

September 2010 Årgång 31 Nr 3

Filosofisk tidskrift

- MARTIN LEMBKE
3 Frågan om varat
- LARS BERGSTRÖM
10 Kompatibilism på villovägar
- GUNNAR SYRÉHN
21 Emil – existentialisten från Lönneberga
- TEODOR GARDELLI
25 Kunskapsbegreppet och ett svar till skeptikern
- GUNNAR ANDERSSON
33 Rättfärdigande och kritisk prövning
- LARS SAMUELSSON
40 Transitivitet och omärkbara effekter
- BJÖRN ÖSTBRING
51 Richard Rortys pragmatism
- 57 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 2010

Grafisk form: Ulf Jacobsen

Satt med Indigo Antiqua

Tryck: Carlshamn Tryck & Media AB, Karlshamn 2010

ISSN 0348-7482

1. INLEDNING

Det verkar finnas ting som inte måste finnas. Men om sådana ting finns, *varför* finns sådana ting? Eller i mängdteoretiska termer: varför är mängden av kontingenta element inte tom? Denna till synes enkla fråga har beskrivits som själva varats gåta och det mörkaste av alla filosofiska problem; låt oss tala, kort och gott, om ”Fråga X”. Syftet med det som följer är att undersöka vilket av de tänkbara svaren som är bäst.

Först ett par detaljer av terminologiskt slag. I modalsemantiska ordalag kan vi definiera ett kontingent ting som ett ting som existerar i den faktiska världen men inte i alla metafysiskt möjliga världar (hur nu dessa ska förstås), medan ett nödvändigt ting definieras som ett ting som existerar i alla metafysiskt möjliga världar (inklusive den faktiska världen). Fråga X förutsätter alltså att det finns ting i den faktiska världen som *inte* existerar i alla metafysiskt möjliga världar: en förutsättning som verkar rimlig, även om den inte är av a priori art.

Låt oss tillägga att Fråga X är en strikt metafysisk fråga i den bemärkelsen att den inte *kan* besvaras på naturvetenskaplig väg. För att alludera på Wittgenstein (1992, 121): icke *hur* de kontingenta tingen är, är mystiskt, utan *att* de är. Kanske finns det oändligt många kontingenta ting och kanske har dessa ting (eller någon underliggande kontingent substans) existerat i evighet, men frågan varför mängden av kontingenta element inte är tom har därmed inte fått ett svar.

2. FYRA TÄNKBARA SVAR

Vår uppgift är alltså att försöka avgöra vilket av de tänkbara svaren på Fråga X som är bäst. Men tänk om *Faktum X* – det (förmodade) faktum att mängden av kontingenta element inte är tom – är ett *factum brutum*, ett absolut oförklarligt faktum? *Att* det finns sådana *facta bruta* tas ju ofta för givet numera givet vår kunskap om kvantmekaniska fenomen; till exempel verkar det inte finnas något som förklarar varför en enskild

atom sönderfaller vid en viss tidpunkt. Kanske finns det heller inget som förklarar varför det finns kontingenta ting överhuvudtaget?

Kanske inte. Men i sådana fall, som Lars Bergström noterar i *Döden, livet och verkligheten* (2004, 198), ”bör vi försöka förklara *varför* det är oförklarligt”. Vår utgångspunkt måste ju vara att försöka förklara så mycket som möjligt; om vi tvingas till slutsatsen att något är ett factum brutum bör vi på ett epistemologiskt metaplan försöka förklara *varför* så är fallet.

Men en sak i sänder. Först måste vi undersöka saken så gott vi kan, givet att Fråga X faktiskt har ett svar. Det verkar då som att vi har att välja mellan fyra mycket generella alternativ.

- (A) Minst ett av de kontingenta tingen har aktualiserat sig självt.
- (B) Minst ett av de kontingenta tingen har aktualiserats av minst ett nödvändigt ting.
- (C) Det *måste* finnas minst ett kontingent ting.
- (D) Det är en *slump* att det finns minst ett kontingent ting.

Dessa svar är uttömmande (utan att därmed utesluta varandra) eftersom varje enskilt ting i den faktiska världen antingen är kontingent (sådant att det inte finns i alla möjliga världar) eller nödvändigt (sådant att det finns i alla möjliga världar).

Alternativ (A) bör vi nog lägga åt sidan, ty att ett icke-aktualiserat ting skulle aktualisera sig självt – att ett ting i ontologisk mening skulle lyfta sig självt i håret – verkar ju inte alls rimligt. Alternativ (B), å andra sidan, ska vi ägna en hel del uppmärksamhet; det kan med visst fog kallas det *teistiska* alternativet, även om det också skulle kunna förespråkas av dem som menar att kontingenta ting blott är tillfälliga formationer av en nödvändigt existerande grundsubstans.

Alternativ (C) skiljer sig från (A) och (B). Det gör ju gällande att det *måste* finnas minst ett kontingent ting (vilket som helst); att en värld *utan* minst ett sådant ting är metafysiskt *omöjlig*. Metafysikern David Armstrong menar att tanken på en dylik ”distribuerad nödvändighet av kontingenta ting” kan ”avfärdas” (2006, 282); en annan metafysiker, Bede Rundle, har dock ägnat en hel bok, *Why there is Something rather than Nothing*, åt att förorda detta alternativ. Erik Olsson förutspår i en recension (2005) att de flesta av oss nog kommer att finna Rundles argument ”förbryllande och paradoxala”, men det hindrar ju inte att vi listar (C) bland de *tänkbara* svaren på Fråga X.

Alternativ (D), slutligen, verkar mycket konstigt; vad förklarar man genom att hänvisa till slumpen? Vi får tillfälle att återkomma till denna fråga i ett senare skede, och då kommer vi också att se att (D) – liksom (C) – inte utgör något reellt alternativ.

Alla fyra svaren på Fråga X har en minst-ett-karaktär som gör att de kan smulas sönder i oändligt många varianter. Framför allt bör detta noteras i anslutning till (B), som ju till exempel kan tänkas innebära att minst ett av de kontingenta tingen har aktualiserats gemensamt av exakt 335 nödvändiga ting. Det modalfilosofiskt mest intressanta med (B) är dock att det postulerar *minst* ett nödvändigt ting. Om vi använder oss av Ockhams rakkniv, dvs. principen att vi inte bör postulera fler entiteter än vad som krävs för att förklara det som ska förklaras, bör vi heller inte postulera fler nödvändiga entiteter än nödvändigt för att besvara Fråga X.

3. ATT VÄLJA MELLAN (B) OCH (C)

Om vi lämnar *bootstrapping*-alternativet (A) och slumpalternativet (D) därhän står vårt val mellan alternativen (B) och (C). Bede Rundle, som vi redan har nämnt, förordar (C) med motiveringen att det är mer "ekonomiskt" (2004, 109) än (B), ty medan (B) postulerar ett *ting* postulerar (C) blott en abstrakt *princip*: en princip som dock räcker (enligt Rundle) för att besvara Fråga X.

Låt oss titta närmare på detta argument. Vad (C) påstår är att det måste finnas minst ett kontingent ting, medan (B) påstår att minst ett ting är sådant att det måste finnas. Det verkar inte alls tydligt på vilket sätt (C) skulle vara mer "ekonomiskt" än (B), om vi formulerar saken på detta sätt. Varför skulle en värld där minst ett kontingent ting måste finnas vara ontologiskt enklare än en värld där minst ett ting är sådant att det måste finnas?

Här kan vi göra en intressant iakttagelse. Såväl (B) som (C) implicerar ett påstående som är sant med *nödvändighet* – om det är sant överhuvudtaget, vill säga. Om (B) är sant finns det ju minst ett ting som existerar i alla möjliga världar, och i sådana fall är påståendet *att* detta (eller dessa) ting existerar sant i alla möjliga världar. På motsvarande sätt med (C); om det måste finnas minst ett kontingent ting är påståendet *att* det finns minst ett kontingent ting sant i alla möjliga världar. Vare sig vi väljer (B) eller (C) tvingas vi alltså acceptera ett påstående såsom nödvändigt sant.

När vi talar om nödvändigt sanna satser tänker vi nog ofta på tautologier som "alla cirklar är runda" och logiska syllogismer som "om alla P är Q , och x är P , så är x Q ". Ibland avses endast satser av dessa slag när man talar om *logiskt* nödvändiga satser. Men därmed är ju inte sagt att satser som är sanna i alla möjliga världar måste vara av dessa slag. Andra slags kandidater som ofta diskuteras är av *metafysiskt* slag – satser som till exempel "ur intet kommer intet", "alla kroppar har utsträckning" och "värme är molekylär rörelse". Men även *normativa* satser som "det är fel att döda för nöjes skull", *axiologiska* satser som "människor är mer värda

än schimpanser” och *teologiska* satser som ”Gud är treenig” är exempel på påståenden som vissa anser vara sanna i alla möjliga världar även om de ju alls inte är av tautologiskt eller syllogistiskt slag.

De båda påståenden som impliceras av (B) och (C) är inte heller av tautologiskt eller syllogistiskt slag. Likväl är ett av dem en nödvändig sanning, om antingen (B) eller (C) är korrekt. (Kanske att det i sig förklarar varför vissa hellre ser Faktum X som ett *factum brutum*.) Alternativ (B) implicerar påståendet att det finns minst ett nödvändigt ting, dvs. att *det finns minst ett ting som i sig är sådant att det måste finnas*. (C), å sin sida, implicerar påståendet att det finns minst ett kontingent ting, dvs. att *det finns minst ett ting som i sig är sådant att det inte måste finnas*.

Återigen är det svårt att se varför (C) skulle vara mer ”ekonomiskt” än (B). Snarare frestas man (nästan) till ett slags ontologiskt argument till förmån för (B). Till skillnad från påståendet ”det finns minst ett ting som i sig är sådant att det inte måste finnas” kan ju påståendet ”det finns minst ett ting som i sig är sådant att det måste finnas” tolkas som en tautologi; alltså ligger det närmare till hands att tänka sig att implikationen av (B) är sann med nödvändighet.

Dock är (C) mer ”ekonomiskt” i åtminstone *en* bemärkelse, ty det påstår ju inget annat än att det måste finnas minst ett ting som i sig är sådant att det inte måste finnas. (B) påstår däremot, inte bara att det finns minst ett ting som i sig är sådant att det måste finnas, utan också att *detta (eller dessa) ting har aktualiserat minst ett kontingent ting*. Medan (C) reser *en* följdfråga, nämligen *varför* det måste finnas minst ett kontingent ting, reser (B) minst *två* följdfrågor: dels *varför* minst ett ting är sådant att det måste finnas, dels *varför* detta (eller dessa) ting har aktualiserat minst ett kontingent ting.

Men inte heller i denna bemärkelse verkar (C) vara att föredra framför (B), ty dess bättre ”ekonomi” vinnas ju på bekostnad av sämre substans. Det faktum att (C) genererar färre följdfrågor än (B) beror helt enkelt på att det inte säger lika mycket – om det överhuvudtaget säger något alls. Frågan (Fråga X) är ju varför det finns kontingenta ting. (C) svarar att det *måste* finnas kontingenta ting. Men vad är det för slags svar, egentligen?

Bergström behandlar en parallell frågeställning och lyfter fram det problematiska i ett svar som ”upprepar det som ska förklaras och dessutom tillägger något, som inte förklarar det som ska förklaras” (2004, 192). Det som ska förklaras är ju varför det i den faktiska världen finns kontingenta ting; (C) svarar att det finns kontingenta ting, inte bara i den faktiska världen, utan i *alla* möjliga världar. Men på vilket sätt förklaras därmed varför det i den faktiska världen finns kontingenta ting?

Den här intetsägande karaktären blir än tydligare om vi återvänder till

det till synes poänglösa svarsalternativet (D): Det är en *slump* att det finns minst ett kontingent ting. Anta nämligen att vi ”besvarar” Fråga X på detta sätt. Vad vi då gör är att tillägga något, som inte förklarar det som ska förklaras, utan att ens upprepa det som ska förklaras! Vad (D) säger är ju att det finns minst en möjlig värld där det *inte* finns kontingenta ting, men på vilket sätt skulle det förklara varför det i den faktiska världen *finns* kontingenta ting? Problemet för (C) är att det verkar ha samma intetsägande karaktär som (D).

4. EN NÄRMARE ANALYS AV (B)

Eftersom (A) är osannolikt på gränsen till omöjligt och (C) och (D) är intetsägande står vårt epistemologiska hopp till alternativ (B): minst ett av de kontingenta tingen har aktualiserats av minst ett nödvändigt ting.

Låt oss alltså anta att (B) är sant. Då uppstår följdfrågan *varför* minst ett av de kontingenta tingen har aktualiserats av minst ett nödvändigt ting. Att svara i termer av *slump* – att minst ett av de kontingenta tingen *av en slump* har aktualiserats av minst ett nödvändigt ting – vore lika intetsägande som att besvara Fråga X med (D), och att svara i termer av *nödvändighet* – att minst ett av de kontingenta tingen *med nödvändighet* har aktualiserats av minst ett nödvändigt ting – vore lika intetsägande som att besvara Fråga X med (C). Men vad återstår då för alternativ, om varken slump eller nödvändighet förslår?

Det enda som verkar återstå är en förklaring i termer av *fri vilja*: minst ett av de kontingenta tingen har *av fri vilja* aktualiserats av minst ett nödvändigt ting. Detta svar förutsätter förstås att det finns något sådant som en genuint libertariansk (inkompatibilistisk) fri vilja. Här finns inte utrymme att närmare gå in på den saken. Låt det vara nog sagt att om det *inte* finns en sådan vilja verkar det heller inte finnas något rimligt svar på frågan varför minst ett av de kontingenta tingen har aktualiserats av minst ett nödvändigt ting, och i sådana fall bör vi snarast konstatera att alternativ (B) leder oss till ett factum brutum.

Ett sådant factum brutum verkar i alla hänseenden oundvikligt, givet att (B) är sant. En annan följdfråga är ju varför minst ett ting är sådant att det existerar med nödvändighet; vad är det som gör att detta (eller dessa) ting är nödvändiga? Återigen vore det intetsägande (och i detta fall till och med självmotsägande) att svara i termer av slump, liksom det vore intetsägande att svara i termer av nödvändighet. Men det går inte heller att svara i termer av fri vilja, ty om ett ting vore nödvändigt av fri vilja vore ju dess nödvändighet i själva verket *inte* nödvändig. Här verkar vi alltså oundvikligen nå fram till en slutstation där vi epistemologiskt sett måste hoppa av. Frågan varför minst ett ting är sådant att det existerar

med nödvändighet (givet att det finns ett sådant ting) kan helt enkelt inte besvaras.

Om (B) är sant – och vad är alternativet? – måste vi alltså i den logiska förlängningen acceptera tanken på ett *factum brutum*: det nödvändiga tingets nödvändighet låter sig inte förklaras. Men om så är fallet, varför då inte ta ett steg tillbaka och i stället låta Faktum X – det (förmodade) faktum att det finns kontingenta ting – vara ett *factum brutum*? Varför inte låta ”varats mysterium” förbli ett mysterium i stället för att lösa det med hjälp av ett *annat* mysterium?

5. ATT FÖRKLARA VARFÖR NÅGOT INTE KAN FÖRKLARAS

Svaret har att göra med själva utgångspunkten för kunskapsteoretiska framsteg: vi måste ju försöka förklara så mycket som möjligt. Om ett *prima facie* mysterium visar sig kunna förklaras kan vi inte avstå från den möjligheten bara för att förklaringen i sig implicerar ett *nytt* mysterium. Vi kan inte *välja* vad som ska vara ett mysterium; antingen är ett faktum oförklarligt eller också inte.

Här är det läge att recapitulera Bergströms metaepistemologiska princip: om något är oförklarligt bör vi försöka förklara *varför* det är oförklarligt (jfr 2004, 198). Vad gäller det *factum brutum* som impliceras av alternativ (B) verkar vi åtminstone ha kunnat ge en *induktiv* metaförklaring av detta slag, ty vi har ju kunnat *utesluta* alla tänkbara metafysiska förklaringar i termer av slump, nödvändighet och fri vilja. Om Faktum X däremot hade varit ett *factum brutum* hade vi tvingats till den märkliga slutsatsen att det var oförklarligt *trots* att en till synes möjlig förklaring stod till buds. En sådan slutsats hade nog fått sägas strida mot andemeningen i Bergströms metaepistemologiska princip.

6. SLUTSATSER

Det verkar som att Fråga X – Varför är mängden av kontingenta element inte tom? – bör besvaras i linje med alternativ (B): minst ett av de kontingenta tingen har aktualiserats av minst ett nödvändigt ting. I enlighet med analysen i avsnitt (4) kan vi precisera detta alternativ i termer av fri vilja (ty det verkar ju vara den enda möjligheten): minst ett av de kontingenta tingen har *av fri vilja* aktualiserats av minst ett nödvändigt ting. (Eller mindre tvetydigt: minst ett nödvändigt ting har av fri vilja aktualiserat minst ett kontingent ting.)

Vad vi tycks ha kommit fram till är alltså inget mindre än början på ett kosmologiskt kontingensargument för Guds existens. Om vi accepterar dess metafysiska grundantagande, nämligen att det i den faktiska världen

finns minst ett kontingent ting, tycks vi ju ha skäl att postulera minst ett nödvändigt ting, ett ting som inte bara är utrustat med en genuint fri vilja utan som också är kapabelt att skapa kontingenta ting ur intet. Det får ju sägas vara en bra början på en beskrivning av Gud, definierad på vanligt teistiskt vis.

Alternativet verkar vara att betrakta Faktum X – det (förmodade) faktum att det finns kontingenta ting – som ett *factum brutum*: ett oförklarligt faktum, ett mysterium i absolut bemärkelse. Problemet med detta alternativ är *inte* att det hänvisar oss till en epistemologisk ändstation, ty det gör i förlängningen även alternativ (B); dess problem är snarare att det påstår att något, som verkar kunna ges en förklaring, inte kan ges en förklaring.

LITTERATUR

- Armstrong, David. 2006. "Reply to Efrid and Stoneham". *Australasian Journal of Philosophy* 84 (2):281–3.
- Bergström, Lars. 2004. *Döden, livet och verkligheten: Essäer*. Stockholm: Thales.
- Olsson, Erik J. 2005. Recension av Bede Rundle, *Why there is Something rather than Nothing*, i *Notre Dame Philosophical Reviews*, 2005-03-05. <http://ndpr.nd.edu/review.cfm?id=2081>
- Rundle, Bede. 2004. *Why there is Something rather than Nothing*. Oxford: Oxford University Press.
- Wittgenstein, Ludwig. 1992. *Tractatus logico-philosophicus*. Stockholm: Thales.

1. INKOMPATIBILISM

Kompatibilism är den doktrin som innebär att handlingsfrihet (eller fri vilja) är förenlig – kompatibel – med determinism. Inkompatibilism förnekar detta. Den kände filosofen Peter van Inwagen, som själv är inkompatibilist, rapporterar att enligt en kvalificerad bedömning är de flesta analytiska filosofer kompatibilister medan de som särskilt sysslat med frågor om fri vilja och handlingsfrihet numera oftast är inkompatibilister.¹ För sin del trodde van Inwagen annars att de flesta var kompatibilister.

Standardargumentet för inkompatibilism går ungefär så här. Antag att jag vid en viss tidpunkt t inte lyfter min hand.² Determinismen säger att allting som händer i världen är orsaksbestämt. Detta innebär att händelser och sakförhållanden i världen före t enligt naturlagarna orsakar att jag inte lyfter handen vid t . Om F är en beskrivning av det totala tillståndet i det förflutna före t – eller, om man så vill, före min födelse – och L är en konjunktion av alla naturlagar, så följer det logiskt från $F \& L$ att jag inte lyfter handen. Kunde jag då ha lyft handen? Nej, kan man tycka, ty det skulle ha inneburit att antingen F eller L (eller bägge) hade varit falska, och jag har inte i min makt att ändra vare sig det förflutna eller lagarna. Alltså kunde jag inte ha lyft handen. Och alltså kan jag av liknande skäl – mer generellt – *aldrig* göra något annat än det jag faktiskt gör.

Detta argument – ofta kallat ”konsekvensargumentet” – kan utfor-

¹Se van Inwagen 2004, fotnot 21, s. 350. Enligt en opinionsundersökning bland filosofer nyligen (*PhilPapers Surveys*) var 550 av 931 filosofer (mer eller mindre övertygade) kompatibilister, dvs. 59 procent (av metafysiker 55,5 procent och av handlingssteoretiker 53,4 procent). Cirka 12 procent av filosoferna ansåg att vi inte har någon fri vilja. Men det finns ingen garanti för att dessa siffror är representativa.

