

Erik Ryding

Parfit och den personliga identiteten

I sin uppsats "Proust och personlig identitet" (FT 1:1986) gör Hans Mathlein i anslutning till Derek Parfit en distinktion mellan absolut och relativ identitet. Och enligt hans mening är personlig identitet av den senare, graderbara typen. Emellertid tillägger han: "Eftersom vi ständigt förändras fysiskt och psykologiskt är det egentligen inte *identitet* vi är ute efter, när vi försöker ta ställning till om vi har att göra med samma person över tiden." Det kan verka en smula förvirrande — den relativa identiteten är egentligen inte alls någon identitet. Några konsekvenser för Mathleins resonemang får denna terminologiska inadvartens nu knappast, men den tycks mig ha intresse som en reflex av den oklarhet som hos Parfit råder kring identitetsbegreppet.

I *Reasons and Persons*, Parfits huvudarbete, heter det att personlighetens enhet framförallt beror på två minnesrelationer. 1) Vi har direkta ihågkomster av upplevelser i det förflutna. Det mesta är visserligen bortfallet men vissa händelser erinrar vi oss. 2) Det sträcker sig en sammanhängande kedja av minnen från varje tidpunkt i det förflutna (med undantag för de tidigaste barnåren) och till nuet. Förmodligen minns jag inte nu vad jag var med om en viss dag för trettio år sen, säg den 12 augusti. Men den 13 augusti 1956 kom jag ihåg en mängd ting jag upplevt föregående dag och den 14 augusti en rad upplevelser från den 13e etc. Denna relation kallar Parfit psykisk kontinuitet.

Av de båda minnesrelationerna sägs den första — den direkta erinringen — vara viktigast både från praktisk och teoretisk synpunkt, men det är i alla fall kontinuiteten som åstadkommer den personliga identiteten. Motiveringen är att identitet utgör en transitiv relation, och kontinuitet men inte direkt erinring är transitiv. Om dag ett är kontinuerligt förbunden med dag två och dag två med dag tre, så måste också dag ett och dag tre ha denna typ av förbindelse. Däremot kan jag komma ihåg händelser, som inträffade en viss dag för tio år sedan utan att för den skull nu komma ihåg, vad jag den dagen mindes från ännu tidigare ske-

den i mitt liv (s 205 f).

Mot bakgrunden av dessa förklaringar måste man emellertid känna sig frågande inför kommentaren till ett tankeexperiment som Parfit utför några sidor längre fram. Han antar att en nervkirurg efter behag kan manipulera en persons minnen och karaktär och har bestämt sig för att göra sitt offer till en andlig avbild av Napoleon. Förvandlingen sker i en mängd omgångar, så att individen efter varje ny behandling bara är ett litet stycke närmare målet än efter den föregående. Parfit förklarar nu att den manipulerade — vi kan kalla honom A — under behandlingens första fasen utan tvekan fortsätter att vara A men vid dess slut lika tveklöst inte är det. Någonstans under vägen upphör den ursprungliga personen att existera, fast vi inte kan säga *exakt när* lika litet som vi kan säga när en sandhög som berövas ett korn i taget upphör att vara en hög. Vid mitten av proceduren finns det över huvud taget inte något bestämt svar på frågan, om A fortfarande är i livet eller ej (s 231 ff).

Framställningen är inte helt klar. Den kan tolkas *så*, att kirurgen från varje uppnått stadium i metamorfosen *direkt* fortsätter till nästa, eller *så*, att han efter varje behandlingstillfälle återställer offret i ursprungligt skick och alltså var gång startar om från utgångsläget. Enligt den första tolkningen kommer det att råda kontinuitet genom hela proceduren. Eftersom varje enskilt ingrepp medför en obetydlig förändring, kommer försökspersonen att på varje stadium minnas åtskilligt av vad han tänkt, känt och förnummit på det närmast föregående. Då Parfit likväl säger att A inte existerar i slutfasen, så innebär det att han här övergett kontinuiteten till förmån för den direkta erinringen, när det gäller att fastställa, vad som är en och samma person; att den manipulerade individen i slutfasen inte längre sägs vara A måste framförallt bero på att han inte har kvar några av A:s minnen.

Enligt den andra tolkningen bryts däremot kontinuiteten mot slutet. Individen förs i ett språng från sin begynnelsepersonlighet till Napoleonstadiet och blir därmed en ny person.

Emellertid är ju slutprodukten exakt densamma vare sig den ena eller andra metoden har använts; likväl skall den enligt kontinuitetskriteriet i det ena fallet vara identisk med A, i det andra vara en ny människa. Läkaren tar livet av A genom att mellan turerna låta honom återvända till sin forna personlighet. Intuitivt verkar slutsatsen svår att acceptera.

En granskning av vad minneskontinuiteten innebär ger också vid handen att den är otillräcklig som grund för personens enhet. Den inte

bara tillåter så stora förändringar att det inte finns några hågkomster som är gemensamma för olika tidsfaser; den stämmer också dåligt med vad som avses, då man tänker på sig själv som samma person från födelsen och till nuet. Den uppfattningen skulle inte rubbas, om t ex en chock raderade ut de närmaste tolv timmarna ur minnet. Att man upphört att existera under den "borttappade" tiden skulle betraktas som ett bisarrt påstående.

