normsystem — exempelvis ett som inkluderar att man bör ha attityden *A* och ett annat som innefattar en plikt att ha *B*. Vidare antog jag att dessa system kunde rangordnas i stränghetshänseende: systemet som begär *B* av agenten är strängare än ett som fordrar *A*, eftersom det förra

nödvändiggör offer av det egna livet.

Det är denna senare föreställning som jag numera avvisar som ohållbar: att hylla *A* som moraliskt riktig är knappast mindre uppfordrande än att hylla *B*, ty det innebär att acceptera en värld där vars och ens chanser att överleva är, såvitt man kan bedöma, mindre. Faller denna rangordningsidé, faller också min tidigare tes att den som intar hållning *B*, och handlar därefter, exekverar en överkvalificerad handling i den bemärkelsen att han, i och med han utför precis vad han är förpliktigad att utföra likmätigt sitt moralsystem som är strängare än det som omfattar påbudet om *A*, gör *mer* än vad det senare systemet fordrar.

Ryding tycks förbise att min analys av s k överkvalificerade handlingar var grundad på denna relativisering eller individualisering av moralbegreppet. Han tangerar idén på ett par ställen (s 31, 32), men gör inte mycket mer än att ge luft åt sitt ogillande. Sin huvudsakliga energi ägnar han åt att genomdriva en annan tes som han pådyvlar mig: "Det överkvalificerade momentet består alltså i att han /soldaten/ låter sitt eget starkare intresse vika för andras svagare" (s 32). Jag känner mig inte alls träffad av detta försök till bestämning av min position och den efterföljande salvan. Från min universalistiska ståndpunkt är det lika felaktigt att gynna en annan på ens egen bekostnad därför att han är *en annan*, som det är att gynna sig själv därför att man är *sig själv*.

Nu till den väsentliga frågan: hur skall man förstå s k överkvalificerade handlingar? Jag menar att man måste fatta detta fenomen mot bakgrund av det förhållandet att människors moralnormer (d v s de viljeattityder som de kan ha av rent universella skäl) är så betungande att de flesta inte alltid förmår att leva upp till dem i praktiken. Tänk er (vilket inte är orimligt) att en majoritet av moralagenter skulle anse att de bör inta hållningen *B*, av det egoistiska skälet att denna policy ger dem störst chans att överleva, enligt en sannolikhetskalkyl på rent universella premisser. När en sådan individ hamnar i en situation vilken tillför honom den icke-universella informationsbiten att *han själv* är agenten som har att offra sig (och inte någon som kan bli räddad), kommer samma egoism, på basis av denna fullständigare information, att styra honom mot attityd *A*. Endast den som är besjälad av ett altruistiskt sinnelag — och som betraktar *B* som moralisk åtminstone delvis av dylika skäl — kommer att kunna vara sin moral trogen i realiteten. Om nu merparten av moralagenter är lagda åt egoism snarare än altruism, kommer de att söka värna om sin egen livhank, fastän deras moral dik-

terar motsatsen. Det fåtal som lyckas efterleva normen att inta *B*, och handla därefter, begår en överkvalificerad akt i den meningen att *de för-
mår uppfylla sin* — som, *ex hypothesi*, också är majoritetens — *morals
fordringar i större utsträckning än vad som är genomsnittligt*. Att utföra
en överkvalificerad handling är således inte att göra något *utöver* det
ens moral kräver, utan att göra något vars underlåtande hade varit
orätt, i omständigheter där endast en minoritet är motivationsmässigt i
stånd till att verkställa denna morals fordringar.

Invändningen mot denna tolkning är given: den strider uppenbarligen
mot den gängse karakteriseringen av överkvalificerade handlingar som
handlingar vars utförande är rätt och berömvärt, men vars uraktlåtande
inte är fel och klandervärt. Man får emellertid inte ta för givet att denna
standardbeskrivning är sakrosankt och måste behandlas som ett datum
vilket teorier om vad som är moraliskt rätt måste vara förenliga med.
Det är alltid möjligt att ett intuitivt moraliskt omdöme framstår som
förkastligt och i behov av revision i ljuset av en teori som strävar efter
att systematisera moralintuitioner, vare sig detta sker genom att härleda
dem från någon enstaka, fundamental princip (som t ex utilitetsprinci-
pen) eller genom att utforma en särskild beslutsprocedur som generar
dem (vilket jag här har försökt mig på). Och som R M Hare betonar i
Moral Thinking (Clarendon Press, 1981, s 199) intuitionen, att den som
iscensätter en överkvalificerad handling utför en handling vars uraktlå-
tande inte vore fel och klandervärt, kan höra till denna kategori. (I min
tidigare uppsats var jag blind för denna möjlighet, liksom Ryding ver-
kar vara.) Jag är dock inte beredd att gå till sådana extremer som Hare
när han tycks luta åt åsikten (se t ex *Moral Thinking*, kap 1.3) att man
kan göra moralfilosofi utan att alls appellera till andra intuitioner än
lingvistiska. När en teori gör rent hus med alltför många och djupt ro-
tade moralföreställningar, förverkar den sin trovärdighet, och man ser
sig om efter alternativ. Meningarna om s k överkvalificerade handlingar
går dock alltför mycket i sär för att någon beskrivning av dem skall
kunna betraktas som oomkullrunkelig. Betänk t ex att det hör till van-
ligheten att agenten själv i all anspråkslöshet bara säger sig ha gjort sin
plikt.

Därtill kommer att man med bistånd av ett konsgrepp inspirerat av
Hare kan förklara standardbeskrivningens plausibilitet utan att anta
dess riktighet. Låt mig särskilja två moraliska nivåer vilka jag skall
döpa till *idealmoralen* respektive *realmoralen*. Den förra har jag redan
försökt skissera ett kriterium för: den består av vad som *är* moraliskt

rätt och fel (gott och ont etc). Realmoralen i en viss social gemenskap, däremot, består i vad folk i denna gemenskap *hävda* vara moraliskt rätt och fel *i syfte att, givet agents begränsade moraliska motivation, nå ett resultat som är en så god approximation som möjligt till ett perfekt efterlevande av idealmoralen*. Det befrämjar inte alltid detta syfte att städse säga att något är rätt (eller fel) när det *är* rätt (eller fel) och att säga att något är rätt (eller fel) endast om det *är* rätt (eller fel). Man inser kanske att utförandet av en moraliskt korrekt handling i vissa omständigheter förutsätter ett stort mått av viljestyrka, som en viss agent saknar, och att konsekvenserna riskerar att bli fatala om agenten uppmuntras att ge sig i kast med denna handling. Då kan det vara lämpligt att för honom förtiga vilken handling som är den riktiga och i stället förvisa honom om att det räcker om han gör en betydligt mindre ambitiös handling. På samma sätt, om någon har försummat en överkvalificerande handling, kan det vara oklokt att fästa hans uppmärksamhet på detta förhållande och klandra honom, ehuru han enligt min tolkning egentligen är klandervärd. Han är ju inte *speciellt* klandervärd, d v s inte mer klandervärd än folk i allmänhet, och av, låt oss säga tidsbrist, måste man inskränka sig till att ta i upptuktelse dem som gör grövre förseelser. Att en person faktiskt undslipper klander kan lätt tas till intäkt för att han inte är klandervärd. Den överkvalificerade handlingen, vars underlåtande inte *sägs* vara fel och inte väcker (påtagligt) klander, får då sken av att vara en handling vars underlåtande inte *är* fel och klandervärd.

En del filosofer bestämmer det moraliskt rätta som det som är i enlighet med något i stil med vad jag har betecknat som realmoral, nämligen ett normsystem som instiftats för att styra det samhälleliga beteendet i en riktning, som befördrar harmoni och välbefinnande hos de individuella medlemmarna. En sådan konception finns antydd i denna enda bok jag känner till som ägnas överkvalificerade handlingar, nämligen David Heyds *Supererogation* (Cambridge U P, 1982), och den utvecklas med speciellt stor kraft av R B Brandt i *A Theory of the Good and the Right* (Clarendon Press, 1979). Denna approach involverar den syn på överkvalificerande handlingar som kommer till uttryck i standardbeskrivningen. Men vid en närmare granskning av Brandts analys framträder bristerna i en sådan approach.