²Handlingen att lyfta handen förekommer i van Inwagens version av argumentet. Det kan tyckas vara en lite väl trivial handling, men den kan vara nog så betydelsefull i ett visst sammanhang, t.ex. när man ska rösta genom handuppräckning eller när man bjuder en viss summa pengar på en auktion. Samtidigt är det en enkel kroppsörelse, som inte kräver någon större ansträngning eller speciell kompetens.

mas på lite olika sätt. De på senare tid mest kända formuleringarna har utarbetats av van Inwagen.³ För egen del är jag starkt benägen att godta någon variant av det. Den mest kända invändningen mot konsekvensargumentet har framförts av David Lewis.⁴ Jag ska återkomma till det. Enligt min mening är Lewis motargument inte särskilt övertygande. Men först kan det vara lämpligt att kortfattat begrunda några andra möjliga invändningar mot konsekvensargumentet.

2. KLASSISK KOMPATIBILISM

För det första kan man undra om inte konsekvensargumentet omedelbart undermineras av ett gammalt och välkänt argument för kompatibilism, som går ut på att ordet ”kunde” ska tolkas på ett sådant sätt att det är förenligt med determinism att vi ofta kunde ha gjort något annat än det vi gör.⁵ En variant av detta argument innehåller antagandet att ”jag kunde ha lyft handen vid *t*” ska uppfattas som synonymt med ”jag skulle ha lyft handen vid *t*, om jag hade försökt (eller valt) att göra det”. Uppenbarligen utesluter inte determinismen att jag kunde ha lyft handen i denna *hypotetiska* mening. Men det hjälper knappast, ty konsekvensargumentet kan då återigen användas för att visa att jag inte kunde ha försökt att lyfta handen, om determinismen är sann.⁶

Att jag skulle ha utfört en handling av en viss sort, om jag hade försökt göra det, innebär bara att jag har (eller hade) en generell *förmåga* (eng. ”ability”) att utföra den sortens handlingar och att jag också hade *tillfälle* (eng. ”opportunity”) att göra det vid den aktuella tidpunkten. Ibland saknar man förmåga eller tillfälle. Jag har t.ex. förmåga att lyfta handen, men jag har inte förmåga att lyfta 100 kg med en hand. Jag har inte tillfälle att lyfta handen om den vid den aktuella tidpunkten är fastbunden vid kroppen. Jag har förmåga att köra bil – det ingår så att säga i min repertoar – men jag saknar tillfälle att göra det när jag inte har tillgång till någon bil eller när jag har tappat bilnyckeln.

³Se t.ex. van Inwagen 1975, 1983 och 2004.

⁴Lewis 1981; i fortsättningen citerad från den svenska översättningen 2000.

⁵En klassisk variant av detta argument finns exempelvis i kapitel 6 av Moore 1912. G. E. Moore tänker sig där att ”I could” ska tolkas som ”I should, if I had chosen”.

⁶Redan Moore insåg denna svårighet, och att det inte hjälper att jag skulle ha försökt lyfta handen om jag hade försökt (eller valt) att försöka göra det. Osv. Osv. Här öppnar sig uppenbarligen en infinit regress. Om man vill undvika den kan man ersätta ”försökt” med något ord som inte betecknar en handling, t.ex. ”önskat”. Men det hjälper inte, ty jag önskade ju inte lyfta handen och kunde inte heller ha framkallat denna önskan (om vi inte återigen ska råka in i en regress).

Att jag har både förmåga och tillfälle att utföra en viss sorts handling, som jag i en given situation faktiskt inte utför, visar emellertid inte att jag kunde ha utfört den i den aktuella situationen – utom just i den hypotetiska betydelsen. En inkompatibilist, som dessutom är determinist, skulle påpeka att jag i den situationen var förutbestämd att inte försöka utföra handlingen. Och därmed också förutbestämd att inte utföra handlingen i just denna situation.⁷

3. NATURLAGARNA

Konsekvensargumentet förutsätter att man inte kan ”bryta mot” eller ”ändra” naturlagarna, men hur vet vi att detta är en riktig förutsättning? Det hänger delvis på vad naturlagar är för något och det finns det olika uppfattningar om.

Enligt en klassisk uppfattning – som brukar förknippas med David Hume – är naturlagarna bara vissa regelbundenheter i verkligheten. Men om lagarna bara är regelbundenheter,⁸ så kan de inte *påverka* det som händer och de kan inte heller *tvinga* oss att handla på något särskilt sätt. De ”bestämmer” inte vårt handlande. Snarare då tvärtom: det är (bland annat) våra handlingar som bestämmer vilka naturlagarna är. Och i så fall kan man kanske ”bryta mot” eller ”ändra” naturlagarna. Visserligen inte så att naturlagarna först är på ett sätt och sedan på ett annat, utan så att man *kunde* ha handlat annorlunda än man gör och att om man hade gjort det så *skulle* lagarna ha varit annorlunda än de faktiskt är.

Men det finns också en annan uppfattning om naturlagarna, nämligen att de är regelbundenheter som råder med en sorts *nödvändighet*. Det är inte bara sant att ingenting rör sig snabbare än ljuset, utan detta är i någon mening också nödvändigt. Men det är oklart vad detta är för en sorts nödvändighet. Det är ju inte så att en regelbundenhet är ”nödvändig”, i den relevanta betydelsen, om den gäller i varje möjlig värld. Ty vi kan inte utesluta att det finns möjliga världar i vilka föremål rör sig snabbare än ljuset.⁹ Möjligen kunde man tänka sig att en regelbundenhet är nödvändig i den relevanta betydelsen om ingen människa (eller annan levande varelse) har både förmåga och tillfälle att upphäva den i något

⁷Den som vill skulle då kunna säga att jag i *en viss mening* saknar antingen ”förmåga” eller ”tillfälle” att utföra handlingen – t.ex. att lyfta handen – i just denna situation, trots att jag i en mer vardaglig mening av orden har både förmåga och tillfälle att göra det.

⁸Eller generella beskrivningar av samband.

⁹Sådana världar är logiskt och – kanske – metafysiskt möjliga även om de inte är fysikaliskt möjliga.

enskilt fall.¹⁰ Med denna tolkning blir det förstås omöjligt för oss att bryta mot naturlagarna.

Just denna tolkning har jag aldrig träffat på i den filosofiska litteraturen om naturlagar. Men vi måste ändå tänka oss att konsekvensargumentet förutsätter någon sorts nödvändighetstolkning. Annars tappar ju konsekvensargumentet helt sin förmåga att övertyga. Vi kan emellertid inte förutsätta den tolkning som jag just föreslagit, ty då blir argumentet å andra sidan alldeles för starkt; då förutsätter det alldeles för uppenbart det som skulle visas. Den filosofiska diskussionen om vad naturlagar är har inte utmynnat i något allmänt accepterat resultat. Konsekvensargumentet har därför på denna punkt en viss svaghet.

4. DET FÖRFLUTNA

Konsekvensargumentet förutsätter också att man inte (i någon relevant mening) kan ”ändra” eller ”påverka” det förflutna.¹¹ Men hur vet vi egentligen att vi inte kan det? Kanske vårt orsaksbegrepp är sådant att orsaker alltid föregår sina effekter i tiden.¹² Våra handlingar kan i så fall inte *orsaka* några händelser i det förflutna. Men det kan kanske ändå gälla i vissa fall att om vi hade handlat *annorlunda* än vi gör, så *skulle* det förflutna ha varit annorlunda än det är. I sådana fall skulle vi då i denna mening kunna ”ändra” det förflutna – om vi kunde ha handlat annorlunda.

Men här, om inte förr, måste man göra en distinktion mellan två olika betydelser av ”kunde”, nämligen mellan vad vi kunde kalla en *aktiv* och en *passiv* betydelse. Den passiva betydelsen kan illustreras av händelser som kunde ha inträffat som ett resultat av tidigare händelser, eller möjligen av en ren slump. En lampa som är släckt ”kunde” t.ex. ha varit tänd och en kastad tärning som visar en etta kunde ha visat en sexa. Men det är väl klart att det inte är denna passiva betydelse som är av intresse i diskussionen om fri vilja och handlingsfrihet. Lampan skulle ha varit tänd om någon hade vridit om strömbrytaren och i den meningen

¹⁰Jag antar att hur mycket vi än försöker så kommer vi aldrig att lyckas accelerera en partikel till en hastighet som är större än ljusets.

¹¹I van Inwagens version av argumentet motsvaras det förflutna – dvs. *F* i sektion 1 ovan – av ”världens tillstånd” vid en tidpunkt innan den handlande personen är född.

¹²I själva verket har vi nog inte *ett* orsaksbegrepp, utan obestämt många olika. Men kanske gäller det för de flesta av dessa att ingenting kan orsaka händelser i det förflutna. Detta skulle då möjligen kunna anses vara ”en begreppslik sanning”. För egen del är jag rätt skeptisk till att det skulle finnas begreppslika sanningar – och särskilt till antagandet att filosofiskt intressanta påståenden skulle kunna vara begreppslika sanningar.

”kunde” den ha varit tänd. Men den kunde ju inte själv ha vridit om sin strömbrytare, den kunde inte ha tänt sig själv. Den hade inte i sin makt att göra något alls. Den ”kunde” visserligen ha varit tänd i den passiva betydelsen, men den kunde inte ha varit tänd i det jag kallar den aktiva betydelsen av ”kunde”.

Att jag ”kunde” ha lyft min hand vid tidpunkten *t* kan tolkas antingen passivt eller aktivt. Med den *passiva* tolkningen innebär det att det finns en möjlig värld i vilken jag¹³ lyfter handen och i vilken min handling är orsakad av vissa tidigare händelser, som inte inträffade i den aktuella världen.¹⁴ Det förflutna är alltså annorlunda i den möjliga värld där jag lyfter handen. Men det är inte jag eller min handling som orsakar de händelser i det förflutna (eller i lagarna) som i den alternativa världen orsakar min handling. Snarare tvärtom.

Om jag däremot i den *aktiva* meningen ”kunde” ha lyft handen vid *t*, så innebär det att min handling inte bara är en konsekvens av någonting i omvärlden, något som jag inte har någon kontroll över. Det kan t.ex. bero på att min handling orsakar de förflutna händelser som krävs för att jag ska lyfta handen i den möjliga värld i vilken jag gör det. Det skulle alltså innebära att jag genom att lyfta handen kan ändra det förflutna (eller lagarna) – om nu detta är nödvändigt för att jag ska lyfta handen. Det låter rätt häpnadsväckande! Kan man verkligen åstadkomma sådana förändringar bara genom att lyfta på handen? Det är nog inte alldeles självklart att jag verkligen *kunde* ha lyft handen vid *t* i den aktiva betydelsen av ”kunde”.

Eventuellt kan skillnaden mellan en aktiv och en passiv tolkning av ”kunde” också illustreras med följande exempel. Jag kastar en tärning och får en tvåa. I den situationen ”kunde” jag (i en passiv mening) ha fått en sexa; däremot ”kunde” jag inte (i aktiv mening) ha fått en sexa. Likaså ”kan” jag (passivt) dra en vinstlott i ett visst lotteri, men jag ”kan” inte (aktivt) dra en vinstlott i det lotteriet. Det är inte självklart att denna distinktion är precis densamma som den jag först antydde.¹⁵

¹³Eller kanske snarare min ”motsvarighet” (eng. *counterpart*), eftersom jag inte gärna kan vara identisk med min motsvarighet i andra möjliga världar, eftersom vi ju har delvis olika egenskaper. Om det nu inte finns en individ, som vi alla är olika delar av.

¹⁴Jag förenklar lite här. Att jag lyfter handen i den alternativa världen kunde också bero på att naturlagarna i den världen är annorlunda än i den värld där jag inte lyfter handen. Det kunde också vara ett resultat av att både lagarna och det förflutna är annorlunda.

¹⁵Jag har förmåga att lyfta handen, men jag har inte förmåga att slå en sexa eller att dra en vinstlott, åtminstone inte i den meningen att jag har kontroll över om jag lyckas eller inte. – Distinktionen mellan ett passivt och ett aktivt ”kunde” kan också jämföras med van Inwagens distinktion mellan ”might have done” och

Hur man exakt ska definiera ”kunde” och ”kan” i dessa sammanhang är inte lätt att säga. Men tre saker verkar rätt klara: (1) hur dessa ord tolkas är av central betydelse för frågan om handlingsfrihet, (2) det finns inte någon entydig och precis betydelse av dessa ord i vanligt språk; bland annat finns åtminstone en tvetydighet beroende på distinktionen mellan aktiva och passiva betydelse, och (3) i konsekvensargumentet måste ”kunde” och ”kan” tolkas med en aktiv innebörd, ty det verkar uppenbart att man i en passiv mening kunde ha gjort något annat än det man gör.

Kan vi aktivt ändra det förflutna? Egentligen är det kanske inte svårare att ändra det förflutna än att ändra det som händer i framtiden. Antingen finns det (ännu) ingen sanning om vad som händer i framtiden eller också finns det en sådan sanning. I det första fallet finns det ingenting att ändra. I det andra fallet är situationen i princip densamma som i fråga om det förflutna: det som är sant är sant. Ur objektiv synpunkt är det knappast någon skillnad mellan framtiden och det förflutna. Denna skillnad kan bara vara meningsfull ur ett subjektivt perspektiv. Den förutsätter ett ”nu” relativt till vilket vissa händelser är förflutna och andra är framtida och nuet förutsätter ett subjektivt perspektiv. Ur ett sådant perspektiv kan det *tyckas* som att framtiden är mer föränderlig eller påverkbar än det förflutna – men att det *tycks* så innebär inte nödvändigtvis att det *är* så.

5. DETERMINISM

Determinismen innebär som sagt att allting som händer i världen är orsaksbestämt. Detta brukar också kallas ”den tillräckliga grundens lag”: alla händelser har en tillräcklig orsak. Annorlunda uttryckt: varje händelse är en nödvändig konsekvens av tidigare händelser i kombination med naturlagarna.

Man kan undra om determinismen är riktig. Detta förutsätts visserligen inte av konsekvensargumentet – som ju bara går ut på att *om* determinismen är riktig, *så* har vi ingen handlingsfrihet. Men argumentet kan kanske tyckas bli betydligt mindre intressant om determinismen skulle vara falsk.¹⁶

”was able to do” (van Inwagen 2008, s. 332); riktigt hur de två distinktionerna förhåller sig till varandra är jag osäker på. Man kan också jämföra med Lewis distinktion mellan en svag och en stark betydelse av ”kunde ha falsifierat ett påstående” (se Lewis 2000, s. 396); en distinktion som enligt Lewis också framförts av Keith Lehrer i en uppsats från 1980 (se Lewis, s. 397). Jag ska återkomma till detta i avsnitt 6 nedan.

¹⁶David Lewis är kompatibilist, men hans kritik av konsekvensargumentet bygger inte på att determinismen skulle vara falsk. I själva verket anser emellertid Lewis också att determinismen *är* falsk, se Lewis 2000, s. 390.

Det brukar ju ofta sägas att kvantfysiken har visat att världen inte är deterministisk och att en del subatomära händelser är slumpartade.¹⁷ Själv har jag ingen klar uppfattning om riktigheten i detta, men mitt intryck är att hjärnforskare i allmänhet anser att slumpfenomen på mikronivå inte har någon inverkan på våra handlingar – och van Inwagen påstår för sin del att ”obestämdhet på kvantnivå spelar uppenbarligen ingen roll i den kausala förhistorien till våra handlingar”.¹⁸ Men denna uppfattning tycks inte vara oomstridd. Det har hävdats att det kan finnas kaotiska processer i hjärnan och i sådana processer skulle då subatomära slumpändelser kunna förstöras upp betydligt, så att de påverkar mentala händelser och beteenden.¹⁹

Vissa formuleringar av determinismen – som t.ex. van Inwagens och den klassiska formulering som brukar tillskrivas Laplace – använder sig av uttryck av typen ”världens totala tillstånd vid tidpunkten x ”. Man kan undra om det finns något sådant. Dels tycks det förutsätta någon bestämd och objektiv uppsättning egenskaper och relationer, vilket är svårt att tro på om man inte tror på Platons idévärld. Och dels tycks det förutsätta att *samtidighet* har en objektiv innebörd – något som strider mot den speciella relativitetsteorin, där ju *samtidighet* är relativ till koordinatsystem.

Jag tror inte att det är så lätt att ge en definition av determinismen, som är både exakt och intressant. Men kanske kunde man säga ungefär så här: för *varje* händelse H så *finns det* en uppsättning naturlagar, L , och en uppsättning händelser i det förflutna, F , sådana att varje möjlig värld som innehåller L och F också innehåller H (medan det däremot finns möjliga världar som innehåller L men inte H och världar som innehåller F men inte H). Så länge man inte riktigt vet vad ”naturlagar” är för något, så finns det ändå en viss oklarhet här. Men kanske är det ändå tillräckligt klart för våra syften.²⁰

¹⁷Detta kan dock ifrågasättas; se t.ex. Carl Hoefer 2010, sektion 4.4. Å andra sidan påpekar också Hoefer att determinismen inte går särskilt bra ihop med den allmänna relativitetsteorin; se sektion 4.3.

¹⁸Min övers. av van Inwagens formulering: ”quantum indeterminacy obviously plays no part in the causal genesis of human action” (2004, s. 350).

¹⁹Se t.ex. Robert Kane, 1994, avsnitt 10. Här finns också vidare litteraturhänvisningar. Å andra sidan finns det tydligen också helt deterministiska tolkningar av kvantfysiken.

²⁰En ytterligare en komplikation kan vara den att naturlagar bara uttalar sig om fysikaliska och andra ”naturvetenskapliga” händelser och därmed inte direkt om handlingar, som ju vanligen identifieras genom mer vardagliga beskrivningar. Men vi kan då lägga till det rimliga villkoret att för varje handling H finns det en ”naturvetenskaplig” händelse H^* , sådan att om H inte hade utförts, så hade inte heller H^* inträffat.

Våra handlingar måste väl antas vara bestämda av vad som händer i våra hjärnor – eller mer allmänt i våra kroppar – och det som händer i våra kroppar är antagligen bestämt av fysiologiska och elektrokemiska processer, tillstånd och händelser; kort sagt, av fysikaliska orsaker. Och dessa fysikaliska orsaker har i sin tur fysikaliska orsaker. Och så vidare. Orsakerna bildar orsakskedjor bakåt i tiden. Våra handlingar bestäms sålunda av dessa orsakskedjor. Handlingarna är resultatet av vad som tidigare har inträffat. Enligt konsekvensargumentet kunde vi därför aldrig ha gjort något annat än det vi faktiskt gjorde.

6. KOMPATIBILISM ENLIGT LEWIS

Enligt David Lewis är däremot determinismen förenlig med att vi ofta kunde ha handlat annorlunda än vi gjorde. Han hävdar att konsekvensargumentet blandar ihop två olika betydelser av att man ”kunde ha falsifierat” ett påstående. Men jag tror att man kan uttrycka hans invändning mot argumentet på ett enklare sätt.

Låt oss återigen se på exemplet från avsnitt 1 ovan. Lewis går med på att om jag hade lyft handen vid t , så hade jag falsifierat $F \& L$. Det är okontroversiellt. Men Lewis menar att jag *kunde* ha falsifierat $F \& L$, eftersom jag i en viss mening ”kunde” ha falsifierat L .

Här måste jag infoga en reservation. Jag vet inte om Lewis verkligen själv anser att vi någonsin kan handla annorlunda än vi gör. Jag tror inte han säger det någonsin. I sin kritik mot konsekvensargumentet vill han bara visa att det är *förenligt med determinismen* att vi ibland kan handla annorlunda. För resonemangets skull godtar han därför – eller låtsas godta – detta. Han säger: ”Jag kan göra något som är sådant att, om jag gjorde det, så skulle en lag ha brutits” (s. 391–2) och ”Jag kan lyfta min hand trots att det är förutbestämt att jag inte kommer att göra detta” (s. 393).

