

Februari 2014 Årgång 35 Nr 1

Filosofisk tidskrift

- CAJ STRANDBERG
3 Moralisk motivation – ett dilemma och dess lösning
- LARS BERGSTRÖM
18 Döden som förlust
- HEMO QUARON
30 Nietzsches ”Fallet Wagner”
- DANIEL RÖNNEDAL
35 Aktualism, possibilism och handlingsriktighet
- RECENSIONER
42 Sören Häggqvist om *Tankar som ändrar allt* av Sören Holst
47 Frans Svensson om antologin *Dygdernas renässans*
52 Lars Bergström om *Kort om spelteori* av Ken Binmore
58 Jens Johansson om *Kort om döden* av Lennart Lundmark
- 60 NOTISER

Filosofisk tidskrift

Utges av Stiftelsen Filosofisk tidskrift och Stiftelsen Bokförlaget Thales

Redaktör och ansvarig utgivare: Lars Bergström

Filosofisk tidskrift utkommer med fyra nummer per år

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

Lösnr: 55 kr (inkl. moms)

För prenumerationsärenden kontakta:

Nätverkstan ekonomitjänst, Box 31120, 400 32 Göteborg

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

ekonomitjanst@natverkstan.net

Förlagsadress: Bokförlaget Thales, Box 50034, 104 05 Stockholm

info@bokforlagetthales.se www.bokforlagetthales.se

Copyright © Respektive författare 2014

Grafisk form: Ulf Jacobsen

Satt med Indigo Antiqua

Tryck: Carlshamn Tryck & Media AB, Karlshamn 2014

ISSN 0348-7482

Moralisk motivation – ett dilemma och dess lösning

1. INTRODUKTION

En av de frågor som diskuteras mest i metaetik – det teoretiska studiet av moralen – är relationen mellan vårt moraliska språk och vår motivation, eller benägenhet, att utföra olika handlingar. Det verkar ganska uppenbart att det finns en nära koppling mellan dessa fenomen. Om en person säger att en handling bör utföras – till exempel att man bör ge en del av sina pengar till välgörenhet – så verkar vi förutsätta att hon åtminstone i någon utsträckning är motiverad att utföra handlingen ifråga. Om det skulle visa sig att hon inte alls är motiverad att utföra handlingen, så behövs en förklaring, till exempel att hon egentligen inte menar det hon säger därför att hon är ironisk eller av någon anledning inte talar sanning. Debatten handlar framför allt om huruvida det finns ett nödvändigt samband mellan det moraliska språket och vår motivation att handla. Enligt internalism, som jag skall förstå den, finns en begreppsligt nödvändig koppling mellan moraliska omdömen och motivation.¹ Internalismen innebär att om man anser att en handling bör utföras, så måste man vara motiverad att utföra den – om man inte är motiverad så anser man egentligen inte att handlingen bör utföras. Enligt externalism, däremot, finns det inte någon intressant nödvändig koppling mellan moraliska omdömen och motivation. Huruvida internalism eller externalism är korrekt anses betydelsefullt eftersom uppfattningarna har implikationer för frågor som har att göra med om moraliska omdömen kan vara sanna eller falska och om moralen är objektiv eller subjektiv.

Enligt en utbredd uppfattning i metaetiken har internalism en viktig fördel i jämförelse med externalism, eftersom den verkar kunna förklara våra språkliga intuitioner beträffande kopplingen mellan moraliskt

¹Det finns också exempel på internalister som anser att kopplingen mellan moraliska omdömen och motivation inte är begreppsligt nödvändig. För en översikt över olika typer av internalism, se Björklund m.fl. 2012, s. 123–37 och Strandberg 2011, s. 341–69. Vad som i den engelskspråkiga debatten kallas ”moral judgment” benämner jag här ”moraliskt omdöme”, men möjligen hade ”moralisk uppfattning” hade varit lämpligare.

språk och motivation. Den kan förklara varför det skulle vara konstigt att säga något som ”Man bör avstå från att äta kött” utan att vara motiverad att avstå från sådana handlingar ens i någon utsträckning. Men internalismen finns i olika varianter som varierar i styrka. Enligt stark internalism gäller den nödvändiga kopplingen mellan moraliska omdömen och motivation för alla personer. Enligt svag internalism gäller den nödvändiga kopplingen endast personer som uppfyller vissa villkor.

Vad jag skulle vilja argumentera för i denna artikel är att internalismen hamnar i ett dilemma – och att externalismen kan lösa detta dilemma.² Stark internalism kan förklara våra språkliga intuitioner beträffande kopplingen mellan moraliskt språk och motivation, men den kan inte förklara varför en person under vissa förhållanden kan hysa ett moraliskt omdöme utan att vara motiverad i enlighet med det. Och omvänt: svag internalism kan förklara varför en person kan hysa ett moraliskt omdöme utan att vara vederbörligen motiverad, men den kan inte förklara våra språkliga intuitioner vad gäller kopplingen mellan moraliskt språk och motivation. Jag kommer att avsluta med att argumentera för att externalism undslipper detta dilemma eftersom den kan förklara båda.

2. STARK INTERNALISM OCH SPRÅKLIGA INTUITIONER

Den enklaste versionen av internalism kan formuleras på följande sätt:

Stark internalism: Det är begreppsligt nödvändigt att om en person anser att hon moraliskt bör utföra x , så är hon motiverad att utföra x (åtminstone i någon utsträckning).³

Det vanligast argumentet för internalism är, som just nämnts, att det kan förklara våra språkliga intuitioner beträffande kopplingen mellan moraliskt språk och motivation. Stark internalism innebär att det är en del av den språkliga meningen hos en sats som ”Man bör avstå från att äta kött” att en person som accepterar, eller asserterar, den är motiverad att handla i enlighet med sin moraliska uppfattning. Detta gäller vidare *alla* personer som hyser moraliska omdömen. Eftersom det är begreppsligt nödvändigt, och därmed handlar om språklig mening, är det något som vi har kunskap om i kraft av att vara kunniga språkanvändare. Det betyder att om en person sanningsenligt yttrar satsen så vet vi, i och med den kunskap vi har som kunniga språkanvändare, att hon är motiverad på ett visst sätt. Om då en person säger att man bör avstå från att äta kött,

²För ett utförligare försvar av de tankegångar som framförs i denna artikel, se Strandberg 2013, s. 25–51.

³Jag kommer fortsättningsvis att ta förbehållet ”åtminstone i någon utsträckning” för given.

men inte är motiverad att avstå från sådana handlingar, blir vi rimligtvis förvånade om det inte finns någon naturlig förklaring till hennes beteende – hon säger ju något som inte stämmer med den språkliga meningen hos satsen hon yttrar.

3. RATIONALISTISK INTERNALISM

Vi har sett att stark internalism har en enkel förklaring på våra språkliga intuitioner. De flesta internalister har emellertid kommit till slutsatsen att denna version av internalism är för stark. Anledningen är att det verkar vara möjligt för en person att anse att en handling bör utföras, utan att vara motiverad att utföra den, nämligen om hon har vad man skulle kunna kalla ”mentala problem”. Det verkar vara möjligt att en person hyser ett moraliskt omdöme utan att vara vederbörligen motiverad om hon lider av tillstånd såsom apati, djup depression, känslomässig störning, neuros, tvångstankar, osv. Anta att en person anser att hon bör betala sin hyra i tid, men att hon är djupt deprimerad av en eller annan anledning. Det verkar möjligt att hon faktiskt anser att hon bör betala sin hyra i tid, men att hon inte alls är motiverad att göra det just därför att hon är så deprimerad. Hennes depression har slagit ut hennes motivation att följa sin moraliska uppfattning i detta fall.

Detta problem har gjort att många internalister har antagit svag internalism:

Svag internalism: Det är begreppsligt nödvändigt att om en person anser att hon moraliskt bör utföra x , så är hon motiverad att utföra x , givet att hon uppfyller villkoren V .

Formulerad på detta sätt är svag internalism en typ av internalism där villkoren ” V ” kan formuleras på olika sätt. Grundtanken är att detta villkor skall utesluta ”mentala problem” av den typ jag just nämnde – apati, depression, känslomässiga störningar, etc. – som kan få konsekvensen att en person inte är motiverad att handla i enlighet med sitt moraliska omdöme. Det har emellertid visat sig vara svårt att formulera ” V ” på ett sätt som gör att internalismen inte blir cirkulär och därmed ointressant. Den mest lovande versionen av denna uppfattning verkar vara vad jag kallar ”rationalistisk internalism”:⁴

Rationalistisk internalism: Det är begreppsligt nödvändigt att om en person anser att hon moraliskt bör utföra x , så är hon motiverad att utföra x , givet att hon är *rationell*.

⁴För anhängare av denna uppfattning, se t.ex. Smith 1994, särskilt kap. 3 och 5, och Korsgaard 1996, s. 311–34.

Tanken i rationalistisk internalism är att om en person lider av ett mentalt problem som depression eller apati, så är hon inte är fullt rationell. Och i så fall kan det vara så att hon anser att hon bör utföra en handling utan att hon är motiverad i enlighet med sitt moraliska omdöme. Hon är inte motiverad att göra det som hon anser att hon bör göra därför att hon inte är rationell och denna brist på rationalitet beror på att hon är till exempel djupt deprimerad eller apatisk.

Man kan nu fråga sig: Varför skulle det vara så att man är rationell endast om man är motiverad att göra det som man anser att man bör göra? Det verkar som det måste finnas något med moraliska omdömen som förklarar denna nödvändiga koppling mellan moraliska omdömen och motivation. Det rimligaste svaret verkar vara att moraliska omdömen består i omdömen om vad vi har *skäl* att göra. Det verkar nämligen rimligt att tänka sig att om man anser att man har skäl att utföra en viss handling, men inte alls är motiverad att utföra den, så är man inte fullständigt rationell. Om man inte är motiverad att göra det som man själv anser att man har skäl att göra, så är man inte fullständigt rationell enligt detta synsätt. Anta att en mamma anser att det hon har starkast skäl att göra är att hjälpa sin son att sluta knarka, men att hon inte är motiverad att göra det därför att hon är djupt deprimerad. Det verkar rimligt att säga att hon inte är helt rationell eftersom hon inte är motiverad att göra det som hon själv anser att hon har starkast skäl att göra.

För att förstå denna tankegång lite tydligare så är det lämpligt att identifiera två olika teser som tillsammans leder fram till rationalistisk internalism. De två teser som tillsammans implicerar rationalistisk internalism kallar jag ”rationalism” och ”normativ internalism”.

- (1) *Rationalism*: Det är begreppsligt nödvändigt att om en person anser att hon moraliskt bör utföra x , så anser hon att hon har ett normativt skäl att utföra x .

Rationalism fångar tanken att moraliska omdömen består i omdömen om vad man har normativa skäl att göra.

- (2) *Normativ internalism*: Det är begreppsligt nödvändigt att om en person anser att hon har ett normativt skäl att utföra x , så är hon motiverad att utföra x , givet att hon är rationell.

Normativ internalism fångar tanken att om man anser att man har normativa skäl att utföra en viss handling, men inte är motiverad att utföra den, så är man inte fullständigt rationell. Rationalism och normativ internalism implicerar rationalistisk internalism:

- (3) *Rationalistisk internalism*: Det är begreppsligt nödvändigt att om en person anser att hon moraliskt bör utföra x , så är hon motive-
rad att utföra x , givet att hon är rationell.

I det följande kommer jag att rikta fokus på (2), normativ internalism, och kritisera denna uppfattning. Vi kommer sedan att se att denna kritik leder fram till att internalister bör anta en svagare version av (3), rationalistisk internalism, och att detta i sin tur leder fram till det internalistiska dilemmat som jag nämnde i inledningen.

4. PROBLEM MED NORMATIV INTERNALISM

Låt oss börja med en typ av fall som jag tror utgör problem för normativ internalism.⁵ Anta att en person anser att det hon har absolut starkast skäl att göra i en viss situation är att utföra en viss handling x . Anta vidare att hon anser att hon har mycket svagare skäl att utföra en annan handling y . Anta också att hon vet att hon inte både kan utföra x och y samtidigt – x och y är praktiskt oförenliga. I ett sådant fall så verkar det rimligt att anta att personen måste vara motiverad att utföra den handling som hon anser att hon har starkast skäl att göra – x – för att vara helt rationell.⁶ Om man anser att det man har absolut starkast skäl att göra i en viss situation är att utföra en viss handling, men inte är motiverad att utföra den, så verkar man brista i rationalitet. Följande tes verkar alltså rimlig:

Det är begreppsligt nödvändigt att om en person anser att hon har starkast normativa skäl att utföra x , så är hon motiverad att utföra x , givet att hon är rationell.

Men normativ internalism är mycket starkare än denna tes. Normativ internalism innebär att det är begreppsligt nödvändigt att en rationell person är motiverad att utföra *varje* handling som hon anser att hon har skäl att utföra – *oavsett* hur *svagt* hon anser att detta skäl är. I vårt exempel så måste personen inte bara vara motiverad att utföra den handling som hon anser att hon har *starkast skäl* att utföra – x – för att vara rationell. Hon måste också vara motiverad att utföra den handling som hon anser att hon har ett *mycket svagt skäl* att utföra – y – för att vara rationell. Om hon inte är motiverad att utföra även y , så är hon inte fullständig

⁵Det är naturligtvis möjligt att rikta kritik mot de argument jag framför i denna sektion; jag försöker besvara en del av denna kritik i Strandberg 2013, s. 25–51.

⁶Det verkar också rimligt att hävda att hon måste vara *mest* motiverad att utföra x för att vara helt rationell. Detta är emellertid inte självklart; se Arpaly 2000, s. 488–513.

rationell. Närmare bestämt är hon irrationell i den utsträckning hon inte är motiverad att utföra y .

Ett problem med normativ internalism är att det inte verkar uteslutet att en person som anser att hon har absolut starkast skäl att utföra handling x , och ett väldigt svagt skäl att utföra en annan handling y , kan vara fullständigt rationell, även om hon inte alls är motiverad att utföra den handling y som hon anser att hon har ett väldigt svagt skäl att utföra.⁷ Låt oss föreställa oss en person som är allvarligt sjuk och som vet att det enda sättet att överleva är att genomgå en viss operation. Hon anser följaktligen att det hon har absolut starkast skäl att göra är att genomgå behandlingen ifråga. Men anta också att hennes läkare informerar henne om att *om* hon genomgår operationen så kan hon inte dricka kaffe under en timme. Hon anser följaktligen att hon har ett mycket svagt skäl att inte genomgå operationen. Enligt normativ internalism så är det *begreppsligt nödvändigt* att hon är motiverad att avstå från operationen för att vara fullständigt rationell och att hon är *irrationell* i den utsträckning som hon inte är motiverad på detta sätt. Det verkar inte uppenbart. Det verkar som det är *begreppsligt möjligt* för henne att vara fullständigt omotiverad att inte genomgå operationen och ändå vara helt rationell.

Låt oss ta ett annat exempel på samma problem. Anta att en person ser tio personer som precis har blivit överkörda av en spårvagn och att hon inser att hon måste ringa alarmcentralen för att de skall ha någon chans att överleva. Hon anser följaktligen att det hon har absolut starkast skäl att göra i denna situation är att ringa alarmcentralen. Men anta att hon också inser att *om* hon gör det så riskerar hon att komma lite sent till ett möte. Hon tror alltså att hon har ett mycket svagt skäl att inte ringa alarmcentralen. Normativ internalism innebär att det är begreppsligt nödvändigt att hon är motiverad att avstå från att ringa alarmcentralen för att vara fullständigt rationell och att hon är irrationell i den utsträckning hon inte är motiverad på detta sätt. Det verkar inte uppenbart. Det verkar som det är fullt möjligt för henne att vara fullständigt rationell även om hon inte alls är motiverad att avstå från att ringa alarmcentralen.

Det finns också åtminstone ytterligare en typ av fall som ger upphov till problem för normativ internalism. Anta att en person anser att hon har ett skäl att utföra en handling x och att hon har ett skäl att utföra en annan handling y . Anta vidare att hon anser att hennes skäl att utföra x och att utföra y är exakt lika starka. Anta slutligen att hon vet att hon inte kan utföra både x och y – x och y är praktiskt oförenliga. Normativ internalism innebär att det är begreppsligt nödvändigt att personen är

⁷Flera filosofer har noterat detta och liknande problem. Se Copp 1997, s. 45–46; Wallace 2006, s. 187–88, och Gert 2008, s. 16–17. Se även Schroeder 2007, s. 166–67.

motiverad att utföra både x och y för att vara fullständigt rationell och att hon är irrationell i den utsträckning hon inte är motiverad att utföra både x och y . Men det är inte uppenbart. Det verkar som hon *kan* vara fullständigt rationell och samtidigt sakna motivation att utföra x eller y under vissa förhållanden.

Låt oss återigen illustrera med ett exempel. Anta att du funderar på vad du ska äta till lunch och att du väljer mellan pizza och pasta. Låt oss föreställa oss att du anser att du har exakt lika starka skäl att äta pizza som pasta, men att du vet att du inte kan välja båda eftersom du inte har råd. Normativ internalism säger då att det är begreppsligt nödvändigt att du är motiverad att äta både pizza och pasta för att vara fullständigt rationell – och att du är irrationell i den utsträckning du inte är motiverad att äta båda. Men det verkar inte uppenbart. Jag är benägen att säga att i ett sådant fall så skulle du kunna vara fullständigt rationell om du bara är motiverad att äta pizza (men inte pasta) eller bara är motiverad att äta pasta (men inte pizza). I ett dylikt fall så verkar vi acceptera att du kan vara helt styrd av vad du för tillfället känner för – av dina momentana fysiska böjelser – och ändå vara rationell.

Hur skall en försvarare av rationell internalism ställa sig till de problem för normativ internalism som jag påpekat ovan? Som jag ser det är det bästa att försvara en svagare variant av normativ internalism som undkommer dessa problem. Den version av normativ internalism som man då kommer fram till är följande:

- (2*) *Ny normativ internalism*: Det är begreppsligt nödvändigt att om en person anser att hon har ett normativt skäl att utföra x , så är hon motiverad att utföra x , givet att hon (i) är rationell och (ii) inte tror att det finns starkare, eller minst lika starka, skäl att utföra en annan handling y , som är praktiskt oförenlig med x .

Ny normativ internalism undkommer de problem för den ursprungliga versionen av normativ internalism som jag nämnde tidigare. Den innebär inte att en person som anser att hon har absolut starkast skäl att utföra en handling x , och ett väldigt svagt skäl att utföra en annan handling y , behöver vara motiverad att utföra y för att vara fullständigt rationell. Och den innebär inte att en person som anser att hon har lika starka skäl att utföra en handling x som att utföra en annan handling y måste vara motiverad att utföra både x och y för att vara fullständigt rationell.

Låt oss nu se vilka konsekvenser den nya, svagare, varianten av normativ internalism har för rationalistisk internalism. Vi såg tidigare att rationalism i (1) tillsammans med den ursprungliga versionen av normativ internalism i (2) implicerade rationalistisk internalism i (3). Den situation vi nu har är att rationalism i (1) tillsammans med den nya, svagare,

varianten av normativ internalism i (2*) implicerar en svagare variant av rationalistisk internalism:

- (3*) *Ny rationalistisk internalism*: Det är begreppsligt nödvändigt att om en person anser att hon moraliskt bör utföra x , så är hon motiverad att utföra x , givet att hon (i) är rationell och (ii) inte tror att det finns starkare, eller minst lika starka, skäl att utföra en annan handling y , som är praktiskt oförenlig med x .

Det är alltså möjligt för rationalistiska internalister att undvika argumenten mot den ursprungliga versionen av normativ internalism i (2) genom att anta en svagare variant av denna uppfattning, ny normativ internalism i (2*). Men detta har ett visst pris: De måste då acceptera en svagare variant av rationalistisk internalism, ny rationalistisk internalism i (3*).*

5. ETT INTERNALISTISKT DILEMMA

Tidigare nämnde jag att stark internalism verkar kunna förklara våra språkliga intuitioner beträffande kopplingen mellan moraliskt språk och motivation, men att den inte kan förklara varför en person inte behöver vara motiverad i enlighet med sitt moraliska omdöme. Vi såg att en svagare variant av internalism, rationalistisk internalism, kan förklara detta fenomen. Men vi såg sedan att man kan rikta kritik mot en av de teser som leder fram till rationalistisk internalism, nämligen normativ internalism. För att undkomma kritiken mot normativ internalism så behöver man anta en svag variant av rationalistisk internalism. Men som vi nu kommer att se så verkar inte den nya versionen av rationalistisk internalism i (3*) kunna förklara våra språkliga intuitioner beträffande kopplingen mellan moraliskt språk och motivation.

