

Filosofisk tidskrift

Nr 4 Årgång 42 November 2021

3 Hotet från medvetna maskiner

Lars Bergström

12 Problemet med kollektiv påverkan

Andrea S. Asker

28 Heidegger mot moderniteten – en instruktiv förbrytelse?

Erik Hallstensson

Kommentar

35 Mycket sent om Hedenius och Mathlein

Johan E. Gustafsson

Recensioner

37 Hans Teke om *Why Solipsism Matters* av Sami Pihlström

41 Anders Hansson om antologin *Filosofi och pandemi*

45 Jesper Jerkert om *Forskningsetik för humaniora*
av Håkan Salwén

50 Martin Lembke om *Är Gud annorlunda?* av Lennart Nordenfelt

53 Robert Callergård om *Novum Organum* av Francis Bacon

55 Björn Lundberg om antologin *Oumbärliga samtal: Humaniora
och samhällsvetenskap i svensk offentlig debatt*

58 Lars Bergström om *Medvetandets mysterium* av Annaka Harris

Filosofisk tidskrift

Utges av Stiftelsen Filosofisk tidskrift och Stiftelsen Bokförlaget Thales

Redaktörer: Jens Johansson (ansvarig utgivare) och Olle Risberg

Redaktionsassistent: Karin Enflo

redaktion@filosofisktidskrift.se www.filosofisktidskrift.se

Filosofisk tidskrift utkommer med fyra nummer per år

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

Lösnummer: 55 kr (inkl. moms)

För prenumerationsärenden kontakta:

Nätverkstan ekonomitjänst, Box 31120, 400 32 Göteborg

Telefon 031-743 99 05

ekonomitjanst@natverkstan.net

Förlagsadress: Bokförlaget Thales, c/o Filosofiska institutionen,

Stockholms universitet, 106 91 Stockholm

info@bokforlagetthales.se www.bokforlagetthales.se

Copyright © Respektive författare 2021

Grafisk form och produktion: Ulf Jacobsen

Satt med Arnhem Blond och Joanna Sans Nova

Tryck: Carlshamn Tryck & Media AB, Karlshamn 2021

ISSN 0348-7482

Hotet från medvetna maskiner

Lars Bergström

Många framstående futurologer och dataloger varnar för att utvecklingen inom området artificiell intelligens (AI) ganska snart kan leda till superintelligenta maskiner (AGI), som eventuellt tar kontroll över världen och utrotar eller förslavar oss människor. Även datalogins fader Alan Turing tycks ha tänkt i de banorna. Om varningarna är befogade bör vi nog snart försöka vidta motåtgärder. Den svenske matematikern Olle Häggström har i sin senaste bok *Tänkande maskiner* (2021) förslag på vad som bör göras. Häggström skriver bra, medryckande och engagerat, och boken förefaller mycket initierad. Litteraturlistan omfattar runt trehundra referenser, inklusive en del filosofi.

Förkortningen AGI står för "artificiell generell intelligens", närmare bestämt på minst mänsklig nivå. Hittills har något i den vägen inte synts till, de mest spektakulära AI har bara varit framgångsrika på mycket speciella områden, som t. ex. schackspel, bilkörning och ansiktsigenkänning. Det råder ingen enighet bland experterna om, och i så fall när, AGI kan förväntas uppstå, men somliga tror att det kan ske redan i vårt århundrade.

Som stöd för uppfattningen om AGI:s farlighet hänvisar Häggström huvudsakligen till det han kallar Omohundro-Bostrom-teorin (s. 139). Såvitt jag förstår består den i att varje AGI har ett övergripande "slutmål" och att den är beredd att tillgripa vilka medel som helst för att uppnå detta mål – inklusive sådant som innebär mänsklighetens utplåning eller förslavning. Jag ska återkomma till detta.

En förutsättning tycks alltså vara att det uppstår maskiner som har uppnått en intelligens som motsvarar eller överskrider människans. Man kan undra om en sådan AGI också är medveten. Kan den förstå vad den gör? Kan den inte bara tänka, utan också förstå vad den tänker? Eller hur man nu ska uttrycka det. Detta är något som också har intresserat filosofer.

Häggström diskuterar detta i kapitel 9 av *Tänkande maskiner*. Han har tidigare varit inne på ämnet i flera andra sammanhang, bland annat i sin bok *Here Be Dragons* (2016). Han är benägen att godta det som brukar

kallas ”beräkningsteorin om medvetande” (*the computational theory of mind*, förkortat CTM), vilken lite förenklat innebär att medvetande består av beräkningar eller symbolmanipulation, något som även kan produceras digitalt i en dator.

En mycket känd invändning mot CTM har konstruerats av filosofen John Searle, i ett tankeexperiment som kallas ”Det kinesiska rummet”. Häggström invänder i sin tur mot Searles resonemang, och jag ska här försöka reda ut vem av dem som har rätt.

* * *

I det kinesiska rummet sitter Searle själv i ett rum där skriftliga meddelanden på kinesiska skickas in genom ett fönster och där Searle – som inte kan kinesiska – med hjälp av instruktioner i en regelbok, och ett stort förråd av kinesiska tecken, skriver ned andra kinesiska tecken på papperslappar som han sedan skickar ut genom fönstret som svar på de budskap som skickats in. Reglerna är konstruerade av någon som kan kinesiska, så utbytet av papperslappar utgör en helt begriplig konversation. Begriplig för den som kan kinesiska, alltså.

Det enda medvetande som finns i rummet tycks vara Searles, men han kan alltså inte kinesiska. Hur kan då konversationen fungera? Ja, det svar man omedelbart tänker på är förstås att det är rummet som helhet, med alla dess ingredienser, som kan kinesiska. Detta svar kallar Searle ”systemsvaret”. De mest kända företrädarna för systemsvaret är Douglas Hofstadter och Daniel Dennett, och de har formulerat svaret i sin bok *The Mind's I* (1981), där också Searles artikel är omtryckt.

Men Searle invänder redan i sin ursprungliga artikel mot systemsvaret. Han menar att rummet med dess hjälpmedel i princip kan elimineras genom att han själv memorerar regelboken och förrådet av tecken och därefter tillämpar reglerna med hjälp av sin egen hjärna. Då kan han alltså konversera på kinesiska, trots att han inte kan kinesiska.

På detta svarar Hofstadter och Dennett att Searle faktiskt kan kinesiska när regelboken är så bra att konversationen fungerar. Häggström ansluter sig också till systemsvaret, men han säger att han vill ge det en annan tolkning. Han hävdar att det uppstår ”en multipel personlighetsstörning” i Searle, när han memorerar regelboken.

Måhända har Searle rätt i att han fortsatt inte kommer att begripa kinesiska, men i så fall talar han för bara den ena av de båda personligheter vi andra utifrån kan observera och samtala med och som tycks bebo hans

kropp: å ena sidan Searle-engelsktalaren som (av allt att döma sanningsenligt) bedyrar att han bara begriper engelska och inte ett ord kinesiska, och å andra sida Searle-kinesisktalaren som obehindrat (fastän en smula långsamt) för konversationer på kinesiska. (2021, s. 235)

Ja, så kan man tänka, men jag har inte lyckats begripa hur detta är något annat än Hofstadter och Dennetts uppfattning. Även de skiljer på två personer, en som talar engelska och en kinesisktalande (som de till och med föreställer sig är en kvinna; jfr 1981, s. 377); det är alltså fråga om två personer i Searles kropp, inte en person med två personligheter. Från och med nu ska jag kalla dem Searle-E och Searle-K.

Faktum är för övrigt att något liknande redan finns hos Searle, som talar om att "det finns i själva verket två delsystem i människan: det ena förstår engelska, det andra kinesiska" (ibid. 359, min övers.). I sin bok misstänker Häggström däremot, på grund av vad jag sagt i min recension av hans tidigare bok (Bergström 2016), att jag inte godtar existensen av Searle-K (2021, s. 236).

Den centrala frågan är förstas om Searle-K verkligen förstår vad som sägs i den kinesiska konversation som utspelar sig. Är det inte snarare så att hen bara manipulerar kinesiska tecken utan att veta vad de betyder? Den som utarbetade regelboken och kompilerade förrådet av kinesiska fraser förstår naturligtvis kinesiska, men den personen finns ju inte i Searles kropp, och inte heller i vare sig Searle-K eller Searle-E – eller ens i det ursprungliga rummet. Samma fråga kan man ställa sig när det gäller ett datorprogram som klarar ett Turingtest: förstår programmet (tillsammans med den hårdvara det körs på) det som sägs? Det är just det som Searle förnekar.

Hofstadter och Dennett tycks däremot anse att det är själva den enorma *mängden* av regler som skapar förståelse: "nästan all förståelse måste ligga i miljarderna av tecken i regelboken" (1981, s. 375, min övers.; jfr Häggström, s. 234). Några få regler räcker förstås inte för att den som hanterar reglerna ska förstå vad tecknen betyder, men Hofstadter och Dennett tycks anse att när reglerna blir så många att konversationen fungerar, så uppstår på något sätt även en förståelse av de kinesiska tecknens betydelser. Riktigt hur detta mirakel går till förklarar de dock inte.

Jag tycker det är betydligt rimligare att uppfatta situationen så att Searle-K visserligen "kan" kinesiska i den meningen att hen kan föra en konversation på kinesiska (som i ett Turingtest), men att hen ändå

inte förstår vad konversationen handlar om eller vad de kinesiska tecknen betyder.

Jag har i ett annat sammanhang (Bergström 2021) sagt att detta bevisas av att Searle-K inte kan översätta de kinesiska tecknen i regelboken till engelska. Detta argument underkänner Häggström (på Facebook) med motiveringen att cirka en miljard kineser ju förstår kinesiska, utan att kunna översätta det som sägs på kinesiska till engelska. Det har han förstås rätt i. Och han anser tydligen att Searle-K inte kan engelska. Men om Searle-K inte kan engelska, så har han ingen nytta av regelboken. Den är nämligen skriven på engelska, eftersom den ju ska kunna begripas av Searle, som bara kan engelska (jfr Häggström 2021, s. 232).

Häggström invänder kanske att det inte är Searle-K, utan Searle-E, som tillämpar regelboken. Och att Searle-K därför inte behöver kunna engelska. Det låter i och för sig rimligt. Men om man utesluter Searle-E och dennes kunskaper i engelska från Searle-K, så finns det väl inget annat kvar av Searle-K än själva regelboken (och förrådet av fraser). Därmed är hen inte en person över huvud taget. Och en regelbok kan väl inte förstå något alls, inte ens de regler den innehåller. Det skulle nog inte ens Häggström påstå. Men han tycks vara övertygad om att Searle-K är medveten och förstår kinesiska (s. 238–41).

Searle själv använder inte det översättningsargument jag här föreslår. Han säger bara att systemet (eller Searle-K) inte vet vad konversationen handlar om, trots att den flyter på utan problem. Searle-E vet vad engelska ord refererar till, men Searle-K vet inte vad de kinesiska tecknen refererar till (1981, s. 359).

Någon skulle kanske vilja invända mot översättningsargumentet att Searle-K skulle kunna gå ut på nätet och där hitta ett översättningsprogram, med vars hjälp hen kan översätta de kinesiska tecknen till engelska. Men då har vi ju lämnat Searles tankeexperiment. Med ett sådant kriterium på förståelse, så förstår ju även Searle själv kinesiska. Det kan inte gärna vara relevant.

Över huvud taget är det väldigt konstigt att påstå att det inom en och samma kropp, i Searles tankeexperiment, finns två olika personer. Är det inte betydligt mer korrekt att säga att det bara finns *en* person, nämligen Searle, som utan regelboken endast kan engelska, men som med hjälp av regelboken kan simulera en konversation på kinesiska – dock utan att förstå vad konversationen handlar om och vad de kinesiska tecknen betyder? Searle säger själv att ”så som situationen har beskrivits är det kinesiska delsystemet helt enkelt en avdelning av det engelska

delsystemet, en avdelning som ägnar sig åt meningslös manipulation av tecken i enlighet med regler på engelska” (1981, s. 360, min övers.).

Men Häggström anser tvärtom att vi ”utifrån kan observera” två olika personer som bebor Searles kropp (2021, s. 235). Jag tror inte att andra observatörer i allmänhet skulle hålla med om det. Vid första anblicken verkar det kanske som att Searle kan både engelska och kinesiska. Men när han sedan försäkrar att han inte förstår kinesiska, utan bara konverserar på kinesiska med hjälp av en regelbok, så skulle man väl godta detta. Man skulle revidera sitt första intryck, snarare än att tro att Searles kropp även rymmer en helt annan person, förutom Searle-E. Och om man å andra sidan tror att Searle ljuger och att han faktiskt förstår kinesiska, så skulle man nog fortfarande bara observera en person, nämligen Searle.

* * *

Att det i alla händelser finns *en* person, en medveten varelse, i det kinesiska rummet gör ju att det inte är ett exempel på artificiell intelligens. Hur kan det då användas som ett argument mot CTM, som implicerar att medvetande kan produceras digitalt i en dator? Ja, tanken är väl att Searle i det kinesiska rummet i princip kan ersättas med en icke-mänsklig mekanism som tillämpar regelboken – som i så fall får antas vara skriven på någon lämplig maskinkod. En sådan mekanism i kombination med regelboken och frasförrådet skulle då i någon mening förstå kinesiska, eller bete sig som om den gjorde det. Skulle den också vara medveten?

Den tes Alan Turing driver i sin berömda artikel från 1950 tycks vara att datorer är intelligenta och kan tänka i den mån de kan konversera som människor i Turingtest. Han bemöter diverse invändningar mot denna tes, bland annat invändningen att datorer inte är medvetna. De kan, enligt denna invändning, inte känna något, som t. ex. smärta, stolthet, ilska, depression osv.

Eftersom Turing tar upp och bemöter denna invändning, som rör medvetenhet, så menar han tydligen att datorer inte bara kan tänka, utan att de också kan vara medvetna. Kanske anser han att tänkande förutsätter medvetenhet – något som i och för sig kan låta helt rimligt.

Den som hävdar att maskiner inte är medvetna kan enligt Turing inte gärna godta den ”extrema” tesen att det enda sättet att *veta* att någon annan aktör – människa eller maskin – är medveten är att *vara* den aktören. Ty den tesen leder till solipsism, dvs. åsikten att endast man själv

är medveten, vilket är en orimlig åsikt. Så den som hävdar att maskiner inte är medvetna skulle säkerligen inte godta den ”extrema och solipsistiska” tesen, utan antagligen godta Turings test för att en maskin är medveten (jfr 1950, s. 60–61). Simsalabim!

Felet med Turings argument är att han inte skiljer mellan den ”extrema” tesen (som är rimlig) och solipsismen (som är orimlig). Att vi inte kan *meta* att någon annan är medveten, är ju fullt förenligt med att vi kan *tro* det. Vi tror att andra människor är medvetna, men vi tror inte att maskiner är medvetna. Och Turings test visar inte att det är något fel på den åsikten.

Men det framgår faktiskt inte helt klart om Turing verkligen tror att datorer kan vara medvetna, eller att de måste vara medvetna för att klara ett Turingtest. Searle förnekar för sin del inte att en maskin skulle kunna tänka. Han påpekar att han själv är en sorts maskin som kan tänka, men att han är inte en dator (s. 367). Poängen med Turings maskiner är däremot att de inte är människor.

Att vi inte kan *meta* att andra människor är medvetna, är i alla händelser fullt förenligt med att vi *tror* att de är det. Att vi inte vet att andra är medvetna tvingar oss alltså inte att godta solipsismen. Så Turings resonemang verkar på denna punkt helt misslyckat.¹

* * *

Jag har försökt visa att Häggströms argument mot Searles tankeexperiment inte är övertygande. Searles tankeexperiment tycks visa att en dator i princip skulle kunna konversera på kinesiska – och alltså klara ett Turingtest – utan att förstå kinesiska och utan att vara medveten. Däremot är detta inget argument mot CTM. Det är nämligen fullt förenligt med att en dator kan vara medveten.

En mänsklig programmerare kan knappast konstruera ett program som klarar ett utdraget Turingtest, men ett neuralt nätverk skulle kanske genom maskininlärning tillämpad på massor av konversationsdata kunna göra det. Och någon kunde kanske misstänka att en sådan inlärningsprocess på något sätt skulle dra med sig medvetande och förståelse som en sorts bieffekt.

Det verkar dock inte troligt att något sådant skulle kunna inträffa om inte datorn dessutom har någon sorts perceptionsapparat, som tillåter

1. I en redaktionell kommentar till denna text har dock Olle Risberg hävdat att det är något fel med att tro något som man tror att man inte vet. Det är jag inte beredd att hålla med om.

den att uppleva andra intryck från omvärlden. Det räcker då inte med kameror, mikrofoner och andra sensorer, som ersättning för sinnesorgan. Vad som krävs är dessutom sinnesintryck, och jag har väldigt svårt att tro att sådana skulle kunna åstadkommas av programmering eller maskininlärning. Häggströms initierade redogörelse för AI-forskning innehåller inte heller någon hypotes om hur verklig varseblivning – eller ens någon digital beräkningsstruktur med motsvarande funktion – skulle kunna uppstå i en dator. Själv lutar Häggström åt att datormedvetande är möjligt, men han medger också att frågan är ”fortsatt öppen” (s. 228 och 241).

De flesta av oss tror väl att medvetande är orsakat av fysikaliska processer, vi vet bara inte hur. Det är fortfarande ett mysterium. Kanske uppstår medvetandet som en bieffekt när vi lär oss tala och på annat sätt interagerar med vår omgivning. Vi kan i så fall knappast utesluta att maskininlärning i kombination med varseblivning skulle kunna ge samma effekt, alltså medvetande. Men vi kan aldrig veta om det är så, eftersom vi inte har något objektivt test på om medvetande föreligger. Turingtest duger exempelvis inte. Vi vet inte om det känns på något särskilt sätt att vara en dator som klarar ett Turingtest.

* * *

Men spelar det egentligen någon roll om en AGI är medveten eller inte för problemet om den kan vara farlig för oss människor? Det finns ju mycket som är farligt, utan att vara medvetet. Exempelvis massförstörelsevapen av olika slag.

De som tror att en AGI kan bli farlig uttrycker sig ofta som om den har fri vilja och som om den kan vilja göra något som kan skada oss. Det kan tyckas kräva medvetande. Häggström säger att man bland olika AI kan skilja mellan *agenter* och *verktyg* (s. 157–59). Man får intrycket att han tror att agenter kan bli farligare än verktyg.

Det är klart att det finns verktyg som inte är agenter, t. ex. hammare och bilar. Men när det gäller anordningar som betar sig på något lämpligt sätt för att uppnå ett givet mål är nog distinktionen svårare att upprätthålla. Är en termostat en agent eller ett verktyg? Är en självkörande bil, ett mästarprogram i schack, en autopilot i ett flygplan eller en självgående gräsklippare en agent eller ett verktyg?

Anhängare av Omohundro–Bostrom-teorin tycks anse att en AGI som har som slutmål att tillverka gem är en agent, och potentiellt en mycket farlig agent. Den kan nämligen få för sig att optimal gemproduktion

kräver att mänskligheten röjs ur vägen. Men jag har svårt att inse varför den inte lika gärna är ett verktyg. Kan det bero på om den är medveten eller inte? Eller på om den har en "egen vilja", vad nu det kan betyda?

