

November 2013 Årgång 34 Nr 4

Filosofisk tidskrift

- LARS-GÖRAN JOHANSSON
3 Naturalism, transcendentalism eller fenomenologi?
- STINA BÄCKSTRÖM
19 Hur hamnade Sanaz på festen?
- PATRIK RIBE
33 Den liberala autonomins förutsättningar
- HANS RUIN
42 Wittgenstein och filosofins tilltal
- KARL EKENDAHL
46 Döden igen
- RICHARD SWINBURNE
50 Hägström om Swinburne och sannolikheten för Guds existens
- LARS BERGSTRÖM
57 Teismen som förklaring
- 62 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 2013

Grafisk form: Ulf Jacobsen

Satt med Indigo Antiqua

Tryck: Carlshamn Tryck & Media AB, Karlshamn 2013

ISSN 0348-7482

Lars-Göran Johansson

Naturalism, transcendentalism eller fenomenologi?

1. INLEDNING: TRE POSITIONER I KUNSKAPSTEORIN

Jag skall i denna uppsats diskutera de tre kunskapsteoretiska positionerna naturalism, transcendentalism och fenomenologi. Dessa tre filosofiska riktningar är alla reaktioner på skeptiska tvivel. Diskussionen i denna uppsats handlar uteslutande om kunskap om yttervärlden, inte om logisk eller matematisk kunskap.

Med ordet ”naturalism” menar jag i denna uppsats kunskapsteoretisk naturalism, ett sätt att svara på skepticism som återfinns hos bl.a. Hume och Quine. Det finns många andra användningar av detta ord, t.ex. som beteckningen på tesen att världen inte är skapad av någon gud, utan har uppkommit och utvecklats genom naturliga processer, eller som en komponent i moralfilosofin. Detta är inte vad som avses med ordet naturalism i denna uppsats.

2. NATURALISM

Hur får vi reda på saker och ting? Via våra sinnen, säger empirister. Hur vet empiristerna det, frågar då kritikerna. Empiristerna svarar att vetenskaplig kunskap om oss själva, våra sinnesorgan, hjärnans funktioner osv., ger en rätt bra beskrivning av hur olika typer av signaler träffar våra sinnesorgan och hur hjärnan sedan bearbetar dessa så att det blir representationer av den yttre verkligheten. Visst, det är många detaljer som vi inte känner till, men i det stora hela har vi en trovärdig bild av hur sinnesorganen tar emot signaler och hur de processas av hjärnan. Det som fattas är en teori om hur vissa tillstånd i hjärnbarken kan bli till medvetna perceptioner, men knappast någon betvivlar att det är vissa hjärntillstånd som är de biologiska korreleten till våra medvetandetillstånd. De flesta är ense om att medvetandetillstånd supervenerar på fysiologiska tillstånd; dvs., att om en persons medvetandetillstånd ändras, så ändras också det fysiologiska tillståndet hos vederbörande.

Kritikerna hävdar att detta är förutsätta vad som skulle bevisas; man

kan inte bygga på empirisk evidens när frågan är varför man skall tro på empirisk evidens.

Många empiristers drag är nu att bli naturalister, dvs., att avvisa kravet på ett principiellt, filosofiskt rättfärdigande av empiriska kunskaper. Detta krav härstammar från föreställningen att filosofin, speciellt kunskapsteorin, skall utgöra en apriorisk grund för empiriska kunskaper, ja för all vardaglig och vetenskaplig kunskap (utom matematik och logik). Många empirister menar att i stället för att försöka finna en säker grund för all kunskap skall vi starta vår filosofiska reflektion där vi befinner oss, vi skall börja med att utgå ifrån vad vi för tillfället håller för sant och successivt utvärdera och revidera dessa försanthållanden. Vi behöver ingen yttersta grund. Hume uttryckte det som att vi i allt filosoferande måste utgå ifrån hur vi är funtade. Ett av hans argument (*Om förståndet*, s. 219) var att även om vi inte kan ge något hållbart argument för att det finns yttre ting som orsakar våra sinnesintryck, så är vi människor inte förmögna att på allvar tvivla på detta. Det tillhör vår *natur* (därav beteckningen ”naturalism”) att anta att det finns yttre ting och att våra vardagliga uppfattningar om åtminstone lätt tillgängliga ting och händelser i stort sett är sanna. När filosofen försöker ge argument för dessa trosuppfattningar så förefaller argumenten mycket mer dubiösa än den slutsats som de är tänkta att stödja. Hume konstaterar att vi i praktiken inte har något val annat än att utgå ifrån våra naturgivna uppfattningar. På ett annat ställe (s. 250) skriver Hume att man för stunden kan bli överväldigad av skeptiska tvivel, men inom mindre än en timme efter att man slutat fundera på om vi har några goda skäl för våra trosuppfattningar, så har allt tvivel försvunnit; var och en kommer att återgå till att tro allt som man brukar tro utan att besväräs av tvivel. Denna iakttagelse förefaller mig träffande och Hume tar den alltså som en utgångspunkt.

Quines uppsats ”Epistemology Naturalized” är den moderna uppföljaren till denna hållning. I uppsatsen börjar Quine med en kritisk genomgång av Carnaps projekt i *Der Logische Aufbau der Welt*. Carnap försökte där visa att vår kunskap om yttervärlden kan reduceras till vissa enkla element, av Carnap benämnda ”elementarupplevelser”, som uttrycktes med ”elementarsatser”. Tanken var att alla meningsfulla påståenden kan reduceras till påståenden om dessa elementarupplevelser och när det har lyckats kan empiristen säga att vi har en säker grund bestående av dessa elementarupplevelser för all empirisk kunskap.

Detta projekt kollapsar, vilket Carnap själv insåg ganska snart; man kan inte säga att en teoretisk sats, t.ex. lagsats, är ekvivalent med en viss ändlig uppsättning satser om elementarupplevelser: varje all-sats går utöver varje ändlig uppsättning av observationsatser. Så reduk-

tion av teorier till observationer är inte möjlig. Carnaps nästa steg var då att överge idealet om reduktion av teorier till elementarupplevelser, för att i stället hävda att teoretiska påståenden implicerar påståenden om elementarupplevelser, men inte omvänt. Quine påpekar nu att i och med detta steg har hela poängen med en rationell rekonstruktion av vår kunskap byggd på en tänkt säker bas fallit bort: för om man inte genom en definition kan eliminera teoretiska påståenden, så har man inte eliminerat: vi har ingen säker grund. (Den intensiva diskussionen om teoriers underbestämning av evidensen utgör en god illustration.) Vore det inte bättre, frågar sig Quine retoriskt, att i stället försöka beskriva hur vi faktiskt kommer fram till våra uppfattningar? Låt oss vända oss till kognitiv psykologi i stället för att spekulera, skriver Quine.

Om målet för kunskapsteorin är att visa giltigheten hos empirisk kunskap genom att härleda den från observationsdata, så är det cirkulärt att använda kognitiv psykologi. Men om vi överger målet om att härleda empirisk kunskap från observationsdata, så finns det ingen anledning att avhålla sig från att använda all tillgänglig information när vi försöker förstå sambandet mellan våra observationer och teoretiska kunskaper. Quine avvisar således cirkularitetsargumentet; det biter inte på naturalisten.

Vi behöver inget aprioriskt rättfärdigande av kunskap menar naturalisterna. Vårt tänkande, både vad gäller vardagliga och vetenskapliga frågor, är en ständigt pågående process av revidering av våra uppfattningar, och denna process drivs inte i första hand, eller kanske inte alls, av några abstrakta filosofiska principer, utan av gjorda erfarenheter av misslyckanden och, någon gång, framgångar. ”We are faced with the problem of error” skriver Quine i sin sista bok, *From Stimulus to Science*. Vad vi kan göra, och gör, är att ställa upp och testa hypoteser. Men inga hypoteser kan slutgiltigt bevisas; allt är i princip möjligt att betvivla. Dock inte på en gång; varje tvivel grundar sig på något argument som i sin tur bygger på att vi håller vissa andra saker för sanna.

Det är inte ovanligt att från tesen ”Vilken som helst av mina trosuppfattningar kan vara falsk” dra slutsatsen ”Alla mina trosuppfattningar kan vara falska”. Men det är en felaktig slutledning, vilket lätt kan visas med hjälp av ett motexempel.¹ Fallibilism är inte detsamma som skepticism.

Vi har succesivt lärt oss att förfina kraven på vad som kan godtas. Ett välkänt och klassiskt exempel är när medicinarna insåg att man behöver

¹Davidson ger följande motexempel: Satsen ”vilken som helst av sedlarna i min plånbok kan ha det högsta serienumret” är sann. Därav följer inte ”alla sedlar i min plånbok kan ha det högsta serienumret”, eftersom det är en falsk sats. Så det följer inte logiskt från satsen ”vilket som helst av mina trosuppfattningar kan vara falsk” att ”alla mina trosuppfattningar kan vara falska”.

göra dubbelblindtester för att vara rimligt säkra på att en viss behandling har verkan. Denna insikt är tämligen ny. Ingen filosof kom på detta, utan det var resultatet av gjorda empiriska erfarenheter bland medicinarna själva. Man kan sammanfatta saken med att vetenskaplig utveckling också innefattar kunskapsteoretisk reflektion och revidering av kunskapssteoretiska principer; de senare har ingen epistemisk prioritet.

Att vara kunskapsteoretisk naturalist innebär inte att man avvisar möjligheten av misstag, eller att man påstår att dagens vetenskapliga kunskap är säker. Men den totala skepticismen ("Det är möjligt att allting jag tror är falskt") bekymrar inte naturalisten; han ser inga skäl för en sådan ståndpunkt.

3. TRANSCENDENTALISM

Ett alternativ till naturalismen är Kants transcendentalism. Kant skriver i *Prolegomena* att han "väcktes ur sin dogmatiska slummer" av Hume. Det är välbekant att han syftar på Humes kritik av försöken att lägga en rationell, apriorisk grund för vår empiriska kunskap om yttervärlden. Kant erkänner oförbehållsamt att Hume hade rätt i sin kritik. Men Kant var inte tillfredsställd med Humes utväg, att ge upp hela idén om ett aprioriskt rättfärdigande och att inta den naturalistiska hållningen; detta uppfattade han som ett misslyckande. Men hur skall man då argumentera för att det vi tror om världen är sant om vi varken kan ge något aprioriskt eller empiriskt argument för tesen att vi har sanna uppfattningar om de yttre tingen? Som bekant löste han problemet genom att skilja på två nivåer av diskurs, den transcendentala och den empiriska. Den transcendentala frågan är: vilka är de nödvändiga betingelserna för att överhuvudtaget ha någon kunskap? Kants svar är, som bekant, att dessa betingelser ges av hur medvetandet arbetar, vilka dess fundamentala strukturerande funktioner är. Formerna på våra varseblivningar ges av hur medvetandet är beskaffat och dessa former är att yttre fysiska objekt varseblivs såsom ting ordnade i tid och rum. Sedan faller vi omdömen om dem. För att fälla omdömen måste man ha begrepp. Bland dessa finns enligt Kant ett antal fundamentala sådana, "kategorierna". Dessa kategorier är nedlagda i oss; de är konstitutiva för medvetandet.

I detta perspektiv bortfaller frågan om hur vi kan veta att våra förnimmelser av yttre ting är representationer av yttre ting, av det enkla skälet att varseblivningsinnehållet i Kants teori inte är en representation av något som finns oberoende av medvetandets strukturerande aktivitet. I en rimlig tolkning av Kants teori säger den att både rationalisters och empiristers projekt, att rättfärdiga uppfattningen att våra intryck av yttre

ting är veridikala representationer av dessa ting, bygger på misstaget att blanda ihop den transcendentala och den empiriska diskursen.²

Jag tolkar inte Kant som att han förnekar att det finns yttre, av medvetandet oberoende ting, han är ingen idealist i traditionell mening. Skälet är att ett sådant förnekande förutsätter distinktionen mellan yttre ting och deras mentala representationer, och denna distinktion går inte att göra i Kants teori. Men invänder kanske läsaren, Kant skiljer ju på ”fenomen”, dvs. *tingen såsom de ter sig för oss*, och ”noumen”, *ting såsom de är i sig själva*; är det inte likvärdigt med att skilja på innehållet i en varseblivning, dvs. något mentalt, och något yttre oberoende ting?

Nej. Kants ståndpunkt är, som jag tolkar honom, att kunskapsobjekten *konstitueras* av kunskapsakten och kunskapsakten är ontologiskt primär i relation till kunskapsobjektet. I kunskapsakten organiseras flödet av intryck till ett objekt. Kunskapsobjektet finns inte som ett objekt, som en entitet, oberoende av medvetandets organiserande process. Det är viktigt att observera att det inte är frågan om en process i tiden, utan om en ontologisk analys av relationen mellan kunskapsakten och objektet.

Det förefaller många självklart att anta att objektet finns oberoende av kunskapsakten, och att den, ifall den är lyckad, genererar en representation av detta objekt. Och detta verkar passa in i Kants tal om ”tinget i sig”. Men detta är inte, tror jag, Kants uppfattning.

Kant fullständiga fras är ”das Ding an sich selbst betrachtet” och kontrasten är ”das Ding als Phaenomen”. Det förefaller mig klart att det inte är två olika föremål han talar om, utan två olika *perspektiv* på ett och samma ting, det föremål om vilket vi fäller omdömen. Det ena perspektivet är ”utsikten från ingenstans”, för att låna en fras från Thomas Nagel, det andra perspektivet är vårt mänskliga perspektiv såsom varelser som har sinnliga erfarenheter och ett medvetande som arbetar så som Kant tänker sig saken. I det första perspektivet kan vi per definition inte fälla några omdömen om det; själva akten att fälla ett omdöme om ett ting förutsätter ett mänskligt medvetande som utför akten, dvs., omdömesfällandet sker alltid från vårt mänskliga perspektiv på tingen som fenomen. Så fort vi försöker tänka på tinget i sig självt, oberoende av hur det ter sig för oss, så försöker vi lyfta oss i håret. Det går inte.

På detta sätt anser sig Kant kunna lösa, eller snarare upplösa, det problem som Hume påvisade: det är medvetandets aktivitet som åstadkommer det som vi uppfattar som yttre objekt, och empiristernas

²Kant förnekar förstås inte att medvetandet även kan bli föremål för empiriska undersökningar; Kant skiljer på medvetandet såsom objekt för transcendental analys, det transcendentala medvetandet, och medvetandet såsom objekt för empiriska undersökningar, det empiriska medvetandet. I ett empiriskt perspektiv är det möjligt och legitimt att skilja på veridikala och icke-veridikala varseblivningar.

distinktion mellan det yttre tinget och dess representation i medvetandet, idén, upphävs. Det är alltså helt centralt att inse att *tinget i sig* inte är Kants motsvarighet till det yttre fysiska objektet så som filosoferna före Kant (och Kant själv då han fortfarande var lärjunge till Leibniz/Wolff) föreställde sig saken. För vore *tinget i sig* något skilt från *tinget som fenomen* skulle han ha kvar problemet, nämligen att ge en teori för relationen mellan tinget i sig och det uppfattade tinget, en teori som inte leder till skepticism.

Det förefaller mig träffande att säga att Kants transcendentalfilosofi bl.a. är en kritik av rationalismen för att den har perspektivet ”utsikten från ingenstans”, det som religiösa människor uttrycker som ”guds perspektiv”, och för att den försöker beskriva världen från detta perspektiv.

Utan tvivel såg Kant något viktigt när han lade märke till att kognitiva akter i hög grad är aktiva processer i våra medvetanden. På något sätt konstruerar medvetandet helheter, gestalter, av de stimuli som i varje ögonblick träffar våra sinnen. Han hade förstås ingen kunskap om sinnesfysiologi, den disciplinen fanns inte på hans tid. Men den teori man hade på 1700-talet om ljus, nämligen Newtons teori i *Opticks*, att ljus bestod av små partiklar, räcker för att dra slutsatsen att synintrycken består av ett flöde av inkommande partiklar vars effekter sedan sinnet på något sätt sätter samman till gestalter. (Enligt modern kvantteori hade Newton rätt i att ljus består av en sorts partiklar.)

Men talet om ett *ting i sig*, som vi inte kan veta något om, väcker undran hos många. Vilken funktion har detta begrepp i teorin? Om vi inte kan säga något om det, varför då överhuvudtaget ha detta begrepp i teorin? Kan man inte ta bort denna del av Kants filosofi och nöja sig med att tala om hur tingen ter sig för oss? Men då är man tillbaka i traditionell idealism, som blev en framträdande rörelse under 1800-talet. Ett tänkbart alternativ är fenomenologi. Men innan vi kommer till Husserl och fenomenologin är det fruktbart för diskussionen att säga något om Brentano och Frege; för Husserls och Freges teorier kan ses som olika svar på vad som kan kallas Brentanos problem.

4. BRENTANOS PROBLEM

Michael Dummett har i uppsatsen ”Thought and Perception: Two Philosophical Innovators” gjort en intressant jämförelse mellan Frege och Husserl. Dummett menar att båda utvecklade sina respektive ståndpunkter såsom lösningar på ett centralt problem som de båda hade ärvt från Brentano.

Franz Brentano är ju berömd för sin analys av medvetandeprocesserna såsom bestående av två delar, *akter* och *aktobjekt*. Brentano menade att

något är ett medvetandetillstånd om och endast om det består av dessa delar, som är så relaterade att akten är *riktad* emot sitt aktobjekt; detta är konstitutivt för ett medvetandetillstånd. (När det gäller propositionella tillstånd är aktobjektet en proposition, som kanske hellre bör kallas ett innehåll och inte ett objekt. Jag bortser i fortsättningen från denna skillnad och kallar även det propositionella innehållet för ett aktobjekt.)

Dummett menar att Brentano lämnade över ett olöst problem till sina efterträdare, nämligen vad det är för slags objekt som våra medvetandeakter är riktade emot. I vissa medvetandeakter, veridikala perceptioner, förefaller det självklart att det objekt som akten är riktad emot är det verkliga, yttre objektet som vi förnimmer. Men ibland har vi icke-veridikala medvetandetillstånd, såsom hallucinationer eller drömmar, vilka för oss ter sig kvalitativt precis likadana som veridikala perceptioner, men dessa akter är ju vanligen inte riktade emot verkliga objekt. Samma sak gäller emotionerna; de är ju vanligen riktade emot verkliga personer eller handlingar, men det är ett välkänt fenomen att man kan vara arg, glad, förbannad osv. på saker som inte finns, t.ex. en sedan länge död anförvant eller en ren fantasivarelse.

Man är frestad att säga att aktobjekten i dessa fall är mentala objekt, mentala representationer. Skillnaden mellan veridikala och icke-veridikala mentala akter är då att i de veridikala akterna finns det ett verkligt objekt som den mentala representationen är en representation av, medan i icke-veridikala akter så saknas det ett sådant yttre objekt, den finns enbart en tom representation (som väl rätteligen i så fall inte borde kallas en representation). Det förefaller ju vara en rätt naturlig förklaring, som är analog med hur vi ser på tomma namn och beskrivningar, dvs., uttryck som saknar referens. Beskrivningar kan ha betydelser, vara meningsfulla, utan att ha referens och mentala objekt kan vara objekt för akter men sakna sitt korresponderande yttre objekt.

Brentano själv avvisade denna lösning med det slagkraftiga argumentet att när man lovar en kvinna att gifta sig med henne är det den verkliga kvinnan man lovar att gifta sig med; man har ju inte tänkt gifta sig med en mental representation i sitt eget medvetande.

Men hur tänkte sig då Brentano att förklara det faktum att medvetandeakterna ibland är riktade emot objekt som inte finns? Är det inte närmast en självmotsägelse att säga så?

Dummett menar att Brentano inte hade någon lösning på detta dilemma. Det är just detta problem som Frege och Husserl försökte lösa på olika vis. Och dessa olika lösningar markerar separationen mellan analytisk filosofi och fenomenologi.

