

KRISTER BYKVIST

Är de flesta utilitarister deontologer?

Om preferentialismens relativisering av värde

1. Introduktion

Det är populärt att hävda att den viktigaste skillnaden mellan deontologer och konsekventialister är att de förra men inte de senare kan ta *agentrelativa* värden eller mål i beaktande. (Se t ex Nagel (1986) kap 8 och 9, Pettit (1993) s xii-xiv och Scheffler (1987).) Till de agentrelativa mål som deontologen vill försvara hör agentrelativa *restriktioner*, dvs vi är förbjudna att utföra vissa typer av handlingar trots att de skulle maximera värdet i en viss situation, och agentrelativa *friheter*, dvs vi är ibland fria, men inte förpliktade, att välja den handling som maximerar det goda. Det är agentens intima relation till sina egna handlingar som är det viktiga här. Eftersom det är omöjligt att, utan att göra våld på begreppet konsekvens, konstruera en agentneutral konsekventialism som inkorporerar dessa deontologiska intuitioner, drar vissa moralfilosofer slutsatsen att vi måste överge den agentneutrala konsekventialismen. Inte nödvändigtvis så att vi ska bli renodlade deontologer. Vi kan hålla fast vid att utfallens värde bestämmer handlingens normativa status. Däremot måste vi överge kravet på agentneutralitet. Om vi inför agentrelativa värderingar av utfallen, så kan vi fånga några av de viktigaste deontologiska intuitionerna.

Vad jag vill visa är att vi har skäl att införa agentrelativa värderingar för att kunna fånga vissa viktiga och vanliga *utilitaristiska* intuitioner. Detta får konsekvenser för hur vi ska se på distinktionen mellan konsekventialism och deontologi. En definition av konsekventialism som utesluter agentrelativitet tycks vara alltför snäv då den inte rymmer några av de vanligaste formerna av utilitarism.

2. Agentrelativitet

Ibland sägs det att en teori är agentrelativ om den ger olika mål till olika agenter. En teori som ger olika agenter samma mål sägs då vara agentneutral. Visst ger denna karakterisering ett hum om vad agentrelativitet och agentneutralitet handlar om. Men vi kan inte nöja oss med denna oklara formulering. En något bättre formulering ges i Samuel Schefflers ofta citerade definition av konsekventialism:

[den specificerar] en princip som rangordnar totala sakförhållanden från det bästa till det sämsta från en opersonlig ståndpunkt. Med andra ord, de rangordningar som genereras av den utsedda principen är inte agentrelativa; *de varierar inte från en person till en annan beroende på vilken bestämd situation man befinner sig i* (Scheffler (1987) s 1, min översättning och kursivering).

Inspirerade av Scheffler skulle vi kunna säga att en teori är *agentrelativ* om och endast om den implicerar att rangordningen av en viss mängd utfall varierar från en agent till en annan. En rangordning av en mängd utfall varierar om något utfall byter plats i rangordningen. En teori är då *agentneutral* om och endast om den implicerar att rangordningen av en mängd utfall måste vara densamma för alla agenter.

Kanske bör vi inte trycka så hårt på att det ska vara en annan *agent* för att vi ska få den relevanta agentrelativiteten. Det är väl relativiseringen till olika *agentpositioner* och inte till olika personer som är det viktiga. En agentposition kan vi då identifiera med en tidpunkt i en värld och en mängd alternativa handlingar. En person kan inte vara olika agenter, men en och samma agent kan vara vid olika agentpositioner.

3. Preferentialism och agentrelativitet

Många moderna utilitarister är preferensutilitarister. Deras axiologi, *preferentialismen*, är den teori som säger att värdet av ett utfall uteslutande baseras på preferenser. Om vi kan visa att preferentialismen är agentrelativ, så har vi alltså visat att många moderna utilitarister förespråkar en teori som är agentrelativ. Är nu preferentialismen agentrelativ i den mening som jag skisserat ovan? Något klart besked får vi först om vi tittar lite noggrannare på vad som menas med preferentialism. Som påpekats av Wlodek Rabinowicz, Bertil Ström-

