

# Från hedoniskt värde till moraliskt moras

*Ingmar Persson*

## 1. Introduktion

I ett tidigare nummer av *Filosofisk tidskrift* (nr 3, 2023) skisserade jag min syn på förhållandet att det är dåligt i sig att känna smärta och gott i sig att känna välbehag för den individ som känner dessa förnimmelser med utgångspunkt i en kritisk diskussion av Derek Parfits syn på samma förhållande i *On What Matters* (2011). Den smärta och det välbehag vi bägge har i åtanke antas ogillas respektive gillas av subjektet (vilket jag för övrigt inte betraktar som en begreppslig sanning). Jag instämmer i Parfits påstående: ”Det är vårt hedoniska gillande och ogillande [...] som gör dessa medvetande-tillstånd [av välbehag och smärta] bra/goda eller dåliga” (2011, vol. 1, s. 3 och s. 55; alla översättningar är artikelförfattarens egna). Men jag avviker från honom genom att ge en *subjektivistisk* analys av de hedoniska attitydernas ”värdegörande” av sensationerna. Jag vill också betona att det rör sig om värde *för dem* som erfar sensationerna, ett *individ-relaterat värde*.

Därför skrider jag här vidare till att generalisera denna analys av hedoniskt individ-relaterat värde till en analys av *intrinsikalt* individ-relaterat värde i allmänhet. En komplikation dyker upp när denna typ av analys ska utvidgas till att gälla även individ-relaterat värde *på det hela taget* eller *totalt*, vilket tillika innefattar vad som är av *extrinsikalt* värde beträffande ett objekt i kraft av dess relationer till andra objekt som exempelvis är konsekvenser av objektet. Men denna komplikation utesluter inte att principiellt samma slags subjektivistiska analys är giltig för individ-relaterat värde såväl totalt som intrinsikalt. Det är en fördel ur ett systematiskt eller teoretiskt perspektiv om en enhetlig analys av alla former av individ-relaterat värde är möjlig.

Ett annat krux på vår färd är steget från individ-relativt värde till *absolut* värde, varav moraliskt värde är ett exempel, om tilltro fästs till en förhärskande intuition. Mitt förslag är att absolut värde kan alstras ur värde *för alla*. I så fall rör det sig inte heller här om artskilda typer av värde, som ett subjektivistiskt och ett objektivistiskt värde eller ett naturalistiskt och ett icke-naturalistiskt värde. Men i den subjektivistiska

tolkning som förespråkas här, finns inga absoluta värden och således inga absoluta värden av moraliskt slag ity inget har värde för alla. Därmed glider vi ned i något som jag liknar vid ett *moraliskt moras*, men detta är ett öde som vi inte skulle undgå fast det existerade absoluta moraliska värden i objektivistisk mening, emedan det uppenbarligen skulle vara diskutabelt vilka dessa värden är.

## 2. Individ-relativt värde

Parfit påstår att "de viktigaste användningarna av 'god' and 'dålig' är [...] skälsimplicerande" (2011, vol. 1, s. 39), och det är denna användning som är aktuell här, för "vi har alltid skäl att vilja undvika intensiv smärta (*agony*)" (s. 31). Att denna användning av "dålig" är skälsimplicerande innebär att när smärta sägs vara något intrinsiskt dåligt, sägs smärtan vara en egenskap som ger oss skäl att vilja undvika den för dess egen skull. Eftersom det är "vårt hedoniska gillande och ogillande [...] som gör dessa medvetande-tillstånd goda eller dåliga", måste det vara så att dessa tillstånd förser oss med skäl genom att vara föremål för vårt gillande eller ogillande. Dessa hedoniska attityder är däremot inte själva baserade på skäl (s. 53).

Denna tankegång skapar emellertid krångel för Parfit, ty han invänder mot teorier vilka uppfattar skäl som vilje-beroende (*desire-dependent*) att, likmätigt dessa teorier, måste viljeattityder i slutändan "få sin normativa kraft från någon viljeattityd eller målsättning som vi inte har något skäl att ha, och sådana viljeattityder eller målsättningar kan inte försvarligt påstås ge oss några skäl" (2011, vol. 1, s. 91). Men hedoniskt gillande och ogillande kan vi ju enligt hans egen utsago inte ha skäl för, vilket dock inte hindrar att han betraktar dem som skälsskänkande.

Nu vill han visserligen inte kalla hedoniska gillanden och ogillanden viljeattityder (*desires*), men jag argumenterar i min tidigare uppsats (FT nr 3, 2023) för att detta är ett misstag: de är ju trots allt attityder som driver oss att förlänga förnimmelser av välbehag och förkorta förnimmelser av smärta och som blir tillfredsställda om vi lyckas härmed och otillfredsställda om vi misslyckas. I så fall inställer sig följande problem: om hedoniska gillanden och ogillanden kan göra välbehag och smärta till något för oss gott respektive dåligt i sig samt tillhandahålla skäl för att förlänga välbehag respektive förkorta smärta, fastän dessa gillanden och ogillanden inte kan understödjas av skäl, varför kan då inte andra sorters viljeattityder göra detsamma även om de inte kan vara skälsba-

serade? Slutsatsen måste bli att det inte kan vara det förhållandet att en viljeattityd omöjligen är skälsunderstödd som gör den oförmögen att bli skälsgrundande, som berövar den "normativ kraft". Jag ska strax försöka förklara vad som i stället för skälsunderstöd kan göra en viljeattityd skälsgrundande.