5. FREGES LÖSNING PÅ BRENTANOS PROBLEM

Frege var ju som bekant primärt intresserad av matematikens filosofi, inte av medvetandeakter i största allmänhet. Så man kan hos Frege inte finna något direkt svar på Brentanos problem när det gäller perceptionsakter. Men Frege diskuterar propositionella tillstånd och deras innehåll, propositionerna eller tankarna. Dummett menar att Freges semantiska teori utgör ett sätt hantera problemet.

Frege gjorde som bekant en distinktion mellan ett uttrycks *referens* (ty. "Bedeutung" och dess *mening* (ty. "Sinn"), och denna distinktion tillämpade han generellt, på såväl singulara termer, generella termer som fullständiga satser. Meningen hos ett uttryck är inte något objekt i Freges mening, utan det sätt på vilket vi identifierar eller "fattar" uttryckets referens. När det gäller en singular term är referensen det föremål som termen "står för" eller "pekar ut", när det gäller en generell term det *begrepp*³ som termen uttrycker och när det gäller en fullständig sats är referensen dess sanningsvärde.

Kan man nu säga något om statusen hos de objekt som medvetandeakterna är riktade emot med användning av denna teori? Ja.

Ett problem som upptog Frege var hur man skulle hantera semantiken för satser som beskriver innehållet i trosföreställningar, dvs. satser i indirekt anföring. Problemet är att extensionalitetsprincipen⁴ inte verkar gälla inom omfånget för de psykologiska verben "tror", "vet", "hoppas" osv. Anta att utsagan "Kalle Karlson vet inte att Jens Stoltenberg är Norges nuvarande statsminister" är sann. Om vi byter ut termen "Norges nuvarande statsminister" mot en annan term med samma referens, nämligen "Jens Stoltenberg", så får vi "Kalle Karlsson vet inte att Jens Stoltenberg är Jens Stoltenberg.", vilket torde vara en falsk sats. Således har vi erhållit en falsk sats från en sann sats genom att byta ett uttryck i satsen mot ett annat med samma extension. Extensionalitetsprincipen gäller alltså inte i indirekt anföring, vilket är en högst ovälkommen konsekvens som man gärna vill undvika.

Freges lösning var som bekant att säga att i indirekt anföring har namn och bestämda beskrivningar inte sin vanliga referens utan det är

³Ett begrepp är enligt Frege en funktion som avbildar föremål på sanningsvärden. Funktioner är inte objekt.

⁴Extensionalitetsprincipen säger att extensionen hos ett uttryck är en funktion av extensionen av uttryckets semantiska delar. Det innebär bl.a. att om vi i en sats, vars extension är sanningsvärdet, byter ut en term mot en annan term som refererar till samma föremål, så skall satsens extension, dvs. dess sanningsvärde, inte ändras. Det är en mycket övertygande princip: sanningshalten i vad jag säger borde inte bero på vilka namn jag använder för att tala om ett föremål, utan vad jag säger om föremålet.

uttryckens *mening* som i sådana kontexter utgör referensen. Satsen ”Jens Stoltenberg är Norges nuvarande statsminister” har en mening, som är den *tanke* som satsen uttrycker, och satsen ”Jens Stoltenberg är Jens Stoltenberg” har en annan mening, dvs. uttrycker en annan *tanke*. *Tankar*, är för Frege ett slags abstrakta objekt.

Kan man tillämpa Freges teori även på perceptionsakter och emotioner? Kan man säga att då jag ser min bil på parkeringen, så är objektet för min observationsakt inte själva bilen, utan meningen hos uttrycket ”min bil på parkeringen”. Det förefaller orimligt: inte kan man säga att jag, när jag ser bilen på parkeringen, riktar min uppmärksamhet på innebörden av orden, eller på det sätt jag identifierar min bil. Däremot verkar det rimligare säga att *hela uttrycket* ”ser min bil på parkeringen” är ett begrepp i *Freges mening*, dvs., en funktion som avbildar föremål på föremål, i detta fall avbildar personer på *Det Sanna* eller *Det Falska*.⁵ Mig veterligt har Frege inte diskuterat analysen av satser som uttrycker perceptioner eller emotioner, men man kunde kanske utvidga Freges teori och säga att man inte skall betrakta substantiv som termer, dvs. ord som i kontexten har referens, när de står inom omfånget för perceptions- och emotionsverb. Ett och samma ord kan alltså i en kontext vara en term, dvs. ett ord som refererar, och i en annan kontext inte vara en term.⁶ Det innebär att man avvisar analysen av mentala tillstånd såsom akter riktade emot objekt.

Det viktiga i sammanhanget är inte Freges lösningsförslag, eller dess möjliga tillämpning på Brentanos problem, utan att Frege startade en process bort från den direkta fokuseringen på medvetandetillstånd och i stället riktade uppmärksamheten på hur vi omtalar dem, det som brukar kallas den ”lingvistiska vändningen” i filosofin. Dummett hävdar (1991b, s. 111) att i § 62 i *Grundlagen der Arithmetik* börjar Frege med att fråga efter talens natur men slutar paragrafen med att fråga om meningen hos uttryck som innehåller talord. ”There is the linguistic turn” utbrister Dummett.

Man bör inte tolka detta som att Frege var den förste som vände uppmärksamheten från medvetandetillstånden till deras språkliga uttryck, för så var inte fallet. Quine (1981, s. 67) citerar John Horne Tooke som 1786 skriver om Locke: ”The greatest part of Mr Locke’s essay, that is, all which is related to what he calls abstraction, complexity, generalisation, relation, etc., of ideas, does indeed merely concern language.” Quine påpekar även att medeltidens nominalister hade samma grepp. Quine menar att steget från att tala om idéer till att tala om de språkliga uttrycken för dem var ett

⁵Sanningsvärdena är för Frege ett slags abstrakta objekt.

⁶Denna ståndpunkt vore helt i linje med kontextprincipen, som var en central idé hos Frege.

viktigt steg framåt och en klar förbättring av empirismen. Men det är ändå en rimlig bedömning säga att Frege startade den lingvistiska vändning som skedde omkring förra sekelskiftet.

Om vi anammar detta synsätt, så kan man omformulera Brentanos problem till: hur skall vi *semantiskt analysera utsagor* om perceptioner och emotioner? Detta är enligt min mening det bestående värdet av Freges bidrag, inte hans konkreta lösning. Det var ett stort steg framåt därför att våra utsagor är intersubjektivt tillgängliga, medan våra mentala tillstånd inte är det.

Det tog som bekant rätt lång tid innan den lingvistiska vändningen hade fått fullt genomslag i det som kom att bli den analytiska traditionen. Exempelvis har Russell inte tagit detta steg i sin berömda uppsats ”On Denoting” från 1905 och han avfärdar, även långt senare, på några ställen explicit studiet av hur vi använder språket som filosofiskt ointressant; för Russell, Moore och flera andra bland Cambridgefilosoferna var det studiet av begrepp som var filosofins uppgift. Men det är oklart vad de exakt menade med begrepp.

6. HUSSERLS LÖSNING PÅ BRENTANOS PROBLEM: FENOMENOLOGI

Enligt Dummett (”Thought and Perception”, s. 270) var Husserl och Frege eniga om att ett språkligt uttryck har objektiv referens endast i kraft av att ha en viss mening (*Sinn*). Men till skillnad från Frege hade Husserl ingen systematisk teori för hur mening bestämmer referensen för olika typer av uttryck. Detta är en fundamental skillnad, hävdar Dummett, eftersom för Frege är referensteorin det primära, det som motiverar meningsteorin.

Husserl och Frege var också eniga om att förekomsten av meningsfulla uttryck för vilka det saknas objektiva korrelat, dvs. vilka saknar referens, inte är något problem; det är själva poängen med att göra distinktionen mellan ett uttrycks mening och dess referens att kunna säga om uttryck som saknar referens att de ändå är meningsfulla.

Husserls fundamentala idé var, enligt Dummett (ibid. s. 271), att generalisera begreppet mening till att gälla inte bara språkliga akter utan även mentala akter. Det på detta sätt generaliserade begreppet är vad Husserl kallar ”noema”. Således menar Husserl att objektet för en mental akt ges av aktens noema: mentala akter är intrinsikalt noematiska och det är noemat som åstadkommer de mentala akternas intentionalitet. Genom att på detta sätt generalisera begreppet mening ansåg sig Husserl ha löst Brentanos problem. Precis som språkliga uttryck kan ha mening och ändå sakna objektiv referens, så kan mentala akter vara intentionala,

vara riktade emot ett objekt, utan att objektet existerar. Riktadheten ligger i akten, analogt med att meningsfullheten ligger i det språkliga uttrycket och inte i att det finns ett objektivt korrelerat.

Noema har således följande egenskaper enligt Husserl:

- (1) Det är noemat som gör att en förnimmelse är en förnimmelse *av* ett objekt.
- (2) Ett noema kan vara en gemensam ingrediens i flera förnimmelser, analogt med att olika språkliga uttryck kan ha olika mening men samma referens.
- (3) Noema kan variera mellan olika akter medan dessa akter är riktade emot samma objekt, också analogt med att satser med olika mening kan ha samma referens.

Husserl anser att vi vanligen inte uppmärksammar noemat, men att det inte är omöjligt. Precis som vi när vi talar inte tänker på satsernas (eller satsdelarnas) mening, utan på de ting och händelser vi talar om, så uppfattar vi inte noemat i de mentala akterna utan endast de föremål som akterna är riktade emot. Men det är inte omöjligt att varsebli noemat; genom träning skulle man kunna göra det till föremål för ens uppmärksamhet.

Dummett hävdar att Husserls idé att generalisera begreppet mening (*Sinn*) till något som även kan appliceras på mentala akter kräver att man har en klar uppfattning om vad *Sinn* är och kan se i vilka avseenden det går att generalisera. Att bara *säga* att man generaliserar ett begrepp är inte att *göra* det.

Det är just på denna punkt som separationen mellan analytisk filosofi och fenomenologi sätter in, enligt Dummett. Hans värdering är att den språkliga vändningen som Frege startade ledde till fruktbar utveckling av många filosofiska frågeställningar, men att många viktiga begrepp försvann från horisonten. Husserls fenomenologi å andra sidan, kan behandla dessa begrepp, men är i sin kärna oklar. Dummett skriver: "Husserl turned philosophical attention away from language, but left an essential vagueness at its centre" (s. 287).

7. EVIDENS FÖR PÅSTÅENDEN OM DE MENTALA TILLSTÅNDENS FORMER

Fenomenologins beskrivning av de mentala akterna grundar sig inte på empiriska iakttagelser utan på väsensskådande, enligt Husserl. Vad är det? Hur åskådar man en perceptionsakts väsen? Kan det vara introspektion? I så fall har man anledning att vara skeptisk: det finns i psykologisk forskning en omfattande diskussion om introspektion och den generella

slutsatsen är att den är högst opålitlig. Vanliga observationer med våra sinnen är även de ibland opålitliga, men introspektionens osäkerhet är av en helt annan dignitet.

Det kan vara på sin plats att påminna sig hur Wittgenstein (i *Filosofiska undersökningar*) kritiserar föreställningen om att vi genom introspektion kan få en klar uppfattning om våra sinnestillstånd. Anta att jag har en ny sorts erfarenhet vid en viss tidpunkt, en upplevelse som inte är relaterad till något objekt som är intersubjektivt tillgängligt, utan är frikopplad från yttervärlden. Jag hittar på ett ord för denna upplevelse och antecknar det i min almanacka. Vid en senare tidpunkt har jag samma sorts upplevelse och antecknar ordet igen. Men, frågar sig Wittgenstein, hur kontrollerar jag att det faktiskt är *samma* sorts upplevelse de båda gångerna, dvs. att jag korrekt använder samma ord för de två upplevelserna? Det enda kriterium man har att gå på är att en viss upplevelse vid en viss tidpunkt *ter sig* för mig som likadan som minnet av en tidigare upplevelse. Men minnet av den tidigare upplevelsen är inte samma sak som den tidigare upplevelsen själv. Och även om man kunde garantera minnets korrekthet, så har man ändå inget annat att gå på än att den aktuella upplevelsen och minnet av en tidigare upplevelse förefaller subjektet likadana. Dvs. kriteriet för korrekthet är att det *ter sig* för subjektet korrekt att använda samma ord. Men att något *ter sig* korrekt och att det är korrekt är inte samma sak. Det finns således inga kriterier för korrekt användning av ordet. Wittgensteins slutsats är att en sådan rent privat stipulation av ett ords betydelse inte kan räknas som en form av språk-användning och han drar slutsatsen att det inte kan finnas något sådant som ett privat språk.

Men hur lär sig då små barn användningen av känslord såsom ”arg”, ”ledsen” eller ”rädd”? Som all språk-användning så måste inläringen av dessa ord bygga på intersubjektivt observerbara omständigheter. Alla som har eller har haft småbarn och funderat på deras språkutveckling kan nog fylla i detaljerna i en beskrivning av hur det går till. Det centrala är att känslorna är reaktioner på yttre iakttagbara händelser och knappast någon förälder behöver fundera över vilka känslor barnen har; de uttrycks vanligen mycket tydligt, även utan språk.

Jag finner Wittgensteins argument övertygande och drar slutsatsen att fenomenologers beskrivningar av strukturen hos perceptionsakter också måste vara underkastade intersubjektivt kontrollerbara kriterier för korrekthet. Hur går detta ihop med väsensskådande? Det kan uppenbarligen inte vara en rent subjektiv process.

Min gissning är att trovärdigheten i beskrivningen av perceptionernas struktur såsom bestående av en akt riktad emot ett aktobjekt bygger på att det är en beskrivning av strukturen hos de *satser* som vi använder för att

tala om våra perceptioner av yttre publikt tillgängliga objekt. Såväl dessa yttre objekt, som de satser vi använder för att tala om dem, är publika och gemensamma för alla deltagare i språkspelet. Detsamma gäller emotioner och propositionella attityder.

De verb vi använder för att beskriva våra mentala akter kan användas både transitivt och intransitivt, men den transitiva användningen är den primära; man kan säga ”jag ser” eller ”jag är arg” i vissa kontexter där objektet för akten är så uppenbart för lyssnaren att det inte behöver nämnas. Men om det av sammanhanget inte framgår vad man ser, eller på vem man är arg, så uppfattar vi utsagan som ofullständig. Således har de predikat med vilka vi beskriver våra mentala tillstånd formen <verb+substantiv> (perceptioner), <verb+preposition+substantiv> (emotioner) eller <verb+att+fullständig sats> (propositionella attityder). Så min hypotes är att den struktur som Brentano och Husserl fann i medvetandeakterna fann de i språket. Väsenskådande är i så fall tyst tal med sig själv.

När vi tänker på en mental akt, dvs. tyst talar med oss själva och så att säga betraktar våra perceptioner, emotioner och tidigare tankar, använder vi samma kognitiva processer som då vi producerar en sats som uttrycker vår tanke; den enda skillnaden mellan att tänka och att tala är att då vi tänker utan att säga något så aktiveras inte motorcortex, den del av hjärnan som bl.a. styr talet, medan detta sker då vi utsäger tanken.

Man kunde hävda att den språkliga strukturen hos våra satser som beskriver våra mentala akter reflekterar strukturen hos de mentala akterna, och att observationer av satser är indirekt evidens för strukturen hos de mentala tillstånden. Men hur vet vi det? Har vi någon oberoende evidens, dvs. av språket oberoende evidens, för att de mentala tillstånden har denna struktur? Jag kan inte se hur det skulle kunna vara möjligt.

8. MEDVETANDETS AKTIVITET

Såväl naturalism, som transcendentalfilosofi och fenomenologi är eniga om att perceptionsakter är aktiva processer i vilka våra medvetanden bearbetar och organiserar vad som kommer från yttrevärlden. Så vad är då skillnaderna? Den ena skillnaden är att naturalismen anför empiriska, vetenskapliga argument för denna tes, medan transcendentalfilosofi och fenomenologi anför icke-empiriska argument. Den andra skillnaden är att för naturalismen är medvetandet ett objekt i den empiriska världen och inget annat; något transcendentalt perspektiv på medvetandet som Kant menar är möjligt, eller möjligheten av att det kan bli föremål för icke-empiriska kunskapsakter, väsensskådande, som Husserl hävdar, erkänner inte naturalismen.

9. REPRESENTATIONALISM OCH DESS ALTERNATIV

Lockes empirism är en representationsteori; våra idéer representerar tillstånd och fakta i yttrevärlden. Men varken Humes naturalism, Kants transcendentalism eller Husserls fenomenologi är representationsteorier; dessa utgör olika utvägar för dem som ser varje form av representations-teori som ohållbar. Men deras olika alternativ delar två svagheter: alla tre tar för givet att man direkt har tillgång till sina egna medvetandetillstånd och alla tar för givet att man kan generalisera dessa iakttagelser till att gälla allas medvetanden. Alla talar om medvetandet i bestämd form och läsaren förstår det som något allmänt. Ingetdera antagandet är välgrundat: tvärtom, empirisk forskning säger snarast det motsatta vad gäller det första antagandet⁷ och för det andra antagandet ger man inget argument.

Min uppfattning är att det finns starka skäl emot representationsteori, dvs. antagandet att det i medvetandet finns en sorts mentala objekt som representerar de yttre observerade objekten. Perceptioner är perceptioner av yttre objekt, och perceptioner studeras bäst empiriskt, inte genom introspektion och filosofisk analys.

10. VILLKOR FÖR KUNSKAP

Kants transcendentalism gör en skarp åtskillnad mellan analyser av medvetandets funktioner och av empiriska fenomen. När Kant analyserar villkoren för kunskap och identifierar dessa som strukturegenskaper hos medvetandet är detta medvetande inget empiriskt objekt, utan ett transcendentalt objekt, dvs. ett objekt för transcendental analys. (Man kan också studera medvetandet rent empiriskt; då är det ett empiriskt objekt.) Åtskillnaden mellan empiriska objekt och medvetandet som transcendentalt objekt är en följd av Kants åtskillnad mellan två nivåer av diskurs: den empiriska diskursen och den transcendentala diskursen, dvs., analysen av villkoren för kunskapens möjlighet.

För en naturalist är det inte främmande att diskutera vilka villkor som måste vara uppfyllda för att man skall kunna veta något. Men det strider emot naturalismen att säga att det är en diskurs på en annan nivå. Medvetandeprocesserna är naturliga processer som försiggår i den fysiska världen (ehuru beskrivna med mentala predikat), som påverkas av fysiska händelser och som orsakar fysiska händelser. När vi utforskar dessa empiriskt kan vi dra slutsatser om villkoren för kognition. Och om vi förkastar introspektion som metod och utgår ifrån att evidens för påståenden om medvetandetillstånd är empirisk evidens, dvs. intersubjektivt tillgänglig evidens, finns det inget transcendentalt

⁷Någon, jag vet inte vem, har vitsigt uttryckt detta som "Hur skall jag veta vad jag tänkt innan jag hört vad jag sagt?"

studieobjekt. Med den utgångspunkten försvinner möjligheten av en diskurs på två nivåer.

Modern naturalism och modern empirism utgår i stället från att observationer av intersubjektivt tillgängliga ting och händelser utgör grunden för kunskapsanspråk. Mer precist, det är observationssatser, vars sanningsvärden vi är eniga om, som utgör den intersubjektiva grunden för vad som kan räknas som de yttersta skälen för att tro något. Ingen helt säker grund, förvisso; men den bästa grund vi har.

Många har hävdats att det inte finns något sådant som en helt teoriberoende observationssats. Det är förvisso sant att mycket av det som vanligen kallas observerbart i hög grad är teoriberoende; man säger t.ex. att man observerar galaxer som befinner sig flera miljarder ljusår bort. Dessa observationer görs med avancerad teknisk utrustning vars konstruktion och funktion är avhängig av tillämpningen av många avancerade teorier.

Men det går att identifiera en klass av helt teoriberoende observationssatser. Quine definierar en observationssats såsom en sats som beskriver en här- och nu-situation, och där beskrivningen är sådan att alla som är närvarande och som talar samma språk, omedelbart är överens om dess sanning (eller falskhet).