berg och Jan Österberg klyver sig preferentialismen i två huvudgrupper alltefter vad som ses som värdebärare. Enligt den ena huvudgruppen, som vi kan kalla *objektpreferentialism*, är det det önskade eller oönskade objektet som har värde. Ofta lägger man här till ett antagande om värdets uppkomst som säger att det är *genom* att vara önskat eller oönskat som objektet *får* värde. Enligt den andra huvudgruppen, som vi i brist på bättre kan kalla *tillfredsställelsepreferentialism*, är det i stället *sakförhållandet* att en preferens är tillfredsställd eller frustrerad som har värde. Skillnaden mellan dessa huvudgrupper kan enkelt belysas med följande exempel. Anta – för att hämta ett fall från Hägerström – att jag önskar att jag får en tunna sill. Att denna önskan är tillfredsställd innebär att det föreligger ett sammansatt sakförhållande som består av sakförhållandet att jag önskar att jag får en tunna sill *och* sakförhållandet att jag får en tunna sill. Enligt tillfredsställelsepreferentialismen har *hela* detta sammansatta sakförhållande värde medan enligt objektpreferentialismen så är det endast en *del* som har värde, nämligen sakförhållandet att jag får en tunna sill.

4. Objektpreferentialism

Nu tillbaka till frågan om preferentialism och agentrelativitet. Jag vill hävda att *alla* rimliga former av objektpreferentialism och *många* rimliga former av tillfredsställelsepreferentialism är agentrelativa. Jag börjar med objektpreferentialismen.

Anta att jag just nu kan springa ut på torget och köpa en tunna sill. Värdet av att jag får en tunna sill bestäms av preferenser, men vilka? Är vi riktigt generösa så tittar vi på alla *möjliga* preferenser, både faktiska och blott möjliga, som är riktade mot sakförhållandet att jag får en tunna sill. Men detta tycks betyda att sakförhållandet att jag får en tunna sill varken är bra eller dåligt, eftersom man för varje möjligt gillande med en viss styrka av att jag får en tunna sill kan hitta ett möjligt ogillande med samma styrka av att jag får en tunna sill, och det positiva värde som gillandet ger uppvägs av det negativa värde som ogillandet ger. Nu kanske man kan hålla med om att sakförhållandet att jag får en tunna sill varken är bra eller dåligt, men eftersom värderingsmetoden är tänkt att vara generell och det gäller för *varje* sakförhållande p att man för varje möjligt ogillande av p hittar ett möjligt gillande av p som har samma styrka, så får vi det absurda

resultatet att inget sakförhållanden är bra och inget sakförhållande är dåligt. (Detta har påpekats av Wlodek Rabinowicz och Jan Österberg.) Observera att dessa värderingar gäller oberoende av vilken agentposition vi intar; alltså är denna objektpreferentialism agentneutral. Agentneutraliteten har dock köpts till ett alldeles för högt pris. Vi måste alltså begränsa mängden värdegrundande preferenser för att få en rimlig objektpreferentialism.

En radikal begränsning, som föreslagits av bland andra Wlodek Rabinowicz, är att man ska tillämpa en form av *aktualism* och bara utgå från de preferenser som förekommer i den aktuella världen. Värdet av sakförhållandet att jag får en tunna sill bestäms endast av de *aktuella* preferenser som är riktade mot sakförhållandet ifråga. Om vi dessutom inte gör någon skillnad i när en preferens föreligger utan räknar med förflutna, nuvarande och framtida aktuella preferenser undviker vi agentrelativism. Värdet av att jag får en tunna sill fixeras av de faktiskt existerande förflutna, nuvarande och framtida preferenser som är riktade mot sakförhållandet i fråga. Det är klart att detta värde inte varierar över agentpositioner.

