

Torbjörn Tännsjö

Död och personlig identitet

I otålig väntan på en diagnos av mina magplågors orsak ligger jag på Roslagstulls sjukhus. Då inte läkarna bökar med mina tarmar läser jag Mats Furbergs artikel "Om dödsbegrepp och definitioner" i *Filosofisk tidskrift* 1/87. Furberg ställer läsaren inför en utmaning. Han påstår, att det inte är möjligt att definiera dödsbegreppet, om man vill att definitionen ska vara på en gång en "realdefinition", ge vägledning i frågor om transplantationer, dödshjälp o s v, vara icke-normativ och ge en hygglig beskrivning av vårt vardagliga språkbruk. Vem vill inte göra det omöjliga? Jag behöver också något att sysselsätta tankarna med. Jag antar utmaningen!

1. Personlig död

Vad Furberg är intresserad av är problemet med mänsklig död, eller kanske mera bestämt, med vad det innebär att en *person* dör. Hans utgångspunkt är en statlig utredning, som jag inte ska bry mig om i det här sammanhanget. Jag håller med Furberg om att *den* inte frambragt någon särskilt lyckad definition av dödsbegreppet. Men hur ska det då förstås att en person dör? Svaret är generande enkelt. Det förslag jag vill framföra är att *en person dör vid en viss tidpunkt om och endast om hon vid denna tidpunkt, oavsett på vilket sätt och av vilka orsaker, oåterkalleligt upphör att finnas till.*

Rimmar detta med vardagligt språkbruk? Är denna definition religiöst neutral? Om den accepteras, hur kan då personer som tror på ett evigt liv obesvärat tala om döden? Men vänta nu, talar verkligen de som tror på evigt liv obesvärat om döden? Säger de inte, snarare än att moster har "dött", att hon har "gått hädan", att hon inte mer "är ibland oss", o s v. Min definition ger vid handen att den som tror på evigt liv måste förneka döden. Och det är just vad de tycks göra.

Men jag kan ändå ana, att Furberg inte är övertygad. En invändning

jag nog måste räkna med, är att min definition ger upphov till fler frågor än den besvarar. Också om den hyggligt skulle ansluta till vårt vardagliga språkbruk är det omöjligt att utan vidare avgöra, om den verkligen är icke-normativ. För att den ska ge vägledning i vår medicinska praxis krävs att frågetecken kring detta med personlig identitet rätas ut och det är sannerligen inte lätt att innan dess heller avgöra om den verkligen är en "realdefinition". Mitt svar på dessa tänkta invändningar är att min definition reser *de rätta* frågorna.

2. Död och identitet

Att en person dör vid en viss tidpunkt innebär att hon vid denna tidpunkt för gott upphör att existera. Det innebär alltså att frågan om den personliga döden och den personliga identiteten hänger nära samman. I stället för dödsbegreppet står vi alltså med två traditionella filosofiska begrepp om existens och identitet samt med en fråga om vad det innebär att en person har upphört att existera. Det är den senare frågan, som är den intressanta.

Begreppen om existens och identitet borde inte behöva vålla oss några bekymmer i det här sammanhanget. De är tillräckligt skarpa för den typ av diskussion som förs här. Vi kan helt enkelt nöja oss med att säga, att det är fråga om "vanlig" existens och "vanlig" identitet. Att säga att en person inte mera finns eller existerar är inte märkvärdigare än att säga att en viss byggnad är borta. Och vi kan tryggt acceptera vad vi lärt i våra elementära logikstudier om identitetsrelationen. Relationen är reflexiv: A är identiskt med A. Relationen är symmetrisk: om A är identiskt med B så är B identiskt med A. Relationen är transitiv: Om A är identiskt med B och B identiskt med C, så är A identiskt med C.

Den fundamentala frågeställningen när det gäller problemet om personlig identitet är denna. I kraft av vad är *två* rumstidliga "bitar" av en person bitar av *samma* person? Svaret på den frågan kan formuleras i termer av ett *kriterium* på när rumstidliga (i fortsättningen underförstår jag detta klumpiga uttryck) Personbitar är Bitar av samma Person, vad jag ska kalla ett PBP-kriterium. Låt oss se på ett lurigt exempel där frågan blir akut.

