

FILOSOFISK TIDSKRIFT

4/96

MATS FURBERG

Filosofi på svenska

WLODEK RABINOWICZ *Värde grundat på preferens*

BJÖRN ERIKSSON

Utilitarism för syndare

INNEHÅLL

| | |
|--|----|
| <i>Filosofi på svenska</i> | 3 |
| MATS FURBERG | |
| <i>Värde grundat på preferens</i> | 21 |
| WLODEK RABINOWICZ | |
| <i>Utilitarism för syndare</i> | 37 |
| BJÖRN ERIKSSON | |
| <i>Recension</i> | 49 |
| ÅSA CARLSON om Anette C Baiers <i>A Progress of Sentiments</i> | |
| <i>Notiser</i> | 59 |

Filosofisk tidskrift nr 4 1996 årg 17. Filosofisk tidskrift utges av Stiftelsen Filosofisk tidskrift och Bokförlaget Thales.

Redaktör och ansvarig utgivare: Lars Bergström

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

Produktion och prenumerationer:

Bokförlaget Thales, Djupdalskvior 5, 620 12 Hemse, förlagsredaktör Ulf Jacobsen, tel 0498-48 48 93

Prenumeration: Tidskriften kommer ut med 4 nummer/år, pris för 1997 är 150 kronor, pg 50 79 91-8 Filosofisk tidskrift

Lösnummer: 35 kronor

Satt av Ord och Grafik, Stockholm

Tryckt hos Lagerblads tryckeri AB, Karlshamn, 1996

ISSN 0348-7482

MATS FURBERG

Filosofi på svenska

Titeln är flertydig. Den kan avse filosofiska *produkter* utgivna på svenska. Många av dessa är gjorda på grekiska eller latin eller franska eller engelska (eller amerikanska som knappast längre är blott en dialekt) men överflyttade till vårt mål. En första avdelning handlar därför om översättningar, deras målsättningar och problem.

Snävare kan rubriken avse filosofiska produkter tänkta på svenska och av svenskar, vilket dessvärre kan vara aktstycken mödosamt hoppplitade på engelska eller andra för skribenten hopplöst främmande språk. Endast indirekt säger jag något om denna betydelse.

Det som ligger mig om hjärtat är filosofi på svenska i en mening där båda huvudorden har omyttts. Temat är då den verksamhet, det filosoferande, som bäst sker på ens modersmål, oavsett om man talar svenska, polska, engelska eller ett icke-indoeuropeiskt språk som finska. Det är inte givet att filosoferandet avsätter skriftliga produkter och tvivelaktigt om det alltid bör göra det; Sokrates kan ha avhållit sig från plitande av andra skäl än skrivkramp.

Temat är inte blågult. Det är att filosofera på modersmålet och handlar om generella förbindelser mellan tänkande, språk och kultur. Jag månar inte om intellektuell isolationism. Men jag är bekymrad över de strävanden till internationalisering som inte spörjer sig om en filosofisk enhetskultur är en alltigenom god sak.¹

1 När ett filosofiskt äktapar ständigt gnabbas som Ulla M Holm och jag är det svårt att sära den enes bidrag från den andras; det växelverkar. Betonningen av bourdieuska habitueringar, dragningen åt praktikalism samt sneglanden åt filosofisk antropologi passar mitt gamla intresse för Wittgenstein och halvgamla för Heidegger men kommer från Ullas *Modrande och praxis* (Göteborg 1993); jag har till och med tillåtit mig återge hennes älsklingscitat från Proust om imperativer som sitter i armar och ben.

1. Att översätta till modersmålet

Att översätta en invecklad filosofisk text är alltid att tolka. Det översätta jämkas därtill efter läsarnas förutsättningar, särskilt när tilltänkt läsekrets är lekmän med kunskaper och intressen som är långt ifrån författarens. Avståndet till originaltanken kan bli betydande.

A. I den samling texter av Martha Nussbaum som har kallats *Känslans skärpa tankens inlevelse* (Stockholm / Stehag 1995) blir 'Aristotelian Social Democracy' 'socialdemokrati enligt Aristoteles' och 'putative' 'allmänt accepterad'; 'minor premise' förvandlas till 'sammansatt premiss'; 'commensurability' återges utan vidare spisning med 'jämförbarhet'; 'universal' och andra filosofihistoriska facktermer omskrivs oigenkännligt; den viktiga och kontroversiella uppsatsen "The Discernment of Perception" görs till intellektuellt men dessvärre inte språkligt nonsens av att 'perception' kallas 'förnimmelse' och att 'känsla' får täcka både 'emotion' och 'feeling', termer Nussbaum är mån om att hålla i sär.

Somliga av dessa missar gör ingen kompetent översättare. Andra vittnar mer om grav filosofisk än om grav språklig okunnighet och skulle ha avlägsnats av en fackgranskare.² Jag ryser och går vidare.

B. Anders Wedbergs och Lars Hertzbergs överflyttningar av Wittgenstein till svenska är trognare än någon av de engelska. Skälet är enkelt: språken är varandra närmre än engelska och tyska. *Vorstellung* och *wollen* vållar inte oss besvär. Men *ett* ord ställer till krångel. Det betecknar ett nyckelbegrepp i Wittgensteins sena tänkande: 'Sprachspiel'.

Blå boken hade sagt 'language game' och därmed satt mönster för översättningen. I 30-talets början, kanske under Ramseys olycksaliga inflytande, var *Sprachspiele* analoga med schack, bridge och tennis. De har konstitutiva regler, och gott utövande underlättas av strategiska (tum)regler. Däremot finns inga idéer av den art Gadamer flörtar med i *Wahrheit und Methode*, de som går ut på att *Sprachspiele* påminner om *Lichtspiele*, om ljusspel under brospann i vägglitter. *Spiele* är för Wittgenstein ting *vi* har för oss, medan Gadamer menar att "agenten" är språket självt.

2 Får man hoppas. Men i en recension ger sig Gunnar Fredriksson till att *prisa* översättningen. Brukar kulturskribenter inte läsa mot originalet? Eller försöka fatta vad de anmäler?

I begynnelsen var *Sprachspiele* alltså analoga med spel. Mot slutet finner Wittgenstein detta sätt att se löjligt och fördärligt. Ett femtiotal paragrafer i *Philosophische Untersuchungen* med början kring § 190 polemiserar mot den ramseyska tanken. Nu jämför han hellre med lekar. Har en lek regler kan dessa, säger han på engelska, vara något "we make up as we go along" (*PU*, § 83). De är då knappast konstitutiva: leken definieras ej i termer av dem. Inte heller är de särdeles strategiska; det finns inga segrar eller förluster i *Reigenspiele* [långdanser], omnämnda i *PU* § 66; och många lekar är samspel mer än motspel. Språket är inte kommet ur resonemang, påstår § 475 av *Über Gewißheit*; och det uttalandet från hans allra sista månad ska nog sammanföras med § 25 i *PU*: "Befehlen, fragen, erzählen, plauschen gehören zu unserer Naturgeschichte so, wie gehen, essen, trinken, spielen".

Definitioner, konstitutiva regler, språket som kalkyl var för den mogne Wittgenstein vanställningar av vad *Sprachspiele* ibland är; men 'language game' och 'språkspel' hart när tvingar tanken ditåt och därmed till missuppfattningar av hans sena filosofi. Dock skulle 'språklek' ha irrat för långt åt formlöshet. 'Spiel' har ett omfång utan motsvarighet i målspråken, så översättarna skulle ha behövt hissa varningsflagg.

C. Richard Matz begår inte tabbar av typ A men råkar ut för bekymmer av typ B. En ytterligare faktor gör hans försvenskning av Heideggers *Sein und Zeit* så misslyckad.

Det ges ett filosofiskt skäl till Heideggers idiom. Han tänker sig att bl a genom hans föregångares insatser har det språk han ärvt blivit så utformat att den verklighet vi alla delar har styckats upp på sätt som skymmer viktiga sammanhang. En ny indelning måste till. Hur förmedla den? Genom det traditionella, vanställande språk han vill diskreditera! Han kan inte överta det sånt det är; han måste kompensatoriskt deformera det.

Till den ändan arbetar han som en filosofins Gunnar Ekelöf, han som krossar bokstävlarna, bönar på latrin och kräkiska, associerar fosterland med mosterland och fasterland, ordleker ihop företeelser som för invanda uppfattningar står långt från varann och hoppas att läsaren i det mötet ska skymta sammanhang som vanligt språk inte mäktar uttrycka. Utan det fenomen litteraturteoretiker benämner 'semantisering' bleve boken obegriplig. Analogt semantisering kan här

och var göras på främmande språk, men inte ens mästerliga translatorer kan hoppas på att alltid skapa en god överensstämmelse. Och Matz saknar Wedbergs och Hertzbergs hand med tyska och svenska.

Lars Gustafsson har med en term från Hans Magnus Enzenberger resonerat som så:³

poesin har fler *frihetsgrader* än andra typer av språkanvändning. [...] Teoretiska eller verkliga kroppar kan vikas och tänjas på olika sätt och ju större denna repertoar är, desto större frihetsgrad har kroppen. En sadelyta kan till exempel inte vikas nedåt men uppåt – om den är orienterad i rummet som en riktig sadel. En handled har fler frihetsgrader än en armbågsled och så vidare. // Om man försöker göra saker med en kropp som går utanför den repertoaren går den sönder, det vill säga kan inte återställas i ursprungsläget. // [...] i den poetiska verksamheten kan vi göra fler saker med språket utan att det går sönder än vi kan göra i ett affärsbrev, ett matrecept eller en lagtext.

Kan detta vara ens diktarens sätt att se? Goda lyrikläsare är fostrade att inte skria 'Nonsens!' och kasta boken ifrån sig. De letar mening och sammanhang långt bortom den punkt där de skulle ha givit tappat med annan text. I vetskap om att denna ihärdighet är ett konstitutivt drag i lyrikläsandet kan och bör författaren tillåta sig antydningar och pauser som till vardags går spårlöst förbi men som i långsam och uppmärksam läsning blir meningsbärande, ja, måste så bli. Dikten medger *mindre* språklig frihet än affärsbrev och matrecept. Om diktaren bekymrar sig om hur det han säger mottas, borde hans inställning alltså vara motsatt Gustafssons: lyrik är ofriare än annan språkanvändning, just emedan så mycket mer har gjorts meningsbärande. Detsamma gäller den bundenhet som inte är "blott" tolkning utan troget vill återge textens alla nyanser.

Denna bundenhet beror nog på ett symboliskt kontrakt mellan läsare och text (snarare än författare). Mindre mystifierande: en genrekonvention bjuder att varje frasering, varje överklivning, varje paus lyriskt räknas och jämförs med aldrig yppade alternativ. En sidoeffekt är att enskilda ord gör meningsfulla syntaxuppror och ökar sin enzen-

3 Lars Gustafsson: "Poesins oändliga frihet", *Framtider* 1/95, s 11.

bergerska-gustafssonska frihetsgrad medan dikten som helhet minskar sin.

Min misstanke är att den som vill förstå *Sein und Zeit* måste läsa likt en lyrkläsare. Det innebär att hon eller han för stunden måste lägga åt sidan något nära filosofins hjärta för dem fostrade i "min" tradition – kravet att det filosofiskt fattade står *klart*, helst in i minsta vinkel och vrå. Det är ett pris jag är villig betala, under en övergångstid. Permanent dunkel ger jag inget för.

Heidegger begär, tror jag, att vi läser honom med lyrkläsarens tålmodiga långsamhet i att bilda sig ett fast omdöme. I gengäld förbinder han sig att i semantiseringar osv göra texthelheten meningsbärande. Han förtjänar att översättas. Jag skulle inte bli förvånad om han lever kvar i filosofins historia då Schock-pristagarnas namn är glömda. Men han betvivlade själv att översättning är möjlig. De engelska, franska och svenska smakprover jag har sett vederlägger inte hans pessimism.

Att översätta *Sein und Zeit* är att följa Heidegger efter, behandla målspråket med liknande distortioner som han har åsamkat tyskan samt se till att inte ge entydighet (nåja!) åt termer han ville hålla ännu mer polysema än de till vardags är men inte heller ge dem större mångtydighet. Följsamheten är knappast möjlig utom då språken är närstående nog att tillåta samma sorts ordlekar, oväntade sammansättningar, semantiserande rytmiseringar osv – villkor som knappast någonsin är uppfyllda. En återgivning av en dikt från ett annat språk är alltid en tolkning, har tappat eller avstått från en del och lagt till nytt. En återgivning av en text av Heideggers typ kan inte nöja sig med det. Ska det filosofiska innehållet bevaras, *måste* översättningen ha textens mångtydigheter och luckor, entydigheter och slutenheter. Vilket styrker Heideggers åsikt: SZ är oöversättlig. Så obegriplig som ryktet säger är den inte; men vill du förstå, måste du till originalet. Ingen genväg ges. Matz' fiasko är strukturellt. Företaget är omöjligt, hur skicklig translatorn än är.

Bör alltså dylika klassiker inte översättas? För många av oss är tolkningar enda sättet att stifta bekantskap med lyrik på främmande språk, och för snart sagt alla är tolkningar svårumbärliga hjälpmedel. Något liknande gäller de filosofiöversättningar som inte är av typ A. Naturligtvis vore det fatalt att *lita* på dem; ett skräckexempel är George Pitchers på sin tid flitigt lästa *The Philosophy of Wittgenstein* (1964), koncipierad utan misstanke att Wittgenstein inte alltid har sagt

vad översättare lägger i hans mun. Men läser man originalet jämsides med överflyttningarna, lär man sig sannolikt mer än om man tragglar sig fram på egen hand. Förståelsen skulle öka om fler översättningar funnes, också av Matz' art. Det skulle kräva att läsarna lärde sig använda dem alla och jämföra dem inbördes för att komma till rätta med originalet. Svenska filosofiinstitutioner uppmuntrar sällan till så intensiva klassikerstudier. Det är *ett* av felet med dem.

2. Att filosofera på modersmålet

Att översättningar av låt oss säga Kierkegaard, Nietzsche och Heidegger tappar påtagligt mycket mer än översättningar av låt oss säga Kant, Frege och Quine förklaras ibland med att de förra är mer "litterära". Nussbaum är sannerligen inte den första att söka en djupare förklaring i att de lyfter fram sin filosofisyn genom att utnyttja sitt språk på fler plan, kanske med andra och (tror jag) väsentligt *lägre* "frihetsgrader" än de senare. Hän mot en sådan djupare förklaring vill jag nu famla.

Det är, förmodar jag, en truism att allt tänkande över vad det vara månne kan ha filosofiska ankanger som förblir just ankanger tills de har tematiserats och blivit föremål för analys och argument.

Däremot är det knappast en truism att med Cavell kalla filosofi "a willingness to think not about something other than what ordinary human beings think about, but rather to learn to think undistractedly about things that ordinary human beings cannot help thinking about, or anyway cannot help having occur to them".⁴ Åtskillig akademisk filosofi verkar yvas över att vara esoterisk och strunta i huruvida det som försiggår i lärosalarna är relevant för problem som vanligt folk tänker på, problem som nykomlingar tror är filosofiska. Broad låter *Five Types of Ethical Theory* utmynna i att moralfilosofierna i hans bok saknar praktisk nytta men stor sak: "to try to understand in outline what one solves *ambulando* in detail is quite good fun for those people who like that sort of thing".⁵

Dock finns akademisk filosofi som har Cavells recept. Till den ska jag hålla mig. Jag ska driva tesen att den missgynnas om det blir kutym att modersmålet används bara i filosofiska kåserier. Skälen är nog i början tänkvärda även för dem som inte delar min filosofisyn.

4 Stanley Cavell: *Themes Out of School* (Chicago och London 1984), s 9.

5 London 1930, s 285.

Efterhand knyter jag an till en konception som är rätt kontroversiell. Jag har inte lust att utnämna denna till Den Sanna Konceptionen. Men något allvarligt och rentav kulturellt ödesdigert skulle inträffa den dag den motades ut från svenska universitet.

Modersmål och gemensam bakgrund. Troligen är det ingen slump att tänkare av Broads och i fonden Platons typ ser enskilda fall mest som illustrationer av allmänna principer, medan tänkare av Cavells och i fonden Aristoteles' typ ger de enskilda fallen egenvärde men håller ett öga på hungern efter det generella som halvt om halvt misstänks vara sjuklig. Skillnaden är knappast blott i tonvikt. Från respekt för enskildheter ska min tes dock utgå.

Anta att du inbjuds till ett främmande land för att föreläsa över valfritt filosofiskt ämne! Säkerligen börjar du fundera över vad som kan tänkas intressera auditoriet och hur väl du kan förmedla dina tankar. Sjuk av din hemrepertoar faller bort eller ändrar kraftigt karaktär. Sånt bortfall och sån förändring kan göra väl. Men nu ska inte det sköna i att inte vara hemma stå i förgrunden. Befrielsen har sitt pris. Det kan vara värt att betala. Men det tål att tänka på då du vet med dig att du på ett främmande språk ska rikta dig till folk med annan allmän bakgrund.

Med 'annan allmän bakgrund' avser jag något betydligt bredare än 'delvis annan filosofisk inriktning och beläsenhet'; den skillnaden finns ju gudskelov fortfarande mellan olika svenska universitet. Nej, jag tänker på icke-filosofiska skillnader i erfarenhet och kultur, bakgrundsfaktorer vi förlitar oss på då vi ställer och förklarar filosofiska problem för varandra. Anspelningar på väder och vind, på ljus och dagar, på mat och dryck, på mänskor, skeenden och historiska förhållanden jämte traditionella värderingar av sådana fenomen, ting som här hemma är självklara och gör vår mening tydligare, är ting som där är okända eller förbryllande eller fattas på sätt som i våra ögon är omöjliga. Minns amerikanska diskussioner av Ingmar Bergmans filmer!

Nu är det knappast givet att den allmänna bakgrund jag så vagt åberopar är just modersmålets. Minst lika sannolikt är det sätt att leva; och förmodligen lever en filosof i Alingsås mer likt en filosof i Stanford än en småbrukare i Burträsk. Nordliga indianer, inuiter och samer är språkskilda men delar vissa grundläggande erfarenheter som gör det möjligt för en svensk same att förstå inuitens problem men kanske näst intill omöjligt att förstå den stockholmske kulturskribentens. Ett

gemensamt modersmål är alltså inte nödvändigt för en gemensam bakgrund.