Kvintessensen av Brandts propå (se t ex s 194) är att det att en handling är moraliskt fel betyder att den är förbjuden av varje beteendereglerande system (som uppfyller vissa villkor) vilket fullt rationella perso-

ner skulle tendera att stödja, framför alla andra dylika system eller inget alls, för ett samhälle i vilket hägn de förväntar sig att tillbringa hela sitt liv. Brandt påpekar (s 191—192) att en rationell person inte skulle finna det mödan värt att ge sitt stöd åt ett regelsystem som med största sannolikhet inte skulle vinna några andra rationella individers bifall, som exempelvis ett system skräddarsytt för att ge en själv maximal tillfredsställelse. Vad Brandt emellertid inte tillräckligt uppmärksammar är att, eftersom han inte stipulerar att hans rationella individer måste vägledas av någon välvilja visavi alla medvetna varelser, finns det en risk att dessa individer enas om regelsystem som förtrycker och utnyttjar varelser som, i kraft av sina bristande förståndsgåvor, inte kan sätta sig till motvärn — varelser som mentalt retarderade, icke-mänskliga djur eller t o m framtida generationer som kommer att leva efter den generation som f n åtnjuter rationalitet. Jag tror inte att jag är ensam om att tycka att det tar emot att klassificera en handling som moraliskt rätt när den är i överensstämmelse med ett normsystem som ignorerar dessa gruppers väl och ve. Det tillstånd, som resulterar om ett normsystem omsätts i praktiken, måste vara moraliskt godtagbart om handling i enlighet med systemet skall kunna vara moraliskt rätt.

Den generella lärdom som man kan dra av detta resonemang är: om konformitet med normer som är avpassade efter tillstådesvarande betingelser — företrädesvis den motivationsmässiga — skall kunna garantera en handlings moraliska halt, måste inpräntandet av dessa normer (med rätta) betraktas som det mest effektiva, realiserbara medlet att producera att mål bestående i *ett tillstånd vilket det är moralagenters skyldighet att förverkliga så långt det står i deras makt*. Givet förutsättningen att vad som står i en moralagents makt bl a är en funktion av hans moralmotivations styrka, kan man av den givna målspecifikationen härleda, att en agent som agerar i samklang med normer, vilka tänjer hans moralmotivation till det yttersta, gör sin skyldighet. Den nämnda förutsättningen är emellertid i högsta grad dubiös: om den håller streck, kan människor reducera sina etiska plikter genom en omoralisk vandel som stegrar deras depravering och moraliska oförmåga, vilket ter sig paradoxalt. Mer naturligt är då att tänka sig saken som idealmoralisten föreslår: om moralisk degeneration sätter in, tvingas moralens väktare att se genom fingrarna i större utsträckning än tillförne, alltmedan de moraliska plikterna står fast.

Jag konkluderar därför att standardbeskrivningen av s k överkvalificerade handlingar är oacceptabel, när den grundar sig på att moraliska

plikter eller skyldigheter måste definieras på ett *realmoraliskt* sätt, som avhängiga den förhärskande moraliska dådkraften. I stället måste dessa plikter och skyldigheter fixeras enligt *idealmoraliska* riktlinjer som negligerar motivationssvagheter, med påföljd att den som undandrar sig en överkvalificerande handling handlar felaktigt och klandervärt (om än i mindre grad) och att den som inte smiter från den ådagalägger den *övergensnittliga* styrkan i sin moralvilja genom att även i dessa bistra omständigheter genomföra vad hans moral begär.

Avslutningsvis (s 34—35) tar Ryding upp den typ explication jag just har linjerat. Han avfärdar den med anmärkningen att människor som regel inte spontant känner dåligt samvete vid tanken på att de själva eller indignation över att andra är ur stånd att prestera överkvalificerande dåd och att en etik som saknar ”emotionell resonans” inom en själva blir ”skugglik”. Min utläggning av distinktionen mellan ideal- och realmoral bemöter delvis denna invändning. Folk i allmänhet kan vara helt okänsliga för det dåliga samvete och den indignation som Ryding talar om, helt enkelt därför att de saknar den reflektionsförmåga som tarvas för att urskilja idealmoralen bakom realmoralen. Däremot tror jag inte — och nu extrapolerar jag från egna upplevelser — att moraliskt insiktsfulla människor kan undgå att erfara åtminstone en lätt anfäktelse av dåligt samvete och en viss brist på aktning gentemot sig själva och andra när deras otillräcklighet inför överkvalificerande uppgifter avslöjas, såvida de inte är indifferentia även gentemot grövre moraldefekter. □

Erik Ryding

Parfit och den personliga identiteten

I sin uppsats "Proust och personlig identitet" (FT 1:1986) gör Hans Mathlein i anslutning till Derek Parfit en distinktion mellan absolut och relativ identitet. Och enligt hans mening är personlig identitet av den senare, graderbara typen. Emellertid tillägger han: "Eftersom vi ständigt förändras fysiskt och psykologiskt är det egentligen inte *identitet* vi är ute efter, när vi försöker ta ställning till om vi har att göra med samma person över tiden." Det kan verka en smula förvirrande — den relativa identiteten är egentligen inte alls någon identitet. Några konsekvenser för Mathleins resonemang får denna terminologiska inadvartens nu knappast, men den tycks mig ha intresse som en reflex av den oklarhet som hos Parfit råder kring identitetsbegreppet.

I *Reasons and Persons*, Parfits huvudarbete, heter det att personlighetens enhet framförallt beror på två minnesrelationer. 1) Vi har direkta hågkomster av upplevelser i det förflutna. Det mesta är visserligen bortfallet men vissa händelser erinrar vi oss. 2) Det sträcker sig en sammanhängande kedja av minnen från varje tidpunkt i det förflutna (med undantag för de tidigaste barnåren) och till nuet. Förmodligen minns jag inte nu vad jag var med om en viss dag för trettio år sen, säg den 12 augusti. Men den 13 augusti 1956 kom jag ihåg en mängd ting jag upplevt föregående dag och den 14 augusti en rad upplevelser från den 13e etc. Denna relation kallar Parfit psykisk kontinuitet.

Av de båda minnesrelationerna sägs den första — den direkta erinringen — vara viktigast både från praktisk och teoretisk synpunkt, men det är i alla fall kontinuiteten som åstadkommer den personliga identiteten. Motiveringen är att identitet utgör en transitiv relation, och kontinuitet men inte direkt erinring är transitiv. Om dag ett är kontinuerligt förbunden med dag två och dag två med dag tre, så måste också dag ett och dag tre ha denna typ av förbindelse. Däremot kan jag komma ihåg händelser, som inträffade en viss dag för tio år sedan utan att för den skull nu komma ihåg, vad jag den dagen mindes från ännu tidigare ske-

den i mitt liv (s 205 f).

Mot bakgrunden av dessa förklaringar måste man emellertid känna sig frågande inför kommentaren till ett tankeexperiment som Parfit utför några sidor längre fram. Han antar att en nervkirurg efter behag kan manipulera en persons minnen och karaktär och har bestämt sig för att göra sitt offer till en andlig avbild av Napoleon. Förvandlingen sker i en mängd omgångar, så att individen efter varje ny behandling bara är ett litet stycke närmare målet än efter den föregående. Parfit förklarar nu att den manipulerade — vi kan kalla honom A — under behandlingens första fasen utan tvekan fortsätter att vara A men vid dess slut lika tveklöst inte är det. Någonstans under vägen upphör den ursprungliga personen att existera, fast vi inte kan säga exakt *när* lika litet som vi kan säga när en sandhög som berövas ett korn i taget upphör att vara en hög. Vid mitten av proceduren finns det över huvud taget inte något bestämt svar på frågan, om A fortfarande är i livet eller ej (s 231 ff).