11. SLUTORD

Vi har inget behov av en första filosofi, en säker grund för andra vetenskapers anspråk på att uppnå kunskap, eller för den delen, någon apriorisk grund för våra anspråk på att veta något överhuvudtaget. Vetenskaplig och vardaglig kunskap behöver inga garantier.

Kunskapsteoretisk naturalism är intimt förbundet med ett vetenskapligt förhållningssätt. Man skall kunna ge empiriskt stöd för sina påståenden och för att något skall räknas som empiriskt stöd skall det vara intersubjektivt tillgängligt. Varken transcendentalism eller fenomenologi uppfyller detta krav och därför förefaller dessa doktriners påståenden, om t.ex. åskådningsformer, kategorier, väsensåskådande, medvetandeakternas riktadhet, noemat osv., så mycket mer tvivelaktiga än det de är tänkta att stödja, dvs. anspråken på att vi vet saker, såsom att Vätern är Sveriges största sjö.

LITTERATUR

- Carnap, R. 1928. *Der logische Aufbau der Welt*. Berlin: Weltkreis.
Dummett, M. 1991a. "Thought and Perception: The Views of Two Philosophical Innovators". I *Frege and Other Philosophers*, s. 263–88. Oxford: Oxford University Press.

- Dummett, Michael. 1991b. *Frege: Philosophy of Mathematics*. London: Duckworth.
- Frege, Gottlob. 1884. *Die Grundlagen der Arithmetik: Eine logisch mathematische Untersuchung über den Begriff der Zahl*. Breslau: W. Koebner. Sv. övers. *Aritmetikens grundvalar*. Övers. Fredrik Stjernberg. Stockholm: Thales, 2002.
- Hume, David. 2002. *Avhandling om den mänskliga naturen*, Bok 1, *Om förståndet*. Övers. Robert Callergård. Stockholm: Thales.
- Kant, Immanuel. 2002. *Prolegomena*. Övers. Marcel Quarfood. Stockholm: Thales.
- Quine, W. V. 1981. "Five Milestones of Empiricism". I *Theories and Things*, s. 67–72. Cambridge, Mass.: Harvard University Press.
- Quine, W. V. 1969. "Epistemology Naturalized". I *Ontological Relativity and Other Essays*, s. 69–90. New York: Columbia University Press.
- Quine, W. V. 1995. *From Stimulus to Science*. Cambridge, Mass.: Harvard University Press.
- Wittgenstein, Ludwig. 1992. *Filosofiska undersökningar*. Övers. Anders Wedberg, rev. av Lars Hertzberg. Stockholm: Thales.

Stina Bäckström

Hur hamnade Sanaz på festen?

*Hornsby och Stewards kritik av
Davidsons handlingsmetafysik*

De grundläggande utgångspunkterna i Donald Davidsons handlings-teori – att handlingar är händelser orsakade av mentala tillstånd, och att de relevanta orsakssambanden kan beskrivas i termer av strikta orsakslagar – har länge fungerat som ett slags standarduppfattning inom analytisk handlingfilosofi. Mycket arbete har lagts ned för att bemöta problem som uppstår inom detta ramverk utan att ge upp de grundläggande teserna. Men en växande skara filosofer menar att problemen kan lösas endast genom att ifrågasätta Davidsons utgångspunkter. I denna uppsats diskuterar jag hur två av dessa kritiker, Jennifer Hornsby och Helen Steward, ifrågasätter de orsaks- och händelsebegrepp som utgör två av stöttepelarna i Davidsons handlingsteori.

Först kommer jag ge en överblick över den plats Davidsons orsaks- och händelsebegrepp har i hans handlingsteori. Efter det kommer jag återge en kritik av dessa begrepp, som återfinns hos både Hornsby och Steward. Jag kommer också beskriva några element i Stewards ambitiösa projekt att skapa en handlingsteori utifrån en annan metafysik än den Davidson förutsätter. Jag avslutar med en fråga till Steward.

1. DAVIDSONS HANDLINGSMETAFYSIK

Hur ser Davidson på orsaker, och vilken roll spelar hans orsaksbegrepp i hans handlingsteori? Principen om Kausalitetens Nomologiska Karaktär är en av de centrala utgångspunkterna i Davidsons handlingsteori. Enligt denna princip är *orsak och verkan händelser som följer på varandra i ett strikt lagbundet mönster*.¹ Det är en av de tre principer som ger upphov till det trilemma som han formulerar i uppsatsen "Mental Events".² Att förstå detta trilemma och hur Davidson försöker upplösa det ger en

¹En utläggning av vad denna princip innebär för Davidson finns i uppsatsen "Causal Relations" (Davidson 2001a). Han leder principen tillbaka till Hume, och Humes begrepp om orsak och verkan som konstant konjunktion av händelser.

²Davidson 2001a, i svensk översättning i Davidson 2004.

grundläggande översikt över hur Davidson ser på handlingar, deras upphov och förklaringar.

Om den första principen bakom trilemmat är Principen om Kausalitetens Nomologiska Karaktär, så är den andra principen bakom detta trilemma att mentala händelser orsakar fysiska händelser (avsiktsförklaringar av handlingar anses av Davidson vara ett exempel på detta).³ Den tredje principen är att det inte finns några psykofysiska lagar.⁴ Hur utgör då dessa principer ett trilemma? Låt oss tänka på saken genom ett exempel. Sanaz har lust att dansa, och hon tror att det kan bli dans på festen hon är inbjuden till. Så hon går till festen. Om vi tar påståendet ”Sanaz gick till festen för att hon hade lust att dansa”, så anser Davidson att vi där beskriver ett orsakssamband mellan händelser, vilket exemplifierar den andra principen att mentala tillstånd orsakar fysiska händelser.⁵ Och där det finns orsakssamband finns det också en lag, enligt den första principen om orsakssambandens lagbundenhet. Men här tycks det ju inte finnas någon lag, eftersom det enligt Davidsons tredje princip inte finns några *psykofysiska* lagar. Det finns ingen lag som binder samman händelser av typen *känna lust att dansa*, och händelser av typen *gå på fest*. Hur kan då påståendet ”Sanaz gick på fest för att hon hade lust att dansa” beskriva ett orsakssamband?

Davidsons lösning är att hävda att det finns en lag, men inte en lag som innefattar mentala termer, utan en lag som beskriver *samma händelser* men med helt andra begrepp. Tricket är, enligt Davidson, att vi kan förstå en enskild avsiktsförklaring – exempelvis ”Sanaz gick till festen för att

³Denna principens betydelse och giltighet diskuterar Davidson minst av de tre principerna. Hornsby och Steward skulle ifrågasätta, av skäl som kommer bli tydliga så småningom, att denna princip är giltig så som Davidson förstår den.

⁴Davidson 2001a, s. 208–209, Davidson 2004, s. 154–155. Davidson diskuterar denna princip på många ställen, bland annat i flera av uppsatserna i *Inquiries into Truth and Interpretation* (Davidson 2001b). Ett viktigt led i dessa diskussioner är tanken att det ramverk inom vilket mentala begrepp får sin betydelse är styrt av *rationella principer* snarare än naturlagar. Kathrin Glüers systematiska översikt över Davidsons filosofi ger en utmärkt ingång till denna och Davidsons andra grundläggande teser (Glüer 2011).

⁵Jag förenklar här Davidsons syn på handlingars orsaker. Han menar att det är karakteristiskt för handlingar att de har en förklaring i agentens primära skäl (*primary reason*). Att ha ett primärt skäl är enligt Davidson att ha en proattityd till en handling av en viss typ och en trosföreställning att en partikulär handling faller under denna typ. Dessa attityder är inte enligt Davidson mentala *händelser*, utan snarare dispositioner. Men en förklaring av en handling i termer av personens skäl syftar ändå till en mental händelse som är handlingens orsak, nämligen händelsen att agenten formade eller kom att ha dessa attityder (Davidson 2001a, s. 12).

hon hade lust att dansa” – som att den hänvisar till händelser som här beskrivs med mentala termer, men som också kan beskrivas med helt andra begrepp. I synnerhet kan vi beskriva dem med de naturvetenskapliga begrepp som vi använder när vi formulerar naturlagar. När vi gör dessa ombeskrivningar, menar Davidson, så kommer ett lagbundet kausalsamband att visa sig, och då ser vi att de händelser som förut var beskrivna med mentala begrepp står i en orsak-verkan relation till varandra. Den relevanta lagen formuleras alltså i helt andra termer än de mentala begrepp som ingår i avsiktsförklaringen. För att ”Sanaz gick till festen för att hon hade lust att dansa” ska vara en orsaksförklaring krävs *inte* att det finns en lag som binder samman händelser av typen ”ett festbesök” med typen ”känna lust att dansa”. Vad som krävs är att de händelser som beskrivs i syftesförklaringen kan ges en helt annan beskrivning, och givet den helt nya beskrivningen så finns det en lag. Mentala händelser och handlingar är således identiska med orsaker och verkan som kan beskrivas i lagbundenheter. När vi ger en syftesförklaring gör vi, enligt Davidson, ett minimalt antagande, nämligen att det existerar en regelbundenhet – som kan beskrivas i en lag – under vilken dessa händelser faller (även om vi inte känner till denna lag).⁶

Denna lösning på trilemmat kallar Davidson för *anomal monism*. Anomal monism är, som termen antyder, en position som förenar två element: monism och anomal. *Monismen* står i primär kontrast till kroppsjäl *dualism*, och består i ett ontologiskt antagande; nämligen att mentala händelser (t.ex. tankar, känslor och avsikter) är identiska med fysiska händelser, och på grund av detta kan ges en beskrivning enligt vilken de är bundna av strikta kausala lagar. Detta är inte en identifikation av *typer* av mentala händelser med *typer* av händelser som förekommer i formuleringen av lagen (det som brukar kallas för *type-type identity*). Snarare är det en identitet på de enskilda händelsernas nivå: en enskild mental

⁶Jag återger här en standardtolkning av Davidsons position vad det gäller relationen mellan avsiktsförklaringar och orsaksrelationer. Fred Stoutland argumenterar i uppsatsen ”Intentionalists and Davidson on Rational Explanation” (Stoutland 1999) för en annan tolkning av Davidson. Enligt Stoutland menar Davidson inte att en vardaglig syftesförklaring innehåller ett antagande om att personens skäl är handlingens orsak och därmed kan beskrivas i fysiska termer. Det implicita antagandet är snarare, enligt Stoutlands Davidson, att det *finns* en händelse som är *associerad* med personens skäl och som är handlingens orsak. Stoutland anser att om vi förstår Davidson på detta sätt så blir skillnaden mellan hans och von Wrights position – som ofta framställts som motsatta poler – till en stor del terminologisk. Hur vi förstår Davidson på denna punkt har dock liten betydelse för diskussionen i denna uppsats eftersom Hornsby och Steward skulle finna samma skäl att ifrågasätta den Davidson som Stoutland framställer.

händelse är identisk med en fysisk händelse (så kallad *token identity*). Att denna monism är *anomal* innebär, som Davidson förklarar saken, att våra beskrivningar av människors handlande inte kan inordnas i eller reduceras till beskrivningar av strikta lagbundenheter mellan händelser. Våra beskrivningar av människors avsikter och handlingar begripliggör väsentligen människor i relation till rationella normer och ideal, inte som instanser av strikta naturlagar.⁷

En väsentlig del av Davidsons anomala monism är således att tankar, avsikter, likväl som handlingar är händelser (eller är intimt förknippade med händelser) som kan ges en beskrivning utan att använda mentala begrepp. En viktig fråga blir då hur Davidson ser på händelser. Han behöver ett händelsebegrepp som är sådant att han kan säga att *en och samma händelse* kan beskrivas i mentala termer och i fysiska termer. Det gör han genom att definiera händelser på det mest neutrala och abstrakta sätt som är möjligt. Händelser har, enligt Davidson, ingen natur annat än att de står i kausala samband med andra händelser. Identifikationskriteriet för händelser är, enligt Davidson, att om två händelser har samma orsak och samma verkan är de två händelserna identiska. Händelser är således en kategori given endast i termer av det nomologiska orsaksbegreppet, och det ska inte förutsätta något annat om naturen hos *olika typer* av händelser.⁸ (Nedan kommer jag att uttrycka detta som att Davidson förutsätter ett *odifferentierat* händelsebegrepp.⁹)

Den förståelse av vad en handling är som Davidson formar inom ramen för den anomala monismen är då (i förenklad form) följande. ”Sanaz gick till festen för att hon hade lust att dansa” beskriver ett orsakssamband mellan två händelser, en mental händelse (Sanaz kände lust att dansa), och en fysisk händelse – en kroppsrörelse (eller en serie kroppsrörelser) inklusive dess konsekvenser (Sanaz gick till festen). Den mentala händelsen orsakar kroppsrörelsen (som i sin tur ger upphov till vidare konsekvenser), men det är endast under helt andra beskrivningar av dessa händelser som vi kan se under vilken lag dessa händelser faller.¹⁰

⁷Davidson 2001a, s. 209–15, Davidson 2004, s. 156–70.

⁸Davidson 2001a, s. 179.

⁹I sin ”Reply to Quine on Events” säger Davidson att detta sätt att identifiera händelser är otillräckligt. Men det är tydligt att han söker efter ett kriterium som liknar det första på det sättet att det bör vara neutralt och inte förutsätta något om naturen hos sina underkategorier (2001a, s. 309–11).

¹⁰Se Davidson 2001a. Framför allt i uppsatserna ”Actions, Reasons and Causes” och ”Agency”.

2. HANDLINGAR, HÄNDELSER OCH DERAS ORSAKER

Davidsons anomala monism har kritiserats på en rad olika punkter. En vanlig kritik är att positionen kollapsar till en form av epifenomenalism: att avsikter, tankar och känslor som sådana inte gör någon skillnad överhuvudtaget, eftersom det är på en helt annan beskrivningsnivå – där vi identifierar helt andra egenskaper hos händelser än de mentala egenskaperna – som vi finner det som gör skillnad. Tankar, känslor och avsikter tycks bli epifenomen – inte *som sådana* relevanta för händelseförlopp.¹¹

Denna invändning är intressant – jag har alltid varit förbryllad över hur vi, enligt Davidson, kan få en genuin förståelse av orsakssamband i världen genom våra vardagliga avsiktsförklaringar. Människor och deras avsikter och förmåga att göra skillnad tycks irrelevanta, när det är i en beskrivning i helt andra begrepp (där termer som ”person” och ”avsikt” inte förekommer) vi får syn på hur saker *egentligen* hänger ihop. Men det är inte denna invändning jag vill diskutera här utan en variant på en annan vanlig kritik, kritiken att Davidson *utelämnar agenten*. Den variant av kritiken jag är intresserad av är Hornsby och Stewards tanke att Davidson genom att förutsätta Principen om Kausalitetens Nomologiska Karaktär samt det händelsebegrepp som är dess korrelerat – att händelser är givna i termer av deras plats i orsakskedjor – inte lämnar något utrymme vare sig för *handlingar* eller för *djur och människor som handlar* (jag kommer hädanefter följa Hornsby och Steward och kalla djur och människor som handlar för ”agenter”). Tanken här är att Davidsons orsaks- och händelsebegrepp *utesluter* både att det finns agenter och att det finns handlingar.

Stewards och Hornsby variant av kritiken att agenten utelämnas skiljer sig från hur exempelvis David Velleman (1992) artikulerar agentproblematiken. Velleman anser att Davidsons handlingsteori är otillräcklig, men han vill inte ge upp de centrala förutsättningarna bakom den. Han menar att ”Standardhistorien [dvs. Davidsons handlingsteori] kan inte anklagas endast för att den hänsyftar på mentala tillstånd och händelser som sker i agentens medvetande [*mind*]. Vilken förklaring som helst av mänskligt handlande kommer att tala i termer av sådana händelser, därför att händelser är de grundläggande elementen i förklaringar generellt”.¹² Hornsby och Steward är mer radikala i sina avsteg från Davidsons teori. Enligt dem är det lönlöst att försöka hitta ett utrymme för agenten och agenter handlingar inom ramen för Principen om Kausa-

¹¹Brian P. McLaughlins uppsats ”On Davidson’s Response to the Charge of Epiphenomenalism” (1993) är en bra ingång till debatten om epifenomenalism.

¹²Velleman 1992, s. 468. Detta och resten av citaten i denna uppsats är oversatta av mig.

litetens Nomologiska Karaktär. Vi behöver ett helt annat orsaksbegrepp för att rimliggöra vad en handling är.

Vi kan börja med Hornsby. I en nyligen publicerad artikel ställer Hornsby Davidsons syn på handlingar och handlingars orsaker mot det hon betraktar som Elizabeth Anscombes handlingsteori. Hon hävdar att det finns en radikal skillnad i deras syn på vad handlingar är och hur de är orsakade (och det råder inga tvivel om att Hornsby föredrar Anscombe). Hornsby skriver:

Davidson tillskrev Anscombe inspirationen till sin egen redogörelse för handling. Men Davidson antog att mänskligt agentskap finns i en värld där kausal verksamhet är begränsad till relationer mellan saker i kategorin händelser eller tillstånd; och detta garanterar att hans uppfattning om agentskapets och handlingens natur är ytterligt olik Anscombes. (Hornsby 2011, s. 105)

Vad är det då vi finner hos Anscombe som inte finns hos Davidson, enligt Hornsby? Anscombe, menar Hornsby, framställer våra vanliga beskrivningar av handlingar på ett sådant sätt att en annan typ av kausalitet än den nomologiska händelsebaserade varianten *tas för given*. Detta illustrerar Hornsby med en rad handlingsverb som Anscombe kallar för "kausativa" (*causative*) verb. Exempel på kausativa verb är "skrapa", "bära", "dra", "välta" och "tillverka". När vi tillämpar dessa verb för att beskriva en persons handling så förutsätter vi en typ av orsakssamband, men inte den typ som Davidson tänkte sig. Säg att vi påstår att "Josefin välter stenen". Här påstår vi att Josefin orsakar en förändring, nämligen en förändring av stenens läge. Och vi förutsätter att hon gjorde den genom att, just, *välta* den. Att välta något är enligt Hornsby's Anscombe en process som innefattar *en agents förändring av någonting annat*. Att förändring sker *och* varför, är så att säga inbegripet i själva processen; kausaliteten är som Hornsby uttrycker det, "innefattad i en händelse".¹³

Davidson skulle nu naturligtvis hålla med om att en handlingsbeskrivning förutsätter ett orsakssamband. Det var ju den första principen i hans trilemma. Men han menar att det inte är något annat än det händelsebaserade orsakssambandet som förutsätts. Han skriver: "Kausalitet är centralt för handlingsbegreppet, men det är vanlig kausalitet mellan händelser som är relevant."¹⁴ Så det är där motsättningen finns mellan Anscombe och Davidson, enligt Hornsby. Om vi lyssnar på Hornsby's Anscombe så kan vi inte reducera den form av kausalitet som handlingsbeskrivningen förutsätter, till den händelsebaserade kausaliteten. Vi måste förutsätta att *agenter* kan ge upphov till

¹³Hornsby 2011, s. 107. Originalets engelska lyder: "internal to an event".

¹⁴Davidson 2001a, s. 53.

förändring, och detta står i konflikt med tanken att den enda relevanta kausaliteten är den som artikuleras i strikt lagbundna sambanden mellan händelser.

Detta sista påstående är en av de centrala teserna i Stewards nyutkomna monografi *A Metaphysics for Freedom* (2012). Hur argumenterar då Steward för detta påstående? Jag rekonstruerar ett av resonemangen med ett ganska fritt förhållningssätt till Stewards framställning, eftersom argumentationen sträcker sig över många sidor i hennes bok. (Kapitel 8 är det mest relevanta stället att titta på.) Detta arguments slutsats är att en beskrivning av en händelse i termer som inte förutsätter agentkausalitet, vilket Davidsons teori tycks kräva, inte längre är en beskrivning av en handling. *Agentkausalitet* är i sin tur ett begrepp som Steward ägnar mycket tid åt att förklara, men vi kan nöja oss så länge med att säga att det innebär att agenter ger upphov till förändring genom att handla.