Inte heller är det tillräckligt. Kvinnor har stundom roligt åt andra saker än män och "kan bara inte förstå" hur karlar kan resonera som de gör; och omvänt. Syntax och semantik sätter sällan käppar i hjulet; besvären är av annan art än dem som exempelvis ungdomsspråk vållar. Snarare är habituering och praxis så annorlunda att orden i sitt sammanhang får associationer och övertoner som går den ena sidan förbi.

Denna fläckvisa oförståelse är just fläckvis, även om fläckarna kan visa sig ha samband. På det stora hela talar vi ett *gemensamt* modersmål. Detta kräver en gemensam bakgrund; annars blir det problem för den språkdimension som rör liknelser och metaforer. Denna kärvar om det förklarande ledet i liknelsen är okänt eller dåligt känt. Detsamma gäller då metaforleden är oförstådda. Sånt vore lappri om "retoriska figurer" vore idel dekorationer. Så är det ju inte. Bildspråk torde vara det första vi griper till inför oväntade situationer på vad område det vara må. Hur mycket skulle kunna sägas om själsliv utan "döda" metaforer om förståelse, begripande, fattande, upprördhet osv? Hur döda är de? Hur allmänmänniska, alltså hur överflyttbara från en kultur till en annan? Är irreducibla metaforer meningslösa? Utmärks inte en metafor av att utgå från något enskilt och plötsligt vidga till nya områden? Fungerar tekniken då det enskilda i fråga är okänt för åhöraren?

I fortsättningen tar jag för givet att åtskilligt i modersmålet binder samman oss svenskar, att rottrådarna grenas mer och tvinnas tätare än de som håller ihop de flesta intresse- och praktikgrupper samt att implikaturer och de kommunikativa "mekanismer" som skapar dem inte kan försummas då vi trevande vill artikulera filosofiska problem.

Implikatur. Ifall liknelser och metaforer för sin begriplighet kräver att talare och åhörare har en gemensam icke-filosofisk erfarenhet och kultur men denna gemenskap fattas, så drabbas den meningsförmedling och, viktigare, den meningsdaning eller semantisering som Grice har kallat för implikatur. Undersökningar av implikaturer rör hur ord i kraft av sin vanliga mening samt förhållandet att de fälls i en avvikande situation pragmatiskt förmedlar något mäktigt skilt från vad deras syntax och semantik bjuder. Ironi av den typ som inverterar ett yttrandes rättframma innebörd är ett banalt exempel. Implikaturer av

metaforiska yttranden illfänas redan i ens eget modersmål och tenderar att ändra sig över tid; Geijers odalman ”går som jag / och längtar efter sol och vår”. På främmande språk blir implikaturer oregeliga; vad vi har sagt fattar vi inte men att det måste vara galet framgår av auditoriets munterhet.

Du står alltså med inbjudan i hand och undrar vad du ska föreläsa över. Tänker du dig för, ratar du ämnen där skillnader i allmän bakgrund kan spela dig spratt. Du skyr områden med metaforer och liknelser som inte är stendöda. Du vägrar ta teman där du måste lita till implikaturer. Du vill så lite utomtextlighet du kan. Samtidigt vet du att det inomtextliga vållar dig sina egna svårigheter, att orden tappar i både konkretion och laddning: även inom det metafor- och liknelösa område du vill inskränka dig till vet du mindre vad du säger än om du kunde hålla dig till ditt hemvanda språk.

En lösning är att söka dig hän mot formlernas värld. Men formler måste förklaras i annat än formler och blir knappast intressanta utan nya frågor. Dessa kräver motivering och formulering. Sånt blir vanskligt utan pekanden, utan demonstrationer, utan metaforer, utan analogier, utan implikaturer. Tar du till något sånt medel, kan du befara relevanta skillnader i din och auditoriets allmänna bakgrund. Du bör därför akta dig för nya frågor och klamra dig till redan givna.

Anta att det i en framvällande internationaliserings spår är främst skrifter på främmande språk som meriterar; de kan ju bedömas av folk utifrån! Kommer inte du att snegla åt denna tänkta domarkår och resonera på samma sätt som nyss, alltså söka dig till formaliserade områden och färdiga frågor? Vad du gör på ditt modersmål men inte kan upprepa tillnärmelsevis lika säkert på utrikiska räknas inte riktigt. Så småningom tar du det inte ens själv på samma allvar som ”äkta”, dvs internationaliserbar, filosofi.

Nästa steg är att även hon, han och jag resonerar på ditt sätt. Stora områden som hittills har setts som filosofi och fortfarande räknas dit av dem som har fötts till ett världsspråk läggs i svensk akademisk lägervall. De glöms inte ens av oss universitetsutbildade filosofer. Men de tas av hand av andra – av sociologer, psykologer, litteraturvetare, biologer och anhängare av olika frälsningsläror. Detta är en framtid som i mycket redan är här.

Lingua franca. Så länge svenska talas, bör vi se till att det filosoferas på det. Detta ”kulturella” patos ligger på annat plan än det som

här är mitt. Jag vill vädra misstanken att filosofi i Cavells anda råkar i farozonen då vi övergår till utrikiska.

Farhågan är knappast uppenbart plausibel. För inte så länge sen var latin västerlandets lingua franca, men cavellskt filosoferande frodades. Att ett levande världsspråk övertar latinets roll borde väl inte vara så farligt, helst om stigande folkbildning gör att nästan envar nödtorftigt förstår språket ifråga?

Scenariot skiljer sig emellertid från det medeltida. Där hade den lärda världen visserligen övergivit folkspråken, men gemensam bakgrund var bibelns värld av gränssituationer, beskrivna på medvetet bildspråk; och skolastiskt latin användes för att diskutera sådana situationer på sätt som höll metaforerna levande. Till yttermera visso tränades ju predikanten i att tillämpa bibelns ord på nuet, vilket får betraktas som en övning i konsten att styra implikaturer.

Vårt scenario ger gränssituationer en god dag. Det är skapt för en tid då även filosofer har specialiserat sig och slutat att på arbetstid tänka på vad vanligt folk inte kan låta bli att tänka på: vad det är att leva, vad det är att dö och hur vi bör vara under vår stund på jorden. Sådana tankar formas i analogier, bilder och implikaturer. Dessa är inte utan vidare överspråkliga. De får därmed dåligt plats i överspråklig filosofi som med nödvändighet abstraherar från det regionala. Jag fruktar att vi som inte har lärt oss hantera andra språk med munkarnas bild- och implikaturmedvetenhet visserligen blir goda experter på ett eller annat fackområde men skruttiga filosofer. Vi blir fina på att besvara problem men usla på att ställa "första frågor", dem som ännu söker ord och formulering, och framför allt usla på att ställa "första frågor" utanför vår lilla specialitet.

Farhågan är blott löst bunden till motvilja att överge modersmålet. Desto mer knyts den till motvilja att göra filosofen till en fackman bland fackmän, fördjupad i det som så hemskt och så oklart av de franskinspirerade benämns "filosofins [?!] diskurs".

3. En bekännelse till filosofin

Så långt kan jag kanske påräkna viss sympati. Den övergår nog i sin motsats när jag nu spekulerar vidare över "icke-filosofisk bakgrund" och knyter an till en kontroversiell uppfattning om filosofins uppgift som bland annat innebär att ju mer vi förvandlar oss till ett skrå av problemlösare, desto mindre av filosofer blir vi. Det blir en bekännelse

till filosofi i sokratisk-aristotelisk anda, inte i Platons scientistiska.

Filosofi, tänker jag mig, är en förvetenskaplig verksamhet. Ur den växer vetenskaper. Självtänkan kan inte bli vetenskap. Ty då förlorar den sin själ, den intellektuella oro som söker en fråga som ännu inte är artikulerad. Är artikulationen gjord, är det dags för annat folk. Filosofi är *frågande* mer än *frågor*, en undrande hållning mer än ett spörsmål; och filosofi är avgjort mer frågor än svar. Den är stadiet före klara frågor och före utarbetade metoder att besvara dem. Den är med vid konceptionen men ger därmed sitt mest distinkta bidrag. Nu kan den vandra vidare, som den i *en* försvarbar läsning av filosofins historia alltid har gjort. Den attraheras av vetenskapernas säkrare gång men vet, med all respekt för Kant, att deras lunk och hållning är annorlunda och utgör en frestelse den inte får falla för.

Förbegrepp och oro. Filosofisk verksamhet är förvetenskaplig. Dess prefix framför andra är 'för-' eller 'pre-'. Den har, tror jag, ett starkt band till modersmålet, inte för att verksamheten är språklig (det var si och så med den saken även under oxfordfilosofins glansdagar) utan för att inläring av modersmålet ger en urbild av hur artikulation tillgår och vad den åstadkommer.

Tänk lite till på hur viktig den icke-filosofiska bakgrunden är för att vi ska förstå varandra! Rikta in dig på något empiriskt begrepp och undra hur det har förvärvat! Känns inte behov av en företeelse med det filosofiska prefixet, av ett förbegrepp eller (synonymt) en prekonception?

Som namnet säger är ett förbegrepp ett ännu-icke-begrepp. Sånt rynkar vi näsan åt, förståeligt. Men verkar det inte bestickande att gränsen mellan olika empiriska företeelser dras av hur man i en kultur är habituerad att se och hantera "dem"? Är det inte sannolikt att denna habituering är prekognitiv, går före inhämtandet och skapandet av vad vi kallar begrepp?

Går du med på att det finns begrepp hämtade ur erfarenheten men formade av människor är det lockande postulera något i stil med prekonceptioner. Ty ostensiva definitioner är inte basala; de kräver de underförståddheter som gör att barnet så småningom lär sig att vi pekar på ballongen och inte på fjärilen som fladdrar förbi den, på ballongens färg men inte på dess form eller på dess form men inte på dess färg.

Samtidigt kan man misstänka att förbegreppstraditionen kognitiviserar alltför mycket, trots att detta är vad Heidegger vill undvika med

talet om *Vorhabe*, *Vorgriff* osv. Man kan misstänka att den mot sina egna syften manar fram en alltför statisk bild av något Wittgenstein bättre fångar i sina *Sprachspiele* vilka normalt inte definieras och inom sin *Lebensform* inte kan definieras om de är fundamentala för den livsformen. Detta något kommer Bourdieu än bättre åt i termer av synnerligt ointellektuella habitueringar som på olika sätt fortsätter livet igenom.

Habitueringar, förbegrepp eller vad vi nu vill kalla dem har vi utan att märka dem. Den stund de anas ändrar de karaktär, tar form och kan sättas i fråga. En oro sprider sig om det slår oss att vi måste ha ett eller annat förbegrepp som bromsar något vi har tagit för givet och otänkande accepterat. Filosofisk blir oron när både sedvänja och uppstickare är plausibla och vi inte ser varför de skär sig och vilken sida som bör ge vika.

Artikulation och fråga. Artikulation löser inget filosofiskt dilemma men gör dilemmat tydligt. Dittills har det mest varit en ordlös oro. Nu inställer sig den befrielse som består inte i ett svar men i att vi äntligen vet vad vi frågar efter. Mycket till vetande är det sällan; förbegrepp är just såsom förbegrepp vaga och undflyende. Ändå skönjer vi, om än nått och jämt, t ex varför vi har varit så missnöjda med att en "ostensiv definition" skulle kunna lära barnet vad rödhet är. Tanken på förbegrepp är inte den klaraste som tänkts. Men ibland är oklarhet det bästa ämnet bjuder.

Maximen att inte söka större klarhet än ämnet medger står ju i början av *Nikomachiska etiken*. Där förbinds den med tanken att fastän filosofiska undersökningar av hur livet ska levas är bland de viktigaste vi kan ge oss på, har vi inga tydliga och självklara föreställningar att utgå ifrån och bland annat därför inget välgrundat hopp om att nå så värst tydliga och välgrundade resultat. Fortsättningen läser jag som ett argument för att det vid starten inte heller finns mycket till frågor utan mest en oro. Mycket av arbetet består i en genomgång av vardagliga föreställningar om goda och dåliga liv och försök att på basis av dem *göra* frågor (dvs artikulera oron), granska dem och besluta vilka som är värda att ställa. Kanske idealiserat ser jag *Nikomachiska etiken* som en tidig studie av färden från filosofisk oro till den någorlunda ro som den blotta utformningen av en någorlunda hanterlig fråga skänker. Hade Aristoteles inte fortsatt till svar, om än tentativa, hade han givit ett praktexempel på filosofi som artikulation, som en verksamhet vars

mål är nått i och med att en tämligen välartad fråga föds fram.

För temperament som Russells och troligen Platons ligger något komiskt i att filosofisk verksamhet skulle stoppa där. För temperament som Moores och troligen Sokrates' är framvärkandet av en hanterlig fråga stort nog. Att besvara den är inte nödvändigtvis filosofens sak, ty han eller hon är ingen vetenskapare. Om Moore och Sokrates likafullt är filosofer, tillerkänns en väsentlig roll åt artikulation och med den åt förbegrepp.

Artikulation kräver förutsättningar. Dessa ligger inte fixa och färdiga att röja. De är mer konstellationer av förbegrepp (illa uttryckt, men jag återkommer). Artikulering är att konceptualisera det förbegripna så att preciseringar och kritik blir möjliga. Sådant arbete går knappast att utföra på andra förbegrepp än dem man själv har i egen livsform, dem man är inhabituerad i.

En sådan livsform är den man får i och med att man lär modersmålet. I dagens samhälle är den mindre specifik än i gårdagens. Vi lever mindre utsatta för väder och vind, mindre traditionsbundna och kanske mer multietniskt. Ändå fostras vi in i svenskheter där vissa fenomen är så självfallna att vi aldrig tänker på dem förrän vi möter utlänningens häpnad men då förstår på sätt som merendels går honom eller henne förbi. Förmodligen är jag inte ensam om att ha större utbyte av svenskspråklig lyrik än annan; inte heller om att vara mindre dålig på att uttrycka och kommentera känslor och stämningar på mitt modersmål än på andra språk. De nästan tvåspråkiga lär i frågor som emotionellt djupt angår dem helst svara sig själva på modersmålet: där vet de bättre om det de känner är låt oss säga kärlek, förälskelse eller kättja. Mycket filosoferande kring liv och död, kärlek och hat kretsar kring situationer som bäst har uttryckts lyriskt och väsentligt består av känslor och stämningar.

Är denna iakttagelse riktig, är den livsform eller de livsformer man habitueras in i då man lär sitt modersmål inte något att saklöst försumma. Ty inte ens ett proffs kan alltid avhålla sig från att som privatperson undra över frågor av typen 'Vadan och varthän?' Skäl mot att göra sånt i ett andraspråk med dess inflätade livsform(er) är att förbegrepp avslöjas bäst i att vissa ordlekar, analogier, metaforer och implikaturer är naturliga eller någorlunda möjliga och andra långsökta eller omöjliga, och prøvostenen för vad som är naturligt, möjligt, långsökt och omöjligt är infödningens reaktion. Utbölingen ser snabbare

”naturliga” reaktioner, eftersom de inte är naturliga för honom. Just därför begriper inte han förbegreppen. De fattas blott av den som har haft dem. (Måhända en stipulativ skärpning av en tanketendens, men den ger i så fall en av de mer tänkvärda tautologierna.) Förbegrepp märker man då de börjar slinka ifrån en. Första åtgärden vid upptäckten brukar vara att sätta dem i fråga.

Då en artikulation är genomförd och förbegrepp alltså har blivit begrepp är infödingen inte så privilegierad, trots att han har en bekantskap utbölingen saknar och därmed en förståelse denne inte har. Du och jag kan med behållning läsa Aristoteles, med de reservationer som gäller översättningar och de reservationer som gäller flytt från en tid och kultur till en annan. Men vi kan inte artikulera hans problem för honom, medan du ibland lyckas artikulera mitt problem och jag ditt. Han och vi delar inte förbegrepp, och därför blir implikaturena andra. Här är en förklaring till att klassikerna säger delvis nya ting till varje ny generation, stundom ting de inte kan ha menat. I förklaringen ligger även en påminnelse om att förbegrepp inte är filosofiskt försumbara och att somliga av dem har sin hemvist i modersmålet.

Jag har nu *sagt* att filosofi som artikulation riktar in sig på att konceptualisera förbegrepp i den ”modersmålsliga” in habituering vi har på vissa områden. Det menar jag inte. Förbegrepp är blotta ansatser till begrepp, så den bestämda beskrivningen ’de förbegrepp som’ är tom. Att tala om förbegrepp i pluralis eller singularis har svårigheter analoga med dem som vidlåder att räkna antalet icke-inträffade olyckor. I mitt fall går besvären djupare. Man kan inte individuera sånt som förutsätts för *möjligheten* att individuera. Där måste vi tala oegentligt. Med tur kan man tala oegentligt på flera olika sätt med olika typer av missvisningar som bitvis tar ut varann. Min enda fäbless för det substantivsjuka talet om förbegrepp betingas en vag önskan att påminna om att filosofiska insikter finns i andra traditioner än de i Sverige knäsatta och en ännu vagare önskan att påminna om att Wittgenstein kanske är mer befryndad med Heidegger än med Ramsey. Sannolikt är det dock rekommendabelt att låta Bourdieu ta över och ersätta förbegrepp med habitueringar.

Filosofisk antropologi och bekännelse. En lärdom vi arvtagare till en ”analytisk” tradition får från början av vår utbildning är den lafontaineska att den som gapar efter mycket ofta mister hela stycket. Småbitar är vår specialitet. Vem törs idag tillstå att hon eller han drivs

av önskan att bättre förstå oss människor eller vår tid eller vår kultur? Dels klingar bekännelsen väl humanistisk för dem som drömmer om filosofi som en hardware science. Dels låter den så dignande; klyftan mellan ambitionen och vad vi rimligtvis kan uppnå är förnedrande stor.

Ändå kan många studier inom de områden jag bäst känner till fattas som skott på grenen filosofisk antropologi. Denna växer ihop med en gren som aldrig har blivit riktigt erkänd på denna sidan Östersjön, nämligen hermeneutik. Hermeneutik fattas då som studiet och kritiken av sätt på vilka vi som har hermeneutiskt intresse habitueras att se, tänka på och tala om tillvaron.

Filosofisk antropologi kan i denna sammanväxt väsentligt karakteriseras som ett kritiskt studium av förbegrepp eller habitueringar. Ofta tar den form av artikulation, av en inte alltid välvillig utarbetning inifrån av habituerade ansatser till handlingsmönster och synsätt. För dem som fattar filosofi som syntes, en synoptisk vy av till exempel ett samtida sätt att se mänskans plats i tillvaron, är förmodligen sådana ansatser tecken på att "analytiker" äntligen har börjat slå in på Sanningens Väg. I mina ögon är *sån* antropologi, den världsbildsframställande, en filosofi på dekadans, liksom all annan filosofi som sysselsätter sig med att besvara frågor mer än att formulera och ställa dem.