Derek Parfit gör i boken *Reasons and Persons* ett berömt tankeexperiment där en person "teletransporteras" från jorden till Mars. Teletransport — t ex från jorden till Mars — går normalt till så här. Personen som ska "transporteras", låt oss kalla honom Pettersson, sövs först

och kartläggs därefter noga i en registrator här på jorden. Han mals sedan ned i en köttkvarn. Mätresultatet sänds per radio till Mars där en kopian i nytt material konstruerar en exakt kopia av honom. Kopian väcks till medvetande, säger sig vara personen som sövdes ned på jorden, kan återge dennes barndomsminnen o s v.

Pettersson på jorden, före teletransporten, är en personbit. Kopian på Mars är en annan personbit. Varje sådan personbit är identisk med sig själv och med ingen annan. Men några är möjligen dessutom bitar av samma person. I fallet då teletransporten fungerar normalt har vi nog en benägenhet att säga att kopian och Pettersson på jorden före teletransporten är delar av samma person, Pettersson. Parfit menar att detta språkbruk är rimligt. Men Parfit komplicerar bilden. Antag att mekanismen i något enstaka fall klickar. Pettersson kliver in i registratorn, sövs ned och kartläggs. Radiosignalen sänds till Mars där en kopia byggs. Men Pettersson på jorden mals aldrig ner. I samma ögonblick som kopian på Mars vaknar till liv vaknar Pettersson till liv på jorden. Vad ska vi säga om den situationen? Vi står nu med tre personbitar: Pettersson på jorden före teletransporten, Pettersson på jorden efter teletransporten och kopian på Mars. Är de alla tre delar av samma person, Pettersson?

Detta kan inte säkert utslutas. Det kan vara frestande att resonera så här. Identitet är en transitiv relation. Pettersson före teletransporten kan kanske tyckas vara identisk med Pettersson på jorden efter teletransporten. Och Pettersson på jorden före teletransporten kan kanske tyckas vara identisk med kopian på Mars. Men kopian på Mars och Pettersson på jorden, efter teletransporten, är inte identiska. Därför kan inte Pettersson före teletransporten vara identisk både med Pettersson efter (den halvt misslyckade) teletransporten och med kopian på Mars. Det skulle strida mot transitivitetskravet. Men resonerar vi så har vi lättit oss förledas av vårt eget språkbruk. Normalt refererar vi till *hela* personer med deras namn. Hela personer är identiska med sig själva och inga andra. Men då vi talar om Pettersson före teletransporten, efter teletransporten och kopian på Mars åsyftar vi alltså inte hela personer utan vad jag har kallat personbitar. Det finns alltså inga *formella* problem med åsikten, att Pettersson är en helhet, identisk med sig själv och med ingen annan, som består av Pettersson på jorden före teletransporten, Pettersson på jorden efter teletransporten och Petterssons kopia på Mars. Inget i identitetsbegreppet hindrar att personer på detta sätt förgrenar sig.

Men Parfit vill inte desto mindre hävda, att personer *faktiskt* aldrig förgrenar sig. Det leder honom till den häpnadsväckande slutsatsen, att om teletransporten lyckas överlever Pettersson, han har verkligen rest från jorden till Mars med ljusets hastighet. Men "misslyckas" teletransporten så att vi får två fortsättningar av Pettersson, en på jorden och en på Mars, *dör* han. Det förefaller mig bisarrt. Men det är en följd av hans syn på personlig identitet som en fråga om psykisk kontinuitet utan förgrening.

Parfits PBP-kriterium är bara ett i mängden av föreslagna sådana kriterier. Nozick löser i mitt tycka mera elegant problemet hur vi ska utsluta att personer förgrenar sig i rummet. Hans PBP-kriterium lägger också vikt vid mentalt sammanhang men han tillför också den finurliga idén, att två successiva personbitar är bitar av samma person om den som är senare i tiden är den "närmaste" efterföljaren till den förra. Hur ordet "närmaste" ska förstås lämnas utan precisering. Det får avgöras från fall till fall. I exemplet ovan överlever Pettersson teletransport. Pettersson på jorden före transporten och kopian på Mars är delar av samma Pettersson. I fallet då teletransporten halvt misslyckas så att det finns två kontinuerliga fortsättningar av Pettersson före transporten är det den på jorden som är den "närmaste". Den och Pettersson före transporten är nu delar av samma person medan kopian på Mars är första delen av en helt ny person.