Framställningen är inte helt klar. Den kan tolkas *så*, att kirurgen från varje uppnått stadium i metamorfosen *direkt* fortsätter till nästa, eller *så*, att han efter varje behandlingstillfälle återställer offret i ursprungligt skick och alltså var gång startar om från utgångsläget. Enligt den första tolkningen kommer det att råda kontinuitet genom hela proceduren. Eftersom varje enskilt ingrepp medför en obetydlig förändring, kommer försökspersonen att på varje stadium minnas åtskilligt av vad han tänkt, känt och förnummit på det närmast föregående. Då Parfit likväl säger att A inte existerar i slutfasen, så innebär det att han här övergett kontinuiteten till förmån för den direkta erinringen, när det gäller att fastställa, vad som är en och samma person; att den manipulerade individen i slutfasen inte längre sägs vara A måste framförallt bero på att han inte har kvar några av A:s minnen.

Enligt den andra tolkningen bryts däremot kontinuiteten mot slutet. Individen förs i ett språng från sin begynnelsepersonlighet till Napoleonstadiet och blir därmed en ny person.

Emellertid är ju slutprodukten exakt densamma vare sig den ena eller andra metoden har använts; likväl skall den enligt kontinuitetskriteriet i det ena fallet vara identisk med A, i det andra vara en ny människa. Läkaren tar livet av A genom att mellan turerna låta honom återvända till sin forna personlighet. Intuitivt verkar slutsatsen svår att acceptera.

En granskning av vad minneskontinuiteten innebär ger också vid handen att den är otillräcklig som grund för personens enhet. Den inte

bara tillåter så stora förändringar att det inte finns några hågkomster som är gemensamma för olika tidsfaser; den stämmer också dåligt med vad som avses, då man tänker på sig själv som samma person från födelsen och till nuet. Den uppfattningen skulle inte rubbas, om t ex en chock raderade ut de närmaste tolv timmarna ur minnet. Att man upphört att existera under den "borttappade" tiden skulle betraktas som ett bisarrt påstående.

Över huvud taget förefaller det mig svårt att förstå, vad Parfit menar med "personlig identitet". Det kan ju inte vara fråga om strikt identitet, eftersom en människa aldrig är exakt lik från minut till minut. Men inte heller är det bara ett annat namn för minneskontinuiteten. Detta framgår av ett annat av de många tankeexperiment, som fyller *Reasons and Persons*. En individ antas bli fördubblad genom att dela sig som en amöba. Båda de sekundära individerna är beträffande minnen, karaktär, attityder etc direkta fortsättningar av den ursprungliga, odelade, men, sägs det, ändå inte identiska med honom, eftersom de inte är identiska med varandra och identitet är en transitiv relation. Något mer än blott kontinuitet skall den personliga identiteten alltså innebära — men vad? Det enda vi får veta är, att den dessutom skall vara transitiv. Detta är ju självklart i fråga om den strikta identiteten. Om det för varje egenskap gäller att *a* har den, om och endast om *b* har den, och samma förhållande råder mellan *b* och *c*, så måste också *a* och *c* ha fullständig egenskapsgemenskap. Men någon motsvarande överförbarhet följer inte ur de sammanhang som skall konstituera den personliga identiteten. Två företeelser kan båda vara kontinuerligt förbundna med en tredje utan att vara det sinsemellan. Mina ben är kontinuerligt sammanhängande med min bål utan att för den skull hänga samman med varandra. Fordran på transitivitet kommer alltså som ett fristående tillägg, för vilket Parfit inte ger någon motivering.

Gissningsvis har han inte velat helt släppa termen "identitet", eftersom de allra flesta människor har en djupt rotad föreställning om att de är samma varelser livet igenom, och sedan har *ordet* dragit med sig transitivitetskravet. Kanske medverkar också förnimmelsen av att de sammanhang, som enligt *Reasons and Persons* håller samman personligheten, inte kan motivera vår normala attityd gentemot framtiden. Det heter att man i stället för "Jag kommer att vara död" bör säga "Det kommer inte att finnas några framtida erfarenheter, som på visst sätt är relaterade till mina nuvarande", och i stället för "Jag kommer att lida" bör säga "Det kommer att finnas lidande, som på visst sätt är relaterat

till mina nuvarande erfarenheter” (s 281 f). De relationer som åsyftas är givetvis minnesrelationerna. Skillnaderna mellan att jag kommer att drabbas av plågor och att en främmande gör det, skulle alltså vara att offret i det ena fallet har minnen i första person av en del av mina nuvarande upplevelser och i det andra inte är utrustad med sådana. Det är svårt att förstå, varför man under dylika omständigheter speciellt skall frukta smärtor som drabbar en själv, helst som man — vilket Parfit på ett annat ställe medger — vid starka akuta plågor knappast lär syssla med att minnas sitt förflutna.

Förhållandet kan belysas med ett tänkt fall: En man får det dystra beskedet att han inom en snar framtid skall utsättas för svåra fysiska plågor och att han när smärtan släppt kommer att finna att han har förlorat minnet. Ifall Parfit har rätt, kommer tidpunkten för minnesbortfallet att bli avgörande för frågan, *vem* det är som råkar ut för lidandet. Inträffar den förändring i hjärnan som utplånar minnet i början av plågoperioden eller först vid dess slut? I det senare fallet är den drabbade samma person som fick meddelandet, i det förra någon annan. Visserligen aktualiseras hur som helst inga minnen, eftersom medvetandet är helt fyllt av stundens pina, men man är densamme, om man skulle kunna göra det, dvs om minnesmekanismen är intakt. (Inte heller i normala fall aktualiserar vi ju ständigt personliga minnen.) Emellertid blir upplevelsen precis densamma, vare sig den plågade skulle kunna minnas eller ej, vilket betyder att det för den som får beskedet är likgiltigt, om det blir han själv eller en annan person, som utsätts för lidandet.

Kanske skulle Parfit vilja svara, att vi inte *har* något skäl att för egen räkning frukta kommande olycksöden (och naturligtvis då inte heller att glädja oss åt behagligheter, som väntar i framtiden). Människorna skulle alltså, så långt man kan följa vittnesbörden tillbaka, i onödan ha bekymrat sig för sin egen framtid.

Det kan vara frestande att utan vidare avvisa en åsikt, som så radikalt strider mot vår naturliga inställning och som inte har någon chans att bli tillämpad i praktiken — man vågar nog med bestämdhet påstå att inga filosofiska teorier om personlighetens identitet kommer att få folk att sluta upp med att hysa ett alldeles särskilt intresse för sina privata framtidsutsikter. Emellertid kan även en outrotlig attityd bygga på en illusorisk grundval, och om man förkastar Parfits teori, bör man kunna ange något alternativ. Finns det alltså någon annan faktor än minnesrelationerna, som gör det rimligt att säga att den som disponerar en viss

kropp är en och samma person vid olika tidpunkter?

Några filosofer såsom Bernard Williams och Thomas Nagel vill göra själva den fysiska kontinuiteten (hela kroppens eller hjärnans) till kriterium på personlig identitet, men jag skall här inte diskutera detta förslag, mot vilket jag tror att avgörande invändningar kan riktas, utan skall i stället se på situationen från upplevelsesidan.