Ta Sanaz och den handling som hon utförde, nämligen att gå till festen. Enligt Stewards resonemang är det inte möjligt att korrekt beskriva just detta förlopp i termer som inte förutsätter en agent och agentens handling. Vi kan förstås beskriva förloppet som en händelse "Sanaz förflyttning av sig själv till fots hemifrån till festen mellan kl. 19.30 och 19.45". Men denna beskrivning *inbegriper* en handlingsbeskrivning: "Sanaz förflyttning av sig själv till fots." Handlingsbeskrivningen är det som erbjuder oss möjligheten ringa in den relevanta händelsen. Skulle beskrivningen inte inbegripa en handlingsbeskrivning, menar Steward, så skulle vi inte längre kunna ringa in den relevanta händelsen. Det vill säga: om vi väljer att beskriva detta förlopp som en händelse, så *inbegriper* antingen denna beskrivning en handling – och förutsätter således ett annat begrepp om kausalitet än det händelsebaserade – eller så gör den inte det, i vilket fall beskrivningen då inte längre fångar in den relevanta händelsen.

Vi kan notera, först, att för att beskrivningen "Sanaz förflyttade sig själv till festen" ska vara sann så krävs inte bara att Sanaz befann sig hemma kl. 19.30, och sedan på festen kl. 19.45, utan det krävs att Sanaz *förflyttade sig själv* till festen. "Förflytta sig" tycks vara ett exempel på det som Hornsby via Anscombe kallade ett kausativt verb, i likhet med "välta". Det är väsentligt för att *just detta* ska ha ägt rum att det var *Sanaz* som orsakade en förändring genom att förflytta sig till fots. Men om det stämmer, så blir beskrivningen av handlingen som en händelse en *abstraktion*, en abstraktion vars giltighet förutsätter Sanaz och Sanaz handling. Det är så att säga så att det inte finns någon händelse att beskriva om det inte fanns en handling som utfördes. Och handlingen kräver någon som handlar.

Kan vi inte då bryta ner det vi kanske först ville betrakta som en hel-

het, *handlingen*, till en serie mer grundläggande händelser? Davidson menade, som jag nämnde ovan, att det vi vanligen beskriver med ett handlings-verb inbegriper en kroppsrörelse inklusive dess vidare effekter. Kan då inte "Sanaz gick till festen" förstås som en sekvens av på varandra följande kroppsrörelser inklusive deras konsekvenser? En sådan serie händelser kan beskrivas så här: "Sanaz vänstra knäled böjde sig, vilket fick hennes fot att förflytta sig från A till B, sedan böjdes hennes högra knäled, och så vidare; och detta resulterade i att Sanaz rumsliga position ändrades till att sammanfalla med festlokalens."

Stewards reaktion på detta försök från Davidsons sida att försvara tanken att handlingar är händelser orsakade av andra händelser, skulle vara att denna beskrivning är helt förenlig med att det som hände *inte* var att en handling utfördes, och att den således skulle kunna ringa in en helt annan händelse än den vi är ute efter. Sanaz ben kan röra sig utan att *Sanaz rör dem*. (Hon skulle också påpeka att denna beskrivning av en serie händelser innehåller termerna "Sanaz" och "ben", vilkas giltighet som begrepp, anser hon, inte kan antas utan att förutsätta agenter och därmed agentkausalitet. Men att visa hur hon tänker här skulle innebära ett avsteg från det resonemang vi nu för.) Och om inte Sanaz har rört sina ben, så är det heller inte längre sant att *hon har gått till festen*. Beskrivningen "Sanaz ben rörde sig från A till B ..." stämmer lika bra (om inte bättre) som en beskrivning av hur Sanaz blev manipulerad av en dockspelare som fick Sanaz ben att röra sig genom att dra i trådar. Således har vi inte ringat in den *handling* som också är en händelse på detta sätt.

Vad skulle då Davidson säga om detta? Det omedelbara svaret skulle vara att Sanaz kroppsrörelse *måste vara orsakad av hennes mentala tillstånd* för att rörelsen skall vara en handling. Om det är dockspelaren som rör Sanaz ben, skulle Davidson hävda, så är kroppsrörelserna inte orsakade av Sanaz mentala tillstånd, vilket, skulle han säga, var förutsett i handlingsbeskrivningen. Så frågan blir då om vi kommer undan problemet genom att utöka beskrivningen på följande vis: "Sanaz lust att gå till festen orsakade hennes vänstra knäled att böja sig, vilket fick hennes fot att förflytta sig från A till B, etc." Detta tycks undanröja oron att vi beskriver något som lika gärna skulle kunna vara en dockspelares manipulation av Sanaz kropp. Men, och detta insåg Davidson själv, det är inte så enkelt. För Sanaz lust att gå till festen kan orsaka Sanaz ben att röra sig utan att Sanaz har handlat. Detta är det klassiska problemet för Davidsons teori, nämligen det som kallas "avvikande kausalkedjor" (*deviant causal chains*).

Davidson noterade själv att avvikande kausalkedjor var ett problem för tillräckligheten hos karakteriseringen av handlingar som händelser

orsakade av mentala tillstånd.¹⁵ Där beskriver han en klättrare som håller en annan person i ett rep. Klättraren känner lust att göra sig av med vikten och faran det innebär att hålla i den andra personen, och han vet att han kan göra sig av med faran och vikten genom att släppa taget. I förskräckelsen över denna hemska tanke tappar han taget om repet. Här har en kroppsrörelse orsakats av hans lust att släppa taget, men klättraren har inte därmed utfört *handlingen* att *släppa* taget om repet. Davidson skriver: "Denna trosföreställning och vilja kan göra honom så nervös att det orsakar honom att tappa sitt grepp. Ändå kan det vara så att han aldrig *valde* att tappa sitt grepp, eller att han gjorde det avsiktligt."¹⁶ Det Davidson ansåg att detta visade var att det inte räcker att en kroppsrörelse är orsakad av vissa mentala tillstånd, den måste vara orsakad "på rätt sätt". Själv ansåg han att det inte skulle gå att ge någon substantiell redogörelse för vad det innebär att en handling är orsakad på rätt sätt. "För mig ter det sig oöverstigit", skriver Davidson om problemet med avvikande kausalkedjor.¹⁷ Mycket bläck har sedan förbrukats på att försöka visa att det inte alls är oöverstigit. (Vellems teori är ett exempel på ett sådant försök.)

Enligt Steward är avvikande kausalkedjor ett symptom på hur det inte är möjligt att göra reda för handlingar inom ramen för tanken att handlingar är händelser orsakade av mentala tillstånd. Åtminstone så länge vi förutsätter Principen om Kausalitetens Nomologiska Karaktär, och den syn på händelsers identitetskriterier som är dess korrelat. För Steward vill inte förneka att vi kan tala om orsakssamband här, så väl som om händelser. Men Davidson förutsätter ett *odifferentierat* händelsebegrepp och ett odifferentierat orsaksbegrepp, medan Steward vill att vi ska anamma *differentierade* begrepp.

Vad innebär då detta? Det centrala problemet med avvikande kausalkedjor botas inte, enligt Steward, genom att skapa en teori om vad det innebär att en handling måste vara orsakad på "det rätta sättet". Snarare måste vi, enligt Steward, erkänna agent-kausalitet, och vi måste se handlingar som agentens *göranden* inte som händelser beskrivna i termer av deras mentala orsaker. En korrekt beskrivning av ett förlopp som att Sanaz gick till festen förutsätter både ett begrepp om handling, där handlingen som en särskild form av händelse redan är förutsatt, och ett orsaksbegrepp, där agentkausalitet som en särskild form av orsakssamband är förutsatt. Ett differentierat begrepp är karaktäriserat av att i den abstrakta bestämningen ("händelse", "orsak") är begreppets mer specifika underkategorier redan förutsatta. Om ett begrepp är differen-

¹⁵Davidson 2001a, s. 78–79.

¹⁶Davidson 2001a, s. 79.

¹⁷Ibid.

tierat, så är inte den abstrakta bestämningen den mest grundläggande, utan det är de mer specifika underkategorierna som ger abstraktionen dess innehåll.¹⁸ Steward säger,

För även om jag antar, för diskussionens skull, att handlingar *är* händelser, så är de händelser av ett väldigt speciellt slag. De är händelser som är orsakanden [*causings*] av kroppsrörelser och förändringar *av agenten*. Handlingskausalitet är därför just kausalitet genom orsakanden av agenter. [...] Idén om ett orsakande av en agent finns kvar i hjärtat av handlingsbegreppet. (Steward 2012, s. 205)

(Vi ska återkomma till varför Steward använder sig av den lite otympliga formuleringen ”kausalitet genom orsakanden av agenter”. Den operativa termen här är ”orsakanden”, vilken ska signalera att handlingar är orsakande processer.)

Även Hornsby skulle hålla med om att ett odifferentierat händelsebegrepp i samklang med principen om kausalitetens nomologiska karaktär inte kommer vara användbart för att förstå handlingar. Hornsby skriver: ”Enligt Anscombes uppfattning kan inte en generell princip för individuering av händelser av vilket slag som helst tjäna som individueringssprincip för handlingar.”¹⁹ Och att behandla handlingar i enlighet med sina egna individuationsprinciper är också, enligt Hornsby, att förutsätta att personer kan utföra ”kausalt arbete”. Hon säger: ”Poängen här är endast att det finns exempel där vi är bundna att tänka att i den mån en person har burit något, eller klämt något, eller knuffat till någonting, så är det särskilda kausala arbetet den personens arbete.”²⁰

Stewards sätt att använda termen ”agenkausalitet” – Hornsby använder inte termen, men hon delar tanken – skiljer sig från en vanlig föreställning om agentkausalitet. Begreppet agentkausalitet förknippas ofta med uppfattningen att agenter orsakar sina handlingar. Men att agenter skulle orsaka sina handlingar är, enligt dessa författare, en ohållbar uppfattning. Problemet blottas när vi ställer frågan *hur* agenter i så fall skulle orsaka sina handlingar. Svarar vi att de orsakar sina handlingar *genom att*

¹⁸ Anton Ford (2011) argumenterar för tanken att det finns olika sätt på vilket en generell bestämning kan förhålla sig till sina mer specifika underkategorier. Ett sådant sätt är vad han kallar för kategorisk generalitet (*categorical generality*). Kategorisk generalitet är sådan att den specifika bestämningen är mer fundamental än den generella. Ford hävdar att handlingar som art av den mer generella kategorin händelser är ett exempel på kategorisk generalitet. Fords uppsats skapar således ett ramverk för att uttrycka tanken att det händelsebegrepp vi måste förutsätta för att förstå handlingar är differentierat, även om han inte uttrycker saken i dessa termer.

¹⁹ Hornsby 2011, s. 117.

²⁰ Hornsby 2011, s. 109.

handla, så har vi ett cirkelresonemang eftersom det vi ville veta var just *hur* agenten orsakar sina handlingar. Men det tycks inte finnas något vetligt alternativ till cirkelresonemanget. Vi kan inte betrakta agenten som ett mystiskt första led i en orsakskedja där händelser följer på händelser, som om agenten hade en förmåga att *utan att göra något* knuffa igång en orsakskedja. För då ter sig tanken på agentkausalitet som ytterligt paradoxal, som en knuff som samtidigt inte är någon knuff.²¹

Detta är varför det är viktigt att använda formuleringen ”kausalitet genom orsakanden av agenten” – som Steward gjorde i citatet ovan. En handling är en process som utförs av agenten. Vår förståelse av en sådan process, t.ex., välta, dra, och förflytta sig, är det mest grundläggande i vår förståelse av den relevanta typen av kausalitet. ”Sanaz gick till festen” är alltså en formulering som bäst uttrycker den form av orsaksförhållande vi här har att göra med, enligt dessa författare.

Vi kan således inte, enligt Hornsby och Steward, förstå agentkausalitet genom att förstå begreppet ”orsak” frånkopplat från vår förståelse av begreppet ”handling”, en förståelse som vi sedan försöker tillämpa för att uttrycka vad en handling är. För att undvika detta misstag vill de att handlingar och agentkausalitet (och således också det relevanta begreppet ”orsak”) ska *förstås tillsammans*. Hornsby skriver: ”Det generiska begreppet om orsakande [*causing*] [...] är något vi snappar upp när vi samlar samman [de kausativa] verben: vi förstår det genom att känna igen den kausativa karaktär som förenar dem. Vi kan inte anta att vi kan förstå det genom att veta vad det innebär för en händelse att orsaka en annan händelse.”²² Enligt min terminologi så menar Hornsby här att ”orsak” är ett differentierat begrepp: de kausativa verben och de kausala processer de beskriver är mer grundläggande än ett abstrakt begrepp om orsak och verkan. Här framstår olikheten med Davidsons uppfattning tydligt: den kausalitet som handlingar exemplifierar kan enligt honom förstås endast i termer av de abstrakta orsaks- och händelsebegrepp som ges i termer av varandra och inte förutsätter något om naturen hos sina underkategorier. I skarp kontrast anser då Hornsby och Steward att det relevanta orsaksbegreppet endast framträder i det att vi förstår vad en handling är.

Nu, en omedelbar fråga till Hornsby och Steward är hur vi ska förstå detta att handlingar är orsakande processer. Är det inte något mystiskt, ett slags magiskt brott i den annars helt lagbundna materiella verkligheten? Hornsby, via Anscombe, menar tvärtom: detta slags kausalitet är – kontra Humes antagande att vi inte kan observera något annat än på varandra följande händelser – tydlig för blotta ögat. Hornsby skriver:

²¹Davidson själv kommer i sitt försök att analysera begreppet ”agentskap” själv fram till denna paradox (2001a, s. 52).

²²Hornsby 2011, s. 107.

”Genom att se denna händelse [någon som bär en resväska], så ser du verksam kausalitet, och det är för att du ser henne arbeta – henne orsaka det ena och det andra.”²³

Poängen är inte heller att det nomologiska orsaksbegreppet inte har *något* tillämpningsområde eller giltighet. Snarare vill dessa författare (i synnerhet Steward) skapa utrymme för *fler än en form av kausalitet*. Steward skriver,

Kausalitet betraktas bäst som en kategori: ett stort och ontologiskt flexibelt paraplybegrepp under vilket vi tar in en vid mångfald av ontologiskt spridda relationer och relations samband, förenade endast av sina kopplingar till vårt intresse för förklaring, förutsägelse och kontroll av fenomen. [...] Kan inte den enkla sanningen vara att vi behöver oreducerbart skilda ontologiska kategorier för att göra rättvisa åt totaliteten av kausala fenomen? (Steward 2012, s. 210)

Nu kan man undra om inte Stewards och Hornsby's tanke om agentkausalitet ger upphov till något som Davidson skulle se som ett olyckligt avsteg från den monism han ansåg vara den enda rimliga metafysiska utgångspunkten. Svaret är nja. Steward och Hornsby är inte monister i den bemärkelsen att de anser vi kan göra reda för alla orsakssamband genom att endast förutsätta ett odifferentierat händelsebegrepp och Principen om Kausalitetens Nomologiska Karaktär. Vi behöver handlingar som en särskild form av händelse med sina egna individuationsprinciper, och fler former av kausalitet än den nomologiska händelsebaserade. Men det innebär inte att de därmed är *dualister*. Alternativet till Davidsons form av monism är inte att postulera en immateriell men kausalt verksam mental substans. Utgångspunkten för Steward är att hitta ett begrepp om agentkausalitet som inte kräver något sådant. Hon säger: ”Jag ska försöka berätta en historia om hur vi kan tänka om agentskapets kausala metafysik som är naturalistiskt acceptabel och inte innehåller något spöklikt, magiskt, dualistiskt eller på något annat vis metafysiskt problematiskt.”²⁴

Denna formulering skulle också passa som en beskrivning av Hornsby's utgångspunkt, en utgångspunkt hon har artikulerat i en tidigare samling uppsatser och benämnt som en *naiv naturalism*. (Se introduktionen till Hornsby 1997.) Substansdualism enligt den Descartes som läroböckerna presenterar är således inte det som Hornsby och Steward förespråkar. Alternativet är snarare att se att inom ramen för det naturliga och det

²³Hornsby 2011, s. 108. Tanken att vi kan se kausala processer med blotta ögat hämtar Hornsby från Anscombes uppsats ”Causality and Determination” (Anscombe 1981). Anscombe menar att detta gäller inte bara för agents handlande, utan även icke-animaliska processer.

²⁴Steward 2012, s. 229.

materiella ryms djur och människor som genom att handla ger upphov till förändring, en förändring som inte var på förhand given.

Nu är ”naturalism” ett av de mest hala och flertydiga begreppen i den samtida filosofin, och det skulle krävas en noggrann begreppsutredning för att få en mer fullödlig förståelse av vad Steward och Hornsby på respektive sätt lägger in i detta ord förutom ett förnekande av substansdualism. (Här skulle nog också skillnader mellan Hornsby och Steward uppenbara sig.)

En huvudtes i Stewards bok är emellertid att agentkausalitet genomryrar hela djurriket, människor är inte de enda handlande varelserna. Så för henne är ett adekvat handlingsbegrepp helt centralt för att förstå *animalitet*. På detta sätt försöker Steward omformulera frågan om agentkausalitetens berättigande. Om vi inte anser att agentkausalitet är ett berättigat begrepp, så förlorar vi mycket mer än vi kanske från början tänkte oss. Vi förlorar inte bara mänskligt handlande, vi tappar också vårt grepp om stora delar av av djurriket, enligt Steward. (Steward har också en intressant diskussion om gränserna för agentkausalitet.)

Så enligt Steward och Hornsby hamnade Sanaz på festen inte på grund av att hennes mentala tillstånd orsakade hennes ben att röra sig i riktning mot festen, utan för att hon utförde en serie handlingar i enlighet med sina syften; kanske *begrundade* hon sin lust att dansa, *utvärderade* sina möjligheter, *bestämde* sig, och *gick* till festen.

* * *

Jag har här utelämnat många aspekter av både Steward och Hornsby tankgångar. Jag har inte överhuvudtaget inte – annat än indirekt – kommit in på det som faktiskt är Stewards huvudtema för boken, nämligen frihet och determinism. Detta för att mitt huvudsyfte har varit att visa hur dessa författare utmanar en grundläggande tes i Davidsons handlingsteori. Som avslutning tänkte jag ställa en fråga som (framförallt) Stewards framställning ger upphov till. Hur ska skillnaderna mellan djurs och människors handlande förstås, mot bakgrund av antagandet att både djur och människor är agenter? Många har menat att människor är agenter i en väsentligen annorlunda mening än andra djur. Ibland uttrycks detta som tanken att människor är det djur som *bestämmer sina egna syften*, medan djurs syften är givna av deras natur. Om jag skulle tilllåta mig att spekulera skulle jag gissa att Steward, i förlängningen av det resonemang hon driver i denna bok, skulle kunna hävda att människor visserligen kan bestämma sina egna ändamål, men att detta faktum är givet av, just, människans natur.²⁵

²⁵Tack till Martin Gustafsson för synpunkter på en tidigare version av denna text.