Men begränsningen till de aktuella preferenserna ger oss en oacceptabel "aktualistisk diktatur". För plocka ut ett sakförhållande p som någon i den aktuella världen önskar. Tänk på en annan värld där p föreligger och alla personer (eller djur) har så gott som alla preferenser frustrerade. Kontrastera denna helvetesvärld med ytterligare en annan tänkt värld – en paradiskvärld – i vilken p inte föreligger men alla personer har så gott som alla preferenser tillfredsställda. Anta dessutom att förutom p så gäller det för varje sakförhållande som någon aktuell person önskar att antingen föreligger detta i både helvetesvärlden och i paradiskvärlden eller så föreligger detta varken i helvetesvärlden eller i paradiskvärlden. Om vi nu tänker oss en situation där agenten kan välja att realisera helvetesvärlden eller paradiskvärlden, så säger den aktualistiska objektpreferentialismen att agenten ska värdera helvetet högre än paradiset. Varför? Enligt denna teori är ju den enda relevanta skillnaden mellan helvetet och paradiset att det finns ett aktuellt önskat sakförhållande p som föreligger i helvetet men inte i paradiset. Det är endast beroende på denna skillnad som helvetet får högre värde än paradiset. Att helvetet verkligen innebär ett helvete för dem som lever där och att paradiset innebär ett paradisk för dem som lever där har ingen direkt betydelse för värderingen av dessa världar.

Detta är oacceptabelt. (Observera att jag här inte behöver förutsätta att illvilliga preferenser ska beaktas. p behöver inte vara ett sakförhållande att någon lider.) Således är även här priset för agentneutralitet alldeles för högt.

Man skulle kunna invända att värdet av en värld v ska bestämmas av de preferenser som är aktuella i v . I helvetesvärlden är de aktuella preferenserna de preferenser som existerar där, och värdet av världen ska därför bestämmas av dessa preferenser. I paradysvärlden är de aktuella preferenserna de preferenser som existerar där, och värdet av denna värld ska därför bestämmas av dessa preferenser. Men om "aktuella preferenser" förstås på detta relativa sätt, så är det ju klart att värdet kommer att variera över världar. För anta att i en värld v föredrar alla p framför q och att i en annan värld v' föredrar alla q framför p . Då kommer p att vara bättre än q i världen v , eftersom alla aktuella preferenser i världen v är till p 's fördel. Men q kommer att vara bättre än p i världen v' , eftersom alla aktuella preferenser i världen v' är till q 's fördel. Om man vill undvika att värdet varierar kan man alltså inte tolka de aktuella preferenserna på detta relativa sätt.

En annan invändning är att vi faktiskt här i den aktuella världen bryr oss om andra människors lidande och lycka. Vi önskar inte andra ont, vi önskar dem gott. Bortsett från att detta inte kan gälla alla aktuella personer – alla aktuella personer är ju inte änglar – och att de som bryr sig om andras väl och ve bryr sig betydligt mer om den egna välfärden, så är detta ett misslyckat sätt att försvara objektpreferentialismen. Vi vill ju inte säga att det negativa värdet av helvetesvärlden är avhängigt av om någon *annan* bryr sig om deras lidande. Att personerna i helvetesvärlden har nästan samtliga preferenser frustrerade måste ha någon *direkt* påverkan på värdet av världen.

Vi har alltså sett att vare sig vi når agentneutralitet genom att beakta alla möjliga preferenser eller genom att endast beakta de aktuella preferenserna så är kostnaderna för stora. Jag kan nu inte se hur man skulle kunna uppnå agentneutralitet genom att begränsa mängden av värdegrundande preferenser på något annat sätt utan att vi därmed lämnar preferentialismens centrala idé att värde ska baseras endast på preferenser. Därför tror jag att vi kan slå fast att inte någon av de rimliga formerna av objektpreferentialism kan undkomma agentrelativitet.

5. Tillfredsställelsepreferentialism

För att en tillfredsställelsepreferentialist ska kunna bestämma värdet av de alternativa utfallen måste han besvara frågan: Vilka tillfredsställelser och frustrationer ska vara med och bestämma värdet hos utfallen? Svarar han "alla tillfredsställelser och frustrationer som förekommer i utfallet!", så avslöjar han sig vara en agentneutral preferentialist. Värdet av ett sakförhållande varierar inte över agenter i hans teori, eftersom agentpositionen inte kan påverka vilka tillfredsställelser och frustrationer sakförhållandet innehåller. Men många tycker att denne preferentialist visar en alldeles för tillåtande attityd. Vissa typer av preferenser bör diskvalificeras. I det följande ska jag ta upp två kriterier som kan användas för att diskvalificera preferenser. Enligt det första kriteriet, *det temporala kriteriet*, är preferensernas förhållande till *tiden* viktig. Vi bör vara tidspartiska i någon mening. Enligt det andra kriteriet, *det modala kriteriet*, är det viktigt att skilja på preferenser vars existens är avhängig av våra val och de preferenser som existerar oavsett vad vi gör.