Lewis skiljer mellan en stark och en svag betydelse av ”kan falsifiera” och ”kunde ha falsifierat”. Han menar (eller låtsas mena) att jag i den *svaga* meningen kunde ha falsifierat L .²¹ Den definieras så här:

Låt oss säga att jag kunde ha falsifierat ett påstående, i den svaga meningen, om och endast om jag kunde ha gjort något som är sådant att om jag hade gjort det så skulle påståendet ifråga ha varit falsifierat (fast inte nödvändigtvis av min handling eller av någon händelse orsakad av min handling). (s. 396)

I vårt exempel tänker han sig att det kan gå till på följande sätt:

Om jag hade lyft min hand, så skulle en lag redan ha brutits. Händelseförloppet skulle ha divergerat från det faktiska vid en tidpunkt strax innan jag lyfte min

²¹ Detta håller van Inwagen med om; se van Inwagen 2004, s. 345.

hand, och vid tidpunkten för denna divergens skulle ett lagbrott ha inträffat. Ett sådant lagbrott kallar jag ett divergensmirakel. [...] Men detta divergensmirakel skulle inte ha orsakats av att jag lyfte handen. Snarare skulle kausaliteten ha gått i andra riktningen. (s. 393)²²

Poängen tycks alltså vara den, att jag kunde ha lyft handen i vad jag tidigare, i avsnitt 4 ovan, kallade en *passiv* betydelse av ”kunde”. Med andra ord: det finns en möjlig värld i vilken jag lyfter handen. Detta är knappast kontroversiellt. Det kommer inte i konflikt med determinismen och det kan säkert godtas av van Inwagen och andra inkompatibilister. Men det tycks knappast innebära någon vidare *handlingsfrihet*. Vilken handling jag än utför, så är den enligt Lewis – när han nu låtsas vara determinist – bestämd av det förflutna och naturlagarna.

Men Lewis laborerar som sagt också med det han kallar en *stark* mening av ”kunde ha falsifierat”. Den definieras så här:

Och låt oss vidare säga att jag kunde ha falsifierat ett påstående, i den starka meningen, om och endast om jag kunde ha gjort något sådant att om jag hade gjort det så skulle påståendet ifråga ha falsifierats antingen av min handling i sig själv eller av någon händelse orsakad av min handling. (s. 396)

Lewis anser att om jag kunde ha lyft handen, så kunde jag ha falsifierat *F & L* i den starka meningen. Ty om jag hade lyft handen vid *t*, så hade *F & L* varit falsifierat *av min handling* \bar{n} även om varken *F* eller *L* hade varit falsifierat av min handling.²³ Han skriver:

Om jag kunde ha lyft min hand trots det faktum att determinismen är sann och jag inte lyfte den, så är det sant i den svaga så väl som i den starka meningen att jag kunde ha falsifierat konjunktionen [*F & L*] av det historiska påståendet och lagpåståendet. (s. 396)

Kunde jag då, enligt Lewis, ha lyft min hand? Och vad menar han i så fall med ”kunde”? Det är inte alldeles lätt att säga. Men han tycks inte vilja hävda att jag kunde ha lyft handen i det som jag tidigare kallat en *aktiv* betydelse av ”kunde”.

Här måste man observera att min ”aktiva” betydelse inte är detsamma

²²Lewis menar att handlingen inte *orsakar* miraklet; ty om handlingen hade utförts, så skulle visserligen något mirakel ha inträffat, men inte nödvändigtvis just *detta* mirakel. På motsvarande sätt gäller *inte* i den värld där handlingen utförs, att om den inte hade utförts, så skulle just det aktuella miraklet inte ha inträffat, ty det kan ha varit något annat som inte hade inträffat (jfr Lewis 2000, s. 394).

²³Lewis hävdar att vi däremot inte i den starka meningen kan falsifiera lagar (2000, s. 396). Som Helen Beebe påpekar är detta lite underligt, eftersom Lewis i Humes anda uppfattar lagar som regelbundenheter utan nödvändighet; se Beebe 2003, s. 273–4.

som Lewis ”starka” mening.²⁴ Visserligen kunde jag, i Lewis starka mening, ha falsifierat både *F & L* och påståendet att jag inte lyfter handen – om jag kunde ha lyft handen. Men det verkar som om han menar att jag bara kunde ha gjort *det* i en passiv betydelse. Ty han menar ju att i den värld där jag lyfter handen är det ett divergensmirakel²⁵ – inte jag eller min handling – som falsifierar *L* och därmed *F & L*.

Det verkar med andra ord som om Lewis position inte kommer i konflikt med van Inwagens och andra inkompatibilisters! De senare menar att vi aldrig i en *aktiv* mening kan handla annorlunda, om determinismen är sann. Lewis hävdar att vi i en *passiv* mening ändå kan det. Inkompatibilisterna kan utan vidare gå med på det. Det råder ingen oenighet.

7. KÄNSLAN AV ATT HA KUNNAT GÖRA NÅGOT ANNAT

Kompatibilismen kan te sig attraktiv om man är benägen att tro att våra handlingar är orsaksbestämda men ändå känner sig alldeles övertygad om att man ofta kunde ha gjort något annat än det man gjorde. Denna känsla av att kunna göra något annat har nog de flesta av oss. Men jag för min del tror att den är illusorisk.

Lewis och många andra – även van Inwagen – anser att man kan testa eller underbygga sina uppfattningar genom att försöka passa ihop dem till en koherent enhet, ett så kallat ”reflektivt ekvilibrium”. Men de uppfattningar man inledningsvis har måste då kritiskt granskas. Även om man är mycket starkt övertygad om att man ofta kunde ha gjort något annat, så måste man fråga sig om det verkligen finns något som talar för att det är så. Och det gör det väl inte, när allt kommer omkring.

Man kan lätt *förklara* att vi har känslan av att kunna utföra handlingar vi inte utför. Våldigt många handlingar – dvs. handlingstyper – är ju sådana att vi har förmåga, och ofta tillfälle, att utföra dem. Det vet vi därför att vi ofta faktiskt utför dem. Vi lyfter händer och armar, vi öppnar dörrar, vi dricker vatten, vi tar en promenad, vi läser en bok, osv. Alltså har vi förmåga att utföra handlingar av dessa slag. Vi har också lärt oss att när vi försöker utföra sådana handlingar, så lyckas vi så gott som alltid – utom i vissa undantagsfall när vi t.ex. drabbats av tillfällig förlamning. Förmåga är inte något som kommer och går, när vi har lärt oss att göra en sak, så har vi en relativt konstant förmåga att göra det. Vi har aldrig haft någon

²⁴Man kan säga att Lewis distinktion mellan stark och svag mening gäller ordet ”falsifiera”, medan min distinktion mellan aktiv och passiv betydelse gäller orden ”kunde” och ”kan”.

²⁵Talet om ett ”mirakel” kan vara lite vilseledande. Lewis menar inte att det innebär ett brott mot naturlagarna i den värld där det inträffar; det strider bara mot lagarna i vår värld, men där inträffar det ju inte; se Lewis 1979, s. 469.

upplevelse som tyder på motsatsen. *Därför* tar vi också för givet att vi kan utföra handlingar av dessa slag även i de situationer när vi inte gör det.

Men vid närmare eftertanke inser man att det egentligen inte finns någon evidens för att förmågan ”finns” eller ”är påkopplad” även i de situationer där vi inte utför den aktuella handlingen. Vi har ju aldrig testat förmågan i just sådana fall. Det enda vi vet är att *när* vi testar förmågan, så utför vi också handlingen. (Hittills, i alla fall.)

Så *känslan* av att kunna handla annorlunda är fullt begriplig. Men när man sedan kommer att tänka på att allt som sker kanske är orsaksbestämt – vilket verkar rätt rimligt och i överensstämmelse med våra teorier, även om det nog är rätt svårt att precisera och bevisa – så inser man att vi egentligen inte har någon grund för uppfattningen att vi ofta kunde ha handlat annorlunda, men att vi tvärtom har vissa skäl att tro att denna uppfattning är falsk. Kompatibilister av olika slag har försökt visa att dessa skäl är otillräckliga. Men såvitt jag kan se har de ännu inte lyckats. Och jag tror inte att de någonsin kommer att lyckas. Rätta mig om jag har fel.

LITTERATUR

- Beebe, Helen. 2003. ”Local miracle compatibilism”. *Noûs*, vol. 37, nr 2.
- Hofer, Carl. 2010. ”Causal determinism”. *Stanford Encyclopedia of Philosophy*.
- van Inwagen, Peter. ”The Incompatibility of Free Will and Determinism”. *Philosophical Studies* 27, 1975,
- van Inwagen, Peter. 1983. *An Essay on Free Will*. Oxford: Clarendon Press.
- van Inwagen, Peter. 2004. ”Freedom to break the laws”. *Midwest Studies in Philosophy*, vol. 28.
- van Inwagen, Peter. 2008. ”How to Think about the Problem of Free Will”. *Journal of Ethics*, vol. 12.
- Robert Kane. 1994. ”Free Will: The Elusive Ideal”. *Philosophical Studies*, vol. 75.
- Lewis, David. 1979. ”Counterfactual Dependence and Time’s Arrow”. *Noûs*, vol. 13, nr. 4.
- Lewis, David. 1981. ”Are we free to break the laws?”. *Theoria*, vol. 47, 1981; svensk övers. i *Filosofin genom tiderna: Efter 1950*, utg. av K. Marc-Wogau. L. Bergström och S. Carlshamre, s. 390–398. Thales, 2000.
- G. E. Moore. *Ethics*. 1912.

Gunnar Syrén

Emil – existentialisten från Lönneberga

Ett av de mest betagande avsnitten i den värld som befolkas av Emil, lilla Ida, Alfred och de andra i Astrid Lindgrens Lönneberga, är det där Emil räddar sin vän drängen på gården – ”du och jag, Alfred” – från en hotande död i blodförgiftning, genom att mot alla odds köra honom till doktorn i Mariannelund. Alla som har läst boken *Än lever Emil i Lönneberga* från 1970, vet vad jag menar.

Alfred har skurit sig på kniven när han satt i en trivsamt arbetets gemenskap med husbondfolket och pigan Lina och snidade räfsen på kullen vinterkväll. Såret blir allt elakare, och när det börjar synas röda ränder långt upp på armen, förstår alla att det står illa till med Alfred. Något som författaren låter Krösa-Maja uttala i en fundering som läsaren måste uppfatta som både kusligt och på samma gång distansnerande, eftersom den yttras av den skrockfulla gamla gumman: Hon undrar om dörren till drängkammaren är bred nog för att släppa in en likkista!

Emil blir förtvivlad när hans böner om att man skall köra hans vän till doktorn i Mariannelund blir avvisade. Man nekar honom inte av elakhet eller hårdhet, utan därför att vädret med snö och kyla är sådant att ingen tror att det går att ta sig fram på vägen.

Då sker det. Emil står inte ut med tanken att Alfred skulle lämna honom, utan spannar med stor möda hästen för en släde, och släpar med ännu större besvär Alfred till släden. Sedan kör han mot Mariannelund, kämpar mot snömassorna, hamnar i diket, blir räddad av en bonde, kör vidare, men fastnar till slut i snön. När allt hopp tycks vara ute, kommer en snöplog och banar väg, och Emil lyckas köra ända fram till doktors mottagning där Alfred tas om hand. Att han törs se på när Alfreds finger skärs upp, finner doktorn naturligt; har han klarat att ta sig dit, så klarar han naturligtvis också detta.

Emil utför i en förtvivlad situation det till synes omöjliga. Helt enkelt därför att han *måste* göra det. Och därför att han som den ende insåg inte bara situationens allvar, det gjorde de nog alla, men att det var *nu* som gällde, och att det fanns en enda möjlighet och att denna *måste* prövas.

När jag första gången läste den bok där historien berättas, vilket jag

gjorde som högläsning för min lille son en gång i tiden, hade jag nyligen för mig själv läst en bok av skenbart helt annan karaktär, nämligen Albert Camus' *Pesten* från 1947. Men jag slogs ändå av att Astrid Lindgren i sin berättelse gestaltar moraliska och existentiella frågor av allmänmänsklig betydelse, desamma som Camus hanterar i sin roman, och att bådas genialitet visar sig i att de kan skriva om dessa inom helt realistiska ramar.

Pesten utspelar sig i ett kolonialt Algeriet, närmare bestämt i staden Oran, där plötsligt ett stort antal råttor med blödande nosar visar sig på gatorna. Befolkningen förvånas och förskräcks, men varken den eller de styrande vill förstå vad detta kan komma att betyda, nämligen att en svår sjukdom kan komma att drabba staden, en pest. När så småningom sanningen ändå går upp för dem, isoleras Oran från omvärlden.

En av de första som inser allvaret i situationen är han som skall bli bokens hjälte, doktor Rieux, en klok för att inte säga lärd läkare som vet mycket om pestens härjningar i äldre tider, men som nu måste kämpa för att få omgivningen att förstå vad som håller på att ske.

Några andra för handlingen och för den moraliska diskussionen inte oviktiga personer i boken, är journalisten Rambert som genom stadens isolering blivit skild från sin älskade, och Monsieur Tarrou som är något gåtfull, men som uppenbarligen har något ouppklarat med sig i sitt livsbagage.

Albert Camus är den kanske viktigaste representanten i vår tid för den filosofi som brukar kallas den existencialistiska, och som annars räknar namn som Kierkegaard och Sartre. Kierkegaards filosofi är ibland motsägelsefull, och dess grund är en kristendom som inte alla bekänner sig till. Och Sartre har korrumperats i sin förkunnelse av att även vad vi nu känner som mördarregimer kan ha ett berättigande som byggstenar i något som skulle vara större än vi, nämligen Historien.

Camus däremot står för mig med böcker som *Pesten* och kanske också *Främlingen* som alltmer betydelsefull i en tillvaro som i kampen mot fundamentalism och religiös fanatism söker moraliska hållpunkter när kristendomen förlorat sin plats som det självklara moraliska rättesnöret.

Men tillbaka till Emil i Lönneberga. Jag hörde en gång Gunnar Brandell, lysande i sin kunskap om europeiskt 1900-tal, sammanfatta existencialismens budskap i orden: Förtvivla, men förtvivla icke! Astrid Lindgrens Emil ser förtvivlan och uppgivenhet runt omkring sig, och känner den naturligtvis också själv när han står vid Alfreds sjuksäng. Men till skillnad från de andra låter han inte denna förtvivlan utmynna i handlingsförlamning, utan vänder den i dess motsats och utför det till synes omöjliga.

Som invånarna i Oran är Emil och hans familj instängda, en instängdhet som man kan men inte behöver tolka symboliskt eller filosofiskt. Det

är en rent fysisk instängdhet, påminnande om man så vill om allas vår instängdhet i vår tillvaro; intressant vid en jämförelse är att i *Pesten* sker ingen utbrytning ur denna instängdhet; kampen mot pesten sker *inom* de gränser som karantänen påbjuder. Medan det hos Astrid Lindgren sker en faktisk utbrytning ur instängdheten när Emil ger sig iväg!

Mot dem som vägrar att inse faran hos Camus, och mot dem som låter sig bindas av sin förtvivlan och sin, som det skall visa sig, falska insikt hos Astrid Lindgren, står insikten hos Rieux och Emil att handling är nödvändig. Emil äger naturligtvis inte Rieux' kunskaper, men visar sig ha större insikt än samtliga de övriga i hans familj. Med risk för att bli alltför subtil, kunde man säga att det motstånd mot att förstå som finns runt läkaren i *Pesten* motsvaras av ett liknande motstånd hos människorna runt Emil. Och att detta motstånd är släkt med det "falska medvetande" som den som existentialismen talar om med hänvisning till Sartres *Varat och intet* från 1943.

Doktor Rieux' liksom Emils insikt leder till handling, ett centralt begrepp inom den existentialistiska filosofin. Vi är alla dömda till frihet, men denna frihet ger oss inte bara ångest, utan gör det också möjligt för oss både att välja och att välja handlingen. Och detta val har vi även i synbarligen helt hopplösa situationer; Sartres exempel att den dödsdömde på väg till galgen i alla fall kan sparka bödeln på benet, är kanske ett övertydligt sådant. Läkaren hos Camus väljer att kämpa mot pesten trots risken för att bli smittad och dö, och i den kampen väljer också journalisten Rambert sida och stannar, trots möjligheten att bli smugglad ut ur stan för att återförenas med sin älskade.

Det är människans elementära villkor som gestaltas i båda böckerna. Emils koleriske pappa blir inte arg när Emil vredgat talar till honom på ett sätt som annars skulle ha renderat honom ett besök i snickarbo. För även om han inte ser den möjlighet som sonen ser, så förstår han att just nu är allt annat än Alfreds infekterade tumme betydelselöst. Journalisten Rambert ser inte framåt, grubblar inte över framtiden, lika litet som Emil bekymrar sig för annat än att meter för meter komma närmare Mariannelund och räddningen. Det är *nu* som gäller. Det är en "moralisk minimalism" och en "värderealism" av det slag som Lars Gustafsson tillskriver Camus i en insiktsfull essä från 1997.

Emil i Lönneberga är en barnbok och som en sådan tydligare än *Pesten* vad gäller det moraliskt riktiga i den handling som hjälten väljer att utföra. Den annars som problembarn betraktade Emil hyllas enstämmigt när han kommer hem, och inte bara av familjen. Dessutom har doktorn i Mariannelund skrivit ett brev som prisar gossens mod och företagsamhet, och som vi alla vet så fick Emil med tiden ett politiskt förtroendepdrag, en framtid som också antyds när han kommit hem med Alfred.

Boken är alltså, som varje god barnbok bör vara, optimistisk. Men även *Pesten* slutar med optimistiska ackord: huvudpersonen ger inte upp men utför sitt yrkes kärleksgärningar, Rambert visar sig vara en hjälte även han och dessutom klarar han skilsmässan från sin älskade bättre än han trodde att han skulle göra. Monsieur Tarrou är något mer svårbedömd, men vill man också låta honom indikera hoppfullhet, kunde man säga att han anpassar sig till situationen eftersom den liknar det liv han alltid levat.

Existentialismen formuleras gärna i korta, slående imperativer; någon har jag använt ovan. Men starkare än varje imperativ tycks mig Emils sakliga konstaterande när han med krafter som en ox skottar snö för att komma fram med sin släde. Han vänder sig då till sin häst Lucas och förklarar: ”Man blir så stark när man *måste*.”

Sista ordet skall dock gå till Emils lillasyster Ida. När familjen efter Emils bragd sitter och utmålar hans framtid, reserverar sig fadern mot alltför högtflygande planer. Han erkänner dock att det nog kan bli något stort av Emil om han får leva och ha hälsan – och om Gud vill. Modern nickar instämmande: Ja, om Gud vill! De har inte förstått någonting; nyss har det skett något märkligt därför att Emil gjorde ett fritt och modigt val, men ingen runt honom har insett detta till fullo. Ingen utom lilla Ida. Hon sammanfattar vad som skett och vad detta kommer att innebära för framtiden i ett lakoniskt: ”Och om Emil vill.”

REFERENSER

- Camus, Albert. 1942. *L'Étranger*. Sv. övers. *Främlingen*, övers. Jan Stolpe. Stockholm, 2009.
- Camus, Albert. 1947. *La Peste*. Sv. övers. *Pesten*, övers. Elsa Thulin. Stockholm, 2003.
- Gustafsson, Lars. 1997. ”Camus’ *Sisyfos*’ och *Pestens* budskap”. I densammes *Vänner bland de döda: Essäer om litteratur*, s. 103–111. Stockholm.
- Lindgren, Astrid. 1970. *Än lever Emil i Lönneberga*. Stockholm.
- Sartre, Jean-Paul. 1943. *L'Être et le néant: Essai d'ontologie phénoménologique*. Paris. Sv. övers. *Varat och intet: I urval och med inledning av Dag Østerberg*, övers. Richard Matz och Suzanne Almqvist. Göteborg, 1992.

Kunskapsbegreppet och ett svar till skeptikern

De skeptiska argumenten föranleder visst tvivel på att vi kan ha kunskap över huvud taget. En rimlig kunskapsanalys bör kunna bemöta dessa argument för att visa att kunskap är möjlig. I denna artikel utreder jag hållbarheten hos två skeptiska argument givet Nozicks kunskapsanalys. Först diskuterar jag *argumentet med hjärnor i näringslösning*, och visar att det inte håller givet Nozicks analys. Sedan visar jag att ett skeptiskt argument som Jonathan Dancy framlägger mot Nozicks analys inte visar att vi inte kan ha kunskap. I artikeln diskuterar jag även kunskapsbegreppet som sådant, och kommer därmed in på debatten om *internalism* kontra *externalism*. Slutligen hävdar jag att vi använder ordet ”kunskap” på flera olika sätt (även sett till propositionell kunskap), och att vi därför bör förändra vårt språkbruk.