Både Parfit och Nozick tillgriper vad man kunde kalla mentala PBP-kriterier. Också andra typer av kriterier har varit i svang. Grovt kunde man skilja mellan två huvudtyper, metafysiska och empiriska. Bland de empiriska kan vi urskilja mentala och fysiska.

En anhängare av fysiska kriterier är Thomas Nagel, som anser att två personbitar är bitar av samma person om de har samma hjärna. Enligt Nagel kan vi alltså inte överleva ens lyckad teletransport. Ett annat fysiskt kriterium för Erik Ryding fram i samma nummer av *Filosofisk tidskrift* där Furbergs uppsats finns tryckt. Enligt Ryding är det åtminstone tillräckligt för att två personbitar, med var sin subjektiva jagupplevelse, ska vara bitar av samma person, att de avsatt anteckningar i samma dagbok!

Specifikt för de metafysiska kriterierna är att de gör det i princip omöjligt att empiriskt avgöra om två personbitar är bitar av samma person — och att det alltså blir empiriskt omöjligt att avgöra, om en viss person vid en viss tidpunkt upphört att existera eller ej. Specifikt för dem är också tanken att frågan om en viss person överlevt en viss för-

ändring alltid, oavsett om vi kan veta vilket, har ett bestämt svar. Anhängarna av metafysiska kriterier är beredda att göra tankeexperiment av följande typ. Om jag inte blivit född till världen men Descartes i mitt ställe blivit född av mina föräldrar och delat alla mina empiriska kännetecken, vad skulle han då ha haft att säga om dödsbegreppet? Anhängarna av empiriska kriterier skulle säga att den person som på detta sätt omtalas *är* jag.

Anhängarna av metafysiska kriterier har ofta anklagats för att i själva verket inte ha några kriterier alls, deras tänkande är konfyst, har det sagts. Att de inte ger några kriterier kan det kanske ligga något i, i så motto att de brukar hävda att personbegreppet är enkelt och oanalyserbart. Om den uppfattningen är konfys är en omstridd fråga. Anhängarna av det metafysiska kriteriet brukar för sin del göra anspråk på att företräda det sunda förnuftets ståndpunkt.

3. Realdefinition

Att dö vid en viss tidpunkt är att vid den tidpunkten för gott upphöra att existera, har jag hävdad. Nu har vi sett att den formuleringen döljer betydande oklarheter. För att förstå den måste vi veta, om inte vad en person *är* för något, så åtminstone vad det innebär för två personbitar att vara bitar av samma person. Och här tycks intuitionerna gå vitt isär. Flera olika PBP-kriterier ligger hyggligt i linje med vårt vardagliga språkbruk, men inget pekas ut framför andra av det. Och inför särskilt utstuderade fall är det osäkert om vi har några bestämda intuitioner om vilka utslag kriterierna *bör* ge. Det är här tanken om en *realdefinition* kan komma till vår hjälp. Skulle vi inte, bland alla möjliga PBP-kriterier som är hyggligt förenliga med vårt vardagliga språkbruk, kunna välja ut ett, som ger den *naturliga* innebörden åt termen "person"?

Vi kommer nog inte ifrån talet om mer eller mindre naturliga egenskaper och slag av företeelser. Det går kanske inte här att dra skarpa gränser. Allra naturligast är kanske de slags elementarpartiklar med deras begränsade uppsättning egenskaper, som fysikerna omtalar. Men också färgen grön har en viss grad av naturlighet, den är t ex naturligare än Goodmans "grue", o s v.

Paradexemplet på ett "naturligt" slags företeelse är en sådan som det görs essentiell hänvisning till i någon naturlag. Och en realdefinition, som jag här ska använda uttrycket, är en definition, som gör en naturlig avgränsning av en företeelse, t ex — och det är paradfallet — så att

det blir möjligt att på ett essentiellt sätt hänvisa till den i en naturlag. Kan personer på detta sätt ha en naturlig, reell existens?