Det finns två skäl, som jag ser det, att betvivla att ny rationalistisk internalism kan förklara våra språkliga intuitioner beträffande sambandet mellan moraliskt språk och motivation.

*Det finns olika sätt för anhängare av rationalistisk internalism att försöka undvika dilemman i nästa sektion. Ett sätt skulle vara att hävda att en starkare version av (1), rationalism, än den jag bygger mitt resonemang på ovan. Enligt detta alternativ skulle rationalism formuleras på följande sätt: Det är begreppsligt nödvändigt att om en person anser att hon moraliskt bör utföra x , så anser hon att hon har *starkast* normativa skäl att utföra x (inte endast *något* skäl att utföra x). Denna version av rationalism skulle i sin tur kombineras med en rimligare version av normativ internalism än (2*). Kombinationen av dessa teser skulle i sin tur leda till en rimligare version av rationalistisk internalism än (3*). Jag argumenterar för att den nämnda versionen av rationalism är för stark i Strandberg 2013, s. 25–51.

För det första är det tvivelaktigt om rationalistisk internalism är tillräcklig för att erbjuda en sådan förklaring. Enligt ny rationalistisk internalism är det begreppsligt nödvändigt att om person anser att en handling bör utföras, så är hon motiverad att utföra den under förutsättning att två villkor är uppfyllda: för det första att personen är rationell och för det andra att hon inte anser att hon har starkare, eller minst lika starka skäl, att utföra en annan handling som är praktiskt oförenlig med den första handlingen. Det betyder att rationalistisk internalism kan förklara varför det är konstigt att en person säger att en handling bör utföras utan att vara motiverad att utföra den *endast* i de fall vi kan *förutsätta* att båda dessa två villkor är uppfyllda. Men det verkar rimligt att anta att vi skulle anse att en persons brist på motivation i ett sådant fall är konstigt *även* om vi inte har någon uppfattning om huruvida det andra villkoret är uppfyllt. Anta att någon säger ”Man bör ge en del av sina pengar till välgörenhet”, men att det visar sig att hon inte alls är motiverad att göra det. Vi skulle förmodligen anse att kombinationen av hennes yttrande och brist på motivation är konstigt även om vi inte har någon uppfattning över huvud taget om huruvida hon tror att hon starkare eller lika starka skäl att utföra en annan handling. Om en person yttrar att en handling inte bör utföras, men inte är motiverad att utföra handlingen, så skulle vi, tror jag, tycka att det är konstigt även om vi inte har någon uppfattning över huvud taget om huruvida hon anser att hon har starkare eller lika starka skäl att utföra en annan handling. Den typ av komplexa överväganden som skulle krävas för att vi skulle avstå från att dra denna slutsats verkar inte vara för handen när vi tar ställning till personens yttrande. Det skulle betyda att ny rationalistisk internalism inte är tillräcklig för att förklara våra språkliga intuitioner beträffande kopplingen mellan moraliskt språk och motivation.

För det andra är det tvivelaktigt om rationalistisk internalism är nödvändig för att erbjuda en sådan förklaring. Enligt denna version av internalism är det, som vi har sett, begreppsligt nödvändigt att om person anser att en handling bör utföras, så är hon motiverad att utföra den under förutsättning att två villkor är uppfyllda: för det första att personen är rationell och för det andra att hon inte anser att hon har starkare eller minst lika starka skäl att utföra en annan handling som är praktisk oförenlig med den första handlingen. Det betyder att denna uppfattning inte kan förklara varför vi skulle tycka att det är konstigt att en person säger att en handling bör utföras utan att vara motiverad att utföra den även i de fall där vi *vet* att det andra villkoret inte är uppfyllt. Men det verkar rimligt att tänka sig att det är precis så. Anta att en person säger ”Man bör ge en del av sina pengar till välgörenhet”, men inte är motiverad att göra det. Anta vidare att vi vet att hon anser att hon starkare skäl att utföra en an-

nan handling än den som hon anser att hon moraliskt bör utföra. Det verkar som vi ändå skulle tycka att hennes yttrande är mycket märkligt. Vi skulle förmodligen undra varför hon säger att en handling bör utföras om hon inte har någon som helst motivation att utföra handlingen. Vi skulle förmodligen undra vad *poängen* med hennes yttrande är om hon nu inte har någon som helst benägenhet att faktiskt utföra den. Om en person säger att en handling bör utföras, men inte är motiverad att utföra den, så skulle vi förmodligen tycka att det är konstigt. Och vi skulle sannolikt tycka att det är konstigt även om vi vet att hon anser att hon har starkare skäl att utföra en annan handling. Det skulle betyda att ny rationalistisk internalism inte är nödvändig för att förklara våra språkliga intuitioner beträffande kopplingen mellan moraliskt språk och motivation.

Om det jag har sagt ovan stämmer så står internalismen står inför ett dilemma. Stark internalism kan förklara våra språkliga intuitioner beträffande kopplingen mellan moraliskt språk och motivation, men den kan inte förklara varför en person kan hysa ett moraliskt omdöme utan att vara motiverad i enlighet med det. Den rimligaste formen av svag internalism, ny rationalistisk internalism, kan förklara varför människor kan hysa moraliska omdömen utan att vara motiverade, men den kan inte förklara våra språkliga intuitioner beträffande kopplingen mellan moraliskt språk och motivation.

6. EN EXTERNALISTISK LÖSNING

Avslutningsvis skulle jag kortfattat vilja förklara hur externalismen kan undvika det dilemma som internalismen står inför.

Externalismen, som jag förstår den, är helt enkelt förnekandet av internalismen:

Externalism: Det är inte begreppsligt nödvändigt att om en person anser att hon moraliskt bör utföra x , så är hon motiverad att utföra x (inte ens om villkor V är uppfyllda).

För att externalismen skall kunna undvika det internalistiska dilemmat så måste den kunna göra två saker. Dels måste den kunna förklara våra språkliga intuitioner beträffande kopplingen mellan moraliskt språk och motivation. Och dels måste den kunna förklara varför människor kan anse att handlingar bör utföras utan att vara motiverade att utföra dem.

Låt oss börja med att se hur externalister kan förklara våra språkliga intuitioner, och framför allt varför det är så konstigt om en person faller ett moraliskt yttrande utan att vara motiverad i enlighet med det.

Externalister kan, enligt min uppfattning, förklara detta genom att använda sig av en pragmatisk teori om språk. Sådana teorier handlar om

hur yttranden kan uttrycka viss information utan att det är en del av den strikt språkliga meningen hos de satser som yttras. Närmare bestämt tror jag att externalister kan använda sig av språkfilosofen Paul Grices pragmatiska teori om konversationella implikaturer. Grice menar att vi följer vad han kallar "The Principle of Cooperation" när vi talar med varandra: "Make your conversational contribution such as is required [...] by the accepted purpose [...] of the talk exchange in which you are engaged."⁹ När man talar med andra människor skall man med andra ord endast säga sådant som stämmer med det gemensamt accepterade syftet med den konversation man deltar i. Grice menar att vi uppfyller denna princip genom att följa ett antal olika maximier. En maxim, "Maxim of Relevance", säger: "Be relevant." Med andra ord skall man, när man är involverad i en konversation, endast följa yttranden som är relevanta givet syftet med konversationen. Grice hävdar att eftersom vi följer "The Principle of Cooperation", och de olika maximerna, så kommer våra yttranden att uttrycka information som inte är del av det semantiska innehållet hos de satser vi yttrar, även om vi ibland kan frestas att tro att så är fallet. Han kallar den extra information som yttranden uttrycker på detta sätt för "*konversationella implikaturer*".

Jag tänker mig att externalister kan argumentera för följande tes:

Ett pragmatiskt förslag: Det är generellt fallet att om en person yttrar en sats av typen "Handlingen x bör utföras", så implicerar hennes yttrande konversationellt att hon är motiverad att utföra x .

Om en person yttrar satsen "Man bör ge en del av sina pengar till välgörenhet", så implicerar hennes yttranden konversationellt, enligt detta förslag, att hon är motiverad att utföra sådana handlingar, och denna koppling mellan moraliska yttranden och motivation gäller generellt. Detta förslag är förenligt med externalism eftersom det endast säger att det är en persons moraliska yttrande som uttrycker att hon är motiverad att utföra handlingen ifråga. Det säger inte att det är begreppsligt nödvändigt att om en person anser att en handling bör utföras, så är hon motiverad att utföra den. Jag har argumentet för detta pragmatiska förslag utförligt i andra sammanhang och kommer endast att ange dess konturer här.¹⁰

⁹Grice 1981, s. 26.

¹⁰Se Strandberg 2012, s. 87–122, och Strandberg 2011, s. 341–69. Den typ av implikatur det pragmatiska förslaget refererar till kallar Grice "generalized conversational implicatures", som utmärks av att det finns en generell, standardiserad, koppling mellan yttrande av en viss sats och en viss implikatur (Grice 1981, s. 37–40). För två konkurrerande pragmatiska förklaringar, se Finlay 2004, s. 205–23 och Copp 2009, s. 167–202.

Det pragmatiska förslaget vilar på tre enkla och tämligen okontroversiella antaganden. För det första: En sats av typen ”Handlingen x bör utföras” implicerar strikt att det finns ett moraliskt skäl att utföra x . Om en handling moraliskt bör utföras så följer det logiskt att det finns ett moraliskt skäl – ett skäl enligt moralen – att utföra handlingen ifråga. För det andra: Ett generellt syfte med moraliska konversationer är att påverka människors handlingar. Generellt sett är en viktig funktion med moraliska samtal att vi försöker få människor att utföra vissa handlingar och få dem att undvika att utföra vissa handlingar. För det tredje: Ett grundläggande sätt att påverka människors handlingar är att låta dem få veta vilka attityder vi har till dessa handlingar, till exempel om vi vill att dessa handlingar utförs eller om vi vill att dessa handlingar inte utförs.

Givet dessa tre antaganden så är det rimligt att tänka sig att en persons yttrande av en sats av typen ”Handlingen x bör utföras” konversationellt implicerar att hon är motiverad att utföra x . Låt oss ta ett enkelt exempel. Anta att en person deltar i ett samtal om moraliska aspekter av köttätande med andra människor. Vid ett tillfälle i samtalet säger hon ”Man bör avstå från att äta kött”. Det verkar då rimligt att tänka sig att hon vill att människor skall avstå från att äta kött. Varför verkar det rimligt? Jag tror att svaret är att om hon *inte* vill att människor skall avstå från att äta kött, så skulle hennes yttrande inte vara relevant givet det generella syftet med moraliska konversationer – nämligen att påverka människors handlingar. Huvudtanken är alltså följande: Satsen ”Man bör avstå från att äta kött” implicerar strikt att det finns ett moraliskt skäl att avstå från köttätande. Om en person yttrar denna sats i en moralisk konversation, som har som generellt syfte att påverka andra människors handlingar, så verkar det rimligt att förutsätta att hon vill att människor skall avstå från sådana handlingar. Anledningen är helt enkelt att om hon inte vill att människor skall avstå från att äta kött så, skulle hennes yttranden inte vara relevant givet det generella syftet med moraliska konversationer – att påverka människors beteenden. Mer specifikt så skulle en person som säger att man bör avstå från att äta kött, men inte vill att människor avstår från köttätande, bryta mot en av Grices maximer, ”The Maxim of Relevance”, som säger att man endast skall säga sådant som är relevant givet syftet med den konversation man deltar i.

Hittills har jag försökt argumentera för att en persons moraliska yttrande konversationellt implicerar att hon vill att vissa handlingar skall utföras. Vi kan nu sätta denna tankegång i samband med en persons motivation att utföra handlingarna ifråga. När en person säger att en handling bör utföras menar hon att hon själv bör utföra handlingen under förutsättning att det inte finns några indikationer på motsatsen – givet att hon inte explicit säger att just hon inte behöver utföra handlingen el-

ler att det framgår av kontexten att hon anser det. Hon gör med andra ord i allmänhet inget undantag för sig själv om hon inte låter detta framgå på något sätt. Det betyder att hennes moraliska yttrande konversationellt implicerar att hon vill att hon själv utför handlingen, med andra ord att hon är motiverad att utföra handlingen. Det verkar därför rimligt att hävda att en persons yttrande såsom ”Man bör avstå från att äta kött” konversationellt implicerar att hon är motiverad att avstå från sådana handlingar.

Det är också viktigt att notera att den pragmatiska kopplingen mellan moraliska yttranden och attityd enligt detta förslag gäller generellt, vilket innebär att det finns en standardiserad korrelation mellan yttranden av vissa typer av satser och attityder. Som jag nämnde tidigare så bygger det pragmatiska förslaget på uppfattningen att moraliska konversationer generellt har syftet att påverka människors handlingar. Det är alltså ett syfte som moraliska konversationer standardmässigt har – det är ett syfte som systematiskt återkommer i moraliska konversationer. Vi kan tänka på kopplingen som ett slags bakgrundsantagande som finns i moraliska konversationer, som styr hur vi uppfattar vad som sägs, men som vi i allmänhet inte är medvetna om. Det innebär i sin tur att det generellt är fallet att våra yttranden av satser av typen ”Handlingen x bör utföras” konversationellt implicerar att vi är motiverade att utföra x , även om vi inte är medvetna om att moraliska konversationer har det nämnda syftet. Det innebär också att en persons yttrande av en viss moralisk sats gör att vi automatiskt kan sluta oss till att hon är motiverad på ett visst sätt.

Vi kan nu se att externalismen kan förklara våra språkliga intuitioner beträffande kopplingen mellan moraliskt språk och motivation med hjälp av det pragmatiska förslaget. Enligt detta förslag så är det generellt fallet att en persons yttrande av en sats av typen ”Man bör utföra handling x ” konversationellt implicerar att hon är motiverad att utföra handlingen x . Det är därför rimligt att tänka sig att när en person yttrar en sådan sats, så har vi en stark förväntan att hon skall vara motiverad att utföra handlingen ifråga. Om hon faller ett sådant yttrande, men inte är motiverad att utföra handlingen, så tycker vi följaktligen att detta är märkligt. Vi kan också förklara denna reaktion genom att konstatera att personen har brutit mot ”The Cooperative Principle”, och särskilt mot ”The Maxim of Relevance”, genom att fälla ett yttrande som överensstämmer med ett generellt syfte med denna typ av konversation. Följaktligen undviker externalismen, givet att den antar det pragmatiska förslaget, den del av dilemmat som vidlåder rationalistisk internalism.

Grice poängterar att i de fall yttranden av en viss typ av sats generellt konversationellt implicerar en viss proposition är det lätt att felaktigt dra

slutsatsen att den är en del av meningen hos satsen.¹¹ Enligt det pragmatiska förslaget finns en generell, standardiserad, koppling mellan yttranden av moraliska satser av typen ”Handling x bör utföras” och informationen att personen ifråga är motiverad att utföra x . Det innebär att det är lätt att felaktigt tro att det finns en begreppsligt nödvändig koppling mellan dem. Det faktum att vi automatiskt kan sluta oss till att en person är motiverad på grundval av hennes moraliska yttrande bidrar också till att ett sådant felslut är möjligt. Det pragmatiska förslaget kan med andra ord ge en förklaring till varför internalismen kan förefalla så rimlig.

Vi kan nu också se att externalismen har förmågan att förklara hur det är möjligt att en person kan anse att en handling bör utföras utan att vara motiverad att utföra den om hon lider av till exempel apati, djup depression, etc. Enligt externalismen så gäller generellt att det inte finns någon nödvändig koppling mellan moraliska omdömen och motivation. Det är alltså möjligt, enligt externalismen, att en person som anser att en handling bör utföras inte alls är motiverad att utföra den. Men en förklaring till att en person inte är moraliskt motiverad kan i sin tur vara att hon lider av något ”mentalt problem” såsom apati eller depression. Följaktligen undviker externalismen den del av det internalistiska dilemmat som vidlåder stark internalism.

7. SAMMANFATTNING

I denna artikel har jag argumenterat för att internalism om moraliska omdömen och motivation står inför ett dilemma. Den kan antingen förklara våra språkliga intuitioner eller förklara hur det kommer sig att en person inte behöver vara motiverad i enlighet med sitt moraliska omdöme, men den kan inte förklara båda. Jag har också argumenterat för att externalismen undviker dilemmat – externalismen kan förklara båda.

LITTERATUR

- Arpaly, Nomy. 2000. ”On Acting Rationally against One’s Best Judgment”. *Ethics* 110, s. 488–513.
- Björklund, F., G. Björnsson, J. Eriksson, R. Francén Olinder, C. Strandberg. 2012. ”Recent Work on Motivational Internalism”. *Analysis* 72, s. 123–37.

¹¹Grice 1981, s. 37. Enligt Grice utmärker detta generella konversationella implikaturer. En möjlig invändning är att mitt pragmatiska förslag inte kan förklara kopplingen mellan moraliska omdömen, moraliskt tänkande, och motivation, utan endast mellan moraliska yttranden och motivation. Jag menar emellertid att det kan förklara även den senare kopplingen; se Strandberg 2011, s. 347–55, 361–66 och Strandberg 2013, s. 47–49.

- Copp, David. 1997. "Belief, Reason, and Motivation: Michael Smith's *The Moral Problem*". *Ethics* 108, s. 33–54.
- Copp, David. 2009. "Realist-expressivism and Conventional Implicature". I *Oxford Studies in Metaethics*, vol. 4, utg. R. Shafer-Landau, s. 167–202. Oxford: Oxford University Press.
- Finlay, Stephen. 2004. "The Conversational Practicality of Value Judgment". *The Journal of Ethics* 8, s. 205–23.
- Gert, Joshua. 2008. "Michael Smith and the Rationality of Immoral Action". *The Journal of Ethics* 12, s. 1–23.
- Grice, Paul. 1981. "Logic and Conversation". *Studies in the Ways of Words*, s. 22–40. Cambridge, Mass.: Harvard University Press.
- Korsgaard, Christine. 1996. "Skepticism About Practical Reason". I *Creating the Kingdom of Ends*, s. 311–34. Cambridge: Cambridge University Press.
- Schroeder, Mark. 2007. *Slaves of the Passions*. Oxford: Oxford University Press.
- Smith, Michael. 1994. *The Moral Problem*. Oxford: Oxford University Press.
- Strandberg, Caj. 2011. "The Pragmatics of Moral Motivation". *The Journal of Ethics* 15, s. 341–69.
- Strandberg, Caj. 2012. "A Dual Aspect Account of Moral Language". *Philosophy and Phenomenological Research* 84, s. 87–122.
- Strandberg, Caj. 2013. "An Internalist Dilemma – and an Externalist Solution". *The Journal of Moral Philosophy* 10, s. 25–51.
- Wallace, R. J. 2006. "Moral Motivation". I *Contemporary Debates in Moral Theory*, utg. J. Dreier, s. 182–96. Oxford: Blackwell.

Döden som förlust

Andra människor är nog inte medvetna om att alla som ägnar sig åt filosofin på det rätta sättet självmant förbereder sig för en enda sak: att dö och vara död.

Platon, *Faidon* 64a

1. FÖRLUSTTEORIN

Det jag här kallar ”förlustteorin” tycks numera vara den vanligaste uppfattningen bland filosofer om varför döden är något ont för den som dör.¹ Thomas Nagel formulerar den så här: ”Det är klart att om döden över huvud taget är ett ont, så kan det inte vara på grund av dess egentliga egenskaper utan endast på grund av vad den berövar oss” – och ”om döden är ett ont är det *förlusten av livet* snarare än tillståndet att vara död, icke-existerande eller medvetlös, som är oacceptabelt”.²

Genom döden går man alltså miste om något, man berövas något, man *förlorar* något av värde, som man skulle ha haft om man inte hade dött. *Vad* – och *hur mycket* – man förlorar är kanske inte helt klart, men enligt en vanlig uppfattning är det något i stil med mellanskillnaden i värde mellan det liv man skulle ha levat om döden inte hade inträffat och det liv som faktiskt tog slut med döden.

Fred Feldman ger en precisering i denna riktning som bygger på att man jämför olika *möjliga världar*, dvs. olika sätt på vilka världen hade kunnat vara beskaffad. Vi kan säga att värdet av en möjlig värld w för en person s , förkortat $V(s, w)$, är allt det som är bra för s i w minus allt det som är dåligt för s i w .³ Feldman antar vidare att det som kan vara bra,

¹Se t.ex. Feldman 1991, s. 206 och de författare han där hänvisar till, t.ex. Harry Silverstein som skriver: ”According to the standard argument, death is simply an evil of deprivation, an evil consisting in the loss, or lack, of a positive good, namely, life” (1980, s. 404). Förlustteorin i ungefär denna utformning finns även hos många senare författare. Se t.ex. Ben Bradley 2009, s. 113. Och John Broome (2012, s. 232) säger exempelvis: ”When you die, what you lose is [...] the rest of your life. The badness of this loss is [...] the goodness of the rest of your life”.