Eftersom Parfit anser att förhållandet att smärta är något intrinsikalt dåligt betyder att smärtan förser oss med *skäl* att hysa aversion mot den genom att uppväcka ogillande, så förnekar han att smärtans intrinsikala dålighet helt enkelt *består i* att ogillas (för sin egen skull). Trots detta förfäktar han att i "varje möjlig värld skulle smärta vara dåligt i sig" (2011, vol. 2, s. 489), givet att smärtan ogillas i den. Han insisterar på att smärtans dålighet "inte följer av meningen av orden 'dålig' och 'smärta'" (s. 490). Hans avståndstagande från metaetisk naturalism implicerar att "inga sådana normativa sanningar kan vara analytiska" (s. 490). Inte desto mindre hävdar han att vi kan ha kunskap om dessa nödvändiga sanningar *a priori*, dvs. "enbart genom tänkande" (s. 490) och oberoende av sinneserfarenhet och introspektion.

Med "analytiska" sanningar i trängre bemärkelse kan avses konceptuella sanningar som är *självklara*, vars sanning är så uppenbar att man blir övertygad om den så snart man begriper sanningarnas innebörd, tautologier som att en helårsprenumeration är en prenumeration, men i sin förnekelse av metaetisk naturalism måste Parfit förläna "analytisk" en vidare betydelse som innefattar konceptuella sanningar vilka är så svårfångade att de kan inses vara dylika först efter ingående analys. Om han emellertid brukar "analytisk" på detta vidare sätt, förefaller det mig mysteriöst hur vi kan veta *a priori* att det är sant i varje möjlig värld att en smärta som ogillas för dess egen skull är dålig i sig, fast det inte är analytiskt eller i vidare mening konceptuellt sant att vad som ogillas är intrinsikalt dåligt. Lyckas vi begripa att detta är sant genom rent tänkande, måste det väl vara konceptuellt sant eftersom tänkande sker med rent konceptuella medel. Är det däremot konceptuellt sant att (ogillad) smärta är dålig för oss, blir det begripligt att detta kan vara nödvändigt sant och *a priori* bekant för oss, men det sker till priset av att Parfit trots allt hamnar i värdenaturalismens garn.

Han gör, som påpekats, gällande att den viktigaste användningen av värdetermer är "skälsimplicerande", definierbar i termer av skäl. Antag att man i stället accepterar en icke-naturalistisk teori som går ut på att "god" och "dålig" betecknar *odefinierbara* icke-naturliga egenskaper; skulle detta vara ett botemedel för trubblet Parfit trasslar in sig i? Den

mest kända talesmannen för denna typ av teori är G. E. Moore som förfäktade att termen "god" betecknar något enkelt och odefinierbart, så att försöka definiera den med hänvisning till naturliga egenskaper är att göra sig skyldig till ett "naturalistiskt misstag" (1903, s. 9–10).

Denna teori råder dock inte bot på nämnda problem, för relationen mellan denna oreducerbara värdeegenskap och naturliga egenskaper kan inte vara en kontingent relation som vi upptäcker *a posteriori* eller erfarenhetsmässigt, ty i så fall äventyras återigen värdeegenskapens icke-naturliga status. Den måste vara en nödvändighetsrelation som är känd *a priori*, men som inte är konceptuell, ty detta skulle innebära återfall i naturalism – just precis den knipa som Parfit sitter i. (Senare ska jag dessutom påpeka att vi inte upplever tings värde som en inneboende egenskap hos dem.)

Det är brydsamt att hitta en värdeutsaga vars sanning vi kan vara säkrare på än den att det är i sig dåligt för oss att känna pina. Ifall vi accepterar en naturalistisk analys av denna utsaga enligt vilken förhållandet att pina eller stark smärta är i sig dåligt för oss att förnimma inte betyder annat än att vi ogillar att förnimma sådan smärta för dess egen skull, har vi en förklaring till denna epistemiska säkerhet, för vi har lättillgänglig kunskap om vad vi ogillar (och gillar) för dess egen skull. Om å andra sidan Parfit har rätt i att det relevanta värdeförhållandet implicerar existensen av ett icke-naturligt, objektivt eller vilje-oavhängigt skäl, kan vi inte vara rättmätigt säkra på att det är dåligt i sig för oss att bli pinade, för vi kan inte vara så säkra på att det finns något sådant icke-naturligt, objektivt skäl. Detta är i högsta grad metaetiskt kontroversiellt. (Dessa kritiska synpunkter ingår i min granskning av Parfits teori om praktiska skäl i Persson 2013, kap. 12.1 och 12.2.) Dock skulle det vara högst förmätet att ge sken av att denna argumentation konklusivt vederlägger metaetisk objektivism, men den uppmuntrar oss inte desto mindre att styra kosan annorstädes.

Jag har framkastat hypotesen att intrinsikala viljeattityder kan vara skälsgrundande, fastän de enligt vilje-baserade teorier inte kan vara skälsbaserade eftersom hedoniska gillanden och ogillanden, som också är en sorts viljeattityder, kan vara skälsgrundande, trots att de inte är skälsbaserade. Mitt förslag är att intrinsikala viljeattityder kan vara skälsgrundande därför att de kan vara *stabilt sanningsbaserade* precis som hedoniska attityder är. Ogillande av smärta är en viljeattityd som är baserad på eller utlöst av hur en smärtförnimmelse faktiskt upplevs och som är inriktad på att få förnimmelsen att upphöra, medan gillande

av välbehag är baserat på eller utlöst av hur en välbehagsförnimmelse faktiskt upplevs och är inriktat på att förlänga den. Det finns ingen ytterligare information att införskaffa om dessa förnimmelers natur som kan underminera dessa viljeattityder.

En vilja att undvika en *hotande* smärta för dess egen skull kan naturligtvis inte vara baserad på hur smärtan *känns för närvarande*, men den kan vara baserad på en åskådlig trosföreställning om hur den *kan komma* att kännas. En sådan trosföreställning, förvärvad genom tidigare upplevelser av smärta, kan ge praktiskt taget lika pålitlig information om den hotande smärtans karaktär som en aktuell förnimmelse av den. En intrinsikal aversion mot hotande eller annalkande smärta, som skulle ogillas när den anländer, kan därför också vara en stabilt sanningsbaserad viljeattityd som ingen ytterligare information kan rubba.