1. NOZICK OCH SKEPTICISM

Robert Nozick analyserade 1981 kunskap som *sann tro som spårar sanningen*. Han menade att en person S vet att p om och endast om följande fyra villkor är uppfyllda:

- (1) p är sant
- (2) S tror att p
- (3) Om p inte vore sant, skulle S inte tro att p
- (4) Om p vore sant, skulle S tro att p .¹

¹Villkor (3) och (4) är s.k. ”konjunktiva konditionaler”, vilket innebär att de är sanna om och endast om det som ligger närmast till hands om förledet vore sant är att också efterledet är sant. Uttrycket ”ligger närmast till hands” är vagt och behöver definieras, vilket Nozick gör i termer av möjliga världar. Kortfattat kan man säga att vi ska undersöka huruvida efterledet är sant i de möjliga världar där förledet är sant som ligger närmast (dvs. mest liknar) den faktiska världen. Utöver ovan nämnda villkor presenterar Nozick villkor som tar hänsyn till genom vilken metod S kommer att tro p , men för enkelhets skull använder jag mig i denna artikel bara av (1)–(4). Se Gardelli (2010) för närmare förklaring av villkoren och av Nozicks analys.

Eftersom de skeptiska argumenten behandlar vår möjlighet till *kunskap*, är det relevant för både deras innehåll och hållbarhet hur vi analyserar ”kunskap”. Det kan vara så att en skeptisk tes är hållbar givet en viss kunskapsanalys, men inte givet en annan. Nozick själv säger att han angående sin analys inte ämnar motbevisa skeptikern, ej heller övertyga denne. Om man ska kunna övertyga skeptikern, måste det finnas en öppenhet och en klarhet hos denne som ingen annan kan garantera. Därför vill Nozick endast visa på kunskapens möjligheter; Nozick försöker att ta skeptikerns argument på allvar och ändå övertyga *sig själv* om att han kan ha kunskap.

1.1. *Hjärnor i näringslösning*. Kan du givet Nozicks analys veta att du inte är en hjärna i ett kar med näringslösning som via en dator matas med upplevelser (exakt identiska med dem du har nu) av ett antal vetenskapsmän? Anta att det är sant att du inte är en hjärna i näringslösning och att du tror att du inte är en sådan. Då är (1) och (2) uppfyllda. Är (3) uppfyllt? Nej, ty om du vore en hjärna i näringslösning, skulle du ändå tro att du inte var det, eftersom du *ex hypothesi* skulle uppleva precis det som du nu faktiskt upplever. Alltså kan du inte veta att du inte är en hjärna i näringslösning, givet Nozicks analys. Men kan du då veta att du är en människa som sitter (sitt gärna ner!) och läser en artikel? Låt oss undersöka om villkoren (1)–(4) kan vara uppfyllda. Anta att det är sant att du är en människa som sitter och läser en artikel och att du tror att du är en människa som gör det, så att (1) och (2) är uppfyllda. Om det då inte vore sant att du var en människa som satt och läste en artikel, ligger det närmast till hands att du skulle vara en människa som stod eller låg och läste en artikel, eller precis hade tagit en paus, och då skulle du inte trott att du satt och läste. Alltså är (3) uppfyllt. Dessutom är (4) uppfyllt, eftersom det ligger närmast till hands om du är en människa som sitter och läser en artikel att du tror att du är en människa som gör det. Alltså har du enligt Nozicks analys kunskap om att du är en människa som sitter och läser en artikel (om du nu sitter). Det som gör att man inte kan säga att du inte vet på grund av att du skulle kunna vara en hjärna i näringslösning, är att detta inte är ett relevant alternativ i den situation du faktiskt befinner dig i (om det, som vi har antagit, är sant att du är en människa som sitter och läser en artikel), eftersom en möjlig värld där du är en hjärna i näringslösning är väldigt avlägsen från den faktiska världen där du är en människa som läser en artikel.

Att du inte kan veta att du inte är en hjärna i näringslösning, medan du kan veta att du är en människa som sitter och läser en artikel, är ett brott mot (eller motbevis av) att kunskap är sluten under känd logisk implikation (”closed under known logical implication”) givet Nozicks analys. Att kunskap är sluten under känd logisk implikation betyder,

enligt Dancy (1985), att om jag vet att p , och jag vet att p implicerar q , så vet jag att q . (Man bör tillägga, vilket Dancy inte gör, att det finns ett uppenbart motexempel mot denna formalisering av principen. Motexemplet är att jag vet att p och jag vet att p implicerar q , men jag ändå inte tror att q . Om jag inte tror att q , kan jag nämligen inte veta att q . Vi ska därför diskutera om principen håller givet att den också förutsätter att jag tror att q .) Men givet att (1)–(4) definierar kunskap, kan det gälla att jag vet att p (att jag är en människa som sitter och läser), och jag vet att p implicerar q (att om jag är en människa som sitter och läser så är jag inte en hjärna i näringslösning), men att jag ändå inte vet att q (att jag inte är en hjärna i näringslösning). Nozick ser det som en förtjänst att kunskap inte är sluten under känd logisk implikation givet hans analys, eftersom många skeptiska argument bygger på denna princip, som Dancy kallar "the principle of closure". Man kan kräva av Nozick att han ska förklara varför en så pass intuitiv princip som denna inte håller. Nozick konstaterar att det för att man ska ha kunskap krävs att ens trosföreställning på något sätt varierar med sanningen. Givet en kunskapsanalys som innehåller villkor som uppfyller detta, kommer, enligt Nozick, kunskap inte vara sluten under känd logisk implikation. Villkor (3) är ett exempel på en sådan typ av variation, och det har just denna effekt. Det följer nämligen inte ur att (3) är uppfyllt och att S vet att p implicerar q , att om q inte vore sant så skulle S inte tro att q . Anledningen till att detta inte följer är att de möjliga världar som skulle föreligga om p vore falskt respektive om q vore falskt inte nödvändigtvis är desamma, även om p implicerar q . (Ett exempel som illustrerar detta är när " p " står för "S är en människa som sitter och läser en artikel" och " q " står för "S är inte en hjärna i näringslösning".) Om villkor (3) är nödvändigt för kunskap, är alltså kunskap inte sluten under känd logisk implikation.

Nozick påstår att skeptikern är inkonsistent i frågan om det är nödvändigt för kunskap att tron varierar med sanningen. För att visa att du inte vet att du inte är en hjärna i ett kar med näringslösning, använder sig skeptikern av ett argument som tycks förutsätta att en viss typ av variation hos tron är ett nödvändigt villkor för kunskap. Argumentet lyder: "Om du vore en hjärna som flöt i ett kar, skulle du ändå tro att du inte var det, så du vet inte att du inte är det" (Nozick 1981, s. 209, min övers.). När skeptikern sedan sluter sig till att du inte vet att du är en människa som läser en artikel, använder sig skeptikern av "the principle of closure", som inte är valid om man förutsätter att kunskap kräver denna typ av variation med sanningen hos kunskapssubjektets trosföreställningar. Om man accepterar Nozicks analys håller man med skeptikern om att vi inte kan veta att vi inte är hjärnor i näringslösning. Det intressanta i vardagen är dock (åtminstone för dem där vardagen

inte innehåller alltför djupa studier i epistemologi) om vi kan veta vad vi gjorde i går och vad olika handlingar kommer få för konsekvenser, vilket Nozicks analys tillåter.

1.2. "The argument from error." Jonathan Dancy presenterar ett skeptiskt argument, "the argument from error", som han menar utgör ett problem för Nozicks analys genom att det är giltigt även givet denna analys (Dancy 1985). Utgångspunkten för argumentet är att vi tidigare har begått (eller hade kunnat begå) misstag i situationer där vi har hävdats att vi har vetat. Utifrån detta kan man sedan argumentera på två olika sätt, och Dancy blandar de olika varianterna friskt. Det går att se argumentet antingen som ett argument för att vi, eftersom vi tidigare har begått misstag, inte (utan att bryta mot en viktig princip) kan *hävda* att vi har kunskap, eller som ett argument för att vi inte kan *ha* kunskap.

["The argument from error"] kan inte förkastas på grund av att det argumenterar från det faktum att du kanske har fel till slutsatsen att du inte vet. Ett sådant argument skulle vara ogiltigt, som Nozick kan visa. Men argumentet går inte till väga på detta felaktiga sätt, utan indirekt via principen om universaliserbarhet. (Dancy 1985, s. 46, min övers.)

Detta citat illustrerar att Dancy tänker sig att man ska använda sig av en analogi till R. M. Hares universaliserbarhet i argumentet. Hares universaliserbarhet säger att vi motsäger oss själva om vi faller olika moraliska omdömen om två, sett till alla relevanta egenskaper, identiska situationer. En relevant egenskap skulle, om vi faller ett moraliskt omdöme om en handling, kunna vara vilka konsekvenser handlingen får för dem som påverkas av den. Exempel på egenskaper som *inte* är relevanta är vid vilken tidpunkt och plats som handlingen äger rum (om inte detta påverkar handlingens konsekvenser eller dylikt). Man kan alltså tala om "universaliserbara" respektive "icke universaliserbara" egenskaper, där tidpunkt och plats för en handling tillhör kategorin "icke universaliserbara egenskaper". Liknande distinktioner skulle kunna göras angående en situation där någon faller ett omdöme om huruvida någon (t.ex. hon själv) har kunskap om någonting. Principen om universaliserbarhet kan då tillämpas på *hävdande* av kunskap, men inte på samma sätt på *innehav* av kunskap (ty för att veta någonting behöver man inte fälla något omdöme om att man vet), vilket Dancy stundtals påstår att argumentet behandlar:

[. . .] det är något riktigt med ett argument som medför, som "the argument from error" gör, att för varje slags anledning till att du inte vet att du inte är en hjärna i näringslösning, vet du inte, t.ex., att *The Times* kommer att publiceras i morgon, ej heller huruvida du sitter och läser. (Dancy 1985, s. 45, min övers.)

Eventuellt kan Dancy med "the argument from error" visa att vi aldrig kan *hävda* att vi vet utan att bryta mot en epistemologisk variant av Hares moraliska universaliserbarhet. Han kan som jag nämnt ovan dock inte på samma sätt utifrån universaliserbarheten visa att vi inte kan *ha* kunskap, eftersom universaliserbarheten behandlar fällandet av omdömen, och vi inte behöver fälla några omdömen om att vi vet för att vi ska ha kunskap (åtminstone inte enligt Nozicks kunskapsanalys). Inte heller kan man sluta sig direkt från att vi aldrig (utan att bryta mot universaliserbarheten) kan hävda att vi har kunskap till att vi aldrig kan ha kunskap.

Låt oss se om Dancy kan argumentera enligt följande för att visa att vi inte kan ha kunskap: jag har haft fel i situationer som ur mitt perspektiv är identiska med en aktuell situation i alla relevanta avseenden, och därför kan jag inte nu veta att det jag tror är sant verkligen är sant; jag har förut haft falska trosföreställningar om vädret, och alltså kan jag inte nu veta hur vädret kommer att bli. Genom att peka på att jag har haft fel i liknande situationer kan man visa att jag inte kan vara helt säker på att jag har rätt, att jag inte kan garantera med full säkerhet att det jag tror är sant verkligen är sant. Att påstå att detta innebär att jag inte kan veta, är att påstå att ett nödvändigt villkor för kunskap är att kunskapssubjektet kan garantera med full säkerhet att hon vet. Om ett sådant villkor antas blir kunskap väldigt sällsynt, och Dancy skriver uttryckligen (se ovan) att det inte är denna väg han vill gå. Dancy lyckas alltså inte finna något stöd för att vi inte kan ha kunskap hos "the argument from error". Ändå tänker han sig att argumentet visar just detta, vilket framgår av att han i slutet av sin diskussion av argumentet sammanfattar det på följande sätt:

[*"The argument from error"*] argumenterar för [...] att eftersom vi har gjort misstag, [...] eller skulle göra misstag i tänkbara liknande situationer, så vet vi inte nu. (Dancy 1985, s. 14, min övers.)

Även om argumentet inte visar att vi inte kan ha kunskap, kanske vi måste ge upp vår möjlighet att hävda att vi har kunskap, förutsatt att vi accepterar Dancys epistemologiska universaliserbarhet och att hans argumentation är riktig. Detta utgör dock inget större problem för Nozicks analys, eftersom vi fortfarande verkar kunna hävda att vi har goda skäl att tro att vi har kunskap, eller att vi troligtvis vet. I vilket fall så inverkar inte argumentet på Nozicks målsättning att visa på kunskapens möjligheter, eftersom Dancy med "the argument from error" inte lyckas visa att vi inte kan ha kunskap givet Nozicks analys.

2. KUNSKAPSBEGREPPET

2.1. *Internalism och externalism.* De första två villkoren i Nozicks analys återfinns i många kunskapsanalyser, inte minst i den s.k. ”klassiska kunskapsanalysen”, som presenterades av Platon i dialogen *Teaitetos*. Sedan Edmund Gettier 1963 visade att den klassiska analysen inte är tillräcklig har man främst diskuterat om man ska ändra den klassiska analysens tredje villkor, som är att trosföreställningen ska vara rättfärdigad, eller om analysen ska kompletteras med ytterligare villkor. Ytterligare ett alternativ är att som Nozick ersätta rättfärdigandevillkoret med ett eller flera andra villkor. En fråga som har varit omdebatterad de senaste decennierna är om kunskapsbegreppet är *internalistiskt* eller *externalistiskt*. Jag ska här lite förenklat redogöra för distinktionen mellan internalism och externalism. Ett sätt att definiera internalism är enligt följande: en kunskapsanalys är internalistisk om och endast om den kräver att allt som behövs för att det tredje villkoret ska uppfyllas (eller snarare för att alla villkor utöver den klassiska analysens två första villkor ska uppfyllas) måste vara *kognitivt tillgängligt* för personen i fråga (Bonjour 2002, s. 234). Analysen är externalistisk om detta inte är uppfyllt, dvs. om den tillåter att något som krävs för att det tredje villkoret ska uppfyllas inte måste vara kognitivt tillgängligt för personen i fråga. Den är alltså externalistisk om den tillåter att någon nödvändig faktor hålls extern för kunskapssubjektets kognitiva perspektiv. Nozicks kunskapsanalys är ett exempel på en externalistisk kunskapsanalys. Jag måste nämligen inte ha kognitiv tillgång till de faktorer som behövs för att (3) och (4) ska uppfyllas. Dessa faktorer har att göra med hur yttervärlden är beskaffad, eftersom detta avgör vilka konjunktiver som är sanna, och hur yttervärlden ser ut är inte något som jag med nödvändighet har kognitiv tillgång till.

En viktig fråga att ställa sig är: vad menar vi när vi säger att någon har kunskap? Jag är beredd att argumentera för att vi ibland menar en sak, och ibland en annan. Vi använder ordet ”kunskap” i flera olika bemärkelser. Dessutom tillämpar olika människor ordet på olika situationer. Jag har i ”Om Nozicks kunskapsanalys” (Gardelli 2010) argumenterat för att våra intuitioner kring om någon har kunskap eller ej inte ger samma svar i alla situationer. Dessutom har vi intuitioner om vilken kunskapsanalys som är riktig, där vissa intuitioner talar för att den klassiska kunskapsanalysen (tolkad internalistiskt) är riktig, medan andra intuitioner talar för att Nozicks analys är riktig. Detta ger upphov till att vi, om vi ska få ett kunskapsbegrepp som är gemensamt för oss alla, måste förändra vårt språkbruk. Ett sätt att göra detta är att hitta en analys som någorlunda väl stämmer överens med de flestas intuitioner samtidigt som den har andra praktiska konsekvenser, och sedan konsekvent börja tillämpa denna. Om man definierar kunskap i enlighet med No-

zicks analys, fångar detta våra externalistiska intuitioner, medan begrepp som "rättfärdigande" och eventuellt "evidens" tolkat internalistiskt kan användas när vi vill beskriva att en person har en viss kognitiv koppling till sin trosföreställning. Man kan i så fall tala om "rättfärdigad kunskap" eller "välgrundad kunskap" i de fall där man vill tala om att någon har en sann trosföreställning som spårar sanningen, och dessutom har en viss typ av kognitiv koppling till sin trosföreställning.

2.2. *Kunskap och sannolikhet.* En av de mest kontraintuitiva konsekvenserna av Nozicks kunskapsanalys är att jag kan veta att jag var på kalas på jorden i går (jag överlåter åt läsaren att visa att det är möjligt för mig att veta att jag var på kalas på jorden i går, givet att jag var det), men jag inte kan veta att jorden fanns i går (ty för detta är inte (3) uppfyllt; om jorden inte hade funnits i går, hade den skapats i dag med alla fossil och alla människors "minnen", och då hade jag fortfarande trott att jorden fanns i går). Det problematiska är att det inte kan vara sant att jag var på kalas på jorden i går utan att det är sant att jorden fanns i går. Däremot kan det senare vara sant utan att det förra är det. Därför är det minst lika sannolikt (skulle man kunna säga) att det senare är sant som att det förra är sant. Detta rimmar illa med att det är omöjligt att veta att jorden fanns i går men att det är möjligt att veta att jag var på kalas på jorden då. Vad man ska göra i denna situation är kanske att skilja på sannolikhet och kunskap. Vi behöver inte ordet "kunskap" för att tala om att någonting är sannolikt sett ur någons perspektiv, eftersom vi har "(subjektiv) sannolikhet" för att beskriva detta.

2.3. *Att veta att man vet.* En intressant egenskap hos Nozicks analys är att vi givet denna kan veta att vi vet någonting. (Detta lämnar jag som övning åt läsaren att visa.) Är det rimligt att jag vet att jag vet att jag åt gröt till frukost i morse på jorden, men att jag inte kan veta att jorden fanns då? Som svar på denna fråga måste jag ge en motfråga: vad menar vi när vi säger att jag vet att jag vet någonting? Menar vi inte egentligen något i stil med att jag är *säker* på att jag vet någonting? Enligt Nozicks analys kan jag inte vara (helt) säker på att jag vet att jag åt gröt i morse, eftersom det är möjligt att jorden inte fanns då och det är möjligt att jag hela tiden har varit en hjärna i näringslösning. Kanske borde vi alltså i vårt språkbruk skilja på när vi menar att någon har kunskap, dvs. (om vi accepterar Nozicks analys) att hon har en sann trosföreställning som spårar sanningen, och när vi menar att någon är säker på någonting. Om vi gör detta, inser vi att vi inte så ofta vill tala om att någon vet att hon vet någonting, givet Nozicks kunskapsanalys. Detta skulle nämligen svara mot att hon har en sann trosföreställning om att hon har en sann trosföreställning som spårar sanningen, och att hon dessutom inte hade trott att hon hade en sann trosföreställning som spårade sanningen om

hon inte hade haft det, samt att det ligger närmast till hands att hon tror att hon har det så länge hon har det. Vi har inte särskilt lätt att intuitivt avgöra om detta är sant eller inte. Det är i stället mer intressant att tala om att hon har goda skäl för att tro att hon vet, eller om att det är sannolikt att hon har rätt i sin trosföreställning. Så om vi accepterar Nozicks kunskapsanalys, är det viktigt att vi använder ord som ”säker”, ”rättfärdigad” eller ”sannolik” för att fånga de meningar av ordet ”kunskap” som skiljer sig från den mening som Nozicks analys tar fasta vid. Kanske kan vi då dessutom lösa debatten om internalism kontra externalism genom att konstatera att vi använder ordet ”kunskap” på flera olika sätt, och att vi ur detta bör urskilja flera olika begrepp som vi kan benämna med olika termer.

LITTERATUR

- Bonjour, L. 2002. ”Internalism and Externalism”. I *The Oxford Handbook of Epistemology*, red. P. Moser, s. 234–263. Oxford: Oxford University Press.
- Dancy, J. 1985. *An Introduction to Contemporary Epistemology*. Oxford: Blackwell Publishers.
- Gardelli, T. 2010. ”Om Nozicks kunskapsanalys”. *Filosofisk tidskrift*, nr 2, s. 19–28.
- Gettier, E. 1963. ”Is Justified True Belief Knowledge?”. *Analysis* 23 (6): 121–123.
- Hare, R. M. 1981. *Moral Thinking: Its Levels, Method, and Point*. Oxford: Oxford University Press. Sv. övers. *Moraliskt tänkande: Dess nivåer, metod och syfte*. Stockholm: Thales, 1994.
- Moser, P. 2002. *The Oxford Handbook of Epistemology*. Oxford: Oxford University Press.
- Nozick, R. 1981. *Philosophical Explanations*. Cambridge, Mass.: Harvard University Press.

Rättfärdigande och kritisk prövning

Nyligen hävdade jag i *Filosofisk tidskrift* (2009b) att det är förnuftigt att tro på hypoteser som klarat av sträng kritik. Mitt försök att därigenom finna en kritisk medelväg mellan skepticism och dogmatism har kritiserats av Ingemar Nordin och Lars Bergström. Nordin (2010) frågar ”Varför måste vi tro på något?” och menar att det är överflödigt att införa en princip för förnuftig tro i den kritiska rationalismen och finner dessutom en sådan princip svår att förena med en odogmatisk inställning. Bergström (2010) betvivlar att jag lyckats övervinna skepticismen och menar att min princip för förnuftig tro leder till logiska problem.