Jag hade hoppats knyta tanken på filosofi som artikulation till en tanke på filosofi som bekännelse, men en antydning får räcka. Wittgenstein, Wisdom och Cavell menar att somliga filosofiska svårigheter inte är rent och kanske inte heller primärt intellektuella utan är en perversion av viljan. (Vi traktar efter säkerhet men utger oss för att leta efter kunskap. Att så är ligger mitt för näsan på oss och är inget vi mödosamt letar oss fram till så som vi letar oss fram till våra förbegrepp.) Perversionen kan bekämpas endast om den tillstås.

Vanligen sägs trion se filosofi som psykoterapi. Det är ingen lyckad karakteristik om psykoterapi hör samman med individualpsykologi. Ty Wittgenstein & Co tror att svårigheterna ifråga är åtkomliga genom att granska och ändra de sätt varpå vi – inte just du, inte just jag utan vi alla inom en viss gemenskap – är habituerade att se, tänka på och tala om världen.

Man [dvs vi] prädiiziert von der Sache, was in der Darstellungsweise liegt. Die Möglichkeit des Vergleichs, die uns beeindruckt, nehmen wir für die Wahrnehmung einer höchst allgemeiner Sachlage.

Att läsa detta uttalande (*PU* § 104) och omgivande paragrafer som om de rörde språk och språkliga missuppfattningar är att trivialisera. Vi är inte i samma läge som kungen i *Alice i Underlandet*, han som när kuriren skrytsamt utbrast 'Ingen kommer fram fortare än jag' påpekade att i så fall hade kuriren inte varit framme före Ingen. Men långt, mycket långt innan vi hann vuxen ålder har vi vant oss vid och blivit fostrade till att handlande betrakta och umgås med vår omvärld på vissa sätt och inte andra. Dessa sätt underlättar vissa analogier, liknelser och implikaturer. De försvårar andra till den grad att det krävs ett stort viljeuppåbud för att ens börja skönja en ny möjlighet.

§ 115 i *PU* påstår att en *bild* [*Bild*] håller oss fångna. Snarare gör dock habitueringar det. De "ligger i" vårt språk i betydelsen att de från våra tidigaste år har givit de förbegrepp som gör våra ostentioner osv möjliga. Att se (eller uppfatta) världen annorlunda kräver att gå emot alla traditioner, inte främst de intellektuella eller ens de språkligt överförda utan de som på habitueringars sätt har satt sig i armar och ben, heideggerskt prekognitivt.

Vad Wittgenstein och hans folk är sysselsatta med är alltså endast på ytan språkkritik. Via studiet av förbegrepp sker en *traditionskritik* som varken är ett knäfall för vad gammalt och fornt är eller en strävan att göra oss urarva. Sådana studier är primärt undersökningar av förbegrepp man själv har i sin habituering. 'Man själv' är den individuella undersökaren – men som representant för en livsstil eller livsform, den "man" bibringas med modersmålet.

4. Slutkläm

Mycket har inte sagts om filosofi på just svenska. Men avsikten har varit att ge skäl för tron att då vi som har svenska till modersmål övergår till att filosofera på till exempel engelska har vi gjort mer än språkbyte. Vi har i samma rörelse avhämt oss möjligheten till vissa typer av undersökningar.⁶

För några av er är försakelsen obetydlig; ni tänker er att ni sysslar med transkulturella företeelser, kanske betingade av den mänskliga hjärnans beskaffenhet. Min misstro mot både metafysik och genetik är

6 Argument för de undersökningarnas vikt ges i min "Philosophical Polyglottism and First Person Perspectives" (*Theoria* LIX 1 – 3, 1993) som detta föredrag delvis förutsätter och delvis kompletterar.

otidsenligt större. Skalar man bort allt lokalt och tidsbundet har man tagit bort färg och doft och ljud men också, befarar jag, allt som ger tillvaron halt och meningsfullhet. För mig vore det att offra det mesta som gör ett filosofiskt liv värt att leva.

Naturligtvis är det lockande se strukturer och principer som Det Filosofiskt Väsentliga. Men inte är *Pride and Prejudice*, *Sense and Sensibility* och *Emma* tre påklädningsnummer av en och samma företeelse, den vars nakna form Vikram Seth antyder i titeln på sin janeaustenska mastodontroman om Indien strax efter självständigheten, *A Suitable Boy!* Inte är de moraliska dilemman vi möter där eller i vår vardag blotta illustrationer till konflikter mellan universaliserbara principer! Inte kan detaljer lätt sällas bort tills vi når en kärna som låter sig återge i formler om handlingen H i situationen S! Nej, romaner och vardagserfarenheter ger det stoff varur strukturer och principer möjligen kan *vinnas* och då som abstraktioner. Aristoteles anklagar indirekt Platon för att med grottfolket ha förväxlat skuggor med det som kastar dem: han har gjort abstraktioner viktigare än de företeelser de abstraheras från.

Har anklagelsen fog för sig, är konkreta detaljer och regionalismer inget filosofen skyndsamt bör se bort ifrån. Detta är fatalt lätt hänt när man filosoferar på ett främmande språk, tömt på modersmålets associationer. Ännu lättare än annars faller man för frestelsen att glida från att tänka på och tala om en konkret och därför ofta svårkaraktäriserad handling i en konkret och därför ofta svårkaraktäriserad situation till att tala abstrakt om handlingen H i situationen S och inte tänka på någon detaljerad handling i någon detaljerad situation överhuvud. Kvar blir då fackjargong samt teknik men, misstänker "aristotelikern", näst intill inget att tillämpa på, i vart fall inget som "ordinary human beings cannot help thinking about, or anyway cannot help having occur to them". Kvar blir alltså en väldig kvarn som mal med imponerande akademiskt oljud men utan mäld.

Denna skiss till motsättning mellan "platonism" och "aristotelianism" kräver förståss en diskussion om flera ting: om det partikulära versus det generella; om "praktiska slutledningar" (är de slutledningar i bemärkelsen att slutsatsen är en sats, inte en handling? är de entymematiska, med undertryckt överpremiss? hur stor vikt bör läggas vid Aristoteles' åsikt att agenten mindre drar en slutsats än så varseblir en

situation att varseblivning *involverar* handling?)⁷; samt en djupdykning i den sorts philosophy of mind där 'mind' inte är cartesianskt-husserlskt medvetande och knappast heller själsliv utan mer sinnelag, håg eller rentav personlighet. Så diskussion skulle ta mig ännu längre från mitt ämne, så jag sätter punkt här.⁸

7 Mina ansatser till svar finns i kap III och IV av *Nedom vara och bära?* (Nora 1993) samt Nussbaumrecensionen i *Kvinnovetenskaplig tidskrift* 1995:4.

8 Ett första tack till deltagarna i diskussionen vid Thales' jubileum den 24 november 1995 och ett andra tack till deltagarna i "mitt" seminarium i Göteborg där en bearbetning av texten ventilerades!

WLODEK RABINOWICZ

Värde grundat på preferens

Om två tolkningar av preferensutilitarism*

1. Preferensbaserat egenvärde – två tolkningar

Det som skiljer preferensutilitarism (PU) från andra former av utilitarism är värderingskomponenten. Preferensutilitaristen har ett eget förslag om vad det är som besitter egenvärde: Egenvärde eller, som man ibland kallar det för, *intrinsikalt* värde baserar sig på preferenser. Intrinsikalt värdefulla tillstånd kopplas enligt denna uppfattning till tillfredsställelse av våra önskningar.

Följande bör noteras: (a) Vad som anses vara värdegrundande är "våra" preferenser, och inte bara "mina". (b) Hänsyn tas bara till så k "intrinsikala" preferenser; det intressanta är vad vi föredrar för dess egen skull. (c) Eventuellt kan de relevanta preferenserna antas vara "idealiserade" på ett eller annat sätt. T ex kan det röra sig om de

* Denna korta uppsats baserar sig på en längre förlaga författad tillsammans med Jan Österberg. Som kommer att framgå, är Österberg och jag oeniga vad beträffar tolkningen av preferensbaserad axiologi. Denna oenighet har vi försökt utveckla och klargöra, men inte upplösa, i den gemensamma uppsatsen, "Value Based on Preferences: On Two Interpretations of Preference Utilitarianism", *Economics and Philosophy*, april 1996. I den föreliggande korta versionen tar jag tillfället i akt att i första hand framlägga min syn på saken. Medan jag tar hänsyn till några av Österbergs invändningar och försöker besvara dem, samt tar upp ett par kritiska synpunkter som har framförts av Sven Danielsson och Torbjörn Tännsjö, avstår jag att bemöta ett antal invändningar mot min ståndpunkt som bl a har förts fram av Krister Bykvist (i en opublicerad uppsats "Utilitarian Deontologies? On Preference-Utilitarianism and Agent-Relative Value"), John Broome och Derek Parfit (de senare i korrespondens). Broomes och Parfits synpunkter bemöts i viss utsträckning i den längre förlagan. Diskussionen av Bykvists kritik får anstå till ett annat tillfälle.

(intrinsikala) preferenser vi skulle ha om vi vore rationella, välinformerade om fakta och om vi hade tillfälle att noga överväga saken.

Vad är det nu egentligen som PU tillmäter intrinsikalt värde? Är det

(i) det förhållande *att* våra (intrinsikala) önskningar och preferenser tillgodoses? Eller är det snarare (ii) de förhållanden som är *föremål* för våra (intrinsikala) önskningar och preferenser? Observera att våra preferenser tillgodoses i och med att dessa förhållanden realiserar. De vanliga formuleringarna av PU antyder närmast den förra tolkningen, men även den senare tolkningen är tänkbar.

Här är ett exempel som illustrerar den distinktion som jag är ute efter. När vi önskar, säg, att regnskogen bevaras (inte, eller åtminstone inte bara, för något externt syfte utan för regnskogens egen skull), är det då intrinsikalt värdefullt att vi får som vi önskar eller är det snarare regnskogens fortlevnad som, i ljuset av vår önskan, framstår som något intrinsikalt värdefullt?

Även om vi får som vi önskar om, och endast om, regnskogen bevaras, är det uppenbart att det här är fråga om två olika förhållanden. Den ekvivalens som råder dem emellan är kontingent med tanke på att vi kunde ha andra önskningar.

Min uppsaliensiske kollega, Jan Österberg, har såvitt jag vet varit den förste att uppmärksamma denna skillnad mellan två tolkningar av preferensbaserat intrinsikalt värde (i en opublicerad uppsats "A Defence of the Preference Satisfaction Theory of Intrinsic Value"). Låt oss kalla dem för *preferenstillfredsställelse-* resp. *preferensobjekt*tolkningen. Österberg pläderar för den förra, medan jag själv vill plädера för preferensobjekttolkningen i stället. Åtminstone vill jag göra det villkorligt: Om PU överhuvudtaget är en rimlig ståndpunkt, så är den rimligare givet objekttolkningen.

2. *Betraktaren och deltagaren*

Mot dessa två tolkningar av preferensbaserad syn på värde svarar, såvitt jag förstår, två skilda sätt att se på utilitarism: *betraktarmodellen* och *deltagarmodellen*. Enligt den förra förkroppsligas den utilitaristiska inställningen i en opartisk välvillig betraktare som bedömer situationen objektivt och "utifrån". En vanlig människa kan approximera denna inställning genom att distansera sig från sitt personliga engagemang i situationen. Hon uppger, så långt det nu är möjligt, sitt partikulära

perspektiv och försöker anlägga "the view from nowhere", för att använda Thomas Nagels välkända formulering.

Deltagarmodellen uppställer däremot som ett utilitaristiskt ideal en attityd av deltagande med andra: det gäller att se på situationen "inifrån", men inte bara ur mitt eget utan även ur det andres perspektiv. Deltagarmodellen förutsätter att jag i stället för att distansera mig från mig själv identifierar mig med andra subjekt: vad den rekommenderar är inte objektivitet utan universaliserad subjektivitet.

Deltagarmodellen har en klart Humesk proveniens: den moraliska inställningen vilar på *sympati* – på vår förmåga till medkänsla och identifiering med andra. Som bekant, ansåg Hume att sympati, genom sin solidaritetsskapande karaktär, "produces our sentiment of morals". Den förklarar vår uppskattning av dygderna.

Att deltagarens identifiering med andra och betraktarens opartiska välvilja kan vara svåra att skilja åt framgår av följande citat från Rawls' *Theory of Justice* (s 26f), där de två modellerna betraktas som en:

The most natural way, then, of arriving at utilitarianism ... is to adopt for the society as a whole the principle of rational choice for one man. Once this is recognized, the place of the impartial spectator and the emphasis on sympathy in the history of utilitarian thought is readily understood. For it is by the conception of impartial spectator and the use of sympathetic identification in guiding our imagination that the principle of one man is applied to society. It is this spectator who is conceived as carrying out the required organization of the desires of all persons into one coherent system of desire; it is by this construction that many persons are fused into one.

Jag skulle nu vilja klargöra hur kontrasten mellan de två modellerna förhåller sig till vår distinktion mellan de två tolkningarna av preferentialisk syn på värde.

Deltagarmodellen innebär att det moraliska subjektet identifierar sig med den andre och därigenom övertar hans preferenser. Den andres intrinsikala preferens blir min egen: objektet för hans intrinsikala önskan blir därigenom något jag föredrar för dess egen skull. Deltagarmodellen leder därför på ett naturligt sätt till att intrinsikalt värde tillmäts objektet för preferensen.

Betraktarmodellen innebär däremot att det moraliska subjektet vill den andre väl. Jag vill dig väl, men detta behöver ju inte innebära att jag vill det du vill. Det kan till och med vara så att den distansering som utmärker betraktarperspektivet gör att jag i detta perspektiv förlorar anledning att vilja sådant som vi människor, med vår partikulära situation i världen, brukar vilja. Det kan då vara frestande att göra det *kantianska* steget och dra slutsatsen att det vi råkar vilja – viljans partikulära föremål – saknar ovillkorligt, absolut värde.

Enligt Kant är det inte viljans föremål – dess mer eller mindre kontingenta "materia" – utan snarare själva vårt predikament att vara viljande varelser som är relevant för frågan om det gudas natur. Om man sedan (inte längre i likhet med Kant, för vilken viljans "form" var det väsentliga) lokaliserar det ovillkorliga värdet i viljeuppfyllelse som sådant snarare än i viljans partikulära föremål, då har man nått en ståndpunkt som påminner om PU i dess tillfredsställelsetolkning.

Mycket förenklat kan man alltså säga att valet står mellan två moralfilosofiska traditioner – mellan impulser från Hume och från Kant.

3. Invändningar mot tillfredsställelsetolkningen

Jag skall nu presentera några argument som talar mot preferenstillfredsställelse- och för preferensobjekttolkningen.

(i) *Värdeinternalism*. Det förefaller vara ett rimligt krav på en moralisk teori att de värden den uppställer bör vara relativt intimt kopplade till våra handlingsmotiv. *Hur* intim denna koppling bör vara kan man säkert tvista om, men kanske kan följande villkor uppställas som ett minimikrav på en (rent värdebaserad, "teleologisk") moralteori.

VÄRDEINTERNALISM: När ingen annan än jag själv berörs, och jag är välinformerad, rationell och har tänkt noga på saken, då sammanfaller de intrinsikala värden som jag enligt den moraliska teorin bör förverkliga med de målsättningar som jag själv är angelägen om att nå.

Med andra ord: under idealiska omständigheter och när jag är den ende som berörs, bör intrinsikala värden sammanfalla med de målsättningar som motiverar mig att handla.

PU i dess objekttolkning har inga svårigheter med detta villkor.

Ponera att jag önskar att förstå relativitetsteorin (och min önskan är av den rätta intrinsikala sorten – jag vill nå denna förståelse för dess egen skull, av ren intellektuell nyfikenhet). Anta att situationen är sådan att ingen annan berörs och det är fråga om en övervägd önskan som jag skulle ha kvar om jag verkligen visste vad jag ger mig in på. PU i dess objekttolkning implicerar att det jag är angelägen om, dvs föremålet för denna min önskan, har ett intrinsikalt värde. PU i dess tillfredsställelsetolkning implicerar däremot att det endast är uppfyllelse av min önskan, vilken denna önskan än må vara, som äger intrinsikalt värde. Den senare målsättningen behöver jag emellertid *inte* vara angelägen om (åtminstone inte intrinsikalt, för önskeuppfyllelsens egen skull). Jag *kan* vara det, vid sidan om mycket annat jag eftersträvar, men det är inte nödvändigt att jag har denna önskan ”av andra ordningen”. I synnerhet kan det mycket väl vara så att min önskan att förstå relativitetsteorin är ”kategorisk” i den meningen att objektet för min önskan är att förstå teorin ifråga och inte ”att förstå den förutsatt att jag önskar mig denna förståelse”. Min önskan behöver med andra ord inte vara betingad av sin egen existens. Men i så fall *kan* jag vara likgiltig inför det faktum att jag genom att realisera objektet för min önskan även får min önskan uppfylld. Att min preferens tillfredsställs behöver inte höra till mina motiverande målsättningar.

Jämför med följande citat ur Stephen Darwalls *Impartial Reason* (Cornell UP, Ithaca and London 1983, kap 3, s 37):

From the fact that any intentional action must be explainable by the agent's desires it does not follow that the agent's reason for acting must itself make reference to his desires. Usually when we act as we desire we are not moved to act by our awareness of the fact that we desire. Rather, a desire itself partly consists in a disposition to be moved by certain facts connected with its *object*. ... Desires are for their objects, and the locus of the desiring person's concern is most typically the object itself and the facts connected with it rather than the fact that she has the desire.

Väsentligen samma tanke förs fram i en inflytelserik uppsats av Philip Pettit och Michael Smith, ”Backgrounding desire”, *Philosophical Review* 99 (1990), 565–92.

(ii) *Butlers problem*. Joseph Butler's berömda kritik av hedonismen tycks i minst lika hög grad träffa värdeuppfattningar som tillmäter in-

trinsikalt värde åt preferenstillfredsställelse snarare än åt objekt för preferenser. Preferenstillfredsställelse kan endast vara möjlig som en sekundär målsättning: om vi hade blott en enda preferens – att våra preferenser blev tillgodosedda – skulle vi ju inte ha några preferenser att tillgodose. I Butler's terminologi, "self-love" – egenkärleken – behövs input. "Take away these affections [= particular passions and appetites], and you leave self-love absolutely nothing to employ itself about; no end or object to pursue..." (Förordet till *Fifteen Sermons*, § 15).