Detta är fullt förenligt med att information om denna smärtas *externa relationer*, som dess konsekvenser, kan tillkomma, information som gör det rationellt att, *allt taget i beaktande*, välkomna eller vilja erhålla smärtförnimmelserna, men detta undergräver inte motviljan mot smärtan *för dess egen skull*. Min ståndpunkt är således att intrinsikala viljeattityder kan vara skälsgrundande därför att de kan vara stabilt sanningsbaserade och att man kan definiera intrinsikal godhet och dålighet för oss i termer av vad som satisfierar eller frustrerar sådana viljeattityder (eller andra "pro-attityder" som önskan, glädje etc.) hos oss.

Mina exempel på intrinsikal godhet och dålighet har varit de hedoniska förnimmelserna av välbehag och smärta. Enligt *hedonismen* är sådana positiva känslotillstånd av välbehag, njutning, lycka, tillfredsställelse etc. och negativa känslotillstånd av smärta, lidande, olycka, otillfredsställelse etc. *det enda* som är intrinsikalt gott respektive intrinsikalt dåligt för oss. Jag förutsätter inte att denna doktrin är korrekt, men dessa känslotillstånd är de minst kontroversiella exemplen på vad som är intrinsikalt gott och dåligt för oss, och de är därför lämpade att tjänstgöra som illustrationer av individ-relaterade intrinsikala värden.

Som kontrast till situationen för intrinsikala viljeattityder är stabil sanningsbaserad en mer problematisk företeelse angående viljeattityder som tar i beaktande inte bara hur ett objekt är i sig utan också dess externa relationer, för då kan man sällan vara säker på att ytterligare information inte kommer att flyta in och förändra viljeattityden. Om man upptäcker att ett objekt O som man hyser en aversion mot för dess egen skull, har en konsekvens som man vill ha för dess egen skull, kan detta medföra att man sammantaget vill ha O. Därefter kan man bli varse att

O har en annan konsekvens som man har en motvilja mot för dess egen skull eller att den första konsekvensen i sin tur har en konsekvens gentemot vilken man har en motvilja för dess egen skull. Då kan en motvilja mot O allt taget i beaktande uppstå. Det finns ingen ände på hur vitt eller långt man på dessa sätt kan spåra ett objekts externa relationer. Detta är ett problem som i praktiken blir mer akut i den samtida världen där människan har i sin hand en välutvecklad teknologi som möjliggör global påverkan vilken till på köpet sträcker sig långt in i framtiden.

I fråga om ett förhållandes värde på det hela taget – dess *totala* värde – får man följaktligen som regel hålla till godo med att hävda att det *temporärt framstår* som ägande ett visst sådant totalt värde och att denna bedömning kan vara i behov av revision i ljuset av ny information. I denna bemärkelse blir värdeomdömen gällande vad som totalt är gott eller dåligt för någon i praktiken individ-relativa inte bara med avseende på individers vilje-attityder utan också med avseende på deras epistemiska situation: givet att alla fakta som är relevanta för subjektets viljeattityd rörande O är som de nu framstår för subjektet, är O nu gott/dåligt totalt sett för subjektet.

Att vi här får ge oss till tåls med denna dubbelsidiga individ-relativitet förebådar den kommande kapitulationen inför omöjligheten att etablera att något har värde för alla och därigenom absolut värde vilket, enligt en allmänt rådande uppfattning, moraliska värden besitter. Det sistnämnda innebär att vi hamnar i ett moraliskt moras som, vi ska se, inte kan undvikas även om det skulle existera absoluta moraliska värden i objektivistisk betydelse emedan det inte skulle vara okontroversiellt vilka dessa värden är.

Det kan invändas att sakförhållandet att vi ogillar en smärta för dess egen skull inte utesluter att smärtsensationen också är utrustad med någon egenskap som vi gillar för dess egen skull. Därför garanterar förhållandet att vi ogillar en smärta för dess egen skull bara att den *såttillvida* är intrinsikalt dålig för oss, inte att den är intrinsikalt dålig för oss på det hela taget. Det är helt riktigt att denna omständighet är logiskt möjlig, men det är ingen realistisk möjlighet. Smärtsensationer har en del andra kvaliteter än sin smärtsamhet: de kan vara brännande, stickande, skärande etc. och de har en viss kroppslig placering och varaktighet. Detta är emellertid inga kvaliteter för vilka vi gillar dem för deras egen skull. Så, även om det strängt taget måste tillfogas att en för sin egen skull ogillad smärta inte har någon annan egenskap för vilken den är gillad för sin egen skull för att det ska följa att smärtan är intrin-

sikalt dålig för oss allt taget i beaktande, kan detta tillägg betraktas som underförstått. Följaktligen är omdömen om intrinsikalt värde inte epistemiskt problematiska så som omdömen om extrinsikalt värde och totalvärde är.

Slutligen bör det understrykas att ett objekts totalvärde för någon härrör från det intrinsikala värdet för vederbörande av objekt som står i relationer som understödjer det förstnämnda objektets totalvärde, såväl interna relationer som att vara en del av objektet som externa relationer som att vara en effekt av objektet. Därav följer att den subjektivistiska analys som har givits av värde av det intrinsikala slaget också måste äga giltighet för totalt och extrinsikalt värde. Detta framgår tydligast när ett objekts värde totalt är *identiskt* med dess intrinsikala värde eller egenvärde eftersom objektet inte står i någon relation till något externt objekt som förser objektet med extrinsikalt värde och på så vis spår på dess totala värde.