Här tänker jag försöka visa att den odogmatiska inställning som Nordin utgår från när han betonar att allt vetande är felbart leder till frågan vilka av alla felbara hypoteser vi skall hålla för sanna. Utan ett svar på den frågan har vi inte övervunnit skepticismen. Bergströms kritik berör grundläggande frågor om hur man skall rättfärdiga hypoteser och positioner. Jag kommer att argumentera för att varje försök att rättfärdiga hypoteser och positioner med tillräckliga skäl leder till den typ av logiska problem som Bergström anför mot min position. Men om vi försöker *kritiskt pröva* hypoteser och positioner i stället för att *rättfärdiga* dem, så uppträder inte dessa problem. Även om vi inte kan ge absoluta rättfärdiganden av hypoteser och positioner, så är det möjligt att ange när det är förnuftigt att hålla dem för sanna.

1. FELBARHET OCH FÖRNUFTIG TRO

Enligt min princip är det förnuftigt att tro på de hypoteser som klarat sträng prövning. Samtidigt hävdar jag att alla hypoteser, också de som klarat av sträng prövning, är felbara och kan vederläggas vid nästa prövning. Nordin (2010, s. 36) menar att det är svårt att samtidigt både tro på och vara skeptisk till en hypotes, och undrar om det kan vara förnuftigt att vara splittrad på detta sätt. Nordin uppfattar min princip för förnuftig tro som en princip för absolut tro och bergfast övertygelse av det slag som ibland rekommenderas i religiösa sammanhang. Men med ”tro”

menar jag inte det, utan ett felbart försanthållande som kan revideras efter ytterligare prövning. Ofta säger vi att vi inte vet något säkert, men att vi tror att något förhåller sig på ett visst sätt, att vi med nuvarande information håller en viss uppfattning för sann. Sådana försanthållanden kan vara förnuftiga eller oförnuftiga beroende på den nuvarande informationen. Insikten att allt vetande är felbart är fullt förenlig med att vi för tillfället håller vissa trosföreställningar för sanna. I själva verket gör vetandets felbarhet det till ett viktigt problem att fastställa vilka av alla felbara trosföreställningar som det för närvarande är förnuftigt att hålla för sanna givet nuvarande information. Det är visserligen möjligt att hävda en fullständig skepsis och påstå att det inte finns några trosföreställningar som det är förnuftigt att hålla för sanna, som t.ex. David Miller (1999) gör, men det är svårt att förena en sådan uppfattning med en kritisk rationalism, som betonar inte bara vetandets felbarhet, utan också förnuftets roll i kunskapsökandet. En radikal skepsis innebär att vi avstår från att med det kritiska förnuftets hjälp försöka finna en medelväg mellan dogmatism och skepticism. Nordin (2010, s. 40) misstas sig när han uppfattar min princip för förnuftig tro som en oskön fläck på den kritiska rationalismens vapensköld. Utan en sådan princip står den kritiske rationalisten utan någon vapensköld alls när han konfronteras med relativister och dogmatiker.

2. GOODMAN'S PROBLEM

Nordins nästa invändning mot min princip för förnuftig tro är inspirerad av Goodmans (1955) problem. Nordin (2010, s. 36) frågar om det är lika förnuftigt att tro att alla smaragder är gröna som att tro att alla smaragder är "gröda" (definierat som gröna fram till en viss tidpunkt i framtiden och därefter röda). All hittillsvarande erfarenhet är förenligt med båda hypoteserna och bekräftar enligt Goodman och Nordin båda hypoteserna lika mycket. Därför menar Nordin att den ena hypotesen hittills klarat av sträng kritik lika bra som den andra och att det enligt min föreslagna rationalitetsprincip för närvarande är lika förnuftigt att tro på den ena hypotesen som på den andra. Samtidigt tycker de flesta av oss att det är oförnuftigt att tro att alla smaragder är gröda.

Är det verkligen så att hypotesen att alla smaragder är gröda vid nuvarande tidpunkt klarat av sträng prövning lika bra som hypotesen att alla smaragder är gröna? För att strängt pröva en hypotes måste man pröva de konsekvenser av hypotesen som är mest osannolika vid prövningsögonblicket och som utgör hypotesens svagaste punkter. Den mest osannolika konsekvensen av hypotesen att alla smaragder är gröda är att en smaragd efter en viss framtida tidpunkt skulle vara röd. Den strider mot

den väl bekräftade bakgrundskunskapen att alla smaragder är gröna. Nu har hypotesen ”alla smaragder är gröda” egenskapen att det hittills inte varit möjligt att pröva dess mest osannolika konsekvenser. Det innebär att hypotesen inte kunnat prövas strängt. För närvarande finns det därför inga förnuftiga skäl att tro att alla smaragder är gröda, vilket är helt i enlighet med vår intuition.

3. ÄR PRINCIPEN FÖR FÖRNUFTIG TRO ÖVERFLÖDIG?

Nordin anser att kritiska rationalister inte behöver en princip för förnuftig tro för att kunna handla praktiskt eller för att undvika kunskapsrelativism. Som exempel anför han det praktiska problemet att starta en bil. De flesta av oss tror att bilen kommer att starta om vi vrider om startnyckeln. Nordin anser att det är förnuftigt att vrida om startnyckeln för att starta bilen, eftersom den handlingen ”tidigare visat sig fungera i praktiken”. I jämförelse ”med andra otestade (eller testade och ej fungerande) handlingar har vi ett fullgott skäl att vara mindre skeptiska till *h*” (handlingen att vrida om startnyckeln) eftersom vi falsifierat den ”ultraskeptiska hypotesen” att *h* aldrig fungerar (Nordin 2010, s. 38). Nordin förutsätter här att när det gäller praktiskt handlande, så har vi goda skäl att vara mindre skeptiska till den hypotes som tidigare fungerat i praktiken. Han menar därför att han har goda skäl att förvänta sig att bilen kommer att starta när han vrider om startnyckeln. Detta är ett exempel på ett försanthållande som Nordin finner förnuftigt i ljuset av tidigare erfarenheter. Också Nordin förutsätter, om än stillatigande, en princip för förnuftigt försanthållande. Men han formulerar inte någon princip för sådana försanthållanden, varför hans position inte är så öppen för kritik som min.

Nordin (2010, s. 39) argumenterar på liknande sätt när det gäller konsekvenserna av Goodmans problem för praktiskt handlande. För praktiskt handlande anser han att det är förnuftigt att vara något mindre skeptisk till hypotesen att alla smaragder är gröna, eftersom den hypotesen tidigare fungerat bra i praktiken och eftersom den ”ultraskeptiska hypotesen” att smaragder aldrig är gröna har falsifierats. Men man kan argumentera precis på samma sätt för hypotesen att alla smaragder är gröda. Den har *hittills* fungerat bra i praktiken. Dessutom har den ultraskeptiska hypotesen att smaragder aldrig är gröda också falsifierats. Enligt Nordins resonemang har vi ingen anledning att vara mera skeptiska till hypotesen att alla smaragder är gröda. Tar vi däremot hänsyn till betydelsen av *sträng* prövning, som jag föreslår i mitt rationalitetskriterium, så är det oförnuftigt att tro att alla smaragder är gröda, eftersom den hypotesen hittills inte kunnat prövas strängt. Därför finns det enligt mitt kriterium för förnuftig tro för närvarande inte några goda skäl för att

(felbart och reviderbart) tro att alla smaragder är gröda eller för att basera praktiskt handlande den hypotesen. Nordin (2010, s. 38–39) kan därför som kritisk rationalist känna sig lugn när han ställs inför det praktiska problemet om han skall välja en smaragdprydd slipsnål till morgondagens fest. Han har inga förnuftiga skäl att befara att den gröna smaragden vid festen skulle ha förvandlats till en röd.

Om vi lämnar praktiskt handlande och övergår till en diskussion om vetenskaplig forskning, så behöver vi inte heller någon princip för förnuftigt tro enligt Nordin (2010, s. 39). Han anför att vi väljer att jobba vidare med den hypotes som bäst stått emot sträng prövning, eftersom den är ”ändå den bästa kandidaten för sanning”. Här förutsätter Nordin stillatigande rationalitetsprincipen att den hypotes som bäst stått emot sträng prövning är den bästa kandidaten för sanning, medan jag uttryckligen föreslår att det är förnuftigt att hålla den för sann. Skillnaden är inte så stor.

I min artikel hävdade jag (2009b, s. 15–16) att principen för förnuftigt tro kan tillämpas också vid diskussionen av testsatser, som t.ex. testsatserna ”Här är nu ett glas vatten”. Om testsatser klarar empirisk prövning, så är det enligt min princip förnuftigt att tro på dem. Nordin (2010, s. 39) invänder att vi inte behöver någon princip för *absolut* tro på testsatser, vilket jag håller med om, eftersom min princip är en princip för felbara och reviderbara försanthållanden, inte för absolut tro.

Nordin (2010, s. 39–40) anför vidare att testsatser accepteras genom *metodologiska beslut* och att vår eventuella tro på den ena eller andra testsatsen är irrelevant i det sammanhanget. Men Nordin beslutar sig inte utan vidare för att acceptera testsatser, utan diskuterar ”välkorroboretrade testsatser”, dvs. testsatser som klarat sträng prövning, och anser att vi har goda skäl att vara ”mindre skeptiska” mot sådana satser. Även här förutsätter Nordin stillatigande en rationalitetsprincip, nämligen att vi har goda skäl för det metodologiska beslutet att acceptera välkorroboretrade testsatser. Detta kan enklare uttryckas som att det är förnuftigt att hålla testsatser för sanna som klarat av sträng prövning, vilket är en tillämpning av min rationalitetsprincip på testsatser (jfr Andersson 2006; Andersson 2009a, s. 25–26).

4. ÄR DET FÖRNUFTIGT ATT ACCEPTERA DEN KRITISKA RATIONALISMEN?

Lars Bergström (2010, s. 29–30) kritiserar min motivering för principen att det är förnuftigt att tro på de hypoteser som vid en viss tidpunkt bäst stått emot kritisk prövning (*KR*). Min motivering var: ”Om *KR* klarar av kritisk prövning, inklusive prövningen om den uppfyller sina egna krav, så är det förnuftigt att acceptera *KR*” (Andersson 2009b, s. 17).

Bergström kallar denna sats *KKR* och anser att jag därmed infört en ny rationalitetsprincip och frågar om det är förnuftigt att acceptera den. För att besvara frågan om det är förnuftigt att acceptera *KKR* krävs enligt Bergström ytterligare en rationalitetsprincip *KKKR* och så vidare, vilket leder till en oändlig regress.

Avgörande för Bergströms argument är hur *KKR* uppfattas. Om *KKR* är en ny och oberoende rationalitetsprincip får vi en oändlig regress, men om *KKR* är en speciell tillämpning av *KR* så får vi det inte.

Till diskussion står följande rationalitetsprincip: Det är förnuftigt att tro att *P* (vid tidpunkten *t*) om och endast om *P* är den hypotes som (vid tidpunkt *t*) bäst har stått emot sträng kritik (*KR*). Om vi ersätter *P* med *KR* ovan får vi: Det är förnuftigt att tro att *KR* (vid tidpunkten *t*) om och endast om *KR* är den hypotes som (vid tidpunkt *t*) bäst stått emot sträng kritik (*KKR*). Det visar att *KKR* inte är en ny och oberoende rationalitetsprincip, utan en tillämpning av *KR* på *KR* själv. Frågan om *KR* uppfyller sina egna krav är ett sätt att pröva *KR*, inte ett försök att bevisa *KR* med hjälp av någon annan rationalitetsprincip. Försök att bevisa *KR* leder till en oändlig regress, eftersom vi hela tiden måste fråga efter ytterligare rationalitetsprinciper. Försöket att pröva om *KR* uppfyller *sina egna krav*, kräver däremot inte ytterligare rationalitetsprinciper och leder därför inte till en oändlig regress, utan är en undersökning om *KR* är fri från inre motsägelser.

Leder försök att pröva *KR* till en logisk cirkel? Bergström (2010, s. 30) menar att frågan om *KR* kan tillämpas på sig själv på sin höjd ger oss ett cirkelbevis för *KR*. Vad följer av att *KR* kan tillämpas på sig själv? Till att börja med betyder det att *KR* klarat av en speciell typ av kritisk prövning. Om vi vidare antar att *KR* klarat av övriga kända typer av kritisk prövning, så följer med hjälp av en tillämpning av *KR* (på *KR* själv) att det är förnuftigt att hålla *KR* för sann.

Om jag strävade efter att *bevisa* att det är förnuftigt att hålla *KR* för sann med hjälp av *KR* själv, så hade Bergström med rätta kunnat invända att mitt resonemang är ett cirkelbevis. Men jag strävar inte efter att *bevisa att* utan efter att *pröva om* det är förnuftigt att hålla *KR* för sann. Strävan efter att *bevisa* leder till en oändlig regress, eftersom vi alltid kan fråga om premisserna för beviset i sin tur kan bevisas. Den oändliga regressen kan bara avbrytas genom att vi dogmatiskt beslutar oss för att utan ytterligare argument acceptera vissa typer av premisser eller genom att vi inför en logisk cirkel. Därför ställer oss strävan efter att rättfärdiga inför valet mellan tre oönskade alternativ, nämligen att välja mellan oändlig regress, dogmatiskt avslut eller logisk cirkel, vilket innebär att vi ställs inför ett rättfärdigandetrilemma (jfr Albert 1991, avsnitt 2 om Münchenhausen-trilemmat). Strävan efter att *pröva* leder däremot inte till något rättfärdigandetrilemma,

eftersom vi inte hela tiden ständigt måste fråga efter ytterligare rättfärdigande *förutsättningar*. Vid prövning av en sats undersöker vi om satsens *konsekvenser* är accepterbara. Har vi goda skäl för att finna att en sats har en oacceptabel konsekvens, så har vi goda skäl för att inte acceptera den prövade satsen (Albert 1991, avsnitt 1 och 5; Andersson 2009a).

5. ETT FILOSOFISKT TRADITIONSBROTT

Den västerländska filosofin har allt från Platon och Aristoteles fram till den moderna empirismen varit behärskad av sökande efter visshet (Wedberg 1970, 104), vilket lett till uppfattningen att kunskap bör bestå i rättfärdigad sann tro. Strävan efter rättfärdigad kunskap har lett de flesta konsekventa tänkare fram till rättfärdigandetrilemmat. Många är de försök som förgäves har gjorts att undvika detta problem. Ett försök är t.ex. att ersätta visshet med sannolikhet. Men om vi strävar efter att nå visshet om graden av sannolikhet, så uppträder ändå rättfärdigandetrilemmat. Samtidigt har en del filosofer, som t.ex. Hume (1748, sektion X, avsnitt 1), insett att vi inom vetenskapen inte når fram till rättfärdigad sann tro och ansett att det är förnuftigt att låta graden av tro vara proportionell mot evidensen. Men Hume kunde inte rättfärdiga sin uppfattning om förnuftig tro, varför han förblev en skeptiker.

Den kritiska rationalismen bryter med traditionen som uppfattar kunskap som rättfärdigad sann tro. I stället för att sträva efter rättfärdigande och visshet, strävar den kritiske rationalisten efter kritisk prövning och förnuftiga försanthållanden. Den kritiske rationalisten begår inte misstaget att efter att ha gett upp kravet på visshet och rättfärdigande inom vetenskapen ändå kräva visshet och rättfärdigande inom filosofin. Också inom filosofin kan vi inte nå längre än till förnuftiga försanthållanden efter kritisk prövning. Har man väl insett det, så når man fram till en position som i varje fall inte drabbas av rättfärdigandetrilemmat.

Tar vi hänsyn till tillståndet av den kritiska diskussionen i dag, så är det förnuftigt att acceptera den kritiska rationalismen. Detta är viktigt i dessa postmoderna tider när så många berättelser florerar som aldrig underkastats kritisk granskning.

LITTERATUR

- Albert, Hans. 1991. *Traktat über kritische Vernunft*. Femte upplagan. Tübingen: Mohr Siebeck.
- Andersson, Gunnar. 2006. Test Statements and Experience. I *Karl Popper: A Centenary Assessment, Selected Papers from Karl Popper 2002*, red. Ian Jarvie, Karl Milford and David Miller, 2: 177-183. Aldershot: Ashgate.

- Andersson, Gunnar. 2009a. Critical Rationalism and the Principle of Sufficient Reason. I *Rethinking Popper*, red. Zuzana Parusniková och Bob Cohen, 21–30. New York: Springer.
- Andersson, Gunnar. 2009b. Förnuftig tro och intellektuell moral. *Filosofisk tidskrift* 30 (4): 13–18.
- Bergström, Lars. 2010. Kritisk rationalism? *Filosofisk tidskrift* 31 (2): 29–34.
- Goodman, Nelson. 1955. *Fact, Fiction, and Forecast*. Fjärde uppl. Cambridge MA: Harvard UP, 1983.
- Hume, David. 1748. *An Inquiry Concerning Human Understanding*.
- Miller, David. 1999. Being an Absolute Skeptic. *Science* 284 (June): 1625–1626.
- Nordin, Ingemar. 2010. Varför måste vi tro på något? *Filosofisk tidskrift* 31 (2): 35–40.
- Wedberg, Anders. 1970. *Filosofins historia: Nyare tiden till romantiken*. Andra uppl. Stockholm: Bonnier. Ny utgåva 2003. Stockholm: Thales.

Lars Samuelsson

Transitivitet och omärkbara effekter

En invändning mot ett argument
i Folke Tersmans bok *Tillsammans*

1. INLEDNING

I sin intressanta och mycket läsvärda bok *Tillsammans: En filosofisk debattbok om hur vi kan rädda vårt klimat* presenterar Folke Tersman ett argument som han tänker sig visar att en handling kan vara rätt eller fel av någon annan anledning än att den har bra eller dåliga konsekvenser (Tersman 2009, s. 132-133).¹ Argumentet är i stora drag hämtat från Derek Parfits *Reasons and Persons* (Parfit 1987, s. 78-82), men Parfit drar, som vi kommer att se, en annan slutsats av det än den som Tersman drar. I den här uppsatsen ska jag försöka visa att Tersmans slutsats inte följer av detta argument. Men alldeles obeaktat den frågan så finns det anledning att rikta uppmärksamhet mot den synnerligen aktuella problematik som det här argumentet belyser, nämligen hur vi ska hantera det faktum att många handlingar som de flesta av oss anser att det finns moraliska skäl att utföra, eller inte utföra (såsom handlingen att köra bil till jobbet trots att det skulle gå utmärkt att gå eller cykla) är sådana att deras effekter (positiva eller negativa) är omärkbara (dvs. inte gör någon skillnad alls ”i det stora hela”). Varför ska jag ta cykeln i stället för bilen, om effekterna av just min bilkörning saknar betydelse?

Tersmans version av argumentet går ut på att om vi inte accepterar slutsatsen att handlingar kan vara rätt eller fel av någon annan anledning än att de har bra eller dåliga konsekvenser, så tvingas vi i stället godta den till synes orimliga uppfattningen att relationen *bättre än* inte är transitiv, dvs. uppfattningen att trots att A är bättre än B, och B är bättre än C, så kan C vara bättre än A. (Att uppfattningen verkar orimlig ser vi om vi betraktar följande resonemang: Anta att jag har C, och att jag sedan får välja mellan alternativen att behålla C och att i stället få B. Jag bör då välja B eftersom B är bättre än C. Anta nu att jag får välja mellan att behålla B och att i stället få A. Jag bör då välja A eftersom A är bättre än B. Slutligen får jag välja mellan att behålla A och att få tillbaka C, som

¹För att finna en sådan anledning sneglar Tersman i riktning mot varianter av tanken ”tänk om alla gjorde så”. Han nämner i detta sammanhang Kants kategoriska imperativ och regelutilitarism (Tersman 2009, s. 133-138).

jag hade från början. Jag bör då välja C, eftersom C är bättre än A. Jag har gått i cirkel. Att acceptera en sådan cirkulär värdeordning verkar inte särskilt rationellt. Om vi lägger till antagandet att jag är beredd att betala en skälig (men mycket liten) summa pengar för att få byta ut ett sämre alternativ mot ett bättre, så inser vi ganska snabbt att jag skulle kunna bli ett lätt byte för den som vill sko sig på min bekostnad – givet det kanske osannolika villkoret att denna person har i sin makt att gå på gång avgöra vilket av alternativen som ska realiseras.)