Varför utgör detta egentligen ett problem? Ja, det förefaller att en värdeuppfattning enligt vilken inget annat än preferenstillfredsställelse är intrinsikalt värdefullt å ena sidan "omyndigförklarar" alla våra vanliga intrinsikala strävanden, genom att fränkänna deras föremål intrinsikalt värde, men å andra sidan lever den på dessa omyndigförklarade strävanden. För att uttrycka det drastiskt har vi här att göra med en uppfattning som parasiterar på de attityder som den samtidigt desavouerar. Den skulle ha "absolutely nothing to employ itself about" om vi tog den ad notam och slutade bry oss om annat än att våra preferenser tillfredsställs. (Att vi inte kan ta den ad notam, att sådant är psykologiskt omöjligt, gör väl inte saken bättre?)

Denna uppfattning tycks alltså inte ta på allvar våra önskningar och preferenser samtidigt som den förutsätter att vi fortsätter att ta dem på allvar.

(iii) *Svårbegriplighet*. Att förbehålla allt intrinsikalt värde åt preferenstillfredsställelse ter sig svårbegripligt: Vad kan det finnas för värde i att tillgodose våra strävanden om föremålen för dessa strävanden systematiskt fränkänns ett värde? Varför är det värdefullt att våra önskningar uppfylls om det vi önskar genomgående är värdelöst?

Att folk får det de vill, även i de fall när det de vill saknar värde, kan eventuellt anses vara intrinsikalt värdefullt, vid sidan om mycket annat. Det kan ses som ett värde bland andra. (Många människor tycks lägga en oberoende vikt vid viljeuppfyllelse som sådan och det är inte uteslutet att de skulle behålla denna intrinsikala preferens även om de vore välinformerade och övervägde saken noga.) Men det är inte lätt att begripa hur sådant som preferensuppfyllelse kan framstå som det enda värdet.

(iv) *Hypotetiska preferenser och preferenser avseende hypotetiska fall*. Tänk på en vanlig invändning mot PU: Preferensutilitaristens

värderingar *förefaller* stämma överens med common sense. Han tycks lägga vikt vid sådant som de flesta av oss lägger vikt vid, t ex vid orörd natur, estetiska upplevelser, frihet, kunskap, etc. Men denna överensstämmelse med common-sense värderingar vilar i själva verket på lösan sand. För tänk om vi hade andra preferenser än de vi faktiskt har. Tänk om vi vore likgiltiga inför, säg, naturens rikedom eller inför estetiska upplevelser. Då skulle det enligt preferensutilitaristen inte finnas något värde i dessa tillsynes önskvärda objekt: de skulle ju inte längre tillgodose våra preferenser! Preferensutilitarismens ointuitiva karaktär blir alltså klar när vi betraktar *hypotetiska* fall (i vilka våra preferenser vore annorlunda).

Såvitt jag förstår är denna invändning riktad mot PU i dess preferenstillfredsställesetolkning: hade preferenserna varit annorlunda så skulle deras tillfredsställelse kräva att andra tillstånd realiserats. PU i denna tolkning rättar sig alltså efter hypotetiska preferenser i hypotetiska fall. Preferensobjekttolkningen av PU leder däremot till väsentligen annorlunda slutsatser. PU i denna tolkning rättar sig efter våra *aktuella* preferenser beträffande hypotetiska fall. Om vi har en (övervägd) intrinsikal preferens för att, säg, regnskogen bevaras *även* för det hypotetiska fallet i vilken vi inte skulle ha brytt oss om regnskogens fortlevnad, då implicerar objekttolkningen att regnskogens fortlevnad skulle ha ett intrinsikalt värde även om dess existens inte tillgodosåg några preferenser. Hypotetiska preferenser är alltså av intresse för bedömningen av hypotetiska fall endast om våra aktuella preferenser för hypotetiska fall är känsliga för sådana hypotetiska preferenser – något som de inte behöver vara. Den beskrivna invändningen är därför uppenbarligen otillämplig givet objekttolkningen.

Svårighet: Men – när vi bedömer ett hypotetiskt fall – borde vi då inte, givet deltagarmodellen, identifiera oss även med de hypotetiska aktörerna och därigenom låta deras preferenser också påverka bedömningen? Jag är benägen att svara nej på denna fråga, men vet inte riktigt hur man kan argumentera för detta svar. Jag skulle emellertid vilja hävda att deltagarmodellen förutsätter att det verkliga ges ett absolut företräde framför det blott hypotetiska. När vi identifierar oss med andra och övertar deras preferenser, är det följaktligen bara de verkliga andra och deras verkliga preferenser som som räknas. Observera: om vi hade behövt identifiera oss med såväl de hypotetiska som de aktuella individerna vad beträffar bedömningen av hypotetiska fall,

då hade det funnits anledning att göra likadant vid bedömningen av det aktuella fallet. Varför inte betrakta det aktuella fallet från alla tänkbara hypotetiska perspektiv? Men gör vi så, blir deltagarmodellen bankrutt! Det finns för många tänkbara konflikterande perspektiv att ta hänsyn till – perspektiv som mycket väl kan tänkas neutralisera varandra. Det är därför som deltagarmodellen enligt min uppfattning måste ges en aktualistisk utformning: det är bara de aktuella preferenserna som skall utgöra grund för värderingen. För vidare diskussion av aktualism, se nedan, sektion 4, punkt (v).

Ett annat alternativ vore att för varje hypotetiskt fall bara ta hänsyn till de hypotetiska preferenser som skulle ha förelegat i fallet ifråga. Detta skulle emellertid innebära att intrinsikala värden skulle tillåtas variera från en möjlig värld till en annan, i takt med preferensvariation. Regnskogen skulle vara intrinsikalt värdefull i vår värld, men skulle sakna värde i en värld där folk vore likgiltiga. Den avgörande nackdelen med en sådan uppfattning är att den strider mot alla våra föreställningar om intrinsikalt värde: ett sådant värde skall ju *ex definitione* vara oberoende av varierande externa omständigheter: ett objekt är intrinsikalt värdefullt endast om det är värdefullt under alla omständigheter, i varje möjlig situation.

(v) *Värdepluralism*. PU kritiseras ibland för sin monistiska axiologi: preferensutilitaristerna, säger man, ignorerar det uppenbara faktum att värden är heterogena till sin natur. Värdenas mångfald byts ut mot ett enda värde: att folk får som de önskar. Igen tycks PU strida mot vår intuitiva värdeuppfattning. (Samma sorts invändning brukar förresten riktas mot hedonistisk utilitarism.)

Givetvis kan man lätt bemöta denna invändning om man är anhängare av objektolkningen av PU. Givet denna tolkning är det naturligt för PU att utmyнна i värdepluralism. Våra intrinsikala preferenser tycks ju ha mycket varierande objekt. Skulle vi behålla dessa preferenser även efter ”moget övervägande”, då följer det att det finns flera oreducerbart distinkta intrinsikala värden.

(vi) *Värden som inte är värden för någon*. Något man ofta betraktar som utmärkande för alla former av utilitarism är värdeindividualism: en uppfattning enligt vilken ingenting kan vara värdefullt utan att vara värdefullt för någon (för någon ”individ”). Hedonistiska utilitarister placerar värdet i lust eller lycka: men om det är bra att människor är lyckliga så är det väl bra därför att det är bra *för* varje människa att

vara lycklig. Likadant förhåller det sig med preferensutilitarismen i dess vanliga tillfredsställelsetolkning: om det är bra att folk får som de vill så är det väl bra därför att det är bra för var och en att få som han eller hon vill.

Dessa individualistiska utformningar av utilitarism kommer därför i konflikt med en vanlig värdeuppfattning enligt vilken det kan finnas mycket som är intrinsikalt värdefullt utan att för den skull vara intrinsikalt värdefullt för någon person eller ens för någon levande varelse: t ex artvariation, personlighetsmångfald, jämlikhet, vackra landskap i öde trakter, osv.

Igen är det uppenbart att preferensobjektolkningen klarar av detta problem utan besvär: föremål för (överbägda) intrinsikala preferenser är intrinsikalt värdefulla även i de fall när de inte är intrinsikalt värdefulla för någon. (I den mån som de är föremål för någons intrinsikala preferenser och i förlängningen även föremål för denne någons värderingar måste de ju vara intrinsikalt värdefulla enligt någon; men det är en annan historia.)

Att preferensobjektolkningen av PU lämnar utrymme för värdepluralism och för sådant som är bra utan att vara bra för någon gör att denna ståndpunkt närmar sig common sense vad beträffar värdefrågor. Dessa karaktäristika innebär emellertid samtidigt att ståndpunkten ifråga fjärrar sig från våra gängse föreställningar om vad som utmärker ett utilitaristiskt synsätt. Vi får en värdeuppfattning som saknar de för standardutilitarismen så vanliga reduktionistiska tendenserna: värdeindividualism och värdeindividualism lyser med sin frånvaro.

4. Invändningar mot objektolkningen

(i) *Den rawlska invändningen.* I det citat som har reproducerats ovan (sektion 2), hävdar Rawls att det utmärkande för utilitarism är "organization of the desires of all persons into one coherent system of desire". Vad som händer är att "many persons are fused into one". Den välkända Rawlska slutsatsen blir följaktligen att "[u]tilitarianism does not take seriously the distinction between persons" (*A Theory of Justice*, s 27). Invändningen träffar närmast deltagarmodellen (och därigenom indirekt objektolkningen som denna modell ligger till grund för). Deltagarmodellen involverar ju preferensövertagande. Som deltagare tar jag över andras preferenser och gör dem till mina. Men därigenom reduceras interpersonella preferenskonflikter till *intra-*

personella motsättningar som löses genom en kraftmätning: en svagare preferens ger vika åt den starkare. Det finns därför inget som hindrar att en persons svagare preferens uppoftas för att tillgodose en starkare preferens från en annan persons sida. Inte heller finns det något som hindrar att en persons starka preferens uppoftas för att tillgodose svaga preferenser hos ett stort antal andra personer.

Är denna analys korrekt? Inte nödvändigtvis. Den baseras på ett tvivelaktigt "endimensionellt" synsätt på individuell rationalitet. De olika preferenser hos individen som tillsammans ligger till grund för hans val ses här i analogi med krafter som drar i olika riktningar. Det enda som räknas är deras relativa styrka och antal: den slutliga riktningen – det rationella valet – bestäms genom att de olika krafterna adderas i proportion till deras styrka.

Ett sådant perspektiv på individuell rationalitet är klart överförenklat. Agentens preferenser behöver inte alla vara av samma skrot och korn. Intrapersonella konflikter är inte lätta att lösa när personens olika divergerande målsättningar är heterogena och därför svårligen jämförbara med varandra. I problem som involverar konflikter mellan t ex frihets- och kunskapssträvan, eller mellan personlig ambition och vänskapshänsyn, finns det inget enkelt recept för konfliktlösning. Att bara jämföra preferensstyrka räcker inte (även om sådana jämförelser lät sig göras, vilket inte alls är säkert). Att jämka samman olika dimensioner i ens personliga strävanden är ett komplicerat företag som inte lämnar mycket utrymme för enkel styrketävlan. Inte heller finns det någon garanti för att en sådan sammanjämkning alltid blir framgångsrik; amalgamering av heterogena preferenser kan ibland vara ogörlig.

Om vi accepterar vi en sådan flerdimensionell syn på individuell rationalitet, kan vi försöka värja oss mot den rawlska invändningen: att jag som deltagare övertar andras preferenser behöver inte innebära att distinktionen mellan de olika personerna försvinner. De andras preferenser kastas inte bara i en gryta och blandas ihop med varandra. Även efter övertagande förblir de relativt åtskilda. De uppträder nu som olika dimensioner inom den utilitaristiska agentens preferensstruktur. Interpersonell konflikt transformeras därigenom i en komplex *flerdimensionell* intrapersonell motsättning. Distinktionen mellan personer tas i så måtto fortfarande på allvar.

(ii) *Möjliga världar*. I en möjlig värld i vilken människornas preferenser hade varit annorlunda, i vilken de hade önskat annat än det vi

faktiskt önskar, skulle anhängaren av objekttolkningen ha kommit fram till andra slutsatser vad beträffar intrinsikalt värde än i den aktuella världen. Detta förefaller vara orimligt, med tanke på att påståenden om intrinsikala värden normalt inte antas vara världsbundna. Om sådana påståenden är giltiga, gäller de med nödvändighet: för varje möjlig värld.

Svaret på denna invändning – om det nu finns något svar – måste gå ut på att objekttolkningen förutsätter en subjektivistisk (eller ”projektionistisk”) syn på värde. Värden är inte en del av världen utan något vi projicierar på världen, eller rättare sagt på hela systemet av möjliga världar. I denna bemärkelse kan våra värdepåståenden, som ju görs inom ramen för vår värld, ses som uttryck för ett partikulärt världsbundet perspektiv – ett perspektiv som bestäms av våra faktiska preferenser. Men denna världsboundenhet vad *perspektivet* beträffar gör dem inte världsbundna i deras *applikationsområde*: vi projicierar värden på alla möjliga världar och inte bara på den värld som råkar vara aktuell. Därigenom finns det inget hinder för uppställande av värdeomdömen av intrinsikalt slag (dvs omdömen som tillmäter samma värde åt ett sakförhållande i varje möjlig värld).

Till skillnad från anhängarna av tillfredsställelsetolkningen vägrar objekttolkaren att höja sig till betraktarens distanserade perspektiv. Från ett sådant perspektiv, hävdar han, blir inga värden synliga längre. Världen töms på värde. Om värden skall räddas måste vi hålla oss kvar inom det subjektiva perspektivet. Ett sådant är med nödvändighet världsbundet och i denna mening kontingent.

Den aktuella invändningen kan alltså besvaras *om värdesubjektivism* är en försvarbar ståndpunkt (något som givetvis är starkt omdebatterat).

(iii) *Objektolkarens opportunist*. Det skulle kunna hävdas att objekttolkningen inbjuder till opportunist; den tillåter mig att påstå att ett sakförhållande, A, är intrinsikalt värdefullt bara därför att många människor föredrar det för dess egen skull. Men detta är ju kontraintuitivt. Om jag betraktar A som *intrinsikalt* värdefullt, måste jag kunna utpeka några *interna* egenskaper hos A som gör det värdefullt. Intrinsikala värden kan inte grundas på *externa* egenskaper, sådana som egenskapen att vara föremål för preferenser.

Men är det uppenbart att objekttolkningen har denna kontraintuitiva konsekvens? När jag tar över andras intrinsikala preferens för A – när

jag gör den till min egen preferens – blir jag i stånd att ange de interna egenskaper hos A som gör det värdefullt. Dessa är precis de egenskaper för vilka A har föredragits av andra, och för vilka jag nu föredrar detta sakförhållande!

Vi bör skilja mellan å ena sidan de egenskaper på vilka intrinsikalt värde "supervenierar" – de karaktäristika hos objektet som vi pekar på när vi ombeds att rättfärdiga vårt värdeomdöme (egenskaper som vi värderar objektet *för*) – och å andra sidan de egenskaper och tillstånd som "konstituerar" – ger upphov till – intrinsikalt värde. Det är bara de förra som måste vara interna egenskaper hos objektet. Enligt objekt-tolkningen konstitueras intrinsikala värden hos sakförhållanden av något externt – av våra (intrinsikala) preferenser. Men dessa värden supervenierar inte på preferenserna. I stället supervenierar de på de interna egenskaper hos sakförhållanden för vilka dessa sakförhållanden intrinsikalt föredras. En analogi med konventioner kan hjälpa här. Att ett visst drag i schack är tillåtet supervenierar på de interna egenskaperna hos draget tillsammans med de interna egenskaperna hos spel-situationen. Men den konstituerande grunden för att draget är tillåtet står att finna i den sociala konvention som upprätthåller schackreglerna.

Schackanalogin är emellertid inte fullständig, såsom har påpekats för mig av Sven Danielsson. Det intrinsikala värdets beroende av dess konstituerande grunder skiljer sig från ett motsvarande beroendeförhållande i schackfallet. Hade schackkonventionen varit annorlunda, skulle draget ifråga inte nödvändigtvis vara tillåtet. Hade däremot preferenserna varit annorlunda, skulle intrinsikala värden ändå förbli desamma. (Se sektion 4 ovan, punkt *(iv)*.) Denna disanalogi har att göra med de intrinsikala värdenas obegränsade applikationsområde. Samtidigt som de kreeras av de aktuella preferenserna, är de *applicerbara* även på möjliga världar i vilka preferenserna hade varit annorlunda. En bättre analogi har föreslagits av Krister Bykvist: Tänk på vår användning av en term som "en meter". Mitt paraply är en meter lång. Den längd "en meter" betecknar fixeras konventionellt genom en relation till en extern standard (sådan som standardmetern i Paris, eller en viss bråkdel av den väg som ljuset färdas under en sekund). Men om vi sedan funderar på en möjlig värld i vilken *standard* hade varit annorlunda och längre, är vi ju inte alls benägna att säga att i en sådan värld skulle mitt paraply ha haft en annan längd än i den aktuella världen. Det skulle fortfarande ha varit en meter lång. (Även om man

i denna tänkta värld hade sagt att det var kortare än så!) I Saul Kripkes terminologi (från *Naming and Necessity*) är en term som "en meter" en s k rigid designator: en term vilkens innebörd konventionellt fixeras på ett sådant sätt att den plockar ut samma objekt eller egenskap i varje möjlig värld, oberoende av vilka konventioner än skulle råda i världen ifråga. Värdepredikat betar sig analogt: De rådande preferenserna bestämmer vilka sakförhållanden som är intrinsikalt värdefulla på det "rigida" sättet – sakförhållanden ifråga förblir intrinsikalt värdefulla även i de världar i vilka andra preferenser skulle råda.

Det bör slutligen påpekas att den intuitiva distinktionen mellan interna och externa egenskaper hos ett objekt, vilken ligger till grund för den aktuella invändningen, inte alls är lätt att klargöra. I svaret på invändningen har den emellertid tagits för given, för argumentets skull.

(iv) *Värden enligt någon*. Objekttolkningen tillåter inga "oupptäckta" intrinsikala värden: allt som är intrinsikalt värdefullt måste intrinsikalt föredras av någon. Detta verkar kontraintuitivt. Kan det inte finnas sakförhållanden som är av värde fast de inte har uppmärksamats av någon, eller åtminstone inte tillräckligt uppmärksamats för att kunna intrinsikalt föredras?