Hur vet en person att det är just han som i det innevarande ögonblicket känner smärta eller erfar vällust eller har en viss tanke? Frågan verkar besynnerlig eller åtminstone onödig. I och med att något uppträder som direkt erfarenhet är det inte möjligt att betvivla att man själv är den som har denna erfarenhet. Minnesrelationerna spelar här ingen roll. Jag behöver inte uppväcka några minnen för att kontrollera att det verkligen är i mitt medvetande som en förnimmelse eller tanke uppträder, för mitt medvetande vid en given tidpunkt är = det som direkt upplevs vid denna tidpunkt. Även den som drabbas av en amnesi, som eliminerat alla personliga minnen, kan säga t ex "Jag har ont". Jagbegreppet och begreppet "direkt upplevelse" hänger samman. I enlighet med detta kan vi i stället för "Jag kommer att känna smärta" säga "Det kommer att finnas en direkt upplevelse av smärta". Invändningen att andra ju kan känna smärta utan att jag gör det, kan undgås genom att man bygger ut satsen till "Det kommer att finnas en direkt upplevelse av smärta, även om alla organismer förintas med undantag för den, vars hand just har skrivit orden "har skrivit orden". Hänvisningen till kroppen bör emellertid bara uppfattas som ett enkelt sätt att ange den individuella medvetandeströmmens avgränsning. Det är ett empiriskt faktum att, såvitt vi vet, varje medvetande är förbundet med en organism och inget med mer än en. Men den kroppsliga anknytningen följer inte analytiskt ur begreppet "medvetande". Och det väsentliga är här att det förekommer separata medvetandeströmmar, dvs serier av direkta upplevelser (förnimmelser, tankar, känslor etc).

Vad jag försöker visa på kan också illustreras på ett annat sätt: Antag, att en persons minne bara sträcker sig över en dag, när det gäller hans privata upplevelser. Varje morgon när han vaknar, är alla personliga hågkomster utplånade. Men i gengäld skriver han varje kväll en dagbok, där han noggrant noterar allt han upplevt under dagen. Anteckningarna skulle då ge en bild av medvetandeströmmen, och min tes är att fast den minneshandikappade inte kommer ihåg något av det han tidigare skrivit ner, så är det i alla fall samma person, som dag för dag skildrar sina tankar, känslor etc. Att han kommer att uppleva smärta

blir, förutsatt att han för protokollet ärligt, ekvivalent med att det kommer att stå en uppgift om en smärtupplevelse i dagboken, och får han veta att han nästa dag skall utsättas för plågor, har han inte mindre skäl att känna fruktan än personer med normal minneskapacitet.

Givetvis kan minnesspannet tänkas reducerat ner till den gräns, där det bara omfattar enstaka upplevelser.

Jag tror att Parfit utan att göra saken klar för sig använder de individuella upplevelsevärldarna om inte som identitets- så åtminstone som disparitetskriterium. I det förut nämnda "amöbafallet" förutsätts som någonting självklart att de sekundära individerna är olika personer, ehuru de till en början har identiska minnessammanhang; båda är ju kontinuerligt minnesförbundna med ursprungsindividerna. Denna deras inbördes självständighet får, eftersom Parfit inte är anhängare av ett rent fysiskt identitetskriterium, förklaras med att de har olika medvetandeströmmar: den ene är inte direkt medveten om vad den andre tänker, känner och förnimmer.

Den enklaste, om också kanske inte enda möjliga, förklaringen till medvetandenas avgränsning från varandra är att våra upplevelser är relaterade till olika iakttagande subjekt, och det skulle i så fall vara dessa, som utgjorde den yttersta grunden för personernas enhet. Intro-spektivt går visserligen inte några sådana centra att upptäcka, men detta är, som Broad framhåller, inte något argument mot deras existens. Att något iakttages skulle enligt denna teori betyda att det står i en viss asymmetrisk relation till subjektet, och ingenting kan stå i en asymmetrisk relation till sig själv (*The Mind and its Place in Nature*, s 586).

Med det sagda har jag naturligtvis inte velat bestrida, att minnesrelationerna har stor betydelse för att hålla samman personen. Utan dem — framför allt utan de direkta minnena — skulle vi inte kunna göra oss någon bild av oss själva, inte tillägga oss några individualiserande egenskaper. Man skulle även i sina egna ögon bli en anonym varelse — fortfarande ett jag men inte en personlighet.

Litteratur

Broad, C D: *The Mind and its Place in Nature*, London 1949.

Mathlein, H: "Proust och personlig identitet", *Filosofisk tidskrift* 1:1986.

Parfit, D: *Reasons and Persons*, Oxford 1984.

Williams, B: *Problems of the Self*, Cambridge 1973.

Björn Eriksson

Artisten och djuren — en filosofisk cirkus

Alltsedan Bentham har utilitaristerna i sin hållning till djuren haft ett propagandistiskt vapen mot de tätt duggande anklagelserna för "okänslighet". I motsats till rättighetsetiker och plikttänkare, som ofta bygger sina idéer om etisk relevans på rationalitet och förmåga till begrepps-bildning, har utilitaristerna hävdad att "frågan är inte, kan de tänka? eller, kan de tala? utan, kan de lida?" Men om man får tro Anders Tolland i "Människan över djuren" (FT 2/86) är det slut på det jämlika förhållandet mellan utilitarister och djur.

Det är emellertid inte helt lätt att skönja strukturen i Tollands argumentation. I det följande ska jag försöka klarlägga och kritisera denna.

På en armslängds avstånd tycks argumentet vara detta: (i) Handlings-utilitarismen är korrekt. (ii) Denna kan dock inte tillämpas direkt, utan talar i stället för att det i vårt samhälle bör råda en moral uppbyggd på prima facie förpliktelser. (iii) Dessa utilitaristiskt underbyggda förpliktelser kommer att vara delvis artistiska.

Vissa uttryck i denna argumentation måste klargöras. Vad menar Tolland med: "handlingsutilitarism", "prima facie förpliktelser" och "artism"?

Handlingsutilitarism. "De grundläggande etiska normerna", skriver Tolland, "måste vara av akttteleologiskt slag, enligt min mening eudaimonistiskt". Eudaimonism föredras av Tolland då han tycker att hedonismen lider av att vara associerad med en "simpler njutningsfilosofi". Han anser vidare att eudaimonism, till skillnad från hedonism, knyter (den icke-simpla) njutningen till personers upplevelser och tar fasta på "att man inte ska se upplevelser som någon slags självständiga enheter som kan klumpas ihop hur som helst".

Nu är det inte särskilt meningsfullt att formulera en handlingsutilitarism som inte låter sitt/sina intrinsikala värden kvantifieras. Tolland menar uppenbarligen att lycka kan summeras, dock inte "hur som helst". Vad som menas med "hur som helst" framgår inte — kanske

bara att man måste räkna rätt?

Till de varelser som kan ha den värderade formen av lyckoupplevelser räknar Tolland även "skurkar och grisar".

Den Tollandska akt-teleologin kan då kanske formuleras på detta sätt: En handling H är rätt om och endast om H leder till minst lika stort (inte hur-som-helst-summerat) överskott av icke-simpel njutning över lidande som varje alternativ.

Tolland inser, helt riktigt, att det enligt denna teori kan vara rätt att utplåna människorna till djurens gagn.

Prima facie förpliktelser. Sedan Sidgwick har utilitaristerna i allmänhet varit klara över att moralen sedd som deliberationsmetod, motivationsmönster eller handlingsdispositioner är underordnad den utilitaristiska normen: Vi bör ha de motivationsmönster och använda de deliberationsmetoder som leder oss att åstadkomma de bästa konsekvenserna. Dessa mönster kommer knappast att vara en direkt tillämpning av handlingsutilitarismen. Skälen till att utilitarismen inte anses optimal i sina egna termer är främst följande:

1) Konsekvenskalkylen i en situation är extremt invecklad. Vår brist på kunskap och tid gör handlingsutilitarismen olämplig som deliberationsmetod. Hare gör i *Moral Thinking* reflexionen att få utilitarister skulle komma med sina bussar om de direkt tillämpade sitt etiska system vid varje beslut. Denna kalkylens krånglighet talar för en *enklare* deliberationsmetod, inte något annat riktighetskriterium och inte metoder som skulle vara onödigt icke-utilitaristiska.