²Nagel 1979, s.171 och 172 (kursiv i originalet).

³Se Feldman 1991, s. 209. Vad det är som är bra och dåligt kan man ha olika uppfattningar om. I sin artikel antar Feldman att hedonismen är en korrekt teori, men det tror han i själva verket inte; se s. 210.

eller dåligt, för en person är *sakförhållanden*, dvs. att något förhåller sig på ett visst sätt eller att något är sant.⁴ *Döden* kan alltså inte vara dålig för en person, men *att* personen dör kan vara dåligt för personen (och naturligtvis även för andra, men det bortser vi här från).

Ett sakförhållande p är, enligt Feldman, dåligt för en person s om och endast om det finns två möjliga världar, w och w' , sådana att w är den närmaste värld i vilken p ingår och w' är den närmaste värld i vilken icke- p ingår och w är sämre än w' för s , dvs. $V(s, w) < V(s, w')$.⁵ De möjliga världar som Feldman talar om är en sorts totala beskrivningar av hur världen skulle kunna vara.⁶

Döden är med andra ord dålig för en person om den närmaste värld där den inte inträffar är bättre för personen än den värld där den inträffar. Detta låter kanske rätt rimligt. Och det är som sagt en vanlig uppfattning bland filosofer.

2. TILLÄMPNING AV FÖRLUSTTEORIN

Men det är ändå något konstigt med förlustteorin. Anta att s är en person som dör i vår vanliga värld. Vilken är då den närmaste alternativa värld i vilken personen s inte dör? Man kan gissa att det är en värld i vilken s inte ens är född och i vilken han inte alls existerar. Eller är det en värld i vilken s existerar och är odödlig?

Eller ska man ta fasta på den tid vid vilken personen faktiskt dör? Om s dör vid en viss tidpunkt, t , så är den närmaste värld där denna händelse inte inträffar, dvs. där s inte dör vid t , antagligen en värld där s dör antingen någon sekund tidigare eller senare än t . Skillnaden i värde mellan dessa två (eller tre) världar kan i allmänhet antas vara obetydlig. Döden

⁴Feldman 1991, s. 211–12. Det är sålunda inte t.ex. föremål eller händelser som är bra eller dåliga för en person, men det kan vara bra för s att han äger ett visst föremål vid en viss tidpunkt eller att en viss händelse inträffar.

⁵Feldman skriver (1991, s. 215): "p would be bad for s if and only if ($\exists w$) ($\exists w'$) (w is the nearest p-world & w' is the nearest ~p-world & $V(s, w) < V(s, w')$). Ordet "närmast" ("nearest") syftar inte här på någon rumslig relation, utan på graden av en sorts likhet. För enkelhetens skull kan vi här anta att det bara finns *en* möjlig värld som är "närmast" (eller mest lik) en viss värld.

⁶Han skriver: "a possible world is a huge proposition fully describing some total way the world might have been, including all facts about the past, present and future" (1991, s. 207). I och för sig finns det väl *oändligt* många sanna satser som skulle kunna ingå i en sådan "total" beskrivning, men kanske kan man hitta någon ändlig beskrivning som är tillräckligt innehållsrik för att kunna kallas en möjlig värld. Man kan notera att vår vanliga värld *inte* är en möjlig värld i Feldmans mening; vår vanliga värld är inte en beskrivning eller en "huge proposition". Men ibland ska jag ändå använda ordet "världar" om det som beskrivs.

skulle därmed varken vara bra eller dålig för *s*, alldeles oberoende när han eller hon råkar dö. Det låter inte så plausibelt.⁷

När Feldman ska tillämpa sin förlustteori tycks han emellertid ha glömt bort kravet att man ska jämföra med den *närmaste* alternativa världen. Han tänker sig bland annat att han dör under en flygtur till Europa och han jämför då den värld i vilken han gör det med en värld i vilken han i stället lever vidare under många lyckliga år.⁸ Men detta kan knappast vara den närmaste alternativa världen. Enligt sitt eget förslag borde han snarare ha jämfört med en värld där han tar ett lite senare (eller tidigare) flyg till Europa och dör under den resan i stället – och som för övrigt är maximalt lik den verkliga världen. Men då blir nog värdeskillnaden inte särskilt stor. Kanske blir den ingen alls.

I någon mån diskuterar Feldman problem av det slag jag här har antytt. Han tänker sig att en persons faktiska död kan beskrivas på många olika sätt och att värderingen av den – och den närmaste alternativa värld man då ska utgå från – beror på vilken beskrivning man tar fasta på.⁹ Han diskuterar ett exempel från Krimkriget 1854 där en ung brittisk officer, Herbert, blir skjuten nära staden Balaclava där han deltar i ett anfall med det lätta kavalleriet. Man kan då skilja mellan bland annat följande sanna beskrivningar av Herberts död (vi kan anta att han dog kl. 9.03 på dagen):

- (1) Herbert dör kl. 9.03 den 25 oktober 1854.
- (2) Herbert dör nära Balaclava.
- (3) Herbert dör under det lätta kavalleriets anfall.
- (4) Herbert dör på grund av att han blev skjuten av ryssen Ivan.
- (5) Herbert dör i unga år.

Feldman anser att man ska bortse från (1) och (4), eftersom Herbert i motsvarande närmaste världar skulle dö endast några sekunder senare – eftersom en annan ryss, Boris, skulle ha skjutit honom om Ivan hade missat. Så är exemplet konstruerat. Feldman menar att (5) är den beskrivning man ska ta fasta på, eftersom det är en tragedi att Herbert dör ung, och han menar att vi kan anta att Herbert i motsvarande närmaste alternativa värld kanske såras men klarar sig och sedan lever ett långt och lyckligt liv.¹⁰

⁷Men det stämmer bra med Epikuros tes att döden inte kan vara dålig för den som dör; se Bergström 2013.

⁸Feldman 1991, s. 216. Han säger trots detta att han ska jämföra med "the nearest world in which I do not die en route to Europe on this trip".

⁹Detta är något han tycks ha lärt sig i en artikel av Jeff McMahan (1988). Feldman måste antas syfta på denna artikel, även om han felaktigt kallar den "The Evil of Death", se Feldman 1991, noterna 2, 3, 18 och 19. Exemplet med officeren i Krimkriget kommer just från McMahan 1988, s. 45.

¹⁰Ibid. s. 226.

Men detta är knappast den närmaste möjliga värld i vilken Herbert inte dör i unga år. I den närmaste världen dör han visserligen inte ”i unga år”, men däremot får man väl anta att han dör så snart som möjligt därefter, kanske när han fyllt trettio år. Ju längre han skulle leva i en alternativ värld, desto mer avlägsen från vår skulle ju den världen vara.

Och varför ska vi anta att hans liv i fortsättningen skulle bli *lyckligt* i närmaste möjliga värld? Det verkar inte alls troligt. Ju lyckligare han skulle ha blivit efter 1854, desto mer avlägsen skulle rimligen den världen ha varit från den verkliga världen, i vilken han ju var död efter 1854 och därför inte alls kunde vara lycklig då.

3. RELEVANTA ALTERNATIV

Feldman och hans likasinnade tycks i praktiken alltså inte bry sig om förlustteorins krav att man ska jämföra med den alternativa värld som är *närmast*. Och man kan kanske också fråga sig varför man ska göra det. Vad är det som motiverar det kravet? Borde man inte snarare jämföra med den värld som *faktiskt* skulle ha varit fallet om den döden inte hade inträffat. Varför ska detta antas vara den ”närmaste” världen?

Detta hänger nog ihop med tolkningen av kontrafaktiska villkorssatser. De filosofer som sysslat med sådana satser – exempelvis David Lewis – har såvitt jag vet inte uttryckligen förklarat varför man just ska fokusera på den närmaste alternativa världen, men de menar möjligen att det är så vi faktiskt använder kontrafaktiska villkorssatser. När vi säger att om A inte hade inträffat, så skulle B ha inträffat, så *menar* vi kanske just att i en värld där A inte inträffar men som för övrigt är så lik vår värld som möjligt, så inträffar B.¹¹

Men även om vi faktiskt använder kontrafaktiska villkorssatser på ungefär detta sätt, så behöver det kanske inte vara avgörande för den som vill värdera sin död. Dessutom är det såvitt jag förstår inte uppenbart att vi verkligen använder kontrafaktiska satser på detta sätt. Teorin att det som skulle ha hänt om A inte hade inträffat är sådant som händer i den närmaste värld där A inte inträffar är väl helt enkelt en metafysisk teori, utan något vidare empiriskt stöd.

¹¹Om A inte hade inträffat, så är det oundvikligt att världen även i en del andra avseenden hade varit annorlunda, men man ska då tänka sig att dessa andra avseenden är så obetydliga som möjligt. Enligt Lewis (1979, s. 472) ska man i första hand undvika stora och omfattande avvikelser från naturlagarna, i andra hand ska så stora områden i rumtiden som möjligt vara oförändrade i förhållande till den verkliga världen, i tredje hand ska man undvika små och enkla brott mot naturlagarna och först i sista hand förändringar i enskilda sakförhållanden. Detta ger en viss ledning, men den är ju ganska vag och obestämd.

Om man vill veta vad som skulle ha skulle ha hänt om A inte hade varit fallet, så borde man kanske inte fokusera på det som händer i den närmaste världen där A inte är fallet, utan i stället på de icke-A-världar som är *sannolikast* om A inte är fallet – om man nu inte kan identifiera den värld som *faktiskt* skulle ha varit fallet om A inte varit fallet. Eftersom det man är ute efter är *värdet* av denna alternativa värld (för personen i fråga) kunde man också tänka sig att ta ett *viktat medelvärde* (för personen) av värdet av de världar där icke-A är fallet, där vikterna är sannolikheterna för respektive värld.

Men sådana sätt att modifiera förlustteorin verkar inte särskilt lovande. Framför allt har vi väl ingen möjlighet att på något objektivt sätt bestämma sannolikheterna för olika alternativa världar. Bland annat skulle detta förutsätta att vi kunde bedöma sanningshalten av olika kontrafaktiska villkorsatsar – och om vi kan det så är det ju enklare att gå direkt på frågan om vad som (faktiskt) skulle ha varit fallet om A inte hade varit fallet. Dessutom: det som ”faktiskt” skulle ha varit fallet är nog svårt att skilja från det som ”sannolikt” skulle ha varit fallet.

Enligt förlustteorin – åtminstone i de utformningar som liknar Feldmans – bestäms de relevanta alternativen av hur döden beskrivs. Exemplet med Herbert under Krimkriget illustrerar detta. I det fallet tycks Feldman alltså mena att (5) är den relevanta beskrivningen, vilket medför att den relevanta alternativvärlden är den närmaste värld i vilken Herbert dör när han inte längre är ung. Men även om (5) är mer ”relevant” än beskrivningarna (1)–(4), så finns det ju oändligt många andra sanna beskrivningar av Herberts död och det är oklart varför man just ska ta fasta på (5). Varför är den mer relevant än t.ex. ”Herbert dör innan han fyllt 85 år” eller ”Herbert dör innan han själv vill dö” eller helt enkelt ”Herbert dör”?

Det kunde kanske ligga nära till hands att ta fasta på den beskrivning som personen *själv* uppfattar som mest relevant. Det förutsätter förstås att bedömningen sker innan personen dör. Det innebär också att bedömningen sker så att säga ”inifrån” snarare än ”utifrån”. Denna distinktion kan säkert spela en viss roll. Vi bedömer antagligen andras liv på ett annat sätt än vårt eget. Psykologer har t.ex. experimentellt visat att vi utifrån bedömer lika lyckliga men olika långa liv som lika bra och att ett lyckligt liv bedöms som sämre om det förlängs med några mindre lyckliga, men ändå lyckliga, år. Det är alltså bara den lyckligaste och den sista perioden i livet som spelar roll för bedömningen.¹² Men detta gäller bedömningar utifrån. Inifrån – dvs. ur sitt eget perspektiv

¹²Se Kahneman 2012, s. 433–35. Man kan notera att detta strider mot Nagels antagande att livets värde är ”proportionellt mot dess utsträckning i tiden: mer är bättre än mindre” (Nagel 1979, s. 171–72).

– skulle de allra flesta troligen vilja ha ytterligare några ganska lyckliga år framför sig i stället för att dö.¹³

4. RÄDSLAN FÖR DÖDEN

Det är ju också våra bedömningar *inifrån* livet som hänger ihop med dödsrädsla och med vår föreställning att döden är något ont. Förlustteorin i dess vanliga tappning tycks inte alls kunna förklara dessa föreställningar. Det man oroar sig för, när man oroar sig för döden, är nog oftast något mer generellt eller mer obestämt än specifika sakförhållanden som att man dör vid en viss tidpunkt eller under den flygtur man just ska företa eller under anfallet med den lätta brigaden under Krimkriget – även om sådana specifika rädslor säkert också förekommer.

Det sakförhållande som en mer existentiell dödsrädsla riktas mot skulle väl snarare kunna uttryckas med satsen ”jag kommer att dö”. Den närmaste möjliga värld som är oförenlig med denna sats får då antas implicera ”jag är odödlig”.

Den närmaste värld i vilken jag är odödlig är säkert så olik vår värld att man knappast kan föreställa sig den. Naturlagarna måste rimligen vara helt annorlunda i en värld där människor lever men aldrig dör. Detta gäller även om jag vore den *enda* odödliga människan i den världen. *Finns* det överhuvud taget någon ”möjlig” värld med odödliga människor? Och hur kan man alls värdera något så konstigt? För att värdering av världar alls ska vara möjlig måste man nog hålla sig till världar som är väldigt lika vår, världar som inte skiljer sig mer än i något enstaka värdererlevant avseende.

Men här råder tydligen delade meningar. Jag har väldigt svårt att föreställa mig en värld i vilken jag är odödlig, men Nagel påstår tvärt emot att vi kan ”klart föreställa oss” att en individ fortsätter att existera i obegränsad tid.¹⁴ Jag håller med om att man väl alltid kan föreställa sig att man skulle kunna existera ytterligare någon tid. Men det är inte det samma som att föreställa sig att man fortsätter att existera i obegränsad tid, dvs. i evighet.¹⁵ Bland annat uppstår frågan om vilken ålder man ska

¹³Man kan notera att distinktionen mellan bedömningar *inifrån* och *utifrån* inte är densamma som den som avspeglas i förlustteoriens personrelaterade bedömningar. Att värld *w* är bättre än *w'* ”för personen *s*” betyder att *s* lever ett bättre liv i *w* än i *w'*. Och denna bedömning kan mycket väl göras ”utifrån” eller ”ur evighetens synpunkt”. Hur sedan bedömningen ser ut ”inifrån”, dvs. ur *s:s* eget perspektiv, är en annan sak.

¹⁴Nagel 1979, s. 178.

¹⁵Man kan förstås *säga* att man ska leva i evighet – kanske till och med ”i himmelriket” – men det är inte tillräckligt för att *föreställa* sig detta. Men, som sagt,

ha: ungefär samma biologiska ålder i evighet eller en i oändlighet tilltagande biologisk ålder. Bägge alternativen är avskräckande, men kanske finns det andra möjligheter.

Nagel använder sig också av distinktionen mellan utifrån- och inifrån-perspektiv. Även om vi kan föreställa oss att en individ existerar i evighet, så anser han att detta inte är en rimlig möjlighet. Utifrån anser vi inte att det döden berövar oss är ett evigt liv. Men ”inifrån finns å andra sidan inte denna föreställning om en naturlig tidsgräns”.¹⁶ Tydligt menar han att vi inifrån – ur vårt eget subjektiva perspektiv – uppfattar döden som att den berövar oss ett framtida *evigt* liv. Kanske menar han att det är därför vi är så rädda för döden.¹⁷

Detta är ingen övertygande förklaring. Hur vet Nagel att vi ”inifrån” inte föreställer oss en – visserligen rätt obestämd – tidsgräns? Själv gör jag i alla fall det. Men om ”felet” med döden skulle vara att den berövar oss ett evigt liv, så måste man också visa att det verkligen skulle vara önskvärt att leva i evighet. För egen del skulle jag säga att ett evigt liv förefaller rätt skrämmande.¹⁸ Detta beror bland annat på att det är så svårt att föreställa sig hur det skulle vara att leva i evighet – såvida detta inte bara skulle bestå i att ens vanliga jordiska liv helt enkelt upprepas i det oändliga.¹⁹

5. SYMMETRIPROBLEMET

Ibland har man tänkt sig att förlust av ett möjligt framtida liv efter döden kan jämföras med förlust av ett möjligt tidigare liv före födelsen. Att vara död kan jämföras med att ännu inte vara född. Nagel anser att ett problem med förlustteorin är att vi inte uppfattar prenatal icke-existens som en förlust, trots att den i relevanta avseenden liknar postum icke-existens.²⁰

Såvitt jag förstår är detta så kallade symmetriproblem inget allvarligt problem. Det är ju inget konstigt i att man är mer rädd för postum icke-

andra tycker tydligt annorlunda. Här vill jag dock tillägga att det kanske ändå finns *ett* sätt att föreställa sig hur det är att leva i evighet, trots att man dör, nämligen att anta att tiden är cirkulär, så att man alltid återkommer till sin födelse efter att man har dött; se Bergström 2012.

¹⁶Nagel 1979, s. 179.

¹⁷I slutet av sin artikel säger han: ”Om det inte finns någon gräns för hur mycket liv som det vore värdefullt att ha, så är kanske det slut som väntar oss alla ett dåligt slut” (ibid., s. 181).

¹⁸En känd artikel på detta tema är Williams 1973.

¹⁹Att ens liv upprepas innebär förstås att man inte *minns* något från tidigare instanser av livet.

²⁰Nagel 1979, s. 173–74.

existens än för prenatal icke-existens. Det kan antas bero på att vi är biologiskt programmerade av evolutionen att vilja *fortsätta* leva. De som har en sådan livsvilja kan förväntas klara sig bättre än andra i kampen för tillvaron och sålunda sprida sina gener i högre grad. Och även om man skulle ha kunnat leva ett långt och lyckligt liv före sin födelse – vilket somliga betvivlar, med motiveringen att man i så fall skulle ha varit en annan person²¹ – så är detta inte särskilt intressant när det gäller vår inställning till döden.

Om man nu inte är rädd för prenatal icke-existens, så kan man ändå undra varför vi normalt inte betraktar ett möjligt liv före födelsen som en förlust. Det kunde kanske ha varit trevligt att ha levat ett långt och lyckligt liv innan man faktiskt föddes. Men om man försöker tillämpa förlustteorin på födelsen i stället för på döden, så ska man väl jämföra den aktuella världen med den *närmaste* alternativa värld i vilken man föddes tidigare – och i den världen föddes man väl bara någon sekund tidigare, så den kan väl antas vara ungefär lika bra eller dålig för en själv som den aktuella världen. Om man i stället vill jämföra med en möjlig värld i vilken man föddes *betydligt* tidigare, och levde ett bra liv under tiden fram till tidpunkten för ens faktiska födelse, så vet man ju inget alls om hur långt eller hur lyckligt ens fortsatta liv från den tidpunkten skulle ha blivit i den alternativa världen. En sådan jämförelse blir alltså omöjlig. Den kan inte visa något om hur bra eller dåligt det skulle ha varit att födas tidigare än man faktiskt föddes.

Det klassiska symmetriargumentet, från Lucretius Carus dikt *De Rerum Natura*, kunde kanske formuleras så här: eftersom det inte innebär en förlust att födas när man föddes (snarare än tidigare), så kan det inte heller innebära en förlust att dö när man dör (snarare än senare). Det kan låta förnuftigt – och många framstående filosofer har också tyckt att det är förnuftigt – men i själva verket är det inte alls klart att dessa två möjligheter är symmetriska eller lika i relevanta avseenden. Om man hade fötts tidigare, så hade ens liv säkerligen inte innefattat det liv man lever i den verkliga världen, men om man hade dött senare, så kunde ens liv ha innefattat det liv man lever i den verkliga världen.