LITTERATUR

- Anscombe, G. E. M. 1981. *The Collected Philosophical Papers of G.E.M. Anscombe*, 2, *Metaphysics and the Philosophy of Mind*. Oxford: Basil Blackwell.
- Davidson, Donald. 2001a. *Essays on Actions and Events*. 2 uppl. Oxford: Oxford University Press.
- Davidson, Donald. 2001b. *Inquiries into Truth and Interpretation*. 2 uppl. Oxford: Oxford University Press.
- Davidson, Donald. 2004. *Språk, tanke och handling*. Översättning av Jim Jakobsson. Stockholm: Thales.
- Ford, Anton. 2011. "Action and Generality". I *Essays on Anscombe's Intention*, red. A. Ford, J. Hornsby och F. Stoutland, s. 76–104. Cambridge, Mass: Harvard University Press.
- Glüer, Kathrin. 2011. *Donald Davidson: A Short Introduction*. Oxford: Oxford University Press.
- Hornsby, Jennifer. 1997. *Simple Mindedness*. Cambridge, Mass: Harvard University Press.
- Hornsby, Jennifer. 2011. "Actions in Their Circumstances". I *Essays on Anscombe's Intention*, red. A. Ford, J. Hornsby och F. Stoutland, s. 105–127. Cambridge, Mass: Harvard University Press.
- McLaughlin, B. P. 1993. "On Davidson's Response to the Charge of Epiphenomenalism". I *Mental Causation*, red. J. Heil och A. Mele, Oxford: Clarendon Press.
- Steward, Helen. 2012. *A Metaphysics for Freedom*. Oxford: Oxford University Press.
- Stoutland, Frederick. 1999. "Intentionalists and Davidson on Rational Explanation". I *Actions, Norms, Values: Discussions with Georg Henrik von Wright*, red. G. Meggle and A. Wojcik. Berlin, New York: Walter de Gruyter.
- Velleman, J. D. 1992. "What Happens When Someone Acts?". *Mind* 101, nr 403, s. 461–81.

Patrik Ribe

Den liberala autonomins förutsättningar

Moral är alltså snarare en fråga om känsla än om omdömen, även om denna känsla eller sinnesrörelse vanligen är så lugn och stilla att vi ofta förväxlar den med en idé, i enlighet med vårt vanliga sätt att uppfatta alla ting som starkt liknar varandra som ett och samma.

DAVID HUME

1. INLEDNING

Liberalismen, med sitt ursprung i upplysningen, har kännetecknats av att den eftersträvar en beskrivning av den mänskliga naturen som harmonierar väl med vetenskapens rön. Perspektivet på människan som samhällsvarelse har alltid varit relativt optimistiskt inom liberal samhällsteori. Förnuftet förväntas ge oss en enastående potential. Genom ökat välstånd och utbildning anses människan kunna ”träda ur sin barndom” och på allvar göra bruk av sitt förnuft i syfte att bli herre över sig själv. Modern liberalism utmärks av sin strävan mot en sådan positiv frihet. Autonomi, eller självherravälde, är ett centralt värde inom liberal teori.

Men är ett sådant autonomiideal möjligt att förverkliga? Under senare år har moralpsykologisk forskning tagit en ny och emotionistisk vändning. Den nya forskningen ger starka skäl att betvivla förnuftets betydelse för moraliska omdömen.

En uppmärksammas forskare inom den nya moralpsykologin är Jonathan Haidt. Haidts psykologi är präglad av neurovetenskapliga och evolutionära perspektiv. I förgrunden står våra emotioner, inte förnuftet. I vissa delar kan man hävda att Haidt till stor del bekräftar David Humes emotionistiska moralteori (*sentimentalismen*). Haidts moralpsykologi brukar betecknas som *social intuitionism* (Haidt 2012, s. 48ff).

Inom utvecklingspsykologin har den kognitiva skolan, med prominenta företrädare som Jean Piaget och Lawrence Kohlberg givit stöd för liberalismens antaganden om människan som förnuftsväsen. Moraliska omdömen är enligt Kohlberg produkter av förnuftet. Genom en intentionell, kontextoberoende process konstruerar vi våra värdeomdömen. (Kohlberg, Levine och Hower 1983, s. 69). Kohlbergs och liberalismens rationalistiska bild av människans natur utmanas av Haidts intuitionism.

Syftet med föreliggande artikel är att närmare förstå på vilket sätt den sociala intuitionismen utmanar förverkligandet av liberal autonomi. Autonomi är ett ideal vars praktiska förverkligande är avhängigt människans psykologiska konstitution. Artikeln syftar inte till att övertyga läsaren om att Haidts teori är korrekt.

2. MÄNNISKAN – ETT RATIONALISERANDE DJUR

Haidt har utvecklat en metafor som åskådliggör förutsättningarna för rationalitet. Den tar sin utgångspunkt i hjärnans evolution. Kognitiva processer kan, enligt Haidt, beskrivas med hjälp av att vi föreställer oss en ryttare på en elefant. Elefanten ska representera de äldre delarna av vår hjärna, som vi i hög grad delar med många andra däggdjur. Det limbiska systemet utgör en väsentlig del av hjärnans äldre infrastruktur. Systemet i fråga är med och producerar våra emotioner och spelar en avgörande roll för hur vi intuitivt reagerar på yttre stimuli (Haidt 2012, s. 45).

Vidare står ryttaren som symbol för de nyare delarna av hjärnan, dvs. i synnerhet neocortex; de delar som skiljer oss en del från andra däggdjur. Ryttaren representerar förnuftet. Men för att kunna fatta rationella beslut krävs ett konstruktivt samspel mellan elefant och ryttare, dvs. mellan känsla och förnuft. Utvecklandet av neocortex möjliggjorde en slags ryttare, men det gjorde också att elefanten blev ”smartare” (Haidt 2006, s. 13).

Inte sällan spelar elefanten en mer central roll än vad ryttaren förstår. Ryttaren har en stark tendens att rättfärdiga beslut och handlingar som egentligen inte är produkter av förnuftet i snäv mening. Evolutionärt sett har ryttaren utvecklats för att ”tjäna” elefanten. Aristoteles är känd för att beskriva människan som ett ”rationellt djur”, men människan är enligt Haidt framför allt ”ett rationaliserande djur”, som i efterhand försöker rättfärdiga sitt handlande, sina tankar och sina beslut. Elefanten, med sin betydligt längre evolutionära historia, dominerar ofta sin ryttare på ett sätt som ryttaren inte inser (Haidt 2012, s. 44ff).

Hjärnan utför både automatiska och kontrollerade processer. Elefanten står för de automatiska och ryttaren för de kontrollerade. Flera kognitiva processer pågår parallellt. Vissa av dessa är kontrollerade och medvetna. Kontrollerade processer är språkberoende. Andra processer är automatiserade och okontrollerade. Vi är inte medvetna om dessa. Vårt förnuft försöker i efterhand rättfärdiga, bringa klarhet och ibland ”ta åt sig äran” (Haidt 2012, s. 46). Ryttaren blir en slags försvarsadvokat eller pressekreterare åt elefanten.

Haidt skiljer mellan att fälla omdömen och rättfärdiga omdömen. Han menar att moralomdömen baseras på intuition. Moraliska intuitioner kan sägas utgöra en form av kognition, men de är inte förnuftsbaserade

och kontrollerade. Intuitioner är automatiska och kommer först. Rytaren har ingen kontroll. Intuitioner ”orsakar” värdeomdömena. Moralska resonemang är en slags *post hoc*-konstruktion, dvs. resonemangen är ett sätt att i efterhand rättfärdiga våra intuitioner. Haidt skriver (min översättning): ”[N]är vi ställs inför ett socialt krav på ett verbalt rättfärdigande, blir man en försvarsadvokat som försöker att bygga upp ett försvar snarare än en domare som söker efter sanning”¹ (Haidt 2001, s. 814). Resonemangen är ofta utformade på ett sätt som gör dem effektiva som medel för att påverka intuitioner (och därmed moralomdömen) hos andra individer. Ett intuitivt system illustreras i figur 1 och jämförs med ett resonerande system:

FIGUR 1²

Ett intuitivt system	Ett resonerande system
Snabbt och kräver ingen ansträngning	Långsamt och kräver ansträngning
Bearbetningen är oavsiktlig och automatisk	Bearbetningen är avsiktlig och kontrollerbar
Bearbetningen är otillgänglig för medvetendet, endast resultaten medvetandegörs	Bearbetningen är medvetet tillgänglig och synlig
Kräver inte uppmärksamhet eller koncentration	Kräver uppmärksamhet och koncentration; resurser som är begränsade
Parallell distribution av bearbetningen	Seriell bearbetning
Gemensamt för alla däggdjur	Unikt för människor över 2 års ålder
Kontextberoende	Kontextoberoende
Plattformsberoende (beror på kroppen och hjärnan som rymmer den)	Plattformsberoende (processen kan ”transporteras” till en regelföljande organism eller maskin)

Den sociala intuitionismen understryker betydelsen av den sociala kontexten, inte för att privat moralkognition inte skulle kunna påverka våra omdömen, utan för att tydliggöra hur viktigt det sociala sammanhanget är. Vi påverkas mycket av individer och grupper som har liknade intuitioner och tenderar att dela de uppfattningar som kommer från personer vi uppskattar i övrigt. Moralomdömen växer fram i en komplex social kontext. Moralen fyller en viktig social funktion och hur vi positionerar oss i olika förhållanden får konsekvenser för vår status och roll i gruppen.

¹”[W]hen faced with a social demand for a verbal justification, one becomes a lawyer trying to build a case rather than a judge searching for the truth.”

²Haidt 2001, s. 818.

Haidt beskriver ofta moralen som en ”lagsport”. Människan är en ultrasocial art³ (Haidt 2006, s. 47).

Vår förmåga att resonera har, enligt Haidt, inte utvecklats i syfte att finna sanning. Vi tror i regel på det mesta som ger stöd för de föreställningar och värderingar som finns i just vårt ”lag” (Haidt 2012, s. 71). Våra preferenser är inte heller något vi rationellt väljer. Våra intuitioner formas genom vår sociala interaktion i en viss miljö. Inte sällan har vi motstridiga preferenser och intuitioner som ”företräder” olika sidor av jaget.

Något som gör ”elefant/ryttare”-analogin begränsad är att vi egentligen har att göra med en elefant som ibland bär på motstridiga preferenser (vilket också ofta förbryllar ryttaren). Hjalmar Söderberg lyckas sammanfatta det hela väl i den berömda strofen: ”Jag *vill så mycket* som jag inte skulle vilja att jag ville.” Ryttaren ställer dessutom till det för elefanten ibland genom att ryttaren drar slutsatser (ofta kloka) om olika förväntade konsekvenser. Elefantens preferenser ställs mot varandra och det är ibland svårt att avgöra vilken preferens som är starkast.

Den moraliska mångfald vi ser när vi jämför moralen i olika kulturer och epoker är påtaglig, men variationen är, enligt Haidt, inte arbiträr. Sex fundament med olika evolutionära utvecklingslinjer ligger enligt Haidt till grund för all den moraliska mångfald som människan utvecklat. Människan kan, enligt Haidt, sägas vara utrustad med ett moraliskt ”mixerbord” med sex kanaler. Kanalerna justeras genom våra erfarenheter, men det finns ett slags ”fabriksinställning” som är genetiskt betingad (alltså en medfödd förprogrammering). Sociala och kulturella faktorer spelar naturligtvis en väsentlig roll för ”inställningarna”. Justeringar görs hela tiden, men vissa kanaler är mer dominanta än andra och varje individ har sin någorlunda unika moraliska profil (Haidt 2009, s. 1030ff).

3. SOCIAL INTUITIONISM I PRAKTIKEN

För att konkretisera vad Haidts moralpsykologi säger om förutsättningarna för politisk och moralisk reflektion och diskussion, utgår jag från ett exempel.

Ponera att vi har två väljare A och B. Som kloka demokratiska medborgare försöker de bilda sig en uppfattning om förslagen hos de olika kandidaterna i valet. Efter att ha tagit del av en TV-sänd utfrågning av de olika partiföreträdarna, ställer B följande fråga till A: *Vad anser du om förslag X?* A får först en intuition, dvs. en emotionell respons, och faller sedan ett värdeomdöme. Därefter försöker A motivera varför förslaget är positivt.

³Haidts intuitionism skiljer sig betänkligt från Kohlbergs kognitivism, som bär starka drag av ett ”resonerande system”.

Anta att förslaget handlar om att höja alkoholskatten. A får i detta fall en negativ intuition och faller värdeomdömet att en sådan skattehöjning är orättvis. Därefter säger A att skattehöjningen kommer att drabba konsumenterna. A:s värdeomdöme och resonemang föder i sin tur en intuition hos B som därefter faller ett positivt omdöme om skattehöjningen. Han säger sedan att det är bra med en höjd skatt, eftersom den skulle motverka alkoholkonsumtionen. B:s resonemang väcker en ny intuition hos A som därefter faller värdeomdömet att en minskning av alkoholkonsumtionen förvisso är bra, men att även de konsumenter som inhandlar alkohol sällan riskerar att drabbas om alkoholskatten höjs osv. Om A och B väljer att försöka hänvisa till olika fakta om alkoholskador, skatters effekter på medborgarnas beteende etc. så kommer deras intuitioner att försvagas eller förstärkas. Värdeomdömena kommer eventuellt att preciseras och modifieras i förhållande till de stimuli som den sociala situationen innefattar.

Det jag försökt illustrera är att intuitioner, enligt Haidt, spelar en avgörande roll när vi faller värdeomdömen. Intuitioner genererar värdeomdömen. Resonemangen som sedan följer är en slags *post hoc*-konstruktion, dvs. resonemangen är ett sätt att i efterhand rättfärdiga våra värdeomdömen. I ett socialt komplext sammanhang bombarderas vi med intryck som utlöser reaktioner i vårt automatiserade intuitiva system. För att återknyta till "elefant-ryttar"-metaforen kan man säga att elefanten lägger grunden för våra värdeomdömen och ryttaren försöker sedan rättfärdiga dem och ge sken av att värdeomdömena är en renodlad produkt av medveten kontrollerad kognition.

4. ÄR AUTONOMI MÖJLIG?

Inom samtida liberal politisk teori är autonomi ett centralt normativt ideal. Autonomi har definierats på otaliga sätt. I detta sammanhang eftersträvas en definition som är vid i den meningen att den inbegriper vad som är essentiellt i de flesta autonomibegrepp. Med viss hänvisning till statsvetaren Mats Lundström (och indirekt Jon Elster), hoppas jag kunna hitta en adekvat definition.

Autonomi handlar på en basal nivå om individens möjlighet att "vara sin egen herre". Individens preferenser och ideal skall vara produkter av den egna tankekraften. Den sociala kontexten ska inte direkt eller indirekt forma individens beslut och värderingar. Hennes liv skall genomgående präglas av självstyre (Berlin 1984, s. 142f).

Att vara autonom kan förstås i termer av rationalitet och förutsättningarna för olika val. Rationalitet kan definieras på olika sätt. För att åskådliggöra vad som skiljer de två mest använda rationalitetsbegreppen kan man som Lundström skilja mellan olika val:

- (1) Val av norm
- (2) Val av mål för handling
- (3) Val av handlingar

Enligt gängse definition av rationalitet, som bl.a. förekommer inom nationalekonomin, är aktören rationell när hon har förmågan att maximera förväntad nytta. En sådan förståelse av rationalitet återfinns hos Hume. Endast valet av *medel* kan väljas rationellt. *Målet* kan däremot inte vara föremål för ett rationellt val. I sammanhanget åsyftas *genuina* eller *fundamentala* mål dvs. mål som är självändamål. Att vara förnuftig i fråga om *instrumentella* mål är fullt möjligt. Det handlar inte heller om de val av mål som utgår från eller förutsätter en specifik moralisk eller politisk ”matris”. Att exempelvis se vilka mål som är rimliga med utgångspunkt i utilitarismen är fullt möjligt. Sådan rationalitet är väl förenlig med det tunna rationalitetsbegreppet och Haidts moralpsykologi. Hume ansåg att man är rationell om man har förmågan att dra slutsatser om hur man maximerar den förväntade nyttan. Vi har att göra med ”tunn” rationalitet, som enbart inbegriper val av handling (eller medel) (Lundström 1993, s. 23).

Kant lät formulera ett mer substantiellt eller ”brett” rationalitetsbegrepp. Autonomi har en intim koppling till Kants rationalitetsbegrepp. När aktören är autonom (dvs. när hennes preferenser och normer är produkter av den egna tankekraften) då kan hon också sägas ha möjlighet att vara rationell. Autonomi innefattar substantiell rationalitet. Aktören kan vara rationell ifråga om både självändamål och medel för att förverkliga dessa. Lundström skriver: ”Autonomi förutsätter inte bara frihet från tvång, utan även en ’inre’ frihet; en förmåga att välja och handla utifrån en informerad verklighetsuppfattning och genomtänkt övertygelse om vad som är rätt och fel” (ibid., s. 212).

Att många anser att autonomi är eftersträvsvärt är en sak, men är den överhuvudtaget möjlig att förverkliga? Haidt skulle inte ha invändningar mot att vi kan ha en mer eller mindre väl ”informerad verklighetsuppfattning”. Att ryttaren kan konkretisera olika handlingsalternativ och att detta påverkar elefanten är således koherent med Haidts teori. Betydelsefull är också den sociala omgivningens ”ryttare” (och ”elefanter”) för forandet av individens intuitioner. Ryttaren kan således i någon utsträckning reflektera över sina preferenser och värdeomdömen.

Vad Haidt däremot skulle sätta frågetecken för är huruvida vi i snäv och strikt mening kan ha en ”genomtänkt övertygelse om vad som är rätt och fel”. Om värdeomdömen ger uttryck för preferenser som vilar på intuitioner, som Haidt hävdar, verkar förutsättningarna för ”genomtänkta övertygelser” om rätt och fel tämligen begränsade.

Först och främst är det svårt att hävda att intuitionsbaserade värde-

omdömen är genomtänkta övertygelser (eller ens övertygelser i någon teknisk mening överhuvudtaget). Intuitioner är produkter av både biologiska och sociala omständigheter och processer. Intuitioner är inte resultatet av medveten kontrollerad eller logisk reflektion (Haidt 2001, s. 818). I vilken utsträckning vi rationellt kan påverka och välja våra värdeomdömen är, som jag ser det, avgörande för om autonomi i praktiken är möjlig. Autonomi förutsätter att ryttaren kan övertyga elefanten med förnuftsargument, inte bara påverka i någon vag psykologisk mening.

Om elefanten är ansvarig för våra intuitionsbaserade värdeomdömen, kan ryttaren endast välja ”medel”, inte ”mål”. Moraliska omdömen är till syvende og sist inte en produkt av vårt förnuft, även om förnuftet finns med i processen. Haidt landar i konklusioner som tangerar Humes moralfilosofi:

Och eftersom resonemanget inte är källan varifrån de tvistande härleder sina lärosatser, är det förgäves att förvänta sig att någon logik, som inte talar till känslorna, någonsin kommer att engagera honom att omfamna sundare principer.⁴

Personen P kan således inte rationellt välja vad P anser vara gott. P kan inte rationellt välja sina preferenser, som kommer till uttryck i P:s intuitioner. Ställd inför olika alternativ om vad som bör anses gott kommer P att lyda sin intuition i den sociala kontexten.

När P står inför olika politiska alternativ i ett demokratiskt val, kommer P:s intuitioner att spela en väsentlig roll. P kan inte välja sina genuina preferenser rationellt. P kan däremot anpassa förverkligandet av dem till vissa realpolitiska omständigheter och reflektera över dem instrumentellt. Genom att ha en mer fullständig bild av de politiska alternativen kan P kalkylera instrumentellt och få en mer upplyst bild av de realpolitiska förutsättningarna. Någon kunskap om vad som i genuin mening är bättre eller sämre i en ickeinstrumentell mening är emellertid ej möjlig. P:s intuitioner kan inte vara mer eller mindre välgrundade i någon substantiell mening. Deliberation i fråga om genuina mål kan ej uppnås.⁵

Kanske finns det en invändning mot Haidt. Kan man inte tänka sig

⁴Min översättning av: ”And as reasoning is not the source whence either disputant derives his tenets, it is in vain to expect that any logic, which speaks not to the affections, will ever engage him to embrace sounder principles.” (Hume 1948, s.175.)