2) Vår tendens att bry oss om oss själva lite mer än andra skulle lätt bestämma resultatet av den komplicerade kalkylen — om man nu skulle räkna fel, så varför inte till sin egen glädje? Detta innebär att deliberationsmetoden bör utgöra ett hinder för egoistiska snedvridningar.

3) Det är ett numera välkänt faktum att många handlingar små inte alltid gör en stor å. Även om varje agent i någon grupp utför den bästa handlingen så kan konsekvenserna av gruppens handlingar vara sämre än om gruppen hade handlat på ett annat sätt. Konsekvenserna av våra handlingar är delvis beroende av vad andra gör. Handlingsdispositioner och deliberationsmetoder bör främja en koordinering av handlingar så att gruppens handlingar närmar sig optimum.

Tolland tycks hysa uppfattningen att prima facie förpliktelser bör ses som ett slags tumregler för undvikande av ovan nämnda problem. Denna syn på vardagsnormer och prima facie förpliktelser är förmodligen den enda rimliga för en utilitarist.

Artism. Tolland skriver: "Artism innebär alltså att man diskriminerar djuren enbart på grund av deras arttillhörighet." Artismen som här formulerad är ett beteende, men tanken är säkert att artismen innebär att det inte är fel att diskriminera djur enbart på grund av deras djuriskhet. Det är alltså inte tal om att djurens lidande är mindre, eller mindre viktigt, eller att det är några egenskaper hos alla djur utom deras egenskap att vara icke-människor som på något vis rättfärdigar diskriminering.

Vägen till artismen

Jag har hört sägas att utilitarismen är tråkig: Man samlar bara in vissa fakta och vevar dem genom den "utilitaristiska kvarnen" så har man svaret — det finns inget kvar att resonera om. Tollands behandling av utilitarismen är knappast tråkig i denna mening.

När Tolland bekänner sig till den utilitarism som skisserades inledningsvis kunde man tro att han i denna skulle stoppa in vanliga människor, grisar och skurkar, veva ett varv och ut skulle artismen komma. Men detta händer inte. Hur skulle det kunna hända? Han ser ju att utilitarismen, till och med hans egen version, inte ger det resultatet. Det enda svar som ges är att det är en empirisk fråga om det i någon situation är djuren snarare än människan som bör diskrimineras. Då Tolland är ute för att bryta en lans för artismen har han hamnat i ett dilemma. Hans lösning är lika enkel som förvirrad. "Jag kan helt enkelt inte acceptera att människor offras för djurens skull", skriver han. Och nu omvandlas denna intuition till ett argument mot utilitarismen: "Det kan mycket väl vara en icke-artistisk akt-teleologs objektiva plikt att verka för att människosläktet så gott som utplånades för djurens skull. Detta kan jag inte acceptera och därför bekänner jag mig till artismen." Efter ett kort mellanspel som utilitarist har Tolland återfallit till en artistisk övertygelse utan stöd av argument. I synnerhet utan stöd av utilitarismen.

Men detta kan inte ha gått rätt till, ty i det följande tror Tolland att han fortfarande är utilitarist. Vi får försöka glömma att han förkastat utilitarismen och tro att han ansluter sig till teorin och att artismen kommer att dyka upp som en följd av denna teori på det Tolland kallar moralens "andra nivå". Detta skulle innebära att de utilitaristiskt sett bästa *prima facie* förpliktelserna skulle vara av artistisk art. I ljuset av det tidigare förda resonemanget om skälen för olika tumregler måste tanken vara att artismen skulle förhindra förlamande komplicerade över-

väganden, verka mot våra egoistiska tendenser (!), eller bidra till att koordinera handlingar. Helt kort: Jag kan inte se *något* rimligt argument för att så skulle vara fallet, varken i Tollands text eller någon annanstans.

Men det här kan inte ha gått rätt till, ty Tolland är ju artist. Ett tredje sätt att förstå hans argument är att han inte är utilitarist, utan "artistisk akt-teleolog". Artismen var nödvändig redan på den grundläggande handlingsutilitaristiska nivån. Tolland har bara skojar: det är inte ur en klassisk handlingsutilitarism artismen följer, utan artismen följer ur en artistisk utilitarism — det är denna icke närmare redovisade etik som talar för att vi bör hålla oss med artistiska *prima facie* förpliktelser. Föga överraskande. Den artistiska utilitarismen kan knappast användas som försvar för artismen då artismen knappast kan försvara sig själv.

En sak är emellertid helt klar. Tolland tycker inte att det är fel att diskriminera djuren enbart på grund av deras icke-mänsklighet. Några rimliga argument för denna hållning har inte framförts — minst av allt argument förenliga med en något så när klassisk utilitarism. Den artistiska "intuitionen" för tanken, som vanligt, till den uråldriga artistiska fördomen. □

Tollands artistiska magplask

Det är inte långsökt att tänka sig att människans artism eller artpartiskhet, hennes benägenhet att favorisera sina artfränder just därför att de är artfränder, är en biologiskt betingad instinkt som inte vilar på någon rationell grund. De egendomligaste krumsprång som möter en i Anders Tollands försök att rättfärdiga artpartiskheten i "Människan över djuren" (*Filosofisk tidskrift* 1986:2) bekräftar denna hypotes. Hans huvudtes är:

- (1) "att artismen måste accepteras på den *grundläggande* etiska nivån" (s 21).

Detta förhindrar honom inte att också vilja göra gällande:

- (2) "att de *grundläggande* etiska normerna måste vara av akt-teleologiskt slag, enligt min mening eudaimonistisk (d v s lycka är det som har egenvärde) akt-teleologi" (s 24), samt
- (3) "att åtminstone de högre däggdjuren . . . har upplevelser som är väl så behagliga eller obehagliga som några hon /människan/ har" (s 21).

Om emellertid de grundläggande etiska principerna går ut på att man skall maximera lycka och minimera olycka, och åtminstone de mera högtstående djuren är bärare av lycka och olycka, så kan människor prioriteras på dessas bekostnad på den fundamentala nivån endast i den mån detta befrämjar den totala — "mellanartliga" — maximeringen av lycka och minimeringen av olycka, och inte enbart på att de tillhör den mänskliga arten, vilket artismen kräver. Med andra ord, om (2) och (3) är korrekta, måste (1) förkastas. Det är till yttermera visso just på basis av premisser i stil med (2) och (3) som den kände förespråkaren för djurens rättigheter, Peter Singer, fördömer artpartiskhet.

Tolland tycker sig skönja en "total" konflikt mellan artpartiska och artneutrala intuitioner på den grundläggande nivån, en konflikt som han vill lösa till förmån för de förra (s 24—25). En sådan lösning förutsätter dock att Tolland kompletterar sin teleologiska eudaimonism med

en artpartisk princip som deklarerar t ex att endast människans väl och ve är etiskt relevant. Att det är denna strategi som Tolland implicit tillgriper bestyrks av hur han resonerar sig fram till lämpligheten på den s k praktiska nivån av "normer som tillgodoser även djurens intressen" (s 27). Som han själv observerar, är inte förmåga till lycko- och olycksupplevelser "avgörande för att komma i åtanke på den praktiska nivån" (s 27): människans enda skäl för att ta hänsyn till djuren, liksom till den övriga naturen, är att hon själv annars far illa.

Eftersom Tolland inte tycks klart medveten om att han bekänner sig till en grundläggande artpartisk norm, inser han heller inte behovet av att försvara den, ett behov som är speciellt trängande mot bakgrund av hans eudaimonism. När jag i min artikel "Kan djur ha moralisk inställning?" (*Filosofisk tidskrift* 1983:4) drog parallellen mellan artism, å ena sidan, och rasism och sexism å den andra, var poängen denna: det tycks lika godtyckligt att fästa större avseende vid lycka och olycka därför att de upplevs av en individ av den mänskliga arten som det är att tillmäta dem större vikt emedan de erfars av en människa av en viss ras eller ett visst kön. Ingenting i Tollands resonemang är ägnat att rubba denna parallell; därför faller det platt till marken. □

Recension

Göran Möller: *Risker och människolivets värde. En etisk analys*. Uppsala 1986.