Om man t.ex. hade fötts hundra år tidigare, så skulle man ha varit hundratio år vid den tidpunkt då man i sitt verkliga liv var bara tio år – om man hade levat så länge. Om man hade fötts tio år tidigare, så skulle man antagligen haft helt andra andra vänner och andra upplevelser än i sitt verkliga liv. Kanske också andra föräldrar. Och man skulle troligen ha dött tidigare. Osv. Men om man hade dött senare än man faktiskt gör i den verkliga världen – dvs. om livet helt enkelt hade blivit längre – så

²¹Se Nagel 1979, s. 178.

skulle man inte ha förlorat något av sitt verkliga liv. Åtminstone är det nog så man normalt föreställer sig saken. I och för sig *kunde* – eller skulle – kanske livet före det verkliga dödsögonblicket också ha varit delvis anorlunda, t.ex. hälsosammare eller mer händelsefattigt och därmed tråkigare, om man hade dött senare. Men i så fall är det ju inte säkert att det som helhet hade varit bättre ens om det extra tillskottet hade varit bra. Och då blir jämförelsen återigen omöjlig.

6. FÖRLUST AV MEDVETANDET

Förlustteorin tycks inte alls kunna göra rättvisa åt den negativa inställning till döden – eller den negativa värdering av döden – som folk oftast har.²² Men det kan ligga nära till hands att tycka att det man förlorar när man dör är *medvetandet* och att det framför allt är detta man är rädd för. I och för sig är det ju inte så allvarligt att förlora medvetandet; det gör man ju varje gång man somnar eller när man svimmar. Det allvarliga är förstas i så fall att förlusten är *permanent*.

Emellertid skulle det väl vara lika skrämmande om medvetelslösheten inte vore helt permanent, utan att man t.ex. återfick medvetandet under en minut varje sekelskifte i framtiden – eventuellt i evighet eller till universums död.²³ Å andra sidan är ju detta fullt förenligt med att det ändå i första hand är permanent medvetelslöshet man är rädd för.

Men det kan ändå inte gärna vara själva tillståndet, att vara medvetlös, som är något dåligt.²⁴ Snarare måste det vara att förlora något som medvetelslösheten förhindrar, till exempel att man är vid medvetande och kanske lever ett någorlunda normalt liv. Sätillvida har nog förlustteorin rätt. Men som vi har sett kan det inte gärna vara det extra liv man skulle ha levat i den ”närmaste” alternativa världen som man är rädd för att

²²Folk är förmodligen olika i detta avseende. Nagel skriver: ”Vissa människor anser att döden är fruktansvärd; andra har inget att invända mot döden som sådan men de hoppas att deras egen varken ska komma i förtid eller vara smärtsam” (Nagel 1979, s. 170). Observera att den rädsla som det här är fråga om är självcentrerad. Vi bortser nu från den rädsla man kan känna för att t.ex. ens barn eller andra närstående kan påverkas negativt av att man dör.

²³En liknande spekulatioon finns hos Draper 2012, s. 304–5. Han tänker sig att man skulle vara medvetlös i evighet, utom under en sekund varje natt, vilket han tycker är sämre än att vara permanent död.

²⁴Det är denna insikt som Nagel uttrycker i en passage jag redan har citerat i början av denna uppsats: ”om döden är ett ont är det *förlusten av livet* snarare än tillståndet att vara död, icke-existerande eller medvetlös, som är oacceptabelt”. Men i den avslutande fotnoten säger han ändå: ”Min slutsats blir att någonting, som har att göra med den framtida *utsikten* av ett permanent ingenting, inte fångas i analysen i termer av förlorade möjligheter” (1979, s. 181).

förlora. Ty detta extra liv skulle troligen vara väldigt kort och relativt neutralt ur värdesynpunkt.

Så vad är det då man inte vill förlora, om nu felet med döden är att den innebär en förlust?

7. EN BRA DÖD

Om man är rädd för döden, så kan man ju fråga sig vad man skulle vilja ha i stället. Ett *evigt* liv verkar inte lockande, utan snarast väldigt skrämmande. Hellre då ett *fortsatt* liv. Eller mer anspråkslöst: att leva lite till.

Nej, ett bättre förslag vore nog ett *lagom* långt liv. Jag tänker mig då att ett sådant liv vore ett liv som sträcker sig fram till dess att man inte längre vill leva. Även den mest fanatiska dödsvägrare skulle väl välja ett sådant liv framför ett evigt liv – oavsett om det utspelar sig i vår vanliga verklighet eller i någon sorts himmelrike.

Men att man vid en tidpunkt ”inte längre vill leva” måste då tolkas så, att man inte heller *efter* denna tidpunkt *skulle* ha velat leva vidare, om man inte hade dött då. Detta krav gör visserligen att man aldrig kan veta om det liv man har levat är lagom långt. Men det verkar ändå klart att det extra villkoret måste vara med. Åtminstone om livet ska bedömas från en utomstående perspektiv. Den som bedömer sitt eget liv kan ju också vara rädd för att det extra villkoret vid en given tidpunkt inte är uppfyllt, men om den rädslan är tillräckligt stor, så utgör den ju också ett skäl för att inte vilja dö vid denna tidpunkt.

Ett i denna mening ”lagom långt liv” verkar vara precis vad man kan önska sig, men det är sannolikt inte detta man förlorar genom att dö. Det är i allmänhet inte vad som i närmaste alternativa värld skulle ha förverkligats, om man inte hade dött när man dog. Det är alltså inte det man förlorar genom att dö.

8. DEN RELEVANTA BESKRIVNINGEN AV DÖDEN

Här skulle någon kanske vilja invända mot mitt sätt att tillämpa förlustteorin. Jag har ju tidigare noterat att en persons död sanningsenligt kan beskrivas på oändligt många olika sätt och det finns alltså ingen ”närmaste alternativa värld” (där beskrivningen av döden negeras) att jämföra med så länge man inte har identifierat den ”relevanta” beskrivningen av personens död.²⁵ Hittills har jag dock mest hållit mig till be-

²⁵Den som betvivlar att det verkligen finns *oändligt* många sanna beskrivningar av en persons död kan ju betänka att om t.ex. jag dör innan jag blir 100 år, så är det också sant att jag dör innan jag blir 101 år, och sant att jag dör innan jag blir 102 år, och sant att jag dör innan jag blir 103 år ... och så vidare i all oändlighet.

skrivningar av typen ”Personen dör” och ”Personen dör vid tidpunkten t”.²⁶ Men om jag har rätt i att det man rimligen kan önska sig är att leva ett lagom långt liv – och om det man rimligen kan vara rädd för att inte leva ett lagom långt liv – så ligger det ju nära till hands att säga att den relevanta beskrivningen av en död som inte inträffar precis i slutet av ett lagom långt liv är ”Personen dör för tidigt eller för sent”.²⁷ Och det alternativ man då ska jämföra med, enligt förlustteorin, är den närmaste värld där personen i fråga varken dör för tidigt eller för sent. Och det kan ju tyckas uppenbart att den alternativa världen är bättre för personen än den i vilken han eller hon dör för tidigt eller för sent. Och då är det väl ändå ett lagom långt liv man förlorar genom att dö för tidigt eller för sent.

Ja, men förlustteorin ger inte heller med denna tolkning någon bra förklaring av varför döden är något ont eller något att vara rädd för. Ty den *närmaste* värld i vilken man varken dör för tidigt eller för sent skiljer sig nog inte särskilt mycket från den verkliga världen – alldeles oavsett när man faktiskt dör. Och därmed är dess värde ungefär detsamma som värdet av den faktiska världen, vilket gör att förlusten är närmast negligerbar.

Anta att jag dör vid femtio års ålder och att det är alldeles för tidigt. Den närmaste värld i vilken jag varken dör för tidigt eller för sent är väl då den i vilken jag några sekunder senare *vill* dö och även dör (och även senare skulle ha velat dö). Den världen är väldigt lik den verkliga och har därmed ungefär samma värde. Så det blir ingen större förlust.

9. DÖDENS VÄRDE

Förlustteorin är riktig såtillvida att, som Nagel uttrycker det, ”om döden över huvud taget är ett ont, så kan det inte vara på grund av dess egentliga egenskaper utan endast på grund av vad den berövar oss”. Men förlustteorin kan inte förklara varför döden är något så ont, som väldigt många tycks anse.

Döden *måste* inte innebära en förlust. Den behöver t.ex. inte ses som en förlust om den innebär slutet på ett lagom långt liv. Hur ofta ett liv är lagom långt är svårt att säga, men det är nog inte särskilt ofta.

Den förlust som döden åstadkommer, i jämförelse med *närmaste* möjliga alternativa värld, är i allmänhet inget att oroa sig för. Och det är inte heller, som Nagel säger, förlusten av livet som är oacceptabel. Den

²⁶Karl Ekendahl tycks mena att det problem med förlustteorin jag pekat på (i Bergström 2013) försvinner om man utgår från lämpligare beskrivningar av döden; jfr Ekendahl 2013, s. 46.

²⁷Att få leva ett lagom långt liv är förstås inte *det enda* man rimligen kan önska sig; man kan t.ex. också önska sig att resten av livet blir lyckligt.

”förlust” som är betydelsefull – och kanske rätt vanlig – är snarare förlusten av ett lagom långt liv. Men denna förlust drabbas man normalt inte av genom att dö.

LITTERATUR

- Bergström, Lars. 2012. ”Death and Eternal Recurrence”. I *The Oxford Handbook of Philosophy of Death*, red. B. Bradley, F. Feldman och J. Johansson. Oxford: Oxford University Press.
- Bergström, Lars. 2013. ”Döden och tiden”. *Filosofisk tidskrift*, nr 1.
- Bradley, Ben. 2009. *Well-Being and Death*. Oxford: Oxford University Press.
- Broome, John. 2012. ”The Badness of Death and the Goodness of Life”. I *The Oxford Handbook of Philosophy of Death*, red. B. Bradley, F. Feldman och J. Johansson. Oxford: Oxford University Press.
- Draper, Kai. 2012. ”Death and Rational Emotion”. I *The Oxford Handbook of Philosophy of Death*, red. B. Bradley, F. Feldman och J. Johansson. Oxford: Oxford University Press.
- Ekendahl, Karl. 2013. ”Döden igen”, *Filosofisk tidskrift*, nr 4.
- Feldman, Fred. 1991. ”Some Puzzles About the Evil of Death”. *The Philosophical Review* 100, nr 2.
- Kahneman, Daniel. 2013. *Tänka, snabbt och långsamt*. Stockholm: Volante.
- Lewis, David. 1979. ”Counterfactual Dependence and Time’s Arrow”. *Noûs* 13, nr 4.
- McMahan, Jeff. 1988. ”Death and the Value of Life”. *Ethics* 99, nr 1.
- Nagel, Thomas. (1970) 1979. ”Döden”. I *Frågor om livets mening*, utg. L. Bergström. Stockholm: Prisma.
- Silverstein, Harry. 1980. ”The Evil of Death”. *The Journal of Philosophy* 77.
- Williams, Bernard. 1973. ”The Makropulos Case: Reflections on the Tedium of Immortality”. Kap. 6 i hans bok *Problems of the Self*. Cambridge: Cambridge University Press.

Friedrich Nietzsche skrev *Fallet Wagner* sent i sitt författarskap.¹ Även om den inte skriven i traditionell brevform är undertiteln ”Brev från Turin i maj 1888”. Det är en egendomlig text som behandlar Bizets *Carmen* och Wagners kompositioner. Nietzsche skriver att *Carmen* ”träder fram så lätt, så smidigt, så sirligt artig. Den är älskvärd, den svettas inte”. Wagners verk ger däremot upphov till motsatt reaktion: ”[Wagners orkes-terklang] brutal, konstlad och samtidigt ’oskuldsfull’ [...] En föreglig svett bryter ut över mig” (FW, s. 17). Den här typen av uttryck är ofta förekommande i *Fallet Wagner*. Frågan är hur man kan närma sig dessa anmärkningar och iakttagelser. Är de mer eller mindre av ”biografiskt” eller ”samtidskulturellt” intresse? Eller är de berättigade i Nietzsches filosofi? Argumentet jag kommer att föra i denna artikel är att det finns en filosofisk kontext för den här typen av beskrivningar. Jag inleder med att undersöka polariteten mellan Wagner och Bizet, sedan Nietzsches ”fysiologiska” förståelse av konst, för att avslutningsvis skissera några implikationer från *Fallet Wagner*.

1. BIZET OCH WAGNER

Nietzsche inför flera kontraster och motsatsförhållanden i *Fallet Wagner*: mellan det norra och det södra, mellan det som känns ”grått och fult och kallt” och ”de lätta fötterna, kvickhet, eld” (FW, s. 36). Denna kontrast speglar också skillnaden mellan Wagner och Bizet.² Det första kapitlet i *Fallet Wagner* inleds följande: ”Igår lyssnade jag – ni må tro mig eller ej – för tjugonde gången på Bizets mästerverk. Jag stod än en gång ut med en stilla andakt, jag sprang än en gång inte därifrån” (FW, s. 17). Nietzsche skriver att ”Wagners konst utövar ett tryck av hundra atmosfärer: varsågod och krök på nacken, vad annat kan man göra” (FW, s. 30). Och: ”Med *Carmen* tar man farväl av det *fuktiga* Norden, av det wagnerska idealets

¹*Fallet Wagner* (”FW”) skrevs 1888, svensk översättning 1992.

²Mer strikt är det inte Bizet mot Wagner, utan Bizets *Carmen* mot Wagner. Nietzsche talar inte om några andra av Bizets verk.

alla dunster. [...] Här talar en annan sinnlighet, en annan sensibilitet, en annan munterhet.” Denna sinnlighet hos Bizets verk har en ”sorts sensibilitet som hittills saknat röst i Europas konstmusik – denna sydligare, mera brunbrända, mer förbrända sensibilitet ...” (FW, s. 18). Wagners karaktäriseras av det överväldigande, det ansträngande och tunga: ”Jag tror jag måste vädra ut ett slag. Luft! Mera luft! – ” (FW, s. 23). Bizets opera blir en form av kur mot det tunga, feberaktiga och nästan sjukliga i Wagners verk. Det är med detta i åtanke Nietzsche formulerar sitt eget bud för estetiken: ””Det goda är lätt, allt som är gudomligt rör sig på mjuka fötter”: första budet i min estetik” (FW, s. 17). Nietzsche gör även ett antal andra anmärkningar, som vid första anblicken kan verka egendomliga: ”Så fullkomnad man blir av ett sådant verk! Man blir på köpet själv ett ’mästerverk’” (FW, s. 17). Och lite senare: ”Bizet gör mig fruktbar” (FW, s. 18). Frågan är förstås vad det innebär för Nietzsche när han skriver att man blir ”fruktbar”, blir ett ”mästerverk”, genom ett verk som *Carmen*.

2. FYSIOLOGI OCH KONST

När Nietzsche skriver om sina reaktioner till Wagners och Bizets verk är det deras *effekter* som beskrivs. Det är *ovillkorliga* reaktioner Nietzsche beskriver. Det är en avgörande punkt: kroppen reagerar *innan* tänkandet. Medvetandet, omdömet, blir sekundärt. När tänkandet äger rum har redan effekter och konsekvenser ägt rum; och dessa första intryck är mer ursprungliga, mer nyanserade än de man gör i ett normalt, distanserat, reflektivt omdöme. Därigenom försöker Nietzsche bilda ett ursprungligare förhållningsätt till konstverket i en biologisk, fysiologisk bemärkelse. Inte bara i mottagandet av konstverket men också i skapandet det är en sådan dimension närvarande: Nietzsche kommer tala om *ruset* som en betingelse för skapandet av konst.³ Men hur leder detta ursprungli-

³Rusets roll förklaras mer ingående i *Avgudaskymning* där han menar att konsten skapas från av ett tillstånd av överflöd som i sig initieras av ett rus. Det är ett överflöd som går över sig själv och på tingen och blir därför ett villkor för den livsintensifierande, bejakande konsten. Återigen är det överbrygningar som sker: mellan artisten och konstverket, mellan konstverket och åskådaren, mellan artisten och tingen omkring en, mellan åskådaren och tingen omkring en, osv. Det är ett *stegrande* av intensitet, starkt bunden till *överflöd*. Det är ett tillstånd – och det är från ett sådant tillstånd man berikar allt med ”sitt eget överflöd”. Följden av en sådan stegring är en skiftning i relationen till omgivningen, till tingen. Nietzsche beskriver det som att man tvingar tingen att formas efter en, att man våldför sig på dem, ”*tvingar* dem att ta emot från oss” (*Avgudaskymning*, s. 130). Denna våldförande akt benämns som *idealisering*. Idealisering, som Nietzsche definierar

gare förhållningssätt till ett värderande av konstverket – dvs. hur avgör Nietzsche vilket konstverk som har ett högre värde, att upphöja eller angripa, varför värderas *Carmen* högre än Wagners konst? Värderingen sker genom att han inför en distinktion mellan det som han benämner som ”uppåtgående livets dygder” och ”förfallsdygderna”.⁴ Dessa bildar polaritet. Medan förfallsdygderna är ett *nekande* av krafternas överflöd är de uppåtgående dygderna ett *bejakande* av dem. Det är genom att placera den sinnliga effekten – eller kanske till och med *eftersomaken* – av konstverken i förhållande till den modellen som de kan värderas. Nietzsche menar därutöver att det inte bara finns en intim relation mellan konstverket och åskådaren av konstverket, konsten har även en inverkan på tingen omkring en. Förfallsdygderna, ”dekadensens” estetik, ”förfular” och ”utarmar tingens värde”, ”berövar dem deras färg”, ”förnekar världen”. Den uppåtgående estetiken ”förklarar och förskönar världen” (FW, s. 45). Implicit i detta förskönande av världen genom konstverket sker ett förfinande och mognande av en själv. Att man ”blir” ett mästerverk är en följd av de uppåtgående dygderna i konsten. Det lätta och det behagligt muntra sporrar till en förvandling, förfining, till och med förädling, inte bara hos världen omkring en, utan också hos en själv.

3. MELLAN KONST OCH FILOSOFI

Det är anmärkningsvärt att man inte hittar något verk eller något kapitel specifikt avsatt för estetik i Nietzsches verk. Närmast är *Tragedins födelse*, ett verk som ledigt rör sig mellan vitt skilda discipliner. Likväl är konsten och konstanalysen djupt närvarande i Nietzsches författarskap – i synnerhet i början och i slutet. Linjerna och förbindelserna Nietzsche gör mellan till synes skilda teman kan förefalla vara godtyckliga (mellan klimat, Wagner som ”teatergeni”, Tysklands ungdom, etc). Men bakom dessa förbindelser kan man se ett mönster eller en stil som gör att de ofta naturligt flyter in i varandra. Detta hör till Nietzsches sätt att skriva och tänka. Och som Sven-Olov Wallenstein skriver utmärks Nietzsches tänkande just genom stilen: det är ett försök att bryta emot implicita

eller omdefinierar det, är det som följer av ett pådrivande och påtvingande av en så att andra drag successivt avlägsnas. Att man skapar, så att säga, inifrån och ut och inte utifrån in, är en markant förändring från utläggningarna i *Tragedins födelse*. Där menar Nietzsche att man som konstnär mer eller mindre endast *härmar* naturen, förmedlad genom dionysiska och apolliniska konsttendenser.

⁴”Estetiken är oupplösligen förenad med dessa biologiska förutsättningar: det finns en *décadence*estetik och en *klassisk* estetik – det ”sköna i sig” är ett hjärnspöke, liksom allt vad idealism heter” (FW, s. 45).

såväl som explicita krav på hur man skriver inom filosofin.⁵ I Nietzsches behandlande av konsten ser man hur det konstant sker ett försök till att sammanfläta distinktioner man tenderar att (explicit eller implicit) uppställa. Redan från *Tragedins födelse* finns försök till att överbrygga dikotomin mellan kören och åskådaren, artisen och världens kreativitet. Sådana överbryggingar och överträdelser finns även i *Avgudaskymning*; mellan konstnären och konstverket, mellan konstverkets inverkan på en och tingen omkring en.⁶ Vad som blir tydligt i Nietzsches fall är att han inte alls är intresserad av att upprätthålla konstens autonomi i Kants mening. Konsten är allt annat än ett intresselöst betraktande av konstverken. Inte heller är den menad att vara en fritidssyssla eller förströelse.⁷ Konsten – dvs. konsten i en *vid* mening – får sitt starkaste syfte när konstverk som är skapade från ett överflöd av liv, intensifieras och stegras till ett överflöd av liv för den som tar del av den.⁸ Med denna nödvändigtvis okänsliga översikt i åtanke kan man återvända till den ursprungliga frågan för denna artikel.

4. AVSLUTANDE ANMÄRKNINGAR

Fallet Wagner kan förstås läsas från många perspektiv. Enligt min mening är Nietzsches personliga reaktioner och omdömen till operor och konstverk i *Fallet Wagner* filosofiskt berättigade. Nietzsche använder sig själv

⁵Wallenstein, *Bildstrider*, s. 106.