En AGI som har som slutmål att tillverka gem är väl just ett verktyg för att tillverka gem – om någon annan har givit den detta slutmål. Har den däremot på egen hand bestämt slutmålet, så kan den betraktas som en agent. Men kan verkligen en AGI bestämma sitt eget slutmål? Såvitt jag förstår innebär Omohundro–Bostrom-teorin att AGI:n inte gör det, men att den givet sitt slutmål bestämmer sig för de mest effektiva medlen att uppnå det. Då är den ingen agent, utan ett verktyg – på samma sätt som t. ex. en autopilot eller en självgående gräsklippare.

I själva verket är det kanske bara verktyg som kan ha specifika slutmål. Häggström säger själv att han inte har något slutmål och han hänvisar också till Patrik Lindenfors, som tycks ha samma åsikt (s. 148). Jag är benägen att hålla med – men det hänger till stor del på om en aktör bara har ett slutmål, vilket framställningen hos Häggström och Bostrom ger vid handen, eller om en aktör kan ha flera slutmål (som i så fall på något hittills oförklarad sätt måste balanseras mot varandra). Men även om en AGI endast är ett verktyg, så kan den förstås ändå vara väldigt farlig. Farliga verktyg är ju ingen nyhet i en värld full av bilar, droger, kärnkraftverk och vapen av olika slag.

Lindenfors tycks för övrigt ha en uppfattning om "slutmål" enligt vilken "våra föregångare på savannen faktiskt hade ett, nämligen maximering av antalet avkomma" (s. 148). Men om slutmål tolkas på det sättet, har nog vem som helst slutmål, nämligen – förenklat uttryckt – en konsekvens av det man gör. En sådan konsekvens behöver ju inte vara avsedd eller ens medveten. Om slutmål i stället är något man avsiktligt eftersträvar, är det förstås ingalunda säkert att det också blir en konsekvens av det man gör.

Avsikter kan tyckas förutsätta medvetenhet, men det är nog inte nödvändigt. Flygande autonoma vapen har en sorts inbyggda avsikter, utan att – såvitt vi vet – vara medvetna. Huruvida Omohundro–Bostrom-teorin räknar med avsiktliga slutmål eller slutmål av konsekvenstyp har jag inte kunnat utläsa av Häggströms framställning. Och inte heller av Bostroms bok *Superintelligens*. Men jag antar att de ska vara avsiktliga. Eftersom en AGI är ett verktyg är det någon utomstående som tilldelar den ett slutmål. Slutmålet är förstås medvetet för den som inplanterar det i en AGI, men AGI själv behöver inte vara medveten om något alls.

För min del tror jag inte att människor har någon fri vilja. Jag tror att

det vi vill och gör bestäms av fysikaliska orsaker som i sin tur har fysikaliska orsaker och så vidare. Av liknande skäl tror jag inte heller att en AGI har någon fri vilja. Men jag utesluter inte att den kan vara medveten.

Frågan är då om dess farlighet är beroende av om den är medveten. Häggström citerar en författare som anser att om en AGI är farlig, så spelar det ingen roll ur just farlighetssynpunkt om den dessutom är medveten (s. 226). Det verkar riktigt. Men hur ska man se på risken för att en AGI blir farlig? Är den risken beroende av om den är medveten?

Om människor och andra kännande biologiska varelser utrotas av AGI och ersätts av datorer, så kan detta vara mindre katastrofalt om datorerna är medvetna än om de inte är det, nämligen om de skulle ha huvudsakligen behagliga upplevelser. Skulle datorerna däremot lida svåra plågor vore katastrofen å andra sidan större. Så kan man spekulera.

Men är själva *sannolikheten* för att en superintelligens ska utrota eller förslava mänskligheten beroende av om den är medveten? Varken Bostrom eller Häggström tycks ha någon särskild uppfattning om detta. Såvitt jag förstår har dock medvetenhet ingen som helst effekt i detta avseende, vare sig positiv eller negativ.

Detta är välkommet, eftersom vi nog aldrig kommer att ha anledning att tro att vissa datorer är medvetna. Det enda skäl vi har för att tro, eller kanske snarare orsaken till att vi tror, att andra människor – och vissa djur – är medvetna är att de liknar oss så mycket. Vi tycker inte att datorer liknar oss tillräckligt mycket. De liknar snarare andra maskiner och verktyg, som vi inte tillskriver medvetande.

Litteratur

- Bergström, Lars. 2016. Recension av Olle Häggström, *Here Be Dragons. Filosofisk tidskrift* 37, nr 4.
- Bergström, Lars. 2021. "Läsarbrev". *Sans* 10, nr 3, s. 97.
- Bostrom, Nick. 2017. *Superintelligens: Vägar, faror, strategier*. Stockholm: Fri Tanke.
- Hofstadter, Douglas R. och Daniel C. Dennett, red. 1981. *The Mind's I*. New York: Basic Books.
- Häggström, Olle. 2016. *Here Be Dragons: Science, Technology and the Future of Humanity*. Oxford: Oxford University Press.
- Häggström, Olle. 2021. *Tänkande maskiner: Den artificiella intelligensens genombrott*. Stockholm: Fri Tanke.
- Searle, John R. (1980) 1981. "Minds, Brains, and Programs". I Hofstadter och Dennett, s. 53–67.
- Turing, Alan. (1950) 1981. "Computing Machinery and Intelligence". I Hofstadter och Dennett, s. 353–73.

Problemet med kollektiv påverkan

En utvärdering av Julia Nefskys generella lösning

Andrea S. Asker

I. Problemet med kollektiv påverkan

Som individer ställs vi ofta inför handlingssituationer där ett kollektiv påverkar utfallet (KP-situationer). Det är situationer där människor tillsammans orsakar ett moraliskt betydelsefullt utfall, antingen ett fördelaktigt eller ett skadligt, genom att var och en agerar på ett visst sätt, men där varje individuell handling verkar göra mycket liten eller ingen skillnad. Utfallet tycks bli detsamma (eller åtminstone inte väsentligt annorlunda) oavsett om individen utför sin handling eller ej. Att en individuell handling gör skillnad med avseende på utfallet innebär i det här sammanhanget alltså att utfallet blir väsentligt annorlunda beroende på om den individuella handlingen utförs eller inte (Nefsky 2017, s. 2756).

Vanliga exempel på KP-situationer är klimatpåverkan, köttkonsumtion och röstning i allmänna val (Nefsky 2019; Parfit 1984, s. 198). När många människor flyger, kör bil och använder luftkonditionering bidrar det till klimatförändringar, vilket får omfattande skadliga konsekvenser. Men det tycks inte göra någon skillnad om jag flyger till Frankrike i sommar eller tar ett nöjesåk med bilen på söndagsmorgonen, eftersom klimatförändringarna kommer att vara desamma oavsett hur jag handlar (Nefsky 2019, s. 2). På liknande sätt bidrar utbredd köttkonsumtion till att djur far illa på fabriker, men givet köttindustrins storskalighet är det högst osannolikt att färre djur blir lidande till följd av att jag avstår från att köpa kött till middag i kväll (Nefsky 2019, s. 2). När det gäller röstning kan folket tillsammans rösta fram en bra eller dålig kandidat, vilket får betydande konsekvenser för samhället, men valets segrare blir densamma, oberoende av någon enskild röst (Nefsky 2019; Parfit 1984).

Ett vanligt förekommande resonemang i debatten om kollektiv påverkan är att om en individuell handling inte gör någon skillnad så är handlingen överflödigt och därmed finns det inga moraliska skäl att utföra den. Detta argument tycks felaktigt. Intuitivt tänker vi oss att det finns

moraliska skäl att utföra en handling i denna sorts situation (Nefsky 2017, s. 2744).¹ Däremot är det inte uppenbart vilka dessa skäl är, vilket har gett upphov till en omfattande debatt. Att identifiera moraliska skäl för individuellt agerande i KP-situationer har visat sig vara en stor utmaning. Vi kallar denna utmaning för ”problemet med kollektiv påverkan”. Det ska noteras att vi i detta sammanhang enbart diskuterar individers val vid enstaka tillfällen, och inte deras beteende över tid. Utmaningen är alltså att identifiera ett moraliskt skäl för en enstaka individuell handling, som att rösta i ett specifikt val, avstå från att köra bil till jobbet en morgon eller avstå från att köpa kyckling till middag en kväll.

Julia Nefsky har nyligen presenterat ett förslag till en generell lösning på problemet med kollektiv påverkan, ett förslag som syftar till att identifiera det viktigaste skälet för individuellt agerande i de flesta, om inte alla, KP-situationer. Nefskys lösningsförslag tar fasta på den intuitiva idén att en individuell handling kan bidra, om än bara lite grann, till att realisera ett gott utfall eller att förhindra ett skadligt. Hon menar att denna möjlighet att bidra utgör individens primära skäl för att handla i KP-situationer. Vi kommer här att undersöka Nefskys förslag och utvärdera huruvida det ger en generell lösning på problemet med kollektiv påverkan. Det visar sig att det finns tre centrala problem som påverkar lösningens generaliserbarhet. Dessa problem har att göra med (1) när en handling kan bidra, (2) vilka handlingar som kan bidra och (3) styrkan hos de skäl som möjligheten att bidra ger upphov till. Slutsatsen är att Nefskys förslag inte utgör någon generell lösning på problemet med kollektiv påverkan. Dessutom är det oklart om en generell lösning är rimlig eller ens önskvärd, vilket indikerar att en pluralistisk ansats kan vara att föredra.

II. Nefskys lösning – Att bidra utan att göra någon skillnad

Nefsky börjar med att presentera två viktiga påståenden om problemet med kollektiv påverkan. Det första är att problemet vilar på antagandet att om en handling inte gör någon skillnad för ett utfall så är handlingen kausalt överflödigt med avseende på utfallet. Det andra är att vi måste överge detta antagande för att lösa problemet. Det viktigaste påståendet i Nefskys lösning är att en individuell handling kan bidra till att realisera ett utfall utan att göra någon skillnad där ”att bidra till” definieras

1. Det kan förstås finnas skäl för *eller emot* individuella handlingar beroende på sammanhanget. För enkelhetens skull pratar vi oftast om skäl för en individuell handling, där en individuell handling kan vara att göra något eller att avstå från att göra något.

som att göra ett icke överflödigt kausalt bidrag till utfallet. Det är denna möjlighet att bidra som utgör individens primära skäl för att handla i KP-situationer, enligt Nefsky (2017, s. 2746).

Grundläggande för Nefskys lösning är förhållandet mellan tre begrepp som beskriver de olika sätt på vilka en handling kan stå i relation till ett utfall: att vara en del av orsaken till utfallet, att göra skillnad med avseende på utfallet och att bidra till att åstadkomma utfallet. Enligt Nefsky kan en handling vara en del av orsaken till ett utfall utan att göra någon skillnad för utfallet. Hon anser att detta är något som vi måste anta för att lösa problemet och kallar detta antagande för *den kausala förutsättningen*. Själva problemet är ju att människor som agerar kollektivt orsakar ett utfall tillsammans, även om ingen enskild handling tycks göra någon skillnad för utfallet, så om vi inte trodde att en individuell handling skulle kunna vara en del av orsaken utan att göra någon skillnad så skulle en sådan situation inte ge upphov till problemet med kollektiv påverkan (Nefsky 2017, s. 2750–51).

En handling kan även vara en del av orsaken till ett utfall utan att bidra till det, det vill säga att handlingen kan vara en del av orsaken till ett utfall utan att göra ett icke överflödigt kausalt bidrag till utfallet (Nefsky 2017, s. 2751). För att förstå detta, ta följande exempel: en stor grupp människor befinner sig på ett sjunkande skepp. De är utrustade med hinkar och försöker skopa ut tillräckligt med vatten innan det är för sent att rädda skeppet från att sjunka. Emellertid finns det pumpar ombord som kommer att sätta igång och dränera tillräckligt mycket vatten innan det är för sent. En individuell handling kan då vara en del av orsaken till att skeppet inte sjunker, om människorna lyckas skopa ut tillräckligt mycket vatten i tid, men eftersom pumparna skulle ha satt igång i vilket fall som helst så var utfallet redan garanterat och därmed var den individuella handlingen överflödig. En handling kan alltså vara en del av orsaken till ett utfall utan att vare sig göra någon skillnad för eller bidra till utfallet. Slutligen menar Nefsky att om en handling gör skillnad med avseende på utfallet så bidrar den också till att åstadkomma det, men det omvända gäller inte – en handling kan bidra till att åstadkomma ett utfall utan att göra någon skillnad för det (Nefsky 2017, s. 2746).

Emellertid kan en handling bidra till ett utfall endast under vissa förutsättningar. Enligt Nefsky kan en individuell handling av en viss sort bidra till att åstadkomma ett utfall utan att göra någon skillnad för det om och endast om det vid tidpunkten för handlingens utförande är obestämt huruvida utfallet kommer att inträffa eller ej, samt att det är

möjligt vid denna tidpunkt att utfallet inte inträffar, åtminstone delvis på grund av ett otillräckligt antal handlingar av samma sort (Nefsky 2017, s. 2753). Ta röstning i ett allmänt val som exempel. Säg att valet ännu inte är avgjort och att det är möjligt att ett otillräckligt antal röster på en viss kandidat kommer att innebära att denne förlorar. Under dessa omständigheter kan en individuell röst bidra till att åstadkomma seger. Mer formellt uttryckt: en individuell handling av en viss sort – en individuell X-handling – kan bidra till att realisera utfall Y i en KP-situation om och endast om följande två villkor är uppfyllda:

- (1) Det är möjligt att Y inte kommer att inträffa åtminstone delvis på grund av ett otillräckligt antal X-handlingar.
- (2) Det är möjligt att Y kommer att inträffa åtminstone delvis på grund av ett tillräckligt antal X-handlingar.²

När bägge villkor är uppfyllda är den individuella handlingen inte överflödigt och kan därmed bidra till att åstadkomma utfallet. För att se hur Nefskys lösningsförslag kan appliceras ska vi titta på ett exempel som ursprungligen presenterades av Derek Parfit (1984) och som Nefsky använder genomgående i sin diskussion av kollektiv påverkan:

Vattendroppar. En stor grupp människor, låt oss säga 10 000, är strandade i öknen och lider av svår törst. En lika stor grupp står i utkanten av öknen med en halvliter vatten vardera. Varje person i utkanten av öknen kan välja att donera sin halvliter genom att hålla vattnet i en vagn. Vagnen kommer sedan köras ut i öknen och allt vatten kommer att delas lika mellan de törstande. För varje halvliter som doneras får varje person i öknen ytterligare en droppe vatten (en tiotusendels halvliter), så lite att det inte märks. Därmed verkar en individuell donation inte göra någon skillnad med avseende på det relevanta utfallet, nämligen att minska människornas lidande. Människorna kommer varken att lida mer eller mindre på grund av någon enskild individs donation. (Exemplet hämtat från Nefsky 2017, s. 2743–44 och Parfit 1984, s. 76.)

Enligt Nefskys teori kan en enskild donation bidra till att lindra människornas lidande om och endast om det vid tidpunkten för donationen är möjligt både att lidandet lindras, åtminstone delvis på grund ett till-

2. Nefsky uttrycker inte villkor 2 explicit men hon stipulerar det implicit (Nefsky 2017, s. 2754, not 23).

räckligt antal vattendonationer, och att lidandet inte lindras, åtminstone delvis på grund ett otillräckligt antal vattendonationer. Utifrån detta menar Nefsky att så länge det är obestämt huruvida människornas lidande kommer att lindras (eller huruvida det kommer att lindras till en högre grad) så finns det ett moraliskt skäl för individen att donera sin halvliter, trots att en enskild donation inte kommer att göra någon skillnad för huruvida eller till vilken grad lidandet lindras.

Det bör poängteras att när villkoren för att kunna bidra är uppfyllda *är* den individuella handlingen inte kausalt överflödigt och *kan* därmed bidra till utfallet, men huruvida handlingen faktiskt bidrar till utfallet beror på vad som faktiskt händer. En utförd handling bidrar inte till utfallet om utfallet ifråga inte inträffar eller om handlingen faktiskt inte är en del av orsaken till utfallet (Nefsky 2017, s. 2754). I *Vattendroppar* kan det till exempel hända att människornas lidande inte lindras på grund av för få vattendonationer eller på grund av att en sandstorm förhindrar leveransen av vattnet. Det skulle också kunna vara så att lidandet lindras av något helt annat än vattendonationer, till exempel genom att de strandade människorna snabbt evakueras. Emellertid är det den individuella handlingens *potential* att bidra som är skälgivande, enligt Nefskys lösningsförslag.³

III. När kan en individuell handling bidra?

En invändning som kan framföras mot Nefskys teori är följande: beakta återigen fallet *Vattendroppar*. Anta att 9999 halvlitrar redan har hållits i vagnen som ska köras ut i öknen. Kan jag då bidra genom att donera min halvliter? Det kan jag inte, för vid tidpunkten för donationen finns det ingen möjlighet att människornas lidande inte kommer att lindras på grund av ett otillräckligt antal vattendonationer, såvida inte en enstaka halvliter kommer att göra skillnad för människornas lidande, vilket vi har antagit att den inte gör. Om vi räknar baklänges från min handling kan det hävdas att eftersom en enstaka halvliter inte kan göra skillnad för människornas lidande, enligt vad vi har antagit, så kan personen som överväger att donera sin halvliter före mig inte heller bidra med något och så vidare.⁴ I så fall verkar det som om ingen kan bidra.

3. Nefsky har poängterat att det som är skälgivande är att handlingen inte är överflödigt och att den gör framsteg mot det önskade utfallet, inte att den faktiskt hjälper (J. Nefsky, maj 2020, personlig kommunikation).

4. Magnus Jedenheim-Edling framförde denna invändning under en presentation av Nefskys lösningsförslag vid det högre seminariet i praktisk filosofi vid Uppsala universitet den 15 maj 2020.

Detta resonemang tycks dock vila på antagandet att en handling inte kan bidra till ett utfall utan att göra någon skillnad, vilket ju Nefsky avvisar. Anta istället att 5 000 halvlitrar redan har donerats när jag ska hålla mitt vatten i vagnen. Om inga fler donationer görs så kommer människorna i öknen att få en fjärdedels liter vatten var, vilket skulle lindra deras lidande något, även om deras lidande förmodligen skulle lindras betydligt mer om de fick en halvliter var. Om det vid denna tidpunkt inte är avgjort hur många som kommer att donera sina halvlitrar, så är det fortfarande möjligt att människorna lidande lindras i mindre grad, åtminstone delvis på grund av ett otillräckligt antal vattendonationer. Min individuella donation åstadkommer inte någon ytterligare lindring, men det faktum att ett otillräckligt antal av samma slags handling kan vara det som förhindrar ytterligare lindring gör att det finns ett moraliskt skäl för mig att utföra handlingen, enligt Nefsky.