### 3. Skäl för trosföreställningar och för viljeattityder

Det är lärorikt att jämföra hedoniska gillanden och ogillanden med omdömen som är direkt baserade på sinneserfarenhet, som t. ex. omdömet att något ser rätt ut för mig som är baserat på mitt synintryck av en tomat. Detta omdöme rättfärdigas av synintrycket, men ett sådant rättfärdigande består inte i *skälsgivande*: att något ser rätt ut för mig ger mig inte ett skäl att tro att något ser rätt ut för mig. Ett skäl för ett omdöme är något man kan anföra till stöd för omdömets sanning, men jag kan inte anföra att något ser rätt ut för mig till stöd för att omdömet att något ser rätt ut för mig är sant. Ett sinnesintryck är inget man kan *anföra*; man kan inte *resonera* från ett sinnesintryck till ett omdöme. Det är förvisso korrekt att säga att skälet till att jag tror att något ser rätt ut för mig är att något ser rätt ut för mig, men "skälet" betyder här förklaringen. Det är inte fråga om den *rättfärdigande* användningen av termen enligt vilken skäl är stöd av varierande styrka som kan vägas mot motstridiga skäl i en beslutsprocess.

Rättfärdigande skäl kan vara stöd för att tro något eller för att vilja och handla. Vi saknar rättfärdigande skäl för hedoniska gillanden och ogillanden precis som vi gör för sinnesbaserade omdömen, men precis som det finns skäl eller förklaringar till att vi fäller sinnesbaserade omdömen, finns det skäl eller förklaringar till att vi hyser hedoniska attityder. Den intressanta skillnaden är att medan det finns något som i

avsaknad av rättfärdigande skäl kan rättfärdiga sinnesbaserade omdömen – nämligen den sinneserfarenhet som gör dem sanna – finns det ingenting som kan *rättfärdiga* hedoniska gillanden och ogillanden. Det finns ingenting som *rättfärdigar* att vi reagerar med ogillande snarare än gillande på smärta, ingenting som gör det *rätt* att reagera med ogillande och *fel* att reagera med gillande. Man kan visserligen hänvisa till att det skulle ha *dåliga konsekvenser* för oss ifall vi reagerade med gillande på smärta – eftersom smärta signalerar att något inte står rätt till i vår kropp – men detta implicerar inte att det är fel att gilla smärta *för dess egen skull*.

Denna differens mellan trosföreställningar och viljeattityder har sin källa i att de har olika "anpassningsriktning" (*directions of fit*): trosföreställningars funktion är att anpassa sig till en utomstående verklighets fakta och därigenom bli sanna – det är den evolutionära förklaringen till att vi har dem; viljeattityder har motsatt funktion, att aktivera oss till handlingar som får fakta att anpassa sig till attitydernas objekt, så utfallet blir som vi vill. Då inställer sig andra pro-attityder som tillfredsställelse eller glädje om kännedom om utfallet föreligger. Enligt meta-etisk objektivism är dock denna kontrast mellan trosföreställningar och viljeattityder inte så tydlig, för den antar att det finns något som viljeattityder har att anpassa sig efter: objektiva värden eller normer. Hur denna anpassning ska förstås undandrar sig min fattningsförmåga: hur kan det vara en a priori vis, nödvändig sanning att det är passande att vilja något om detta inte är något som begreppsligt kräver att vara velat?

Jag vill alltså göra gällande att intrinsikala viljeattityder kan vara skälsgrundande och definiera vad som är i sig gott och dåligt för oss under förutsättning att de är stabilt sanningsbaserade, fastän detta inte rättfärdigar dem – vilket inte innebär att de är "orättfärdiga", dvs. att mot dem kontrasterande attityder *kan* rättfärdigas. Det följer att en skälskonstruktion som vilar på intrinsikala viljeattityder vilar på en vacklande grund. Denna vacklan förstärks av det tidigare nämnda faktumet att vi sällan kan vara säkra på vad vi bör vilja på det hela taget; sådana bedömningar bär ofrånkomligen det preliminära prägel. Då har vi det bättre förspänt ifråga om trosföreställningar i så måtto som sinnesbaserade trosföreställningar kan rättfärdigas på basis av sinnesintryck.

Men denna sinnesverifiering lämnar oss med en väldigt inskränkt kunskapsfär, ty den överväldigande delen av våra trosföreställningar är "sinnesöverskridande", dvs. de sträcker sig bortom vad vi varseblir för

tillfället. Våra trosföreställningar om det förflutna är till syvende och sist grundade på minnet, våra trosföreställningar om framtiden på induktiva extrapolationer från vad vi tidigare har varseblivit och våra trosföreställningar om en fysisk verklighet oavhängig av vår perception på inget annat än just vår perception. Det är kunskapsteoretiskt knivigt att begripa hur skälen för dessa sinnesöverskridande trosföreställningar i sista hand kan undgå att sakna rättfärdigande, antingen på grund av att skälen för dem slutligen blir hängande i luften eller på grund av att de rör sig i cirkel. Detta ger ammunition åt skepticism med avseende på minnet, induktionen och den fysikaliska verkligheten i helhet.

Det är ur epistemologisk synvinkel beklagligt om våra skäl för dessa trosföreställningar i slutändan är grundlösa, ty om de stod på fast grund, skulle de kunna bibringa oss en försäkran om att vi har påträffat sanningen. Denna försäkran är emellertid ingenting som nämnvärt skulle kunna förstärka vår redan stenhårda övertygelse om de sinnesöverskridande trosföreställningarnas sanning (se Persson 2013, kap. 11, för vidare diskussion av detta problem). Detta är en stenhård övertygelse som vi säkerligen delar med högre stående djur som över huvud taget inte förmår ställa sig spørsmålet om dessa övertygelsers rättfärdigande.