Den tidspartiskhet som det temporala kriteriet påbjuder kan yttra sig på lite olika sätt.

(1) Man kan vara tidspartisk vad gäller *preferensens* datering. Här kan vi skilja mellan *futuristen*, som inte kan se några skäl att tillfredsställa preferenser som har funnits men som inte längre finns, och som därför anser att endast nuvarande och framtida preferenser är värdegrundande, och *presentisten*, som går ett steg till och inte ser varför man ska tillfredsställa preferenser som kommer att finnas men som inte finns nu, och som därför menar att endast nuvarande preferenser är värdegrundande.

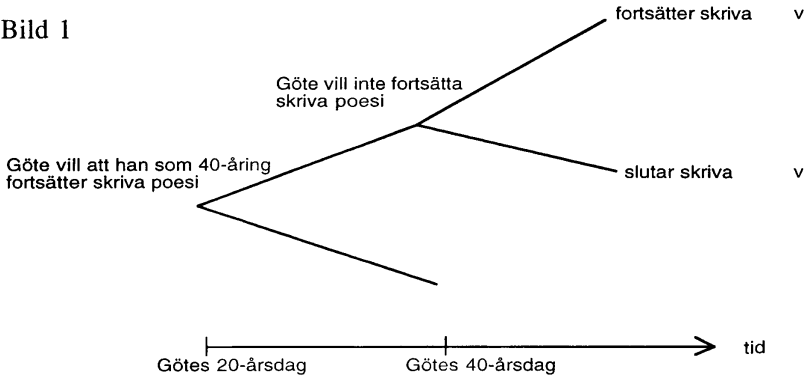
(2) Man kan vara tidspartisk vad gäller *objektets* datering. Även här kan vi skilja mellan presentisten, som inriktar sig på det intresse som finns för nuet och som därför menar att endast preferenser vars objekt föreligger i nuet är värdegrundande, och futuristen som inriktar sig på det intresse som finns för nuet och framtiden och därför menar att endast preferenser vars objekt ligger i nuet eller i framtiden är värdegrundande.

Vad jag nu vill visa är att alla tillfredsställelsepreferentialismer som är tidspartiska vad gäller preferensdatering men som är tidsneutrala vad gäller objekt datering är agentrelativa. Med "alla" menar jag de

presentistiska och futuristiska versionerna. Observera att dessa teorier inte är uppenbart orimliga. De säger något i stil med att vi bör göra det så bra som möjligt för de personer som existerar nu (presentisten) eller de som existerar nu och i framtiden (futuristen). Att göra det bra för en person innebär att tillfredsställa hans nuvarande eller framtida preferenser, oavsett vilken "riktning" i tiden dessa preferenser har.

Jag låter exemplet med poeten Göte illustrera de tidspartiska preferentialismernas agentrelativitet. (Se bild 1).

Bild 1



Göte, en ung poet, vaknar upp på morgonen till sin tjugonde födelsedag. Han håller på att avsluta diktsviten "Det mänskliga träskets sång". Även denna morgon känner han sitt poetiska kall men han vänder nu blicken även framåt och önskar intensivt att han som 40-åring fortfarande skriver poesi. Han är så frälst på poesi att han anser att även om han som 40-åring inte vill skriva så bör han ändå göra det. Det finns inget annat som kan göra livet så värdefullt. Det viktiga är dock inte att man lyckas skriva mästerverk utan att man arbetar med poesi och försöker skriva. Anta att v1 och v2 kan realiseras av både den 20-åringe och den 40-årige Göte. v1 är en värld i vilken han mellan 20 och 40 skriver poesi och efter 40 fortsätter att skriva poesi. v2 är en värld där han mellan 20 och 40 skriver poesi men där han som 40-åring slutar skriva poesi. Anta nu att Göte som 40-åring i v1/v2 kommer att sakna all lust att skriva poesi. Han tycker att det är helt OK att andra håller på med det men han kan *själv* inte finna någon glädje i det.