Det finns givetvis inga fysiska lagar, som hänvisar till personer. Och det är omstritt om det alls finns psykologiska lagar — kanske kunde annars sådana referera på ett essentiellt sätt till personer. Men hur är det med de moraliska "lagarna", eller, som vi väl snarare är vana att säga, "principerna"? Finns det någon sann moralisk princip, som essentiellt hänvisar till detta med personer? Om en sådan finns har vi rätt att tala om personbegreppet som naturligt. Den princip som kunde komma ifråga skulle i så fall vara följande:

(P) För alla personer gäller, att det är något dåligt för dem då de dör, eller, annorlunda uttryckt,

(P*) För alla personer gäller, att det är något dåligt för dem då de oåterkalleligt upphör att existera.

Åsikten omfattas oreflekterat av de flesta. En filosof som försvarat att den är rimlig är Thomas Nagel. Enligt Nagel (se hans uppsats "Death") är det något dåligt, som drabbar mig, då jag oåterkalleligt upphör att existera. Det är något dåligt i sig, oavsett konsekvenserna (med ett förbehåll som ska nämnas senare). Invändningen att detta dåliga inte kan vara så illa eftersom det är något jag rimligen inte kan *uppleva* avfärdar Nagel. Det faktum att jag inte mera finns, trots att det hade varit möjligt att jag hade fortsatt att existera (och njuta livets goda, lägger Nagel till), *är* dåligt för mig — på ett tidlöst sätt. Det är en olycka som drabbar mig och alla människor. Ju förr den drabbar en människa dess värre, förutsatt att hon i annat fall fått leva ett fortsatt liv värt att leva. Det är själva detta faktum, som är dåligt, inte bara att därmed en del goda upplevelser går om intet (vilket ju i princip kan kompenseras av att andra människor eller jag själv *medan jag lever* får en förhöjd livskvalitet). Att olyckor som drabbar oss kan vara genuina, trots att de inte upplevs av oss, är inget märkvärdigt, menar Nagel. Det är t ex en olycka om andra förtalar oss i det fördolda, även om vi inte själva någonsin får kunskap om det. Att slippa det kan vara värt en viss minskning av livskvaliteten. *Hur* dåligt det är att dö, försöker inte Nagel avgöra. Men det är alltså värre ju tidigare i livet döden drabbar oss.

Jag har svårt att ta ställning till om Nagel har rätt eller ej, men jag håller det i vart fall inte för *uteslutet* att han har rätt. Hans ståndpunkt har ett kraftfullt stöd i det sunda förnuftets tänkande på det här områ-

det. Åsikten är *försvarlig*. Men den är aningen obestämd. Ty vad *innebär* det att dö? Vad innebär det för en person att upphöra att existera? I en mening är svaret enkelt att formulera. Att en person vid en viss tidpunkt upphör att existera innebär att det inte vid någon senare tidpunkt finns några personbitar som är bitar av henne. Men hur avgör man om så är fallet? Märkvärdigt nog diskuterar inte Nagel själv den frågan, trots att han, som nämndes ovan, i *andra* sammanhang försvarar en bestämd typ av PBP-kriterier. Han sätter dem emellertid aldrig i samband med frågan om värdet av att leva vidare. Men här kan ju lösningen av problemet med hur en realdefinition av döden ska se ut stå att finna. En realdefinition av döden är en definition, som gör P* sann. Och även om aldrig Nagel själv ställt frågan på det sättet bör vi alltså söka efter de PBP-kriterier, som gör P* så rimlig som möjligt. Finner vi dem får vi en realdefinition av dödsbegreppet på köpet.

4. Vad är det som gör det till en olycka att dö?

Vad är det vi fruktar, då vi fruktar att dö? Är det att det inte ska finnas någon framtida personbit, som delar vårt *metafysiska* jag? Rimligen inte, eftersom det inte finns några empiriska möjligheter att avgöra om det är fallet. Vem vet, kanske förlorade jag vid gårdagens rektoskopering mitt metafysiska jag. Det bryr jag mig inte om!