Vi bör hålla denna vidare skälsvy som inkluderar epistemologi för ögonen när vi sätter oss till doms över praktiska skälskranglighet. Det är otvivelaktigt av metaetiskt intresse ifall det existerar objektiva viljeoberoende praktiska skäl, men utifall sådana lyser med sin frånvaro, är det mindre skakande än att skälen för sinnesöverskridande trosföreställningar ytterst hänger i luften, ty på det praktiska fältet finns ingen sanning att fånga. Objektiva praktiska skäl skulle kunna vara betydelsefulla i praktiken ifall de både skiner som i neon, dvs. tydligt låter sig avläsas, och är så vittomfattande specifika att de ger vägledning i en majoritet av livssituationer. De exempel på objektiva skäl som vanligen presenteras tillmötesgår dock långt ifrån dessa krav – exempel i stil med att vi har skäl att dra oss undan smärta för dess egen skull och att låta bli att tortera spädbarn för vårt höga nöjes skull.

Det finns både naturalistiska och icke-naturalistiska teorier om objektiva värden och skäl, alltså både teorier som behandlar värden och skäl som reducerbara och som icke-reducerbara till naturliga eller empiriska egenskaper av de slag som vetenskap och beprövad erfarenhet umgås med. Jag har försökt visa att såväl Parfits som Moores icke-naturalistiska teorier misslyckas med att klargöra objektiva värden eller skäls natur, och jag är inte bekant med någon naturalistisk teori

som bättre lyckas med detta konststycke än icke-naturalistiska teorier. Ingen objektivistisk teori presenterar värden eller skäl av ett slag som kan utöka vår praktiska vägledning och leda oss till en mer omfattande viljemässig enighet. Uppfattningen att något är ett objektivt värde eller skäl är mer av ett "släp" som följer på observationen att det redan föreligger vidsträckt viljemässig samstämmighet än en "dragkrok" som drar oss samman i sådan endräkt. Parfits paradexempel att det finns skäl att vilja undvika smärta för dess egen skull illustrerar detta fenomen till perfektion.

Han levererar bland annat ett "pino-argument" (*agony argument*) mot subjektivistiska eller vilje-baserade teorier om praktiska skäl: vi har skäl att vilja och försöka undvika all framtida pina, men dessa teorier förnekar att vi har sådana skäl; alltså är dessa teorier falska (2011, vol. 1, s. 76). Det stämmer att enligt dessa teorier kan vår orubbliga intrinsikala aversion mot den pina som hemsöker oss inte vara skälsbaserad, men vilken roll spelar detta? Skälsbaserad skulle näppeligen kunna öka styrkan av vår aversion som redan är orubbligt stark. Inte heller är det klart hur den skulle kunna bidra till att denna aversion erhåller ett bättre eller mer auktoritativt rättfärdigande.

Svagheten i Parfits teori framträder i sitt klaraste ljus om vi tittar på pina som vi redan utsätts för. Denna pina väcker ett hedoniskt ogillande av pinan som vi har konstaterat är ett slags motvilja mot pinan för dess egen skull. Enligt Parfit ger detta ogillande oss ett skäl att hysa en "meta-hedonisk" (2011, vol. 1, s. 54) viljeattityd gentemot pinan. I klartext betyder detta att vi har skäl att *dubblera* vår nuvarande motvilja mot pinan. Detta förefaller, för det första, motivationellt oekonomiskt och överflödigt. Det är rimligare att tänka sig att den meta-hedoniska viljeattityden endast uppväcks i de av Parfit fokuserade situationerna i vilka den hedoniska förnimmelsen och ogillandet av den inte är lokaliserade till nuet. Den uppväcks då av den anledningen att den åskådliga föreställningen om smärta, härstammande från tidigare förnimmelser av smärta, framkallar en motvilja i likhet med den som själva förnimmelsen kan utlösa.

För det andra kan denna vilje-dubbling inte skänka oss ett rättfärdigande av motviljan mot pinan som vi i förstone saknade av ett förhållande som Parfit själv pekar på: skälet för den meta-hedoniska motviljan mot pinan antas härröra från den hedoniska motviljan mot pinan, men det skulle vara mysteriöst om den förra därigenom erhåller ett rättfärdigande som den senare själv saknar, det skulle vara som att

lyfta sig själv i håret. Det skulle också vara egendomligt om det uppkom ett annat slags dålighet, en "skälsimplicerande", objektiv dålighet ur den subjektiva, hedoniska dåligheten. Enligt mitt synsätt rör det sig här bara om ett slags dålighet, av subjektiv art.

#### 4. Absolut och moraliskt värde

Parfit talar om smärtans och lidandets "dubbla dålighet": dels är de i sig dåliga för de varelser som upplever dem till följd av att de ogillar hur de känns, dels är det moraliskt dåligt att de varelser som känner smärta och lidande upplever något som är i sig dåligt för dem (2011, vol. 2, s. 565). Jag har i föregående sektioner undersökt den förra föreställningen, att det är i sig dåligt att känna smärta för dem som känner den. Lidande är något som en sådan smärta kan framkalla hos de som känner den, exempelvis kan vi lida av den smärta vi förnimmer när vi bränner oss. Sådant lidande är i sig dåligt för oss både på grund av att smärtan som är lidandets objekt är ogillad och på grund av att lidandet är en känsla av otillfredsställelse som också den är ogillad.

Den fråga jag nu vill ställa är ifall den subjektivistiska analys av den individ-relativa dåligheten och godheten av något som utvecklats på grundval av en analys av den hedoniska dåligheten och godheten av smärta och välbehag för dem som känner dem kan generaliseras till en analys av dålighet och godhet som inte är individ-relativ utan absolut, som exempelvis moralisk dålighet och godhet vanligen anses vara. Ett nära tillhands liggande förslag är att absolut värde kan genereras via individ-relativt värde *för alla* som i sin tur garanterar absolut värde. Resonemanget kan då uppdelas i två moment: (1) en analys av vad som har (positivt eller negativt) värde för alla i termer av vad som är föremål för allas (positiva eller negativa) viljeattityder, och (2) en tes att det som har värde för alla har absolut värde.