Över huvud taget förefaller det mig svårt att förstå, vad Parfit menar med "personlig identitet". Det kan ju inte vara fråga om strikt identitet, eftersom en människa aldrig är exakt lik från minut till minut. Men inte heller är det bara ett annat namn för minneskontinuiteten. Detta framgår av ett annat av de många tankeexperiment, som fyller *Reasons and Persons*. En individ antas bli fördubblad genom att dela sig som en amöba. Båda de sekundära individerna är beträffande minnen, karaktär, attityder etc direkta fortsättningar av den ursprungliga, odelade, men, sägs det, ändå inte identiska med honom, eftersom de inte är identiska med varandra och identitet är en transitiv relation. Något mer än blott kontinuitet skall den personliga identiteten alltså innebära — men vad? Det enda vi får veta är, att den dessutom skall vara transitiv. Detta är ju självklart i fråga om den strikta identiteten. Om det för varje egenskap gäller att a har den, om och endast om b har den, och samma förhållande råder mellan b och c , så måste också a och c ha fullständig egenskapsgemenskap. Men någon motsvarande överförbarhet följer inte ur de sammanhang som skall konstituera den personliga identiteten. Två företeelser kan båda vara kontinuerligt förbundna med en tredje utan att vara det sinsemellan. Mina ben är kontinuerligt sammanhängande med min bål utan att för den skull hänga samman med varandra. Fordran på transitivitet kommer alltså som ett fristående tillägg, för vilket Parfit inte ger någon motivering.

Gissningsvis har han inte velat helt släppa termen "identitet", eftersom de allra flesta människor har en djupt rotad föreställning om att de är samma varelser livet igenom, och sedan har *ordet* dragit med sig transitivitetskravet. Kanske medverkar också förnimmelsen av att de sammanhang, som enligt *Reasons and Persons* håller samman personligheten, inte kan motivera vår normala attityd gentemot framtiden. Det heter att man i stället för "Jag kommer att vara död" bör säga "Det kommer inte att finnas några framtida erfarenheter, som på visst sätt är relaterade till mina nuvarande", och i stället för "Jag kommer att lida" bör säga "Det kommer att finnas lidande, som på visst sätt är relaterat

till mina nuvarande erfarenheter” (s 281 f). De relationer som åsyftas är givetvis minnesrelationerna. Skillnaderna mellan att jag kommer att drabbas av plågor och att en främmande gör det, skulle alltså vara att offret i det ena fallet har minnen i första person av en del av mina nuvarande upplevelser och i det andra inte är utrustad med sådana. Det är svårt att förstå, varför man under dylika omständigheter speciellt skall frukta smärtor som drabbar en själv, helst som man — vilket Parfit på ett annat ställe medger — vid starka akuta plågor knappast lär syssla med att minnas sitt förflutna.

Förhållandet kan belysas med ett tänkt fall: En man får det dystra beskedet att han inom en snar framtid skall utsättas för svåra fysiska plågor och att han när smärtan släppt kommer att finna att han har förlorat minnet. Ifall Parfit har rätt, kommer tidpunkten för minnesbortfallet att bli avgörande för frågan, *vem* det är som råkar ut för lidandet. Inträffar den förändring i hjärnan som utplånar minnet i början av plågoperioden eller först vid dess slut? I det senare fallet är den drabbade samma person som fick meddelandet, i det förra någon annan. Visserligen aktualiseras hur som helst inga minnen, eftersom medvetandet är helt fyllt av stundens pina, men man är densamme, om man skulle kunna göra det, dvs om minnesmekanismen är intakt. (Inte heller i normala fall aktualiserar vi ju ständigt personliga minnen.) Emellertid blir upplevelsen precis densamma, vare sig den plågade skulle kunna minnas eller ej, vilket betyder att det för den som får beskedet är likgiltigt, om det blir han själv eller en annan person, som utsätts för lidandet.

Kanske skulle Parfit vilja svara, att vi inte *har* något skäl att för egen räkning frukta kommande olycksöden (och naturligtvis då inte heller att glädja oss åt behagligheter, som väntar i framtiden). Människorna skulle alltså, så långt man kan följa vittnesbörden tillbaka, i onödan ha bekymrat sig för sin egen framtid.

Det kan vara frestande att utan vidare avvisa en åsikt, som så radikalt strider mot vår naturliga inställning och som inte har någon chans att bli tillämpad i praktiken — man vågar nog med bestämdhet påstå att inga filosofiska teorier om personlighetens identitet kommer att få folk att sluta upp med att hysa ett alldeles särskilt intresse för sina privata framtidsutsikter. Emellertid kan även en outrotlig attityd bygga på en illusorisk grundval, och om man förkastar Parfits teori, bör man kunna ange något alternativ. Finns det alltså någon annan faktor än minnesrelationerna, som gör det rimligt att säga att den som disponerar en viss

kropp är en och samma person vid olika tidpunkter?