Som den står kan invändningen bemötas genom en *idealiserad* tolkning av intrinsikal preferens (se ovan, sektion 1). Objekttolkningen kan mycket väl tillåta värdefulla sakförhållanden som inte föredras av någon. Det räcker att vi *skulle* föredra dem om vi noga reflekterade över deras (interna) beskaffenhet.

Men det måste ändå erkännas att objekttolkningen i sista analys endast godtar värden som som uppskattas av någon eller skulle uppskattas under lämpliga omständigheter. "Oupptäckbara" värden godtas inte. Är detta erkännande katastrofalt? Eller är det något vi kan leva med?

(v) *Hypotetiska perspektiv*. Deltagarmodellen och objekttolkningen kan kritiseras för utpräglat "aktualistisk" eller "realistisk" orientering: Det ideala utilitaristiska subjektet förväntas identifiera sig med andra aktuellt existerande subjekt, men däremot inte med blott hypotetiska personer. Även när jag betraktar ett hypotetiskt fall antas jag identifiera mig med aktuella subjekt och ta över deras aktuella preferenser beträffande fallet ifråga. Jag förväntas inte ta över de hypotetiska preferenser som skulle föreligga i detta hypotetiska fall. Varför det? Hur kan man rättfärdiga en sådan favorisering av det verkliga och ak-

tuella på bekostnad av det blott möjliga?

Det är en svår fråga att besvara, främst av det skälet att favoriseringen av det verkliga är en så fundamental inställning hos oss människor. Det är som om någon undrade varför vi endast bryr oss om sådant som händer i den aktuella världen men tar med ro alla de hemskheter som utspelar sig i blott möjliga världar. Varför är jag nöjd sålänge den aktuella världen är tillfredsställande? Varför bryr jag mig inte om de tänkbara världar som är fulla av obeskrivligt lidande och smärta? Det enda svar som jag kommer på är: Varför skulle jag bry mig om sådant som inte är fallet?

Anhängaren av objektolkningen kan svara likadant. Varför skulle jag bry mig om subjektiva perspektiv som inte är verkliga? Varför skulle jag ta hänsyn till preferenser som ingen hyser (eller ens skulle hysa efter moget övervägande)? Gränsen mellan det verkliga och det överkliga tycks vara av helt central vikt, även om vikten i denna gränsdragning kanske bara är ett "brute fact", som inte låter sig ytterligare rättfärdigas.

(vi) *Preferensoberoende?* Men i vilken mening är egentligen objektolkningen preferentialistisk i så fall? Den innebär ju att samma sakförhållanden skulle förbli intrinsikalt värdefulla även om folk hade andra preferenser! Hur kan man då hävda att den låter värdet vara beroende av (aktuella) preferenser? Beroendeförhållanden mellan två faktorer borde ju innebära att om en faktor varierar blir även den andra faktorn förändrad. Inget beroende utan hypotetisk variation! (Invändningen har formulerats av Sven Danielsson.)

Såvitt jag förstår bör objektolkaren ta denna invändning på allvar och i viss utsträckning modifiera sin ståndpunkt. Hypotetisk variation kan ges utrymme via *värderelativism*. Hittills har vi rört oss med en ansats i vilken det aktuella preferensperspektivet har förutsatts som givet. Därför var det möjligt för oss att inte låtsas om värdets relativisering till detta perspektiv. Så länge perspektivet hålls konstant, är det möjligt att tala om intrinsikalt värde i absoluta termer, utan relativisering. Men om man så vill, kan man säga att relativiseringen till det aktuella perspektivet har förelegat implicit. En sådan implicit relativisering kan göras synlig om man i stället laborerar med en modellering i vilken vi avstår från att förutsätta ett bestämt aktuellt preferensperspektiv, utan istället tillåter det aktuella perspektivet att hypotetiskt variera. Relative *ett* sådant perspektivval blir då vissa sakförhållanden

intrinsikalt värdefulla. Låter vi i stället ett annat preferensperspektiv fungera som det aktuella – dvs låter vi en annan möjlig värld fungera som den aktuella – får vi en annan uppsättning av intrinsikalt värdefulla sakförhållanden. När värderelativiseringen görs explicit på detta sätt synliggör vi den hypotetiska variation som krävs för ett beroendeförhållande mellan intrinsikala värden och aktuella preferenser. Det bör emellertid understrykas att för varje val av ett aktuellt preferensperspektiv blir de intrinsikala värden som existerar relative detta perspektiv fortfarande invarianta över olika möjliga världar: ett sakförhållande som är intrinsikalt värdefullt i en värld relative ett givet preferensperspektiv förblir intrinsikalt värdefullt relative detta perspektiv även i de världar där preferenser skulle ha varit annorlunda! Ty varje sådant preferensperspektiv bestämmer ju en uppsättning *intrinsikala* egenskaper hos sakförhållanden som intrinsikalt värde (relativiserat till detta perspektiv) supervenierar på: de egenskaper för vilka sakförhållanden föredras för deras egen skull, oberoende av externa omständigheter, utifrån det valda preferensperspektivet.

De läsare som är bevandrade i möjligvärldssemantik för modallogik med rigida designatorer är väl förtrogna med samma dubbelhet vad beträffar relativiteten hos nödvändig sanning: I gängse möjligvärldsmodeller antar man den aktuella världen (eller mera generellt: den aktuella kontexten) som given och detta på förhand gjorda val, som fixerar referensen hos rigida designatorer, gör det möjligt att tala om nödvändiga sanningar i absoluta termer, även vad beträffar påståenden som innehåller rigida designatorer. Men en modell med en fix aktuell värld kan alltid inbäddas i en modell där den aktuella världen inte fixeras: i stället laborerar man där med ett sanningsbegrepp som relativiseras till olika val av aktuella världar (kontexter). I en sådan modell blir olika påståenden nödvändigt sanna (dvs invariant sanna över alla möjliga världar) relativt olika aktualitetsval.

Normalt finns det ingen anledning att göra värderelativiseringen explicit; vi tar ställning till värdefrågor i absoluta termer, utifrån de preferenser vi faktiskt hyser. Men objektolkaren kan alltid falla tillbaka på en sådan relativisering när beroendet mellan värden och preferenser sätts ifråga.

(vii) *Preferentialistisk axiologi?* Men innebär inte det ovan sagda att objektolkningen i själva verket inte alls är en form av preferensutilitarism utan snarare får ses som en s k objective list-teori om intrin-

trinsikalt värde kombinerad med en projektionistisk preferensbaserad metaetik? För objektolkaren spelar preferenserna en viktig roll endast vad värdekonstitution beträffar men inte vad beträffar värdeinnehållet: i princip behöver de inte omnämnas alls när vi i en "objektiv lista" specificerar vilka sakförhållanden som är intrinsikalt värdefulla. Objektolkarens axiologi, till skillnad från hans metaetiska värdeteori, behöver därför inte alls vara preferentialistisk. (Invändningen ifråga har framförts av Torbjörn Tännsjö i ett bokmanuskript med den preliminära titeln *Classical Hedonistic Utilitarianism: A Defence*, kap 6.)

Även om denna invändning är bestickande bör den ändå tillbakavisas. Den bygger på ett antagande att det går att dra en skarp gräns mellan metaetik och substantiell etisk teori. Men kan en sådan skarp gränsdragning upprätthållas? Även om vi ignorerade Quines välkända argument och insisterade på en strikt åtskillnad mellan substantiella frågor och frågor om mening, förblir gränsen mellan metaetik och axiologi oklar. Visserligen kommer rent semantiska teser om meningen hos värdepåstående tillhöra den förra, medan rena värdekataloger kommer att räknas till den senare. Men teser av den typ som görs inom objektolkningen förblir svåra att kategorisera. Att värde konstitueras av preferenser är en tes om värdets natur och inte en utsaga om meningen hos värdepåstående. Som sådan är den en filosofisk tes med viktiga substantiella implikationer. Den tycks överbrygga gränsen mellan det metaetiska och det axiologiska.

Dessutom, även om objektolkningen ändå kategoriserades som primärt en metaetisk ståndpunkt, kan invändningen fortfarande bemötas genom hänvisning till rådande praxis. Det förefaller som om såväl filosofer som icke-filosofer i sina diskussioner av preferensutilitarism ofta (om än ej alltid) har i tankarna en eller annan version av objektolkningen, med värden tillkommande objekten för preferenser. (Jämför t ex Sen och Williams' inledning till deras välkända antologi *Utilitarianism and Beyond*, 1982, eller Dan Egonssons lundsiska avhandling om preferensutilitarism *Interests, Utilitarianism, and Moral Standing*, 1993.) Dessutom bör man ju inte glömma att den deltagarmodell som ligger till grund för objektolkningen har en klart utilitaristisk karaktär. Den metaetiska karaktären hos objektolkningen skulle därför inte diskvalificera den från att vara en version av preferensutilitarism. I stället skulle preferensutilitarismen själv visa sig vara en metaetisk uppfattning, i en av sina vanliga tolkningar.

BJÖRN ERIKSSON

Utilitarism för syndare

Om utilitarismen är riktig så är antagligen de flesta handlingar som utförs moraliskt felaktiga. I de flesta situationer skulle man nog ha kunnat göra något bättre, och det är, i grova drag, tillräckligt för att vi ska handla moraliskt fel enligt utilitarismen. Blotta mängden felaktiga alternativ jämfört med mängden riktiga alternativ (normalt kanske endast ett) gör det nog klokt att gissa att man handlar fel, om man nu skulle undra. Dessutom bryr vi oss nog sällan om att maximera värde, även i de stunder vi bryr oss om mängden producerat värde. Vi nöjer oss med att "satisfiera" i stället för att maximera. Kort sagt kan man nog vara lugn för att få eller inga handlingar är riktiga om utilitarismen är det.

Många filosofer ser här stoff till en vederläggning av utilitarismen. Inte nog med att vi sällan eller aldrig handlar rätt enligt utilitarismen, de flesta av oss ser inget problem i det. Moralen, sägs det, bör avspegla vår moraliska psykologi och den tycks inte vara inte särskilt utilitaristisk. (Se exempelvis Rawls 1972: 145; Mackie 1977: 133; Scheffler 1982: kap. 3; Nagel 1984: 201.)

Andra, som jag, lutar nog åt att felet finns hos våra handlingar och inte hos utilitarismen. Men utilitarismen går nog inte helt fri från klander i det här sammanhanget. Den är oacceptabelt *tyst* vad gäller alla våra felaktiga handlingar. Den tycks inte ha mer att säga om dem än att de är felaktiga.

Det finns mer att säga om våra felaktiga handlingar, t ex att vissa av alla dessa handlingar verkar vara betydligt värre eller mer fel än andra och att vissa orätta handlingar är nästan inte orätta alls.

Det har varit märkvärdigt tyst bland filosofer på den här punkten. Vissa antydningar i rätt riktning finns dock. John Stuart Mill skrev exempelvis i sin essä om utilitarismen att den som "accepterar Nyttan eller Den Största Lyckans Princip, anser att handlingar är riktiga i

proportion till deras tendens att producera lycka och orätta till deras tendens att producera lyckans motsats" (1969: 210 [Min översättning och kursivering]). Mill utvecklar inte detta vidare, men tanken att handlingar är *mer* orätt om de producerar mer lidande, allt i övrigt lika, antyds tveklöst av formuleringen. Några mer samtida exempel på antydningar i denna riktning kan man finna i Dancy (1993: 47), Mackie (1977: 205), Parfit (1984: 33) och Rawls (1972: 176). Ingenstans utom i min egen *Heavy Duty* (1994), kap 6, har jag sett tanken ges en utförligare behandling. (Tyvärr representerar en del av det jag där hade att säga om detta ett passerat stadium inom forskningen på området. (Se Tersman 1995.))

Utilitarismens tystnad i det här avseendet beror naturligtvis på att den i sin klassiska utformning är en teori endast om de diskreta deontiska kategorierna "rätt", "fel" och "plikt". Utilitarismen kan dock utvidgas till en teori som även tillåter omdömen om *grader av moralisk oriktighet*. Och sådana omdömen kan tråkigt nog alltså antas vara de vi normalt har anledning att fälla. I den här artikeln ska jag lägga fram en teori om just grader av moralisk oriktighet avsedd att komplettera den klassiska handlingsutilitarismens riktighetskriterium.

Vissa av våra faktiskt utförda handlingar är speciellt intressanta: de som har en mycket krävande handling som riktigt alternativ. Litteraturen är full av exempel på sådana handlingar: Du har att välja mellan (a) kasta dig över en tappad handgranat och rädda dina kamraters liv eller (b) kasta dig ned i skyttegravan och rädda ditt eget skinn. Det uppoffrande valet (a) skulle medföra bäst konsekvenser. Då man normalt utgår från att det inte vore moraliskt fel att avstå från (a) fungerar dessa sådana exempel som en vederläggning av utilitarismen. (Exemplet kommer från Urmson 1958.) En förhoppning jag har på en utvidgad teori om handlingars grad av felaktighet är att den låter oss visa att handlingar som (a) i exemplet tenderar, just därför att de kräver så stora uppoffringar, att vara orätta i lägre grad. Om den förhoppningen kan infrias, så skulle vederläggningar av utilitarismen som har sin grund i exempel som det ovan förlora något av sitt brett. I så fall skulle ju utilitarismen i viss mån vara sensibel för våra responser på exempel som det med handgranaten.

En annan förhoppning är att teorin ska kunna belysa det för utilitarister gåtfulla begreppet *klandervärdhet* (Till skillnad från frågan om vem och när man bör klandra): att en handling är mer klandervärd ju

felaktigare den är låter ju inte så dumt. Jag kommer att visa att dessa två förhoppningar mycket väl kan komma att infrias.

Först ska jag presentera idéerna bakom teorin om grader av felaktighet i grova drag. Därefter kommer jag att utveckla en teori om grader av felaktighet, först för förenklade två-alternativs fall, dvs fall där agenten endast har att välja mellan en riktig och en orätt handling, sedan för normala fler-alternativsfall.

1. Grunder

Vad är det som bestämmer grader av oriktighet? Jag har funnit följande två variabler lovande: relativ svårighet att utföra handlingar och skillnader i realiserad värdemängd för alternativa handlingar. De här två variablerna framstår som lovande eftersom de anknyter till två typer av överväganden som utilitarister brukar fästa vikt vid, nämligen mängden värde som produceras av olika handlingsalternativ och principen att "bör" implicerar "kan". Förbindelsen mellan "'bör' implicerar 'kan'" och grader av felaktighet kanske inte är uppenbar. Tanken är lite grovt att på samma sätt som du inte alls handlar fel om du "avstår" från att utföra en handling skulle ha haft de bästa konsekvenserna men som du inte kan utföra, så verkar det tilltalande att säga att om du nästan inte kunde utföra en handling (den var mycket svår att utföra) så handlade du inte så väldigt fel när du avstod från den.

Den intuitiva idén som ska ges ett mer precist uttryck är alltså helt enkelt att en handling är mer fel ju mer värde som gick förlorat i och med att den utfördes och mindre fel ju svårare det hade varit att handla bättre.

Att jämföra handlingars värden är inget nytt för utilitarister, men relativa svårigheter är en nyhet i sammanhanget. Jag ska inte ge någon analys av det svårighetsbegrepp som här antyds. Jag har ännu inte sett, eller själv utformat, någon sådan. Jag hoppas att det kan göras. Tills dess kan jag i varje fall säga några saker om hur svårighetsmättet inte får vara konstruerat.

Mättet kan inte vara baserat på de värden som tillfaller eller berövas agenten så att svårigheten att utföra en handling exempelvis består i att den orsakar agenten lidande. Alla sådana vinster och förluster är räknade och bokförda under rubriken "Handlingens värde". Att ta ytterligare hänsyn till dem genom att också föra upp dem under rubriken "svårighet" gör teorin smygegoistisk, genom att vi räknar

upp, eller helt enkelt räknar fler gånger, det värde agenten själv kan inkassera. Det här hindrar inte att en värdeförlust, t ex lidande, som drabbar agenten vid utförande av en handling kan vara orsak till och förklara att handlingen är så svår som den är.

Måttet kan inte heller vara baserat på relativa frekvenser av t ex lyckade utföranden: det är partikulära handlingars svårighet vi är intresserade av. Det här hindrar inte att relativ frekvens av lyckade utföranden kan vara evidens för hypoteser om grad av svårighet för en partikulär handling av en viss typ.

2. Teorin

Vi tummar inte på den klassiska utilitarismens kriterium för riktighet och felaktighet: En handling är rätt om och endast om det inte finns någon alternativ handling som leder till bättre konsekvenser, i annat fall är den felaktig.

För en handling h gäller att dess grad av felaktighet, Fh , bestäms av den relativa svårigheten att utföra h och av hur underlägsen h är sina alternativ med avseende på värdeproduktion, h 's "värdedifferens". Graden av oriktighet är alltså en funktion av värdedifferensen och den relativa svårigheten att utföra handlingen, men vilken funktion? För att besvara den frågan måste vi veta lite mer om hur värdedifferens och relativ svårighet ska förstås.

3. Den intuitiva idén bakom värdedifferens

Till en början håller vi oss som sagt till enkla situationer med endast två alternativ, ett riktigt och ett orätt.

Betrakta följande exempel. Vi har tre personer A , B och C . De har var och en att välja mellan två handlingar h_1 och h_2 som resulterar i värden enligt följande:

| | A | B | C |
|-------|-----|-----|-----|
| h_1 | 100 | 11 | 101 |
| h_2 | 10 | 10 | 100 |

Anta att de alla gör sin h_2 -handling; att h_2A , h_2B och h_2C utförs, kort sagt. Vilken av dem är mest felaktig? Det verkar vara h_2A ; den leder till att 90 värdeenheter går förlorade medan h_2B och h_2C medför att endast 1 förslösas. h_2A och h_2B resulterar båda i lika mycket värde, men deras alternativ är mycket olika vad gäller värde. Min uppfattning

att h_2A är en värre försyndelse än h_2B kan förslagsvis förklaras av att denna differens mellan den utförda handlingen och dess riktiga alternativs värde är större för h_2A . Men h_2C då? Är den mer, mindre eller lika fel som h_2B ? Lika fel, tycks det. Eftersom det är förlorat värde som anges i matrisen spelar inte vikande marginal nytta någon roll för exemplet. Och eftersom det är förlorat värde som räknas är det rimligt att det just differensen som har betydelse. Jag ska beteckna det relevanta måttet "Vrh" (värdedifferens mellan den riktiga handlingens, r , värde och den utförda handlingens, h , värde), och uttrycka resultat som det ovan på följande vis: $Vrh_2A > Vrh_2B = Vrh_2C$.