Emellertid kan det uppstå problem med att avgöra när omständigheterna är sådana att en individuell handling kan bidra till ett utfall. Hur kan vi avgöra när någon grad av lindring redan är garanterad på så sätt att det inte längre är möjligt att lindring inte uppnås på grund av att vattendonationerna är för få? Nefsky skulle förmodligen mena att ju fler som håller i sitt vatten, desto närmare kommer man till garanterad maximal lindring (dvs. den lindring som kan uppnås med 10 000 halvlitrar). Fast vi kan knappast säga att det inte finns någon sådan garanti förrän alla har hållt i sitt vatten och vagnen är full, för om vi säger det när 9 999 halvlitrar har hållits i, är det samma sak som att säga att ytterligare en halvliter skulle kunna göra skillnad för graden lindring. Detta problem kvarstår när vi räknar bakåt från 9 999 halvlitrar i och med att vi inte kan avgöra de exakta förhållandena (med avseende på antalet halvlitrar som donerats) där handlingen kan bidra, utan att hänvisa till chansen att handlingen kommer att göra skillnad.

Nefsky kan bemöta denna invändning genom att hävda att vi inte behöver känna till de exakta förhållandena för att en individuell handling ska kunna bidra för att ha ett moraliskt skäl att agera. Hon menar att när villkoren för att bidra inte är uppfyllda har individen inget *objektivt* skäl att agera, men så länge individen inte vet huruvida villkoren är uppfyllda har hon ändå ett *subjektivt* skäl att agera, baserat på att hennes handling, så vitt hon vet, skulle kunna bidra, även om den faktiskt inte kan det (Nefsky 2017, s. 275). Här blir det relevant att fråga sig vilken definition av begreppet "möjlighet" som Nefsky använder. Nefsky menar att begreppet avser "standardtolkningen av möjlighetsbegreppet som

människor använder för praktiska överväganden” (Nefsky 2017, s. 2760, min övers.). Det rör sig alltså inte om metafysisk eller logisk möjlighet. I övrigt lämnar hon definitionen öppen, så när som på två begränsningar: (1) det som faktiskt är möjligt kan skilja sig från vad individen har skäl att tro är möjligt och (2) möjlighetsbegreppet är inte deterministiskt, vilket innebär att vad individer kommer att göra i framtiden inte är förbestämt av det förflutna och naturlagarna (Nefsky 2017, s. 2761–62). Denna definition av ”möjlighet” tillåter subjektiva skäl på basis av vad individen tror är möjligt.

Emellertid får vi problem om vi antar att vilka trosföreställningar som helst kan ge upphov till de subjektiva skälen ifråga. I så fall skulle uppenbart falska eller motstridiga trosföreställningar om vad en enskild handling kan åstadkomma kunna ge en individ skäl att agera. Därför kan det vara relevant att göra vissa avgränsningar, exempelvis genom att ta hänsyn till vad individen har för skäl att tro som hon gör. I så fall är det viktigt att hålla isär huruvida individens subjektiva skäl avgörs av vad individen *faktiskt tror* eller av *vad hon har goda skäl att tro*. Nefsky verkar mena att det är de trosföreställningar som individen faktiskt har som avgör hennes subjektiva skäl, men hon tycks också anse att det spelar roll för individens subjektiva skäl vad hon har goda skäl att tro (Nefsky 2017, s. 2761), även om det är oklart vilken roll. Eventuellt tänker hon sig att av de trosföreställningar som individen har, så är det endast de som hon också har goda skäl att ha som ger upphov till subjektiva skäl. Men Nefsky säger ingenting explicit om en sådan begränsning.

Det är inte heller uppenbart när individen har goda skäl att tro att villkoren för att bidra är uppfyllda, av samma anledning som det inte är uppenbart när villkoren faktiskt är uppfyllda. Återigen tycks vi inte kunna specificera detta utan att falla tillbaka på idén att en individuell handling kan göra skillnad. Det verkar som om vi måste förlita oss på en vag uppfattning om de specifika förhållanden som måste råda för att individen ska ha goda skäl att tro att villkoren för att bidra är uppfyllda, trots att hon är osäker på det. Detta gör Nefskys ansats otydlig, åtminstone i vissa KP-situationer, vilket kan göra hennes lösning mindre generaliserbar.

Det här problemet behöver dock inte utgöra ett stort hinder för Nefskys teori i praktiken. Det är framför allt om vi antar att individer agerar var och en efter varandra i en gemensam, observerbar process, med en tydlig början och ett tydligt slut, som problemet med att identifiera de rätta förhållandena uppstår. Om vi inte antar en sådan process

blir det enklare att föreställa sig de förhållanden som Nefsky verkar ha i åtanke när hon menar att individen åtminstone har subjektiva skäl att agera. I *Vattendroppar* kan vi tänka oss att dessa förhållanden råder när en individ som kan donera vatten inte vet huruvida det redan är avgjort till vilken grad människornas lidande kommer lindras, och hon också kan se att det är många som inte har hållt i sitt vatten, vilket ger henne skäl att tro att utfallet fortfarande är obestämt, att ett (o-)tillräckligt antal donationer kan avgöra utfallet och att hennes donation därmed kan bidra.

I de flesta verkliga KP-situationer är det ännu mer intuitivt att föreställa sig dessa förhållanden. I handlingssituationer som har att göra med exempelvis välgörenhetsdonationer, klimatförändringar och konsumtion kan individen inte veta till vilken grad utfallet är obestämt och därmed har hon i de flesta fall subjektiva skäl för att agera. Dessutom utspelar sig många KP-situationer under mycket lång tid, utan något tydligt slut, och utfallen realiserar sig stegvis. Fram till dess att det inte längre finns någon fattigdom, klimatkris eller djurplågeri kan det anses obestämt huruvida och till vilken grad det kommer att ske en förbättring. I så fall är villkoren för att kunna bidra faktiskt uppfyllda och därmed finns det objektiva skäl för individuell handling i dessa situationer, enligt Nefskys teori. Även i situationer där det ser ut som om utfallet är så gott som avgjort är det inte omöjligt att utfallet blir ett annat. I ett politiskt val där förundersökningar visar ett övervägande stöd för en kandidat är det exempelvis inte omöjligt att folk ändrar sig eller att fler människor än förväntat röstar. Därmed är det fortfarande *möjligt* att det som hindrar en annan kandidat från att vinna är ett otillräckligt antal röster och i så fall kan en enskild röst bidra till den andra kandidatens seger (Nefsky 2017, s. 2762–63).

IV. Vilka handlingar kan bidra?

Ett annat problem för Nefskys lösning har att göra med begreppet *X-handling* och rollen som det spelar i Nefskys teori. Begreppet betecknar en handling av en viss sort X som individen överväger att utföra i en given KP-situation (Nefsky 2012, s. 79), men det är oklart hur en handlingskategori X ska definieras och avgränsas: hur ska vi avgöra vilka handlingar som ska räknas som X-handlingar i olika KP-situationer? Frågan är central eftersom en individuell handlings förmåga att bidra och därmed våra skäl för individuellt handlande, enligt Nefskys teori,

hänger på huruvida det är möjligt att det önskade utfallet (inte) inträffar åtminstone delvis på grund av ett (o-)tillräckligt antal av *samma sorts handling* (Nefsky 2017, s. 2753). Att en individuell handling är av den relevanta sorten är därmed en del av det som är skälgivande. Därför är det viktigt att vi förstår vad som räknas som en X-handling i olika KP-situationer.

Vi kan anta att en X-handling måste vara sådan att den inte förstör strukturen av en KP-situation, i vilken ett kollektiv handlar och där ingen enskild handling tycks göra någon skillnad. Handlingar som uppenbarligen gör skillnad kan rimligen inte räknas som X-handlingar. Ta fallet *Vattendroppar* igen. Vi antog att en möjlig individuell handling var att hålla *en* halvliter vatten i vagnen. Nefsky å andra sidan talar löst om "att hålla i vatten" och "vattendonationer" (Nefsky 2017, s. 2753), fastän man knappast kan räkna en donation på flera tusen liter vatten åt gången till samma handlingskategori som en donation på en halvliter. Uppenbarligen är det möjligt att det som förhindrar att människors törst lindras delvis är ett otillräckligt antal stora vattendonationer, men detta ger knappast något skäl att donera en halvliter, eftersom detta inte visar att en enskild halvlitersdonation är den typ av handling som kan bidra till att åstadkomma det önskade utfallet. Problemet med kollektiv påverkan är ju att det inte tycks finnas något skäl att utföra en X-handling, men om X-handlingen är att på en och samma gång hålla flera tusen liter vatten i vagnen råder det inget tvivel om att det finns skäl för den, eftersom en sådan handling uppenbarligen skulle minska människornas lidande väsentligt.

Således bör vi i fallet *Vattendroppar* anta att en X-handling är en halvlitersdonation, vilket är ett litet bidrag som inte tycks göra någon skillnad för utfallet. Situationer som liknar *Vattendroppar* är enkla att förstå eftersom det är tydligt vilken sorts handling som individen bör överväga att utföra och varje handling av samma sort åstadkommer samma sak, det vill säga att ytterligare en halvliter hamnar i vagnen (allt annat lika). Samma sak gäller i röstningsfallet: varje person står inför valet att rösta eller inte rösta och varje röstningshandling åstadkommer samma sak, nämligen att ytterligare en röst tillfaller den kandidat som personen röstar på.

I andra KP-situationer är det mindre tydligt vad en X-handling är och ibland skiljer sig de enskilda handlingarna åt så till den grad att det inte finns någon naturlig handlingskategori X dit de olika handlingarna rimligen hör. I fallet med klimatförändringar nämner Nefsky

(2019, s. 1, 9) olika handlingar som minskar eller undviker utsläpp som exempel på X-handlingar, såsom att avstå från en biltur, att avstå från en flygresa eller att undvika att sätta på luftkonditioneringen. Dessa handlingar skiljer sig från varandra, inte bara i sin karaktär utan också i form av vad de åstadkommer, eftersom vissa handlingar kan förhindra större utsläpp än andra. Om till exempel en biltur till jobbet producerar mindre utsläpp än en timmes luftkonditionering, kan man åstadkomma en större utsläppsreduktion genom att avstå från den senare snarare än den förra handlingen.

Det kan vara svårt att avgöra vilka av dessa handlingar som ska grupperas tillsammans i en handlingskategori X och vilka som inte platsar för att de åstadkommer för mycket eller för lite. Att hålla andan i 30 sekunder kan innebära en minimal minskning av mängden koldioxid i luften (åtminstone för ett ögonblick), men vi skulle knappast placera en sådan handling i samma kategori som att avstå från en timmes biltur, och sedan hävda att underlåtet av denna sorts handling bidrar till klimatförändringar. Inom ramen för Nefskys teori skulle ett sådant resonemang leda till det besynnerliga resultatet att en individ har ett moraliskt skäl att hålla andan i 30 sekunder för klimatets skull. Inte heller skulle vi placera handlingen att avstå från en timmes biltur i samma kategori som att plantera flera kvadratkilometer stora skogar, eftersom skogsplantering i stor skala kan förväntas åstadkomma en väsentlig minskning av mängden koldioxid i atmosfären (åtminstone på lång sikt) och därmed potentiellt mildra klimatförändringarna. Däremot är det oklart var gränsen går, i synnerhet som det råder stor osäkerhet kring effekterna av olika åtgärder för klimatet. Det finns också handlingar som inte reducerar individens egna utsläpp alls, men som ändå genererar en förväntad reduktion, till exempel politisk aktivism och lobbyverksamhet – ska sådana handlingar räknas som X-handlingar i fallet med klimatförändringar? Om vi gör en välvillig tolkning kan vi eventuellt på basis av intuition identifiera en lagom bred handlingskategori där de flesta utsläppsreducerande åtgärder som individer kan vidta passar in, men det råder fortfarande osäkerhet och brist på precision kring vilka handlingar som ska räknas som X-handlingar i denna typ av KP-situation. Detta kan försvåra tillämpningen av Nefskys lösningsförslag, vilket kan ge ytterligare problem för dess generaliserbarhet.

V. Skälens styrka

Det mest påtagliga och kanske mest intuitiva problemet för Nefskys lösningsförslag har att göra med styrkan hos de skäl för individuell handling som förslaget identifierar. Nefsky pekar på tre variabler som kan påverka styrkan hos dessa skäl: (1) betydelsen av det aktuella utfallet, (2) den individuella handlingens potentiella kausala bidrag och (3) hur pass obestämt utfallet är, det vill säga hur nära 50 procent sannolikt det är att utfallet inträffar (Nefsky 2017, s. 2763). Den tredje variabeln kan förstås i termer av hur pass obestämt det är huruvida tillräckligt många X-handlingar kommer att utföras för att åstadkomma utfallet, vilket i sin tur avgör hur stor sannolikheten är för att den individuella handlingen ifråga faktiskt kommer att bidra. Beakta återigen fallet *Vattendroppar*. Om en individ överväger att donera sin halvliter, men det är osannolikt att de andra individerna kommer att göra det (till exempel för att de är giriga), så är det osannolikt att de törstandes lidande kommer att lindras alls. I ett sådant scenario är en individuell handling inte överflödigt och kan göra ett bidrag, men sannolikheten för att den faktiskt kommer att göra det är låg, vilket innebär att det moraliska skälet för individen att donera sin halvliter är betydligt svagare än vad det hade varit om utfallet hade varit mer obestämt (Nefsky 2017, s. 2762–63).

Det kan hända att dessa variabler, i synnerhet (2) och (3), kan göra skälet för individuell handling så svagt att det lätt vägs upp av andra faktorer. Om det kausala bidraget som en enskild handling kan göra är mycket litet och sannolikheten för att handlingen faktiskt bidrar är låg, kommer det förväntade värdet av att utföra handlingen inte att ge något avgörande skäl för att utföra den. Detta problem lyfter Nefsky själv fram gentemot de lösningsförslag som bygger på idén att varje individuell handling har en liten chans att göra en väsentlig skillnad för utfallet och att det förväntade värdet av dessa handlingar därmed ger oss skäl att agera (Nefsky 2012, avs. 3.3; 2017, s. 2747–48). Nefsky ser dock ingen anledning att tro att detta skulle bli ett problem för hennes egen teori eftersom hon inte tror att andra faktorer, såsom individens preferenser, kan väga upp möjligheten att göra något som inte är överflödigt, om än bara lite, för att åstadkomma det önskade utfallet i en KP-situation (Nefsky 2017, s. 2765).

Nefsky underskattar dock hur variablerna (2) och (3) kan påverka skälens styrka och vad detta innebär för lösningens generaliserbarhet. Notera först att variablerna (2) och (3) kan uttryckas i termer av relationen mellan det kollektiva handlande som består av X-handlingar och

de moraliskt betydelsefulla utfallen. Villkoren för att bidra kräver, som nämnt, att det ska finnas minst en väg till utfallet som inkluderar X-handlingar, dvs. en väg där det som kan åstadkomma utfallet är, åtminstone delvis, ett tillräckligt antal X-handlingar. Vi säger "minst en väg" eftersom det ofta finns flera olika vägar till utfallet. Formuleringen "åtminstone delvis" är också viktig eftersom det finns många olika faktorer, förutom X-handlingar, som behövs för att utfallet ska inträffa (Nefsky 2017, s. 2754, not 23).

Om det önskade utfallet är nedstängningen av en kycklingfabrik där djuren far illa, skulle en möjlig väg dit kunna vara lagstiftning som förbjuder storskalig kycklingproduktion. En stor försäljningsminskning på grund av att många konsumenter avstår från köp är en annan väg. Det senare alternativet innebär att ett tillräckligt antal individuella handlingar utgör en del av det som kan åstadkomma det önskade utfallet, i kombination med andra faktorer, som att köttproducenterna inte hittar nya marknader att sälja till och inte anammar billigare sätt att producera för att kompensera sin minskade försäljning. Om utfallet inte inträffar så kommer det således att bero delvis på ett otillräckligt antal X-handlingar, delvis på andra faktorer och delvis på att man inte lyckats åstadkomma utfallet på något annat sätt (Nefsky 2017, s. 2754, not 23). Detta tyder på att storleken på det kausala bidrag som en individuell handling kan åstadkomma beror på hur stor kausal roll det kollektiva handlandet (som ju består av individuella X-handlingar) spelar för utfallet jämfört med andra faktorer som också kan vara involverade. Dessutom varierar sannolikheten för att ett tillräckligt antal X-handlingar kommer att utföras, och därmed varierar även sannolikheten för att en individuell handling faktiskt kommer att bidra till ett specifikt utfall.

I vissa KP-situationer tycks det finnas en väg till det aktuella utfallet som är mer sannolik än någon annan väg, nämligen den väg som inkluderar X-handlingar. Dessutom verkar X-handlingarna spela den viktigaste och mest avgörande rollen i utfallets realiserande, jämfört med andra faktorer. Det mest uppenbara exemplet på en sådan situation är röstning. Låt oss föreställa oss ett politiskt val med två kandidater, A och B, där det önskade utfallet är att den maktfullkomlige extremisten B förlorar (Nefsky 2017, s. 2754). I denna situation kan det finnas flera sätt att förhindra B:s seger: B kan till exempel bli mördad (Nefsky 2017, s. 2754). Hursomhelst kan vi anta att det mest troliga sättet att förhindra B:s seger är att tillräckligt många röstar på A så att A vinner och B förlorar. Andra faktorer kan förstås spela in även här,

exempelvis kan korrupktion eller valbedrägeri påverka resultatet, men allt annat lika verkar det som om röstning kommer att avgöra utfallet, vilket i sin tur avgör i vilken grad den enskilda X-handlingen kausalt bidrar till utfallet.

I andra KP-situationer verkar det däremot finnas flera rimliga vägar till det aktuella utfallet och bland dessa är det bara vissa som involverar X-handlingar. Och ibland kan det dessutom vara så att andra faktorer spelar en större, mer avgörande roll för utfallet, även när X-handlingar är involverade. Klimatförändringar och köttkonsumtion är exempel på fall som har denna struktur. Klimatförändringar kan hejdas genom exempelvis överstatliga regelverk, internationellt samarbete och teknologisk innovation. Förbättringar kan också uppnås genom att individer gör omfattande förändringar av sin livsstil. Djurs lidande kan avhjälpas genom exempelvis automatisering, nya handelsavtal och formella sanktioner mot oetisk produktion, alternativt kan industrier tvingas förändra sin praxis eller stänga på grund av att människor förändrar sin konsumtion.⁵

I verkligheten lär det dock vara en kombination av dessa olika vägar och faktorer som bidrar till utfallen ifråga. I så fall kan vi inte på ett meningsfullt sätt skilja mellan de vägar till utfallen som involverar X-handlingar och de som inte gör det i dessa KP-situationer. I vilket fall som helst lär en enskild X-handlings kausala bidrag vara mycket litet, eftersom andra faktorer än individers kollektiva handlande, såsom omfattande strukturella förändringar, kan ha mycket större inverkan på utfall, som mildare klimatförändringar och minskat lidande i köttindustrier.