Anta att den framåtblickande preferens Göte har som 20-åring för att skriva poesi när han är 40 är starkare än den preferens Göte har som 40-åring mot att fortsätta skriva poesi. Anta slutligen att vad

gäller alla andra preferenser så är v1 och v2 likvärdiga, dvs för varje annan preferens som förekommer i både v1 och v2 så gäller det att den är tillfredsställd i båda världarna eller frustrerad i båda världarna.

Om vi utgår från Götes agentposition på hans 20-årsdag, så ger en *presentistisk* (vad gäller preferensdateringen) tillfredsställelsepreferentialism resultatet att v1 är bättre än v2. v1 innehåller ju tillfredsställelsen av en present preferens medan v2 innehåller frustrationen av densamma, och förutom den 40-årige Götes motvilja för poesiskrivande som inte ska räknas, så är v1 och v2 likvärdiga för alla andra preferenser. Om vi utgår från Götes agentposition på hans 40-årsdag (i världen v1/v2), så ger dock teorin resultatet att v2 är bättre än v1. För vid Götes 40-årsdag så hör Götes tidigare poesipreferens till det förflutna, och ska därför inte räknas längre. Relativt 40-årsdagen innehåller v2 tillfredsställelsen av en nuvarande preferens medan v1 innehåller frustrationen av densamma och för övrigt innehåller v1 och v2 exakt samma tillfredsställelser och frustrationer. Alltså är denna presentistiska preferentialism agentrelativ.

Vad säger en *futuristisk* (vad gäller preferensdateringen) tillfredsställelsepreferentialism om värderingen av dessa världar? Om vi utgår från Götes agentposition vid hans 20-årsdag, så säger teorin att v1 är bättre än v2. v1 innehåller tillfredsställelsen av en nuvarande preferens och frustrationen av en framtida preferens medan v2 innehåller frustrationen av en nuvarande preferens och tillfredsställelsen av en framtida, men eftersom den nuvarande preferensen är starkare än den framtida och världarna är likvärdiga för alla andra preferenser, så är v1 bättre än v2. Om vi i stället utgår från Götes agentposition vid hans 40-årsdag, ger teorin resultatet att v2 är bättre än v1, eftersom relativt denna agentposition är den unge Götes poesipreferens förfluten och därmed överspelad. Den enda preferens vi har att ta i beaktande här är den äldre Götes preferens för att inte fortsätta skriva poesi, och denna preferens är tillfredsställd i v2 och frustrerad i v1. Alltså är också denna futuristiska preferentialism agentrelativ.

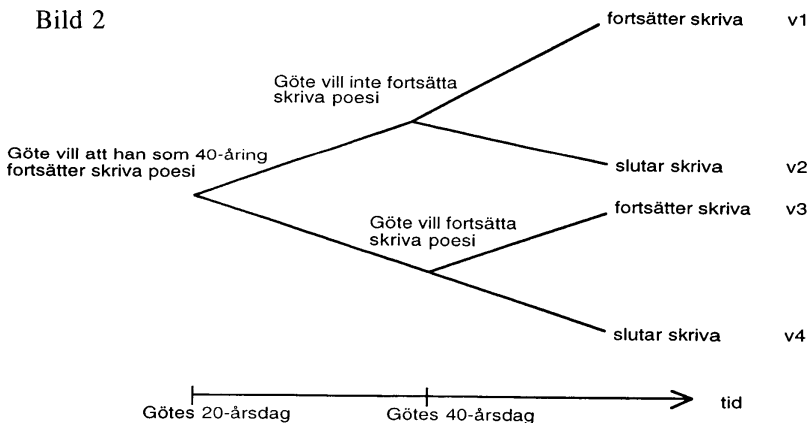
Vi ser alltså att även tillfredsställelsepreferentialismen råkar ut för agentrelativitet om den omfattar en viss typ av tidspartiskhet; en tidspartiskhet som inte alls verkar orimlig.