4. Den intuitiva idén bakom relativ svårighet

I nedanstående matris har handlingsalternativens värden kompletterats med uttryck för olika grader av svårighet att utföra handlingarna.

| | A | B |
|-------|-------------------|-------------------|
| h_1 | 100 (mycket svår) | 100 (mycket lätt) |
| h_2 | 10 (medelsvår) | 10 (medelsvår) |

h_2A och h_2B är de handlingar som utförs. Det tycks mig här tilltalande att betrakta h_2B som mer felaktig och mer klandervärd än h_2A . Skillnaden förklaras så klart av att h_1A är svårare än h_1B .

I fortsättningen skall jag uttrycka grader av svårighet på en skala från 0 (omöjligt) till 1 (oundvikligt), på så sätt att exempelvis "100^{0,5}", associerat med en handling skall förstås så att handlingen resulterar i 100 värdeenheter och är medelsvår (0,5) att utföra.

Att fixera slutpunkterna på detta sätt förtjänar ett par kommentarer. Svårighetsmålet tycks ha några likheter med sannolikhetsmålet. Liksom sannolikheten att *något* händer är 1, så är det oundvikligt (svårighetsgrad 1) att man utför något av de alternativ man står inför i en situation, man kan inte "göra ingenting". Det verkar heller inte otroligt att sannolikheten för att en handling utförs som är mycket svår att utföra är lägre än sannolikheten för att en lätt handling utförs, allt i övrigt lika. Men denna likhet med sannolikhetsmålet ska behandlas med försiktighet. I synnerhet ska den inte lura oss att låta alla alternativs svårighet summera till 1, de summerar till *minst* 1. (Att låta alla alternativs svårighet summera till 1 skulle leda till att en handlings svårighet blev avhängig mängden alternativ, och det skulle sätta stopp för jämförelser mellan handlingar som har olika antal alternativ.)

Förhållandet mellan alternativs svårighet, deras relativa svårighet, tycks mig vara det intressanta för ett mått på felaktighet. Närmare bestämt är det kvoten av det riktiga alternativets svårighetsgrad och det utförda alternativets som tycks vara av intresse. Jag kallar detta för "Srh" d.v.s. handlingen h :s relativa svårighet jämfört med (dividerat med) det riktiga alternativets, r . Vi har nu nått fram till den första, (två-alternativs-) versionen av teorin:

En handling, h , grad av felaktighet, Fh , = $Vrh \times Srh$.

Vrh = Värdedifferensen, dvs den riktiga handlingens värde – h :s värde.

Srh = Den relativa svårigheten, dvs den riktiga handlingens svårighet/ h :s svårighet.

Teorin förutsätter att värdena kan sättas på kvotskala. Det är visserligen ett problem, men det är ett grundläggande problem för de enligt min mening mest tilltalande formerna av utilitarism.

Vrh är utformat så att $Vrr = 0$ vilket medför att det riktiga alternativet inte alls är orätt.

Srh är utformat så att: (a) Ju svårare den riktiga handlingen är jämfört med h , desto mindre fel tenderar h att vara. (b) Om h är lättare att utföra än den riktiga handlingen så ökar detta felaktighetsgraden hos h . (c) Om h är svårare än den riktiga handlingen minskar detta felaktighetsgraden hos h . (d) Om de är lika svåra ($Srh = 1$) så inverkar inte deras respektive svårighet på graden av felaktighet. Det är alltså *skillnader* i svårighet som spelar roll. (e) En handling kan inte vara fel därför att den medför ett sämre resultat än en handling som inte kan utföras. (Detta är som synes, en version av devisen "'bör' implicerar 'kan'".)

Vi kan nu luta oss tillbaka och njuta av att iaktta teorin i arbete i några stiliserade exempel:

MATRIS 1

| | A | B | C | D |
|--------|--|-------------|-------------|--------------------|
| h_1 | $100^{0,5}$ | $100^{0,1}$ | $100^{0,9}$ | $100^{0,1}$ |
| h_2 | $10^{0,5}$ | $-50^{0,9}$ | $10^{0,1}$ | $10^{0,9}$ |
| h_2A | $Vrh_2A = 90; Srh_2A = 0,5/0,5=1$ | | | $Fh_2A = 90$ |
| h_2B | $Vrh_2B = 150; Srh_2B = 0,1/0,9\approx 0,11$ | | | $Fh_2B \approx 17$ |
| h_2C | $Vrh_2C = 90; Srh_2C = 0,9/0,1=9$ | | | $Fh_2C = 810$ |
| h_2D | $Vrh_2D = 90; Srh_2D = 0,1/0,9\approx 0,11$ | | | $Fh_2D \approx 10$ |

Summa summarum kan vi notera att följande gäller: $Fh_2D < Fh_2B < Fh_2A < Fh_2C$

Någon kanske skulle vilja invända att det är orimligt att Fh_2B är mindre än Fh_2C och Fh_2A . h_2B resulterar ju i ett mycket sämre utfall. Ja, vad kan man säga om det? Själv finner jag förmågan att generera resultat *av det slaget* tilltalande: en eftersträvad egenskap hos teorin är att ta hänsyn till grader av svårighet. Att h_2B är mindre fel än h_2C är ett resultat av den förmågan. Men det är ju inte säkert att svårighetsmålet ska viktas mot värdedifferensen som ovan (dvs inte alls). Det är svårt att säga något mer bestämt om vikten hos Srh relativt Vrh innan analysen av svårighet har rots i hamn eller är inom synhåll.

En önskan var ju också att teorin skulle låta oss visa att underlåtelser att utföra krävande riktiga handlingar tenderar att vara i reducerad grad felaktiga. Denna önskan infrias om vi kan anta den rimliga tesen att krävande handlingar systematiskt är svårare att utföra än icke krävande handlingar. Det är svårt att ge ett argument för den tesen nu (eftersom vi vet för lite om svårigheten), men den tycks mig som sagt rimlig.

Den andra önskade konsekvensen av denna utvidgade utilitarism var att den skulle tillåta omdömen om klandervärdhet. En trevlig sak vore om teorin tillät oss att säga att handlingars grad av klandervärdhet varierar med deras grad av felaktighet. Detta tycks mig heller inte orimligt för de fall vi har tittat på, även om teorin måste kompletteras med exempelvis ett tröskelvärde för klandervärdhet om vi ska undvika den moralistiska konsekvensen att varje felaktig handling är klandervärd.

Det är nog (inom parentes sagt) hit debatten om "moralens gränser" och "hjälteedåd" versus "plikt" borde flyttas. Jag föreslår att vi ser dem som frågor om var i de felaktiga handlingarnas rum klandervärdhetsgränsen går. Dvs att vi betraktar som hjälteedåd sådana handlingar som är överlägset värdeproduktiva jämfört med ett alternativ som är mycket mindre krävande och mycket mindre värdeproduktivt men ändå inte faller under klandervärdhetsgränsen.

5. *Många alternativ*

I samma ögonblick som begränsningen till två alternativ, ett rätt och ett fel, lämnas stöter teorin på problem. Den hotas t ex av följande typ av fall, som Folke Tersman (1995) har fäst min uppmärksamhet på:

MATRIS 2

| | A | B |
|-------|--------------|--------------|
| h_1 | $100^{0,01}$ | $100^{0,01}$ |
| h_2 | $0^{0,9}$ | $99^{0,9}$ |
| h_3 | $0^{0,09}$ | $0^{0,09}$ |

Antag att h_3A och h_3B utförs. Eftersom h_1A och h_1B är de riktiga handlingarna ger teorin utslaget att $Fh_3A = Fh_3B$. Men h_3B verkar ju vara betydligt felaktigare än h_3A eftersom den har ett så attraktivt alternativ. Även den lataste samarit skulle kunna tänkas gripa sig an h_2B , den går ju nästan inte att undvika. Dessutom är dess utfall nästan lika bra som h_1B :s.

Ett första steg mot en lösning är att notera möjligheten att relativisera värderingar av grad av felaktighet: h_3B jämfört med h_1B är lika fel som h_3A jämfört med h_1A , men h_3B jämfört med h_2B är värre än h_3B jämfört med h_1B . Detta drag saknar dock brett om det inte kompletteras med en idé om vilken av alla dessa jämförelser som är relevant.

En intressant synpunkt på problemet i matris 2 är att om h_2B och h_3B vore de enda alternativen i B så skulle h_3B vara betydligt mera fel än h_3A , närmare bestämt är förhållandet som det mellan 990 och 11. Vid närmare eftertanke tycks det mig som om den jämförelsen också är den relevanta. Vad hos h_2B gör den till det alternativ som bestämmer Fh_3B ? Den nya versionen av teorin, hädanefter "teorin", ger svaret.

En handling h :s grad av felaktighet, Fh , = $Vmh \times Smh$

m = den handling som maximerar Fh .

m kan, men behöver inte, vara den riktiga handlingen. I matris 2 är t ex m för h_3B h_2B . Den nya teorin ger det önskade resultatet att h_3B är betydligt mer fel än h_3A . Teorin ger naturligtvis också rätt svar i matris 1.

Finns det någon förklaring till att den nya teorin ger rätt svar i matris 2 som ger anledning tro att den gör det även i andra fall? Ett skäl att vara optimistisk på den punkten är att det tycks rimligt, vid närmare eftertanke, att en teori för grader av felaktighet har resurser att ta hänsyn till vilket alternativ som helst, eller i vart fall inte på förhand sållar ut något som irrelevant. Den nya teorin med sin m -variabel har den egenskapen.

Ett annat skäl att tro att teorin är korrekt, är att den uttrycker den rimliga (men lite dystra) principen att *en felaktig handling är så felaktig den kan vara*. Följande analogi kanske klargör tanken. Anta att du har gjort en dålig affär. Du har köpt en tysk trädgårdstomte för 1000 kronor. *Hur dålig är affären?* Ja, det beror på alternativen tycks det. Men det beror inte på *alla* alternativa affärer du kunde ha gjort. Det beror på hur dålig den affär du har gjort framstår vid en jämförelse med *det alternativ som får affären att framstå som mest misslyckad*. Det beror inte på vilken alternativ affär som skulle ha kostat dig minst (du kunde ha behållit alla dina pengar), och det beror inte på vilken affär som skulle ha givit dig mest (du kunde tagit lån och köpt ett gross tomtar). Snarare är det något vi kan kalla "allt-inräknat-tillfredsställelse" (kostnad, nöje, förlorade möjligheter, etc.) som är intressant. Hur dålig din affär är beror på hur lite sådan tillfredsställelse den ger dig jämfört med den affär som skulle ha givit dig mest sådan tillfredsställelse. Det är irrelevant i sammanhanget att det finns en mängd möjliga affärer som faller mellan den affär jämfört med vilken den affär du gjorde framstår som mest korkad och den du gjorde. För affärer verkar det gälla att en dålig affär är så dålig den kan vara. Samma gäller, föreslår jag, moralisk felaktighet.

6. Till vilken nytta?

Det kan hända att någon läsare har glömt bort att funderingarna om grader av felaktighet har försiggått mot bakgrund av ett klassiskt utilitaristiskt riktighetskriterium och att en felaktig handling fortfarande är felaktig och inte bör utföras, oavsett hur lite fel den är.

Detta innebär att det inte klart vad teorin om felaktighet är bra för, och den frågan är av viss vikt, ty det kan vara svårt att finna skäl att acceptera teorin om den inte kan visas vara till någon nytta. Avslutningsvis ska jag därför summariskt försöka ange några funktioner för Teori 2.

(1) *Ökad förståelse*. Vi får resurser att öka vår moraliska förståelse av handlingar och våra bedömningar av dem. Detta är ju ett stående filosofiskt intresse som väl skulle kunna motivera teorin. Men den ger oss ökad förståelse endast om vi har skäl att tro att den är riktig, och det förutsätter enligt en vanlig uppfattning – som jag inte vill ifrågasätta här – att den kan bidra till att förklara något, helst en eller annan observation. Men det kanske den kan, vid närmare eftertanke.

(2) *Förklaring av normativa omdömen.* Det är inte otroligt att teorin kan brukas till att förklara (och hjälpa till att systematisera) omdömen om relativ felaktighet. Mina omdömen om handlingarna i matriserna 1 och 2 tycks t ex kunna förklaras av teorin. Jag antar att mina omdömen inte är helt ovanliga och att teorin alltså skulle kunna förklara även andras omdömen.

(3) *Beslutsmetod.* Ur utilitaristisk synvinkel är det naturligt att ibland fundera över alternativens värden då vi fattar beslut. Tyvärr är det väl ofta så att vi har skäl att tro att ingen av de handlingar vi överväger att ta itu med är riktig, och då verkar det klokt att i varje fall fundera över vilket av de alternativ vi på allvar överväger att utföra som medför den minsta värdeförlusten. Min teoris känslighet för våra alternativs värden kan spela en roll för beslutsfattande av den här typen. Men för en utilitarist kan det tyckas omotiverat att fästa vikt vid *hur svårt* ett handlingsalternativ är att utföra.

Men som tidigare antytts har vi skäl att tro att sannolikheten för att vi ska komma att utföra en viss handling står i proportion till dess grad av svårighet. Svårighet och sannolikhet tycks vara indirekt korrelerade på så vis att svårighet är evidens för osannolikhet. Det verkar klokt att inte planera för en framtid i vilken man utför ett mycket svårt och därmed (antagligen) osannolikt alternativ. (Jag tackar Torbjörn Tännsjö för att ha fäst min uppmärksamhet på detta.)

Vidare tycks det som om svåra, och därmed osannolika, handlingar som leder till förhållandevis mycket värde påkallar "stödåtgärder", exempelvis i form av straff för underlåtelser snarare än bara klander. Sådana åtgärder gör det nog i vissa fall lättare att göra gott. Men här får man vara försiktig. Eftersom de flesta utförda handlingar är fel om utilitarismen är riktig vore det nog ingen god idé att lagstifta mot till exempel felaktiga handlingar.

I det här sammanhanget är det intressant att notera att olika typer av handlingar tenderar att vara olika svåra för olika personer. Det tycks mig till exempel rimligt att tänka sig att det är svårare att utföra en handling man uppfattar som motbjudande. Många icke-utilitarister uppfattar vissa typer av handlingar som kan vara riktiga enligt utilitarismen som motbjudande; att anordna gladiatorspel för många uttråkade, att koka tvål på en av många smutsiga, att döma en oskyldig för att undvika ett upplopp etc. En tilltalande konsekvens av teorin är att icke-utilitarister därför tenderar att handla mindre fel än utilitarister

om de inte utför handlingar som de nämnda (i de osannolika situationer de vore rätt). Samvetets röst bestämmer alltså inte vilka handlingar som är rätt, men den har ett visst inflytande över hur fel våra felaktiga handlingar är.

För övrigt gäller detta även för utilitarister. Enligt många utilitarister, exempelvis Henry Sidgwick (1907) och Richard Hare (1981), så är det viktigt att hålla sig med moraliska beslutsregler som tillåter agenten att fatta snabba beslut och hjälper henne att undvika omotiverade uppvärderingar av sin egen välfärds vikt. Vidare kan sådana regler utformas så att hänsyn tas till att konsekvenserna av ens handlingar ofta beror på vad andra gör eller förväntar sig. För att fungera effektivt bör nog reglerna vara väl integrerade i agentens psyke, i synnerhet Hare har givit emfas åt detta. Det verkar kort sagt rimligt att anta att effektiva beslutsregler för utilitarister gör vissa alternativ (som att bryta ett löfte till en vän) motbjudande. Tråkigt nog är det ibland fel, utilitaristiskt sett, att hålla sina löften till vännerna. Men med glädje kan vi notera att det tenderar att vara i lägre grad fel i kraft av att det framstår som motbjudande och därför är svårt. Åtminstone kan detta gälla för så kallade indirekta utilitarister av Hares slag.

(4) *Retributivistisk rättskipning*. Många anser att ett straff bör stå i proportion till brottet. Detta brukar väl innebära att straffet tänks variera med den skada brottet orsakar (eller den skada brott av det slaget tenderar att orsaka). Denna brottets skadlighet modereras av så kallade förmildrande omständigheter innan straffsatsen avgörs. Att det var svårt att undvika en brottslig handling kan tänkas vara en sådan omständighet. Kanske borde man gå längre och hävda att alla förmildrande är sådana omständigheter som gör det svårt att avstå från brottet. Om detta är rimligt kan den här föreslagna teorin utgöra grunden för en teori om retributiv rättvisa, genom att den (när den har kompletterats med en specifik viktning av *Vmh* relativt *Smh*) just väger samman skadan och de förmildrande omständigheterna. Jag lämnar dessa möjligheter åt retributivisterna.

Litteratur

DANCY, JONATHAN: *Moral Reasons*, Blackwell, Oxford UK and Cambridge USA, 1993

ERIKSSON, BJÖRN: *Heavy Duty. On The Demands of Consequentialism*, Almqvist & Wiksell International, Stockholm, 1994

HARE, RICHARD M: *Moral Thinking*, Clarendon Press, Oxford, 1981

MACKIE, JOHN L.: *Ethics, inventing right and wrong*, Penguin, Harmondsworth, 1977

MILL, JOHN STUART: *Utilitarianism* (1863), i *The Collected Works of John Stuart Mill*, vol. X, University of Toronto Press, Toronto, 1969

NAGEL, THOMAS: *The View From Nowhere*, Oxford University Press, New York, Oxford, 1986

PARFIT, DEREK: *Reasons and Persons*, Clarendon Press, Oxford, 1984

RAWLS, JOHN: *A Theory of Justice*, Oxford University Press, Oxford, 1971

SCHEFFLER, SAMUEL: *The Rejection of Consequentialism*, Oxford University Press, Oxford, 1982

SIDGWICK, HENRY: *The Methods of Ethics*, 7th ed. Macmillan, London, 1907

TERSMAN, FOLKE: "Recension av *Heavy Duty*", i *Filosofisk Tidskrift*, nr. 3, 1995.

URMSON, J. O: "Saints and Heroes" i A. I. Melden: *Essays in Moral Philosophy*, University of Washington Press, Seattle, 1958.