Jag ska inte här försvara ståndpunkten att relationen *bättre än* är transitiv. Men vi kan notera att det inte är särskilt många som har argumenterat för motsatsen. Och vi kan väl i alla fall gå med på så mycket som att uppfattningen att *bättre än* inte är transitiv är klart ointuitiv, och att vi skulle behöva tänka om när det gäller mycket av det vi tar för givet i fråga om värden, moral och rationalitet, om den skulle visa sig vara riktig. Jag ska inte heller försöka ge någon ny lösning på det transitivetsproblem som Tersman målar upp. I stället ska jag ifrågasätta att Tersmans ”lösning” på detta problem verkligen är en lösning, dvs. jag ska ifrågasätta att hans argument ger stöd för uppfattningen att en handling kan vara rätt eller fel av någon annan anledning än att den har bra eller dåliga konsekvenser (eller bra eller dåliga *förväntade* konsekvenser).² För att vara mer precis: jag ska försöka visa att det inte ger något stöd för denna uppfattning alls.

2. TERSMANS ARGUMENT

Tersmans argument bygger på det faktum att vissa handlingar, som tagna för sig har omärkbara effekter, är sådana att när alla, eller många, utför handlingar av den typen, så har det betydande, och i vissa fall mycket allvarliga, effekter. Som utgångspunkt använder sig Tersman av ett tankeexperiment som är lånat från Parfit (1987, s. 76). Här följer Tersmans presentation av det:

Anta att tusen personer försmäktar av törst i en öken någonstans. Anta också att vi är tusen välgörare i en närbelägen stad, och att vi har var sin tillbringare fylld med en liter vatten. Vattnet kan hållas i en tank som sedan förs ut till öknen där vattnet delas lika bland de törstande. För varje glas som hålls i tanken får var och en av de törstande alltså en extra milliliter. Vi tänker oss också att välgörarnas beslut är oberoende av varandra. [...] Det är rimligt att tro att ingen av de törstande skulle märka skillnaden om de fick en extra milliliter. (Den tveksamme

²Många menar ju att frågan om huruvida en handling är rätt eller fel ska besvaras med hänvisning till handlingens realiserande av förväntat värde snarare än med hänvisning till dess realiserande av faktiskt värde. Tersmans (Parfits) argument är lika problematiskt för denna ståndpunkt.

kan i stället anta att det finns en miljon törstande och en miljon välgörare, vilket skulle innebära att varje välgörares bidrag bara skulle utgöra en miljondels liter.) Med andra ord, oavsett hur många välgörare som skulle hålla i sitt vatten utöver de hundra som redan hållt i så skulle de törstande inte märka skillnaden om det hade varit en färre. Men samtidigt gäller alltså att, om *alla* skulle hålla i sina litrar, så skulle de törstande må märkbart bättre ... (Tersman 2009, s. 118–119)

Tersmans resonemang ser nu ut som följer (se också Parfit 1987, s. 78–82):

Ta återigen fallet med de törstande i öknen. Hundra av välgörarna har redan hållt i sitt vatten i tanken, vilket innebär att var och en av de törstande får åtminstone en deciliter. Men det vore givetvis bättre för dem att få en liter, vilket innebär att det är bättre att de övriga nio hundra välgörarna också bidrar än att de inte gör det. [...] Samtidigt gör en extra *tusendels* liter ingen märkbar skillnad. Om vi dessutom antar att det skulle gagna välgörarna att dricka upp vattnet själva [men inte tillnärmelsevis lika mycket som det skulle gagna en av de törstande i öknen att dricka en liter vatten]... så verkar det följa att det gäller för var och en av de återstående nio hundra att hon inte bör bidra. Med andra ord, det är bättre att 100 håller i än att 101 gör det, eftersom ingen av de törstande missgynnas av att hon avstår medan hon själv gynnas. Om vi nu slutligen antar att relationen ”bättre än” är transitiv, vilket innebär att om A är bättre än B och om B är bättre än C så är också A bättre än C, så följer det att det är bättre att ingen av dem håller i sitt vatten än att alla gör det, vilket stridet mot det antagande vi utgick från. (Tersman 2009, s. 132–133)

Låt oss göra resonemanget ännu tydligare. Till att börja med är det viktigt att notera att detta är ett resonemang som handlar om utfall, vilket blir uppenbart om vi betraktar den första premissen. Vad den premissen säger är att det är bättre om de törstande får en liter vatten var än om de bara får en deciliter vatten var. Det vill säga: det utfall i vilket de törstande har en liter vatten var är bättre än det utfall i vilket de bara har en deciliter var. Det är *därför* som det är bättre att de övriga 900 välgörarna håller i sitt vatten än att de inte gör det. Exemplet väddar till våra intuitioner om vad som är bra och dåliga utfall: det är (allt annat lika) bättre att 1000 personer inte utstår ett fruktansvärt lidande orsakat av extrem törst, än att de utstår ett sådant lidande. Det är svårt att förneka.

Resonemanget handlar även fortsättningsvis om utfall, eller om värdet hos utfall. Det kan t.ex. inte vara fallet att det handlar om handlingars värden.³ Det vore visserligen bättre (en bättre handling), givet det resonemang som vi har fört, om den 101:a personen lät bli att hålla i sitt vatten än om hon höll i det, men dessa två möjliga handlingars värden kan

³Att underlåta att utföra en viss handling – t.ex. att inte hålla i sitt vatten – betraktar jag här också som en handling.

vi inte på ett meningsfullt sätt ordna tillsammans med värdena av den 102:a personens möjliga handlingar att antingen hålla i sitt vatten eller låta bli, osv. Den 101:a personens handling att hålla i sitt vatten är varken bättre eller sämre än den 102:a personens handling att hålla i sitt vatten; dessa handlingars respektive värden tillhör inte samma värdeordning. Och om en uppsättning värden inte tillhör samma ordning, så kan det naturligtvis inte heller – vad dessa värden anbelangar – bli fråga om något brott mot transitiviteten hos *bättre än*. Dessutom finns det ingen enskild handling som kan resultera i att 1000 personer i stället för 100 håller i sitt vatten (detta faktum kan ju sägas utgöra en del av vårt problem).

Tanken är således att det utfall som följer om 100 personer håller i sitt vatten är bättre än det utfall som följer om 101 personer gör det, eftersom den enda etiskt relevanta skillnaden mellan dessa två fall är att i det senare går en person törstig i stället för otörstig (nämligen en av de potentiella välgörarna). (Om man av någon anledning oroas av att detta kan ses som ett egoistiskt skäl från välgörarens eget perspektiv, så kan vi i stället tänka oss att han eller hon ger bort vattnet till några törstiga barn som leker i närheten, men vilkas törst – även om den är obehaglig – inte är tillnärmelsevis lika obehaglig som den törst som var och en av de törstande personerna i öknen utstår.)⁴ Låt oss kalla utfallet som följer om 100 personer håller i sitt vatten för U_{100} , och utfallet som följer om 101 personer håller i sitt vatten för U_{101} , osv.

Samma tankegång ger att det är bättre att 101 personer håller i sitt vatten än att 102 gör det, eftersom den enda etiskt relevanta skillnaden mellan dessa två fall är att i det senare fallet går en person törstig i stället för otörstig. Vi får alltså resultatet att U_{100} är bättre än U_{101} , och att U_{101} är bättre än U_{102} . Om vi fortsätter på samma sätt får vi resultatet att U_{102} är bättre än U_{103} och att U_{103} är bättre än U_{104} , ända tills vi kommer till resultatet att U_{999} är bättre än U_{1000} . Om relationen *bättre än* är transitiv så följer det att U_{100} är bättre än U_{1000} . Men vi inledde hela resonemanget

⁴Faktum är att vi kan konstruera argumentet på ett sådant sätt att vi överhuvudtaget inte behöver anta att någon gynnas ifall en ”välgörare” låter bli att hålla i sitt vatten, nämligen om vi fokuserar på den till synes lika uppenbara transitiviteten hos relationer som *åtminstone lika bra som* och *inte sämre än* i stället för hos *bättre än* (det är för övrigt vad Parfit gör (Parfit 1987, s. 78–79)). Om det inte gör någon märkbar skillnad om 101 i stället för 100 personer håller i sitt vatten, så är det ingen som missgynnas ifall 100 i stället för 101 personer håller i sitt vatten, och då är det inte heller sämre att 100 personer håller i sitt vatten än att 101 gör det. På samma sätt är det inte sämre att 101 personer håller i sitt vatten än att 102 gör det, osv. Däremot är det uppenbart mycket sämre att 100 personer håller i sitt vatten än att 1000 gör det, och därmed är transitiviteten hos *inte sämre än* bruten.

med att slå fast att U_{1000} är bättre än U_{100} (eftersom det är bättre att de törstande får en liter vatten var och befrias från sin smärtsamma törst än att de bara får en deciliter var och därmed inte befrias från denna törst). Under förutsättning att detta resonemang är giltigt tvingas vi således ge upp transitiviteten hos *bättre än* (vi behöver förstås inte gå ända till U_{1000} för att få detta resultat, men det verkar vara obestämbart (vagt) hur långt vi behöver gå, eftersom det är obestämbart (vagt) hur mycket vatten som måste hållas i tanken för att (någon av) de törstande i öknen ska märka någon skillnad (jfr Danielsson 1986)).

Tersman fortsätter sin argumentation:

Om en uppsättning påstående implicerar en motsägelse så måste något av dem ges upp. Vilket? Det är uteslutet att uppge antagandet att det är bättre att alla håller i vattnet (utöver de hundra) än att ingen gör det. Att uppge idén om transitiviteten hos relationen "bättre än" är inte heller särskilt lockande. [...] Det som återstår är därför att hävda att det kan vara bättre att hålla i sitt vatten, och att man därför bör göra det, även om det inte innebär att konsekvenserna blir bättre. (2009, s. 133)

3. VARFÖR TERSMANS SLUTSATS INTE FÖLJER

Av skäl som jag nu ska försöka förklara är detta inte alls vad som återstår. Vad som återstår är nämligen att hävda att det finns åtminstone ett par av utfall (U_n, U_{n+1}) sådant att U_{n+1} är bättre än U_n , dvs. sådant att konsekvensen att U_{n+1} realiseras är bättre än konsekvensen att U_n realiseras (se t.ex. Danielsson 1986 för ett förslag på en sådan lösning på den här typen av problem).⁵ Det är också ungefär var Parfit säger. Parfit introducerar ökenexemplet som ett i en rad exempel, med vilka syftet bland annat är att visa att vi har goda skäl att ifrågasätta följande antagande: "Någons smärta kan inte förbättras eller försämrats omärkligt. Någons smärta kan varken bli mindre illa, eller värre, om det inte finns någon möjlighet för personen ifråga att märka någon skillnad" (Parfit 1987, s. 78, min övers.).

Om vi förkastar detta antagande kan vi också undvika vad Parfit refererar till som "det femte misstaget i moralisk matematik", nämligen (det förmodade) misstaget att tro att "om en handling har omärkbara effekter på andra människor, så kan den handlingen inte vara fel *på grund av* att den har dessa effekter" (ibid., s. 75, min övers.). Om man ger upp detta

⁵Det räcker inte att det finns åtminstone ett par av utfall (U_n, U_{n+1}) sådant att U_{n+1} är *åtminstone lika bra som* U_n , eftersom vi då inte undviker resultatet att den relationen (dvs. *åtminstone lika bra som*) inte är transitiv. Utfallet U_{100} är ju absolut inte åtminstone lika bra som U_{1000} , vilket det då skulle behöva vara för att upprätthålla transitiviteten hos denna relation (se föregående fotnot).

antagande så följer det att man accepterar att en handling kan vara fel *på grund av* dess egna effekter, trots att dessa effekter är omärkbara. Dvs., man accepterar att ett utfall, A, kan vara bättre än ett annat utfall, B, även om det inte finns någon i A som kan märka någon skillnad mellan A och B. I vårt ökenexempel innebär det att man accepterar att utfallet U_{n+1} kan vara bättre än utfallet U_n , trots att ingen i U_{n+1} kan märka någon skillnad. Om vi godtar denna ståndpunkt behöver vi inte längre oroa oss för att vi ska tvingas gå med på att relationen *bättre än* inte är transitiv (i alla fall inte med anledning av den typ av fall som vi här diskuterar). Vi kan då vidhålla att U_{1000} är bättre än U_{100} trots att relationen *bättre än* är transitiv, eftersom vi kan hävda att det finns åtminstone något par av utfall (U_n, U_{n+1}), på vägen mellan U_{100} och U_{1000} , som är sådant att U_{n+1} är bättre än U_n .⁶

Om vi däremot, som Tersman föreslår, i stället hävdar att det kan vara bättre att hålla i sitt vatten även om det inte innebär att konsekvenserna blir bättre, så löser vi inget transitivitetsproblem. Skälet är helt enkelt att vi då har övergått från att bedöma värdet hos utfall till att bedöma värdet hos handlingar. I och med att vi har slagit fast att konsekvenserna av handlingen att hålla i sitt vatten inte blir bättre än konsekvenserna av att låta bli, så måste påståendet ”att det kan vara bättre att hålla i sitt vatten” innebära att det är själva handlingen som kan vara bättre. Men att handlingen är bättre förändrar inte värdeordningen av de relevanta utfallen (som är precis lika (förmodat) icke-transitiv som tidigare). Tersmans förslag på lösning kräver i själva verket att vi accepterar att denna ordning kvarstår, eftersom hans slutsats inkluderar antagandet att konsekvenserna av att hålla i sitt vatten inte blir bättre än konsekvenserna av att låta bli (och vi har aldrig ifrågasatt den positiva konsekvensen av att en ”välgörare” låter bli att hålla i sitt vatten, nämligen att hon får dricka upp det själv – eller ge det till någon annan).

Skillnaden mellan Parfits och Tersmans slutsatser är alltså följande: Parfits slutsats innebär att vi ändrar värdeordningen av de relevanta utfallen, medan Tersmans slutsats innebär att värdeordningen kvarstår. Parfits lösning är att förneka att U_n är bättre än, eller åtminstone lika bra som, U_{n+1} . Tersmans ”lösning”, däremot, kräver att vi accepterar att U_n är åtminstone lika bra som, eller bättre än, U_{n+1} , eftersom vi annars inte kan hävda att handlingen som resulterar i U_{n+1} är bättre, *trots att den inte har bättre konsekvenser* än alternativet att inte utföra den. Det enda som kan göra det sant att handlingen *inte har bättre konsekvenser* än alternativet att inte utföra den, är ju att utfallet som följer av att utföra handlingen

⁶Det räcker egentligen inte att det finns ett sådant par. Det måste också ligga på ”rätt ställe”. Dvs., det måste ligga så till att det inte uppstår något nytt brott mot transitiviteten hos *bättre än* mellan U_{n+1} och U_{1000} .

(U_{n+1}) *inte är bättre* än utfallet som följer av att inte utföra den (U_n) . Och därmed kvarstår den icke-transitiva värdeordningen.⁷

För att sammanfatta: Tersmans slutsats är att en handling h_1 kan vara bättre än en annan handling h_2 även om h_1 inte har bättre konsekvenser än h_2 , dvs. även om h_1 inte skulle resultera i ett bättre *utfall* än det som h_2 skulle resultera i. Och skälet till att denna slutsats inte följer av hans argument är alltså att vi genom att hävda att h_1 kan vara bättre än h_2 även om h_1 inte har bättre konsekvenser än h_2 inte på något sätt har förändrat värdeordningen av dessa konsekvenser, som alltså är precis lika (förmodat) icke-transitiv som tidigare.

4. EN ALTERNATIV SLUTSATS

Men kanske skulle man här kunna hävda att den korrekta slutsatsen att dra av Tersmans resonemang är att det helt enkelt inte går att ordna utfall med avseende på deras relativa värde. I så fall är det ju också så att i den utsträckning som handlingar kan vara rätt eller fel, så måste de vara det i kraft av något annat än vilka utfall de resulterar i (dvs. något annat än deras konsekvenser). Men detta tycks i själva verket vara en variant av den lösning som Tersman skriver är utesluten, nämligen att ge upp antagandet att det är bättre att alla håller i vattnet (utöver de hundra) än att ingen gör det. Om vi inte kan ordna utfall med avseende på deras relativa värde så kan det ju inte heller vara riktigt att ett visst utfall är bättre än ett annat. Det kan visserligen inte vara riktigt att det är sämre heller, men nog måste man väl betrakta det som en mycket märklig slutsats att vi inte (korrekt) kan säga att det är bättre att 1000 personer befrias från ett fruktansvärt lidande än att de inte gör det (själv tycker jag denna slutsats vore minst lika märklig som slutsatsen att relationen *bättre än* inte är transitiv).

Nu är det emellertid uppenbart att Tersman inte vill acceptera en sådan slutsats. Han är mycket tydlig med att det faktum att en handling har goda konsekvenser utgör *ett* skäl för att utföra den (Tersman 2009, s. 138). Argumentet skulle också behöva modifieras för att det skulle vara möjligt att dra denna slutsats från det. Som det nu står vädjar det ju till våra intuitioner om just vilka utfall som är bättre och sämre, och förutsätter därmed att utfall kan ordnas med avseende på deras relativa värde. Om vi ville dra slutsatsen att man inte kan ordna utfall med avseende på deras relativa värde fick vi i stället vädja till våra intuitioner om vilka utfall som skulle

⁷Dvs. vi får resultatet att relationen *åtminstone lika bra som* inte är transitiv (se fotnot 4). Om vi dessutom lägger till antagandet att det sakförhållande som består i att en välgörare får behålla sitt vatten har positivt värde (och detta antagande har ingen förnekad), så får vi även resultatet att relationen *bättre än* inte är transitiv.

vara bättre och sämre *om* man kunde ordna utfall med avseende på deras relativa värde. Det är oklart vad detta egentligen skulle innebära, och om vi alls kan sägas ha några sådana intuitioner, samt om dessa intuitioner i så fall skulle ha någon tyngd överhuvudtaget (eftersom vi nu antar att vi faktiskt *inte* kan ordna utfall med avseende på deras relativa värde).

5. EN MÖJLIG UTVÄG?

Det kanske finns ett annat sätt att nå Tersmans slutsats? I diskussionen ovan har jag tagit för givet att det utfall som en viss handling resulterar i består uteslutande i den handlingens konsekvenser. Men, kunde man hävda, själva det faktum att handlingen utförs (eller handlingen själv, om man så vill) kan också ses som en del av utfallet. Utfallet av en välgörarens handling att hålla i sitt vatten skulle i så fall bestå både i handlingens konsekvenser och i själva det faktum att denna handling utförs. Därmed kan man hävda att handlingens värde avspeglar sig i utfallets värde och på så vis förändrar värdeordningen av de relevanta utfallen (bl.a. från att vara icke-transitiv till att vara transitiv). Men inte heller på detta sätt kan vi undkomma transitivetsproblemet.

För det första så kan vi ”mäta” konsekvensernas värde och handlingens värde delvis oberoende av varandra. Till exempel så är ju tanken här att en välgörarens handling att hålla i sitt vatten ska anses som rätt även om inte tillräckligt många andra välgörare håller i sitt vatten (dvs. även om människorna i öknen fortsätter att lida lika mycket som tidigare) – handlingen ska alltså vara rätt oberoende både av dess egna konsekvenser och av konsekvenserna av alla handlingarna tagna tillsammans. Dessutom skulle ju inte konsekvenserna bli sämre om det visade sig att en av välgörarna hade hållt i sitt vatten av misstag, och egentligen avsåg att behålla det, men rimligtvis har inte denna välgörarens handling samma värde som handlingen att avsiktligt hålla i sitt vatten. Vi kan kalla värdet som konsekvenserna av en handling bidrar med till värdet hos ett utfall för V_k , och värdet som själva handlingen bidrar med till värdet hos utfallet för V_h . Vi kan ”mäta” dessa värden var för sig, och alltså ordna utfallen antingen med avseende på deras relativa V_k , eller med avseende på deras relativa V_h (eller med avseende på deras relativa totala värde). Om vi tittar endast på V_k så är det fortfarande fallet att värdeordningen av utfallen i ökenexemplet är (förmodat) icke-transitiv. Resultatet att relationen *bättre konsekvenser än* inte är transitiv är precis lika oönskat som resultatet att relationen *bättre än med avseende på totalt värde* inte är transitiv. Det ser vi tydligt om vi går till nästa punkt.