Till stöd för (2) kan man påpeka att om något har värde för alla, kan ingen utan att vända sig mot sig själv opponera sig mot att det tillskrivs värde. Värdeobjektivister måste däremot insistera på att ett objekt O inte behöver vara gott, trots att O är gott för alla, inklusive dem själva – ett för mitt sinne kontraintuitivt påstående. Däremot är det inte lika övertygande att påstå om det som *inte* har värde för alla att det inte kan ha absolut värde: är det verkligen ogörligt att utpeka någon undergrupp sådan att om något har värde för alla dess gruppmedlemmar, så har det absolut värde, oaktat att det inte har värde

för icke-medlemmar av gruppen? Dyliga grepp verkar dock dömda att brista i normativ neutralitet, precis som andra strategier att garantera vilje-konvergens som senare ska beröras. Enligt mitt förmenande kan man däremot tryggt skriva under på (3) att det som inte har värde för någon inte har absolut värde.

(1) kan förtydligas medelst följande hypotes: (A) om det finns en typ av situation i vilken alla, givet att de har stabilt sanningsbaserade viljeattityder rörande naturliga förhållanden om O, vill ha O, så är O gott för dem alla i denna situation. Enligt (A) är alltså förhållandet att O är föremål för en viss subjektiv attityd hos alla en tillräcklig betingelse för att O är något gott för dem alla och därigenom, på basis av (2), något absolut gott. Tillsammans med (2) utesluter (A) att absolut godhet är något objektivt, något som är oberoende av subjektiva attityder.

Som följeslagare till (3) kan man också hävda att förhållandet att O är föremål för en sanningsbaserad viljeattityd hos en individ som är bekant med O är en nödvändig betingelse för O ska kunna vara gott för vederbörande: (B) om det finns någon typ av situation i vilken ingen som har stabilt sanningsbaserade viljeattityder rörande naturliga förhållanden om O, vill ha O, så är O inte gott för någon i denna situation och därigenom, enligt (3), inte något absolut gott. I motsats till (B) måste värdeobjektivisterna hävda att O ändå skulle kunna vara gott för någon, fastän ingen med stabilt sanningsbaserade viljeattityder vill ha O – återigen något som tycks mig kontraintuitivt.

Problemet med (A) är att det är högst tvivelaktigt om det finns något objekt som alla, i händelse av att de innehar stabilt sanningsbaserade viljeattityder om objektet, vill ha. ((B) dras däremot inte med motsvarande problem: det finns förmodligen saker och ting som ingen vill ha i relevanta omständigheter.) Det är väsentligt att problemet med (A) inte gäller om sanningen av dess försats beträffande viljeattityder garanterar att dess eftersats är sann – att sanningen av (A) dras i tvivelsmål – utan att problemet är att sanningen av dess försats är i högsta grad dubiös. För vad som då är underkastat tvivel är om det existerar något av värde för alla och därigenom, givet (2), något av absolut värde. Tvivlet angår inte om existensen av något med detta värde följer av att det finns något som har värde för alla definierat i termer av deras konvergerande viljeattityder, vilket implicerar att absolut värde inte är en speciell ontologisk kategori artskild från subjektivistiskt individ-relativt värde. Om folks stabila sanningsbaserade viljeattityder faktiskt kan bli unisona eller ej, är ingenting som utgör någon *ontologisk* skillnad, en skillnad

som den mellan vad som är subjektivt eller objektivt eller den mellan vad som är naturligt eller icke-naturligt.

Det är emellertid tänkbart att det vore en ontologisk skillnad som utgjorde förklaringen till att individ-relativt värde inte kan definieras i termer av individers viljeattityder eller att det inte följer att något har absolut värde av att det har värde för alla. Det konstaterades i slutet av sektion 2 att svårigheten att definitivt etablera att något har totalt värde för någon kan bero på svårigheten för vederbörande att definitivt eller slutgiltigt vilja något totalt sett till följd av epistemiska tillkortakommanden. Nu ser vi att svårigheten att etablera att något är av absolut värde kan bero på svårigheten att finna något som har värde för alla.

Antag att ett fysiskt objekt O är källa till ogillad smärta för alla. Antag vidare att inget annat faktum angående O är av positivt eller negativt värde för någon. O är då, allt taget i beaktande, extrinsikalt dåligt för alla. Denna dålighet härstammar från den intrinsikala dåligheten av den smärta som O är upphov till hos envar. Enligt ett resonemang som fördes i slutet av sektion 2 följer av detta att den subjektivistiska analys som värdet av den intrinsikala sorten där bestods med kan utsträckas till O:s dålighet extrinsikalt och totalt för alla och, enligt (2), därigenom till absolut värde. Begrunda också klassen av alla de smärtförmimmelser hos olika individer som O är upphov till. Denna klass är dålig i sig för alla dessa individer i kraft av att den innehåller något som är dåligt i sig för var och en av dem. Därigenom är klassen absolut dålig – i en mening som är förenlig med subjektivism – och enighet kan råda om att det finns moraliskt skäl att eliminera O. Det är emellertid högst otroligt att det förekommer något som i likhet med O vållar alla smärta.

Det är instruktivt att jämföra relationen mellan att något har värde och att det framkallar en viss viljeattityd med relationen mellan att något har en färg och att det ser ut som om det hade denna färg. Föreställningen att fysiska objekt har färg oberoende av vår perception av dem inställer sig mer enträget hos folk i gemen än föreställningen att de har värde oberoende av våra känslö- och viljeattityder. Anledningen till detta är, gissar jag, att vi visuellt varseblir färger som *inneboende* egenskaper hos ting, egenskaper som de besitter oberoende av andra ting, som oss och vår perception, i motsats till exempelvis egenskapen att vara belyst när vi riktar en ljuskälla mot dem. Som kontrast upplever vi inte värde som en inneboende egenskap hos ting; vi upplever inte t. ex. en förmimmeses dålighet så som vi upplever dess smärtsamhet, som en inneboende egenskap hos förmimmesen. Vad vi upplever är i stället

att tinget framkallar en känslö- eller viljeattityd hos oss. Denna relation till oss är självfallet ingen inneboende egenskap hos tinget som vi kan frestas tro att det äger oberoende av oss.