Vi har också anledning att tro att det är mindre sannolikt att tillräckligt många X-handlingar kommer att utföras för att åstadkomma förändring i dessa situationer. För det första skulle det krävas ett oerhört stort antal individuella handlingar för att åstadkomma ett utfall, såsom en substantiell minskning av klimatförändringarna, som ens kan jämföras med vad som kan åstadkommas exempelvis genom den slags innovation som kan göra stora industrier, som cement- och stålproduktion, klimatneutrala (Thornéus 2019). Dessutom kan enskilda X-handlingar i dessa KP-situationer ha vissa negativa sidoeffekter. Vid de

5. Kom ihåg att vi är intresserade av individers val vid enstaka tillfällen, även om ett stort antal enskilda handlingar förstås kan utgöra en generell förändring av människors livsstil och konsumtionsmönster. Det är också värt att förtydliga att det inte behöver finnas ett tillräckligt stort antal olika personer som utför en X-handling vardera. Det viktiga är att det finns tillräckligt många X-handlingar för att uppnå ett visst utfall.

flesta enskilda beslut vägs enbart mindre olägenheter mot möjligheten att bidra till något önskvärt utfall, men i vissa KP-situationer får man räkna med att ens handling också kan bidra till skadliga utfall. Exempelvis kan en omfattande bojkott av vissa industrier leda till arbetslöshet och andra samhällsekonomiska problem, åtminstone på kort sikt. Det tycks rimligt, även om det är spekulativt, att individer kan ha vissa förbehåll när det gäller att utföra sådana handlingar. Sådana avvägningar mellan fördelaktiga och skadliga utfall kan vara problematiska, åtminstone i vissa fall.

Med detta i åtanke kan vi se hur de moraliska skäl som Nefskys ansats identifierar kan försvagas avsevärt. Ta exemplet med enskild köttkonsumtion: om det kollektiva handlandet som består i att avstå från att köpa kött utgör ett av många sätt att minska djurs lidande och det finns andra, strukturella faktorer som kan vara avgörande för huruvida utfallet realiserar och det är osannolikt att ett tillräckligt antal X-handlingar kommer att utföras, så grundar sig det moraliska skälet för en enskild handling på att det finns en liten chans att handlingen kan utgöra ett litet kausalt bidrag till att minska djurs lidande. Ett sådant skäl är mycket svagt och kan lätt vägas upp av andra.

Detta tyder på att de skäl för att utföra en individuell handling som Nefskys lösning identifierar inte lever upp till vad förslaget utlovar. Nefskys teori ska ju identifiera våra primära, mest centrala skäl för individuellt handlande i de allra flesta KP-situationer, men om skälen som lyfts fram av teorin i vissa fall är mycket svaga, kan de knappast fylla denna funktion. Det vore mer lämpligt att utvärdera Nefskys ansats i varje enskild KP-situation för att se om den kan identifiera något skäl som är tillräckligt starkt. Kanske skulle problemet delvis kunna lösas genom att än en gång hänvisa till subjektiva skäl: om individen inte vet eller inte kan veta hur stort hennes potentiella kausala bidrag är, och inte heller kan veta hur stor sannolikheten är för att hennes handling faktiskt kommer att bidra till ett utfall, så kan hon fortfarande ha ett starkt subjektivt skäl att agera, baserat på trosföreställningen att hennes handling kan göra ett betydelsefullt kausalt bidrag. Dock är det inte säkert att individen har den typ av trosföreställningar som genererar starka subjektiva skäl. Kanske är individens intuition snarare att den enskilda handlingen inte kommer att göra någon skillnad, att det är osannolikt att den alls kommer att bidra med något, samt att även om den bidrog med något, skulle det röra sig om ett mycket litet bidrag. Baserat på diskussionen i det här avsnittet kan individen dessutom anses

ha goda skäl för en sådan uppfattning. Dessa överväganden försvagar avsevärt individens eventuella subjektiva skäl att agera. Vi ser därmed att både de objektiva och subjektiva skäl som Nefskys lösning identifierar inte håller vad förslaget lovar.

VI. Slutsats – En generell lösning vs. pluralism

Utifrån den diskussion som förts här kan vi dra slutsatsen att Nefskys förslag inte utgör en generell lösning på problemet med kollektiv påverkan. Hennes lösningsförslag utvecklar och preciserar den intuitiva idén att en enskild handling kan bidra till ett utfall och i många fall kan ansatsen identifiera ett moraliskt skäl för individuellt handlande. Men när vi undersöker saken närmre ser vi att detta skäl inte alltid är tillräckligt starkt. Nefskys ansats har uppkommit i en djupgående debatt som har präglats av ambitionen att hitta en generell lösning på problemet med kollektiv påverkan, men hittills tycks inget förslag kunna generaliseras till alla olika KP-situationer. Detta kan ses som ett tecken på att en generell lösning varken är rimlig eller önskvärd och att det därmed finns anledning att anamma en pluralistisk lösning istället. I så fall kan möjligheten att bidra ses som ett skäl bland flera som tillsammans kan förklara varför individer bör handla på ett visst sätt i en given KP-situation. Dessa skäl kan i sin tur variera från fall till fall. Vi skulle exempelvis kunna tänka oss att det i röstningsfallet finns skäl att delta i ett gemensamt projekt, både för att välja en god politisk kandidat och för att det finns en kollektiv plikt att upprätthålla demokratin. När det gäller konsumtion kan det finnas skäl för en individ att anpassa sina handlingar, inte för att det direkt påverkar produktionen på något betydelsefullt sätt, utan för att det signalerar ett missnöje som skulle kunna påverka andra att åstadkomma förändring. I mindre sammanhang som präglas av lojalitet och återkommande interaktioner, som när människor i ett bostadsområde värnar om en gemensam park genom att nyttja området med måtta, skulle det viktigaste skälet istället kunna vara respekt för det gemensamma. Styrkan och legitimiteten hos dessa skäl kan förstås också ifrågasättas, men syftet är att låta dem belysa ett alternativt förhållningssätt, enligt vilket man inte bör leta efter en enda sorts skäl som ska lösa problemet med kollektiv påverkan i alla instanser.⁶

6. Tack till Krister Bykvist, Jens Johansson och Olle Risberg för insiktsfulla förslag och kommentarer till den här texten.

Litteratur

- Nefsky, Julia. 2012. *The Morality of Collective Harm*. Doktorsavhandling. Berkeley: University of California. <https://escholarship.org/uc/item/9s49q22t>
- Nefsky, Julia. 2017. "How You Can Help, without Making a Difference". *Philosophical Studies* 174, nr 11, s. 2743–67.
- Nefsky, Julia. 2019. "Collective Harm and the Inefficacy Problem". *Philosophy Compass* 14, nr 4. <https://doi.org/10.1111/phc3.12587>
- Parfit, Derek. 1984. "Five Mistakes in Moral Mathematics". Kap. 3 i *Reasons and Persons*. Oxford: Oxford University Press.
<https://doi.org/10.1093/019824908X.001.0001>
- Thornéus, Ebba. 2019. "Chocksiffran: Sveriges nollutsläpp äts upp på åtta veckor". *Aftonbladet*, 6 juli 2019. <https://www.aftonbladet.se/a/lArjby>

Heidegger mot moderniteten – en instruktiv förbrytelse?

Erik Hallstensson

Ur ett högt flygande fågelperspektiv på det korta 1900-talet kan det verka märkligt att publiceringen av Martin Heideggers anteckningsböcker, de så kallade svarta häftena, fått så mycket uppmärksamhet. Innehållet i de svarta vaxhäftena, ett slags renskrivna filosofiska dagböcker, började ges ut 2014 och är den senaste och sannolikt sista större delen av Heideggers samlade verk.¹ Vid det här laget är många filosofers sympatiserande med olika totalitära partier och statsapparater välkända. Heideggers deltagande i det nazistiska maktövertagandet genom att 1934 låta sig utses till rektor över universitetet i Freiburg är inte heller unikt bland filosofer (eller andra intellektuella). Det tio månader långa ämbetsinnehavet, som slutade med att Heidegger sade upp sig, har debatterats i vågor i en filosofisk offentlighet åtminstone sedan det tal Heidegger höll vid sitt tillträde publicerats. Denna debatt har fått delvis ny näring av innehållet i häftena. Mer uppseendeväckande är de samband mellan Heideggers filosofi och den inflytelserika antisemitiska västerländska tanketradition som häftena lägger fram i ljuset. Debatten om dessa samband har varit omfattande och stundtals hätsk, framför allt i Tyskland och Frankrike. Här ska några bidrag till den presenteras, undersökas och värderas.

Fyra volymer, av planerade sju, av häftena har publicerats mellan 2014 och början av 2021. De fyra volymerna täcker åren 1931 till 1941. Det är kvantitativt och kvalitativt omfångsrika böcker. Därför måste

1. Hänvisningar ges till samlingsvolymernas egen paginering, och alltså inte till de enskilda häftenas ursprungliga paginering, som också finns angiven i den samlade verkutgåvan. De enskilda volymerna förkortas till GA och deras ordningstal i samlingsutgåvan.

Redan i Heideggers enda offentliga intervju efter andra världskriget började hans tidigare skrifers filologiska vederhäftighet ifrågasättas. Av utrymmesskäl har ingen hänsyn till dessa frågor tagits här, men en erinran är på sin plats (se Heidegger 1981, Søbye 2019 och Babich 2016).

Alla översättningar är författarens egna. Inga ansatser att följa Heideggers till svenska översatta terminologi i Ruin 2005 eller i de olika översättningarna av *Vara och tid* har gjorts. Det huvudsakliga skälet är att häftena är skrivna när Heidegger själv anser att han distanserar sig från *Vara och tid*. Terminologin i häftena kan alltså inte utan vidare anses vara utbytbar med den i *Vara och tid*.

frågan om Heideggers filosofiska antisemitism preciseras. Många försök att förstå Heideggers anteckningar har redan gjorts sedan häftena började komma ut. Hur mycket nazist var egentligen Heidegger? Har hans antisemitism något, och i så fall vilket, samband med hans publicerade filosofi som inte nämner någon sådan? Mindre uppmärksamhet har ägnats åt de begreppsliga sambanden mellan Heideggers kritik av moderniteten och hans antisemitiska anteckningar i häftena. De bidrag till Heideggerdebatten som här tas upp värderas i ljuset av det här problemet.

Senast efter det första världskrigets industrialiserade krigföring har det västeuropeiska sättet att leva på kommit i kris. Vad som är sant om den värld som människorna både innebor och skapar är satt i fråga. Symptomen på denna kris är kalkylerandets utbytbarhet, organisationstänkandet, den ohanterbara hastigheten, anpassningen av kulturen till masskonsumtion och massestetik. Förhållanden till den mänskliga existensen som sådan vulgariseras. Kalkylerandet är nära förbundet med en närvarometafysik, där naturen betraktas som något som är till hands för människor att mäta och därmed bemäktiga sig. Naturvetenskaperna teknifieras, humaniora blir till instrument för politik och världsåskådning, rättsvetenskapen görs överflödigt, medicinen snävas in till biologisk teknik, teologin blir meningslös och universiteten blir en byråkratisk angelägenhet för senkomna viktigpettrar. Kulturen är lika med inrättandet av gigantiska upplevelsemaskiner. Människans svar på frågor om vem hon är övertas av mekaniseringens effekter. Fler eller färre av de här idéerna kan kännas igen hos många modernitetskritiker.

Det kulturkritiska svepet ovan innehåller många idéer som i grader delas av olika kritiker, i olika ändrar av den politiska skalan. Det är dock Heideggers syn på sin samtid som återges i det förra stycket.² Michael Fagenblat sammanfattar det filosofiska tänkande som Heidegger menar bereder väg för senare dagars kulturkris: Platonismens abstraktion, Cartesianismens grundande av den säkra kunskapen i ett subjekt, nykantianismens individualism och idealism och scientismens utbyte av mänsklig historia och sanning till teknikens historia och mätbarhet (Fagenblat 2016, s. 148).

Heidegger ägnar särskild uppmärksamhet åt vilka som i synnerhet driver fram krisen:

2. Den sammanfattande listan med diagnostiska begrepp är hämtad från Vallega-Neu 2016, s. 131. GA 94, s. 38, s. 48–51, s. 199, s. 365–66; GA 95, s. 1, s. 39.

Judenhetens tillfälliga maktökning har däremot sin grund i att västerlandets metafysik, särskilt i dess moderna utveckling, erbjuder ansatsläget för utbredningen av en annars tom rationalitet och beräkningsfäihighet, som på sådana vägar skaffar fram en tillflykt i 'anden', utan att kunna fatta de förborgade beslutsområdena utifrån sig självt. Ju ursprungligare och ju mer inledande de framtida frågorna och besluten blir, desto mer otillgängliga blir de för denna 'ras'. (GA 96, s. 46. Heideggers egna citationstecken.)

Judarna 'lever' med sin utpräglad räknemässiga begåvning redan sedan länge efter rasprincipen, varför de också som häftigast motsätter sig dess oinskränkta tillämpning. Inrättandet av raslig uppfostran stammar inte ur 'livet' självt, utan ur övermannandet av livet med intrigmakeri [*die Machenschaft*]. Vad de driver fram med en sådan planering är en *fullständig avrasifiering* av folken genom att spänna dem framför ett inrättande av allt varande som är likadant uppbyggt och av samma snitt. Med avrasifieringen går folkens självförfrämligande mot ett enda – förlusten av historien – alltså av Varats beslutsområden. (GA 96, s. 56. Heideggers egen kursivering.)

Hur kan en tänkande människa läsa Heideggers filosofi, inbegripet innehållet i häftena? Heideggers kulturkritik innefattar en antisemitism som identifierar judar som bärare av hans filosofiska problem. Det är den övergripande frågan som den stundtals nebulösa Heideggerdebatten handlar om.

Det kan inte vara så enkelt som att Heidegger borde utgå ur allt filosofiskt frågeställande. Inget av Heideggers egna filosofiska utkast till svar på problemen tas upp här. Det skulle göra en omfattande tolkning av hans projekt nödvändig, något som utrymmet inte tillåter. Det är ändå klart att hans försök att underminera grunderna för en lång tradition av epistemologi och ontologi är både sofistikerade och relevanta. Det finns heller inget självklart svar på hur tänkare med antisemitiska övertygelser ska fortsätta att läsas och undervisas. Många finns i kursplaner och som klassiker inom olika fält, och kommer av allt att döma att finnas kvar där.³ Christian Fuchs har föreslagit att kritisk medie- och kommunikationsvetenskap ska göra sig av med Heideggers inflytande på fältets teori (Fuchs 2015, s. 74–75). Fuchs utgår från analyser av Heideggers filo-

3. Se bara som några exempel Dummett 1973, s. xii om Gottlob Frege, och mycket av Martin Luther. Se också Svenungsson 2017 om Heideggers inflytande på teologin under 1900-talet.

sofi av Karl Löwith, Victor Farías, Tom Rockmore och Emmanuel Faye (Fuchs 2015, s. 56–57). De anser alla, på filosofiskt skilda sätt, att nazism och antisemitism är kärnan i hela Heideggers filosofi. Faye argumenterade redan innan utgivandet av häftena för att Heideggers böcker ska flyttas från filosofihyllorna till "nazismens och hitlerismens historiska arkiv" (Faye 2009, s. 318–19).

Fuchs diskvalificerar flera strategier som han menar har använts för att försvara Heidegger: Den oinitierade kan inte förstå helheten; Heideggers politik och filosofi överlappar egentligen inte varandra; Heidegger var bara anhängare av den nazistiska ideologin under sin tid som universitetsrektor; Heidegger försvarade vissa judar och var begreppsligt emot nazismen (Fuchs 2015 64–65). Citaten ovan ur häftena räcker för att ge Fuchs rätt mot ytliga apologier av den här typen. Det finns å andra sidan också skäl att snarare intensifiera det seriösa studiet av Heidegger. Frågan som Fuchs inte ställer på något förutsättningslöst sätt är den här: Varför, filosofiskt förstått, glider Heideggers antisemitism så lätt in i en kritik av moderniteten som använder begrepp som "rotlöshet", "teknifiering", "utbytbarhet" och "mekanisering"? Som Sander L. Gilman visat har judar åtminstone sedan upplysningen bildat figurer i debatterna om det över- eller utanförnationella, och därmed i föreställningar om vad som är rotat och rotlöst (Gilman 2017, s. 19ff). Idéhistoriskt går det naturligtvis att hävda att just Heideggers kritik har sina rötter i radikalkonservativa tankeströmningar från 1900-talets början.⁴ Därifrån utvecklar Heidegger idéerna vidare i sitt filosofiska projekt. Frågan blir då denna: Om en kritik av moderniteten vill utgå från de här begreppen, utan att Heidegger får ge dem deras betydelse, vad kan de då betyda? Om Heidegger ansåg att till exempel rotlöshet betyder att det finns rotade och rotlösa, hur ser den indelningen ut i en annan kritiks avsiktliga och oavsiktliga användning av sådana begrepp? Vilka driver teknifieringen? Hur går det till? Innehåller kritiken också andra som är sårbara för den?

En kritik av moderniteten som vill ta ett större ansvar för idéernas konsekvenser än vad Heidegger gjorde bör informeras av en läsning av Heidegger. Hans inpassande av sin filosofi i samtidens katastrofala politik borde leda till en för modernitetskritiken självkritisk fråga: Var i denna samtids politiska indelning av människor passar en annan modernitetskritik in? Heidegger ser i häftena runt tiden för rektoratet nazismen som en möjlighet till förverkligande av sina filosofiska idéer

4. Carl-Göran Heidegren (med flera) har kartlagt denna miljö (Heidegren 2019).

(GA 94, s. 111ff). Denna förväntan förbyts snabbt i stor besvikelse. Hela vidden av detta är inte klarlagt, bland annat eftersom Heideggers hela kvarlåtenskap inte är allmänt tillgänglig (GA 94, s. 144–60).

Ett exempel på att inte ta ansvar för en filosofisk modernitetskritik är Friedrich-Wilhelm von Herrmanns påståenden om häftenas roll i Heideggers filosofiska bana. De begrepp Heidegger använder, menar von Herrmann, är sådana med vilka Heidegger "karaktäriserar den *nyaste nya tidens anda* och därmed *nutiden*, såvitt denna tidsålder huvudsakligen förstår sig själv utifrån de matematiska naturvetenskapernas och den moderna teknologins anda" (Von Herrmann 2016, s. 91, Von Herrmanns egen kursivering). Detta är enligt von Herrmann uppenbart för den som noggrant arbetat sig igenom Heideggers huvudsakliga, delvis publicerade, texter. "Med andra ord", fortsätter von Herrmann, "när Heidegger karaktäriserar 'den internationella judendomens' anda inkluderar han den i samtidens dito. Det måste förstås som en reflex av den då dominerande tidsandan att Heidegger specifikt nämner och kritiserar 'världsjudendomen', även om den operationella karaktäriseringen också används om den allmänna moderna andan i nutiden" (Von Herrmann 2016, s. 91–92). Heidegger är alltså inte antisemit, enligt von Herrmann.

Von Herrmann försöker förskjuta frågan om Heidegger och antisemitismen till att handla om filosofers ansvar för och förmåga att tänka emot deras filosofiska och politiska samtid. Det är en svår och omfattande fråga som alltid är värd att ställa, inte minst till Heidegger. Den erbjuder ytterligare en möjlighet till filosofisk självkritik. Även om svaret på den frågan skulle vara att ingen på något sätt kan förväntas vara immun mot sin samtids vidrigheter, fungerar inte von Herrmanns apologi. Låt det vara givet att Heideggers anteckningar är vad von Herrmann säger att de är: en del av hans långa projekt att kritisera hela den samtida moderniteten. Varför är då just judar föremål för utpekande, som modernitetens problems särskilda bärare och iscensättare? Med von Herrmanns språkbruk: Om den moderna tidens anda finns i Heideggers (och vårt) samtida samhälle, varför utpeka just judar? Heidegger anser att det finns ett särskilt judiskt sätt att vara (här använder jag mig inte av hans egen terminologi) som i högre grad än hos andra människor är rotlös, kalkylerande, intrigant och så vidare. Det är inte en karaktärisering av en tidsanda, utan Heideggers egna antisemitiska karaktärologi som Peter Trawny har rätt ut (Trawny 2014).