⁶I *Avgudaskymning* ställer Nietzsche frågan om konstnären skapar i hunger eller i överflöd. Han inför också ett motsatsförhållande mellan "konstnären" och "antikonsnären", som syftar på konstnären som antingen skapar från hunger eller överflöd. Nietzsche menar att tingen omkring en *utarmas* i beröring med "antikonsnären". "Livets uthungrade" är enligt Nietzsche de som "med nödvändighet måste ta tingen till sig, tära på dem, göra dem *magrare*". *Antikonsnären* beskrivs som parasiterande på tingen: "historien är full av sådana antikonsnärer, av sådana livets uthungrade: vilka med nödvändighet måste ta tingen till sig, tära på dem, göra dem *magrare*" (*Avgudaskymning*, s. 130).

⁷Nietzsche: "Låt oss aldrig acceptera att musiken 'vederkvicker', att den 'förströr', att den 'roar'. *Låt oss aldrig roa!* – vi är förlorade om man på nytt börjar betrakta musiken hedonistiskt . . . Det är 1700-talet när det är som sämst" (FW, s. 27). Det är en tanke Nietzsche har redan från början. I förordet till *Tragedins födelse* kritiserar Nietzsche konsten som en "lustig bisak", en "bjällerklang", i våra liv (s. 21).

⁸Nietzsche talar om konst både i en *vid* bemärkelse och konst i *snäv* bemärkelse. När Nietzsche talar om konst i *vid* bemärkelse, dvs. när konsten förhöjs, är det alltid som en följd av ett *överflöd*. Och konst i den snäva betydelsen, som de flesta skulle förstå som konst, i museer och populära böcker är konst som snarare reflekterar ett utarmat liv.

som en mätare för att beskriva närmast fenomenologiska beskrivningar av konstverkens inverknings på honom: att Bizets *Carmen* framkallar en sinnlig lätthet och afrikansk munterhet och att Wagners operor får honom att svettas är ofrivilliga, nästan hälsomässiga reaktioner. Det är dessa reaktioner inför konstverket Nietzsche uppmärksammar eftersom de berättar något värdefullt om konstverket som annars tenderar förloras. Det är en ursprungligare relation till konstverken och också ett sätt för honom att värdera dem. Det förklarar varför Nietzsche inte är intresserad av konstverket *i sig*. Konstverket analyseras alltid i relation *till* något, antingen till skapandet av verket för artisten eller i receptionen av det av åskådaren. Konsten – den *egentliga* konsten, konstverk som *Carmen* – är menad att sammanflätas med livet självt: konstens bejakande av liv, livets överflöd uttryckt genom konst. Således bryter Nietzsche mot idén om estetikens autonomi i Kants mening. Men detta brytande öppnar också för möjligheten att på ett nytt sätt reflektera kring relationen mellan konst och filosofi. För Nietzsche är konsten inte oavhängig eller ens i periferin av hans mer filosofiska utläggningar. Det rätta förhållandet till konst berättar också något om bedrivandet av filosofi. I en av de mest slående meningarna från *Fallet Wagner* beskriver Nietzsche hur han blir mer av en filosof genom *Carmen*: ”Och varje gång jag har lyssnat på *Carmen* tycks det mig verkligen som om jag blivit mer filosof, en bättre filosof än annars: så fördragsam, så lycklig, så indisk, så *rotad*...” (FW, s. 17). Läst från ett sådant perspektiv kan viktiga ledtrådar utvinnas från *Fallet Wagner*, ifråga om hans förståelse av konst och hans filosofi.

LITTERATUR

- Nietzsche, Friedrich. *Avgudaskymning: eller Hur man filosoferar med hammaren*. I *Samlade skrifter*, Bd 8, *Sena skrifter*. Stockholm: B. Östlings bokförlag Symposion, 2013.
- Nietzsche, Friedrich. *Fallet Wagner*. I *Samlade skrifter*, Bd 8, *Sena skrifter*. Stockholm: B. Östlings bokförlag Symposion, 2013.
- Nietzsche, Friedrich. *Tragedins födelse*. I *Samlade skrifter*, Bd 1. Eslöv: B. Östlings bokförlag Symposion, 2000.
- Wallenstein, Sven-Olov. *Bildstrider: Föreläsningar om estetisk teori*. Göteborg: AlfabetaAnamma, 2001.

Aktualism, possibilism och handlingsriktighet

1. INTRODUKTION

Det förefaller vara sant att det inte finns någonting som inte existerar, eller – med andra ord – att allting existerar. Men är denna uppfattning riktig och är det i så fall logiskt sant och därmed nödvändigt att allting existerar, eller är det möjligt att det finns någonting som inte existerar? *Aktualister*, (se t.ex. Plantinga 1974, Adams 1981, Menzel 1991, Fitch 1996), hävdar vanligtvis att allting existerar och att detta är nödvändigt, medan *possibilister* (se t.ex. Linsky och Zalta 1994 och Williamson 1998) i regel förnekar detta.

Olika aktualister och possibilister definierar *aktualismen* och *possibilismen* på lite olika sätt. Vi skall i denna uppsats skilja mellan två tolkningar av den förra tesen, (A) och (NA), och två tolkningar av den senare, (P) och (MP).

- (A) Allting existerar.
- (NA) Det är nödvändigt att allting existerar.
- (P) Det finns ting som inte existerar. (Någonting existerar inte.)
- (MP) Det är möjligt att det finns ting som inte existerar. (Det är möjligt att någonting inte existerar.)

Enligt possibilister som är anhängare av (P) finns det, i någon mening, (icke-existerande) ting som är blott möjliga, kanske mytiska individer, såsom Pegasus (en bevingad häst i den grekiska mytologin), fiktiva karaktärer, såsom Bilbo (en sagofigur skapad av författaren Tolkien), eller hypotetiska entiteter, såsom Vulcan (en hypotetisk planet i vårt solsystem mellan Solen och Merkurius). Possibilistiska anhängare av (MP) hävdar att det är möjligt att det finns sådana ting. Att det *finns* eller *kan finnas* blott möjliga objekt betyder att vi kan tala om objekt av detta slag, referera till dem och kvantifiera över dem, trots att de inte existerar. Vi kan t.ex. tala om och referera till Pegasus, Bilbo och Vulcan. Enligt possibilister inkluderar "allting" mer än bara de ting som

existerar, vi kvantifierar inte enbart över *existerande* ting utan också över icke-existerande *möjliga* objekt. Possibilister gör ibland en distinktion mellan *vara* och *existera*: allt *är*, men allt *existerar* inte. Aktualister förnekar i regel en sådan distinktion; de menar vanligtvis att allt som *är* eller *har vara* existerar. Enligt aktualistiska anhängare av (A) existerar allting och enligt aktualistiska anhängare av (NA) är det nödvändigt att allting existerar; vi kvantifierar endast över ting som existerar. Possibilismen bör inte blandas ihop med tesen att det kunde ha existerat ting som faktiskt inte existerar, vilket är en sats som de flesta accepterar, även de flesta aktualister.

Aktualismen kan förefalla vara uppenbart sann. Hur skulle det kunna finnas något som inte existerar? Och vad är i så fall en icke-existerande entitet? Var finns dessa entiteter? Och vad har de för identitetsvillkor? Man kan också formulera olika argument för aktualismen. Antag att vi inför ett existenspredikat i klassisk predikatlogik och att detta predikat kan definieras i termer av den s.k. existenskvantifikatorn på så sätt att x existerar om och endast om det finns ett y sådant att y är identiskt med x . Då är det logiskt sant att allting existerar. För i klassisk predikatlogik är det logiskt sant att alla x är identiska med något y . Alltså är (A) logiskt sann. Vidare gäller det att alla logiskt sanna satser är nödvändiga enligt klassisk modallogik. Så, om vi kombinerar klassisk predikatlogik med klassisk modallogik, så följer det att det är nödvändigt att allting existerar. (NA) är med andra ord logiskt sann.

Emellertid är det även lätt att hitta argument som ger oss skäl att tro att possibilismen är sann. Vi tycks t.ex. dagligen tala om, tänka på och resonera om ting som inte existerar: mytiska varelser, fiktiva karaktärer i böcker, filmer, TV-serier och teaterpjäser, hypotetiska entiteter som postuleras inom vetenskaperna och filosofin, vilket lätt får oss att sluta oss till att sådana ting finns. Betrakta t.ex. följande argument.

Pegasus existerar inte.

Alltså finns det någonting som inte existerar. (Någonting existerar inte.)

Premissen förefaller vara (kontingent) sann och slutsatsen tycks följa ur premissen. Men i så fall är possibilismen sann.

Man kan också hitta flera, mer tekniska, argument i litteraturen emot aktualismen. Om man kombinerar klassisk predikatlogik med klassisk modallogik så medför aktualismen flera problematiska teser. Här är ett argument av detta slag. Antag att aktualismen är sann.

1. Det är nödvändigt att allting existerar. [Antagande]
2. Om det är nödvändigt att allting är A, så är allting nödvändigtvis A.

[Från klassisk modallogik i kombination med klassisk predikatlogik; detta är den s.k. konversa Barcansatsen (CBF), uppkallad efter Ruth Barcan Marcus]¹

3. Om det är nödvändigt att allting existerar, så gäller det för allting att det är nödvändigt att det existerar. [Instans av 2]
4. Allting existerar nödvändigtvis. [Från 1 och 3 med hjälp av klassisk satslogik]

Premiss 1 är (NA). Om vi antar att existenspredikatet kan definieras i termer av existenskvantifikatorn, följer (NA) ifrån (A), vilket vi sett ovan. Vi kan, i så fall, anta (A) i stället för (NA) i argumentet. Enligt sats 4 så gäller det för allting att det är nödvändigt att det existerar; det finns inga kontingenta ting, det finns ingenting som kunde ha saknat existens. Med detta är absurt. Det finns många kontingenta ting, t.ex. olika levande organismer och mänskliga artefakter. Alltså är aktualismen inte sann. Jag existerar, men det är inte nödvändigt att jag existerar. Om mina föräldrar inte hade träffats, skulle jag inte ha existerat. Eiffeltornet existerar, men det är inte nödvändigt att Eiffeltornet existerar. Om Gustave Eiffel aldrig hade konstruerat Eiffeltornet, så skulle det inte ha existerat. I ljuset av dessa argument är det inte uppenbart att aktualismen är sann. Jag skall dock inte närmare diskutera dessa invändningar här. Jag skall nöja mig med att konstatera att dessa argument tycks visa att aktualismen är oförenlig med klassisk predikatlogik och att aktualisten bör acceptera någon form av s.k. fri logik, eller att det åtminstone är mer naturligt att kombinera aktualismen med någon form av fri logik.

Fri logik är ett alternativ till klassisk logik som förkastar vissa av de antaganden man gör i klassisk logik. I fri logik kan vi inte sluta oss till att en viss individ har egenskapen F från påståendet att alla x har F; vi kan bara sluta oss till att alla existerande ting har F. Det finns huvudsakligen tre varianter av fri logik: positiv, negativ och neutral. I positiv fri logik kan vi inte sluta oss till att det existerar något som har egenskapen F från påståendet att en individ har F. Vi kan bara sluta oss till detta om vi först vet att denna individ existerar. I negativ fri logik kan man emellertid dra denna slutsats.

Om någon form av fri logik är korrekt, så följer det att allting existerar,

¹Jag försöker i den här uppsatsen undvika tekniska symboler. Men för att underlätta förståelsen kommer jag även i några fotnoter att presentera vissa satsers symboliskt. (CBF) kan symboliseras på följande sätt $\Box \forall x A \rightarrow \forall x \Box A$, och följer omedelbart från axiomat $\forall x A \rightarrow A$, regeln att om $A \rightarrow B$ är ett teorem så är också $\Box A \rightarrow \Box B$ ett teorem och klassiska introduktionsregler för all-kvantifikatorn. \Box är den s.k. box-operatorn som läses "Det är nödvändigt att".

dvs. att aktualismen är sann.² Fri logik skulle därför kunna ses som en explikation av aktualismen.

Syftet med denna uppsats är att visa att debatten mellan aktualister och possibilister är relevant för moralfilosofiska riktighetskriterier, samt att diskutera ett argument som talar emot aktualismen (och därmed för possibilismen). Vi skall kalla detta argument ”det normativa argumentet”. Låt oss se närmare på det.

2. DET NORMATIVA ARGUMENTET

Många normativa teorier, som betraktar normativa egenskaper såsom egenskaper som tillkommer enskilda handlingar, går ut på att hitta nödvändiga och tillräckliga villkor för när en handling är riktig, obligatorisk eller förbjuden. Sådana teorier kvantifierar över handlingar. Men inte bara existerande handlingar, utan också framtida handlingar som ännu inte existerar och t.o.m. blott möjliga handlingar som inte har existerat, inte existerar och inte kommer att existera, men som kunde ha existerat. Det tycks åtminstone vara vad de flesta teoretiker tänker sig. Många normativa teorier förefaller t.ex. innehålla en sats av följande form.

(R) En handling är riktig om och endast om den har egenskap S.

Olika normativa teorier fyller i detta schema med olika egenskaper.³ En konkret instans av (R) är följande sats, som ingår i en vanlig version av klassisk utilitarism.

(U) En handling är riktig om och endast om den har minst lika bra konsekvenser som varje alternativ.

(U) tycks, liksom (R), tala om alla möjliga handlingar, inte bara om alla (faktiskt) existerande handlingar. *Allting* är sådant att om det är en riktig

²Med hjälp av symboler kan aktualismen (A) uttryckas på följande sätt: $\forall xEx$, där E är ett speciellt existenspredikat, ”Ex” läses ”x existerar”. Och denna sats är ett teorem i fri logik. (NA) kan symboliseras på följande sätt: $\Box \forall xEx$. $\Box \forall xEx$ är ett teorem i alla system som kombinerar fri logik med klassisk modallogik. I fri modallogik kan man dock inte härleda $\forall x\Box Ex$ från $\Box \forall xEx$, vilket man alltså kan göra i system som kombinerar klassisk predikatlogik med klassisk modallogik.

³Det tycks som om vi kan symbolisera (R) på följande sätt i predikatlogik: $\forall x(Hx \rightarrow (Rx \leftrightarrow Sx))$, där Hx, Rx och Sx är predikat som läses ”x är en handling”, ”x är riktig” respektive ”x har egenskap S”, där S står för olika egenskaper (monadiska eller relationella) i olika teorier. Sx kan t.ex. läsas: ”x gillas av mig (talaren/författaren) lika mycket som varje alternativ”, ”x är inte förbjuden av landets lagar”, ”x är förenlig med Guds befallningar”, ”x skulle accepteras av varje ideell observatör”, för att bara nämna några alternativ.

handling, så har det minst lika bra konsekvenser som varje alternativ; och *allting* är sådant att om det är en handling som har minst lika bra konsekvenser som varje alternativ, så är det en riktig handling.

Men om aktualismen är sann, är detta inte möjligt. Vi kan bara kvantifiera över existerande ting, över handlingar som utförs just nu och kanske över handlingar som har utförts tidigare, men inte över blott möjliga handlingar.

Antag att aktualismen är sann. Då gäller det att det är logiskt sant att endast existerande handlingar har normativa egenskaper. Om en handling är obligatorisk, riktig eller förbjuden (eller liknande) så existerar den. Vidare gäller det att om en handling inte existerar så är den obligatorisk och inte obligatorisk, riktig och inte riktig, förbjuden och inte förbjuden. Allt detta tycks stämma illa överens med hur vi tänker och resonerar kring handlingars normativa status.⁴

Detta argument emot aktualismen bygger på den implicita premissen att aktualisten accepterar någon form av fri logik. Men detta kan ifrågasättas. Det finns de som har menat att aktualismen är förenlig med klassisk predikatlogik (se t.ex. Linsky och Zalta 1994, 1996). Det gäller bara att tolka aktualismen på rätt sätt. Enligt denna tolkning är det sant att allting nödvändigtvis existerar. Men detta är oproblemiskt eftersom inte allting är konkret. Kontingenta ting, såsom levande organismer och mänskliga artefakter, existerar nödvändigtvis, men de är inte konkreta i varje möjlig värld: de är konkreta i vissa möjliga världar men icke-konkreta i vissa andra möjliga världar. En aktualist kunde hävda att detsamma gäller för handlingar.

Detta tycks ha bland annat följande konsekvenser. Redan innan Caesar mördades existerade Caesars mord (men det var inte konkret ännu). Redan innan du fattar ett beslut om vilken möjlig handling du skall utföra i en viss situation, så existerar alla (möjliga) handlingar du överväger, även om de inte är konkreta. När du väl beslutar dig för att utföra en handling och utför denna handling, så innebär det inte att denna handling går från att vara blott möjlig till att existera, den har existerat hela tiden, den går från att vara icke-konkret till att vara konkret.

Om detta är en rimlig position, kan även en moralfilosof acceptera aktualismen. Denna tolkning förefaller inte vara inkonsistent, men är samtidigt intuitivt besvärande. Tolkningen tycks t.ex. medföra att alla objekt som vi vanligtvis brukar betrakta som icke-existerande (och blott

⁴I fri logik är t.ex. följande satser logiskt sanna: $\forall x(Rx \rightarrow Ex)$, $\forall x(\neg Ex \rightarrow Rx)$, $\forall x(\neg Ex \rightarrow \neg Rx)$, där Rx tolkas som ovan och E , som vanligt, är ett existenspredikat. I possibilistiska system med klassisk logik är dessa satser inte logiskt sanna. I positiv fri logik kan vi förvisso tala om enskilda ting som inte existerar och tillskriva dessa egenskaper, men vi kan inte kvantifiera över icke-existerande handlingar.

möjliga), såsom Pegasus, Bilbo och Vulcan, (faktiskt) existerar, det är bara det att de inte är konkreta. Låt oss därför titta på ett annat möjligt försvar av aktualismen.

Ett möjligt aktualistiskt svar på det normativa argumentet går ut på att det inte är kvantifikatoruttrycket ”allting” som skall tolkas possibilistiskt utan att det är hela teorin (U) som i någon mening skall ses som nödvändig. (U) är inte en helt tillfredsställande tolkning av utilitarismen enligt denna tanke. Vi bör i stället använda följande formulering.

(U') Varje handling är nödvändigtvis riktig om och endast om den har minst lika bra konsekvenser som varje alternativ.

Denna formulering tycks i sin tur kunna tolkas på åtminstone tre intressanta sätt beroende på vilken räckvidd vi antar uttrycket ”nödvändigtvis” har. Dessa tolkningar kan formuleras på följande vis.

(U₁) Det är nödvändigt att en handling är riktig om och endast om den har minst lika bra konsekvenser som varje alternativ.

(1) $\Box \forall x (Hx \rightarrow (Rx \leftrightarrow Mx))$.⁵

(U₂) Allting är sådant att det är nödvändigt att om det är en handling, så är det riktigt om och endast om det har minst lika bra konsekvenser som varje alternativ. (2) $\forall x \Box (Hx \rightarrow (Rx \leftrightarrow Mx))$.

(U₃) Allting är sådant att om det är en handling, så är det nödvändigt att det är riktigt om och endast om det har minst lika bra konsekvenser som varje alternativ. (3) $\forall x (Hx \rightarrow \Box (Rx \leftrightarrow Mx))$.

(U₂) och (U₃) är dock problematiska, eftersom vi enligt dessa formuleringar fortfarande endast kvantifierar över existerande ting (och existerande handlingar). (U₃) säger t.ex. att varje handling som (faktiskt) existerar (i vår värld) i alla möjliga världar är riktig om och endast om den har minst lika bra konsekvenser som varje alternativ. Och detta tycks inte vara vad en utilitarist vill hävda. Liknande problem gäller för (U₂). Dessa tolkningar tycks därför inte kunna användas som ett svar på vårt argument. (U₁) däremot säger att det i alla möjliga världar gäller att varje handling som existerar (där) är riktig om och endast om den har minst lika bra konsekvenser som varje alternativ. Och en aktualist tänker sig vanligtvis att olika ting kan existera i olika möjliga världar. Andra möjliga världar kan därför innehålla handlingar som inte existerar i vår värld. Detta är fullt förenligt med att kvantifikatoruttrycket ”allting” tolkas aktualistiskt. Så, en aktualist kunde påstå att en klassisk utilitarist som hävdar (U), egentligen menar (U₁), dvs. att det är nödvändigt att en handling är riktig om

⁵Hx och Rx tolkas som tidigare och Mx läses som ”x har minst lika bra konsekvenser som varje alternativ”.

och endast om den har minst lika bra konsekvenser som varje alternativ.