⁵Haidts antaganden är alltså fullkomligt kompatibla med att vi utifrån vissa normativa principer eller system kan dra rationella slutsatser om vilka handlingar som är önskvärda inom de givna normativa ramarna. Att dra slutsatser om vad exempelvis utilitarismen skulle föreskriva i en viss situation är således fullt möjligt. Att däremot göra genuina (inte instrumentella) rationella val av normativa principer eller system är inte möjligt.

att personen P vid tidpunkten T får positiva intuitioner inför ett normativt system. Därefter försöker P att orientera sig moraliskt och politiskt genom att tillämpa systemet så konsekvent som möjligt. P utgår således från det normativa systemet, inte sina intuitioner. Invändningen är inte särskilt framgångsrik. Valet av system bygger trots allt på intuition, inte deliberation. Haidt skulle sannolikt hävda att P kommer att fortsätta tillämpa systemet så länge de står i samklang med intuitionerna. Om P en dag får mycket negativa intuitioner vid tillämpningen av systemet på ett enskilt område eller en specifik fråga, så skulle P sannolikt ge upp eller modifiera sitt system. Intuitionerna skulle således fortsätta att forma P:s ”ställningstaganden”.

Men kan man inte tänka sig att P fortsätter att klamra sig fast vid sitt system? Det kan man nog tänka sig, men gissningsvis skulle Haidt hävda att P har en motstridig intuition som får P att klamra sig fast. Som individer är vi komplexa och våra intuitioner är ofta motstridiga. Om våra moraliska intuitioner dessutom utgår från sex olika källor (eller emotiva system) kommer vår förmåga att resonera opartiskt och obehindrat att grusas ytterliggare. Risken för motstridiga preferenser är stor. Skälet till varför P ger upp principer är som sagt att ett annat intuitivt system i hjärnan har kickat igång, inte att P genom förnuftsresonemang har nått nya moraliska slutsatser.

Avslutningsvis är det inte alls självklart vilken värdesemantisk teori som harmonierar mest väl med Haidts moralpsykologi. Ingen teori är självskriven. Men den teori som är mest plausibel är nog emotivismen. Värdeomdömen saknar ju, enligt Haidt, kognitivt innehåll (i snäv mening). De uttrycker vår attityd till något. Värdeomdömen innefattar inte någon trosföreställning och uttrycker inte kunskap. Om allt detta stämmer är det svårt att se att genuina värdesatser har sanningsvärde överhuvudtaget.

Förutsättningarna för liberal emancipation och deliberativt meningsutbyte har moralpsykologiskt sett dystra villkor, om Haidts teorier stämmer.

5. SLUTSATS

Den *breda* (eller substantiella) autonomi som vi finner inom liberal normativ teori ter sig svåruppnåelig, om den sociala intuitionismen är riktig. Haidt anser att vi kan tala om moralisk kognition, men det är huvudsakligen inte en medveten kontrollerad rationell process vi har att göra med. Våra värdeomdömen är en funktion av våra intuitioner. Intuitionerna vilar på åtminstone sex olika neurobiologiska moduler, som finjusteras och modifieras i kontakt med den sociala omgivningens stimuli. Autonomi i moralfrågor i någon mer vedertagen mening är inte möjlig.

Om man skulle definiera autonomi i en svagare mening, exempelvis som förmågan att kunna påverka (dock endast i vag psykologisk mening) och i någon utsträckning reflektera över sina preferenser uppstår ingen konflikt med Haidts psykologi. Det är först när autonomi beskrivs som att kunna ”ha genomtänkta övertygelser” i moralfrågor, som bekymmer uppstår. Intuitioner är som sagt inte genomtänkta övertygelser. Förutsett att den sociala intuitionismen är riktig har moderna liberaler skäl att ifrågasätta autonomi som ett realistiskt ideal.

REFERENSER

- Berlin, Isaiah. 1984. *Fyra essäer om frihet*. Övers. Torkel Franzén. Stockholm: Ratio.
- Graham, J., J. Haidt, B. A. Nosek. 2009. ”Liberals and Conservatives Rely on Different Sets of Moral Foundations”. *Journal of Personality and Social Psychology* 96, no. 5, 1029–46.
- Haidt, Jonathan. 2001. ”The Emotional dog and Its Rational Tail: A Social Intuitionist Approach to Moral Judgement”. *Psychological Review* 108, no. 4, 814–34.
- Haidt, Jonathan. 2006. *The Happiness Hypothesis: Finding Modern Truth in Ancient Wisdom*. New York: Basic Books.
- Haidt, Jonathan. 2007. ”Moral Psychology and the Misunderstanding of Religion”, 22/9, http://www.edge.org/3rd_culture/haidt07/haidt07_index.html (2012-04-24)
- Haidt, Jonathan. 2012. *The Righteous Mind: Why Good People are Divided by Politics and Religion*. New York: Pantheon Books.
- Hume, David. 1948. *Moral and Political Philosophy*. New York: Hafner Press.
- Hume, David. 2004. *En avhandling om den mänskliga naturen*, Bok 3, *Om moralen*. Övers. Jim Jakobsson. Stockholm: Thales.
- Kohlberg, L., C. Levine, A. Hower. 1983. *Moral Stages: A Current Formulation and a Response to Critics*. Basel: Karger.
- Lundström, Mats. 1993. *Politikens moraliska rum: En studie av F. A. Hayeks politiska filosofi*. Stockholm: Almqvist & Wicksell.

Till vem talar filosofin? Till vem riktar den sig? Ett spontant svar vore kanske: till alla och envar? Eller kanske till alla förnuftiga varelser. Enligt Aristoteles berömda definition är det just vad vi människor är: förnuftiga djur. För Aristoteles tycks det inte heller ha varit något problem vem filosofin talar till. Den tycks tala till vem som helst som är förmögen att förstå språket och som kan följa ett abstrakt resonemang. Men redan hos Platon är det inte så enkelt. Där kan Sokrates mitt i dialogernas avgörande stund avbryta sig med en reflektion att här närmar vi oss något som inte kan sägas rakt ut, varpå han i stället tillgriper en bild eller allegori. Tänkandet uppträder med ens som chiffer, men därmed också som ett annat slags tilltal, som söker sin speciella läsare. Man kunde hävda att det är något som sker när filosofin närmar sig litteraturen och dess uttrycksmedel. Men det vore en alltför vag definition. För det utmärkande för detta slags tänkande är inte att det arbetar med bilder och metaforer, utan just att det gör sin egen läsbarhet till en fråga.

Om det finns ett urdokument till detta slags skrivande så måste det ändå vara den första mening som skall ha inlett den fragmentsamling som skrevs av Herakleitos 500 f. Kr. Han skriver där: "Inför denna utläggning som är för alltid förblir människorna oförstående, såväl innan de lyssnat som när de väl lyssnat." Det är en oerhörd utsaga. Som sitt första påstående upplyser den läsaren om sin egen obegriplighet. Den berättar att den kommer att förbli oförstådd också efter att läsaren läst den. Vad skall vi kalla ett sådant påstående? Vad skall vi kalla ett filosofiskt tänkande som inleder sig självt på detta vis? Den utpekar sig själv som platsen för en gåta och ett chiffer. Det vore dock fel att tala om det som bara en gåtfull eller "dunkel" filosofi, vilket är det epitet som ända sedan antiken använts om Herakleitos. För det väsentliga är inte att den är svår eller dunkel, utan att den just talar om sig själv och sin begränsade förståelighet. Men samtidigt som detta slags påstående på ytan kan tyckas stöta sin läsare ifrån sig, så innebär det naturligtvis på ett annat plan ett alldeles egenartat tonfall. Det inbjuder ju till en pakt, som inte är för vem som helst.

Vår filosofiska modernitet präglas bland annat av att ha återupptäckt detta chifferade och självreflexiva tilltal. Vi möter det hos Kierkegaard, liksom hos Nietzsche. Hos dem båda finner vi en filosofisk tanke som inte bara uttrycks med en sällsynt litterär stil, utan också en tanke som sysselsätter sig med sin egen begränsade begriplighet. Hos dem båda inbjuds på så vis läsaren att träda in i ett alldeles eget förbund. Den säger till sin läsare: du kanske visar dig vara en av de få som kan förstå mig. Hos Nietzsche finns också en märklig profetia som berör just den gåtfulla skriften. Han skriver i *Bortom gott och ont* att han nu anar framväxten av en ny sorts filosofer, som vill förbli gåtor. Kanske tänker han också på en viss sorts tänkare som liksom han själv inte riktigt känner sig hemma i sin tid; som riktar sitt tal förbi och bortom sin egen tid, mot något kommande, eller kanske något förflutet, men hursomhelst inte något just här. Det säger således också till sin läsare: kom med mig, du som inte känner dig hemma på denna plats, så upprättar vi en förståelse någon annanstans. Vad kunde vara mer lockande.

Är det inte också i denna krets som Wittgenstein hör hemma? Det vill säga, i kretsen av de kringströdda själar som talar ur en hemlöshet i sin tid? Som också talar ur en förlust av språk och av läsare, men som just därigenom når den läsare som vill gå denna gåta till mötes. Den första raden i förordet till hans epokgörande ungdomsverk *Tractatus logico-philosophicus* från 1922 lyder: ”Denna bok kommer kanske att förstås bara av den, som redan tänkt de tankar den uttrycker.” Som ett nästan spöklikt, och sannolikt helt omedvetet, eko av öppningsraden i Herakleitos fragmentsamling inleds den med en reflektion över sin egen läsbarhet och oläsbarhet. Läsaren kommer inte att förstå den efter att ha läst den, om han inte redan har förstått den innan han läst den. Men med detta påstående tilltalar den också läsaren: är du en av de mina, är du som jag – för om du inte är det, så kommer du heller inte att kunna förstå vem jag är, eller varifrån jag talar.

Men vem är denne läsare? I ungdomligt trots förklarar författaren att han inte har några fotnoter eller referenser, eftersom han inte bryr sig om vad någon annan tänkt, eller om de tänkt det samma. De enda två personer som nämns som inspirationskällor är Frege och Russell. När Frege fick boken svarade han i ett älskvärt brev att han inte förstod ett dyft. Och när den alltid så generöse Russell senare ville bistå boken med ett förord blev Wittgenstein så förtvivlad över hans oförståelse att han hotade med att dra tillbaka boken. Ingen av dessa två föregångare och förebilder visade sig vara den läsare han sökte. Inte heller bland Wienkretsen medlemmar som gärna ville värva honom till sin grupp kände han sig förstådd. I stället vände han sig bort i en blandning av högmod och förtvivlan.

I samma förord till *Tractatus* finner vi också den berömda raden: ”vad som alls låter sig sägas, kan sägas klart: och vad man inte kan tala om, därom måste man tiga.” Med dessa ord avslutar han också boken hundra sidor senare. Runtom kring den komprimerade, logiskt-matematiska aforistiska texten öppnar sig en tystnad som omgärdar dess tal. Genom att peka ut denna tystnad har författaren samtidigt också försäkrat sig om sitt tilltal. Han har försäkrat sig om en läsare, som just i denna gåta anar ett återvändande och ett igenkännande. Den har skapat sitt tilltal genom att dra en gräns kring dess fattbarhet. Han säger: jag talar om det som låter sig talas om, men egentligen är detta ju inte så viktigt, eftersom det är vad som inte låter sig omtalas som i grund och botten spelar någon roll. Och därmed talar jag också, om än i en annan mening, också om detta. Men bara till den som själv vet att lyssna till det outtalbara.

Om denna hemlöshet i det omedelbara och det närvarande vittnar också öppningsorden i ett aldrig publicerat förord från 1930, som finns med i boken *Särskilda anmärkningar*: ”Denna bok är skriven för dem som är sympatiskt inställda till den anda i vilken den är skriven.” Här säger han tydligt att han inte skriver för den moderna vetenskapliga kulturen, ja knappt ens för samtiden, utan för ”vänner kringströdda i olika vrår av världen”. Vilka är dessa vänner? Vi skall nog inte tänka oss att han syftar på den ene eller andre. För en vän i denna bemärkelse behöver ännu inte ha gett sig till känna. En sådan vän behöver ännu inte ens finnas till. Den skrivande inbjuder med sitt tal till en vänskap genom att säga till den okände läsaren: du kan visa dig vara den som är min vän. Du kan visa dig vara den som hör mitt tal, och då är du också en av de mina.

Det är viktigt att se, uppmärksamma och tänka över detta egenartade tilltal, om man skall förstå vissa drag i Wittgensteinlitteraturen och receptionen. Det är nämligen slående i vilken omfattning diskussionen kring hans verk handlar om hur han ”missförstås”. Wittgenstein är en sådan författare som man själv förstår, men som de andra inte förstår. Wittgenstein är på något märkligt sätt för var och en alltid just bara min författare. Varje försök att lägga hans rastlösa, irrande tänkande tillrätta i en prydlig struktur – och sådana försök har vi sett många av – tycks ofrånkomligen förgripa sig på den tystnad som han öppnar för sin läsare som dörren till ett hemligt förbund. Det vore fel att kalla det esoteriskt, för det finns ingen hemlig lärdom som vårdas i mindre kretsar. Vad som vårdas är i stället det personligt-opersonliga tilltal som skapas av en text som just i avgörandets stund böjer av för att i stället vittna om sin egen inre omöjlighet.

För hemlängtande läsare kan denna hemlighet bli till en överväldigande lockelse. Texten kommer helt enkelt för nära, identifikationen blir för stark och det egna ljuset utsläcks i dess kalla låga. Men genom att

reflektera över detta tilltal kan vi också se hur det som griper oss också låter oss gå förbi. Ja i en djupare mening kanske det faktiskt uppmanar oss att gå förbi, och att återvända inte till textens egna – utan till vårt egna. Måhända är detta till sist dess lärdom.¹

¹Texten har tidigare framförts som föredrag i samband med en temakväll om Wittgenstein på Kulturföreningen Forum.

I *Filosofisk tidskrift* 2012/3 diskuterade jag det s.k. *tidsproblemet* rörande dödens dålighet. Jag argumenterade för att vi, i ljuset detta problem, bör överge uppfattningen att döden är något dåligt för den som dör. Tidsproblematiken fick fortsatt medieutrymme i form av uppsatsen ”Döden och tiden”, *Filosofisk tidskrift* 2013/1, där Lars Bergström framför ett antal tankeväckande synpunkter som till stor del knyter an till min diskussion. Även om Bergström inte verkar ha några särskilda betänkligheter kring min huvudtes så har hans anmärkningar potential att sätta andra delar av min argumentation i gungning. Min avsikt här är därför att bemöta ett urval av dessa.

För att göra reda för dödens dålighet tillämpas inte sällan någon form av *deprivationsteori*. Enligt denna teori är döden dålig för en person om och endast om personen hade haft det bättre om hon inte hade dött då hon dog. Genom att skilja mellan de olika sakförhållanden som en persons död involverar kan bedömningen av dåligheten hos personens död se olika ut. Sakförhållandet *att jag dör* ung är rimligtvis mycket dåligt för mig eftersom jag hade fått leva ett väsentligt längre liv om jag inte hade dött ung. Däremot verkar sakförhållandet *att jag dör kl. 13.00 den 1 januari 2025* vara varken bra eller dåligt för mig; om jag inte hade dött just då så hade jag antagligen strukit med strax tidigare eller senare. Bergström noterar att ”en lite underlig konsekvens” av deprivationsteorin är att det alltså troligtvis aldrig är dåligt att dö vid en viss tidpunkt.¹ Det må vara sant, men underligheten i fråga tycks vara av *prima facie*-slag, den sorts harmlösa underlighet som lättar i takt med att förståelsen för teorin ökar.

Deprivationsteorin ligger till grund för undersökningens centrala fråga: vilka tidpunkter t är sådana att min välfärdsnivå är lägre vid t än vad den hade varit vid t , om jag inte hade dött då jag dog? Till skillnad från vad Bergström antyder är detta inte tänkt som en ”*hedonistisk* specifikation av tidsproblemet”.² Min egen utgångspunkt är visserligen hedonistisk, men som jag ser det anmäler sig samma fråga oavsett om

¹Bergström 2013, s. 25.

²Bergström 2013., s. 26, min kursivering.

man som hedonist, preferentialist eller pluralist är ute efter att hitta en lösning på tidsproblemet.

Lösningförslaget *när man dör* är ett som jag avvisar ganska hastigt. Alltför hastigt, menar Bergström: ”om döden är en händelse, så inträffar den vid en viss tidpunkt, och då låter det väl rätt rimligt att säga att den är dålig för den som dör vid just denna tidpunkt”.³ Som en princip om *intrinsikala* dåligheter låter Bergströms förslag mycket tilltalande. Det vore konstigt att påstå att t.ex. en smärtupplevelse skulle vara intrinsikalt dålig för en person även efter att smärtan har lagt sig. Men för *instrumentellt* dåliga händelser, såsom döden, gäller knappast att de i regel är dåliga när de inträffar. Tvärtom tycks det ofta röra sig om ett slags värdemässig inkubationstid, att dåligheten bryter ut först senare. När jag skriver att detta är ett ”i bästa fall ointressant” lösningförslag på tidsproblemet menar jag att den (korrekt och trivialt) besvarar en för sammanhanget ointressant fråga, nämligen: vid vilken tidpunkt inträffar den dåliga händelse som döden utgör?⁴

En annan lösning jag granskade är *subsekventismen*, vilken gör gällande att döden är dålig för personen som dör *efter* döden. Problemet med subsekventismen är att den tycks förutsätta att även döda personer – personer som inte längre existerar – har en välfärdsnivå, om än permanent nollställd. Det låter tvivelaktigt. Att t.ex. Sokrates just nu har en välfärdsnivå är en idé som varken faller mig eller Bergström i smaken. Det vore dock inte fullt så motbjudande, menar Bergström, att jämföra två världar vid en viss tidpunkt då personen är död i den ena *men inte den andra* och säga att personens ”välfärdsnivå är lägre i den första, om den är positiv i den andra”.⁵ Detta förslag innebär att t.ex. Olof Palme har en välfärdsnivå just nu, förutsatt att han ännu hade varit i livet om han inte hade blivit mördad. Men en sådan syn har vi goda skäl att betvivla. Det faktum att Palme just nu lever och frodas i en annan möjlig värld tycks nämligen inte ha någon betydelse för hans nuvarande välfärd i den aktuella världen. Att Palmes välfärdsnivå skulle bero på vad som försiggår i andra möjliga världar är särskilt problematiskt ur ett hedonistiskt perspektiv eftersom Palme, i den aktuella världen, helt saknar mentala tillstånd. Palme verkar med andra ord befinna sig i samma ”hedonistiska läge” som skor och andra ting som vi ogärna tillskriver en välfärdsnivå.⁶ I mina ögon förblir subsekventismen därför ohållbar.

Den *atemporalistiska* lösningen på tidsproblemet går ut på att i de fall då döden är dålig för den som dör är den dålig utan att vara det vid

³Bergström 2013, s. 23.

⁴Se Johansson 2013, s. 261.

⁵Bergström 2013, s. 27.

⁶Se Johansson 2013, s. 265.

någon tidpunkt. Här kan det vara lämpligt att påpeka att atemporalismen inte säger att det inte finns någon tidpunkt då den ena världen, där personen dör, är sämre för henne än den andra världen, där personen inte dör. Atemporalismen är alltså förenlig med Fred Feldmans påstående att den ena världen *alltid* är sämre än den andra världen.⁷ Det som utmärker atemporalismen är att den förnekar att det finns någon tidpunkt vid vilken personen hade haft det bättre om hon inte hade dött då hon dog, d.v.s. en sådan tidpunkt som tidsproblemet efterlyser. Att det inte finns någon sådan tidpunkt skulle kunna tolkas som att döden *aldrig* är dålig för den som dör. Detta tenderar dock att kasta en skugga över atemporalismen eftersom det, som Bergström skriver, ”är konstigt att säga att döden är dålig och aldrig dålig”.⁸ Jag tror dock inte att atemporalismen behöver föranleda så märkliga uttalanden, för termen ”aldrig” verkar inte bara syfta till att förneka att det finns en tidpunkt då något är fallet, utan också förneka att något är fallet. En teist som betraktar Gud som en tidlös varelse skulle nog, av just denna anledning, inte vilja påstå att Gud *aldrig* existerar. Givet denna tolkning av ”aldrig” innebär alltså inte atemporalismen att döden aldrig är dålig, utan bara att det inte finns någon tidpunkt då döden är dålig, trots att den är dålig. Atemporalisten förnekar alltså premiss (1), men godtar premiss (2) i följande epikuriska argument:

- (1) Allt som är dåligt för en person är dåligt för henne vid någon tidpunkt.
- (2) Det finns ingen tidpunkt då döden är dålig för personen som dör.
- (3) Alltså är döden inte dålig för personen som dör.