Ett annat exempel, inte på ansvarslöshet utan på hur poängen ovan om modernitetskritikens begreppsvärld kan vara produktiv, är Ba-

bette Babichs framkastade tanke om Heideggers begrepp. Det nyaste nya, säger Heidegger, har övergett kraften att fatta beslut om människans grundläggande frågor till förmån för ett grundlöst påträngande och inträngande i hela livet (GA 96, s. 272). Heideggers kritik i häftena av medietekniken borde läsas och användas för att förstå Twitter och Youtube, säger Babich i en passus (Babich 2016, s. 76). När det här skrivs påstår *Dagens Nyheter* att den kvinna som sköts till döds av poliser när hon stormade Kapitolium i Washington återgav apokalyptiska konspirationsteorier i sina sociala medier (Dahl 2021). Det grundlösa på- och inträngandet (som Heidegger problematiserar på 1930-talet) kan inleda ett filosofiskt utforskande av medier, sociala som asociala. För den som tar sig an ett sådant projekt står en mycket komplicerad filosofisk bakgrund till förfogande, i häftena och i andra Heideggertexter. Där och då måste alltså även frågan ställas om var denna skjutna 35-åring passar in i någon filosofisk slutsats. Är hon någon vars värld invaderats av de sociala mediernas konspirativa lögner? Naturligtvis görs här inga jämförelser mellan vad Heidegger skriver och säger om judar och de som bröt sig in i USA:s parlament. Poängen är teoretisk, och uppmanar till en rörelse från teoretisering till konsekvenser i de sammanhang som modernitetskritiken inte kan undvika, moderniteten själv.

För Heidegger är det inte någon slump att rotlöshet, reduktivt kalkyltänkande och så vidare kan appliceras på särskilt utpekade personer. Heidegger är utan tvivel antisemit åtminstone vid tiden för de hittills utgivna häftena. Frågan är vad dessa modernitetskritiska begrepp kan betyda, och vilka de pekar ut i samtiden.

Litteratur

- Babich, Babette. 2016. "Heidegger's Black Night: The *Nachlass* and Its *Wirkungsgeschichte*". I Farin och Malpas, s. 59–86.
- Dahl, Amanda. 2021. "Kvinnan som sköts i Kapitolium: En krigsveteran som röstat på Obama". *Dagens Nyheter*, 2021-01-09. <https://www.dn.se/varlden/kvinnan-som-skots-i-kapitolium-en-krigsveteran-som-rostat-pa-obama/>
- Dummett, Michael. 1973. *Frege: Philosophy of Language*. New York: Harper & Row.
- Fagenblat, Michael. 2016. "'Heidegger' and the Jews". I Farin och Malpas, s. 145–68.
- Farin, Ingo och Jeff Malpas, red. 2016. *Reading Heidegger's Black Notebooks 1931–1941*. Cambridge, Mass.: MIT Press.
- Farin, Ingo och Jeff Malpas. 2016. "Introduction". I Farin och Malpas, s. ix–xiv.
- Faye, Emmanuel. 2009. *Heidegger: The Introduction of Nazism into Philosophy*. New Haven: Yale University Press.

- Fuchs, Christian. 2015. "Martin Heidegger's Anti-Semitism: Philosophy of Technology and the Media in the Light of the Black Notebooks", *tripleC: Communication, Capitalism & Critique* 13, nr 1, s. 55–78.
- Gilman, Sander L. 2017. "Cosmopolitan Jews vs. Jewish Nomads". I Andrew J. Mitchell och Peter Trawny (red.). *Heidegger's Black Notebooks*. New York: Columbia University Press.
- Heidegger, Martin. 2014. *Gesamtausgabe*, Band 94, *Überlegungen II–VI*. Frankfurt am Main: Vittorio Klostermann.
- Heidegger, Martin. 2014. *Gesamtausgabe*, Band 95, *Überlegungen VII–XI*. Frankfurt am Main: Vittorio Klostermann.
- Heidegger, Martin. 2014. *Gesamtausgabe*, Band 96, *Überlegungen VII–XI*. Frankfurt am Main: Vittorio Klostermann.
- Heidegger, Martin. (1966) 1981. "Intervju med Der Spiegel". *Insikt och handling*, nr 3/4, s. 121–49.
- Heidegren, Carl-Göran. 2019. "Konservatismens, nationalismens och ungdomens radikalisering under Weimarrepubliken". I Storm Torjussen och Berdinesen, s. 221–42.
- von Herrmann, Friedrich-Wilhelm. 2016. "The Role of Martin Heidegger's *Notebooks* within His Oeuvre". I Farin och Malpas, s. 89–94.
- Malpas, Jeff. 2016. "On the Philosophical Reading of Heidegger: Situating the *Black Notebooks*". I Farin och Malpas, s. 3–22.
- Ruin, Hans. 2005. *Kommentar till Heideggers Varat och tiden*. Stockholm: Södertörn Philosophical Studies.
- Storm Torjussen, Lars Petter och Hein Berdinesen, red. 2019. *Heideggers testamente: Filosofien, nazismen og de svarte heftene*. Oslo: Dreyers forlag.
- Svenungsson, Jayne. 2017. "Introduction: Heidegger and Theology after the Black Notebooks". I *Heidegger's Black Notebooks and the Future of Theology*, red. Mårten Björk och Jayne Svenungsson. London: Palgrave Macmillan.
- Søbye, Espen. 2019. "Litterär existens tilendebrakt". I Storm Torjussen och Berdinesen.
- Trawny, Peter. 2014. *Heidegger und der Mythos der jüdischen Weltverschwörung*. Frankfurt, Main: Vittorio Klostermann.
- Vallega-Neu, Daniela. 2016. "The *Black Notebooks* and Heidegger's Writings on the Event (1936–1942)". I Farin och Malpas, s. 127–42.

Kommentar

Mycket sent om Hedenius och Mathlein

Johan E. Gustafsson

I en recension av Ingemar Hedenius bok *Om människans moraliska villkor* hävdar Hans Mathlein att Hedenius har begått ett grovt formellt misstag. Hedenius misstag är, enligt Mathlein, en definition där definiens innehåller en fri variabel som inte förekommer fritt i definiendum (Mathlein 1973a, s. 93). Det eventuella problemet är symbolen "Mkt" i följande definition av "X är ett element, E, i ett konstverk, K":

(Df 1) X är E i $K = X$ är E i K i Mkt . (Hedenius 1972, s. 303)

I en replik protesterar Hedenius att "Mkt" inte skall förstås som en fri variabel utan som ett "individuellt förhållande" (Hedenius 1973, s. 96). Mathlein svarar:

Jag vidhåller att Hedenius begått formella fel. Hans påstående att symbolen "Mkt" (som skall utläsas: kulturmiljön vid tidpunkten t) inte är en variabel, utan betecknar ett (bestämt?) "individuellt förhållande", är inte en rimlig tolkning av texten. Jag har uppfattat det så att Hedenius före de formella definitionerna säger något om det värdeområde som variabeln "Mkt" har och att denna symbol *är* en variabel. Definitionerna blir, uppriktigt sagt, alldeles obegripliga om "Mkt" betecknar en individkonstant. Exakt vilken kulturmiljö vid vilken tidpunkt betecknar i så fall det "Mkt" som står sju rader nerifrån på sid. 303? [(Df 1)] De läsare som har någon logisk-semantisk skolning håller nog med om min kritik ifall de granskar uppsatsen. (Mathlein 1973b, s. 97)

Så har Mathlein rätt? För att reda ut detta behöver vi granska Hedenius förklaring av "Mkt", som lyder så här:

Ett relationsled, som inte framträder i den språkliga formen " X är E i K " men måste tillvaratas i en explikatorisk analys av satser med denna form, är den kulturmiljö, i vilken X är E i K . Jag anger denna miljö med "Mkt", där indexsymbolen " t " avser den tidsbestämning det måste vara frågan

om, när kulturmiljön för ett visst konstverk omtalas på ett meningsfullt sätt. (Hedenius 1972, s. 303)

Med andra ord, Mkt är en kulturmiljö och t är en tidsbestämning sådan att X är E i K i den miljön under t . (Eller, mer exakt: Mkt och t är den *unika* kombination av kulturmiljö och tidsbestämning sådan att X är E i K i den miljön under den tiden. Men denna märkliga komplikation spelar ingen roll för Mathleins kritik.) Detta innebär att definitionen skall läsas så här:

(Df 1*) X är E i K = Det finns en kulturmiljö M och en tidsbestämning t sådan att X är E i K i M under t .

Nu är det tydligt att definiens inte innehåller någon fri variabel som inte förekommer fritt i definiendum. Men, så klart, huruvida Hedenius explikatoriska analys är rimlig (eller huruvida Hedenius behärskade filosofins mer formella metoder) är en annan fråga.

Litteratur

Hedenius, Ingemar. 1972. *Om människans moraliska villkor*. Göteborg: Författarförlaget.

Hedenius, Ingemar. 1973. "Ingemar Hedenius svarar". *Bonniers litterära magasin* 42, nr 2, s. 94–97.

Mathlein, Hans. 1973a. "För och emot Hedenius". *Bonniers litterära magasin* 42, nr 2, s. 93–44.

Mathlein, Hans. 1973b. "Replik". *Bonniers litterära magasin* 42, nr 2, s. 97.

Recensioner

Why Solipsism Matters

Sami Pihlström

Bloomsbury 2020. 240 s. ISBN 9781350126398

Solipsismen är en filosofisk position som sällan diskuteras i fackfilosofiska kretsar, kanske eftersom den är svår att förhålla sig till, intellektuellt men kanske framför allt känslomässigt. I sin ontologiska form går den ut på att det enda som existerar är jag själv inklusive mina upplevelser; att allt som förefaller ha sitt ursprung utanför min upplevelse i själva verket endast existerar inuti denna, och att världen således är identisk med min upplevelse av den. En i grunden mycket enkel, men fullständig konstraintuitiv åskådning.

Själv brottades jag i flera perioder under min uppväxt med solipsistiska grubblerier och försökte utan framgång att hitta en väg ut ur problematiken. När jag senare läste filosofi på universitetet fick jag bekräftat att jag inte var ensam om att inte vilja omfatta solipsismen – de flesta lärare och forskare jag pratade med ville inte ens ta den på allvar, även om de medgav att den rent logiskt var en tänkbar möjlighet. Men till skillnad från dem var jag seriöst intresserad av att diskutera den och vad den skulle innebära om man omfattade den fullt ut. Detta intresse har bestått och jag är därför glad att den finländske filosofiprofessorn Sami Pihlström nu har tagit sig an ämnet i en bok som heter *Why Solipsism Matters* (Bloomsbury 2020). I den diskuteras solipsismen i dess olika former, vilka implikationer den får, inte minst för vår syn på etiken, och hur vi kan förhålla oss till en position som ingen vill bekänna sig till, trots att den teoretiskt verkar omöjlig att motbevisa.

Pihlström klargör tidigt i boken att han själv inte är solipsist men att han, till skillnad från det stora flertalet filosofer, ändå finner solipsismen vara en intressant position som kan lära oss mycket om villkoren för vårt utforskande av världen. Man behöver inte vara ontologisk solipsist och hävda att det enda som *existerar* är den egna upplevelsen, för att vara kunskapsteoretisk solipsist och hävda att det enda som man strängt taget kan ha *kunskaper* om är den egna upplevelsen. Allt jag får veta när mig via min egen mentala apparatur, det vill säga mitt subjek-

tiva perspektiv; därför är solipsismen (eller i vart fall subjektivismen) rent metodologiskt den enda fruktbara utgångspunkten, menar Pihlström. Visst kan man extrapolera och tala om *andra* människors och deras upplevelser, vilka mycket väl skulle kunna existera oberoende av en själv – men det första ledet i kunskapsinhämtningen, som innebär en undersökning av innehållet i den egna upplevelsen (just nu), kommer man inte ifrån. Omvärlden når en på det sättet indirekt.

Detta har så många filosofer genom historien insett; ändå är det få eller inga av dem som bekämt sig fullt ut till solipsismen, även om flera av dem har varit framme och nuddat vid tankegången. Pihlström nämner ett flertal av dem, till exempel 1600-talsfilosofen Descartes som med hjälp av sitt metodiska tvivel kom fram till att det enda absolut säkra, som han ville härleda all annan kunskap från, var att han själv existerade som tänkande varelse. Han anlade därmed ett subjektivistiskt perspektiv på vår kunskap, som sedan förvaltades av Locke, Berkeley och Hume, vilka konstaterade att allt vi tror oss veta måste återgå på sinneserfarenheten. Vad Kant sedan konstaterade var att vi inte kan ha kunskap om den oberoende verkligheten, "tinget i sig", utan bara om världen så som den ter sig för oss sedan vi underkastat den våra kunskapsformer (där de viktigaste är åskådningsformerna tid och rum). En tankegång som Schopenhauer sedan byggde vidare på i en filosofi som definitivt går att tolka solipsistiskt. Eller vad sägs om inledningsorden till hans huvudverk *Världen som vilja och föreställning*: "Världen är min föreställning".

Schopenhauer verkar dock inte mena att man ska, eller behöver, tolka hans filosofi på det sättet; tvärtom menar han att man ska akta sig för det. I sitt nämnda huvudverk (s.174) påpekar han att solipsismen, vilken han benämner som den teoretiska egoismen, visserligen aldrig kan vederläggas med bevis. "... dock har den med visshet ej funnit någon annan användning i filosofien än som en skeptisk sofism ... Som allvarlig övertygelse däremot kunde den endast återfinnas på ett hospital: och då behövdes det mot densamma ej så mycket en bevisning som en kur; varför vi så till vida ej inlåta oss med den, utan endast betrakta den som ... en liten gränsfästning, som visserligen är ointaglig för all tid, men vars besättning ej heller på något sätt kan komma ur den, varför man kan gå förbi den och utan fara lämna den i ryggen på sig." Med andra ord: låt solipsismen vara för man vet inte vad den kan leda fram till. Jag antar att många nutida filosofer skulle hålla med honom om detta.

Men även Schopenhauers efterföljare har varit framme och nosat på

solipsismen, inte minst 1900-talsfilosofen Wittgenstein. Det är framför allt hans ingång till åskådningen som Pihlström är intresserad av, eftersom den formuleras på ett sätt som får skillnaden mellan solipsism och realism (icke-solipsism) att upphävas. I likhet med Schopenhauer menar Wittgenstein att världen är en föreställning hos ett "jag" och att världen förhåller sig till jaget som synfältet till ögat. Men det "jag" som föreställer sig världen kan inte självt *tillhöra* världen utan utgör en gräns för världen, och frågan är då om jaget alls finns i någon empirisk mening. Ingen har veterligt observerat ett metafysiskt subjekt, och eftersom jaget inte finns *i* världen så är det bara världsföreställningen som i själva verket existerar, menar Wittgenstein. I sitt verk *Tractatus* (5.64) skriver han att "solipsismens Jag krymper till en punkt utan utsträckning och kvar blir den med Jaget samordnade realiteten." I en mening blir detta detsamma som realism, eftersom det innebär att världen faktiskt existerar. Å andra sidan existerar den ju fortfarande inte som något annat än *min* värld, eller *någons* värld, beroende på hur detta tolkas.

I sin bok har Pihlström inga ambitioner att försöka vederlägga solipsismen, i varje fall inte Wittgensteins tappning av den. Han återkommer flera gånger till att vilka argument som än anförs för att solipsismen är falsk eller problematisk, till exempel att vi lever i en språklig gemenskap som till sin natur kräver en interaktion mellan flera olika subjekt, så kan solipsisten alltid inkorporera detta i sin egen världsbild och säga att denna interaktion, liksom allting annat, bara finns i form av upplevelser. Ändå är det vår plikt, menar Pihlström, att någonstans i diskussionen göra halt och *inte* följa med solipsisten in i hans eller hennes omtolkning av allt till fördel för den egna åskådningen. För vad händer annars med etiken? Vad händer med vår omsorg om andra varelser?

Pihlström verkar mena att realism i förhållande till andra människors tankar och känslor, det vill säga en övertygelse om att dessa existerar oberoende av den egna livsupplevelsen, är en förutsättning för att man ska kunna bemöta dem som individer med egna moraliska rättigheter. Om solipsismen skulle omfattas fullt ut är det risk för att etiken i alltför stor utsträckning skulle bli "min" etik som jag kan bestämma över som jag själv önskar, utan att "andra" varelsers perspektiv skulle spela någon egentlig roll. Detta, i sin tur, skulle öka risken för att andra varelser skulle reduceras till instrument för den egna behovstillfredsställelsen, vilket underminerar själva idén med att leva moraliskt. Vi måste därför, verkar han mena, fortsätta att *utgå* från att realismen är sann, även om vi inte på intellektuell väg kan bevisa att det förhåller sig så.

Det betyder dock inte att vi bara ska "bestämma" oss för att vara realist och styvnackat "tro" på existensen av en oberoende omvärld, på samma sätt som en religiös person kan fortsätta att tro på en gud utan att behöva ge några förnuftiga skäl för det. Det vore, menar Pihlström, en filosofisk skandal om vi skulle behöva acceptera världens existens på basis av ett sådant icke-diskursivt *credo*. Vad vi istället behöver är ett väl underbyggt försvar för att vi, trots allt, fortsätter att utgå från att realismen är sann. Ett sådant försvar skulle kunna vara något i stil med: jag vet att det inte slutgiltigt går att "bevisa" en oberoende omvärld, men jag fortsätter ändå av pragmatiska skäl att utgå från att en sådan existerar, inte minst därför att risken är stor att etiken i min relation till omvärlden skulle bli lidande om jag inte utgick från detta. Att vara realist, menar Pihlström, är ett existentiellt val som var och en av oss måste träffa på nytt i varje ögonblick, och vi kan göra det baserat på skäl som vi själva finner förnuftiga utan att en solipsist med nödvändighet behöver bedöma dem som övertygande. Vi väljer därmed att förhålla oss till världen på det sätt som bäst gynnar vår förmåga att interagera med den.

Det är ingen tvekan om att Pihlströms bok ger en god överblick över den diskussion som förts om solipsismen, kanske framför allt om hur Wittgensteins tappning av den ska tolkas. Det är också värt en eloge att Pihlström, trots sin uppenbara ovilja att omfatta solipsismen, inte låter sig övertygas av argument mot den som uppenbart inte håller, utan enbart förtydligar en redan realistisk världsbild. Vad jag dock saknar i boken är en redogörelse för vad solipsismen skulle innebära *om* den var sann – till exempel för hur ordet "jag" skulle få en radikalt ny innebörd om det plötsligt denoterade *allt* som existerar, istället för att som i realismen bara denotera en liten del. Det skulle bland annat innebära att vi fick en helt ny ingång till etiken, eftersom den absoluta gränsen mellan en själv och omvärlden då skulle upphöra att finnas. Vilket i sin tur skulle innebära att man helt eller delvis *identifierade* sig med andra var-elser, till moralens uppenbara fördel. Dessutom kan solipsismen medföra en viss förhoppning om en fortsatt livsupplevelse efter den "fysiska" kroppens död, eftersom den utgår från en helt annan relation mellan medvetande och kropp. Men kanske skulle ett sådant resonemang bli alltför spekulativt eller esoteriskt för en bok som denna.