En possibilist tycks kunna ge åtminstone två möjliga svar på denna invändning. För det första kan hon hävda att (U) och (U₁) inte säger samma sak och att det är (U) vi är intresserade av. För det andra kan hon ifrågasätta att vi bör använda kvantifierad modallogik för att analysera (U) och liknande teorier; det räcker med predikatlogik, kunde hon hävda. Och i så fall är detta argument fortfarande ett problem för aktualisten.

3. SLUTSATS

Jag har i den här uppsatsen diskuterat ett argument som talar för att man med tanke på moraliska riktighetskriterier bör acceptera possibilismen. När vi talar om mängden av allting är det rimligt att anta att det ingår blott möjliga objekt i denna mängd och inte enbart existerande ting. Det argument jag har diskuterat är dock inte konklusivt och vi har nämnt några möjliga sätt en aktualist kan tänkas bemöta det. Men vi har också sett att dessa motargument inte är helt oproblematiska. Detta ger oss skäl att ifrågasätta aktualismen och ta possibilismen på allvar. För att avgöra om aktualismen eller possibilismen är sann, måste vi dock sannolikt ta ställning också till flera andra argument för och emot dessa positioner. Vad som förefaller vara uppenbart är emellertid att många etiker implicit antar att vi kan kvantifiera över blott möjliga handlingar. Och det verkar som om de flesta har ansett att detta är oproblematiskt. Men om aktualismen är sann, är detta dock inte möjligt. Vi kan då endast kvantifiera över existerande ting. Och vi måste då använda oss av något annat logiskt system än klassisk predikatlogik om vi vill förstå den logiska formen hos klassiska normativa teorier såsom utilitarismen, eller acceptera att allting nödvändigtvis existerar. Debatten mellan aktualister och possibilister är därmed relevant för moralfilosofer, oavsett vilken position som i slutändan är den rimligaste.

LITTERATUR

- Adams, R. M. 1981. "Actualism and Thisness". *Synthese* 49, s. 3–41.
- Fitch, G. W. 1996. "In Defense of Aristotelian Actualism". *Noûs*, Vol. 30, Supplement: Philosophical Perspectives, 10, Metaphysics, s. 53–71.
- Linsky, B. och E. N. Zalta. 1994. "In Defense of the Simplest Quantified Modal Logic". *Philosophical Perspectives*, Vol. 8, Logic and Language, s. 431–58.
- Linsky, B. och E. N. Zalta. 1996. "In Defense of the Contingently Nonconcrete". *Philosophical Studies* 84, s. 283–94.
- Menzel, C. 1991. "The True Modal Logic". *Journal of Philosophical Logic* 20, s. 331–74.
- Plantinga, A. 1974. *The Nature of Necessity*. Oxford: Oxford University Press.
- Williamson, T. 1998. "Bare Possibilia". *Erkenntnis* 48, s. 257–73.

Recensioner

Tankar som ändrar allt: Om tankeexperiment och nya världsbilder

Sören Holst

Fri Tanke, 2012. 332 s. ISBN 978-91-86061-51-7

Fysikern Sören Holst gör i denna bok två saker. Dels ger han en populärvetenskaplig introduktion till några särskilt spännande tankespår inom fysik, medvetandefilosofi och moralfilosofi. Dels ger han en partiell introduktion till ett av dessa områdens viktigaste verktyg, nämligen tankeexperiment. Här kommer han också med egna förslag om deras natur och funktionssätt. I denna anmeldelse kommer jag främst att uppehålla mig vid bokens bidrag till pågående debatter om tankeexperiment i allmänhet. Det bör dock framhållas att dessa upptar endast en liten del av innehållet i boken och inte är dess huvudsyfte.

Främst är Holsts välskrivna och lättlästa bok en guidad tur genom en serie spännande hypotetiska fall som varit inflytelserika i olika teoretiska debatter. Men ärendet är i första hand att förmedla de insikter om rum och tid, personlig identitet och moral för vilka exemplen bara varit verktyg. Hur verktygen fungerar och vad de har gemensamt är, som jag läser boken, sekundärt, även om det är i dessa frågor boken bidrar med nya förslag.

De teoretiska ståndpunkter boken presenterar, utan att förutsätta bakgrundskunskaper, är i drastisk sammanfattning följande. Einsteins speciella relativitetsteori, Parfits teori om personlig identitet (och densnas mindre vikt jämförd med psykologisk kontinuitet), samt Ungers och Singers utilitarism. Utöver ett flertal av Einsteins och Singers mest kända exempel, teletransportationer, personsfissioner och trolleyfall får läsaren också en pedagogisk redogörelse för några kända tankeexperiment av Galileo, Newton, Stevin och Huygens (från fysiken) samt av Jackson och Unger (särskilt den senares variationer av Foots och Thomsons trolleyfall). Många av de filosofiska exemplen är nog välkända för de flesta som studerat ämnet; andra läsare får en ytterst klar och i stort sett korrekt uppfattning både om exemplen och om delar av debatten runt dem (även

om vissa detaljer kanske visar att författaren inte själv är fackfilosof). Exempelen från fysiken är nog mindre kända bland de flesta läsare som inte råkar vara fysikhistoriker. Särskilt Holsts förklaring av Einsteins exempel och bakgrunden till dem är i mitt tycke en pedagogisk bragd: läsaren får – i behagliga doser och hart när utan tuggmotstånd – en klar idé om den speciella relativitetsteorins oundviklighet givet två enkla postulat hos Einstein, liksom om teorins nära samband med äldre teorier om elektromagnetism. Som populärvetenskap är boken i långa stycken mycket framgångsrik.

Boken har två delar. Den första sägs (i förordet) handla om tankeexperiment mer allmänt, medan den andra är tänkt att med tankeexperimenters hjälp förmedla de ”nya världsbilder” som omnämns i titeln. Med ”nya” menar Holst här väl helt enkelt att de tankar som förmedlas är nya jämfört med de common sense-föreställningar som en av filosofi och relativitetsteori orörd läsare kan tänkas hysa. Bokens faktiska disposition motsvarar emellertid inte riktigt denna avgränsning. Ett antal generella teser om tankeexperiment framförs här och var boken igenom, och Holsts egentliga syn på hur de fungerar sparas till ett avslutningskapitel. Därtill förekommer ett stort antal kortare redogörelser, både av specifika hypotetiska fall och av empiriska rön, på ett flertal ställen i båda delarna av boken. Detta ger boken en något spretig karaktär för läsare som hungrar efter systematik; för tilltänkta lekmanläsare med allmänt god intellektuell aptit gör det antagligen mindre.

Vad är ett tankeexperiment? Holst säger inledningsvis (s. 15) att begreppet ”fångas förmodligen enklast genom exempel”, varpå han raskt ger just två berömda exempel från fysiken: Galileos om fallande kroppar och Simon Stevins om krafter på lutande plan. I följande kapitel – med titeln ”Vad är ett tankeexperiment?” – ger Holst dock två explicita definitioner. Den första definitionen (s. 29) säger att ett tankeexperiment är ”(1) en hypotetisk aktivitet, (2) i vilken en konkret situation målas upp, (3) som leder till en insikt eller ett resultat av något slag”. Holst säger: ”Detta ringar in begreppet ganska väl” (s. 29), men han noterar att definitionens är för vitt. En utförligare version ställer hårdare krav på vad som görs med den ”konkreta” hypotetiska situationen (dvs. scenariot) och vad ”resultatet” skall vara, men släpper, vilket kanske är värt att lägga märke till, kravet på ”insikt”. Denna nya definition (s. 32) säger att ett tankeexperiment (1) är av hypotetisk karaktär; (2) kretsar kring en konkret situation om vilken en specifik slutsats nås via ett ”mer eller mindre uttalat” resonemang; (3) samt leder till ett resultat som är mer generellt än den specifika slutsatsen om scenariot. Liksom i den föregående definitionen är det rimligt att tolka de tre villkoren som tillsammans tillräckliga och individuellt nödvändiga. Den nya definitionen presenteras dock i mot-

sats till den första som en stipulation: ”Det här är vad begreppet tankeexperiment i denna bok syftar på” (s. 32). I mitt tycke är det klokt att i en text som gör centralt bruk av termen ”tankeexperiment” stipulera vad termen skall betyda. Dess faktiska bruk är för spretigt och skiftande för att tillåta vare sig rent dekriptiva definitioner eller ostension via exempel. Det förefaller också som om Holsts stipulation – trots att den inte är helt precis – är hyfsat deskriptivt adekvat för många fall.

Hur kan tankeexperiment ge ”resultat” eller ny kunskap när de inte bygger på ny empirisk information? Detta har varit en central fråga i debatten särskilt om naturvetenskapliga tankeexperiment. Holst kallar den ”informationsproblemet” (s. 149). Holst ansluter sig (s. 153) väsentligen till John Nortons nyktra tes: tankeexperiment är helt enkelt en sorts argument och ger ”ny” information i samma mening som resonemang i allmänhet kan leda till ”nya” slutsatser (Norton 1991). Samtidigt förespråkar han en version av Roy Sorensens (1992) idé om vilka slags argument vi har att göra med i samband med tankeexperiment. Sorensen lanserade en paradoxmodell: tankeexperiment involverar mängder av tillsammans inkonsistenta men inviduellt (åtminstone prima facie) rimliga påståenden. I Sorensens modell hade dessa mängder fem element; i Holsts förenklade modell (s. 68–69) innehåller de fyra: (i) den testade teorin A; (ii) påståendet att scenariot E leder till slutsatsen S, givet att A gäller; (iii) påståendet att S är omöjlig (eller åtminstone ”oinuitiv eller förnuftsvidrig”, s. 70); samt (iv) påståendet att E är möjligt.

Påstående (ii) är, säger Holst, ”kärnan i tankeexperimentet” (s. 69). Den verkar vara en villkorssats. Men om påståendemängden skall vara inkonsistent kan (ii) inte vara en materiell implikation, t.ex. $(A \& E) \supset S$. Men vilket slags villkorssats är det? Holst säger att ”påståendet utgörs av den logiska slutledningen som tar oss från situationen E och antagandena A till slutsatsen S” – eftersom det uttryckligen rör sig om *ett* påstående, och slutledningen sägs vara ”logisk”, kanske (ii) är tänkt att vara en strikt implikation. Men i många av de fall Holst själv tar upp verkar ”slutledningen” till den specifika slutsats S som dras om den specifika situationen E både fordra och tillåta import av fler premisser än enbart A och E, så det verkar *antingen* svårt att förstå (ii) som en strikt implikation *eller* oklart vad antecedenten skulle vara (liknande svårigheter möter Jonathan Ichikawas och Benjamin Jarvis’ förslag att omdömen eller intuitioner om hypotetiska fall är strikta implikationer; jfr Ichikawa och Jarvis (2009), Williamson (2009) och Malmgren (2011)). Att Holst inte ger några logiska detaljer här är dock förståeligt – de läsare som i likhet med mig är frustrerade över detta torde utgöra en minoritet.

Paradoxmodellen ger en grov taxonomi. Tankeexperiment kan delas in i fyra sorter efter vilken medlem av kvartetten de syftar till att

förkasta. Den ”vanligaste typen” (s. 70) syftar till att förkasta (i), dvs. vederlägga en teori. Vissa ”mindre vanliga” (s. 71) syftar i stället till att förkasta (ii) och därmed till att ogiltigförklara den slutledning som säger att S följer av teori och scenario. De tolkningproblem jag hade med (ii) gäller automatiskt även vad denna typ av tankeexperiment egentligen syftar till (och förenas av). Holsts illustration är tvillingparadoxen för speciell relativitetsteori (s. 71 – fallet beskrivs utförligt i kapitel 10), där den specifika slutsatsen *inte* blir att två tvillingar, åtskilda genom den enas långa rymdresa, båda är äldre än varandra. En tredje typ syftar till att förkasta (iii) och alltså visa att omöjligheten hos S är skenbar medan en sista typ förkastar (iv) och omvänt hävdar att scenariot E:s möjlighet bara är skenbar.

Att vissa tankeexperiment har de respektive syften som utmärker de sista två typerna är helt säkert riktigt. Men det är inte klart hur de passar in med Holsts stipulativa definition. Att slutsatsen S om det specifika scenariot är möjlig verkar inte mer generellt än S själv, vilket ju var villkor (3) i Holsts definition; detsamma gäller de exempel som syftar till att påvisa ett partikulärt scenarios omöjlighet. Detta verkar också bara vara en högst specifik slutsats. Kanske ett mer ”allmängiltigt resultat” kan uppnås om en partikulär möjlighet eller omöjlighet tas som bas för en induktion till mer svepande påståenden. Men ett sådant steg ligger utanför paradoxmodellen och innebär att den överges. Det är tyvärr heller ingenting Holst kommenterar, såvitt jag kunnat se.

Holsts mest ambitiösa bidrag till en allmän teori om hur tankeexperiment fungerar presenteras i avslutningskapitlet ”Goda teoriers sällsynthet”. Han föreslår här att tankeexperiment kanske inte av egen kraft kan finna korrekta teorier – för detta ”krävs alltid ett visst mått av empiri” (s. 316) – men att deras uppgift är att ”gallra bland” presumtiva teorier (s. 316). Gallringen består i att tankeexperiment sällar bort teorier som inte är konsistenta, allmängiltiga och enkla. Teorier som däremot har dessa egenskaper kallar Holst ”goda teorier” (s. 301) och dessa ”utgör kandidater till korrekta teorier” (s. 302). Skillnaden mellan teorier som är goda och sådana som dessutom är korrekta tycks alltså vara att de senare förutom att vara goda också är empiriskt adekvata. En central tes är även blott goda teorier är förvånansvärt sällsynta och att vi kan upptäcka mindre goda teoriers brister med hjälp av tankeexperiment.

Ett problem med detta förslag är att förstå precis vad det går ut på. Om de tre nyckelbegreppen konsistens, allmängiltighet och enkelhet tolkas bokstavligt eller i enlighet med gängse filosofiskt bruk verkar tesen vara falsk: det är inte ont om konsistenta, generella och enkla teorier. Tolkat strikt verkar konsistenskravet dessutom göra paradoxmodellen överflödig – är en teori inkonsistent kan väl detta upptäckas utan hjälp

av påhittade specifika scenarier. Med ”inkonsistens” menar Holst dock något mycket svagare, nämligen att teorin är ”motsägelsefull, antingen i sig själv eller tillsammans med andra försanthållanden” (s. 301). Då är förslaget förenligt med åtminstone den första typen av fall enligt paradoxmodellen: den där teorin förkastas därför att den implicerar något omöjligt eller absurt i ett specifikt scenario. Men detta väcker genast frågan om hur vi får tillgång till de modala försanthållanden som påståendena (ii), (iii) och (iv) i modellen verkar röra, och hur viktig inkonsistens med dem bör vara.

I ”allmängiltighet” i Holsts mening ingår inte empirisk (eller annan) adekvans. Å andra sidan verkar Holst mena att allmängiltighet delvis är modal. Åtminstone bör teorin täcka även de hypotetiska scenarier som används i tankeexperiment. Men hur stark en sådan modal generalitet bör vara hos teorier i allmänhet (eller kanske rimligare, vilken modal styrka som är rimlig att begära av olika slags teorier) klagörs tyvärr inte. ”Enkelhet”, slutligen, utläggs som kravet att en teori ”ska följa entydigt och logiskt av ett litet antal påståenden” (s. 301). Men detta verkar vara att kräva status som teorem. Varför en teori bara kan vara god om den har sådan status är svårt att se.

En allmän fråga är vad som kan motsvara empirisk adekvans hos filosofiska teorier. Det verkar inte som om ens ett stort mått av empiri hjälper för att skilja korrekta teorier inom filosofin från inkorrekta. Att Holst inte tar upp denna (svepande) fråga är helt i sin ordning. Men eftersom hans paradoxmodell, och en välvillig tolkning av hans inkonsistensförslag, inbegriper omdömen om hypotetiska situationer och gör konsistens med dem viktig, är frågan om modal kunskap viktig även för hans förslag.

Särskilt i filosofiska fall tänker man sig ofta att intuition ligger till grund för modal kunskap. Holst säger dock förvånansvärt litet om intuition, trots att partier av boken tar upp psykologisk forskning kring hur konkretion hos scenarier kan vara viktiga heuristiskt och ”*aktivera själva förmågan att tänka logiskt*” (s. 164, kursiv i originalet). Här hänvisar Holst till Driscolls teori om begreppsscheman och föreslår att de ”outtalade antaganden som med nödvändighet finns inbyggda i varje schema” (s. 164) också kan utgöra en felkälla vid tankeexperiment. Utöver detta hänvisar Holst i kapitel 13 och 14 till Peter Ungers hypoteser om grupptänkande och Joshua Greenes hypotes om emotionsaktivering som möjliga psykologiska förklaringar till vissa intuitioner – Holst kallar dem ibland ”instinkter” – om bland annat trolleyfall. Psykologins roll blir här att ge förklaringar av varför vissa har falska intuitioner. Här kan man efterlysa en diskussion om vad som positivt kan anföras för att intuitioner ibland är legitim evidens – alternativt varför tankeexperiment inte behöver dem.

Jag har kritiskt diskuterat några centrala punkter i Holsts allmänna förslag om tankeexperiment. Avslutningsvis vill jag upprepa två saker. Först: boken är inte fackfilosofisk (även om jag här behandlat den så), utan vill först och främst presentera stimulerande och viktiga tankar för nyfikna allmänläsare. För det andra: Holst lyckas mycket väl med detta syfte.

SÖREN HÄGGQVIST

REFERENSER

- Ichikawa, J. och B. Jarvis. 2009. "Thought-Experiment Intuitions and Truth in Fiction". *Philosophical Studies* 142, s. 221–46.
- Malmgren, A-S. 2011. "Rationalism and the Content of Intuitive Judgements". *Mind* 120, s. 265–327.
- Norton, J. 1991. "Thought Experiments in Einstein's Work". I *Thought Experiments in Science and Philosophy*, red. T. Horowitz och G. Massey, s. 129–48. Savage, Md: Rowman & Littlefield.
- Sorensen, R. 1992. *Thought Experiments*. Oxford: Oxford University Press.
- Williamson, T. 2009. "Replies to Ichikawa, Martin and Weinberg". *Philosophical Studies* 145, s. 431–34.

Dygdernas renässans

Red. Eva Österberg, Marie Lindstedt Cronberg och Catharina Stenqvist
Atlantis, 2012. 284 s. ISBN 978-91-7353-596-0

1. *Dygdernas renässans* är en nyutkommen antologi bestående av elva uppsatser, samt en inledande introduktion skriven av redaktörerna. Bland uppsatsförfattarna finns representanter från flera olika – om än tämligen nära angränsande – ämnesområden: historia, idé- och lärdomshistoria, kyrkohistoria, religionsfilosofi och sociologi.

Som titeln antyder, så är det genomgående temat för antologin *dygder*. Åtminstone inom den västerländska moralfilosofin, så var intresset för dygder – liksom för deras motsatser, *lasterna* – under en period, säg från början av 1800-talet fram till 1960- eller 70-talet, tämligen begränsat. Detta har emellertid förändrats. Under de senaste fyra eller fem årtiondena har en allt större skara moralfilosofer kommit att ägna sig åt studier av dygder och laster. Startskottet för detta växande intresse brukar ofta sägas ha varit publiceringen av Elizabeth Anscombes uppsats "Modern Moral Philosophy" i tidskriften *Philosophy*, 1959. Tillståndet inom moralfilosofin kan i och med denna förändring sägas ha återgått åtminstone delvis till det historiskt sett normala. Under hela den period som i antolo-

gin kallas för ”den förmoderna”, vilken inkluderar ”antiken, medeltiden, 1500- och 1600-talen” (s. 9), så är det högst påtagligt att just dygder och laster spelar en central roll i det moralfilosofiska tänkandet. Detsamma kan med viss rätt påstås gälla även för flera av de skotska upplysningsfilosoferna – till exempel David Hume och Adam Smith – under 1700-talet, liksom (om än i något mindre utsträckning) för Kant.

Det här nyvaknade intresset för dygder och laster, vilket märks inte bara inom moralfilosofin utan även på många andra håll både innanför och utanför den akademiska världens ramar, är också vad redaktörerna för *Dygdernas renässans* knyter an till i sin introduktion: ”Hur skall vi . . . förstå detta på ett djupare plan”, frågar de sig; ”*Vari ligger dygdebegreppets förnyade attraktionskraft?* Det är ett spørsmål som vi vill väcka tankar kring” (s. 9; kursiveringar i originalet). Lite annorlunda uttryckt: Hur kommer det sig att så många forskare idag intresserar sig för frågor om dygdernas inneboende natur – om vad dygder egentligen är för någonting, hur vi utvecklar dem, och i vilken utsträckning dygder kan tänkas vara nödvändiga för att leva bra eller goda liv? Att ”väcka tankar kring” detta är vad redaktörerna alltså hoppas att antologin skall bidra till. Huruvida antologin uppfyller denna förhoppning är något som jag skall återkomma till nedan.