Inom den teologiska fakulteten i Uppsala finns tre professurer som är ägnade åt specialområden inom den praktiska filosofin: Etik, religionsfilosofi och tros- och livsåskådningsforskning. De teologiska filosoferna anses ofta förtjäna beröm för att de är präster, Guds ords tjänare, och ändå inte åberopar Guds ord i sitt arbete. Alla är dock inte prästvigda, och de som inte har avlagt prästlöftet får väl hur som helst räkna med att få sitt filosoferande bedömt på samma villkor som vi andra. Till den gruppen hör Göran Möller, som har disputerat i etik, särskilt socialetik, på avhandlingen *Risker och människolivets värde*.

Den övergripande frågeställningen är väl formulerad i inledningen:

”Det finns ett slags spänningsförhållande mellan uppfattningen om människans höga värde och okränkbarhet, och vårt accepterande av att människor dödas i olyckor, vilka i viss utsträckning hade kunnat förhindras. En av huvuduppgifterna i denna undersökning är att analysera detta spänningsförhållande och att därigenom förklara hur förekomsten av vissa (icke minimala) risker etiskt kan legitimeras. . . . Syftet med denna undersökning är att klargöra och ta ställning till de etiska problem som aktualiseras vid sådana samhälleliga beslut, där risken för att människor skall förolyckas är en omständighet som är möjlig att påverka. Undersökningen mynnar ut i ett förslag till en metod som kan användas som underlag vid sådana beslut.”

Avhandlingen har två huvuddelar. Den första kallas ”Etiska grundproblem”. Den andra ”Bestämning av acceptabel risknivå”. Den första delen, som upptar ganska precis halva boken, mynnar ut i fyra normer, av vilka de mest centrala, om man förkortar formuleringarna något, säger:

- (A) Varje oskyldig människa har en negativ, absolut rättighet att inte bli avsiktligt dödad.
- (B) Det är tillåtet att handla på ett sätt som kan leda (och leder) till att andra dödas, om man gör det genom att åstadkomma vissa begränsade risker, och åstadkommer dem i syfte att uppnå vissa legitima värden.

Syftet med den första delen är att visa att dessa normer är förenliga med varandra och att relatera dem till vissa andra normer, som Möller åtminstone ibland verkar uppfatta som mer fundamentala. De andra normer som framför allt är aktuella är: (1) Kants princip att man aldrig får använda andra människor blott som medel, samt (2) en princip som säger att alla människor (människors liv) har samma värde och okränkbarhet därför att människan har förnuft, och vida-

re (3) principen om dubbel effekt, som skiljer mellan två sorters negativa konsekvenser av handlingar: dels sådana konsekvenser som är medel för eftersträ- vade mål i den meningen att målet uppnås *via* konsekvenserna i fråga, dels så- dana konsekvenser som är icke eftersträ vade, men insett oundvikliga. Den för- ra sortens konsekvenser antas ha större relevans än de senare när handlingen skall moraliskt bedömas. Väsentlig roll spelar också idén om en "riskpool" (lä- nad från Fried) och exemplet "överlevnadslotteriet" (Harris). Riskpoolen inne- bär att den enes handlingar, som utsätter andra för risker, legitimeras av de and- ras handlingar, som utsätter den ene för risker, allt för att uppnå legitima mål. Överlevnadslotteriet går ut på att ett antal personer sluter ett avtal om att när ett visst antal av dem skulle kunna räddas till livet genom att en av dem avlivas och används till transplantat, så väljs slumpmässigt en av de anslutna ut och av- livas och fördelas som transplantat på de behövande.

Kants princip sägs vara "en grundläggande premiss i undersökningen", men så förhåller det sig vid närmare granskning inte. Det framgår av Möllers reso- nemang att han förstår Kants princip på ett sådant sätt att den a) förbjuder av- siktligt dödande, b) är förenlig med riskpoolsresonemanget men c) förbjuder överlevnadslotteriet. Varför kan inte överlevnadslotteriet, åtminstone i vissa automatiserade varianter, uppfattas som en riskpool? kan man fråga. Något explicit svar på den frågan kan man inte finna i Möllers bok, men man kan räk- na ut, och det bekräftades av Möller vid disputationen, att svaret hänger på den distinktion som görs i principen om dubbel effekt: I överlevnadslotteriet är off- rets död ett *medel* för de andras överlevnad i den meningen att överlevnaden uppnås *via* offrets död, men i acceptabla riskpools är offrets död en icke efter- strä vad bieffekt. För att Kants princip skall kunna spela den roll den antas spela måste alltså "medel" förstås i denna mycket speciella mening. Men även om vi tolkar "medel" på detta särskilda sätt är det ju inte klart att Kants prin- cip förbjuder överlevnadslotteriet, eftersom principen förbjuder att människor används *enbart* som medel, och det inte är klart att den avlivade *enbart* är ett medel — han har ju dock genom lotteriet fått fördelar, åtminstone i termer av *chans* till räddning, och kanske också faktiskt, genom att först själv ha räd- dats från en tidigare död genom lotteriet. Vidare är det ju inte heller klart att Kants princip skulle förbjuda direkt avsiktligt dödande av oskyldiga, och det av samma skäl: principen förbjuder ju att någon används *enbart* som medel, inte att någon överhuvud används som medel.

Om man inte utan vidare accepterar Möllers egna påståenden om vad som är grundläggande premisser i hans undersökning utan i stället tittar på hur hans re- sonemang faktiskt ser ut, tycks alltså inte Kants princip spela någon roll. Om man skulle försöka formulera den princip som Möller de facto verkar utgå från när han hänvisar till Kant, så skulle den snarare lyda:

- (C) Det är ovillkorligen förbjudet att använda andra oskyldiga människors död som medel i den meningen att man dödar dem för att via deras död uppnå värden.

Och det är viktigt att det handlar om medel i denna mycket speciella mening. För det är enligt Möller inte ovillkorligen förbjudet att använda andra oskyldiga

människors död som medel i den meningen att man dödar dem som en förutsedd bieffekt av en handling som man utför för att uppnå vissa legitima värden. I själva verket är det skillnaden mellan dessa fall som löser det problem som Möller från början ställer, nämligen hur man skall förena tanken på livets okränkbarhet med accepterandet av att människor dödas t ex i trafiken. Det är med andra ord principen om dubbel effekt som gör det egentliga jobbet här, fast Möller inte explicit framställer det så. Han föredrar, och det är kanske på sitt sätt begripligt, att flagga med Kant.

Den andra av de ovan nämnda principerna, den som säger att grunden för människans rättighet att inte avsiktligt dödas är hennes förnuft, framställs med en precision som lämnar mycket att önska. Möller talar omväxlande om en människas värde, värdet av hennes liv och hennes livs okränkbarhet utan att göra någon tydlig skillnad däremellan. Uppenbarligen menar vi ibland när vi talar om värdet av en människas liv värdet av att hon lever och ibland i stället värdet av det liv som hon lever. (Och vi kan ha olika möjliga uppfattningar om hur det ena förhåller sig till det andra.) När vi frågar vad en människa är värd kan det dessutom ibland i stället vara fråga om vad hon rimligen kan kräva, vilka rättigheter hon har (analogt till i "arbetaren är värd sin lön"). Om man gör dessa ytterligt elementära distinktioner är det klart att regeln "Om X har förnuft, så har X värde" liksom principen "Om X har värde och är oskyldig, så har X rätt att inte avsiktligt dödas" kan innebära flera olika saker. Något besked om hur Möller egentligen har tänkt här går inte att utläsa ur hans text. Om man framträder med anspråk på att göra en etisk analys får man inte på detta sätt brista i det elementäraste.