Bergström skriver att ”atemporalismen kan tolkas som premiss (2)”, men att den också kan ”tolkas som negationen av premiss (1)”.⁹ Men är inte detta lite vilseledande? Att tolka atemporalismen som premiss (2) tillåter ju att döden inte är något dåligt för personen som dör, vilket atemporalismen förnekar. Negationen av premiss (1), å sin sida, säger inget som helst om döden, utan endast att det finns åtminstone ett ting som är dåligt utan att vara det vid någon tidpunkt. Atemporalismen bör därför snarast tolkas som premiss (2) och negationen av (3).¹⁰

REFERENSER

Bergström, Lars. 2013. ”Döden och tiden”. *Filosofisk tidskrift*, nr 1, s. 23–29.

⁷Feldman 1992, s. 154.

⁸Bergström 2013, s. 28.

⁹Bergström 2013, s. 29.

¹⁰Tack till Jens Johansson för värdefulla kommentarer.

- Ekendahl, Karl. 2012. "Död i tid och otid". *Filosofisk tidskrift*, nr. 3, s. 3–15.
- Feldman, Fred. 1992. *Confrontations with the Reaper*. New York och Oxford: Oxford University Press.
- Johansson, Jens. 2013. "The Timing Problem". I *The Oxford Handbook of Philosophy of Death*, red. Ben Bradley, Fred Feldman och Jens Johansson, s. 255–73. Oxford: Oxford University Press.

Richard Swinburne

Häggsström om Swinburne och sannolikheten för Guds existens

I sin artikel ”Brister i sannolikhetsargument för och emot Guds existens” (2008) framlägger Olle Häggsström kritiska synpunkter på mina sannolikhetsargument för Guds existens och drar slutsatsen att de ”läcker som ett såll”. Min argumentation är formulerad i sin helhet, med hjälp av Bayesiansk sannolikhetskalkyl, i Swinburne 2004 och sammanfattas på ett informellt sätt i Swinburne 2007. Syftet med denna artikel är att besvara Häggsströms kritik.¹ Mitt resonemang återfinns i Swinburne 2004. Här fäster jag framför allt uppmärksamhet på avsnitt som Häggsström verkar ha försummat.

Essensen i min argumentation är följande. En hypotes troliggörs av evidens i den mån hypotesen [h] är enkel, och är sådan att om den är sann, evidensen (e) är förväntad (dvs. troligen kommer att äga rum), och om den är falsk, evidensen inte är förväntad. Detta kan ges en formell inramning genom Bayes sats, som sammanfattas av Häggsström. Vi har n konkurrerande hypoteser h_1, \dots, h_n med ”apriorisannolikheter”, det vill säga sannolikheter innan evidensen e har beaktats. Apriorisannolikheten för varje hypotes h_i är $P(h_i)$. Vi kan förvänta oss att e inträffar (det vill säga är trolig) med varierande sannolikheter, $P(e|h_i)$, för varje hypotes. Jag använder $P(e|h_i)$ för att beteckna den ”prediktiva styrkan” hos h_i ; på engelska ofta kallad för likelihoodfunktionen för h_i . Bayes sats ger ”aposteriorisannolikheten” för alla dessa hypoteser, givet e , genom:

$$P(h_i|e) = \frac{P(h_i)P(e|h_i)}{P(h_1)P(e|h_1) + \dots + P(h_n)P(e|h_n)}$$

¹Jag ber läsare av denna tidskrift om ursäkt för att mitt svar på en artikel som skrevs för fem år sedan kommer först nu, men jag fick helt nyligen vetskap om Häggsströms artikel. Jag är mycket tacksam till Ola Hössjer för att han översatte Häggsströms artikel till engelska, och sedan översatte mitt svar tillbaka till svenska. [*Red:s anmärkning*: Ola Hössjer är professor i matematisk statistik vid Stockholms universitet.]

Den evidens som är relevant för att avgöra sannolikheten för en hypotes kan delas upp i "bakgrundsevidens", som antas för att fastställa apriorisannolikheten för hypotesen ($P(h_i)$), och "omedelbar evidens" (e), som används för att bestämma $P(e|h_i)$. Uppdelningen mellan bakgrundsevidens och omedelbar evidens är godtycklig och påverkar inte den resulterande aposteriorisannolikheten ($P(h_i|e)$). När en aktuell hypotes har begränsat tillämpningsområde (t.ex. om den uttalar sig om strukturen hos en kemisk substans), ligger det närmast till hands att som bakgrundsevidens använda den logiskt mest relevanta och närliggande informationen (som exempelvis erhålls från allmän kemisk teori för hur ämnen är uppbyggda). Men när vi fokuserar på hypoteser som avser att förklara alla fenomen, eller inskränker oss till alla fysiskt betingade fenomen, är det behändigast att behandla all logisk relevant evidens som omedelbar evidens, och betrakta apriorisannolikheten för en hypotes som dess sannolikhet innan någon relevant evidens har beaktats överhuvudtaget. En sådan apriorisannolikhet för en hypotes, som föregår all relevant evidens, kallar jag för dess "inneboende sannolikhet". För varje evidens e är det möjligt att formulera ett stort antal hypoteser (som är ömsesidigt uteslutande och uttömmar sitt tillämpningsområde, det vill säga – givet e – är precis en av dem sann) som kan prediktera denna evidens med mycket varierande styrka, $P(e|h_i)$. Givet evidensen och två konkurrerande hypoteser med samma förmåga att förutsäga evidensen, så betraktar vi den ena hypotesen som mer sannolik än den andra endast med apriorikriterier, och av dessa kriterier är enkelhet (i vid mening) den överlägset viktigaste. En hypotes inneboende sannolikhet är den sannolikhet den skulle få endast med hjälp av dessa apriorikriterier. När jag formulerar apriorisannolikheter för teismen och dess konkurrerande hypoteser syftar jag på de inneboende sannolikheterna, och jag betraktar alltså all relevant evidens som omedelbar evidens. Såvida det inte finns objektiva kriterier för enkelhet, genom vilka vi tilldelar apriorisannolikheter för vetenskapliga hypoteser, är det helt enkelt inte möjligt att bedriva vetenskap (eller ha någon kunskap om världen, bortom det omedelbart observerbara).

Hägström påstår att vi använder enkelhet som sanningskriterium för att detta visat sig vara en "genom historien utomordentligt framgångsrik vetenskaplig tumregel" (2008, s. 6). Men ett resonemang som använder tidigare vetenskaphistoria som evidens för att troliggöra en teori som går ut på att enkelhet är evidens för sanningen, är cirkulärt – eftersom utan enkelhetskriteriet hade teorin inte varit trolig baserad på denna evidens (se exempelvis Swinburne 2001, s. 100–102).

Så det är ganska viktigt att veta vad enkelhet hos en hypotes innebär. Grovt uttryckt handlar det om en teori som postulerar få entiteter och

typer av entiteter, få egenskaper och typer av egenskaper, och få matematiskt enkla samband (definierade på ett visst sätt) dem emellan. Jag gav en kortfattad version av min redogörelse för detta, erhållen från min tidigare analys av de apriorikriterier vi anser det vara riktigt att använda i vetenskapshistoria, i kapitel 3 i Swinburne 2004. (För en mer fullständig version, se kapitel 4 i Swinburne 2001.) Denna redogörelse påverkar hur vi bör kvantifiera apriorisannolikheten för en konstant eller variabel som ligger i ett intervall av möjliga värden – vilket är högst relevant för påståendet att naturlagarnas konstanter och universums randvillkor måste vara ”finjusterade” för att ge upphov till människor. Konsekvenserna för min redogörelse av enkelhet (i grova drag – för vissa tillägg, se Swinburne 2004, s. 172–82) är att vi bör välja den enklaste formen av den mest troliga bakgrundsteorin (i vilken konstanterna har ett godtyckligt värde i intervallet av möjliga värden), och tilldela samma värde till apriorisannolikheten till konstanter som ligger inom intervall av samma längd. Den enklaste formen av gravitationsteorin består av Newtons tre rörelselagar och gravitationslagen, av formen $F = mm'/r^2$; så därför antar vi samma apriorisannolikhet att gravitationskonstanten G ligger mellan – exempelvis 30 och 40 – som att G ligger mellan 40 och 50. Om vi uttrycker Newtons teori i en mer komplicerad form i termer av G^{-1} , och antar att G^{-1} hade samma trolighet att ligga inom varje intervall av samma längd, skulle vi naturligtvis få ett annat resultat för de intervall inom vilka det är lika troligt att finna G . Men, i motsats till Häggström (s. 9–10), följer det av min detaljerade redogörelse att G^{-1} inte är en ”precis lika naturlig och fundamental storhet som G ” – just därför att G , och inte G^{-1} , förekommer i den enklaste och mest naturliga formuleringen av Newtons teori.² Naturligtvis kan man inte ge ett exakt värde till den inneboende sannolikheten för någon vetenskaplig eller historisk hypotes, men man kan ge grova relativa värden, vilket är tillräckligt för att vi ska kunna bedriva vetenskap.

I Swinburne 2004 tillämpade jag det här ramverket för att uppskatta sannolikheten för teism, definierad som hypotesen (som jag kommer att benämna ” h_1 ”) att det i evighet existerar en allsmäktig, allvetande och

²Jag är rädd för att Häggström har misstagit sig då han att påstår (genom att citera Rees arbete) att G (till skillnad från andra konstanter i naturlagarna) ”inte verkar kräva någon finjustering”. Se Robin Collins 2009 för en ytterst noggrann genomgång av den grad av finjustering som krävs för att människor ska utvecklas. Collins arbete publicerades efter Rees, och Collins citerar även Rees. Collins betonar att G kräver en finjustering relativt andra konstanter och variabler, and drar slutsatsen (s. 214–5) att – givet nuvarande uppskattningar av dessa – den måste ligga inom en på 1060 av sitt nuvarande värde för att människor ska kunna ha utvecklats.

fullständigt fri varelse eller väsen. Jag valde e som de mest generella egenskaperna hos universum och människor. De mest framträdande av dessa egenskaper är universums välordning (dess anpassning till enkla "naturlagar"), och det faktum att dess lagar och begynnelsevillkor (eller, mer generellt, om det inte har någon början, dess "randvillkor") är sådana att de leder till människors existens. Inom ramen för denna korta artikel ska jag, precis som Häggström, endast beakta relevansen av den senare typen av evidens, som jag benämner "e", och jag definierar apriorisannolikheten för en hypotes som dess inneboende sannolikhet. I Swinburne 2004 sökte jag visa, med e och h_1 tolkade på detta sätt, att (i) $P(e|h_1)$ är ganska stor, (ii) $P(h_1)$ är betydligt större än $P(h_i)$ för alla konkurrerande h_i , och (iii) $P(e|h_i)$ är väldigt liten för alla de h_i (förutom h_1) för vilka $P(h_i)$ har ett värde som ligger någonstans i närheten av $P(h_1)$. Av detta följer (givet den tolkning av "ganska" stor, "betydligt" större, och "någonstans i närheten av" som jag sökte att motivera med detaljerade argument), beroende på exempelvis antalet konkurrerande hypoteser n , deras apriorisannolikheter ($P(h_i)$), och deras prediktiva styrkor ($P(e|h_i)$), att $P(h_1|e)$ är betydligt större än $1/2$.

Jag sammanfattar mina argument för (i), (ii) och (iii), och börjar med (i). Jag argumenterade för att ett allvetande väsen vet vilka handlingar som är goda, och en fullständigt fri varelse är endast föremål för *kausal* påverkan, och kommer i sina handlingsval endast påverkas av *förnuft*. Han kommer alltid i varje situation att syfta till att företa det bästa sättet att handla (den som det är mest förnuftigt att företa) eller – om det finns två eller flera val som är "båda är bäst" – företa sig ett av dem, och som allsmäktig kommer han att lyckas med det. Häggström påstod (s. 10) att min definition av Gud som "fullständigt fri" i den meningen att "ingen-ting på något kausalt sätt påverkar hans val" (Swinburne 2004, s. 94) inte var förenligt med min åsikt att Gud ibland svarar på böner. Det avgörande ordet i min definition är "kausal". Mänskliga val påverkas delvis av förnuftet (på så sätt att vi väger skälen för att göra saker, och sedan handlar på ett sätt som är mest förnuftigt), men delvis också av icke-rationella orsaker. Jag drog slutsatserna för en Gud, som inte påverkas av den senare typen av val några sidor senare, och skrev att "det följer av Guds fullständiga frihet att endast förnuftet påverkar honom" att göra ett val (s. 98). Så när det gäller att välja vilka böner som ska besvaras och inte, är Gud influerad av förnuftet för att svara på en viss bön och inte på en annan. Jag är skyldig Häggström en ursäkt för att inte ha gjort denna distinktion mellan orsak och förnuft direkt på den sida som Häggström citerar, utan först några sidor senare.

Jag argumenterade för att vi människor är vad jag benämnde "männsliga fria agenter", och som sådana har en unik typ av godhet (som inte

ens Gud har), förmågan att välja mellan gott och ont, och därför att genom våra val kunna påverka och göra en skillnad gentemot oss själva, varandra och omvärlden. Således kan en god Gud förväntas förorsaka vår existens (eller mer precist, existensen av mänskliga fria agenter, som vår art tillhör) med en betydande sannolikhet. Därefter argumenterade jag för att e (som beskrivits ovan) är ett nödvändigt villkor för vår existens, och därför kan Gud förväntas förorsaka e . Därför är $P(e|h_1)$ ganska hög.

Jag gav argument för påståendet att vi har en unik typ av godhet. Det behövs ofta en längre bevisföring och mycket erfarenhet för att övertala någon om moraliska utsagor – därav behovet av vad Häggström verkar avfärda i termer av något som ”formar sig närmast till en lång predikan” (s. 8). Jag koncentrerade mig inte på universums ”skönhet”, utan snarare på godheten i att moraliskt fria agenter existerar, såsom definierats ovan. Häggström avfärdar allt detta utifrån argumentet att våra synpunkter i dessa frågor formats av ”biologisk” och ”kulturell” evolution. Det har även våra föreställningar om matematik, vetenskap och historia! Men det innebär inte att dessa föreställningar inte är sanna, bara att ”evolutionen” har hjälpt oss att se att de är sanna. Jag försvarade översiktligt moralisk objektivism (2004, s. 113–4, ett synsätt som avser sanningspåståenden av vad som är gott och ont, och som många ”analytiska” filosofer, som inte alls är sympatiskt inställda till teism, förespråkar), och argumenterade utifrån den intuition som många av oss har om vad som är bra och dåligt med mänskliga fria agenter, och att människan därför är en mycket bra typ av varelse. Härav följer att, om det finns en Gud, så är sannolikheten betydande att han skulle frambringa existensen av mänskligt fria agenter, exempelvis människor.

Jag övergår till att behandla (ii), att jämföra apriorisannolikheten för h_1 med motsvarande sannolikheter för ett antal konkurrerande påstådda ultimata förklaringar (det vill säga sådana som inte anser sig vara förenliga med ännu mer ultimata förklaringar). Tvärtemot Häggströms påstående erbjöd jag en ”specifikation av mängden $\{h_1, h_2, \dots\}$ av tänkbara [förklarande] teorier”. Jag påstod att det bara finns två sätt att förklara företeelser, som vi båda ofta använder oss av – ”personliga” förklaringar i termer av personer och deras förmågor, trosföreställningar och avsikter; och ”vetenskapliga” (eller, som jag snarare skulle säga idag, opersonliga) förklaringar i form av naturlagar och randvillkor. Sedan argumenterade jag för att naturlagarna främst kan betraktas som beskrivningar av de förmågor och benägenheter hos individuella substanser (exempelvis elementarpartiklar), från vilket det följer, som ”kalla fakta”, att varje ultimat vetenskaplig förklaring av universums allmänna egenskaper innehåller existensen av en enorm samling individuella substanser med inbyggda förmågor och benägenheter, eller kanske i stället en fysisk sub-

stans med en väldigt komplicerad inbyggd förmåga och benägenhet att producera andra substanser. Så – vilken den senaste teorin i fysiken än är – kommer den att erbjuda en väldigt komplicerad slutgiltig förklaring av e , och därför – i slutändan – ha en väldigt låg apriorisannolikhet. Teorier som innebär att e ges en personlig förklaring varierar i enlighet med det antal gudar de postulerar, och med graden av förmåga, kunskap och frihet som tilldelas dem. Jag argumenterade för att antagandet om en enda Gud med oändliga kvantiteter av dessa egenskaper (det vill säga ett nollgränsvärde vad gäller graden av dessa egenskaper) är ett väsende som är allsmäktigt, allvetande och har fullständigt frihet, och därför utgör ett väldigt enkelt antagande som relativt alla konkurrerande teorier av personkaraktär såväl som teorier av vetenskaplig karaktär, har en väldigt hög apriorisannolikhet.

Oändliga grader av egenskaper kan analyseras i termer av nollgränsvärden av dem, medan ett oändligt antal substanser (till exempel elementarpartiklar) inte kan det. Det beror på att substanser är individuella, och ett stort antal av dem utgör många individuella byggstenar. Egenskaper är inte så beskaffade – de har en hög grad, inte stora kvantiteter – och en hög grad av en egenskap består inte av många individuella egenskaper hos sina komponenter. Jag kommenterade explicit denna skillnad (2004, s. 55). Så Häggströms påstående att jag glömt min syn på enkelhet hos ”kvantiteter” när jag diskuterar multiversum, är oberättigat. De två typerna av ”oändligheter” är radikalt olika.

I ljuset av min omfattande diskussion i Swinburne 2004 om varför jag gav h_1 en hög apriorisannolikhet jämfört med alla dess rivaliserande hypoteser h_i , är det absurt av Häggström att föreslå att jag begått ”bridgespelarens lapsus” genom att tillåta mig att tilldela teism en hög apriorisannolikhet på ett godtyckligt sätt. Jag gav detaljerade argument, baserat på kriterier som används inom vetenskapshistoria, för att tilldela de värden $P(h_1)$ och $P(h_i)$ jag gjorde till h_1 och dess rivaler. Jag tog mig även besväret att explicit (2004, s. 131–32) förklara att jag inte begått just ”bridgespelarens lapsus”.

Slutligen kommer jag till (iii). Medan varje fysikteori som gör anspråk på att ge en ultimata förklaring, skulle möjliggöra e , så upprepar jag mitt påstående att postulering av ett väldigt stort och möjligtvis oändligt stort antal substanser, alla med samma förmåga och benägenhet (det vill säga, som alla ”lyder” under samma ”naturlag”), som en *ultimata* förklaring, är att postulera en teori vars apriorisannolikhet är oerhört låg. Häggström skriver att jag ”avfärdar multiversumteorin”. Jag tydliggjorde att vetenskapen kan komma fram till att vårt universum är ett av många, som bildar ett ”multiversum” (2004, s. 165 och 185–87). Men min poäng var att vi endast kan ha skäl till att tro detta om den enklaste

fysikteorin, som förklarar fenomen i vårt universum, hade konsekvensen att andra universa ”knoppades av” från ett tidigt stadium av vårt universums historia. I så fall skulle vårt hela multiversum lyda under vissa lagar, och eftersom ett ouppräkneligt antal logiskt möjliga multiversa inte skulle innehålla ett universum finjusterat för liv, måste vårt multiversum självt vara finjusterat för att kunna åstadkomma ett sådant universum. Och detta troliggör i sin tur den enklaste förklaringen, i termer av Gud. (För en mer detaljerad analys av den effekt som upptäckandet av ett multiversums existens skulle ha på sannolikheten för teism, se mitt arbete 2004.) En teori med många oberoende gudar kan tyckas vara nästan lika enkel som teism, med det finns inga skäl att anta att sådana gudar skulle söka att producera ett gott universum i fullständigt samarbete. Absolut godhet följer av de tre gudomliga egenskaperna, men så snart vi postulerar en eller flera gudar som inte är allvetande och helt fria, så följer inte godhet utan det måste läggas till som en extra icke härledd gudomlig egenskap – vilket komplicerar hypotesen.