Jag saknar också en redogörelse för det faktum att de flesta av oss, även om vi på ett filosofiskt plan skulle erkänna att solipsismen (möjli-gen) var sann, skulle fortsätta att leva våra liv precis som vi gjorde tidigare. I och med det kanske det inte är så viktigt som Pihlström säger att

vi vägrar följa med solipsisten in i hans eller hennes omtolkningar. Vår tro på en oberoende existerande omvärld är nämligen så instinktiv och orubblig att den sannolikt aldrig lämnar oss, oavsett vilka argument som en opponent anför.

Hans Teke

Filosofi och pandemi

Red. Åke Gafvelin och Lapo Lappin

Artos & Norma 2021. 102 s. ISBN 978-91-7217-129-9

Pandemiåret 2020. Det kommer att bli slitet att säga så, men i detta fall är det påkallat. Tvingande social distansering är nämligen bakgrunden till *Filosofi och pandemi*, en bok redigerad av filosofistudenterna Åke Gafvelin och Lapo Lappin.

Både boken och projektet den bygger på är nog att betrakta som en sorts revolt mot påtvingad ensamhet. Under 2020 samlade nämligen redaktörernas digitala intervjuserie "Det metafysiska laboratoriet" några av landets mest profilerade filosofer för att diskutera filosofiska frågor som bland annat uppstått med anledning av pandemin.

Projektet är inspirerande. Redaktörerna har valt att försöka sig på något nytt och oprövat, och det tillsammans med inbjudna filosofikändisar. Det är vågat, och jag antar att allt måste ha varit ganska diffust i förväg, vilket hedrar alla inblandade som trots allt gav sig in i projektet.

Projektet är dessutom ovanligt brett. Redaktörerna nöjer sig inte med att diskutera aspekter av pandemin, som boktiteln antyder, utan tar även upp frågor som kunskapens natur, experternas roll i demokratiskt beslutsfattande, etisk teori, tillämpad etik, värdeteori och metafysik. Även det är vågat.

Dessutom är boken överlag skickligt redigerad. En risk var ju att den hade kunnat urarta i en serie referat av olika inlägg, men med vissa undantag låter det som en diskussion (även om jag har en del att säga om just detta). En detalj jag uppskattade var att om man vet hur dessa filosofer talar, så "hör" man dem också i texten. Uppenbarligen har man lyckats fånga ordval och meningsbyggnad – ja, till och med flera välkända käpphästar – så att deltagarna låter närvarande.

I förordet reflekterar redaktörerna över om vi kanske kan ha att göra med en renässans för det talade språket. De skriver att "[d]en skriftliga filosofins självklara företrädare har kanske börjat rubbas: nätet skapar fler

och fler utrymmen för filosofiskt utbyte, genom livesändningar, Zoom-konferenser, videoinspelningar ..." (s. 12). De skriver dessutom att

[det är] värt att hålla i tanken att intervjuerna bottenar i just det talade språket. Det är ett format med uppenbara fördelar, som tillåter ett inlevarseregister som det skrivna språket inte förmår fånga på samma sätt.

Men samtidigt framhåller de detta:

Bokformatet öppnar nya dörrar: där kan vi lättare gå fram och tillbaka, jämföra och granska långsamt. Det som framträder [i *Filosofi och pandemi*] är ett samtal som sträcker sig över elva avsnitt.

På sina håll menar jag ändå att det uppstår en krock mellan muntligt och skriftligt. Jag ska lyfta några kritiska poänger som alla, på ett eller annat sätt, har att göra med denna krock.

Till att börja med något om att det i boken sägs framträda ett samtal, och att boken blir ett sätt att agera medlare i en "förlängd dialog" mellan filosoferna. Men en dialog är det ju inte, och kan det knappast vara i och med att boken, visar det sig efterhand, bygger på enskilda intervjuer med en filosof åt gången. Detta är att tänja på idén om vad ett samtal är.

Jag tror för övrigt att man hade vunnit på att klargöra detta upplägg tydligare. Jag besvärades av att jag inte direkt kunde avgöra vad det var jag skulle läsa. Längre undrade jag hur diskussionerna hade gått till, även om jag förutsatte att alla, eller åtminstone flera av, filosoferna hade deltagit samtidigt i en serie digitala seminarier ledda av bokens redaktörer. Och jag fick även intrycket av att man på något sätt utgått från det kombinerade konferens- och lunchrummet på Stockholms universitet som kallas "Metafysiskt laboratorium" i och med att intervjuerien kallas så. Dessvärre framgår det inte förrän i slutordet att projektet bygger på enskilda intervjuer med var och en av filosoferna. Där finns även en hänvisning till projektets webbplats där man kan ta del av intervjuerna, och dessutom en lista med 12 personer som bistått med att sammanställa manuset (de är inte någon av de intervjuade och det är oklart vilken roll de har haft i övrigt i projektet). Denna förvirring hade nog kunnat undvikas om upplägget framgått redan i förordet.

Ibland går även fördelen med bokformatet förlorat. Den filosofiska läsningen är som bekant långsam, noggrann och ifrågasättande. Som redaktörerna också skriver går man ofta fram och tillbaka, man jämför, och stannar upp, och tänker. Men frågan är om det verkliga kan bli

aktuellt med ett sådant läsande här. För det faktum att texten bygger på muntliga samtal tycks i flera fall påverka kvalitén på resonemangen.

Vi kan ta bokens allra första inlägg som exempel, där Fredrik Svenaeus reflekterar över filosofins roll. Han menar att "filosofier ska ägna sig åt frågor som är viktiga för människor i deras vardagsliv" och inte "kapsla in sig och syssla med sina specialproblem i sina bubblor" (s. 17).

Här hade jag önskat några klargöranden. För ärligt talat, hur ofta väljer vi att rådfråga en filosof i vardagslivets viktiga frågor? Hit räknar åtminstone jag en mängd olika saker, alltifrån vilka de bästa löparskorna är, till hur man äter en semla med värdighet och om man ska binda boräntan eller inte.

Men det är nog inte så Svenaeus menar. För filosofers uppgift är, får vi trots allt läsa, att fundera över den djupaste frågan, nämligen "livets mening och hur man ska se sin plats i tillvaron som människa". Som jag trodde alltså, knappast någon vardagsfråga.

Sedan avslutas inlägget med: "Det har implikationer för mitt eget liv också, även om jag tycker att jag kan vara ganska bra på att koppla isär de två sakerna." Det är oklart *vad* som har implikationer, och *vilka* saker som kan kopplas isär. Det hade man velat ha en förklaring av, eller möjlighet att fråga.

Så detta inlägg lämpar sig inte riktigt för noggrann läsning. Och jag tror att problemet är att det var tänkt som ett muntligt inlägg i ett ganska avslappnat samtal, och eventuellt "på volley", kanske utan särskild eftertanke. Problemet är att det har fästs på papper.

Svenaeus inlägg är bokens första, men problemet är genomgående. För om man går vidare och läser Jonna Bornemarks korta inlägg om samma sak så säger hon att filosofins uppgift dels är "att göra oss medvetna om de metafysiska idéer som vi baserar vår verklighet på, och dels, förstås, att bearbeta dem åt ett håll som gör världen bättre, intressantare, roligare" (s. 20).

Ingen borde förneka vikten av att medvetandegöra de idéer som ligger till grund för våra uppfattningar. Men vad betyder det att vi ska *bearbeta de metafysiska idéerna åt ett håll* som gör världen bättre, intressantare, roligare? Jag får känslan av att något saknas. (Samma känsla får jag av Bornemarks mycket längre inlägg om språk och verklighet på s. 91–92, för det inlägget tycks inte handla så mycket om att bearbeta metafysiska idéer.)

Problemet blir tydligt även i etikavsnittet. Bland annat kan vi stanna till vid Marcia Cavalcante Schubacks utgångspunkt i analysen av etiska

ställningstaganden: att begäret att ”kontrollera människan, både för det goda och för det onda, är hemskt” (s. 69). Återigen får jag intrycket av att resonemanget är plockat ur sitt sammanhang. Annars blir det svårt att förstå formuleringar som: ”Om man hade fått kontroll över människan, över den etiska genen, tänker många, hade vi aldrig haft utrotningslägren. Men det betyder att vi hade utrotat något före utrotningen” (s. 69). Här får läsaren fylla i en hel del själv, men jag antar att tanken är något i stil med att vår önskan att förhindra ont kan spåra ur och göra att vi i stället övervakar, förtrycker och stöper alla i samma form. Men det måste finnas en bättre analys av detta än att ”det är trots allt ett nazistiskt begär!” (s. 69).

Möjligen har något försvunnit i övergången från muntligt till skrift. Men det kan också vara att resonemanget inte är särskilt klart eller riktigt genomtänkt. Det går inte riktigt att avgöra.

Sedan måste det förstås tilläggas att det finns en hel del inlägg som fungerar bra trots det korta formatet. Som Tännsjös. Man kan förstås vara oense med honom om en massa saker, men i princip är alla hans inlägg, i bokens alla avsnitt, genomtänkta och konsistenta. Det gäller även flera av Ulf Jonssons inlägg. Han formulerar sig klart och avväpnande: ”Jag är inte etiker, så någon förfinad teori om det här har jag inte. Men jag försöker trots det leva ganska anständigt ändå” (s. 36).

Däremot är jag lite frågande inför att vissa filosofer är frånvarande under stora delar av boken. Exempelvis gör Karim Jebari bara några få inlägg, och inget av dem om exempelvis etik. Det är osannolikt att han inte hade något att tillägga där. Det är hur som helst synd att han saknas i det sammanhanget, för hans bidrag brukar vara tankeväckande. Även Marcia Cavalcante Schuback är tyst under hela coronadiskussionen (s. 19–45). Hade hon verkligen inget att säga om det? Och så kan man tycka att det är ovanligt tyst från Åsa Wikforss. Kanske blir det extra tydligt i avsnittet om språket och verkligheten (s. 88–93), för min gissning är att hon har en del att bidra med på det området.

Men trots min kritik av vissa bristfälliga inlägg, och mina svårigheter att till en början förstå upplägget, är det tydligt att boken är resultatet av ett ambitiöst, originellt, vågat och oväntat projekt som jag gärna ser mer av.¹

Anders Hansson

1. Efter att denna recension skrevs upptäckte jag att projektet fortlöper och att fler intervjuer genomförts. Så kanske blir det en uppföljare till denna bok.

Forskningsetik för humaniora

Håkan Salwén

Thales 2017. 88 s. ISBN 978-91-7235-108-0

Forskningsetik verkar ha blivit vanligare i universitetsutbildningar på senare år, framför allt förstås på forskarnivå. Även på avancerad nivå (masternivå) förekommer undervisning i ämnet. Det är för ett sådant sammanhang som Håkan Salwén har skrivit *Forskningsetik för humaniora*. Den utgör uppenbarligen en del av den obligatoriska litteraturen i en kurs vid Stockholms universitet avsedd för studerande vid masterprogram inom humanistisk fakultet. Boken vänder sig således till icke-filosofier och är av grundläggande karaktär.

Först ska sägas att Salwéns bok överlag är lättläst och trevlig. Den nyttjar flera väl valda exempel från de humanistiska forskningsområdena, t. ex. om uppgrävda samiska kranier, om strävan att återskapa utdöda språk och om litteraturforskning som tvivelaktigt övergår i litteraturkritik. Texten presenterar förtjänstfullt många grundläggande distinktioner och fakta, exempelvis om lagen om etikprovning. Boken har också en väl fungerande disposition.

Boken har fem kapitel, där kapitel 1 är en kort inledning. I kapitel 2 presenteras normer som kan anses finnas i vetenskapen och Salwén diskuterar huruvida de kan ge vägledning i exempelvis fall han ställer upp, och även huruvida normerna kan föras tillbaka på vardagsmoraliska principer. Etiska teorier såsom utilitarism och pliktetik nämns bara i förbigående. Även lagen om etikprovning presenteras kort. Kapitel 3 handlar om vetenskaplig oredlighet, vilket på engelska ofta anses omfatta *fabrication*, *falsification* och *plagiarism*, av Salwén översatt till "fabricering", "förfalskning" och "plagiat". Några exempel illustrerar de tre typerna. Sokals bluff dyker upp som exempel på förfalskning. Kapitel 4 handlar om forskningens inriktning, alltså vilken forskning man bör bedriva och varför. Salwén får därvid bl.a. anledning att diskutera begreppet akademisk frihet samt att redogöra för hur humanistisk forskning brukar finansieras. Philip Kitchers idé om att vi kan identifiera vilken forskning som bör bedrivas genom att hypotetiskt betrakta vetenskapen som "välordnad" (*well-ordered*) ges visst utrymme. Kapitel 5 handlar om värdet av humanistisk forskning.

En övergripande synpunkt är – förstås – att Salwéns bok med sina 88 sidor är väldigt kort. Den komprimerade framställningen kan i och för sig godtas med hänvisning till att texten verkar vara skraddarsydd för en

viss kurs. Men det är inte svårt att nämna forskningsetiska frågor som framstår som ganska vanliga inom humaniora men som diskuteras lite, eller inte alls, i boken:

- En del humanistiska (och samhällsvetenskapliga) forskare bedriver s.k. aktionsforskning. Forskaren tar då politisk ställning i frågor som anknyter till det beforskade, och söker åstadkomma förändring. Vilka forskningsetiska problem reser detta?
- Vilka forskningsetiska frågor uppstår när forskning bedrivs genom deltagande observation? Hur ska man i sådana sammanhang balansera forskningsmässiga intressen mot respekten för personlig integritet eller för andra kulturer? Vilka forskningsetiska problem uppstår när forskaren finner att de observerade personerna betar sig klandervärt (kanske brottsligt)?
- Vilka forskningsetiska frågor uppstår specifikt vid insamling och behandling av intervjudata? Vilka frågor uppstår vid enkäter?
- Mycket humanistisk forskning utgår från skrivna källor (historiska och samtida). För bedömning av sådana källors vetenskapliga användbarhet finns källkritiska principer (närhet, oberoende, äkthet, tendensfrihet), som dock traditionellt är befriade från etiska inslag. Kanske krävs också etiska principer för hantering och bedömning av dylika källor, t. ex. vissa värderingsmässiga attityder hos forskaren. I så fall vilka?

Låt mig också ta upp några ställen i boken som jag själv hade behandlat annorlunda.

En terminologisk synpunkt rör ordet "förfalskning" som översättning av engelskans "falsification". Salwén påpekar att "förfalskning" ofta betyder något helt annat i vardagsspråket än vad ordet betyder i hans bok (s. 40). Just därför kan man undra om han inte borde ha valt en annan översättning. "Förvrängning" har använts av bl.a. Medicinska Forskningsrådet och Vetenskapsrådet (Forsman, 2007, s. 4). I Hermerén, Gustafsson och Pettersson (2011, s. 107) används "förvanskning". Bägge dessa alternativ tycker jag är mer träffande än "förfalskning".

Jag blev förvånad över att finna Sokal-affären i boken. Nästan alla känner väl numera till att fysikprofessorn Alan Sokal år 1996 skickade en nonsensartad artikel till tidskriften *Social Text*, där uttalanden av vissa (så kallade) postmodernistiska forskare om naturvetenskap citerades

och påstods ha naturvetenskapligt stöd. Artikeln blev antagen i *Social Text*. Omedelbart därefter gick Sokal ut i en annan tidskrift och skrev, i korthet, att texten hade varit en bluff, att han ansåg att de citerade inte visste vad de talade om (när de använde begrepp från naturvetenskapen), samt att *Social Text* inte kunde vara en pålitlig tidskrift eftersom hans blufftext blivit införd där. En mycket stor debatt följde, som framför allt har handlat om huruvida Sokals kritik mot "postmodernisterna" var rättvis eller ej. En bidebatt, som inte minst Sokals kritiker har velat föra, har rört huruvida Sokal själv gjort sig skyldig till ett forskningsetiskt övertramp när han lät publicera sin opålitliga text. Det är för denna bidebatts skull som Salwén tar upp fallet, men jag tycker inte att det passar särskilt bra i boken. Sokals blufftext är helt opålitlig (en "förfalskning"), men är dock otypisk genom att upphovsmannen utan dröjsmål avslöjade att texten var en parodi. Texten är också otypisk i den meningen att den knappast kan klassificeras som ren humaniora – den innehåller många påståenden och resonemang som snarare faller under fysik eller annan naturvetenskap. Jag satte även ett frågetecken i marginalen där Salwén diskuterar vilka etiska övertramp Sokal kan ha gjort och därvid skriver: "Publiceringen innebar också att ett flertal forskare fick utstå väldigt mycket negativ publicitet" (s. 46), som om negativ publicitet till följd av vetenskaplig kritik skulle kunna utgöra ett forskningsetiskt övertramp. Publiceringen av originaltexten (bluffen) medförde enligt min bedömning ingen negativ publicitet för någon. Det var först när Sokal strax därefter redogjorde för vad han verkligen tyckte som han kritiserade andra forskare, varvid ytterligare andra instämde i kritiken. Men den kritik som Sokal formulerade, t. ex. i Sokal och Bricmont (1998), anser jag ryms i ett normalt akademiskt meningsutbyte.

Kapitel 5 (om värdet av humanistisk forskning) inleds med distinktionen mellan egenvärde och instrumentellt värde. Salwén konstaterar att många har uttryckt åsikten att kunskap har ett egenvärde, men redogör därefter för argument som visar på svårigheter med denna position. Därmed verkar vägen ligga öppen för åsikten att humanistisk forskning har instrumentellt värde (om den alls har värde). Salwén diskuterar i huvudsak fyra idéer om hur humanistisk forskning skulle kunna ha instrumentellt värde. Kortfattat är de:

- (a) Humanistisk utbildning och forskning spelar en positiv roll i en demokratisk samhällsutveckling. Eventuellt ger humaniora mer empati och inlevelseförmåga.

- (b) Humanistiska forskare kan bidra till förståelsen av nutida (kulturella eller samhällseliga) fenomen.
- (c) Humanistiska forskare kan bidra med expertkunskaper som är mer eller mindre tydligt samhällsnyttiga, t. ex. som underlag för politiska beslut, som underlag för myndighetsbeslut eller i juridiska sammanhang (t.ex. som expertvittnen).
- (d) Humanistisk forskning ger bidrag till den ekonomiska utvecklingen (näringslivets konkurrenskraft).