För det andra. Vi kan lätt konstruera scenarion där endast värde av typen V_k är involverat (dvs. där det totala värdet hos de relevanta ut-

fallen utgörs uteslutande av V_k). I ett sådant scenario finns det alltså ingen (möjlig) handling som är av ett sådant slag att själva det faktum att handlingen utförs skulle kunna förändra värdet hos det utfall som handlingen realiserar. För att illustrera detta kan vi modifiera ökenexemplet. Vi behåller de 1000 törstande människorna i öknen, men vi byter ut de 1000 välgörarna mot 1000 barn. För vart och ett av dessa barn gäller att det antingen har en liter vatten, eller att det inte har en liter vatten. Om ett av dessa barn inte har en liter vatten så innebär det att denna liter i stället har delats lika mellan de 1000 törstande i öknen. Anta nu att av de 1000 barnen är det 100 stycken som inte har en liter vatten i sin ägo. Det innebär alltså att de törstande i öknen har fått en deciliter vatten var. Detta är utfall U_{100} i detta modifierade exempel. Anta vidare att du, på något vis, har i din makt att välja huruvida ett barn till ska fräntas sin liter, så att de törstande i öknen får ytterligare en milliliter vatten var. Detta är utfall U_{101} . Vad bör du göra?

Om du väljer att ta vattnet från barnet så kommer barnet att uppleva en obehaglig törst. Vi kan anta att barnet måste avsluta sin roliga lek och gå en ganska lång sträcka för att få nytt vatten. Men det "lidande" som barnet därmed utstår är inte i närheten av det lidande som var och en av de törstande i öknen utstår, och som de skulle befrias från om de fick dricka varsin liter vatten. De törstande i öknen kommer emellertid inte att märka någon skillnad alls om de endast får ytterligare en milliliter vatten var (antar vi) – de kommer inte att må bättre på något sätt. Naturligtvis ska du låta barnet behålla sitt vatten. Här handlar det inte om att dra sitt strå till stacken, eller om att själv erhålla en fördel genom att inte bidra, utan det enda som rimligen kan göra din handling bättre eller sämre i detta fall är handlingens konsekvenser (eller möjligen din avsikt bakom handlingen, men den aspekten saknar relevans här eftersom du i detta fall bör avse att utföra just den handling som du efter noggrant övervägande tror har bäst konsekvenser). U_{100} är alltså bättre än U_{101} .

Men nu får du i stället (märkligt nog) välja mellan U_{101} (där det är 101 barn som inte har en liter vatten i sin ägo) och U_{102} (där ytterligare ett barn blir av med sitt vatten och får uppleva obehag, men där ingen får det märkbart bättre än i U_{101} , eftersom den enda övriga skillnaden mellan de två utfallen är att i U_{102} har var och en av de törstande i öknen en milliliter vatten mer, vilket inte påverkar deras välmående). Utifrån samma resonemang som i föregående val väljer du därför U_{101} . Sedan får du välja mellan U_{102} och U_{103} , och du väljer följaktligen U_{102} eftersom det är det bästa alternativet, osv., ända tills du kommer till valet mellan U_{999} och U_{1000} , där du utifrån precis samma resonemang som tidigare väljer U_{999} eftersom det är det bästa alternativet. Vid denna punkt får du emellertid möjlighet att välja mellan U_{1000} , där förvisso 1000 barn är törstiga och

temporärt olyckliga, men där samtliga 1000 törstande personer i öknen befrias från sitt fruktansvärda lidande, och U_{100} , där detta lidande fortgår. Självklart väljer du U_{1000} , eftersom du då befriar 1000 personer från en fasansfull smärta. U_{1000} är bättre än U_{100} . Men det innebär att cirkeln är sluten och att vi har fått precis samma (förmodade) icke-transitivitet hos relationen *bättre än* som den vi hade i det ursprungliga exemplet.⁸

Att låta handlingars värde ingå i värdet hos utfall utgör alltså ingen lösning på transitivitetsproblemet. Det kvarstår i sådana fall där det inte finns några handlingar för vilka vi skulle kunna hävda att de har ett värde som förändrar utfallets värde. Dessutom kvarstår det som sagt också om vi endast tittar på värdet av handlingars konsekvenser. Det enda sättet att undvika resultatet att relationen *bättre än* inte är transitiv (i den typ av fall som vi har diskuterat här) är att hävda att det finns åtminstone ett par av utfall (U_n, U_{n+1}) sådant att U_{n+1} är bättre än U_n . Och om vi hävdar det så kan vi också hävda att var och en av de 1000 välgörarna (i det ursprungliga exemplet) bör hålla i sitt vatten i tanken, även om vi förnekar att en handling kan vara rätt eller fel av någon annan anledning än att den har bra eller dåliga (faktiska eller förväntade) konsekvenser.

6. AVSLUTNING

En viktig sak att notera om det modifierade ökenexemplet är att det inte längre är ett exempel på när en persons handlingar har omärkbara effekter medan många personers sammantagna handlingar har märkbara effekter. Det är nu bara en person som handlar⁹ (och om vi vill kan vi presentera exemplet helt utan någon handlande agent, som ett exempel där vi helt enkelt jämför värdena hos olika utfall). Det betyder att den transitivitetsproblematik som jag har diskuterat i den här uppsatsen inte är unik för "kollektivt handlande", och det visar också tydligt att Ters-

⁸Jag måste tillstå att jag har ganska svårt att se hur ett utfall U_{n+1} kan sägas vara bättre än U_n i detta exempel (i alla fall om vi gör antalet törstande respektive barn oerhört stort). Innebär det att jag måste acceptera att relationen *bättre än* är icke-transitiv? Kanske kan transitivitet hos *bättre än* bara garanteras om vi från början har tillgång till alla möjliga relevanta utfall (dvs. om det från början står klart att U_{1000} är ett utfall som det är möjligt att nå (i någon mening av 'möjligt'))? I så fall kanske vi kan tillämpa Parfits resonemang och hävda att U_{n+1} kan vara bättre än U_n trots att de positiva effekterna av U_n är omärkbara. Denna tanke finns det emellertid inte utrymme att utveckla vidare här. Det kan ju också vara så att skillnaden mellan två (märkbara) nivåer av lidande på grund av törst motsvaras av en så liten skillnad i vattenmängd som skillnaden mellan att få i sig och att inte få i sig "en vattenmolekyl". I så fall kan det också finnas minst ett par av utfall (U_n, U_{n+1}) sådant att U_{n+1} är bättre än U_n .

⁹Både Parfit och Danielsson ger också exempel av den här typen.

mans förslag på lösning av den inte fungerar: att hänvisa till handlingars värde är inte relevant i den typ av fall som det modifierade ökenexemplet representerar. Den förklaring som Tersman vill ge till varför en handling kan vara bättre än en annan trots att den inte har bättre konsekvenser – en förklaring som hänvisar till någon variant av tanken ”tänk om alla gjorde så” – täcker inte sådana fall där det endast är en agent som handlar. Därav följer också att även om vi hade kunnat lösa transitivhetsproblemet genom att hävda att en handling kan vara rätt eller fel av någon annan anledning än att den har bra eller dåliga konsekvenser, så hade vi varit tvungna att hitta en annan sådan anledning i de fall där vi inte har att göra med ”kollektivt handlande”.

Alltså: vilka skäl vi än vill anföra för att anamma uppfattningen att en handling kan vara rätt eller fel av någon annan anledning än att den har bra eller dåliga konsekvenser, så kan vi inte anföra skälet att vi på så vis kan undvika resultatet att relationen *bättre än* inte är transitiv (åtminstone inte så länge som vi vill kunna fortsätta att ordna utfall med avseende på deras relativa värde), för det är helt enkelt inte sant. (Observera dock att jag inte har sagt någonting som ger skäl att avfärda idén att en handling kan vara rätt eller fel av någon annan anledning än att den har bra eller dåliga konsekvenser. Jag har bara sagt att Tersmans argument inte ger något stöd för den idén.) Den viktigaste slutsatsen som man kan dra ur allt detta är nog att varje moralteori enligt vilken vi åtminstone ibland bör utföra den handling som har bäst konsekvenser (varje rimlig moralteori, om vi ska tro Parfit 1987, s. 69), och som alltså har ett behov av att ordna utfall med avseende på deras relativa värde, får samma transitivhetsproblem i fall med omärkbara effekter. Det räcker inte att introducera andra principer för vad som gör en handling rätt eller fel vid sidan av en konsekventialistisk princip för att undkomma detta problem. Och om vi kan lösa det så kanske fall av typen ökenexemplet, eller klimathotet, inte längre utgör problem för en konsekventialistisk etik. Att dra sitt strå till stacken och köra mindre bil kanske är precis vad man bör göra enligt en konsistent konsekventialist, och det på grund av effekterna som denna handling har, trots att dessa effekter är omärkbara.¹⁰

¹⁰Ett stort tack till Bertil Strömberg för värdefulla synpunkter.

LITTERATUR

- Danielsson, Sven. 1986. ”Om att gå runt på en gräsmatta”, i *Filosofiska invändningar*. Thales. (En kortare version av artikeln finns i *Filosofisk tidskrift* 1/84.)
 Parfit, Derek. 1987 [1984]. *Reasons and Persons*. Oxford University Press.
 Tersman, Folke. 2009. *Tillsammans: En filosofisk debattbok om hur vi kan rädda vårt klimat*. Bonnier Existens.

Nånstans mellan Lomma och Landskrona, jag minns inte exakt, började bilradion ljuda av tonerna från John Lennons *Imagine*. Rastlöst lutade jag mig fram för att byta kanal – men stannade plötsligt upp.

Några dagar tidigare hade jag läst klart Richard Rortys samlingsvolym *Philosophy as Cultural Politics*, och nu arbetade det undermedvetna – mil efter mil längs E6:an – med att greppa bokens teser. Vad var egentligen kärnan i Rortys filosofi?

”Uttrycket ’Verkligheten som den är i sig själv oberoende av mänskliga behov och intressen’ är med mitt synsätt bara ännu ett av de underdåniga namnen för Gud”, skriver Rorty (2007, s. 134).

Så tog Lennon till orda.

*Imagine there's no heaven
It's easy if you try*

Plötsligt gick hjärnan på högvarv. *No heaven!*

*No hell below us
Above us, only sky*

Jag sjönk tillbaka bakom ratten igen.

Det är dags, säger Rorty, att helt och hållet sluta ”titta uppåt”. Det är dags att sluta tro på en objektiv verklighet som vi har skyldighet – och genom vetenskapen möjlighet – att komma i kontakt med. *Above us, only sky*.

Leder inte detta förkastande av ”Verkligheten” till farlig relativism och irrationalism? Nej, svarar Rorty, sådana hot är endast klerikal propaganda. *No hell below us*.

Pragmatismen som filosofi är i grund och botten, förklarade Rortys föregångare F. C. S. Schiller år 1912, inget annat än ”humanismen tillämpad på kunskapsteorin” (Rorty 2007, s. 134). 1700-talets upplysning var humanistisk i många avseenden, men i och med dess starka förnufts- och vetenskapstro så behöll den det ”vertikala hoppet” om en frälsning från ovan. ”Genuin hängivenhet för realism”, skriver Rorty, bör betraktas

som ”upplysningens version av det religiösa begäret att få böja sig inför en icke-mänsklig makt” (Rorty 2007, s. 134).

Richard Rorty ser alltså pragmatismen som ett fullbordande av humanismen. Och därmed också som en del i ett stort kulturellt omslag i västerlandet: det omslag som innebar att människan började titta framåt i stället för uppåt; den utveckling där förhoppningar och skyldigheter blir ”horisontella” i stället för ”vertikala”, och där människors tankar och känslor rör jordelivet och efterföljande generationer – snarare än Gud, Evigheten och livet efter döden.

Efter ett sådant kulturomslag hämtar intellektuella människor sin kraft och sina känslor av storslagenhet ur andra källor än de traditionella.

Teister finner patos i avståndet mellan det mänskliga och det gudomliga. Realister finner det i avgrunden som skiljer mänskligt tänkande och språk från verkligheten som den är i sig själv. Pragmatister finner det i gapet mellan den samtida mänskligheten och en utopisk mänsklig framtid där själva idén om ansvar gentemot något utöver våra medmänniskor har blivit obegriplig, resulterande i den första verkligt humanistiska kulturen. (Rorty 2007, s. 135)

Liksom Lennon har Rorty en framtidsvision:

*Imagine all the people
Living for today...*

Det var med boken *Philosophy and the Mirror of Nature* som Richard Rorty påbörjade sitt filosofiska projekt. Bokens måltavla var idén om kunskap som korrekt avbildning av verkligheten, och idén om medvetandet och språket som en ”spegel”. Därmed var boken en generell uppgörelse med den moderna analytiska filosofin. Den publicerades 1979 och nu trettio år senare har en jubileumsutgåva av verket utkommit.

När boken först publicerades fick den genast ett stort genomslag. Rortys missnöje med den analytiska filosofin hade växt fram under 70-talet, men för alla som inte följde Rorty på nära håll kom *Philosophy and the Mirror of Nature* som en överraskning. Och framförallt som en utmaning: här kom en analytiskt välrenommerad filosof och ifrågasatte grunderna för hela projektet. På de filosofiska institutionerna runt om i USA startades läsecirklar och diskussionsgrupper – ”Is Rorty Right?” (Gross 2008, s. 228).

En viktig utgångspunkt för Rorty är att ”Det är bilder snarare än propositioner, metaforer snarare än utsagor, som bestämmer de flesta av våra filosofiska övertygelser” (Rorty 2009, s. 12). Den position Rorty för fram är inte vanlig skepticism – att ”spegeln” ofrånkomligen är suddig eller ger förvrängningar – utan snarare att hela metaforen är missvisande och skadlig. Problemet är, skulle man kunna säga, att ”världsbild” är ett bildspråk som vi har anammat alltför bokstavligt.

Rorty förklarar att hans bok är ”lik skrifterna av de filosofer jag mest beundrar, terapeutisk snarare än konstruktiv” (Rorty 2009, s. 7). Detta grundar sig en historicistisk syn på de filosofiska problemen, det vill säga Rorty betraktar dem inte som eviga frågeställningar, utan som något som uppkommer som ett resultat av andra idéer och etablerade metaforer. Idéer som i sin tur har utgjort svar på tidigare upplevda problem. Från Heidegger hämtar Rorty idén att ”[Begäret] efter en ’epistemologi’ är helt enkelt det senaste resultatet av den dialektiska utvecklingen av den ursprungligt valda uppsättningen metaforer” (Rorty 2009, s. 163). När filosoferna väl har kommit att betrakta kunskap som en samling korrekta avspeglingar av verkligheten, så är steget kort till att börja leta efter en särskild klass av representationer som är privilegierade genom sin omedelbara kontakt med verkligheten. Denna klass kommer att utgöra kunskapens grund, ”och den disciplin som för oss i dess riktning – kunskapsteori – kommer att bli kulturens grund” (Rorty 2009, s. 163).

En stor del av *Philosophy and the Mirror of Nature* utgörs således av historien om de intellektuella innovationer som åstadkommit av Descartes, Locke och Kant, och med tillbakablickar till Platon och Aristoteles. En historieskrivning vars syfte naturligtvis är att undergräva känslan av de filosofiska problemens naturlighet och självklarhet. Det historicistiska synsättet bäddar för Rortys terapeutiska favoritfilosofer – Wittgenstein, Heidegger och Dewey. De har visat på ett alternativ: ”De åsidosätter kunskapsteori och metafysik som discipliner. Jag säger ’åsidosätter’ snarare än ’argumenterar mot’ eftersom deras attityd gentemot den traditionella problemställningen är densamma som 1600-talets filosofer hyste mot den skolastiska problemställningen” (Rorty 2009, s. 6).

Likväl är Rortys bok ett verk *inom* den moderna analytiska skolan. Dess poäng är dock att peka på en väg ut ur den. Detta sker genom att spinna vidare på några viktiga resultat hos framstående analytiska filosofer (främst Sellars, Quine och Davidson). Rorty försöker visa att de sammantaget undergräver det analytiska projektet – och pekar just i den historicistiska och anti-representationalistiska riktning som han själv önskar gå. ”När Sellars och Quines doktriner renodlas så framstår de som komplementära uttryck för ett och samma påstående: att ingen ’framställning över kunskapens natur’ kan förlita sig på en teori om representationer som står i privilegierad relation till verkligheten” (Rorty 2009, s. 182). Dessa två filosofers verk utgör ”utmaningar mot själva idén om en ’kunskapsteori’, och därmed mot filosofi i sig självt, betraktad såsom en disciplin centrerad kring en sådan teori” (Rorty 2009, s. 169).

I *Philosophy and the Mirror of Nature* identifierar sig Rorty ännu inte som pragmatist. Men hans grundläggande idéer uttrycks här. I slutkapitlet utvecklar han sin vision om en ”filosofi utan speglar”: ett sätt att

bedriva filosofi utan att sträva att nå en fast punkt, ett fundament eller en fulländad systematik. En filosofi som erkänner att mänsklighetens intellektuella samtal, "the Conversation of Mankind", är sig självt nog. Här är hans centrala perspektivförändring. Medan filosoferna traditionellt har försökt att nå fast mark *bortom* konversationen, försökt att sätta punkt, så förkunnar Rorty det motsatta: er enda skyldighet är att hålla samtalet vid liv.

Det var först i volymen *Consequences of Pragmatism* 1982 som Rorty länkade samman sin filosofi med den klassiska amerikanska pragmatismen och dess förgrundsgestalter Charles Sanders Peirce, William James och John Dewey. För Rortys del innebar detta att en långvarig personlig och filosofisk omvandling hade fullbordats.

Historien om denna omvandling kan man ta del av genom Neil Gross bok *Richard Rorty: The Making of an American Philosopher*. Boken, som kan beskrivas som en blandning av biografi och sociologisk fallstudie, är den första större studien över Rortys filosofiska utveckling. Fram till sin död 2007 samarbetade Rorty med Gross i detta projekt, i form av intervjuer samt tillgång till brev och opublicerade manuskript. Det var också Gross som, bland Rortys efterlämnade papper, hittade texten "The Philosopher as Expert", författad omkring 1961 och nu publicerad som ett addendum i jubileumsutgåvan av *Philosophy and the Mirror of Nature* (Gross, s. 162–4; Rorty 2009, s. 395–421). (Potentiella läsare av Gross bok bör vara medvetna om att en avsevärd del ägnas åt att vidareutveckla Pierre Bourdieus och Randall Collins teorier om intellektuellt fält och kapital, framförallt genom att utveckla begreppet "intellectual self-concept". Jag kommer inte att behandla denna aspekt av boken här.)

En av de viktigaste poängerna med Gross bok är den visar att den vanliga bilden av Rorty – bilden av honom som en genuin analytisk filosof som sedan blir avfälling – bör revideras. Gross visar att Rorty även under sin "analytiska fas" hyste ett antal förbehåll. Ursprungligen var Rorty en klassisk metafysiker: han skrev sin magisteruppsats om Whiteheads filosofi, och hans doktorsavhandling behandlar Aristoteles potentialitetsbegrepp. Det är först efter att ha nått doktorsgraden som han börjar positionera sig som en analytisk filosof. 1966 kan hans mor konstatera, i ett brev till en väninna, att Richard "har upphört att vara en metafysiker" (Gross, s. 189). Detta sker enligt Gross genom ett komplext "identitetsarbete". Det framstår tydligt att Rortys satsning på analytisk filosofi var kraftigt villkorad redan från början; många av de betänkligheter som kom att resultera i *Philosophy and the Mirror of Nature* uttryckte Rorty redan i detta stadium.

Men kanske är det mest intressanta med Gross bok att den visar hur

mycket av den senare Rortys filosofi som kan skönjas i hans uppväxt och tidiga influenser.

Richard Rorty växte upp i ett hem präglat av föräldrarnas sociala engagemang. Föräldrarna, James Rorty och Winifred Raushenbush, var trotskister och båda var framträdande i New Yorks intellektuella kretsar under 30-talet. De tilltalades av John Deweys anti-kommunistiska progressivism. Föräldrarnas generella attityd till intellektuell verksamhet var, återigen likt Dewey, att den hela tiden måste motiveras på praktiska och sociala grunder, som ett medel att nå sociala framsteg. De största intellektuella synder man kunde begå var akademiskt navelskådande och världsfrånvärd klassicism.

Men sonens utbildning ledde i motsatt riktning. Redan som 15-åring lämnade Richard föräldrarna i New Jersey för att gå ett specialprogram vid University of Chicago för talangfulla ungdomar. Det var en utbildning i den västerländska kulturens klassiska verk. I början av sin studietid uppvisar han fortfarande föräldrahemmets skepsis. I ett brev hem beskriver han, med en ironi som förebådar hans vuxna författarskap, ett föredrag han bevistat: ”Jag lärde mig att fram till klockan fem igår, Chicago-tid, hade exakt 102 Stora Idéer tänkts ut av den Västerländska Civilisationen” (Gross, s. 95). Men påföljande sommar kom han att läsa Platon, och hans tänkande formades alltmer av hoppet om att finna tidlösa sanningar och tingens essens. Han lockades av den storslagna metafysiken och revolterar således mot Dewey. 1948 skriver Richard till sin far: ”*Damn it*, människan är mycket mer än ett djur och hon kan inte kan inte behandla sig själv på samma sätt som hon har behandlat naturens alla ting. Att reducera henne till den roll hon faktiskt spelar i det biologiska systemet framstår i mina ögon som både omöjligt och värdelöst” (Gross, s. 97).