Dessutom blir hela företaget att grunda människans rättighet att inte avsiktligt dödas på hennes förnuft tämligen svårförståeligt, när Möller samtidigt säger att de människor som inte har förnuft har samma rättighet. Om rättigheten tillkommer de icke förnuftiga genom "en akt av kärlek och solidaritet", som Möller säger, varför inte då låta även de förnuftiga få sin rättighet på samma sätt?

På liknande sätt löper Möllers resonemang i fråga om riskpoolen. Att vi utsätter andra för risker kan rättfärdigas genom att de utsätter oss för risker. Men hur är det med de risker vi utsätter framtida generationer för? Möller leker med tanken att de risker de i sin tur utsätter sina efterkommande för kan berättiga tal om någon sorts ömsesidighet, men inser att det inte låter särskilt övertygande, och antar i stället att de risker vi utsätter våra efterkommande för rättfärdigas av de positiva saker vi också lämnar efter oss. Men om vi kan resonera så om framtida människor, varför inte om samtida? Och om vi genom att peka på de positiva saker vi ger andra samtida kan försvara att vi utsätter dem för risker, då behövs ju resonemanget om riskpool överhuvud inte. Och det vore väl en fördel? Riskpoolsresonemanget verkar ju överhuvud inte vidare övertygande i detta sammanhang. Det kan ju inte sällan förefalla högst tveklaktigt om t ex ett trafikoffer — betrakta en överkörd gammal gumma på landet som sällan satt sin fot på trafikerad väg — i någon rimlig mening frivilligt anslutit sig till någon riskpool eller utsatt andra för några risker till liv och lem.

De huvudresonemang som förs i avhandlingens första del är alltså varken hållbara eller intressanta. Framställningen skäms dessutom av många brister i detaljerna.

Möller tror t ex enligt vad han säger på s 50 att om livets värde skulle vara lexikaliskt överordnat alla andra värden, så måste livet ha ett oändligt stort värde. Det behöver det förstås inte ha.

I samma veva säger han att om livets värde vore lexikaliskt överordnat alla andra värden, så skulle det följa att det aldrig var rationellt för en människa att ta risken att dö för att uppnå andra värden. Men nu är det rationellt att ta sådana risker ibland, därför är livets värde inte lexikaliskt överordnat. Av detta får man dock inte dra den slutsatsen, säger han, att det någon gång kunde vara rationellt att döda en människa för att förhindra att andra värden gick förlorade, och att man *inte* får dra den slutsatsen beror på att man måste skilja mellan normer och värderingar. "Det är nämligen inte tillräckligt att betrakta de värdemässiga konsekvenserna av en handling för att kunna avgöra om handlingen är rätt eller orätt." (51). Men nyss har ju Möller antagit att det av vissa värdemässiga förhållanden nämligen lexikalisk överordning, följer vissa deontiska. Uppenbarligen måste alltså antingen den förra eller den senare delen av resonemanget vara felaktig, de är klart oförenliga.

I samma avsnitt finns också ett egendomligt resonemang om inkommensurabilitet. Möller förklarar inkommensurabilitet hos *termer* som "att det inte finns någon enhetlig skala med hjälp av vilken man kan uttrycka båda termerna" (51), exemplifierar med *tal*, och övergår till att tala om inkommensurabla *värden*: "alla värden i en värdepluralistisk teori är inkommensurabla med varandra. Vore de inte det skulle ju teorin kunna beskrivas med hjälp av ett färre antal (!) värden" (52). Men så är det förstås inte. Att man t ex menar att både lust och kunskap har egenvärde behöver naturligtvis inte innebära att man menar att lust och kunskap inte kan vägas på ett bestämt sätt mot varandra. (Och på sitt vanliga sätt rycker Möller också längre fram i boken (124) undan vad man trodde vara en förutsättning för resonemanget, nämligen att inkommensurabla värden inte kan vägas mot varandra: "Vi tvingas sålunda gång på gång väga inkommensurabla värden mot varandra för att kunna avgöra vilket eller vilka som bör ha företräde i ett visst sammanhang.")

Samma lättsinniga inställning till de egna definitionerna visar Möller i diskussionen om skillnaden mellan aktiv handling och underlåtelse (56). Han gör en kraftig inskränkning i begreppet underlåtelse: "När vi i fortsättningen använder begreppet underlåtelse kommer vi inte att inbegripa fall där underlåtelsen är en avsiktlig del i P's plan." Och det må förstås stå honom fritt. Men han förefaller ha glömt den inskränkningen redan i nästa stycke: "Det finns en distinktion som är analog till den mellan handling och underlåtelse, nämligen distinktionen mellan positiva och negativa plikter. De positiva plikterna uppmanar oss till att utföra vissa handlingar, medan de negativa plikterna förbjuder oss att handla på visst sätt." Uppfattar man underlåtelse på det nyss föreslagna snäva sättet har man ju ryckt undan grunden för analogin.

Jag har redan nämnt den egendomliga Kantläsning som Möller demonstrerar. Även modernare författare återger han på mindre lyckade sätt. Han visar sig t ex ha fått Dworkin alldeles om bakfoten. Möller tror att hans eget resonemang om förnuftet som grund för människans värde och rättigheter gör att han uppfyller Dworkins krav att "... varje rättighetsbaserad teori måste förutsätta rättigheter som inte rätt och slätt är en produkt av övervägd lagstiftning eller be-

stämda sociala sedvänjor, utan är en oberoende grund för att bedöma lagstiftning och sedvänjor" (Möllers översättning). "Genom att vår normativa teori ytterst kan motiveras av att människan är utrustad med ett förnuft och därmed besitter de olika kvaliteter som detta ger upphov till, så uppfyller den det krav som Dworkin ställer. Dessa egenskaper är nämligen inte något uttryck för konventioner och föreställningar utan de är högst reella."

Detta är en grov missuppfattning. Om Möllers teori uppfyller Dworkins krav eller inte, det beror ju på om de normer som tillerkänner människor rättigheter i Möllers teori uppfattas som konventionella eller inte.

Varken resonemangen i stort eller detaljarbetet är tillfredsställande i bokens första del. Inte heller i den andra, som heter "Bestämning av acceptabel risknivå".

Den delen resulterar i ett "kriterieschema", som Möller kallar det, vilket är en disparat samling påståenden om riskers värde och vad som kan och bör accepteras ifråga om risker. "Icke frivilliga risker bör inte vara alltför koncentrerade till vissa människor eller vissa grupper av människor" lyder t ex kriterium 5. Och kriterium 10 säger: "En risk som riktar sig mot framtida människor är lika allvarlig som en motsvarande risk, riktad mot nu levande människor." Kriterieschemat antas kunna användas av samhälleliga beslutsfattare (osagt på vilken nivå) för att fatta beslut i riskfrågor. Om de olika "kriterierna" i ett enskilt fall kommer i konflikt med varandra så skall beslutsfattaren väga dem mot varandra efter eget skön. Detta kallar Möller "den reflektiva jämviktens metod".

På ett tidigt stadium i andra delen sägs: "När vi i det följande använder oss av uttrycket acceptabel risknivå menar vi den risknivå som föreligger när det råder jämvikt, dvs att de positiva konsekvenserna precis väger upp riskens negativa värde" (86). Detta är ju närmast formulerat som en definition eller meningsförklaring, men så är det nog inte tänkt, utan avsikten torde vara att ge ett generellt kriterium på acceptabilitet hos risker. Detta kriterium specificeras ytterligare på de närmast följande sidorna:

"För att kunna avgöra om en viss risk är acceptabel eller ej, är det därför nödvändigt att relatera riskens negativa värde till totaliteten av handlingens eller verksamhetens (positiva och negativa) värdemässiga konsekvenser. I det följande skall vi använda oss av uttrycket "det sammanlagda nettoutbytet" för att beteckna summan av de värdemässiga konsekvenser, förutom risker, som en viss handling eller en viss verksamhet ger upphov till" (88).