Salwén uttrycker kritik mot den sista idén (d) och även mot tanken om ökad empati (under a). Men jag får ändå sammantaget intrycket att han tycker att dessa fyra idéer är värda att ta på allvar – han ägnar 13 sidor åt dem. För egen del tycker jag att alla fyra är tämligen svaga som argument för humanioras värde. De har ett gemensamt problem, som kunde ha diskuterats. Problemet är följande: humanistisk forskning kan förväntas bidra *väldigt olika mycket* till punkterna (a)–(d). Säkert kan man finna ett antal exempel på humanistisk forskning som uppfyller någon eller några av (a)–(d), och sådana anförs också av Salwén. Under (b) skriver han t. ex. att humanistiska forskare med en mängd olika perspektiv kan bidra till förståelsen av filmen *Avatar* (2009) och dess framgång. Men samtidigt är det knappast svårt att finna humanistisk forskning som högst sannolikt *inte* nämnvärt uppfyller någon av (a)–(d). Tag ett forskningsprojekt om August Strindbergs dramer och deras receptionshistoria. Det är sannolikt att Strindbergsprojektets resultat inte har någon direkt påverkan på en demokratisk samhällsutveckling, inte hjälper oss särskilt mycket att förstå vår samtid, inte är efterfrågade i politiska eller juridiska processer, och inte ger något bidrag till den ekonomiska utvecklingen. Eftersom således viss humanistisk forskning bidrar till (a)–(d), medan annan forskning inte gör detta, så är det rimligt att fråga: bör man påstå att humanistisk forskning i största allmänhet har instrumentellt värde bara för att det finns exempel på humanistisk forskning som har sådant värde? Det förefaller inte självklart, åtminstone inte om det instrumentella värdet analyseras i termer av (a)–(d). Frågan hänger förstås också ihop med hur man ser på humanioras avgränsning och på vad som eventuellt knyter samman all humaniora (eller all vetenskap). Humanioras avgränsning behandlas dock ytterst kortfattat av Salwén (s. 14–15).

Men kanske är det också fel på analysen i form av (a)–(d)? Ja, jag tror att den är alltför grovhuggen. Alla förslagen (a)–(d) handlar om hur

humanistisk forskning skulle kunna ha värde för samhället i stort. De är därmed utomvetenskapliga. Men man kan också tänka sig att humanistisk (och annan) forskning kan ha inomvetenskapligt värde. Strindbergsforskningen skulle t. ex. kunna ha inomvetenskapligt värde genom att nå resultat som befrämjar annan forskning eller genom metodutveckling. I själva verket tror jag att det är en vanlig tanke bland forskare att forskning kan ha sådant inomvetenskapligt värde. Om vi till exempel skulle komma fram till att Strindbergsforskningen har sådant värde, kan vi sedan diskutera om detta i sammanhanget ska förstås som ett egenvärde eller ett instrumentellt värde (eller både och). Distinktionerna sammanfaller alltså inte. Min poäng är att man med hjälp av distinktionen mellan inom- och utomvetenskapligt värde kan säga fler intressanta saker i denna fråga än man kan göra enbart med hjälp av distinktionen mellan egenvärde och instrumentellt värde. (Ytterligare en distinktion som kunde vara hjälpsam är den mellan direkta och indirekta nyttor eller värden.) Kontrasten mellan inom- och utomvetenskapliga perspektiv verkar ju annars vara vanlig i forskningsetiska sammanhang (t.ex. Forsman, 1997) och kunde definitivt ha varit användbar här.

Sammantaget har jag uppskattat Salwéns bok, men om han planerar en omarbetning ser jag gärna att texten blir fylligare.¹

Jesper Jerkert

Litteratur

- Forsman, Birgitta. 1997. *Forskningsetik: En introduktion*. Lund: Studentlitteratur.
- Forsman, Birgitta. 2007. *Begrepp om forskningsfusk*. Rapport på uppdrag av Vetenskapsrådet. Tillgänglig via <www.vr.se>.
- Hermerén, Göran, Bengt Gustafsson och Bo Pettersson. 2011. *God forskningsetik*. Stockholm: Vetenskapsrådet.
- Sokal, Alan och Jean Bricmont. 1998. *Fashionable Nonsense: Postmodern Intellectuals' Abuse of Science*. New York: Picador.

1. Tack till Jesper Ahlin Marceta för värdefulla synpunkter på denna recension.

Är Gud annorlunda?

Lennart Nordenfelt

Appell Förlag, 2020. 174 s. ISBN 978-91-984960-9-3

En livslång dröm som äntligen har gått i uppfyllelse? När Lennart Nordenfelt, seniorprofessor vid Ersta Sköndal Bräcke högskola, professor i Hälso- och sjukvårdens teori och etik vid Linköpings universitet och en både produktiv och internationellt välmeriterad författare, tar sig an ett av religionsfilosofins centrala frågekomplex – *Är Gud annorlunda?* – är det i vart fall med en föredömlig nyfikenhet. I en lättläst, välskrivna, välstrukturerad och typografiskt påkostad bok presenterar och problematiserar han ett antal teorier om det religiösa språkets karaktär i allmänhet och gudsbegreppet i synnerhet. Med avstamp i Russells och Hedenius inflytelserika religionskritik, och med den samtida nyateismen som spänstigt språngbräde, ställer han upp följande fråga (se s. 31): ”Finns det en inomvärldslig gudsbild som är förenlig med den ateism som idag så kraftfullt presenterats av personer som Richard Dawkins, Christopher Hitchens, Sam Harris och Daniel Dennett och som tidigare förespråkades av de kända filosoferna Bertrand Russell och i Sverige Ingemar Hedenius?”

En inomvärldslig gudsbild? Utgångspunkten är med andra ord att de monoteistiska religionernas traditionella och övernaturliga (supranaturalistiska) gudsbilder är överspelade. Nämnade ateister, om jag tolkar Nordenfelt rätt, har framgångsrikt visat att det sannolikt inte existerar någon allsmäktig, allvetande och allgod skapare av universum. Men har religionen därmed nödvändigtvis spelat ut sin roll? Nordenfelt verkar inte vilja dra den slutsatsen. Han lyfter upp vissa religionssociologiska trender från senare årtionden som tycks visa att religionen – eller åtminstone andlighet i vidare mening – alls inte är på tillbakagång, globalt sett, trots all långtgående modernisering och sekularisering av stora delar av världen, och han verkar vilja hålla dörren öppen för att ett religiöst engagemang skulle kunna vara meningsfullt och livsbefrämjande även om teismens traditionella gud inte längre tänks existera. Det finns ju åtminstone två andra möjligheter: dels att det faktiskt existerar en metafysiskt fundamental entitet värd beteckningen ”gud”, om än inte utrustad med teismens traditionella superattribut, dels att gudsidén är praktiskt och/eller teoretiskt oersättlig vad gäller förmågan att symbolisera och medvetandegöra allt som djupast sett är värdefullt i tillvaron.

På jakt efter en inomvärldslig gudsbild, förenlig med samtida (ny)ateistiska utgångspunkter, går Nordenfelt igenom ett antal teoretiska möjligheter. Först ut är 1600-talsfilosofen Baruch Spinozas panteistiska idé att Gud är identisk med själva världssaltet. Därefter följer en intressant presentation av 1800-talets idealistiska panteism sådan den tog sig uttryck i Storbritannien i skuggan av den mer tongivande empirismen. Huvuddragen i denna idealistiska panteism har i våra dagar utvecklats inte minst av moralfilosofen Timothy Sprigge i boken *The God of Metaphysics*, vars tankegångar Nordenfelt presenterar i viss detalj.

Efter genomgången av Sprigges idealistiska panteism byter Nordenfelt spår, aningen abrupt, och diskuterar istället ”annorlunda vägar att nå kunskap om Gud” (s. 96). Eventuell kunskap om Gud förutsätter dock dels att vi vet vad vi pratar om när vi pratar om Gud, dels att denna gud existerar, och det är ju snarast dessa frågor som Nordenfelt egentligen avser att diskutera. Hursomhelst, här presenteras i tur och ordning Anders Jeffers religionsfilosofiska grundmönsterteori, enligt vilken världens samlade empiri kan tolkas utifrån ett transcenderande mönster likaväl som (eller bättre än) utifrån ett icke-transcenderande, naturalistiskt mönster, och den walesiske filosofen D.Z. Phillips fideistiska, av Wittgenstein inspirerade språkspelsteori, enligt vilken det religiösa språket har sin egen inre logik och grammatik, väsensskild från vetenskapens dito.

Att det religiösa språket kan tänkas ha en helt annan karaktär än det vetenskapliga menade även 1800-talsfilosofen och katolske prästen Bernard Bolzano, vars idéer på senare tid tydligt påverkade den svenske religionsfilosofen Eberhard Herrmann. Båda dessa tänkare underströk att vetenskaplig verifierbarhet och existentiell meningsfullhet är två radikalt olika ting. Bolzano gick till och med så långt så att han definierade religion i termer av sådant som positivt bidrar till människors lycka och moraliska utveckling. Herrmann å sin sida menade i mer försiktiga ordalag att religion handlar om att erbjuda adekvata svar på människans existentiella behov och predikament.

Avslutningsvis presenterar Nordenfelt ett antal ”anti-realistiska” teorier om Guds natur. Han ger först den anglikanske biskopen John Robinson, vars teologiskt mycket omdiskuterade bok *Honest to God* – eller *Gud är annorlunda* (!) som den heter i svensk översättning – blev en bästsäljare på 1960-talet, stort utrymme. Enligt Robinson, om jag tolkar Nordenfelt rätt, är Gud själva ”väsensgrunden” för allt i vår värld, eller kanske mer specifikt ”det som i sista hand är viktigt för oss” (s. 128). Här var Robinson tydligt påverkad av sin store läromästa-

re, den tysk-amerikanska protestantiska teologen Paul Tillich, som ju beskrev Gud dels såsom identisk med Varat självt, dels såsom det som för människan är av *ultimate concern*. I detta tankelandskap landar till slut även Nordenfelt själv:

Om vi genuint älskar en medmänniska eller livet i allmänhet, då älskar vi samtidigt Gud. Om vi känner vördnad inför tillvarons storhet, sammanhang och skönhet, då är det Gud vi känner vördnad för. Med Gud menar vi i så fall det adekvata objektet för genuin kärlek eller vördnad. Gud kan därmed finnas överallt. (s. 155-56)

Därmed inte sagt att detta är Nordenfelts egentliga slutsats eller svaret på bokens övergripande frågeställning. ”Jag försöker på allvar förstå vad Robinson och Tillich har sagt”, medger han, ”men jag gör det knappast” (s. 156).

För mitt syfte är det dock tillräckligt att säga: mot den uppfattning av Guds natur, som jag försökt skissera, kan den traditionelle ateisten knappast ha något att invända. Här finns inga supranaturalistiska antaganden. Här finns ingen osynlig verklighet som kan betvivlas. Här finns inget övernaturligt väsen. Här finns ingen Person att rikta sin bön till. Här finns bara de fenomen som ger mening till livet för de allra flesta människor. Stridsyxan mellan de så kallade troende och de icke-troende skulle kunna läggas ner! (s. 156)

Här är det svårt som läsare, åtminstone för undertecknad, att undvika en viss känsla av antiklimax. Vad Nordenfelts många noggranna, personligt hållna och befriande icke-polemiska analyser till slut mynnar ut i är alltså konstaterandet att troende och icke-troende egentligen kan enas, eller till och med redan är eniga, om vi med ”Gud” menar något i stil med ”tillvarons yttersta grund”, ”kärlekens djupaste väsen” eller ”varat självt”. Vilket vi ju alla visste redan från början.

En fråga som Nordenfelt tyvärr alls inte berör är varför ateister som Hedenius och Dawkins, vars position ju är själva utgångspunkten för bokens huvudsakliga frågeställning, överhuvudtaget skulle vara intresserade av att beskriva grundläggande mänskliga värden och fenomen, vars existens och betydelse alla kan enas om, i teologiska termer. På vilket sätt skulle upplevelsen av kärlek, eller av vördnad inför tillvarons till synes oändliga rymd, bli djupare eller mer meningsfull av att identifieras med Gud? Varför inte lika gärna – om man så vill – beskriva samma upplevelser i termer av magi eller förtrollning? Eller vad är det

i själva gudsiden som är omistligt för att på ett adekvat sätt kunna tala om kärlekens och hela existensens djupaste innebörd?

Till sist några saker som inte påverkar bokens övergripande argumentation. I samband med att Nordenfelt diskuterar en evolutionär teori om altruistiskt beteende framhålls att vi i förlängningen "delar en gen-pool med de flesta människor" (s. 55) – men vilka människor (undrar man då) delar vi *inte* en gen-pool med? I en närbesläktad diskussion menar författaren att "kraftfull kritik" kan riktas mot Darwins evolutionsteori: "Det finns svårförklarade hopp i utvecklingen. [...] Hur kom den första fågeln till? Steget från en landkrabba till en helt flygfärdig organism verkar vara mycket stort. Detta gäckar idén om utvecklingen i mycket små steg" (s. 64–65). Jag tänker att Nordenfelt egentligen inte har för avsikt att ifrågasätta evolutionsläran, så detta rör sig nog bara om olyckliga formuleringar.

På ett annat ställe, i samband med en kristologisk diskussion, skriver Nordenfelt i förbigående att Jesus anses vara "en stor profet bland både judar och muslimer" (s. 76), men det är förstas bara bland de senare av dessa två grupper som han i allmänhet anses vara en stor profet. Slutligen, i ett hypotetiskt resonemang om D. Z. Phillips språkspelsteori, tar författaren upp "Gamla testamentets uppfattning att jorden endast är sextusen år gammal" (s. 111). Jag vet ingen akademiskt verksam GT-exeget som skulle hålla med om att detta faktiskt är "Gamla testamentets uppfattning".

Martin Lembke

Litteratur

Robinson, John. 1963. *Honest to God*. London: SCM Press.

Sprigge, Timothy. 2006. *The God of Metaphysics* Oxford: Oxford University Press.

Novum Organum

Francis Bacon. Översättning Pär Svensson

Fri Tanke 2021. 272 s. ISBN 978-91-8858-991-0

Fri Tanke förlag har i en serie kallad "vetenskapliga klassiker" givit ut Francis Bacons *Novum Organum*, i översättning av Pär Svensson. Inför detta faktum inställer sig genast en fråga för varje samvetsgrann läsare av *Filosofisk tidskrift*: Ska jag läsa denna bok? Svar: Självklart! Är det inte varje förnuftig människas främsta plikt att använda den tid som blir över till att förkovra sig i det mänskliga tänkandets logik och his-

toria? Pär Svensson översättning gör ett gott intryck och Bacon är en fantastisk stilist och hans tankar är guld. Här måste jag dock inflika en liten varning: Vem är jag att bedöma detta? Sanningen är att mina kunskaper i latin inte är tillräckliga för att göra en ordentlig granskning av översättningen. Det får andra göra. Men jag kan ämnet hyfsat. Så läs denna bok! Det är definitivt bättre att läsa boken än att inte läsa den. Och vem vill väl inte läsa denna ”nya” uppföljare till Aristoteles *Organon* (av vilket *Första analytiken* finns förnämligt översatt och kommenterad av Per-Erik Malmnäs, Thales 2019).

Är nu allt väl med denna utgåva? Nej, tyvärr finns en hel del att klaga på.

En noggrann läsare vill kanske veta vilken text som har legat till grund för översättningen. Är det London-, Amsterdam- eller Parisutgåvan? Har hänsyn tagits till Bacons eget A- och B-manuskript? Är texten med eller utan censurmyndighetens ändringar? Svar: Läsaren kan vara lugn. Det finns endast en förstautgåva av denna bok (1620) att utgå från i ett seriöst arbete. Så brukar det sällan vara. Men just detta vill kanske läsaren få veta. Det hör till den elementära varudeklarationen vid översättning och utgivning av klassiska verk.

Det hör också till att åtminstone ett fåtal ord sägs av översättaren om arbetets utgångspunkter och ambitionsnivå. I detta fall är det uppenbart inte fråga om att förlaget haft ambitionen att skapa ett lärt arbete som vänder sig till de närmast sörjande, utan i förlagets anda vänder man sig till en bredare allmänhet. (Och det gillar jag. Mer Bacon till folket!) Icke desto mindre vill man veta vad man har i händerna.

Men just på denna på punkt, allmänbildningens, saknas också en hel del. Det finns ingen orientering om verkets plats i vetenskapsfilosofins historia. Det finns ett entusiastiskt förord av en medlem i Kungliga Vetenskapsakademien (märk väl: klass X!). Men Folke Tersman är inte expert på just detta, om än han är ypperlig på andra områden. Efter huvudtexten finns en odaterad, rent biografisk, text om drygt 5 sidor av litteraturvetaren Daniel Helsing. Alltså: inget om logiken eller tankens instrument, om den kritiska vetenskapsteorin och dess plats i den samtida vetenskapsteoretiska diskussionen. (Bacon var samtida med Kepler och Galileo och bör ha påverkat Descartes, Hobbes, Boyle, Locke och Newton! Hume var säkert en av Bacons efterföljare, och Reid beklagade under andra hälften av 1700-talet att Bacon tycktes bortglömd, medan alla i stället hyllade Newton; i själva verket, menade Reid, är Newtons fysik en framgångsrik tillämpning av de bästa idéerna hos Bacon. En i Sverige verksam skotsk filosof menade en gång i

samtal med mig att det antagligen inte finns något alls i John Stuart Mills *A System of Logic* [1843] som inte redan finns hos Bacon. Popper polemiserade mot Bacon [falsifikation mot konfirmation], men hade antagligen fel: ibland tycks de lika som bär när man skrapar på ytan. Och så vidare.)

Vetenskapsakademien har bokstavligen satt sin stämpel på denna utgåva och serie – på framsidan syns den. Men det är ju fel akademi! Vet man inte på Fri Tanke att det finns andra vetenskapliga akademier där nära nog alla ledamöter väl känner till hur man hanterar klassiska texter med sakkunskap och akribi? (Tips: Vitterhetsakademien.)

För dem som gillar heta debatter kan man i *Göteborgs-Posten* finna ett kort intensivt utbyte mellan Moa Ekbom och översättaren Pär Svensson, där den förra väljer ordet "haveri" för denna översättning och utgåva, och den senare försvarar sig. (En annan sak: i sitt svar säger Svensson att översättningen har fackgranskats, och det är ju bra, för att inte säga en självklarhet, att så sker. Men det framgår inte av utgåvan att den är fackgranskad och av vem.)

Trots de brister jag har funnit tycker jag inte att Ekboms ord "haveri" är en särskilt träffsäker karaktäristik. Den träffar inte, för den skjuter över målet: avsikten var aldrig att utge ett lärt arbete med textkritisk apparat och hundratals textkommentarer som vänder sig till experterna. Nej, syftet var ädlare än så: allmänbildning. För vad vi har här är en läs- värd och mycket inflytelserik text som alla borde få möjlighet att läsa. Så läs den!

Robert Callergård

*Oumbärliga samtal: Humaniora och samhällsvetenskap
i svensk offentlig debatt*

Red. Johannes Heuman, Jenni Sandström och Sten Widmalm
Daidalos 2020. 217 s. ISBN 978-91-7173-604-8

Det senaste årtiondet har frågor som rör demokratins överlevnad fått plötslig, närmast skrämmande, aktualitet. Ett av de främsta orostecken rör hastigheten med vilken förutsättningarna för konstruktiva samtal i offentligheten eroderar. Eftersom åsikts- och informationsutbyte är en så viktig komponent i alla demokratier är frågan om hur vi värderar, hanterar och förmedlar kunskap så viktig – och kanske viktigare än någonsin. Till en del kan den nu pågående balkaniseringen, som

historikern Johan Östling beskriver fenomenet i antologin *Oumbärliga samtal*, kanske härledas till de algoritmer som ligger till grund för vår tids digitala medieplattformar. Även om problemet åtminstone delvis är av teknisk natur handlar det dock ytterst om oss människor och den politiska kultur som etableras på sådana plattformar.