RECENSION

Anette C Baier, *A Progress of Sentiments – Reflections on Hume's Treatise*, Harvard University Press, 1991

Hume som dialektiker och inte som skeptiker, det är den lite förvånande bild Anette Baier ger i boken *A Progress of Sentiments – Reflections on Hume's Treatise*. Baiers tolkning blir möjlig när man, som hon, läser *hela Treatise* och inte hoppar över den viktiga men ofta bortglömda andra boken, *Of the Passions*. Men den annorlunda tolkningen får inte bara stöd i Humes analys av passionerna, utan också i verkets komposition: Baier ser *Treatise* som en medvetet komponerad och sammanhållen text.

Det som styr Baiers läsning av *Treatise*, och det som hon tror styr Humes utveckling från solitär rationell tänkare till en känslans, erfarenhetens och gemenskapens försvarare, är hans besvikelse på förnuftet (i smal mening). Hume kan inte få teorierna om kausalitet och realism att gå ihop. Visserligen är de externa objektens (liksom jagets) existens en förutsättning för kausaliteten, men i och med Humes tro på uppdelningen mellan primära och sekundära egenskaper hos tingen utplånar just kausaliteten objekten. Hur skulle de

sekundära egenskaperna, våra kausala intryck av tingen, kunna finnas där när vi inte upplever dem? Och de primära egenskaperna – substansen eller bäraren av de sekundära (eller jaget) – kan vi enligt Hume inte ha någon kunskap om (vi kan ju inte erfara dem). Förnuftets slutsatser är paradoxala: "Thus belief in causal regularities is saved by the postulation of external objects, yet destroys its saviour" (B 6).¹ Humes självpåtagna uppgift blir därför att undersöka, utveckla och utvidga förståndet.

I första boken av *Treatise* plockas kausaliteten isär. Humes förnuft kan inte upptäcka något nödvändigt samband mellan orsak och verkan. Inte förrän erfarenheten eller vanan, analyserad som ett reflekterat intryck, anlitas kan Hume få ihop bitarna igen. Vanan producerar nödvändigheten, den induktiva slutledningen producerar lagen, inte tvärt om. Men det är en nödvändighet och en lag som, enligt Baier, stämmer överens med naturen.

I slutet av samma bok är det jagets, det *andliga* jagets, tur att dekonstrueras. Rekonstruktionen kommer i början av nästa bok. Idén om jaget (nu ett kroppsligt och socialt jag) produceras av en känsla,

nämligen av stoltheten, ett annat reflekterat intryck. Men det är ett jag som är beroende av andra för sin existens. Stolthet är en social känsla. Jaget, som kanske har något att vara stolt över, värderas utifrån, via andras ögon. Detta jag återkommer i den tredje och sista boken, *Of Morals*, både som moralens subjekt och som dess objekt.

Denna bok inleds med den komplicerade, artificiella dygden rättvisa innan den övergår till de naturliga dygderna. Rättvisan, tror Hume, är grunden för samhället och kanske nödvändig för människosläktets fortlevnad. Och det är förstås i samhället, i gemenskapen, som den moraliska känslan har sin största betydelse. Den moraliska känslan är också den ett reflekterat intryck, och liksom vanan och stoltheten är den produktiv: den producerar handling.

Dessa tre teman, kausaliteten, stoltheten och den moraliska dygden, uppfattar Baier som *Treatises* huvudteman. De utgör också Humes fasta grund, eftersom de är, som Baier kallar det, reflexiva: "can bear their own survey". Det är människan, inte Gud eller metafysiken, som Hume undersöker. Förklaringen av kausaliteten innehåller kausalitet, vilket brukar uppfattas som en ond cirkel, men ses av Baier som en dialektisk – eller spiralisk – utveckling mot större klarhet. Stolthet (i Humes mening snarare sund självkänsla) är något att vara stolt över och fungerar som en metadygd –

den stolta människan växer i spiraler. Och för att inta den moraliska ståndpunkten och göra moraliska bedömningar krävs just de moraliska dygderna, till exempel sympati, förståelse, fantasi – *a progress of sentiments*.

Baiers största omtolkning av Humes kausalitetsanalys rör kanske egentligen hans egen attityd till slutsatsen att den nödvändiga relationen mellan orsak och verkan är en projektion av våra medvetanden på yttrevärlden: "necessity is something that exists in the mind, not in objects" (T Bok I 217). Den nödvändiga relationen är varken en "filosofisk" (ungefär logisk) relation, eller något vi kan observera med våra sinnen. Vi har visserligen en humesk idé om nödvändigheten,² men den idén är inte en kopia av ett sinnesintryck, utan en kopia av en känsla, som i sin tur beror på erfarenheten eller vanan. Men att varken sinnena eller förnuftet kan upptäcka nödvändigheten är inget misslyckande som måste leda till tvivel. Tvärt om, känslan och erfarenheten är för empiristen Hume en säkrare grund än förnuftet, tror Baier. Och visst är det en möjlig läsning: Hume förvandlar ju till och med förnuftet (i smal mening) till instinkter eller vanemönster. "To consider the matter aright, reason is nothing but a wonderful and unintelligible instinct in our souls..." (T Bok I 230).

Humes berömda definition av *orsak* är dubbel: dels en, som han

kallar det, filosofisk och dels en naturlig. Enligt den filosofiska är en orsak "An object precedent and contiguous to another, and where all the objects resembling the former are placed in like relations of precedency and contiguity to those objects that resemble the latter" (T Bok I 221), och enligt den naturliga "an object precedent and contiguous to another, and so united with it that the idea of the one determines the mind to form the idea of the other, and the impression of the one to form a more lively idea of the other" (T Bok I 221).

Hume skriver själv att de båda definitionerna innehåller ett "främmande objekt" i den meningen att de refererar till något utanför själva orsaken: den filosofiska till *alla* orsak-verkan-relationer av samma typ, och den naturliga till medvetandets disposition att dra kausala slutsatser. Det är dessa "främmande objekt" som gör definitionerna inkorrekta enligt många tidigare kommentatorer (inklusive Hume själv?). Felet med den filosofiska analysen är att vi aldrig kan erfara *alla* instanser av någon typ av kausal relation, och därför aldrig – enligt Humes egen epistemologi – vara förvissade om dem. Den naturliga definitionen är uppenbart cirkulär, medvetandet är *determinerat*, det vill säga *orsakat*, att fungera på ett visst sätt – nämligen *kausalt*. Så här beskriver Norman Kemp Smith det i sin *The Philosophy of David Hume*:

Now, clearly 'determination' is here more or less synonymous with causation. His /Hume's/ use of it in his definition of causation was, however, unavoidable. What he has set himself to give is a *causal* explanation of our belief in causation as holding *between objects*, by pointing to their connexion, *their causal connexion, in the imagination* (s 401).

Baier däremot uppvärderar de "främmande objekten": de gör definitionerna reflexiva och därmed själv-verifierande. Det kausala tänkandet om kausaliteten ger definitionerna en meta-kausal dimension. Med en blinkning till Hegel skriver hon att

Hume can be confident that his definition does present the truth about our concept of cause, precisely because the definition displays how the concept, under this analysis, can successfully be turned on itself. It is a self-demonstrating and self-verifying real definition, arrived at by causal investigation into our concept of cause, and itself joining together the various results of that long and complex exercise in reflexivity. Empirical thinking is here thinking itself (B 91).

I Baiers ögon är Hume en tidig holist. Men det är också fullt möjligt att läsa Humes egna kommentarer till definitionerna som självkritiska. Definitionerna är de bästa han kan

åstadkomma. De "främmande objekten" kan han inte bli av med, men om någon annan lyckas så är han eller hon välkommen. Kanske vore det därför riktigare att se Hume som en visserligen tidig men omedveten eller ofrivillig holist. Hume själv, fortfarande övertygad om kausalitetens nödvändighet och inneboende kraft, har lite svårt att tro på sin egen naturalistiska definition. Förresten är väl debatten om den nomologiska nödvändigheten fortfarande levande.

Också jag har lite svårt att dela Baiers entusiasm, men av andra skäl (den filosofiska definitionen kan jag acceptera). Naturalismen tycks kräva en meta-meta-kausal relation för att bli självgående. Humes förklaring av kausaliteten, uppfattad som naturalistisk, har också ett finalt drag i aristotelisk mening (trots att han uttryckligen förnekar att det skulle finnas olika typer av orsaker): Hume hänvisar till våra instinkter, människans natur och därigenom indirekt till kausalitetens överlevnadsvärde (förutsägelser, planering etc). Inte undra på att Baier läser Hume som en föregångare till Darwin! Humes tanke, som Baier tolkar den, är att vi uppmärksammar regelbundenheter i världen, våra medvetanden blir av vanan determinerade att dra kausala slutledningar, vi blir känsliga för kausala relationer och uppmärksammar ännu fler regelbundenheter. Men vi kanske inte, som Baier tror, uppfattar naturen som bland annat

kausal för att det passar *den*, för att den är sådan, utan för att det passar *oss* att se den så (om det nu är en meningsfull distinktion). För att tala om regelbundenheter måste vi också kunna avgöra vad som är relevanta likheter mellan olika händelser, vad som gör dem till instanser av samma regel. Att vi inte uppfattar naturen som slumpmässig beror kanske på den antagna finala orsaken – organismens eller artens överlevnad – kanske på något annat. Evolutions teorin med sitt empiriska-naturalistiska sken är, tycks det mig, naturalismens egen hus-metafysik.

Och när Baier går så långt att hon säger att "Hume has the makings of a conventionalist account of mathematical and logical necessity" (B 100), undrar jag om hon inte gör sig skyldig till övertolkning. Visserligen skriver Hume att det bara finns ett slags nödvändighet, men det i en kontext där han diskuterar fysisk och psykologisk nödvändighet. Den logiska nödvändigheten är för Hume beroende av relationer mellan idéer, "filosofiska" relationer som inses av förnuftet. Den är alltså inte en känsla, eller ett reflekterat intryck. Just detta, att den förnuftsmässiga nödvändigheten *inte* är ett reflekterat intryck, är, mycket viktigt för Hume. Genom hela *Treatise* upprepar han gång på gång att förnuftet är improduktivt, medan känslan eller passionen är produktiv. Och till skillnad från den kausala nödvändigheten så finns den logiska *i objekten*, som visserligen,

liksom den kausala nödvändigheten, är mentala. I *den* meningen finns det bara en sorts nödvändighet – all nödvändighet är mental.

Thus, as the necessity, which makes two times two equal to four, or three angles of a triangle equal to two right ones, lies only in the act of understanding, by which we consider and compare these ideas; in a like manner the necessity of power, which unites causes and effects, lies in the determination of the mind to pass from the one to the other (T 217).

Men hos Baier blir all Humes nödvändighet normativ eller konventionell: "All necessity derives from normative necessity, and all the norms available to us are our human norms, the products of our reflection" (B 100). Jag tror det är att modernisera Hume lite för mycket. Hume, som gör obesvärade och upprepade hänvisningar till människans natur och till våra instinkter, tänkte sig nog att naturen och instinkterna var oföränderliga och ofrånkomliga – vilket gör det underligt att kalla dem för konventioner. Åtminstone i mitt språkbruk har *konvention* en medveten, social eller kulturell betydelsekomponent sedd i kontrast till något omedvetet, naturligt. Humes logiska, kausala och för den delen moraliska normer är visserligen *mänskliga*, men knappast enbart konventionella. Det finns en nödvändighet också i den

mentala nödvändigheten. Instinkten är (var?) omedveten och naturlig.

Så till Humes andra viktiga reflekterade intryck: passionen stolthet. Nu ska varken ordet *passion* eller ordet *stolthet* uppfattas så som vi i dag brukar uppfatta dem. Den humeska passionen behöver inte alls vara våldsam och överväldigande, utan kan mycket väl vara lugn och genomtänkt – en reflekterad känsla. Och Humes stolthet har inget med högfärd eller fåfänga att göra, utan är mer vad vi skulle kalla sund självkänsla eller motiverat självförtroende.

Att stoltheten producerar idén om jaget, som Hume skriver i inledningen av andra boken, har många uppfattat som dubbelt paradoxalt med tanke på hur han avslutar första boken: enligt gängse tolkning ödeläggs både kausaliteten och jaget. Som vi redan sett tycker Baier inte att Hume tvivlar på kausaliteten: att stoltheten *producerar* något är alltså inget problem för henne. Idén om jaget då, hur förklarar hon den? I *A Progress of Sentiments* tolkar Baier den av stolthet producerade idén om jaget som det sociala och kroppsliga jaget, som en person bland andra personer.

Jag tror att också det är en ogrundad modernisering av Hume. Jaget är för Hume en ofrånkomlig fiktion, och därför, lika litet som kausaliteten eller de externa objekten, något vi kan klara oss utan. Vi har en känsla av ett jag, av ett sub-

jekt, även om vi aldrig kan observera det med vår perceptionsapparat. I *Treatises* första bok analyserar Hume som bekant jaget som en ström av intryck och idéer, "a bundle or collection of different perceptions", och det är, enligt hans egen text, samma jag-idé som stoltheten producerar:

It is evident, that pride and humility, though directly contrary, have yet the same *object*. This object is self, or *that succession of related ideas and impressions*, of which we have an intimate memory and consciousness (T Bok II 39, min kursivering).

Att förstå den av stoltheten producerade jag-idén som en annan än den dekonstruerade skulle också förstöra Humes omdiskuterade "dubbla relation" mellan stolthetens orsak och dess intentionala objekt – jaget. I ett av Humes exempel är orsaken till stolthet ett vackert hus. Men detta hus måste på något sätt vara kopplat till mig, för att jag ska kunna vara stolt över det – det räcker inte med att det är vackert. Jaget, eller idén om jaget, som ju produceras av stoltheten, ingår alltså redan i orsaken till känslan. Likaså finns det en relation mellan den känsla av välbehag det vackra huset väcker hos mig oavsett att det är mitt, och den stolthet jag känner när jag relaterar huset till mig själv. Stoltheten är precis som skönhetsupplevelsen en lustfylld känsla. Idén om jaget ingår *både* i orsaken till stoltheten

och i dess objekt. Lust eller välbehag är en komponent *både* i upplevelsen av husets skönhet och i stoltheten. Det är den dubbla relationen mellan orsaken, känslan och objektet.

I uppsatsen "Hume's Analysis of Pride" (*The Journal of Philosophy*, No. 75 (1978), s 27–40) analyserar Baier själv relationen mellan jaget i orsaken och i objektet som begreppslig, det är *samma* jag-idé som figurerar i de båda fallen. (Jag tror att hon har fel, men det är alltför komplicerat att gå in på här.) När hon i *A Progress of Sentiments* i stället sätter in ett annat jag-begrepp (det kroppsliga, sociala jaget) i objektet, det som produceras av stoltheten, förstörs den dubbla relationen, *såvida* hon inte anser att det är det kroppsliga, sociala jaget som ingår i orsaken också. Men om hon vill att relationen mellan jag-idéerna ska vara kausal (och det borde hon vilja, det reflekterade intrycket stolthet *producerar*, det vill säga *orsakar*, jag-idén liksom vanan producerar den kausala nödvändigheten) får hon problemet att förklara hur något kan producera sin egen förutsättning. I den nämnda uppsatsen klarar hon sig, som sagt, undan detta genom att *inte* betrakta stoltheten som produktiv (sambandet är rent begreppsligt, inte kausalt). I boken tar Baier dock inte ställning till problemet om Humes "dubbla relation" ska läsas som kausal, begreppslig eller både och.

En riktig iakttagelse hon gör är

däremot att jaget i första boken av *Treatise* är sett ur första persons perspektiv – ett jag som letar efter sig själv – medan det stolta eller skamliga jaget i andra boken ses indirekt i andra persons perspektiv. Det är ett socialt jag som är beroende av andras värderingar för att värdera sig själv. Det går att tvivla på sig själv, på att man existerar, men det är svårt att tvivla på att det finns andra personer. Å andra sidan är *personer* inte så viktiga för Hume, det moralen handlar om är karaktärer och karaktärsdrag, inte personer. Jag tror därför att han nöjer sig med sitt dekonstruerade subjekt, en sammanhållen social, moralisk agent är inte nödvändig. Baiers betoning av det interpersonella hos Hume luktar lite sent nittonhundratalet.

En annan brist hos Baier är att hon inte riktigt poängterar den intressanta analogi som finns mellan kausalitets- och identitetsresonemanget. Både kausalitetens nödvändighet och jagets identitet produceras av reflekterade intryck, de är fiktioner (fast Hume kallar kausaliteten en *maxim*, inte en fiktion). Men de är, som Baier skriver, inte falskheter, utan just nödvändiga fiktioner som inte kan verifieras. Och i båda fallen misslyckas förnuftet med att identifiera de entiteter Hume letar efter. Detta pekar framåt, mot hans – i eget tycke – viktigaste undersökning: den moraliska. Inte heller moralen kan inses förnuftsmässigt. Den är en känsla,

men för den skull inte irrationell.

Hela *Of the Passions* är en förberedelse för moraldiskussionen. Moralerna är som sagt en känsla, Humes tredje viktiga reflekterade intryck. Han börjar dock med att analysera de icke-moraliska passionerna. Stoltheten och dess negativa motpol skammen är reflexiva känslor i den meningen att de är vad vi känner när vi värderar oss själva. Kärlek och hat är vad vi känner när vi värderar andra. Stolthet och skam, kärlek och hat är indirekta känslor, de orsakas av något sammansatt (mitt fina hus, dina vackra ögon) och har en person (ägaren eller bäraren i orsaken) till objekt, i motsats till direkta känslor där orsaken sammanfaller med objektet. Den moraliska känslan är också den en indirekt känsla, men det som värderas är inte personen utan karaktärsdraget han eller hon har. Orsaken är personen plus karaktärsdraget, objektet är karaktärsdraget. Och till skillnad från de icke-moraliska känslorna är det inte jag utan "vi" (det generaliserade, opartiska jaget) som värderar.

Moralisk stolthet, stolthet över min förträffliga karaktär, är reflexiv också i en annan mening. Stolthet fungerar nämligen som en metadygd, den "can bear its own survey". Jag kan vara stolt över att vara stolt – i Humes bemärkelse – över min karaktär. Jag kan däremot inte vara stolt över att känna skam, det förtar skammen. Detta tar Baier

som en intäkt för att Hume strider emot den kristna etiken och dess ingredienser av synd, skam och självförnekelse, och hon beskriver honom som en uppkäftig, modern ung man. En karaktäristik som Stephen Buckle vänder sig emot i sin recension av *A Progress of Sentiments* (*Australasian Journal of Philosophy*, vol. 70, No 3 September 1992). Enligt Buckle är Hume snarare gammaldags, åberopande antikens "naturliga" moral och ideal, vilket stämmer väl med mina tidigare iakttagelser (den finala orsaken, människans natur eller essens).