Några filosofer såsom Bernard Williams och Thomas Nagel vill göra själva den fysiska kontinuiteten (hela kroppens eller hjärnans) till kriterium på personlig identitet, men jag skall här inte diskutera detta förslag, mot vilket jag tror att avgörande invändningar kan riktas, utan skall i stället se på situationen från upplevelsesidan.

Hur vet en person att det är just han som i det innevarande ögonblicket känner smärta eller erfar vällust eller har en viss tanke? Frågan verkar besynnerlig eller åtminstone onödig. I och med att något uppträder som direkt erfarenhet är det inte möjligt att betvivla att man själv är den som har denna erfarenhet. Minnesrelationerna spelar här ingen roll. Jag behöver inte uppväcka några minnen för att kontrollera att det verkligen är i mitt medvetande som en förnimmelse eller tanke uppträder, för mitt medvetande vid en given tidpunkt är = det som direkt upplevs vid denna tidpunkt. Även den som drabbas av en amnesi, som eliminerat alla personliga minnen, kan säga t ex "Jag har ont". Jagbegreppet och begreppet "direkt upplevelse" hänger samman. I enlighet med detta kan vi i stället för "Jag kommer att känna smärta" säga "Det kommer att finnas en direkt upplevelse av smärta". Invändningen att andra ju kan känna smärta utan att jag gör det, kan undgås genom att man bygger ut satsen till "Det kommer att finnas en direkt upplevelse av smärta, även om alla organismer förintas med undantag för den, vars hand just har skrivit orden "har skrivit orden". Hänvisningen till kroppen bör emellertid bara uppfattas som ett enkelt sätt att ange den individuella medvetandeströmmens avgränsning. Det är ett empiriskt faktum att, såvitt vi vet, varje medvetande är förbundet med en organism och inget med mer än en. Men den kroppsliga anknytningen följer inte analytiskt ur begreppet "medvetande". Och det väsentliga är här att det förekommer separata medvetandeströmmar, dvs serier av direkta upplevelser (förnimmelser, tankar, känslor etc).

Vad jag försöker visa på kan också illustreras på ett annat sätt: Antag, att en persons minne bara sträcker sig över en dag, när det gäller hans privata upplevelser. Varje morgon när han vaknar, är alla personliga hågkomster utplånade. Men i gengäld skriver han varje kväll en dagbok, där han noggrant noterar allt han upplevt under dagen. Anteckningarna skulle då ge en bild av medvetandeströmmen, och min tes är att fast den minneshandikappade inte kommer ihåg något av det han tidigare skrivit ner, så är det i alla fall samma person, som dag för dag skildrar sina tankar, känslor etc. Att han kommer att uppleva smärta

blir, förutsatt att han för protokollet ärligt, ekvivalent med att det kommer att stå en uppgift om en smärtupplevelse i dagboken, och får han veta att han nästa dag skall utsättas för plågor, har han inte mindre skäl att känna fruktan än personer med normal minneskapacitet.

Givetvis kan minnesspannet tänkas reducerat ner till den gräns, där det bara omfattar enstaka upplevelser.

Jag tror att Parfit utan att göra saken klar för sig använder de individuella upplevelsevärldarna om inte som identitets- så åtminstone som disparitetskriterium. I det förut nämnda "amöbafallet" förutsätts som någonting självklart att de sekundära individerna är olika personer, ehuru de till en början har identiska minnessammanhang; båda är ju kontinuerligt minnesförbundna med ursprungsindividerna. Denna deras inbördes självständighet får, eftersom Parfit inte är anhängare av ett rent fysiskt identitetskriterium, förklaras med att de har olika medvetandeströmmar: den ene är inte direkt medveten om vad den andre tänker, känner och förnimmer.

Den enklaste, om också kanske inte enda möjliga, förklaringen till medvetandenas avgränsning från varandra är att våra upplevelser är relaterade till olika iakttagande subjekt, och det skulle i så fall vara dessa, som utgjorde den yttersta grunden för personernas enhet. Introspektivt går visserligen inte några sådana centra att upptäcka, men detta är, som Broad framhåller, inte något argument mot deras existens. Att något iakttages skulle enligt denna teori betyda att det står i en viss asymmetrisk relation till subjektet, och ingenting kan stå i en asymmetrisk relation till sig själv (*The Mind and its Place in Nature*, s 586).

Med det sagda har jag naturligtvis inte velat bestrida, att minnesrelationerna har stor betydelse för att hålla samman personen. Utan dem — framför allt utan de direkta minnena — skulle vi inte kunna göra oss någon bild av oss själva, inte tillägga oss några individualiserande egenskaper. Man skulle även i sina egna ögon bli en anonym varelse — fortfarande ett jag men inte en personlighet.

Litteratur

Broad, C D: *The Mind and its Place in Nature*, London 1949.

Mathlein, H: "Proust och personlig identitet", *Filosofisk tidskrift* 1:1986.

Parfit, D: *Reasons and Persons*, Oxford 1984.

Williams, B: *Problems of the Self*, Cambridge 1973.