En brännande fråga i tider av "fake news" och "alternativa fakta" är om, och i så fall hur, kunskap kan bidra till det demokratiska samhällets fortlevnad samt i vilken utsträckning humaniora och samhällsvetenskap kan vitalisera en hårt ansatt offentlighet. Det är just dessa utmaningar som står i centrum för antologin *Oumbärliga samtal*. Till formen, om än inte till namnet, rör det sig om en vänbok till Kay Glans, journalist och författare som i sitt yrkesliv ägnat dessa frågor stort utrymme. Bokens redaktörer har kopplingar till *Respons* och Glans själv bidrar med ett efterord på temat "Vem får säga vad?". Förutom flitiga referenser till tidskriften *Respons*, som Glans var med och grundade 2012 och verkade som chefredaktör för till helt nyligen, samt hans tidigare uppdragsgivare *Svenska Dagbladet* och dess essäsida Under strecket, är boken emellertid vällovligt fri från den sortens ryggdunkningar och introspektiva grubblerier som annars tenderar att ge vänböcker ett snävt tilltal.

I redaktörernas inledning och Glans avslutning behandlas det samhälls- och debattklimat som råder inom och utanför akademien. Tyvärr upplevs dessa avsnitt bitvis som daterade och ensidiga. Glans menar till exempel att dagens radikalism främst är en revolt inom ett vänsterparadigm, medan konservatismen är påtagligt marginaliserad. Det är en beskrivning som helt tycks bortse från det senaste decenniets händelseutveckling. När Glans för många år sedan recenserade Kim Salomons bok om den svenska FNL-rörelsen påminde han om att det politiska klimatet kan förändras mycket snabbt. Det är ord som slår mig vid läsningen. Medan inledning och avslutning i *Oumbärliga samtal* huvudsakligen uppehåller sig vid den form av hot mot humanioras samhällsställning som förknippas med identitetspolitik och vänsteraktivism tycks författarna här närmast helt bortse från den auktoritära och populistiska högervåg som sveper fram över Europa och USA, och som i flera fall redan förorsakat direkta krav på att kringskära forskning, att stämpla teoretiska perspektiv eller hela ämnesdiscipliner som "aktivistiska" och att deplattformera dess utövare. Detta märks också i tendensen att forskare drar sig tillbaka från offentligheten när det personliga priset för att delta blir för högt. Medan det,

tack och lov, är relativt lätt att kritisera så kallad identitetspolitik och studenters vänsteraktivism – här finns dessutom en rik flora av fysiska och digitala publikationer som publicerar sådan kritik – utsätts forskare som granskar den nya högerradikalismen för ett sådant politiskt tryck att både offentligheten och kunskapsproduktionen blir tydligt lidande. En bidragande orsak är att dessa forskare stämplas som aktivistiska och att trycket även på etablerade medier att hålla sig neutrala blir så starkt att frågorna tystas ner. Filosofen Jason Stanley har varnat för detta i *Fascismens metoder*, men Glans ser i stället just denna bok som ett exempel på hur det offentliga samtalet utarmas när meningsmotståndare fascist-stämplas. Kanske är det just här det verkliga problemet ligger. I teorin är det rätt enkelt att skissera förutsättningarna för en välfungerande offentlighet, men demokratier kommer alltid att behöva pragmatiska lösningar för att hantera anti-demokratiska krafter – oavsett ideologisk hemvist. Ska det ske genom tystnad, motstånd, mobilisering eller förbud? Hur ska förment demokratiska krafter skiljas från de verkligt demokratiska? Dessa perspektiv, och frågor som rör hur forskarsamhället och dess företrädare ska agera under trycket från antidemokratiska krafter, saknas tyvärr i bokens inledning och avslutning, men hade bidragit till att ge boken ytterligare samtidsrelevans.

Bokens två huvudteman – offentlighetens villkor och forskarens ansvar – återkommer däremot i alla kapitel och belyses ur flera perspektiv. Frågor om forskarrollen och det ansvar som därmed följer berörs tydligast i Sharon Riders och Anna Victoria Hallbergs bidrag. Med hänvisningar till Max Weber och framför allt Karl Popper jämför Rider forskarens krav på saklighet med journalistikens konsekvensneutralitet. Det är lätt att instämma i uppmaningen att forskarsamhället ska odla den självdisciplin som krävs för att skilja mellan personlig övertygelse och sakförhållanden. Även i känsliga politiska frågor bör forskare kunna mötas för konstruktiv diskussion som skiljer på sak och intresse. Emellertid vilar denna plädering på en tydlig politisk idé om vetenskapens samhällsroll, vars liberala utgångspunkt mer utförligt hade kunnat diskuteras i relation till den eventuella politisering av universitetet som är föremål för diskussion.

Statsvetaren Sverker Gustafssons bidrag utgår istället från en historisk problembeskrivning, enligt vilken efterkrigstidens demokratiska framgångar gradvis har underminerats av *new public management* och nyliberal styrningskonst. I Poppers efterföljd pläderar Gustafsson för

fallibilism som ideal, och betonar därmed vikten av att vetenskapen vågar odla en demokratisk kultur som förstärks av komplexitet och osäkerhet, snarare än tvärtom. Vikten av att inte ge efter för krav på snabba eller enkla svar genomsyrar även Hallbergs bidrag, medan historikerna Östling och Urban Lundberg ger en empirisk inblick i vilken roll essäsidan Under strecket spelat som humanistisk kunskapsarena i Sverige. Det korta svaret är att den haft, och alltjämt har, en stark ställning som offentlig kunskapsarena för svenskspråkig humaniora. Östling påpekar att den genom sin beständighet framstår som alltmer ovanlig i ett medialt landskap där recensioner och essäer får minskat utrymme, medan Lundberg betonar att det krävs inlevelse, nyfikenhet och samtidsrelevans för att syresätta humanistisk forskning i offentligheten.

Oumbärliga samtal behandlar viktiga frågor, vars relevans knappast heller kommer att minska i en nära framtid. Det är inte boken för den som söker snabba eller pragmatiska lösningar på demokratins eller den humanistiska forskningens utmaningar, men den bidrar med perspektiv som är väl värda att belysa. Även om det finns skäl att lyfta ett varningens finger för en bitvis förenklad syn på hoten mot universitetet och den akademiska friheten, råder det ingen tvekan om att diskussionen förtjänar att tas på allvar och hållas levande. Samtalet är, i dubbel mening, oumbärligt.

Björn Lundberg

Medvetandets mysterium: En guide till universums största gåta

Annaka Harris

Översättning Tom Sköld. Fackgranskning Hans Liljenström

Fri Tanke 2020. 127 s. ISBN 978-91-8858-941-5

Annaka Harris är journalist med särskilt intresse för medvetandet och hon har för den här boken intervjuat en lång rad specialister på området. Över fyrtio experter tackas, däribland David Chalmers, Antonio Damasio, Thomas Nagel, Christof Koch och Galen Strawson. (Harris är dessutom gift med neurovetaren Sam Harris, som också har skrivit om medvetandet – speciellt i samband med andlighet, meditation och droger; han är för övrigt, tillsammans med Richard Dawkins, Daniel Dennett och Christopher Hitchens, en av ”ateismens fyra ryttare”.)

Annaka Harris bok handlar alltså om ”universums största gåta”,

nämligen hur en anhopning materia kan vara, eller bli, medveten. Man kan väl inte säga att hon ger oss något särskilt uttömmande svar. Men hon vill visa att flera av våra vanliga uppfattningar om medvetandet är felaktiga. Och speciellt vill hon slå ett slag för den så kallade panpsykismen, dvs. tanken att *all* materia har någon sorts medvetande – något som de flesta nog skulle avfärda som flum eller fria fantasier. Bokens sista kapitel – som är mycket kort och mycket svårbegripligt – handlar om relationen mellan medvetande och tid.

Harris säger att hon i den här boken använder sig av Thomas Nagels definition av ”medvetande”, nämligen:

En organism är medveten om det åtföljs av en upplevelse att vara den organismen. (s. 11)

Denna definition säger ju inte vad medvetande är, utan bara vad det innebär att vara medveten. Och bör vi verkligen utgå från att endast organismer kan vara medvetna? Vad skulle AI-entusiaster säga om det?

Kanske menar Harris helt enkelt att en anhopning materia är medveten om, och när, och endast om, den har *upplevelser*. Närmare bestämt *subjektiva* upplevelser. Man kan förstås tycka att alla upplevelser är subjektiva, men det viktiga är alltså hur något känns eller upplevs inifrån, inte hur upplevelser manifesteras eller konstitueras i yttre symptom och beteenden.

Harris anser kanske också att en organisms (eller ett annat objekts) *medvetande* är dess upplevelser – snarare än något som har eller innehåller upplevelsena. Men i själva verket spelar det nog ingen roll för Harris hur ”medvetande” ska definieras, ty den gåta boken handlar om är framför allt frågan om hur materiella objekt kan ha upplevelser. Hon menar också att våra vanliga intuitioner om vilka materiella objekt som har upplevelser är grundlösa och troligen felaktiga (s. 26). Boken innehåller för övrigt redogörelser för teorier och experiment som förekommer inom medvetandeforskning, och på vissa punkter ger Harris även egna förslag (t.ex. på s. 92 och 126).

Det finns välkända experiment – bl.a. av neurologen Benjamin Libet – som tycks visa att våra hjärnor orsakar våra handlingar *innan* vi själva beslutar att utföra dem (se s. 28f.). Det imponerar på Harris och många har tyckt att det är underligt, men egentligen är det väl vad man kan förvänta sig. Våra beslut orsakas rimligen av något som sker i hjärnan och det som sker i hjärnan har i sin tur orsaker som tidigare inträffat där eller i andra delar av nervsystemet. Och om dessa orsa-

ker kan framkalla ett beslut, så kan de också framkalla motsvarande handling. Upplevelsen av fri vilja är alltså en illusion (s. 30, 33): "medvetandet är en passiv åskådare – som följer föreställningen, i stället för att skapa eller styra den" (s. 42).

Fyller medvetandet alls någon funktion? Skulle vi i princip kunna åstadkomma alla våra yttre beteenden, utan att ha några upplevelser? Kunde världen ha varit annorlunda i enbart det avseendet att vi alla vore zombier? En som anser det är David Chalmers. Harris säger sig acceptera hans resonemang, utom på en punkt: en omedveten robot eller zombie skulle inte kunna begrunda den medvetna upplevelsen (s. 44). Men det skulle nog inte heller Chalmers påstå. Är man inte medveten, så är man ju inte medveten om något och kan därför inte heller begrunda något. Frågan är om det finns yttre beteenden som förutsätter medvetande.

Jag är inte säker på att Harris har förstått Chalmers tankeexperiment. Hon säger att en zombie är "en person som utåt sett ser ut som och beter sig likadant som alla andra men inte har några inre upplevelser" (s. 25). Harris tycks fråga sig om personer i ens omgivning ("en expedient i en butik" eller "en nära vän") skulle kunna vara en zombie. Detta är ju ett klassiskt kunskapsteoretiskt problem. Men det är inte Chalmers problem. Hans problem är metafysiskt och det gäller inte de andra, utan oss alla – och framför allt oss själva. Skulle *jag* (och alla andra) kunna sakna medvetande trots att världen i alla fysikaliska avseenden vore precis densamma? Chalmers lutar åt att det är möjligt och att detta visar att medvetandet inte är något fysikaliskt, utan att någon form av dualism måste antas. (I senare skrifter har Chalmers modifierat detta resonemang.)

Vi betvivlar ju inte att vi själva har upplevelser och vi tror väl också att många (eller kanske till och med alla) djur har upplevelser av något slag. Men kanske kan man gå längre: "Om alla de olika beteenden som djur uppvisar kan åtföljas av medvetande, varför skulle då inte växternas reaktioner på ljus också kunna göra det – eller för den delen elektronernas spinn? Kanske är medvetandet en del av själva materian, och utgör en fundamental egenskap i universum" (s. 65).

Det är detta som kallas "panpsykism". Finns det något som talar för det?

Ett argument för att medvetande endast finns i mycket utvecklade nervsystem (som en hjärna) kan bygga på hypotesen att det har uppstått som en produkt av den biologiska evolutionen (s. 79). Men såvitt jag vet har någon sådan förklaring aldrig vetenskapligt etablerats eller

ens konstruerats. Och att upplevelser skulle ha överlevnadsvärde – öka våra överlevnadschanser – förutsätter att de kan påverka organismens beteende, vilket kan betvivlas (jfr Libet-experimenten ovan).

Det finns, lite förenklat, två alternativ. Antingen finns någon form av medvetande i all materia (*panpsykism*), eller också uppstår medvetande endast i vissa, mer komplicerade materiella objekt, som nervsystem och hjärnor (*stark emergens*). Det som enligt Harris och andra talar för panpsykism är att det är det enklaste alternativet (s. 66). Ty stark emergens förutsätter ett oförklarligt hopp någonstans (var?) och på något sätt (hur?) från materia till medvetande.

Evolutionen skapar vissa *beteenden* som gynnar organismens överlevnad och reproduktion, men vilken funktion kan *upplevelser* ha? För att överleva måste en antilop vara disponerad att upptäcka och fly från ett lejon som närmar sig, men på vad sätt är det ett extra plus för antilopen (eller dess gener) att den i samma veva känner en fruktansvärd dödsångest – om det nu är det den känner? Det vore väl bättre om all dess energi satsades på ett effektivt flyktbeteende. Såvitt jag vet har evolutionsbiologin ännu inte någon förklaring till detta.

Eller tänk på fiskar. Kanske känner de smärta när de fastnar på kroken och obehag när de slängs upp på land för att långsamt dö av syrebrist. Men vilken funktion har i så fall detta lidande? Hjälper det fiskar att fortplanta sig?

Nej, det är svårt att se att upplevelser skulle ha något överlevnadsvärde och hur de skulle ha uppstått under evolutionen. Alternativt kan man då anta att upplevelser i någon form redan finns överallt i naturen, hos alla grundämnen och deras beståndsdelar. En enkel hypotes, som inte så lätt kan vederläggas.

Man kan notera att om panpsykismen är korrekt, så är förstås även en artificiell intelligens (AI) medveten. Men däremot har vi knappast någon anledning att tro att den kan vara medveten på samma sätt som vi. Kanske är dess upplevelser mer lika dem som förekommer hos en fladdermus eller en lyftkran.

Det starkaste argumentet mot panpsykism anses allmänt vara det s.k. kombinationsproblemet (s. 89f.). Detta argument, som först tycks ha formulerats av William James, utgår från att det är svårt att förstå hur två eller fler medvetanden kan kombineras till ett när motsvarande materiella föremål binds samman. Försvinner då de ursprungliga medvetandena, eller finns de kvar tillsammans med det som är resultatet av kombinationen?

Harris och neurovetaren Christof Koch har diskuterat detta och då noterat det kända förhållandet att när förbindelsen har brutits mellan hjärnans två halvor (*split brain*) så uppstår två olika medvetanden. De två halvorna har olika upplevelser. Harris och Koch tänker sig att två hela hjärnor skulle kunna kopplas ihop på motsvarande sätt, men i motsatt riktning, så att det uppstår *ett* nytt medvetande, samtidigt som de två ursprungliga försvinner (s. 91). Själv tycker jag att man lika gärna kunde tänka sig att förbindelsen mellan de två hjärnhalvorna i en delad hjärna kan återställas så att det ursprungliga medvetandet återupprättas. Det är kanske omöjligt i praktiken, men principiellt möjligt.

Harris tycker att kombinationsproblemet försvinner om man inte blandar ihop medvetande och "självet" (s. 93). Att två anhopningar av upplevelser kan kombineras – som när två hjärnhalvor kopplas ihop – verkar ju inte så konstigt. Men om de två hjärnhalvorna har varsitt jag, varsitt själv, så borde väl dessa försvinna om hjärnhalvorna kopplas ihop. Det kan ju knappast finnas tre olika jag i samma kropp. Men det skulle ju fortfarande kunna finnas olika upplevelser hos olika delar av kroppen. "Varför utesluta möjligheten att många former av medvetande existerar samtidigt inom min kropps gränser, bara för att något inte visar sig i mitt upplevelsefält?" (s. 79).

Ja, det ligger kanske något i detta. Och man kan ju undra om vi verkligen har två olika jag när hjärnhalvorna har separerats. Eller är det ett och samma jag som har två olika knippen av upplevelser? Ungefär som när man omväxlande blundar med det ena och det andra ögat. Hur avgör man det?

Harris själv tycks vilja se frågor om jaget som relativt ointressanta. Hon säger att det "själv" eller "jag" som vi ofta uppfattar som medvetandets centrum, är en illusion som kan kortslutas t. ex. i meditation eller med psykedeliska substanser som LSD (s. 49). Och att kombinationsproblemet därmed inte är särskilt problematiskt.

I mitt tycke är detta inte särskilt övertygande. Även om man med kemisk påverkan eller på annat sätt kan få människor att tappa känslan av att vara ett subjekt, så visar inte det att jaget inte existerar. Man kanske kan påverka en människa så att hon tror att hon inte har några ben, men det är ju helt förenligt med att benen ändå finns kvar och kan observeras av folk i omgivningen.

Och dessutom: om jaget bara är en illusion, varför finns den illusionen hos de allra flesta av oss? Och varför tar vi då jaget på så stort allvar? Kan jagmedvetandet ha frambringats av evolutionen? Kombinations-

problemet antyder väl att det inte finns i all materia. Varför uppstår det hos vissa organismer? Kan det fylla någon funktion?

Panpsykismen kan nog inte så lätt avfärdas, men den löser inte frågorna om jaget. Åtminstone inte än så länge.

Lars Bergström

Filosofisk merchandise

Det är nu möjligt att köpa tygpåsar med tidskriftens logotyp. Enligt vad redaktionen erfar kan dessa tygpåsar leda till avsevärda förbättringar av ägarens filosofiska argument. Mer information om hur man köper dessa påsar finns på tidskriftens hemsida, www.filosofisktidskrift.se.

Medarbetare i detta nummer av *Filosofisk tidskrift*: *Andrea S. Asker* är doktorand i praktisk filosofi vid Stockholms universitet, *Lars Bergström* är professor emeritus i praktisk filosofi vid Stockholms universitet, *Robert Callergård* är fri forskare, lärare och översättare knuten till Filosofiska institutionen vid Stockholms universitet, *Johan E. Gustafsson* är universitetslektor i praktisk filosofi vid Göteborgs universitet, *Erik Hallstenson* är doktorand i teoretisk filosofi vid Uppsala universitet, *Anders Hansson* är författare och fil. mag. i praktisk filosofi, *Jesper Jerkert* är forskare i filosofi vid Kungliga Tekniska Högskolan, *Martin Lembke* är forskare i religionsfilosofi vid Lunds universitet, *Björn Lundberg* är forskare i historia vid Lunds universitet, *Hans Teke* är universitetslektor i pedagogik vid Linnéuniversitetet i Växjö.

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. Tidskriften tar gärna emot såväl artiklar som repliker på tidigare texter i tidskriften, recensioner och notiser. Insända texter bör normalt inte överstiga 4 000 ord.

Manuskript till Filosofisk tidskrift:

- skickas med e-post till redaktion@filosofisktidskrift.se
- skall utformas i överensstämmelse med den typografi som är normal i tidskriften, utan onödig formatering
- 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
- införda bidrag honoreras inte
- författaren erhåller gratis 10 exemplar – för recensioner 5 exemplar – av det nummer av tidskriften i vilket bidraget varit infört