En risk är alltså acceptabel om det sammanlagda nettoutbytet (som i den engelska sammanfattningen kallas "net *exchange*"!) väger upp riskens negativa värde eller, enligt en annan tanke som Möller framställer som en precisering av den föregående, en riskminskning är påkallad om den minskning av det sammanlagda nettoutbytet som förändringen medför är mindre än riskminskningens positiva värde. Det låter ju inte i och för sig orimligt. Men det är knappast för enligt med vad som hävdats i bokens första del. Betrakta t ex det positiva värdet av transplanterat av livsviktiga organ från trafikoffet. Det måste rimligen ingå i det sammanlagda nettoutbytet av trafiken enligt Möllers definition. Samtidigt är det ju klart, om inte från annat så från Möllers diskussion av överlevnadslotteriet, att transplanterat av livsviktiga organ är något som man får *via* männi-

skors död. Enligt en av de viktigaste normerna i bokens första del är det alltså förbjudet att döda människor för att få transplanterat av livsviktiga organ. Men enligt resonemanget om sammanlagt nettoutbyte i bokens andra del kan det samtidigt vara så att en viss risknivå i trafiken blir acceptabel just genom värdet av transplantaten från trafikoffren gör det sammanlagda nettoutbytet tillräckligt stort. Vi skulle alltså kunna försvara att vi låter så många dö med att vi får transplanterat, samtidigt som vi förbjuder att människor dödas för att ge transplanterat. Att detta skulle vara Möllers mening förnekade han vid disputationen, men att det följer av vad han skriver är klart.

Inte heller går det att förstå hur idén om vägning av risker mot nettoutbyte skall kunna förenas med "kriterieschemat". Där heter det att man vid beslut om risker inte bör låta sig påverkas av sådana egenskaper hos dem som utsätts för riskerna som nytta för andra eller för samhället. Men antag att vi kan välja mellan att utsätta ett visst antal människor med liten nytta för andra för en viss risk och att utsätta precis samma antal människor med stor nytta för andra för precis lika stor risk. Det verkar vara en fullt möjlig situation. Om allt annat är lika blir ju då nettoutbytet större i det förra fallet, just därför att riskpersonerna är mindre värdefulla för andra och riskerna i båda fallen enligt Möllers sätt att se har samma negativa värde. Enligt idén om nettoutbyte skall vi alltså föredra det förra alternativet, medan det nyss nämnda kriteriet förbjuder oss att göra det. (Men resonemanget om nettoutbyte skall kanske trots att det ser så ut inte fattas som någon övergripande idé utan som en bland andra normer som Möller hävdar med reservation för att de kan komma i konflikt med varandra? Något som tyder på det är att de två första numren i "kriterieschemat" tycks följa ur netto-utbytesresonemanget.)

Alla frågeställningar i Möllers bok skulle kunna stå i en avhandling i praktisk filosofi. I den meningen är det en klart filosofisk bok han har skrivit. Men jag kan inte föreställa mig att de praktiska filosoferna någonstans i landet skulle kunna tänka sig att godkänna den som examensarbete i doktorsexamen. Brister ligger i alltför stor utsträckning på den nivå där en kompetent handledare skulle ha ingripit i ett mycket tidigt skede. □

Sven Danielsson

Notiser

Filosofi i näringslivets tjänst

I SAFs tidning "Chefen — specialtidning för managementfrågor" nr 5/1986 intervjuas Göran Hermerén, professor i praktisk filosofi i Lund. Bakgrunden till intervjun är att Hermerén har börjat hålla kurser i etisk analys för företagsledare. I intervjun motiverar Hermerén detta med att "filosoferna kan hjälpa företagen att systematisera och kritiskt granska etiska problem och medvetandegöra osynliga värderingar".

Här uppstår emellertid genast den klassiska frågan: Går det att sälja sina tjänster åt näringslivet men ändå samtidigt behålla möjligheten att kritisera näringslivet? För Göran Hermerén blir dock detta aldrig något problem. Redan i inledningen av intervjun deklarerar han nämligen: "Det är inte meningen att filosoferna ska moralisera eller skriva företagsledarna på näsan vad som är rätt eller fel." Med denna inställning om filosofernas uppgift finns det gott hopp för Göran Hermerén om att även fortsättningsvis få hålla föredrag för företagsledare och näringslivsrepresentanter.

Senare i intervjun visar Göran Hermerén återigen sin välvilliga inställning till näringslivet genom att fälla följande yttrande: "Man behöver inte förutsätta någon motsatsställning mellan lönsamma och etiskt goda affärer. Tvärtom, kan etiska och ekonomiska synpunkter behöva integreras för den som vill bedriva framgångsrik och långsiktig affärsverksamhet." Säg den aktiespekulant eller börsklippare som inte är nöjd med ett sådant tröstens ord!

I samma nummer av "Chefen — specialtidning för managementfrågor" behandlas för övrigt filosofi på fler ställen. I en artikel om "lästrender" meddelas att den som söker managementlitteratur inte skall glömma att söka bland de filosofiska böckerna. I ytterligare en annan artikel, en slags baksideskolumn, konstateras följande: "Den under senare tid dominerande analytiska filosofin möter uppenbarligen invändningar. Det verkar som om fackfilosofer ligger 10 år efter den allmänna debatten. Metafysiken har vunnit stora segrar inte minst inom de grupper som sysslar med managementutveckling." Här finns kanske en ny fräsch nisch för arbetslösa existentialister, hegelianer och fenomenologer.

Mats Wingborg, filosofistuderande

Biologins vetenskapsteoretiska problem behandlas i en bok med titeln *Biologisk kunskap* av Hans Mohr, Studentlitteratur.

Aant Elzinga har nu efterträtt Håkan Törnebohm som professor i vetenskapsteori i Göteborg.

Claes Edlund, Göran Hermerén och Tore Nilstun har utgivit *Tvärskap*, en bok om samarbete och kunskapsutbyte över ämnesgränser, Studentlitteratur 1986. Syftet med boken är att åstadkomma en mer realistisk attityd till tvärvetenskapligt samarbete i projektgrupper.

Från institutionen för vetenskapsteori i Göteborg (Göteborgs universitet, 412 98 Göteborg) har man börjat utge en ny tidskrift för vetenskapsstudier, VEST, som syftar till att integrera filosofiska, psykologiska, ideologiska, historiska, politiska, ekonomiska, sociologiska och andra aspekter på fenomenet vetenskap. I det första numret skriver Lennart Olausson om "Den mångtydiga universitetsideologin", Thomas Brante om "Vetenskapshistoriesociologiska abstraktioner — scener ur ett äktenskap" och Sven Andersson om projektet "Vetande och vision". Ansvarig utgivare för VEST är John Hultberg.

Montaignes Essayer Bok 1 har utgivits i översättning av Jan Stolpe, Bokförlaget Atlantis 1987.

Hans Rosing har utgivit en ny bok med titeln *Den personliga identiteten och livets mening*, Holger Schildts förlag 1986.

Medarbetare i detta nummer av Filosofisk tidskrift: *Mats Furberg* är professor i teoretisk filosofi i Göteborg, *Ragnar Ohlsson* är lärare i praktisk filosofi i Stockholm, *Ingmar Persson* är forskarassistent i praktisk filosofi i Lund, *Erik Rydning* är docent i praktisk filosofi i Lund, *Björn Eriksson* är doktorand i praktisk filosofi i Stockholm och *Sven Danielsson* är docent i praktisk filosofi i Uppsala.

Omslagsbild av *Christian Swardshammar*.

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 brev vänstermarginal (räkna med 50 nedslag/rad och med 30 rader/sida = 1 500 nedslag)
- 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 artikelförfattaren, gratis, 10 fullständiga exemplar av det nummer av tidskriften i vilket bidraget varit infört.

BOKFÖRLAGET THALES

ISSN 0348-7482