Kan det vara att det inte ska finnas någon framtida personbit, som delar *kropp* med denna min personbit, som sträcker sig i tiden fram till nu? Tja, det kan i så fall inte gälla *hela* kroppen. Om jag erbjuds nya inälvor, som garanterat inte kommer att krångla, har jag just nu inget emot en transplantation. Men en hjärntransplantation ser vi under alla omständigheter som hotfull. Det hänger samman med vår insikt om hjärnans betydelse för de mentala förloppen.

Är det då alltså dessa, som är av betydelse? Det hävdar Parfit. Också Parfit menar, att vi har viss anledning att bekymra oss om vår fortsatta existens. Och jag tror inte att han då *bara* tänker i instrumentella termer, det är inte *bara* fråga om, att en framtida personbit som är kontinuerlig med oss som vi är nu kan vara ett bra instrument när det gäller att få vissa önsksningar om framtiden uppfyllda, som vi har nu. Han tycks också tänka sig att det har egenvärde för oss att vi har framtida personbitar, som är kontinuerliga med oss mentalt. Mera bestämt bör det, för att vi ska ha anledning att känna oss tillfreds, finnas framtida personbitar, vilka är kontinuerliga med oss och som parvis hänger ihop

med hjälp av minnen och avsikter. De ska bilda kontinuerliga sekvenser där du idag minns mycket av vad som hände i går, i morgon minns mycket av vad som hände i dag och en del av vad som hände i går, i övermorgon minns mycket av vad som kom att hända i morgon, en hel del av vad som hände i dag, och åtminstone något av vad som hände i går o s v. Också genomgripande förändringar i vad man minns, vill och strävar efter o s v är acceptabla, men de bör inte inträffa alltför snabbt och omvälvande. Detta är "vad som i grunden spelar roll", hävdar Parfit, ifråga om "vår omsorg om vår egen framtid". Att hysa en sådan omsorg är åtminstone inte irrationellt.

Mot en sådan åsikt kan givetvis den hedonistiska utilitaristen invända, att det är orimligt att man skulle offra något som helst ifråga om sin egen eller andras välfärd bara därför att någon person, man själv eller någon annan, skulle få leva längre. Hur kan det vara rätt att på detta sätt åstadkomma onödigt lidande eller en onödigt dålig balans av lycka och lidande i världen som helhet? Men anhängaren av P* kan ju svara, att detta lidande eller denna *icke maximala* balans av lycka och lidande i världen *inte alls* är onödig. En persons liv räddas ju! Balansen av lycka och olycka kan därmed mycket väl visa sig vara *optimal*.

Jag hävdar inte att denna ståndpunkt är sann, men den förefaller mig försvarlig. Det finns inga dräpande argument emot den — lika litet som det finns dräpande argument emot den utilitaristiskt hedonistiska.

Enligt Parfit vållar hans ställningstagande till vad som är viktigt för oss då det gäller omsorg om vår framtid problem med själva personbegreppet. Ty det slags psykiska samband han menar är viktigt för oss kan förgrena sig över tiden, som vi såg i teletransportexemplet. Men personlig identitet kan inte på detta sätt förgrena sig. Det kan inte under någon längre tid förekomma två spatialt åtskilda personbitar, utan inbördes kommunikation, med var sin utvecklingshistoria, som är bitar av samma person. "Förgrening" av en person är alltså liktydigt med döden men likväl intet att frukta. Hur kan en dubbel succés vara ett fiasko? Det är väl så bra att i framtiden ha *två* personer, var och en psykiskt sammanhängande med mig, som tar vid där jag slutar, som att bara ha en, hävdar Parfit. Men är den ståndpunkten rimlig?

Ser vi det hela rent instrumentellt är den givetvis rimlig. Medan en framtida personbit tar sig an mina barn kan en annan sköta min undervisning, en tredje tar sig an mina politiska uppdrag o s v. Men det vi diskuterar här är något annat, det värde det har för oss att fortsätta att

existera *oavsett* konsekvenserna (förutsatt att fortsättningen innebär ett liv över en viss nivå, ett liv värt att leva). Är det i *det* sammanhanget rimligt att tänka sig, att vi kan vara lika nöjda med att få *två* efterföljare, som en? Jag tror inte det.