Förhållandet att det intuitivt föreligger en intim relation mellan att något är av värde och att det finns skäl att vilja ha det – ett förhållande som vi har konstaterat att Parfit utnyttjar – bestyrker också att värde har att göra med en viljeattityd till ett objekt snarare än med en inneboende egenskap hos det. Alltså, ett objekts värde är inte intuitivt en inneboende egenskap hos objektet eftersom det har en intim relation till skäl att vilja eller välja objektet, vilket uppenbarligen inte är en inneboende egenskap hos objektet. (Detta leder till att den objektivistiska tesen snarare får bli något i stil med att det är ”passande” att hysa en viljeattityd gentemot vissa objekt – en tes som redan har befunnits vara problematisk.)

Låt oss nu titta på förhållanden som bidrar till att göra det osannolikt att det finns något objekt som alla med stabilt sanningsbaserade viljeattityder har samma viljeattityd gentemot och som är relevant för frågan om moraliskt värde kan vara absolut. Eftersom människor i allmänhet är uttalat *partiska* till förmån för sig själva och närstående individer som familj och vänner, divergerar deras viljeattityder allt som oftast. För att stävja denna egoism och framtinga konsensus, har moralfilosofer föreslagit diverse opartiskhetskrav som en viljeattityd måste tillfredsställa för att kvalificera sig som moralisk, bland annat har John Rawls föreslagit att viljeattityden måste anläggas bakom en ”okunnighetens slöja” som döljer vilken position man själv intar i valda utfall, och Richard Hare har förespråkat införandet av ett universaliserbarhetskrav som, grovt uttryckt, går ut på att moraliska agenter måste kunna acceptera de viljeattityder de intar om de lever sig in i de omständigheter som alla de varelser som påverkas av viljeattitydernas realiserande befinner sig i.

Om vi för resonemangets skull antar att hedonismen är riktig och alltså att vad som är intrinsikalt gott för varelser är att de känner välbehag och vad som är intrinsikalt dåligt för dem är att de känner obehag, så skulle en moral kunna proklamera att ett utfall är bättre i så måtto som varelser erfar mer välbehag i det och moraliskt sämre i så måtto som de erfar mer obehag i det. I förstone kunde man inbilla sig att det skulle vara möjligt att nå konsensus om något opartiskhetskrav konformitet med vilket skulle medföra att alla skulle vilja iscensätta utfall med så stor övervikt av välbehag över obehag som möjligt, med andra ord vad en viss form av utilitarism anbefaller. Om ett krav omöjliggör

favorisering av speciella individer som det egna jaget och nära och kära, kan det ge sken av att stödja en opartiskhet som leder till sådana utfall. Exempelvis skulle en sådan opartiskhet medföra att alla som vill att deras eget obehag upphör förbinder sig till att vilja att obehag upphör hos envar som känner sådant obehag. Detta är en viljeattityd som alla som är förmögna att förstå dess objekt skulle kunna tänkas hysa.

Hare insåg emellertid tidigt att hans universaliserbarhet inte nödvändiggör utilitarism. Bland annat diskuterade han i *Freedom and Reason* (1963, s. 162) möjligheten av fanatiska nazister vilka skulle vara villiga att acceptera att de själva fick utså obehag i stället för välbehag i händelse av att de visade sig vara av judisk börd. Det är förvisso logiskt möjligt för någon att favorisera eller missgynna varelser som är utrustade med vissa *universella* egenskaper, oavsett om deras egen identitet är sådan att de exemplifierar dessa egenskaper eller ej. Detta ledde till att han gav sig in på olika försök att förstärka sitt universaliserbarhetskrav, så att det önskade utilitaristiska utfallet skulle bli oundvikligt. Inget av dessa försök är dock lyckat (som jag bemödar mig att visa i Persson 2005, kap. 26 och 27).

Ett principiellt dilemma lurar här. Antingen är ett föreslaget opartiskhetskrav på viljeattityder för att de ska kvalificera sig som moraliska alltför svagt för att nödvändiggöra viljeattityders konvergens eller, om det är tillräckligt starkt för att säkra en dylik konvergens, i t. ex. utilitaristisk skepnad, kan kravet misstänkas för att brista i normativ moralisk neutralitet, för att vara infört i syfte att kratta manegen för en viss moral snarare än som ett krav som alla moraliskt kompetenta oberoende normativ hållning måste erkänna, kanske för att det är inbäddat i begreppet moral. Kravet kan då naturligtvis inte *rättfärdiga* en viss normativ moralisk hållning.

Slutsatsen blir att vi går ner oss i ett moraliskt moras: intuitivt kan vi tänka oss att vad som är moraliskt gott eller rätt och riktigt är bindande för alla som är i stånd att begripa vad det vill säga att handla moraliskt, men hur en sådan absolut moral ska kunna grundas gäcker oss. Vi är ur stånd att rädda oss undan en moralisk relativism som går ut på att vad som är moraliskt rätt och riktigt måste förstås som vad som är moraliskt rätt och riktigt *för någon* och variabelt från person till person, beroende på att personernas sanningsbaserade viljeattityder spretar åt olika håll och att enighet inte kan nås om något opartiskhetskrav kraftfullt nog att tygla dem till konvergens.

Denna viljemässiga divergens är dock inte så omfattande som den

divergens som *faktiskt* föreligger, ty den faktiska divergensen är till stor del betingad av att folks viljeattityder inte är sanningsbaserade utan avhängiga av kunskapsluckor och av falska religiösa eller politiska ideologier eller andra grasserande fördomar som ofta frodas i dessa luckor. Dessa källor till divergens är oberoende av om subjektivistiska eller objektivistiska värdeteorier är korrekta. Det vore övermodigt att tro att metaetisk objektivism konklusivt kan vederläggas, men även objektiva moralföreskrifter som är bindande för alla skulle existera, är det, som anmärkts, odiskutabelt att de inte står klart skrivna för oss som i neoljus.