2. Bo Lindberg skriver i inledningen till sitt bidrag att han ”[i] det följande skall . . . beskriva mildhetens roll i den förmoderna moraldiskussionen och sedan något diskutera dess relevans i nutiden” (s. 66). Bortsett från hänvisningen till just *mildhetens* roll och relevans, så ger Lindberg här uttryck för en ansats som stämmer rätt väl in på det stora flertalet av uppsatserna i *Dygdernas renässans*. Författarna har valt ut var sin dygd – eller i alla fall någonting som vid en eller annan tidpunkt uppfattats som en dygd – vars utveckling de försöker följa under framför allt den förmoderna historien genom att göra olika nedslag i densamma – genom att titta på hur några olika tänkare eller traditioner har föreställt sig vad det innebär att äga och utöva dygden ifråga – för att därefter säga åtminstone någonting också om vilken betydelse den relevanta dygden kan tänkas ha i vår egen tid. I några fall har författarna valt dygder som de i slutändan betvivlar att vi behöver idag. Detta antyds gälla för till exempel mildhet och *lydnad* (vilket är ämnet för tidigare ärkebiskopen KG Hammars bidrag). I de allra flesta fall, så har författarna emellertid valt att rikta sin uppmärksamhet mot dygder som de kommer fram till nog har viss eller kanske till och med stor relevans även i vår egen tid; till exempel *mod*, *lojalitet*, *tålmod*, *barmhärtighet* och *ödmjukhet*.¹

¹Dessa dygder behandlas i tur och ordning av Eva Österberg, Eva Kärfve, Kenneth Johansson, Mohammad Fazlhashemi och Catharina Stenvik.

Som vi skall se i avsnitt 4, så finns det anledning att vara kritisk till bidragens starkt historiska prägel. Men om vi för stunden lägger detta åt sidan, så kan det noteras att flera av författarna lyckas förhållandevis väl, givet det begränsade utrymme de har till sitt förfogande, med att lyfta fram vissa till synes centrala aspekter av olika föreställningar som omhuldas under historiens lopp angående vad det lite mer konkret kan tänkas innebära att äga och utöva olika dygder, liksom med att på ett begripligt sätt åtminstone antyda något om de olika dygdernas eventuella relevans för oss idag. Ordentligt på tok i dessa avseenden går det egentligen bara i Sven-Eric Liedmans fall. Liedman, vars bidrag handlar om (praktisk) *klokhet*, förmår helt enkelt inte begränsa sig i den utsträckning som formatet kräver. På 16 sidor försöker han sig på att först säga något "om ordet 'klok' som sådant och om dess motsvarigheter på en rad andra språk" (s. 28), och därefter att skildra själva begreppets västerländska historia – en skildring i vilken Platon, Aristoteles, Thomas ab Aquino, Niccolo Machiavelli, Thomas Hobbes, David Hume, Immanuel Kant, Henri de Saint-Simon, Auguste Comte, Karl Marx, John Stuart Mill, Karl Popper, Alasdair MacIntyre, Martha Nussbaum, Jürgen Habermas och Paul Ricœur alla svyschar förbi i en svindlande hastighet – för att till sist landa i slutsatsen att man kan "hävda att social kompetens och medborgarkompetens tillsammans utgör ett nutida, [genom EU-rekommendationer] politiskt sanktionerat begrepp om det som tidigare kallades klokhet" (s. 42). Eftersom i stort sett ingenting substantiellt kommer fram i den snabba genomgången av alla de filosofer som räknas upp ovan, och inte heller i den väldigt kortfattade beskrivningen av vad de här EU-idealen om social kompetens och medborgarkompetens egentligen går ut på, så blir slutsatsen att de senare "tillsammans utgör ett nutida ... begrepp om det som tidigare kallades klokhet" omöjlig att riktigt förstå eller ta ställning till i ljuset av Liedmans uppsats.

3. Men lyckas antologin med att "väcka tankar kring" frågan om "*Vari ... dygdebegreppets förnyade attraktionskraft*" ligger? I viss utsträckning, i alla fall, så gör den väl det. I så måtto som författarna har rätt i att det finns ett flertal traditionella dygder som är av betydelse även för oss idag, så framstår ju till exempel det samtida intresset för dygdebegreppet – för att nå en djupare förståelse av dygdernas inneboende natur – som fullt begripligt och viktigt. Vidare väcker väl de bidrag som är ägnade åt dygder, vars nutida relevans tycks mer tvivelaktig, tankar om att vi, givet att vi skulle börja lägga större vikt i samhället vid utvecklandet av dygder, trots allt måste vara noga med att inte uppmuntra utvecklandet av precis allting som någon gång under historien räknats som en dygd;

vissa av dessa traditionella eller historiska dygder förutsätter föreställningar om tillvarons beskaffenhet som många av oss inte längre accepterar – föreställningar om klass- och könhierarkier, till exempel, eller om existensen av en Gud som kräver blind åttlydnad. Detta har förvisso inte på något direkt sätt att göra med frågan om ”*Vari dygdebegreppets ... förnyade attraktionskraft*” ligger, men är åtminstone nära relaterat till det förnyade intresset för dygder och laster.

4. Jag har emellertid flera reservationer mot *Dygdernas renässans*. Här skall jag begränsa mig till att ta upp två av dessa.

Till att börja med, så ger antologin lätt upphov till ett intryck hos läsaren av att det förvisso finns ett nyvaknat intresse bland forskare för dygdebegreppet, men att ingen riktigt har brytt sig om att fundera på om det finns något egentligt behov idag av de saker som traditionellt har uppfattats som dygder (rättvisa, mod, måttfullhet, klokhet, och så vidare); det vill säga de saker vilka man förväntar sig skall falla in under ett rimligt dygdebegrepp. Men detta är ju fullständigt felaktigt. Det finns numer en oerhört omfattande litteratur om såväl dygdeteori i allmänhet som olika mer specifika dygder – inklusive de flesta som tas upp i *Dygdernas renässans* – och i denna litteratur finns exempel både på författare som anlägger ett framför allt historiskt perspektiv och på sådana som koncentrerar sig framför allt på nutiden, liksom en stor grupp som försöker få med båda dessa perspektiv. Denna litteratur är dock i det närmaste helt frånvarande i antologin. Om detta beror på att man inte haft möjlighet att sätta sig in i den, eller att man fann den så bristfällig att det inte var mödan värt att diskutera, kritisera eller ens hänvisa till den, eller på någonting annat, kan jag inte svara på. Men det sätter hela projektet i en lite märklig dager.

Givet vad man enligt redaktörerna vill åstadkomma med antologin, så är det vidare någonting väldigt olyckligt över det stora utrymme som i de flesta bidrag ägnas åt att redogöra för vad det under olika perioder av den förmoderna historien mer konkret eller substantiellt har ansetts innebära att utöva olika dygder. Detta är olyckligt eftersom det naturligtvis inte finns någon som idag intresserar för dygdeteori som på fullt allvar tänker sig att det skulle finnas något samtida behov av att utöva till exempel mod, måttfullhet och rättvisa *på exakt de sätt* som det kanske fanns under en viss period av antiken, medeltiden eller någon gång på 1500- och 1600-talen. Att försöka väcka tankar kring det förnyade intresset för dygdebegreppet genom den typen av historiska genomgångar förefaller väldigt långsökt jämfört med att på ett ordentligt sätt lyfta fram exempel på områden i vår samtid där begrepp som mod (liksom feghet och dumdrighet), måttfullhet (måttlöshet, girig-

het), rättvisa (orättvisa), och så vidare, verkligt används eller framstår som tydligt relevanta. Att komma på sådana exempel tar inte lång tid. Tänk bara på sådana saker som den feighet – ofta i kombination med måttlöshet och grymhet – som utmärker det anonyma näthatet, och på modet hos många av dem som står upp emot det; på den girighet och orättvisa som utmärker företagsledare och politiker som ger sig själva bonusar och/eller pensionsvillkor som vi andra bara kan drömma om; på den grad av måttfullhet som krävs för att undvika ohälsa, att dra på sig stora skulder, med mera; eller för den delen på hur våra vänskapsrelationer förändras – både på gott och ont, efter vad det verkar – i takt med vår ökade användning av olika sociala medier. I ljuset av sådana exempel framstår det knappast som särskilt märkvärdigt att det också finns ett ökat intresse för den inneboende naturen hos dessa egenskaper. Givetvis kan man söka sig bakåt i historien för att se om man där kan finna viktiga komponenter av en rimlig dygdeteori, det vill säga komponenter av en rimlig redogörelse för just dygdernas inneboende natur. Detta är också vad väldigt många dygdeteoretiker har gjort – det finns ju en utbredd debatt mellan dem som menar att vi kan finna flera avgörande komponenter redan hos Aristoteles och dem som i stället försvarar en dygdeteori som mer påminner om vad vi finner hos till exempel Hume, hos Kant, eller hos Nietzsche (en debatt av vilken man för övrigt inte märker ett spår i antologin). Men, återigen, vad vi inte finner är försök att hävda att vi idag med hull och hår skulle ta över någon av de mer konkreta eller substantiella föreställningar om vad det innebär att utöva olika dygder som dessa tänkare försvarade.

5. På grund av dessa, enligt min mening tämligen allvarliga, reservationer (och även på grund av andra, mindre reservationer som jag har utelämnat här), så har jag svårt att känna den entusiasm över *Dygdernas renässans* som jag egentligen skulle vilja göra. En bra antologi om dygder och laster på svenska vore väldigt välkommen. Tyvärr får vi dock vänta ytterligare en tid på detta.²

FRANS SVENSSON

²Jag vill tacka Jens Johansson och Hans Mathlein för att de tog sig tid att läsa och kommentera ett tidigare utkast till denna recension. De bidrog båda med flera hjälpsamma synpunkter.

Kort om spelteori

Ken Binmore. Översättning: Roland Poirier Martinsson
 Fri Tanke, 2012, 234 s. ISBN 978-91-86061-39-5

Spelteori är av intresse inom flera akademiska discipliner – t.ex. matematik, ekonomi, statskunskap, filosofi och biologi – och kanske också i andra sammanhang, t.ex. för politiker och pokerspelare. Flera av de personer som fått Ekonomipriset till Alfred Nobels minne har också ägnat sig åt spelteori, t.ex. Alvin Roth och Lloyd Shapley 2012, Robert Aumann och Thomas Schelling 2005 och John Nash, John Harsanyi och Richard Selten 1994. De ”spel” det är fråga om är interaktion i största allmänhet; enligt Ken Binmore spelar människor ”ett spel varje gång de interagerar med varandra” (s. 5), men han säger också att spelteorin ”bara fungerar när människor spelar *rationellt*” (s. 6), så man kan ju undra hur ofta det sker. Vidare kan tydligen inte bara människor bete sig ”rationellt”; spelteorin kan också förklara beteendet hos spindlar och fiskar, ”eftersom de konkurrenter vars gener programmerades för irrationellt beteende nu är utdöda” (s. 6).

Binmore har genomgående en rätt självbelåten attityd och han skriver inte alltid så pedagogiskt. Ibland är det svårt att förstå vad han menar, t.ex. när han säger att samarbete och konflikt ”är två sidor av samma mynt” (s. 11) – men kanske är detta något som beror på översättningen. Flera av tabellerna och diagrammen verkar felaktiga; jag vet inte om det är likadant i det engelska originalet eller om det har tillkommit i den svenska utgåvan. Det hänvisas ofta till olika spel och det hade därför varit bra om man kunde hitta dem med hjälp av index. Men det går inte alltid. Om man t.ex. slår upp ett spel som kallas ”kampen mellan könen” i index, så hänvisas man till sidan 34, men där beskrivs inte spelet (även om det omnämns), utan man får själv leta sig tillbaka till sidorna 18–19 för att få reda på hur det såg ut. Där är emellertid siffrorna i tabellen felplacerade, såvitt jag förstår; övre vänstra och nedre högra rutan borde byta plats. Och för filosofen John Rawls anger index sidan 197, men Rawls kritiserar redan på sidan 16, vilket index inte noterar.

Vad är spelteori för något? Dess syfte tycks enligt Binmore vara att förutsäga personers agerande i spelsituationer med hjälp av observationer av deras tidigare beteenden i olika valsituationer – eller, alternativt, att påstå att deras beteende är ”motsägelsefullt” i den mån det inte skulle överensstämma med förutsägelserna (jfr s. 12–13). Vi kan illustrera detta med hans diskussion av det kanske mest kända av alla de spel som förekommer i den spelteoretiska litteraturen, nämligen *fångarnas dilemma*.

I sin ursprungliga version handlade detta om två fångar som var för sig kan välja mellan att erkänna eller inte erkänna ett brott. Men Binmore

diskuterar också en lite enklare variant och jag ska här hålla mig till den. Även här har två spelare att var för sig välja mellan två alternativ. Var och en kan antingen *ta* 10 kr till sig själv ur en pott med pengar eller *ge* 20 kr ur potten till den andre. Om bägge är givmilda får de alltså 20 kr var, om bägge är själviska får de 10 kr var, om den ene är givmild och den andre självisk, så får den givmilde ingenting och den själviske 30 kr.

	B ger	B tar
A ger	A får 20 kr – B får 20 kr	A får inget – B får 30 kr
A tar	A får 30 kr – B får inget	A får 10 kr – B får 10 kr

Spelteorin anses då säga att det är rationellt för bägge att ta 10 kronor till sig själva (under förutsättning att de bara bryr sig om pengar). ”Dilemmat” består i att de bara får 10 kr var på detta sätt, trots att de ju kunde ha fått 20 kr var, om bägge i stället hade varit givmilda. Men det är likafullt sant att varje spelare får mer pengar om hen betar sig själviskt, oavsett vad den andre gör.

Binmore anser att många har missförstått fångarnas dilemma. Som exempel på ett sådant ”förelslut” skriver han: ”I vanligt språk säger Immanuel Kants kategoriska imperativ att det är rationellt att göra det du vill att alla andra ska göra. Om det var sant så skulle det vara rationellt att samarbeta [dvs. att vara givmild] i fångarnas dilemma. Men önsketänkande är aldrig rationellt. Det är en konstant källa till förvåning för mig att Kant aldrig ställs till svars för att ha föreslagit en rationalitetsprincip utan att ge några som helst skäl för varför vi bör ta den på allvar” (s. 205).

Detta är ganska typiskt för Binmores sätt att skriva. Det är tydligt att han ogillar moralfilosofer; han berömmar sig t.ex. av att ha retat gallfeber på ”renläriga moralfilosofer” genom att använda spelteori när han diskuterar etik (s. 8) och han påstår, utan argument, att moralfilosofen John Rawls ”försvarade en sund intuition med fel argument” (s. 197). Men hur Binmore kan vara så säker på att Kant ”aldrig ställs till svars” är ett mysterium; har han verkligen läst hela den minst sagt omfattande Kant-litteraturen? Dessutom kan man ju undra om det Kant föreslår verkligen är en ”rationalitetsprincip” i Binmores mening, snarare än t.ex. en *moralisk* princip. Att moralen ibland kan kräva något *annat* av oss än att göra det vi rent personligen råkar föredra, är ju en föreställning som inte är helt okänd i mänsklighetens historia. Men Binmore ser den nog som ett ”förelslut”.

Binmore anser alltså att Kants morallära är oförenlig med spelteorin. Men det kan knappast stämma. Binmore påpekar själv att att spelteorin

”inte har något substantiellt innehåll” (s. 74). Den har inte ”mer egentligt innehåll än aritmetik eller logik” (s. 82). Det har han nog rätt i; spelteorin tycks bestå av en rad teorem som har bevisats följa logiskt från vissa axiom eller antaganden. Men om spelteorin inte har något substantiellt innehåll kan den inte gärna komma i konflikt med en moralisk princip.

Ibland händer det förstås att folk vill göra det moralen kräver och i fångarnas dilemma kunde man kanske anse att moralen, som Kants princip kan tyckas antyda, kräver att bägge spelarna ska vara givmilda (eller samarbeta, i den ursprungliga versionen). Det visar sig för övrigt också att folk i verkligheten ofta väljer att vara givmilda. Förutsättningen att spelarna bara är intresserade av pengar är ju egentligen väldigt konstig och farligt vilseledande, eftersom alla vet att den är felaktig. Spelteoretikerna har också själva visat empiriskt att den inte stämmer, bland annat i så kallade ultimatumspel (jfr s. 63f).

Binmore menar därför att vi mycket väl kan – och egentligen också bör – slopa antagandet att folk bara bryr sig om pengar. Den ”valuta” som spelteoretiker intresserar sig för är så kallad *nytta* (eng. ”utility”). Inte i den hedonistiska betydelse som klassiska utilitarister som Bentham och Mill tog fasta på, utan i den moderna preferensteoretiska meningen (jfr s. 12). Varje (rationell) individ antas då ha en ”nyttofunktion” som tilldelar ett numeriskt nyttovärde till varje utfall som kan bli resultatet av ett spel. Denna nyttofunktion bestäms av hur individen väljer mellan olika alternativ, inklusive val mellan lotterier. De ”preferenser” som kan beskrivas på detta sätt är så kallade ”avslöjade” (*revealed*) preferenser (s. 12–3). Vad folk anser vara bra och dåligt – eller vad folk skulle må bra eller dåligt av – spelar alltså ingen som helst roll för spelteoretiker; det avgörande är hur individerna faktiskt betar sig i valsituationer.

Man måste därför skilja mellan å ena sidan fångarnas dilemma i den ursprungliga ”fysikaliska” eller ”objektiva” meningen, där man spelar om pengar eller antal år i fängelse, och å andra sidan fångarnas dilemma i det ”preferensbeskrivande” eller ”subjektiva” spel, där de olika utfallen definieras med den *nytta* de har för respektive spelare. Det senare kan se ut så här:

	B ”ger”	B ”tar”
A ”ger”	A får a – B får b	A får c – B får d
A ”tar”	A får e – B får f	A får g – B får h

där $e > a > g > c$ och $d > b > h > f$. Vilka numeriska värden dessa åtta bokstäver betecknar beror på vilka nyttoskalor vi väljer. Men tabellen

beskriver ett fångarnas dilemma i den meningen att den undre raden är dominant för A och den högra kolumnen är dominant för B, vilket innebär att den nedre högra rutan ger det utfall som kommer att väljas, trots att den övre vänstra rutan hade varit bättre för båda. Det är den enda Nash-jämvikten i spelet (jfr s. 21 och 26). Men detta spel är inte symmetriskt på samma sätt som i det objektiva spelet. Vi kan t.ex. inte säga att A:s bästa utfall (dvs. *e*) är lika bra som B:s bästa utfall (dvs. *d*), ty vi kan inte göra intersubjektiva jämförelser mellan A:s nytta och B:s. Nyttan, i spelteorins mening, kan bara mätas intrasubjektivt på intervallskalenivå, med godtycklig nollpunkt och enhet.¹

I tabellen ovan har jag satt citationstecken kring ”ger” och ”tar”. Det är oklart vilka objektiva handlingar dessa beteckningar här kan stå för. Fångarnas dilemma i objektiv mening passar kanske inte så ofta. Men det kan finnas många olika objektiva spel som passar in på det subjektiva spelet.

Man kan också notera att subjektiva utfall är mindre ”självcentrerade” än de motsvarande objektiva. Om B väljer den vänstra kolumnen i det objektiva spelet, så är de spelteoretiskt relevanta utfallen som A kan få inte längre 20 respektive 30 kronor; snarare är frågan vilket A föredrar av de två objektiva utfallen [A får 20 kr och B får 20 kr] å ena sidan och [A får 30 kronor och B får ingenting] å den andra. Och det är ju mycket möjligt att A föredrar det första av dessa, dvs. att A i valet mellan dessa två utfall faktiskt väljer det första.

Kort sagt: det är oklart vad som *är* ett ”fångarnas dilemma”, trots att detta är det mest diskuterade av alla spel. Syftar man på ett objektivt eller ett subjektivt spel? Binmore ger inget tydligt svar på den frågan.