Dessa två motsatta tendenser – föräldrarnas respektive utbildningens – skapade en inre spänning. Rorty har i en självbiografisk essä beskrivit hur han därför närdes av drömmen om att, med Yeats ord, ”hold reality and justice in a single vision”. Det var ett uttryck för förhoppning om att hans livsval – filosofin – skulle gå att förena med föräldrarnas sociala engagemang.

Richard Rortys filosofiska utveckling är i mångt och mycket historien om hur han tappade tron på möjligheten att finna en sådan vision. Och hur han därefter, i linje med hans föräldrars ideal, valde att sätta sociala framsteg framför ”verkligheten”. Häri ligger kärnan i Richard Rortys filosofi. Att tvärtemot Yeats (jfr Rorty 1999, s. 7) – och tvärtemot Platon – inte sträva efter att ge sina sociala och politiska förhoppningar en filosofisk uppbackning. Att inte längre ”titta uppåt”.

Det var ett sådant kulturellt omslag som Rorty satte sina egna för-

hoppningar till, och han hämtade sitt patos ur, som han uttryckte det, gapet mellan samtiden och den humanistiska utopin. Till sina kritiker kunde han följaktligen passande ha replikerat:

*You may say, I'm a dreamer
But...*

LITTERATUR

Gross, Neil. 2008. *Richard Rorty: The Making of an American Philosopher*. The University of Chicago Press.

Rorty, Richard. 1982. *Consequences of Pragmatism*. University of Minnesota Press.

Rorty, Richard. 1999. *Philosophy and Social Hope*. Penguin Books.

Rorty, Richard. 2007. *Philosophy as Cultural Politics*. Cambridge University Press.

Rorty, Richard. 2009. *Philosophy and the Mirror of Nature*. Thirtieth-Anniversary Edition. Princeton University Press.

Notiser

JORDAN HOWARD SOBEL (1929–2010) TILL MINNE

En framstående filosof och en god vän, Howard Sobel, avled 26 mars i år. Han var omtyckt och beundrad av många svenska filosofer och hans död blev för dem en svår förlust. Sobel bodde i Toronto i Kanada, men intellektuellt sett var Uppsala hans andra hem. Varje vårtermin från 1999 och framåt höll han kurser och deltog i högre seminarier vid den uppsalianska filosofinstitutionen. Helt följdriktigt utnämndes han 2003 till hedersdoktor vid Uppsala universitet. Till sin 80-årsdag, 2009, fick han från filosofiska institutionen i Uppsala en fin festskrift med bidrag från hans vänner och kolleger i Sverige och utomlands. Festskriftens dubbelbottnade titel, *Logic, Ethics and All That Jazz*, omnämner några av de stora passionerna i hans liv.

Sobel besökte Uppsala första gången i maj 1986. Förutom att ge en serie avancerade föreläsningar i beslutsteori tjänstgjorde han vid detta tillfälle som fakultetsopponent vid Jan Österbergs doktorsdisputation på en avhandling om etisk egoism. Många av oss som var närvarande minns fortfarande disputationen mycket väl. Sobel var ovevlig i sin kritik, trots (eller kanske på grund av) sin höga uppfattning om avhandlingens kvalitet. Senare, 1993, kom han att publicera en detaljerad analys av Österbergs bok i *Erkenntnis*.

Han kom åter till Uppsala det akademiska året 1997/98 samt på våren 1999 – bägge gånger som gästforskare vid Svenska kollegiet för avancerade studier. Givetvis fick han vid flera tillfällen inbjudningar till att hålla föreläsningar vid andra universitet i landet – i Lund, Stockholm och inte minst i Umeå, där han bl.a. gav Burmanföreläsningar 1998.

Hans första kontakter med svenska filosofer går tillbaka till slutet av sextioalet. Han läste då Lars Bergströms banbrytande avhandling, *The Alternatives and Consequences of Actions* (1966), vilket ledde till en livlig korrespondens med författaren. Ännu tidigare, 1965, publicerade han i den svenska internationella filosofitidskriften *Theoria* en uppsats om ”Generalization Arguments”, där han ifrågasatte några teser i Marcus Singers välkända *Generalization in Ethics* (1961). Detta var förresten So-

bels första publikation i en filosofisk tidskrift. Generaliseringsargument inom etiken var ämnet för hans doktorsavhandling som han försvarade vid universitetet i Ann Arbor, Michigan, samma år som Singers bok kom ut. Som handledare hade han Richard Cartwright och William Frankena. Efter en kortare sejour i Princeton tillbringade Sobel sedan några år i Los Angeles, vid UCLA, och därefter i Amherst, vid University of Massachusetts. 1969 flyttade han till Toronto, som "associate professor". Han utnämndes till "full professor" 1976 och stannade kvar vid universitetet i Toronto fram till pensioneringen 1995.

Theoria hade förmånen att även i fortsättningen vara ett forum för hans filosofiska författarskap. Tidskriften har publicerat flera av hans artiklar, om så pass skilda ämnen som utilitarism, beslutsteori, Protagoras' paradox, Sokrates' motivering för beslutet att underkasta sig atenarnas dom, etc. Så sent som 2008 utkom i *Theoria* en av Sobels sista uppsatser, om lögnarparadoxen. Han bidrog även generöst till ett flertal uppsaliensiska och lundensiska festskrifter, för Lennart Åqvist, Frederick Stoutland, Krister Segerberg och för undertecknad.

Mina kontakter med Sobel startade under det tidiga åttiotalet med en omfattande korrespondens om kausal beslutsteori. Ett besök hos Sobel och hans hustru Willa under deras Cambridgevistelse 1985 blev sedan början till en livslång vänskap. Sobels nydanande bidrag inom besluts- och spelteori och hans subtila analyser av de begreppsliga ramarna för utilitarism och kontraktualism inom etiken har varit en central inspirationskälla för mitt eget arbete inom dessa fält. Hans utformning av kausal beslutsteori framstår för mig fortfarande som oöverträffad, fullt i klass med klassiska arbeten av Allan Gibbard & William Harper, David Lewis och Brian Skyrms, samt senare bidrag av James Joyce. Sobels tidiga uppsatser om besluts- och spelteoretiska ämnen finns samlade i volymen *Taking Chances* (Cambridge University Press 1994), vilken jag skulle vilja påstå är en oundgänglig läsning för den som vill fördjupa sig i det rationella valets problematik. På hans hemsida, www.uts.utoronto.ca/~sobel/ finns också ett komplicerat och krävande bokmanuskript *Probability, Chance and Choice: A Theory of Rational Agency* (1978, reviderat 1980). Manuskriptet, som tyvärr aldrig har publicerats, förtjänar ett närmare studium: det ger fascinerade inblickar i beslutsteorins grundvalar.

Sobel fortsatte på nittioalet att skriva om praktisk rationalitet, även efter att hans nära umgänge med kausal beslutsteori hade upphört. Från denna senare period bör man särskilt nämna undersökningar av antaganden som ligger till grund för baklängesinduktion. Detta inflytelserika sätt att identifiera effektiva handlings- och spelstrategier bygger på förutsättningar som kräver ingående filosofisk analys. Sobel var en av de första som på allvar tog sig an denna uppgift. Från samma period

kan man också nämna hans diskussion av cykliska preferenser och penningpumpargument.

Förhållandet mellan rationellt beslutsfattande och den fria viljans problem finns i fokus för hans fascinerande bok från samma period, *Puzzles for the Will: Fatalism, Newcomb and Samarra, Determinism and Omniscience* (University of Toronto Press 1998).

Vad gäller Sobels analyser av utilitarism och kontraktualism får man söka sig till en serie uppsatser från sjuttio- och åttiotalet, samt det tidiga nittioalet, som publicerades i sådana tidskrifter som *Inquiry*, *Noûs*, *Canadian Journal of Philosophy*, *Theoria*, *Dialogue* och *Philosophical Studies*. Hans bidrag inom dessa områden blev tyvärr aldrig samlade i bokform, men deras inflytande på filosofiska debatter under denna period var betydande. De lästes allmänt av moralfilosofer med konsekventalistiska sympatier, inte minst i Sverige, och diskuterades flitigt. En frågeställning som jag särskilt kommer att tänka på i detta sammanhang gäller våra förutsägbara avvikelser från optimala handlingsplaner och dessa avvikelserns normativa relevans för våra nuvarande beslut – ett ämne som redan Bergström hade tagit upp i sin avhandling. För att ge ett exempel som kommer från en senare uppsats av Frank Jackson och Robert Pargetter:¹ bör jag åta mig ett angeläget uppdrag som jag är vällämpad att utföra (säg, att recensera en nyutkommen bok) om jag vet med mig att jag inte kommer att leverera ifall jag ställer upp? Eller bör jag snarare tacka nej på direkten? Många av oss känner tyvärr igen detta problem från vår egen erfarenhet! Sobels behandling av frågan finns i artikeln "Utilitarianism and Past and Future Mistakes" i *Noûs* 1976.

Sobel var alltid djupt intresserad av logiska metoder. Ett av hans speciella intressen var logiska paradoxer. Ett annat var tillämpningen av logiken inom religionsfilosofin. Cambridge University Press publicerade 2004 hans *Logic and Theism: Arguments For and Against Beliefs in God*. Då jag inte är särskilt förtrogen med ämnet tillåter jag mig att översätta några rader ur Nicholas Everitts recension: "Detta är ett framstående religionsfilosofiskt arbete. Det kännetecknas av en kolossal lärdom, noggrannhet i sitt hänsynstagande till detaljer, såväl textuella som filosofiska, enorm förtrogenhet med den relevanta litteraturen. Det är djärvt där det behöver vara djärvt men varsamt nyanserat där förbehåll är på sin plats" (*Philosophical Books* 47, s. 380-3). Everitt är författare till *The Non-existence of God* som utkom samma år som Sobels bok.

Efter volymen om logik och teism utgav Sobel nyligen ett nytt verk, *Walls and Vaults: A Natural Science of Morals (Virtue Ethics According to David Hume)* (Wiley 2009). Hume hade länge varit ett av hans favoritprojekt,

¹F. Jackson och R. Pargetter, "Oughts, options, and actualism", *Philosophical Review* 95 (1986): 233-55.

vid sidan av tre andra centralgestalter i moralfilosofins historia: Platon, Aristoteles och Kant. På Sobels hemsida finns tre bokmanuskript som handlar om dessa filosofer: "No Light Matter: Socrates and Plato On Justice"; "Ends and Means: Aristotle on Happiness and Virtue"; "The Mystery and the Glory: Immanuel Kant's Philosophy of Morals". Där finns också en längre text om metaetikens historia under nittonhundratalet: "Good and Gold: A Judgmental History of Metaethics from G. E. Moore through J. L. Mackie". Man får hoppas att allt detta material hittar en förläggare. Sobels sätt att närma sig dessa moralfilosofins superstjärnor är genomgående stimulerande och subtilt i sin inriktning mot argumentationsstrukturen i deras verk: "djärvt där den behöver vara djärvt men varsamt nyanserat där förbehåll är på sin plats".

Blotta mängden av opublicerat material ger en fingervisning om hur aktiv Sobel har varit ända fram till slutet. Han blir saknad av sina läsare, kolleger och vänner.²

Wlodek Rabinowicz

HUR MAN HITTAR ETT BRA UPPSATSÄMNE

Det finns en amerikansk filosof som brukar säga till sina studenter att om de vill hitta ett bra uppsatsämne så ska de läsa någon berömd filosof och plocka ut satser som innehåller ord som "obviously" eller "clearly". Sådana ord tar man ofta till i brist på argument och det kan därför vara mycket lämpligt att försöka vederlägga eller argumentera mot just satser av detta slag. På svenska finns det många liknande ord och fraser, t.ex. "naturligtvis", "förstås", "självklart", "mycket talar för att", "rimligt att anta", osv. När man hör eller ser sådant ska man dra öronen åt sig.

NYA BÖCKER

Marcus Agnafors har doktorerat på en avhandling med titeln *Justice among us: A philosophical analysis of Michael Walzer's theory of justice*, Linköpings universitet 2010.

Richard Ohlsson har utgivit en bok med titeln *Tio moderna filosofiska klassiker*, BTJ Förlag, 2010. Den behandlar Kuhn, Wollheim, Berlin, Rawls, Nozick, Pirsig, Singer, Bok, Hare och Nagel på lite drygt 100 sidor.

²Dessa minnesord är baserade på en engelskspråkig förlaga som utkommer i *Theoria* nr 3, vol. 76, 2010.

RATIONELLT BETEENDE

I en tidigare notis (1/10) omtalades några intressanta experiment som visar att hundar har ett visst sinne för rättvisa. I samband härmed har Wlodek Rabinowicz påpekat att schimpanser vid Planckinstitutet i Leipzig visar helt andra taktik, närmare bestämt i så kallade ultimatumspel. Han skriver: "Villkoren är högst orättvisa enligt våra mått: den apa som utfärdar ett ultimatum kräver att få ta hand om nästan hela kakan som den inte själv har bakat, medan föremålet för erbjudandet får nöja sig med småsmulorna. Samarbetet etableras ändå, i och med att den är till nytta för alla berörda. Åtminstone i dessa sammanhang tycks alltså schimpanserna vara okänsliga för rättviseaspekter". Här har man alltså äntligen hittat den så kallade *homo economicus* – en varelse som annars har varit svår att påvisa i verkligheten. (Men beteckningen "homo" är förstås lite vilseledande.)

GENUSNYTT

Testosteron gör råttor mer aggressiva och riskbenägna. Därför ligger det nära till hands att tro att hormonet skulle ha samma effekt på människor. Men så är det tydligen inte. *Forskning och Framsteg* rapporterar om försök där man gav antingen en dos testosteron eller ett sockerpiller till drygt 120 kvinnliga studenter. De kvinnor som fått testosteron visade sig vara mer givmilda och hamna i färre konflikter än de som fick sockerpiller. Men det fanns också en tydlig effekt i motsatt riktning! Kvinnor som *trodde* att de fått testosteron blev nämligen mindre givmilda, oberoende av om de trodde rätt eller inte. Förväntningseffekten var starkare än den biologiska effekten.

LYCKAD REKLAM

Den berömde rättsfilosofen H. L. A. Hart från Oxford var en gång inbjuden att ge en föreläsning om "Jurisprudence" vid ett universitet i Israel. Man skulle kunna tro att ämnet bara skulle locka några få specialister och ambitiösa studenter, men det visade sig till Harts förvåning att en stor hörsal var alldeles proppfull med förväntansfulla åhörare. Emellertid upptäckte han senare att på affischen för föredraget hade ämnet angivits som "Jewish Prudence".

AKADEMISK SKEPTICISM

Horace Engdahl, som motsägelsefullt nog inte längre är ständigt sekreterare, har formulerat den tänkvärda åsikten att "om människor bara

ytrrade sig när de hade en välgrundad uppfattning, skulle det mänskliga samtalet avstanna” (DN 14/3). Av detta tycks följa att folk inte har några välgrundade uppfattningar – åtminstone inte tillräckligt ofta för att några samtal om dessa uppfattningar skulle kunna uppstå. Kan det möjligen vara så att Engdahls åsikt är den *enda* som är välgrundad? Eller är även den ett exempel på vad han betecknar som ”det lösa tyckandet”?

EN SORTS KANON

De mest citerade böckerna inom s.k. ”analytisk filosofi” under efterkrigstiden är tydligen (enligt *Google Scholar*) följande, siffrorna inom parentes anger antalet citeringar i on-line artiklar):

1. Thomas Kuhn, *The Structure of Scientific Revolutions* (37 197)
2. John Rawls, *A Theory of Justice* (26 768)
3. Ronald Dworkin, *Taking Rights Seriously* (7 892)
4. Ludwig Wittgenstein, *Philosophical Investigations* (7 169)
5. Karl Popper, *Conjectures and Refutations* (6 516)
6. Alasdair MacIntyre, *After Virtue* (6 579)
7. Gilbert Ryle, *The Concept of Mind* (6 356)
8. John Rawls, *Political Liberalism* (6 352)
9. Robert Nozick, *Anarchy, State, and Utopia* (6 246)
10. H. L. A. Hart, *The Concept of Law* (6 212)
11. Richard Rorty, *Philosophy and the Mirror of Nature* (5 616)
12. John Searle, *Speech Acts* (5 387)
13. Jerry Fodor, *Modularity of Mind* (5 050)
14. Daniel Dennett, *Consciousness Explained* (4 810)
15. Karl Popper, *Objective Knowledge* (4 701)

Därefter följer: Saul Kripke, *Naming and Necessity* (4 535); W.V. Quine, *Word and Object* (4 565); Paul Feyerabend, *Against Method* (4 420); Richard Rorty, *Contingency, Irony and Solidarity* (4 011); Charles Taylor, *Sources of the Self* (3 233); Jerry Fodor, *The Language of Thought* (3 292); Carl Hempel, *Aspects of Scientific Explanation* (3 137); David Chalmers, *The Conscious Mind* (3 065), Daniel Dennett, *The Intentional Stance* (2 985); Derek Parfit, *Reasons and Persons* (2 972). (På grund av speciella omständigheter har böcker som gränsar mot juridik, politik, psykologi och lingvistik här blivit något överrepresenterade. Man måste också komma ihåg att det här är fråga om böcker. Filosofer som mest publicerat sig i artiklar, t.ex. Donald Davidson och David Lewis, saknas därför på listan.)

HÖGA KRAV

Den amerikanske filosofen Robert Paul Wolff, som väl mest är känd för sin bok *In Defense of Anarchism* från 1970, har en läsvärd blog, där han bland annat publicerar sina minnen från olika filosofinstitutioner i USA. Bland annat berättar han hur det gick till på Columbia University när man skulle försöka få en efterträdare till Ernest Nagel som efter många år flyttat till Rockefeller University. Nagel var ju en stor stjärna och man vill gärna ha någon efterträdare i samma klass. Men vem skulle man fråga? Flera förslag framfördes – t.ex. Quine, Davidson, Kripke, Sellars och Goodman – men institutionens prefekt, Justus Buchler, tyckte inte att de var tillräckligt bra. (Buchler, född 1914, är väl numera okänd för de flesta, men han var enligt Wolff en mycket intelligent och kunnig metafysiker, som dock skrev böcker som ingen brydde sig om och som inte heller Wolff lyckades förstå något av, trots att han försökte.) Hur som helst, när Buchler röstat mot alla förslag som kom upp, frågade Wolff honom om det fanns någon i hela världen som kunde anses lämplig, oavsett om de kunde tala engelska. Buchler tänkte en stund och svarade sedan nej. Alltmer desperat frågade då Wolff hur det skulle gå om man släppte kravet på ”nu levande”. Till exempel Descartes? Nej, han hade ett för snävt upplevelsebegrepp, enligt Buchler. Kant, då? Nej, han begrep sig inte riktigt på relationen mellan möjligt och faktiskt. Till sist frågade Wolff: finns det *någon* i hela filosofihistorien, som skulle vara värdig att anställas av filosofinstitutionen på Columbia? Buchler funderade en stund och sa till slut: Aristoteles och Whitehead.

Medarbetare i detta nummer av *Filosofisk tidskrift*: Gunnar Andersson är professor emeritus i vetenskapsteori vid Umeå universitet, Lars Bergström är professor emeritus i praktisk filosofi i Stockholm, Teodor Gardelli är filosofistuderande vid Luleå universitet, Martin Lembke är doktorand i religionsfilosofi vid Lunds universitet, Lars Samuelsson är post. doc. i praktisk filosofi vid Umeå universitet, Gunnar Syrén är docent i litteraturvetenskap i Uppsala och Björn Östbring är fil. kand. i teoretisk filosofi och studerar politisk teori.

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 normalt med datorpost till lars.bergstrom@philosophy.su.se
- kan också skickas till
Lars Bergström, Telegrafgränd 1 B, 111 30 Stockholm
– men i så fall skall det dessutom bifogas på diskett
- 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
- noter och litteraturhänvisningar kan gärna inarbetas i texten
- 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
- ändringar mot manuskript bekostas icke av förlaget; författaren debiteras kostnad för sådana ändringar
- införda bidrag honoreras inte för närvarande
- i stället för särtryck erhåller författaren, gratis, 10 exemplar – för recensioner 5 exemplar – av det nummer av tidskriften i vilket bidraget varit infört