I *praktiken* spelar det därför ringa roll om subjektivism eller objektivism bildar sanningen i värdeteorin; oavsett vilket får vi finna oss i att traska i ett moraliskt moras av djup oenighet där våra moraluppfattningar ytterst avgörs av uppväxtmiljö och personlighetsfaktorer snarare än av kunskap och förnuft. Det ligger en hel del sanning i författaren Jonathan Swifts anmärkning att argument aldrig förmår en människa att rucka på en felaktig åsikt som hon inte har förvärvat på grundval av argument – och vår moral grundläggs i barndomen.

Man kan spekulera i om framväxten av den gängse uppfattningen av moralen som något absolut som alla med moralisk kompetens kan komma överens om har betingats av småskaligheten i de samhällen som människan har levt i under större delen av sin historia. Medborgarna i sådana miniatyr-samhällen tenderar att vara personligen bekanta med varandra och så pass mentalt homogena både viljemässigt och epistemiskt att det kan te sig överkomligt för dem att leva upp till de vagare opartiskhetskrav som en outvecklad moralfilosofi mäktar med att leverera, exempelvis krav i stil med att man ska behandla andra som man själv vill bli behandlad vilka förekommer i flera gamla världsreligioner. Det potentiella moraliska moraset är knappt märkbart under slika yt-skikt av likformighet.

I dagens globaliserade värld med mer heterogena samhällen hårbärgerande miljontals medborgare tränger det moraliska moraset däremot fram med obändig kraft och tydlighet. Att den avancerade teknologin i dessa samhällen möjliggör påverkan över hela jorden för den överskådliga framtiden förstärker ytterligare utmaningen att nå moralisk samklang. Ty den avancerade teknologin medför att det i sektion 2 uppmärksammade förhållandet att spårningen av handlingars externa relationer i princip saknar gräns får betydande praktisk signifikans eftersom konsekvenser för massor av individer under långvariga tids-

perioder är nära nog omöjliga att pålitligt förutsäga. Till detta kommer att en alltmer sofistikerad moralfilosofi oförtrutet radar upp ett flertal plausibla opartiskhetskrav som det är marigt att välja bland och ännu marigare att efterleva enär de ovillkorligen måste tillämpas på enorma mängder av okända individer med vilka samarbete är påkallat och/eller för vilka relevanta handlingsalternativ har kännbara följder.

För att sammanfatta de ovannämnda villkoren för moraliskt moras, dvs. svår- eller olöslig moralisk oenighet:

Uppfattningen av moralen som absolut tvingas vika för moralisk relativism, enligt vilken det inte finns någonting som är moraliskt rätt och riktigt för alla personer oavsett deras viljeattityder utan moralen uppvisar viljemässig avhängighet.

Det finns djupa meningsskiljaktigheter i moralfilosofin, både i den normativa etiken – t. ex. mellan konsekventialism och deontologi eller mellan en moral som är strikt opartisk och en som tillåter viss partiskhet mot jaget och nära och kära – och i metaetiken – t. ex. mellan subjektivism och objektivism eller naturalism och icke-naturalism – som inga filosofiska argument verkar kunna avgöra (vilket jag argumenterar för i Persson 2017, kap. 12).

Många människor har djupgående känslomässiga bindningar till falska religiösa eller politiska ideologier eller andra fördomar förhärskande i deras uppväxtmiljö som mer än förnuftsargument formar deras moraliska ståndpunkter, t. ex. kan deras syn på abort och eutanasi bestämmas av deras kristna tro eller deras syn på egendomsrätt av deras egoism.

I en globaliserad värld ställs rika mångmiljon-stater i besittning av sofistikerad teknologi inför moraliska val mellan handlingar vars konsekvenser är så vittförgrenade att det är praktiskt taget omöjligt att avgöra vilka av dem som har bäst konsekvenser; detta gäller t. ex. hur antropogena klimatförändringar ska hanteras, hur biologisk mångfald ska bevaras eller hur stor skyldighet de har att ge bistånd till de fattigaste länderna.

Förutom moralens subjektivitet, som försvarats här, bidrar alltså segslitna moralfilosofiska kontroverser, otillräcklig empirisk kunskap i vissa moraliska frågor och känslostyrning av moraluppfattningar överlag till det moraliska moraset. Rygggraden i min argumentation är emellertid en demonstration av hur man med utgångspunkt i en subjektivistisk analys av det individ-relaterade, intrinsikala värdet av hedoniska sensationer kan konstruera subjektivistiska analyser av både individ-

relaterat extrinsikalt och totalt värde och av absolut värde. Denna hypotes framställer allt värde som härledbart till eller från individ-relaterat egenvärde. Hedoniska förnimmelser är inte det enda som har egenvärde för oss, men de kan mycket väl vara det vars egenvärde är mest fundamentalt, liksom de kan vara den fundamentala motivationskällan.

## Litteratur

- Hare, Richard Mervyn. 1963. *Freedom and Reason*. Oxford: Clarendon Press.
- Moore, George Edward. 1903. *Principia Ethica*. Cambridge: Cambridge University Press.
- Parfit, Derek. 2011. *On What Matters*, vol. 1 & 2. Oxford: Oxford University Press.
- Persson, Ingmar. 2005. *The Retreat of Reason*. Oxford: Clarendon Press.
- Persson, Ingmar. 2013. *From Morality to the End of Reason*. Oxford: Oxford University Press.
- Persson, Ingmar. 2017. *Inclusive Ethics*. Oxford: Oxford University Press.
- Persson, Ingmar. 2023. "Smärtans och lidandets intrinsikala dålighet", *Filosofisk tidskrift* 44, nr 3, s. 3–10.