Vi bör inte låta oss förledas på den här punkten av Parfits teletransportexempel. I fallet då vi får *två* överlevande, Pettersson på jorden och kopian på Mars, tror jag Nozick har den rimligare språkliga intuitionen. Här bör vi tala om Pettersson på jorden som fortsättningen av Pettersson före teletransporten. Kopian på Mars är en *ny* individ. Pettersson har verkligen överlevt den misslyckade "transporten". Pettersson på jorden efter transporten är den *närmaste* efterföljaren till Pettersson på jorden före transporten. Och det är en sådan Pettersson alltid är angelägen om att ha. Det Pettersson har anledning att frukta *är* död, men inte i Parfits utan i Nozicks mening.

Vi kan emellertid konstruera mer genuina exempel där personer tycks "förgrena" sig, exempel där vi inte kan se några skäl att säga, att någon av en uppsättning framtida personbitar är *den* närmaste och mest självklara "fortsättningen" av en viss person, men där alla hänger hyggligt samman med honom (så att vi inte skulle tveka att säga att han hade överlevt, om det förelegat en enda fortsättning av honom, godtyckligt vilken). Vad ska vi säga om sådana fall? Utgör möjligheten av dem ett hot mot oss? Jag tror att de gör det. Om det *alls* ligger något i P* måste den förstås som så, att det är dåligt för oss, inte bara om det inte finns *någon* personbit i framtiden, som hänger psykiskt ihop med oss på lämpligt sätt, utan också, och lika dåligt, om det inte finns en *unik* sådan. Att det förelåg flera skulle innebära en i bokstavlig bemärkelse "upplösning" av vår person.

Om man kör iväg mig till operationsavdelningen, delar upp min hjärna i olika delar, transplanterar dem i lämpliga kroppar o s v och lyckas väcka alla till liv med påföljden, att de alla är ett slags deformerade fortsättningar av mig, ser jag det som slutet för min del; den olycka det möjligen innebär att dö har drabbat mig. Men behåller man *en* av hjärndelarna i *min* kropp, så att jag, lika deformerad som de andra ovan omtalade personerna, vaknar upp i "min" säng, ansluts till "min" droppflaska här på rum A4, vill jag hävda, att jag förvisso kan ha råkat illa ut i operationen i olika hänseenden, men *en* olycka har ändå besparats mig — jag har inte blivit tagen av daga. I min egen kropp, min egen sjuk-säng o s v finns en erforderlig grund för att skilja ut *en* framtida persondel, som den unika fortsättningen av *mig*.

5. Sammanfattning

Slutsatsen av det ovanstående är inte så kategorisk, som jag kanske hade önskat. Jag inser att jag inte utan vidare kan sägas ha gjort vad Furberg sade vara omöjligt. Men det är *möjligt* att jag har gjort det "omöjliga". Jag har gjort det, om det sunda förnuftet har rätt, då det, i strid med bl a den hedonistiska utilitarismen, ansluter sig till P^* .

Låt oss se på de krav Furberg ställde på en dödsdefinition, de krav han menade vara omöjliga att uppfylla. Den skulle anknyta till allmänt språkbruk, vara icke-normativ, ge vägledning i frågor om transplantationer och dödshjälp och den skulle vara en realdefinition. De kraven uppfyller min definition.

Min definition av vad det innebär att en person vid en viss tidpunkt dör är att hon vid denna tidpunkt oåterkalleligt upphör att existera. Detta ska förstås så att det efter denna tidpunkt inte finns någon unik personbit, som hänger samman med henne psykiskt (på ett sätt som jag antytt mera i detalj ovan).

Den dödsdefinition jag givit ansluter hyggligt till allmänt språkbruk. Inga normativa uttryck ingår i definitionens. Tveklöst bör den kunna ge vägledning om vi till exempel vill utforma kriterier på när respiratorer ska få stängas av o s v. Och, till sist, om P^* är en rimlig normativ uppfattning, kan min definition med fog sägas vara en realdefinition. Den avgränsar dödsbegreppet på ett "naturligt" sätt, nämligen så att P^* , tanken om döden som en olycka, blir så rimlig som möjligt. \square