Som vanligt tar Baier fasta på de reflexiva och reflekterande dragen när hon beskriver Humes moralsyn. Jag tycker att hon har helt rätt när hon argumenterar emot tolkningar som gör Hume till emotivist:

Strictly speaking, Hume is no emotivist. His theory of morality bases it on reflective passions and correlated sentiments, and thought is as essential to them as it is to pure reasoning (B 180).

Moralen är visserligen en känsla, men inte en irrationell, ogrundad eller oreflekterad känsla. Den har kognitiva inslag genom sin orsak och sitt intentionala objekt, och den känsla som räknas som moralisk är den som vi har när vi opartiskt, lugnt och sansat betraktar en situation (som vi måste känna till alla fakta om!) – med andra ord när vi intar den "moraliska ståndpunkten". På ett känt ställe kallar Hume det

reflekterade intrycket eller passionen för "an original existence" (T Bok II 156), något Baier kallar ett olyckligt misstag. Men vad Hume där menar är bara att passionen inte innehåller något representativt element, den är ingen kopia av något. Han menar inte att den uppstår ur intet, eller är riktad mot intet, tvärt om är den "placed betwixt two ideas, of which the one produces it, and the other is produced by it" (T Bok II 40). Om man analyserar hela paketet som passionen, blir den alltså delvis representerande.

Intentionaliteten tar bort en del av känslans förmodade subjektiva karaktär, men som vanligt tycker jag att Baier överdriver det reflekterande draget. För det första finns det, om känslan rationaliseras alltför mycket, risk för att moralen tappar sin motiverande kraft. Enligt Hume, och det är faktiskt hans huvudargument mot dåtidens etiska rationalister, motiverar passionen till handling, vilket förnuftet inte gör. Passionen är aktiv, förnuftet passivt. För det andra är den moraliska känslan visserligen ett reflekterat intryck och den moraliska bedömarens visserligen eftertänksam och opartisk, men känslan är "naturlig" och många av de moraliska dygderna är "naturliga". Hume är förvissad om att det finns en människans natur. Då blir reflexiviteten, som Baier understryker så, självklar; det blir självklart att moralens "subject matter" finns redan före bedömningen. "Naturligheten" ger

alltså Humes "moraliska ståndpunkt" en fast förankring som Baiers tolkning inte har. Hos henne blir Humes moralteori till ett av många möjliga reflexiva ekvilibrier, där såväl föreställningar och åsikter som känslor och reaktioner vägs in och tillåts påverka varandra.

Inte för att jag vill försvara Humes tro på en mänsklig natur och ett opartiskt, generaliserat jag, kompetent att bedöma andras och min egen karaktär. Inte alls. Jag är inte ens villig att skriva under på alla hans universella dygder. Humör är visserligen bra men många sorters vänlighet kan jag klara mig utan. Och jag föredrar större vadmuskler än Hume. (Hans ideal var konformade ben!) Men det här är problem som Baier förtiger. Den moderna etikett Baier möjligen kan tänka sig att klistra på Hume är "intern realist".

Of course it takes something in a human (or at least in a sensitive animal) breast to recognize and evaluate what is in human breasts, and of course it takes a community of mutual moral evaluators to "create" a particular moral status, such as "depraved" or "of perfect character" (B 196).

Att moralen är ett mänskligt "påfund" håller nog många med om, men hur stor kan eller ska gemenskapen vara? Hur omfattande kan eller ska teorin (inklusive åtminstone vissa av våra känslomässiga reaktioner) vara? Min intuition

säger att ju mer vi universaliserar desto mer måste vi rationalisera den moraliska känslan.

Trots mina små invändningar hoppas jag att det framgår att *A Progress of Sentiments* är en lärorik bok. Framför allt har den fått mig att inse hur viktigt det är att läsa *hela* verk. Men när jag tar *hela Treatise* i beaktande är det, som nog redan märkts, snarare produktiviteten än reflexiviteten som slår mig. Vanan, ett reflekterat intryck, producerar idén om den kausala nödvändigheten, orsaken producerar verkan, induktionen producerar empirisk kunskap. Stoltheten, ett annat reflekterat intryck, producerar idén om jaget. Den moraliska känslan, ännu ett reflekterat intryck, motiverar till eller producerar handling och moralisk kunskap. Passionerna, ytterst vilande på lust och lidande, är den aktiva principen, i motsats till förnuftet som är improduktivt och passivt.

Fokuseringen på motsättningen mellan aktivt och passivt gör på sätt och vis min Hume mycket mer föråldrad än Baiers. Men det visar också på en kontinuitet i den västerländska filosofin som ibland glöms bort. Hume kastar visserligen om passionens och förnuftets förtecken i förhållande till antikens uppfattning, men liksom hos Aristoteles är det hos Hume människans natur eller essens – organismens ändamål – som är den yttersta drivkraften. Är det, som sagt, inte likadant än i dag, fast drivkraften nu ibland antas

vara sexualiteten, ibland överlevnaden?

Nog med reflektioner, *A Progress of Sentiments* är en ganska pratig och lite ostrukturerad bok som följer *Treatise* väldigt nära. Men den är bra. Och för den som vill lära känna Hume är den nödvändig. Baiers uppskattning av Hume är bortom allt tvivel:

The glorious title of inventive philosopher should go to one who not merely shows us how to avoid, or at least to reconcile, the contradictions into which we continually stray, but also demonstrates how we may achieve reflexive self-acceptance, agreement with ourselves (B 277).

Åsa Carlson

1. I alla sidhänvisningar står B för Baier, A C: *A Progress of Sentiments - Reflections on Hume's Treatise*, Harvard University Press, 1991; och T för Hume, D: *A Treatise of Human Nature*, Fontana/Collins, 1962.

2. Som bekant delar Hume in medvetandets "perceptioner" (ung. innehåll) i intryck och idéer, de senare är kopior av sinnestryck. Intrycken är antingen primära (sinnestryck, kroppsliga känslor etc) eller sekundära (emotioner, passioner etc), vilka också kallas reflekterade. De reflekterade intrycken uppstår ur primära intryck eller idéer, men de *avbildar* inte sina orsaker. De är både (eller kanske varken ... eller) emotiva och kognitiva (de kan inte vara sanna eller falska eftersom de inte är kopior av något, men de har analyserbara orsaker och intentionala objekt).

NOTISER

¶ Nelson Goodman säger på ett ställe i sin berömda bok *Ways of World-making* från 1978 att "facts are small theories, and true theories are big facts" (s 97). Härav tycks följa – givet att stora fakta är fakta, vilket väl ingen torde förneka – att sanna teorier är små teorier. Och härav följer i sin tur att stora teorier är falska. Man kan undra om Goodman menar att detta är en naturlag. Och *hur* stor kan en sann teori vara? Är t ex den allmänna relativitetsteorin alltför stor? Och hur mäter man storleken på en teori? Det vore intressant att veta.

¶ *Kuvertproblemet: en kommentar 12/3*

I Filosofisk tidskrift 2/86 presenteras följande beslutsteoretiska problem: "Antag att någon låter dig välja mellan två kuvert, som innehåller pengar, och att du får veta att det ena innehåller dubbelt så mycket som det andra. Då kanske man resonerar så här: Antag att det vänstra kuvertet innehåller x kronor. Då innehåller det högra kuvertet antingen $2x$ kronor eller $x/2$ kronor. Bägge dessa fall är lika sannolika, inget talar mer för det ena än för det andra. Det betyder att sannolikheten är $1/2$ i båda fallen. Och det betyder i sin tur att det förväntade värdet av att ta det högra kuvertet är

$$1/2 \cdot 2x + 1/2 \cdot x/2 = 1,25x \text{ kr}$$

Det förväntade värdet av att ta det vänstra kuvertet är ju däremot

$$1/2 \cdot x + 1/2 \cdot x = x \text{ kr}$$

Alltså är det förväntade värdet större för det högra kuvertet. Men på ett helt analogt sätt kan man – om man i stället startar med det högra kuvertet – visa att det förväntade värdet är större för det vänstra kuvertet. Men detta innebär en motsägelse. Var ligger felet?"

1. Om man betraktar detta problem utifrån klassisk sannolikhets-teori, så måste innehållet i det vänstra kuvertet vara en stokastisk variabel ($\text{in}(V)$). Detsamma gäller även för innehållet i det högra kuvertet. Förväntat värde är nämligen

denna teori bara definierat för stokastiska variabler. Enligt den klassiska teorin är stokastiska variabler funktioner som tilldelar objekt (i det aktuella fallet penningbelopp) till varje utfall i ett slumpförsök. Sett från klassisk sannolikhetsteori innehåller citatet ovan åtminstone tre felaktigheter: (i) påståendet att det förväntade innehållet i det vänstra kuvertet alltid är lika med innehållet i det vänstra kuvertet vid ett försök, (ii) påståendet att det förväntade innehållet i det högra (vänstra) kuvertet alltid är $1,25 \cdot x$ kr och att $1,25 \cdot x > x$ givet att innehållet i det vänstra (högra) kuvertet är x kr, och (iii) påståendet att förutsättningarna implicerar att symmetri råder, dvs att $\text{in}(V)$ och $\text{in}(H)$ är ekvivalenta stokastiska variabler eller att de åtminstone har samma väntevärde. För att upptäcka alla felen krävs det viss förtrogenhet med den klassiska teorin men för att hitta det viktigaste av dem räcker det långt med sunt förnuft: Antag nämligen att förutsättningarna för kuvertproblemet är uppfyllda. Då finns det ett största penningbelopp, säg b , som kan ligga i något av kuverten eftersom det bara finns ett ändligt antal penningbelopp tillgängliga. Om b kan ligga i det vänstra kuvertet, så är det förväntade innehållet i det högra kuvertet givet att innehållet i det vänstra kuvertet är b lika med det faktiska innehållet i det högra kuvertet som alltså är ett belopp som är hälften av b . Alltså gäller det då att det förväntade innehållet i det högra kuver-

tet givet att innehållet i det vänstra kuvertet är x kr åtminstone någon gång är lika med $\leq x$ kr. Om b inte kan ligga det vänstra kuvertet, så kan det ligga i det högra kuvertet. Då kan vi på samma sätt visa att (ii) inte gäller.

Kuvertproblemet illustrerar alltså på ett utmärkt sätt hur ett till synes rimligt och ofta använt sätt att resonera om sannolikheter och väntevärden kan leda till orimliga slutsatser. Vidare ser vi hur problemet helt enkelt upplöses om vi ersätter de naiva begreppen med mer precisa sådana. (Hans Mathlein skall ha stort tack för alla kommentarer och förslag till förbättringar. Även Jan Odelstad och Wlodzimierz Rabinowicz, skall ha tack för värdefulla kommentarer. Den läsare som vill se ett formellt bevis för att påståendena (i) – (iii) är felaktiga kan skicka efter ett sådant hos undertecknad på adress: Filosofiska institutionen, Stockholms universitet, 106 91 Stockholm.)

Per-Erik Malmnäs

¶ *Om svensk filosofi.*

Mer än en filosof har dödförklarat Gud och i våra dagar för han en till synes tynande tillvaro. Men som jag ser det har han (eller hon) tagit en annan, mer diffus och mångfasetterad skepnad. Gud eller snarare religionen finns alltså enligt detta synsätt men den är inte längre lika lätt att identifiera.

Men det är som bekant inte bara Gud som har dödförklarats; filosofin, romanen, ideologierna och punk-

rocken har gått samma bistra öde till mötes. Det förefaller finnas någon sorts mänsklig drift som går ut på att i ett slag förpassa de mest skilda förteelser till historiens skräpkammare. Personligen ställer jag mig skeptisk till den här verksamheten som ofta verkar vara ett utslag av önsketänkande och sensationslystnad.

Utan att vilja ta i för mycket skull jag ändå vilja påstå att filosofin mår dåligt, att hon lider av någon form av åkomma som gör att hon har blivit sängliggande under en längre tid. Jag har själv inte studerat tillräckligt mycket filosofi för att våga ställa en säker diagnos men jag skulle vilja resa några frågor som kanske kan vara intressanta både för lekmän och filosofer av facket.

Min första fråga gäller de svenska filosofernas osynlighet; varför tar inte filosoferna den plats i de offentliga debatterna som deras kunskap fordrar. Det verkar finnas långt fler filosofer än Torbjörn Tännsjö som har intressanta tankar om de mest skilda frågor, filosofer vars förmåga att analysera och argumentera vida överträffar många självutnämnda samhällsdebattörers. Kort sagt; varför läser man så sällan en dagstidningsartikel skriven av en filosof, varför hör man aldrig en filosof uttala sig på teve eller i radio?

Det finns ju en hel rad mycket aktuella ämnen som har mycket nära anknytning till den nutida filosofins verksamhetsområden. Allra närmast kommer jag att tänka på det nyligen

avslutade kriget på Balkan som borde ha kunnat intressera moralfilosofer. På samma sätt borde moralfilosofer kunna känna en viss dragning till den debatt som har den söndervittrande samhällsmoralen på dagordningen. För mig personligen är det beklämmande att tex Alf Svensson tillåts vara en av de tongivande när det gäller moralfrågor. Jag tror helt enkelt att det finns många filosofer som har mer intressanta saker att säga om de här frågorna än vad politiker, programledare på teve och kändisar har.

För att provocera läsaren ytterligare skulle jag vilja påstå att orsaken till att filosofin har hamnat i vanrykte i Sverige beror mest på filosoferna själva. Om man jämför med t ex situationen i Frankrike så framstår situationen än mer beklämmande, I Frankrike syns och hörs filosoferna dagligen i den offentliga debatten. Delvis kan det här naturligtvis ges en mängd förklaringar; Frankrike har en makalös filosofisk tradition, fransmän uppskattar livliga diskussioner och de betraktar det inte i allmänhet som suspekt att i stor utsträckning ägna sig åt tankeverksamhet.

Om man i Sverige berättar att man läser filosofi så får man långt i från sällan frågan "Men vad skall du bli då?". När jag under våren vistades några månader i Frankrike märkte jag en avgörande skillnad i synsätt, filosofi är inte så märkvärdigt det är nödvändigt. När jag vid ett tillfälle berättade att jag

studerade filosofi, så fick jag frågan om jag var cartesian. Vad jag svarade skall jag inte avslöja här men jag försäkrar att det var en märklig upp-

levelse att ställas inför en så avgörande frågeställning i Paris tunnelbana, mitt i natten.

Ola Wihlke

Medarbetare i detta nummer av *Filosofisk tidskrift*: *Mats Furberg* är professor i teoretisk filosofi i Göteborg, *Wlodek Rabinowicz* är professor i praktisk filosofi i Lund, *Björn Eriksson* är lektor i praktisk filosofi i Stockholm, och *Åsa Carlson* är doktorand i praktisk filosofi i Stockholm.

Inbjudan till Filosofidagarna i Lund

13-15 juni 1997

Nationell konferens i filosofi

Vartannat år anordnar Svenska filosofisällskapet en nationell konferens i filosofi med syfte att främja intresset för filosofi och filosofisk forskning. 1997 är Filosofiska institutionen i Lund värd för arrangemanget.

Planerade sektioner (*Sektionsordförande inom parentes*)

- | | |
|---|---|
| 1) Beslutsteori (<i>Sven Danielsson</i>) | 8) Medvetandefilosofi (<i>Peter Gärdenfors</i>) |
| 2) Estetik (<i>Göran Hermerén</i>) | 9) Metafysik (<i>Sten Lindström</i>) |
| 3) Feministisk filosofi (<i>Ulla Holm</i>) | 10) Moralfilosofi (<i>Torbjörn Tännsjö</i>) |
| 4) Filosofins historia (<i>Mats Furberg</i>) | 11) Politisk filosofi (<i>Lars Bergström</i>) |
| 5) Filosofisk logik (<i>Krister Segerberg</i>) | 12) Religionsfilosofi (<i>Dick Haglund</i>) |
| 6) Humanistisk och samhällsvetenskaplig teori (<i>Nils-Eric Sablin</i>) | 13) Språkfilosofi (<i>Peter Pagin</i>) |
| 7) Kunskapsteori (<i>Dag Prawitz</i>) | 14) Tillämpad etik (<i>Lennart Nordensfelt</i>) |
| | 15) Vetenskapsfilosofi (<i>Martin Edman</i>) |

Forskare, lärare och doktorander verksamma vid institutioner för främst filosofi och vetenskapsteori inbjuds att delta och bidra med ett föredrag.

Vi behöver få besked om Din medverkan (ange gärna sektionstillhörighet) senast den *31 januari 1997*.

För vidare information, vänd Dig till

konferenssekreterare *Caj Strandberg*

Filosofiska institutionen, Kungshuset, 222 22 LUND

Tfn: 046-222 32 68 Fax: 046-222 44 24

e-mail: Filosofidagarna@fil.lu.se

Internet: <http://www.fil.lu.se/filosofidagarna.html>

Välkomna!

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

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

MANUSKRIFT TILL FILOSOFISK TIDSKRIFT

- sändes till redaktören, Lars Bergström, Reimersholmsg. 39, 117 40 Stockholm
- skall vara försedda med namn och adress
- skall som regel vara skrivna på svenska, och citat från andra språk bör översättas till svenska
- noter och litteraturhänvisningar bör inarbetas i texten
- särskild litteraturförteckning upprättas i alfabetisk ordning och placeras sist i manus
- manus skall vara maskinskrivna på A4-papper med skrift endast på arkets ena sida och med bred vänstermarginal
- manus bör dessutom bifogas på diskett för Macintosh (disketten kommer naturligtvis att returneras)
- manus kan också skickas per datorpost till adress Lars.Bergstrom@philosophy.su.se
- för icke beställt material ansvaras ej
- korrektur läses i regel endast av redaktören
- ändringar mot manuskript bekostas icke av förlaget; författaren debiteras kostnad för sådana ändringar
- införda bidrag honoreras ej för närvarande
- i stället för särtryck erhåller författaren gratis 10 exemplar av det nummer av tidskriften i vilket bidraget varit infört

Tidigare årgångar av Filosofisk tidskrift kan beställas från Bokförlaget Thales, Redaktionen, Djupdalskvior 5, 620 12 Hemse.

För 1980 — 1991 är priset 50 kr/årg, enstaka ex 15 kr

För 1992 — 1995 är priset 120 kr/årg, enstaka ex 35 kr. Nr 4 1995 75 kr

Porto ingår