Så jag hävdar att jag etablerat mitt resultat att sannolikheten för teism, givet de ovan beskrivna evidenserna, är ganska hög, mer precist – av skäl som utvecklats i detalj i Swinburne 2004 – ”betydligt större än 1/2”. Häggströms kritik tenderar överhuvudtaget inte att minska styrkan i mina argument.

LITTERATUR

- Collins, Robin. 2009. ”The Teleological Argument: An Exploration of the Fine-Tuning of the Universe”. I *The Blackwell Companion to Natural Theology*, red. William Lane Craig och J. P. Moreland, s. 202–81. Chichester: Wiley-Blackwell.
- Häggström, Olle. 2008. ”Brister i sannolikhetsargument för och emot Guds existens”. *Filosofisk tidskrift*, nr 4, s. 3–11.
- Swinburne, Richard. 2001. *Epistemic Justification*. Oxford: Oxford University Press.
- Swinburne, Richard. 2004. *The Existence of God*. Andra uppl. Oxford: Oxford University Press.
- Swinburne, Richard. 2007. ”Argument för Guds existens”. *Filosofisk tidskrift*, nr 4, s. 26–40.
- Swinburne, Richard. 2012. ”God, Bayes, and the Multiverse”. I *Probability in the Philosophy of Religion*, red. Jake Chandler och Victoria S. Harrison, s. 103–23. Oxford: Oxford University Press.

Martin Lembke har nyligen presenterat en princip med vars hjälp han anser sig kunna förklara varför det inte kan finnas någon lösning på ”varrats gåta” – dvs. frågan om varför det finns något överhuvud taget – och som dessutom ”låter oss dra flera slutsatser av betydande metafysiskt intresse” och som utgör ett ”kunskapsteoretiskt stöd för en teistisk, snarare än naturalistisk, världsbild”.¹ Principen ser ut så här:

UP För varje klass K gäller, om K inte är tom, att det finns en förklaring till att K inte är tom om och endast om det finns minst en icke-medlem av K som gör så att K inte är tom.

Denna princip tycks säga att den enda förklaring som kan finnas till att en (icke-tom) klass är icke-tom är att något utanför klassen ”gör så att” den är icke-tom.

Men är verkligen UP en vettig princip? Det naturliga sättet att angripa denna fråga är att försöka hitta några motexempel mot UP.

Låt K vara klassen av primtal mellan 2 och 5. Den är inte tom, ty den innehåller talet 3. En förklaring till att den inte är tom är att det finns ett tal mellan 2 och 5, som är ett primtal, nämligen 3. Det är ett primtal eftersom det är ett positivt heltal som inte är produkten av två mindre heltal. Detta är en förklaring, men denna förklaring består inte av att det finns minst en icke-medlem av K som ”gör” så att K inte är tom. Alltså ett motexempel mot UP.

Lembke skulle kanske vilja invända att $\{3\}$, dvs. den mängd som enbart innehåller talet 3, inte är en ”klass” i den meningen att den är extensionen av någon egenskap. Men det är den väl. Den är extensionen av egenskapen att vara identisk med talet 3. Den är också extensionen av flera andra egenskaper, t.ex. egenskapen att vara ett positivt heltal större än 2 och mindre än 4. (De som vill ha en ontologi som innehåller egenskaper – vilket jag för min del gärna skulle vilja undvika – kan

¹Lembke 2013. En utförligare version av hans tankegångar finns i hans doktorsavhandling i religionsfilosofi vid Lunds universitet: *Non-Gods and Gods: A Cosmological Treatise*, 2012.

ju tvista om huruvida de två egenskaper jag här har nämnt är identiska eller inte.)

Men Lembke vill kanske bara godkänna motexempel av icke-matematiskt slag. Ta då följande. Låt K vara klassen av regndroppar som föll på mitt huvud en viss eftermiddag i höstas. Den var inte tom. Men förklaringen till att det föll regndroppar på mitt huvud då borde väl hänvisa till meteorologiska förhållanden samt det faktum att jag befann mig utomhus på en viss plats, inte till att det fanns någon icke-regndroppe, eller någon regndroppe som inte föll på mitt huvud, som "gjorde" så att vissa regndroppar föll på mitt huvud. Eller vill Lembke säga att det fanns ett regnmoln som "gjorde" det? En sådan förklaring skulle i alla fall inte vara tillräcklig, ty molnet "gjorde" knappast att jag befann mig utomhus.²

Ett sista exempel. Lembke påstår att en följsats av UP är följande: För alla komplementära klasser K och icke- K gäller, att även om varken K eller icke- K är tom så är minst en av dem sådan att dess icke-tomhet saknar förklaring. Men låt då K vara extensionen av egenskapen att vara antingen ett bord eller en bagare. Icke- K är då extensionen av egenskapen att vara varken bord eller bagare; där ingår bl.a. en tårta som gjorts av en bagare. Någon snickare i icke- K har gjort ett bord i K , vilket förklarar att K inte är tom. Och en bagare i K har gjort en tårta som ingår i icke- K , vilket förklarar att icke- K inte är tom. Alltså är det Lembke anser vara en följsats av UP falsk. Vilket talar för att även UP är falsk.

Principen UP verkar alltså inte särskilt plausibel. Men låt oss för resonemangets skull ändå anta att den är riktig. Vad händer då?

Lembkes viktigaste tes är väl att UP ger ett stöd åt en *teistisk* världsbild. Kan det verkligen stämma? Hans argument är att teismen kan förklara "mer" än naturalismen. Han säger: "Det kunskapsteoretiska idealet bör vara att förklara så mycket som möjligt" (s. 7).

Det är inte helt klart vad det är som ska jämföras, men jag tror att Lembke menar att "naturalismen" eller "en naturalistisk världsbild" kan anses vara ungefär det som dagens vetenskaper anser vara sant. "Teismen" eller "en teistisk världsbild" kan då sägas vara detsamma som naturalismen *plus* antagandet att det finns en allsmäktig gud.³ Lembke

²Kanske skulle Lembke vilja säga att det är Peanos axiom för aritmetiken som har "gjort" så att det finns ett primtal mellan 2 och 5 och att världens tillstånd i övrigt "gjorde" så att regndroppar föll på mitt huvud. Och att dessa "saker" är element i lämpliga klasser och kan "göra" så att UP satisfieras.

³För enkelhetens skull får vi väl här anta att dagens vetenskap är förenlig med antagandet att det finns en allsmäktig gud – något som kanske inte är självklart. Alternativt kan vi säga att teismen innehåller allt som dagens vetenskap säger, utom sådant som är oförenligt med att det finns en allsmäktig gud. Men det blir i så fall svårare att veta vad teismen innebär.

säger att ”den teistiska världsbilden innehåller ju rimligen *alla* objekt som också figurerar i en naturalistisk dito, *plus* Gud (och eventuellt andra övernaturliga objekt)”. Vi måste väl anta att teismen på detta sätt parasiterar på naturalismen, om det ska stämma att den kan förklara allt som naturalismen kan förklara. (Ty om teismen *bara* säger att det finns en allsmäktig gud, så förklarar den ingenting – eftersom den inte säger något om vad guden använder sin makt till. På motsvarande sätt skulle naturalismen inte förklara något om den *bara* säger att ingenting är allsmäktigt eller övernaturligt.)

Men om teismen tolkas på detta sätt, så förklarar den inte mer än naturalismen, utan snarare mindre. Den förklarar allt som naturalismen förklarar, men inte att det finns en allsmäktig gud. Det senare förklaras förstås inte heller av naturalismen, men det behövs ju inte, eftersom detta enligt naturalismen inte är ett faktum. Så rent ”procentuellt”, så att säga, förklarar naturalismen mer.

Här invänder förstås Lembke att teismen förklarar något som naturalismen inte förklarar, nämligen att allt det som naturalismen handlar om finns. Ty teismen säger att det är skapat av Gud. Men *det* säger ju faktiskt *inte* teismen, som vi nyss beskrev den. Vi kan förstås lägga till detta, som ett extra element i teismen. Så låt oss göra det. Men då innehåller teismen en hypotes, ”Gud skapade världen”, som inte förklarar något – även om den förstås implicerar att världen finns. Ty för att den ska kunna godtas som en *förklaring*, så måste den vara (empiriskt) prövbar och prövad med framgång, vilket teismen inte är.⁴

Jag inser att denna tankegång kan vara lite svårsmält, eftersom väldigt många tycks ha vant sig vid idén att man kan förklara att världen finns med antagandet att en gud har skapat den. Jag försökte illustrera det problematiska i detta med ett exempel av detektiv-typ, men låt mig nu använda ett lite annorlunda exempel.

Anta att vi vill förklara varför Lembke i första meningen i sin artikel bl.a. skriver att ”Lars Bergström och jag är överens i åtminstone två väsentliga ting”. Varför skriver han just så? Låt oss anta att jag försöker ”förklara” detta med att en ärkeängel styrde hans fingrar till att forma dessa ord (i stället för t.ex. ”Bergström och jag har samma uppfattning i åtminstone två avseenden”). Vore detta en förklaring? Nej, den *säger* ju bara att en ärkeängel styrde Lembkes fingrar på detta sätt. Den *förklarar* inte varför fingrarna styrdes på detta sätt. Ty det den säger, utöver att fingrarna styrdes på detta sätt, är något som hänger fullständigt i luften,

⁴I mitt tidigare inlägg citerade jag bl.a. uttalandet ”If a hypothesis is untestable, it cannot tell us anything” från Quine och Ullians lärobok *The Web of Belief*. – Här och i fortsättningen bortser jag från förklaringar i t.ex. matematik och moral. Jag håller mig till förklaringar av hur det förhåller sig i den empiriska verkligheten.

eftersom det inte finns något (annat) som talar för att det är sant. Att det var en ärkeängel som styrde hans fingrar på detta sätt medger inga (andra) förutsägelser som kan kontrolleras. Det finns därför ingen oberoende evidens för detta. Och därför gör det inte att vi vet mer om något än vi redan tidigare visste. Vi kanske *tror* något mer, men vi *vet* inte mer.

En hypotes som kan förklara något måste kunna *prövas*, vilket i sin tur betyder att den kan generera *förutsägelser* om sådant som vi ännu inte känner till. Det räcker alltså inte att hypotesen att en ärkeängel styrde Lembkes fingrar också kan utvidgas till att ärkeängeln även styr *mina* fingrar när jag skriver detta och att den också har orsakat en massa andra saker som vi känner till, t.ex. att det regnade i går. På samma sätt med en allsmäktig gud. Det räcker inte att man antar att hen har orsakat en massa fysikaliska och historiska fakta som vi känner till. Teismen måste också implicera något som vi ännu inte känner till, för att den ska kunna prövas. Om teismen t.ex. var så innehållsrik att den, eventuellt i kombination med diverse kända fakta och naturlagar, implicerade att läkarna under april månad år 2015 kommer att hitta ett säkert botemedel mot alla sorters cancer, så kan vi vänta till 2015 och då se om förutsägelsen stämmer. Skulle den göra det, och skulle hypotesen på liknande sätt göra andra lyckade förutsägelser, utan att någonsin misslyckas, så har vi ett gott – men naturligtvis inte konklusivt – skäl att tro på teismen. Och först då skulle vi ha anledning att tro att den kan förklara det som Lembke pekar på, nämligen att det finns icke-allsmäktiga ting.

Att en hypotes har förklaringskraft endast om den är prövbar innebär dock inte att vi *alltid* måste invänta att dess förutsägelser bekräftas innan vi kan godta den. I vardagslag godtar vi ofta hypoteser för att vi helt enkelt är övertygade om att deras förutsägelser kommer att slå in. Att jag ser en bok på bordet framför mig förklarar jag t.ex. direkt med att det faktiskt ligger en bok på bordet framför mig. Jag väntar inte på att även andra människor ska se det jag ser, innan jag godtar hypotesen. Men det beror på att jag är så säker på att även andra skulle se en bok där jag ser den.

Ett annat exempel på en förklaring av en iakttagelse som vi godtar ganska direkt, utan att kräva ytterligare förutsägelser, kunde vara att vi ser spår av skor i snön på en väg och förklarar det med att en människa har gått där. Detta grundas på att vi tror eller vet mycket annat, t.ex. att människor brukar ha skor när de går i snö, att de brukar lämna spår efter sig, att de är väldigt osannolikt att spåren kunde ha orsakats av något annat (t.ex. en gorilla eller häst i skor eller en ballongfarare som med hjälp av en sko på en lång pinne lyckats åstadkomma dessa spår, etc.).

Men teismen är inte en förklaring av något sådant vardagligt slag. Även om jag har en upplevelse av att en gud finns, så kan jag inte betrakta denna upplevelse som evidens för att guden finns, om inte upplevelsen

skulle uppkomma hos vem som helst i samma situation. Till skillnad från synintryck och andra empiriska upplevelser är ju gudsupplevelser inte alls intersubjektiva. De delas inte av alla. Och jag har inte tillgång till någon oproblematiserad erfarenhet av att gudar i allmänhet brukar skapa världar som ser ut som vår eller till någon teori enligt vilken detta är betydligt sannolikare än att världen har uppstått på något annat sätt, t.ex. att den skapats av en stor och elak fågel som därefter omedelbart upphörde att existera eller av en plötslig och metafysiskt nödvändig explosion som liknar Big Bang.

Lembke tycks dock inte godta detta resonemang. När jag tidigare, på liknande sätt, hävdade att det inte finns någon oberoende evidens för att det finns en allsmäktig gud svarar Lembke att det är just det empiriska faktum att det finns icke-allsmäktiga ting som motiverar antagandet att det finns minst en allsmäktig varelse.⁵ På motsvarande sätt skulle han väl då säga att det faktum att han skrev just de ord han skrev motiverar antagandet att det finns en ärkeängel som styrde hans fingrar. Nog låter det konstigt!

Men såvitt jag förstår är kärnan i Lembkes framställning faktiskt just detta! Att världen finns, och är som den är, motiverar enligt honom antagandet att det finns en allsmäktig gud (som enligt naturalismen inte finns i världen). Detta är han inte ensam om att tycka, men det är ju något som inte alls följer av principen UP. Enligt UP kan man på sin höjd säga att det finns en förklaring till att naturalismens värld finns *om och endast om* det finns något utanför denna värld som skapat den. Men att världen är som den är *motiverar* inte enligt UP ett sådant antagande. Detta antagande, om en övernaturlig aktör utanför – men ändå kausalt i kontakt med – världen, är gripet ur luften.

Det stämmer inte ens, som Lembke säger, att om teismen skulle vara *sann*, så skulle den kunna förklara något som naturalismen inte kan förklara. Ty en förklaring har vi inte innan vi har en teori *som vi har goda skäl att hålla för sann* (och som dessutom implicerar det som ska förklaras – eller åtminstone gör det sannolikt; detta är ett nödvändigt, men inte tillräckligt, villkor). Enbart det faktum att teorin *implicerar* det som ska förklaras är *inte* ett gott skäl att hålla den för sann.

LITTERATUR

- Bergström, Lars. 2012. "Varats gåta", *Filosofisk tidskrift* 33, nr 1.
 Lembke, Martin. 2010. "Frågan om varat", *Filosofisk tidskrift* 31, nr 3.
 Lembke, Martin. 2013. "Varat: en fråga om uteslutning", *Filosofisk tidskrift* 34, nr 2.

⁵Se Lembke 2013, fotnot 13.

Notiser

RÄTT SKA VARA RÄTT

Prefekten för Filosofiska institutionen i Stockholm fick för en tid sedan ta del av följande meddelande:

Hej, skriver till dig för att på något sätt nå en kanal till dem som håller på med Friedrich Nietzsche. Har under en längre tid retat mig på att de flesta uttalar Nietzsche som om det stavades med ”ch”, dvs. som ett främre tj-ljud – t.ex. som ”tjej” i stället för ”sch” som i t.ex. ”dusch”. Det borde inte vara främmande för oss att uttala ett ”sch” korrekt, vårt språk som har mycket tyskt språkligt påbrå. Dessutom har vi ju många ord som har detta ljud. Lyssnar ofta på radioprogrammet Filosofiska rummet och kända filosofer uttalar ofta namnet fel. Inför öppnandet av Munchutställningen på Thielska galleriet där ju även Nietzsche är ett tema uttalade t.f. intendenten namnet med tj-ljud ett flertal gånger i TV. Han är ju ingen expert så han borde inte lastas egentligen men jag skrev ändå till museet och bad dem påtala detta för honom. Men nu vädjar jag till Dig att göra en insats på era kurser och kanske i någon aktuell publikation. Jag tycker verkligen att Nietzsche förtjänar ett bättre öde än att ideligen uttalas felaktigt. De flesta reagerar ju själva på hur ens namn uttalas!

Med vänliga hälsningar, Lillo Sjöberg, Stockholm.

NYA BÖCKER

Mats Furberg, professor emeritus i Göteborg, har utgivit en självbiografisk bok med titeln *Jesu barnbarn. En väg till filosofin*, Thales 2013.

Fredrika Spindler har skrivit *Deleuze, tänkande och blivande*, Glänta Produktion 2013.

Immanuel Kants bok *Himlens allmänna naturhistoria och teori* har utgivits i svensk översättning på Gidlunds förlag.

Jan Sjunnesson har redigerat boken *Kommunitarism: En antologi om individualism och gemenskap i moral, politik och filosofi*. Vulkan.se, Stockholm 2013.

Peter Ryman har publicerat sin doktorsavhandling i praktisk filosofi,

On Free Will as Categorical and Conditional Freedom, Uppsala universitet 2012.

Lundafilosoferna Johannes Persson och Nils-Eric Sahlin har utgivit en tjock bok med titeln *Vetenskapsteori för sanningssökare*, Fri tanke 2013.

Att drottning Kristina ägnade en stor del av sitt liv åt alkemi är kanske inte allmänt bekant. Det är ämnet för Susanna Åkermans bok *Fenixelden. Drottning Kristina som alkemist*, Gidlunds förlag 2013.

Förlaget Anomali ger ut ”samtida filosofisk litteratur av icke-akademisk karaktär”, nu senast Johan Argus, *Filosofins födelse: En kritik av det textuella förnuftet* (2012) och Gunnar Odhner, *Ingenting är fördolt* (2012).

Förlaget Fri tanke har utgivit Eskil Franck, *Giv mig, min son, ditt hjärta* (2013).

Fri tanke har även utgivit Michail Gorbatsjovs mycket intressanta självbiografi *Som jag minns det*, 2013. Bland annat skildras hans försök att, med inspiration från den sene Lenins tänkande och mot nomenklaturans intressen, demokratisera Sovjetunionen.

KATEGORISKA IMPERATIVET SOM ÄNDAMÅL

Det *kategoriska imperativets* två första formuleringar är de följande: (1) Handla endast efter den maxim genom vilken du tillika kan vilja att den blir en allmän lag. (2) Handla aldrig så att du behandlar mänskligheten i din egen eller någon annan person enbart som medel, utan alltid tillika som ett ändamål. Dessa två formuleringar leder till en paradox om en agent A utför en handling som berör personen P för att uppfylla (1). P blir då ett medel för A eftersom dennes ändamål inte är P utan (1). Således leder (1) till att (2) inte kan uppfyllas.

Gustav Alexandrie

Medarbetare i detta nummer av *Filosofisk tidskrift*: Lars Bergström är professor emeritus i praktisk filosofi, Stina Bäckström har doktorerat i filosofi vid universitetet i Chicago och är för närvarande post doc forskare i filosofi vid Åbo Akademi, Karl Ekendahl är doktorand i praktisk filosofi i Uppsala, Lars-Göran Johansson är professor i teoretisk filosofi i Uppsala, Patrik Ribe är gymnsielärare i filosofi i Uppsala, Hans Ruin är professor i filosofi vid Södertörns högskola, och Richard Swinburne är professor emeritus i filosofi i Oxford.

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
- 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