Att folk betar sig så att de tillfredsställer sina preferenser tycks vara en truism, eftersom de preferenser spelteoretiker intresserar sig för är ”avslöjade” preferenser. Men då kan man fråga sig om det överhuvud taget är möjligt att handla altruistiskt i fångarnas dilemma (tolkat subjektivt). Om spelarna skulle göra det, så spelar de väl inte längre fångarnas dilemma! Ty deras preferenser – som ju ”avslöjas” av deras beteende – skulle ju då se annorlunda ut än i den subjektiva tabellen ovan. I själva verket är den tabellen inte en beskrivning av ett spel, utan av ett handlande. Den säger att både A och B ”tar”.

Här kan man tänka sig två olika positioner, beroende på hur spelteorins syfte uppfattas. Som vi har sett tycks syftet enligt Binmore vara att *förutsäga* personers agerande i spelsituationer med hjälp av observationer av deras *tidigare* beteenden i olika valsituationer (jfr s. 12–13). Det skulle alltså innebära att man först skulle fastställa vilken nyttofunktion en

¹Något som Binmore tycks ha glömt bort i sin diskussion av blandade Nash-jämvikter i kampen mellan könen på s. 34.

individ har för att därefter göra förutsägelser. Men sådana förutsägelser förutsätter då att individens nyttofunktion inte förändras under tiden. Binmore tycks anse att det vore irrationellt av en person att inte agera i överensstämmelse med sina tidigare preferenser, dvs. att inte bete sig i överensstämmelse med hur man tidigare har betett sig. Men han ger inget som helst skäl för detta antagande. Och det vore ju också orimligt. Ett sådant antagande strider ju mot den väletablerade teorin att vi lär oss av våra misstag, så att vi inte upprepar dem i framtiden. Det är klart att vi ofta förutsäger en persons beteende på grundval av tidigare beteenden, men i allmänhet anser vi inte att det är *irrationellt* att bete sig på ett nytt sätt. Men det tycker tydligen Binmore och han kallar sådant beteende ”motsägelsefullt” (se t.ex. s. 13).

Den andra möjliga positionen skulle innebära att man accepterar att folks nyttofunktioner ofta förändras över tiden och att det inte är något irrationellt med det. I så fall ska vi i fångarnas dilemma hålla oss till spelarnas nyttofunktioner just vid den tid då spelet spelas. Men då blir det också omöjligt att spela det altruistiska ”ger”. Skulle man göra det, så spelar man inte fångarnas dilemma. Man kan därför inte heller säga att det ena beteendet är mer rationellt än det andra.² Spelarna gör bara det de gör och deras nyttofunktioner bestäms av vad de gör (och vad de skulle göra). Inte tvärtom.

Binmore diskuterar en stor mängd olika spel, mer eller mindre komplicerade. Ofta kan det vara lärorikt för den oinitierade. Men de teoretiska grundvalarna är delvis otydliga, bland annat på det sätt som jag här har försökt illustrera i samband med fångarnas dilemma. Låt mig avslutningsvis nämna ytterligare ett av de ”felslut” som Binmore tar upp i samband med ett spel som, i likhet med fångarnas dilemma, diskuterats flitigt av filosofer, nämligen *Newcombs paradox*.

I Newcombs paradox får man välja mellan att ta en eller två lådor, där den första, som är stängd, innehåller antingen en miljon dollar eller ingenting och den andra, som är öppen så att man ser vad som finns i den, innehåller tusen dollar. En mycket duktig förutsägare har på förhand lagt en miljon i den stängda lådan om och endast om hen har förutsagt att man bara kommer att ta den lådan. Man vet att förutsägaren har spelat

²Såvitt jag vet säger de spelteoretiska teoremen aldrig någonting om vad som är ”rationellt”. Det enda som kanske kunde kallas ”rationellt” i sammanhanget är väl i så fall det att en person uppfyller de antaganden som krävs för att hen alls ska ha en nyttofunktion och i sitt handlande maximera förväntad nytta. Binmore säger (s. 13) att dessa antaganden är ”blygsamma”, men det kan nog diskuteras – och det har också diskuterats, dock inte av Binmore i denna bok. Kanske har de flesta av oss inga nyttofunktioner alls, eftersom de ”blygsamma” villkoren är alltför krävande!

spelet med olika personer väldigt många gånger och nästan alltid har lyckats göra korrekta förutsägelser. Ska man då bara ta den stängda lådan (och därmed troligen få en miljon dollar) eller ska man ta bägge lådorna (och därmed troligen bara få tusen dollar)?

Jag sa en gång till Binmore, när vi bägge deltog i en konferens om spelteori, att jag för min del bara skulle ta den stängda lådan. Han förklarade då att jag inte var riktigt klok. I samma anda säger han i boken att den som säger att man bara ska ta den stängda lådan gör sig skyldig till en ”motsägelse” (s. 213). Och han säger att ”det blir [...] aldrig optimalt att spela en starkt dominerad strategi” (s. 214).³

Om saken vore så enkel som Binmore säger, så skulle det vara obegripligt att Newcombs paradox har diskuterats så mycket och att flera bedömare anser att det vore rationellt att bara ta den stängda lådan. Visserligen är det *dominanta* alternativet att ta bägge, ty oavsett om det ligger en miljon i den stängda lådan eller ej – vilket man inte kan påverka när man väljer, ty man kan (troligen) inte påverka det förflutna – så får man mest om man tar bägge. Men å andra sidan tycks den *förväntade nyttan* av att endast ta den stängda lådan vara större än den förväntade nyttan av att ta bägge. Ty den evidens man har tillgång till pekar ju på att sannolikheten för att det ligger en miljon i den stängda lådan är mycket stor om man bara väljer den. Poängen med Newcombs paradox är just att den tycks visa att två av den vanliga spelteoriens principer kan komma i konflikt.

I en undersökning bland över 3200 filosofer svarade 21,9 procent att de skulle välja – eller lutade åt att välja – två lådor medan 19,5 procent föredrog en låda. Men de flesta, 58,6 procent, valde något ”annat” (*PhilPapers Surveys*)⁴. De som svarade något ”annat” avsåg främst någon variant av ”vet ej”, men övriga 3,6 procent gav svar som möjligen kunde antyda att själva spelet på något sätt är omöjligt. Och kanske är det detta Binmore menar när han säger att Newcombs paradox rymmer en motsägelse. Motsägelsen skulle bestå i att förutsägaren på förhand ganska säkert vet vad man kommer att välja, trots att man har mer än ett handlingsalternativ.

Men är detta en motsägelse? Det verkar inte så. Att någon vet vad man *kommer* att göra utesluter väl inte att man *kan* göra något annat? Och förutsätter verkligen spelteorin att vi har ”handlingsfrihet” i en så stark mening att det är omöjligt att göra mycket goda förutsägelser av våra handlingar? Många filosofer – 51,5 procent enligt den undersökning

³Han anger också två misslyckade försök att undvika ”motsägelsen”, men de är konstiga och såvitt jag kan se innehåller de minst en felaktig siffra.

⁴Bland de filosofer (inklusive studenter och ev. en del icke-filosof) som räkade beslutsteori (antagligen inklusive spelteori) som en specialitet var siffrorna helt annorlunda. 53,8 procent av dessa valde två lådor, 21,2 procent valde en låda och 25 procent svarade något ”annat”.

jag nyss nämnde – anser exempelvis att handlingsfrihet är förenlig med determinism; och om de har rätt, så borde det ”i princip” vara möjligt att förutsäga alla våra handlingar, även i de fall då vi kunde ha handlat annorlunda. Här finns olösta problem som Binmore verkar helt omedveten om.

LARS BERGSTRÖM

Kort om döden

Lennart Lundmark

Fri Tanke, 2012. 148 s. ISBN 978-91-86061-44-9

Fri Tanke förlags *Kort om*-serie är – med inspiration från Oxford University Press-serien *Very Short Introductions* – avsedd att erbjuda initerade introduktioner till angelägna, ofta filosofiska problem. De hittills publicerade titlarna uppvisar en skimrande mångfald i både ämnesval (såsom ångest, virus, yttrandefrihet och spelteori) och kvalitet. Lennart Lundmarks *Kort om döden* är det första svenska originalverket i serien.

Till största delen är boken en redogörelse för några historiskt inflytelserika skolors syn på döden. Lundmark refererar epikuréernas två huvudargument för att döden inte kan vara något ont för den som dör: för det första finns inte personen längre när hon är död och är därför utom räckhåll för att skadas av den; för det andra är döden bara en spegelbild av den harmlösa period under vilken personen ännu inte blivit född. En lite mer kritisk diskussion av dessa viktiga men omstridda tankegångar skulle ha höjt temperaturen i Lundmarks framställning. Han uppehåller sig främst vid stoicismen, med dess rekommendation att med förnuftet vid rodret finna sig i naturens ordning. Denna lära lockade bland andra Cicero, som skrev att den vise inte brinner ”av törstigt begär eller flyger hojtande iväg i larvig munterhet” (s. 24) och att “[e]ländig är den åldrade man som inte under ett långt liv har lärt sig att strunta i döden” (s. 28). Den stoiker som ägnas störst utrymme är Seneca, inte minst dennes rekommendation att träna bort skadlig dödsskräck genom att reflektera över i vilka omständigheter det vore rationellt att ta sig själv av daga. Sådana omständigheter befinner sig enligt Seneca exempelvis en förtryckt individ i: ”Galning, varför klagar du! . . . Vart du än vänder ögonen finns en utväg ur plågorna . . . Frågar du efter frihetens väg? Vilken åder som helst i din kropp!” (s. 39) Lundmark skildrar även den längtan efter martyrdöden – uppfattad som en biljett till himmelriket – som utmärkte kristenhetens första generationer. ”Måtte jag få glädje av vilddjuren som väntar mig”, skrev biskop Ignatios av Antiokia förväntansfullt (s. 55). De

metafysiska problemen med den kristna doktrinen om kroppens återuppståndelse – i synnerhet frågan i kraft av vad det är samma kropp som först ruttnar bort och långt senare åter är vid full vigör – hade förtjänat en lite mer utförlig diskussion än de hastiga anmärkningar Lundmark tillhandahåller.

Efter 1500-talsförfattaren Montaigne, vars tankar om döden var starkt präglade av stoicismen, avslutar Lundmark sin filosofihistoriska exposé. De senaste decenniernas omfattande och högkvalitativa filosofiska debatt om dödens dålighet – en debatt som sattes i gång av Thomas Nagels uppsats ”Death” (i *Noûs* 1970) – får vi inte veta något om. Det är synd, särskilt med tanke på att många idéer i den debatten är lättfattliga nog för en bok av detta slag. (Jag tror knappast att någon skulle skriva en *Kort om biologi-* eller *Kort om psykologi-*bok som inte redogör för vad som sagts av specialister på ämnet efter renässansen.)

Lundmark övergår i stället till några specifika frågor i medicinsk etik, framför allt hjärndödsbegreppet och dödshjälp, med fokus på den svenska offentliga diskussionen. Eller kanske snarare avsaknaden av en sådan: Lundmark visar tydligt hur man från politiskt håll skyr frågorna av rädsla för att väcka obehag hos medborgarna – detta trots att allmänheten ofta är relativt liberal på dessa punkter.

Även om *Kort om döden* inte är alltför rik på filosofiskt innehåll och urvalet som sagt är lite underligt, så innebär varje enskilt kapitel trivsam läsning. Lundmark skriver klart och redigt, inte sällan med en underton av torr humor. I konkurrensen med andra verk i *Kort om*-serien står sig hans bidrag väl.

JENS JOHANSSON

Notiser

DE BÄSTA FILOSOFITIDSKRIFTERNA

I en omröstning nyligen, där filosofer ombads rangordna 25 filosofiska tidskrifter, blev resultatet som följer bland de tio högst rankade:

1. Philosophical Review
2. Journal of Philosophy
3. Noûs
4. Mind
5. Philosophy & Phenomenological Research
6. Ethics
7. Philosophical Studies
8. Australasian Journal of Philosophy
9. Philosopher's Imprint
10. Analysis

PRAKTISK JURIDIK

Göran Lambertz föreslår (i DN 13/12) att man ska ”diskutera vem som har största skulden till att så många recensenter har hyllat Dan Josefssons bok som ett vittnesmål om att psykoterapin har den främsta skulden till att Thomas Quick erkände ett antal mord som han inte hade begått”. Är det något att diskutera? Att boken har hyllats måste väl ändå vara författarens fel. Hade han inte skrivit boken skulle ju ingen ha hyllat den.

PALEONTOLOGISK EPISTEMOLOGI

I Richard Dawkins självbiografi *Nyfikenhet och förundran* (Fri tanke 2013) kan man bland annat lära sig att de gigantiska dinosaurerna hade två hjärnor! Två hjärnor var, alltså. Det berodde på att de var så långa att det blev alltför långt mellan huvudet och bakbenen, som ju var stora och viktiga. Därför fanns även en sorts hjärna i bäckenet, nära bakbenen. Detta förhållande har inspirerat B. L. Taylor till en dikt som börjar så här:

Behold the mighty dinosaur,
Famous in prehistoric lore,
Not only for his power and strength,
But for his intellectual length.

You will observe, by these remains,
 The creature had two sets of brains –
 One in his head (the usual place),
 The other at his spinal base.
 Thus he could reason *A priori*
 as well as *A posteriori*.
 No problem bothered him a bit,
 He made a head and tail of it.

DÖDA 2013

Kända filosofer som dött under förra året är Ronald Dworkin, Fred Dretske, Peter Geach, Arthur Szathmary, David Braybrooke, Avrum Stroll, Keith Gunderson, Arthur C. Danto, Edward McClennen och, i Sverige, Thomas Anderberg.

KRIPKES BORD

I sin berömda bok *Naming and Necessity* hävdar Saul Kripke att ett visst bord som han har framför sig inte kunde ha varit gjort av is. Det är faktiskt gjort av trä och detta, menar han, är en *nödvändig* egenskap hos bordet, en egenskap det har i alla möjliga världar där det alls existerar. Denna tes tycks ha accepterats av Kripkes läsare och åhörare. Åtminstone av de allra flesta. Den avgörande intuitionen är att *detta* bord, som är av trä, inte kunde ha varit av is – ty hade det varit av is, så skulle det inte vara *detta* bord.

Men analogt kan man kanske tycka att *detta* bord, som faktiskt är grönt, inte kunde ha varit vitt – ty hade det varit vitt, så skulle det inte ha varit *detta* bord. Och i så fall skulle distinktionen mellan nödvändiga och tillfälliga egenskaper kollapsa. Vad är den eventuella skillnaden?

Det gröna bordet kunde ha målats vitt för något år sedan. Därför kunde det nu ha varit vitt, trots att det då skulle ha varit samma bord som det som faktiskt är grönt. Men kan man inte resonera på samma sätt med alla dess egenskaper? Bordet kunde t.ex. ha renoverats för något år sedan – på ett sådant sätt att vissa av dess delar, gjorda av trä, byttes ut mot nya delar gjorda av plast. Renoveringen kunde ha pågått under en längre tid, under vilken fler och fler trädelar byttes ut mot plastdelar. Det är lätt att tänka sig. Lite svårare, men fullt möjligt, är det att tänka sig en liknande renovering där trädelar byts ut mot nya delar gjorda av is (som preparerats så att de inte smälter; en sådan renovering är kanske *fysikaliskt* omöjlig, men det är irrelevant, eftersom Kripke är intresserad av vad som är *metafysiskt* möjligt).

Frågan är då: är det ”samma” bord före och efter renoveringen? Detta är en gammal fråga i filosofin, oftast exemplifierad med ”Theseus skepp”. Är skeppet fortfarande Theseus skepp, när alla dess delar har bytts ut? Kripke tycks mena (i en fotnot på s. 51) att vi med det gradvis renoverade bordet har ett fall av vaghet. Han menar tydligen att bordet *inte* är detsamma när de flesta delarna bytts ut, men att det är detsamma när endast några få delar bytts ut. Var gränsen går är obestämt. Så kan man ju säga, men lika gärna kan man väl säga att det är *samma* bord – om än renoverat – när alla eller de flesta av dess delar har bytts ut. Det avgörande tycks då vara att bordet har en historia, där det vid någon eller några tidpunkter hade kunnat förändras på ett sätt som faktiskt inte aktualiserades.

Frågan ”Kunde bordet ha varit gjort av is?” kan också tolkas på ett annat sätt. Hittills har jag tolkat den som ”Kunde det *nu* ha varit gjort av is?”, men den kan också uppfattas som ”Kunde det *tidigare* ha varit gjort av is?”. Svaret är väl fortfarande ja. Vi kan tänka oss att bordet ursprungligen gjordes av is, men att det sedan gradvis renoverades så att isbitar byttes mot träbitar, så att det nu är helt och hållet av trä (bortsett från spikar och skruvar, etc).

Om *alla* egenskaper hos ett föremål kunde ha saknats, så har det alltså inga nödvändiga eller ”essentiella” egenskaper. Det skulle W. V. Quine ha gillat.

HUR MAN BLIR ETT HELGON

Den förrförrre påven, Johannes Paulus II, har nu blivit helgon efter att han har utfört två underverk. De mirakulösa underverken består i att två personer, var för sig, har blivit friska efter svåra sjukdomar. Påven var visserligen död när detta hände, men anhöriga till de sjuka bad den avlidne påven att han skulle ingripa så att sjukdomen skulle botas. När sedan sjukdomen försvann fanns ingen annan förklaring än att påven hade ingripit. Vilket skulle bevisas.

PÅSTÅENDEN OM VETENSKAP

Hjärnforskaren Martin Ingvar har skrivit en bok som heter ”Vad är vetenskap och hundra andra jätteviktiga frågor” (Bokförlaget Langenskiöld, 2012). Den ingår i en serie där olika författare besvarar motsvarande frågor, t.ex. ”Vad är konst . . .”, ”Vad är politik . . .”, och så vidare för sanning, sport, musik, litteratur, och sex. Varje bok innehåller alltså 101 frågor – som alla har en viss anknytning till det ämne som explicit anges – och varje fråga besvaras, på en sida, med ett kort och ett långt svar.

Man får bland annat lära sig att vetenskapen är sann ”om man utgår

från en korrekt modell" (2), att ett stickprov måste vara "tillräckligt" stort för att vara representativt (18), att "vetenskapen utgår från att sanningen förändras" (26), att Kant menade att "verkligheten inte var nåbar utan bara en illusion" (33), att allting har en orsak, "men vi hittar den inte alltid" (38), att "ju enklare förklaring desto mer sannolikt riktig" (44), att ontologin eller läran om varandet framför allt fokuserar på "beskrivningar av det som gör något till ett levande väsen" vilket "betyder att man försöker förstå vad det är som gör en apa till en apa och inte varför just en apa är som den är (accidens)" (61), att "psykologi är vetenskapen om hur vi människor fungerar och parapsykologi är tvärtom" (90), och att "ingen sanning är fullständigt bevisbar" (101).

Ingvars bok får väl anses vara ett exempel på populärvetenskap. Men för den oinitierade är nog svaren ofta rätt obegripliga, eller väldigt otillräckliga för att vara någorlunda rättvisande.

PERSONLIG ICKE-IDENTITET

När den franske författaren Guillaume Appollinaire skulle begravas höll vagnen med hans lik inte på att komma fram på gatorna på grund av jublande människomassor som kastade hattar i luften, hurrade, sjöng och ropade: "Guillaume är väck! Hurra! Ned med Guillaume!" Vännerna fann det groteskt, och på sitt sätt stilenligt. Den Guillaume som åsyftades var emellertid den tyske kejsar Wilhelm, som just flytt till Holland i samband med vapenstillståndet efter första världskriget.

FELET MED VÅR VÄRLD

Bertrand Russell konstaterade vid något tillfälle att "felet med vår värld är att de dumma är så säkra på sin sak och de kloka så fulla av tvivel". Det ligger väl en hel del i detta – även om man nog kan misstänka att det finns något mer som är fel i vår värld. Och hur är det med omvändningen? Kan man utgå från att de som är säkra på sin sak alltid är dumma och att de som tvivlar alltid är kloka? I så fall blir det rätt lätt att vara klok.

Medarbetare i detta nummer av *Filosofisk tidskrift*: Lars Bergström är professor i praktisk filosofi, Sören Häggqvist är docent i teoretisk filosofi i Stockholm, Jens Johansson är docent och universitetslektor i praktisk filosofi i Uppsala, Hemo Quaron är magister i filosofi och estetik och kandidat i idéhistoria vid Uppsala universitet, Daniel Rönnefeldt är fil.dr i teoretisk filosofi i Stockholm, Caj Strandberg är docent i praktisk filosofi i Lund och Frans Svensson är docent och forskarasistent i praktisk filosofi i Stockholm.

Instruktioner till skribenter

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

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

Manuskript till *Filosofisk tidskrift*:

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