Nu kan man ju ändå tycka att tidspartiskhet är något konstigt och att en rimlig tillfredsställelsepreferentialism ska vara tidsneutral i båda avseendena. Men även en tidsneutral tillfredsställelsepreferentialism

råkar ut för agentrelativitet om den är modalt partisk. Ni minns att det modala kriteriet skiljer ut preferenser alltefter som de är kontingenta eller inte. Gemensamt för de modalt partiska teorierna är att de hävdar att ett tillstånd inte kan vara bättre (sämre) än något annat om det inte är bättre (sämre) *för* någon. En variant av modal partiskhet menar att detta betyder att endast *nödvändiga preferenser*, sådana som är oundvikliga i den meningen att de existerar oavsett vad vi gör, är värdegrundande. Vi gör det inte bättre för någon, och därmed inte heller världen bättre, genom att skapa nya preferenser och tillfredsställa dem. Vi ska istället inrikta oss på att tillfredsställa nödvändiga preferenser.

En annan variant – och en populär sådan bland populationsetiker – är vad man brukar kalla ”person-affecting theory”, enligt vilken man inte kan göra det bättre för en person genom att skapa henne. Vi kan bara göra det bättre för *nödvändiga personer*, sådana personer som existerar oavsett vad vi gör. Om ”göra det bättre för nödvändiga personer” förstås som att tillfredsställa deras nödvändiga preferenser är vi tillbaka till teorin som värnar om oundvikliga preferenser. Men vi kan stoppa denna reduktion genom att säga att vi kan göra det bättre för nödvändiga personer genom att *skapa nya* preferenser åt dem och tillfredsställa dessa. Gemensamt för dessa teorier är dock att vi inte gör världen bättre bara genom att skapa nya lyckliga människor.

Oavsett vilken form av modal partiskhet vi väljer så får vi en agentrelativ teori. Låt oss gå tillbaka till poeten Göte, (se bild 2).



Här har vi ändrat exemplet något genom att anta att Göte på sin 20-årsdag kan se till att han fortfarande på sin 40-årsdag är poesifrälst. Vi kan t ex föreställa oss att han kan åka iväg och bosätta sig i ett kollektiv av brinnande poeter och stanna där i 20 år. Resultatet av denna vistelse skulle bli att han på sin 40-årsdag intensivt skulle vilja fortsätta skriva poesi. Detta betyder att den äldre Götes motvilja för att skriva poesi är en kontingent preferens relativt den yngre Götes agentposition. Motviljan finns i endast två av de alternativa världar som det står i den yngre Götes makt att realisera. Anta också att denna motvilja är starkare än den yngre Götes önskan att som 40-åring fortsätta skriva poesi. (v3 är en värld där Göte har åkt till poesikollektivet och han på sin 40-årsdag intensivt vill fortsätta att skriva poesi och faktiskt fortsätter att göra det. v4 är identisk med v3 förutom att den äldre Göte inte fortsätter att skriva.)

Hur värderas v1 och v2 av en teori som är ute efter att tillfredsställa nödvändiga preferenser? Om vi utgår från den yngre Götes agentposition får vi resultatet att v1 är bättre än v2. Den äldre Götes motvilja är ju kontingent relativt den yngre Götes agentposition och ska därför inte räknas utifrån denna position, och den yngre Götes poesipreferens är den enda nödvändiga preferens som berörs av valet mellan v1 och v2. Om vi däremot utgår från den äldre Götes agentposition i världen v1/v2, så får vi resultatet att v2 är bättre än v1. För relativt denna agentposition är ju nu den äldre Götes motvilja en nödvändig preferens. Och eftersom den yngre Götes preferens också är nödvändig men svagare, så får vi resultatet att alla nödvändiga preferenser tagna i beaktande så är v2 bättre än v1.

Vi får samma resultat om vi istället tillämpar "person-affecting theory". Byt ut den äldre Göte i v1/v2 till en Göte junior, och byt ut den äldre Göte i v3/v4 till en *annan* Göte junior. Anta att Göte som 20-åring har valet att skapa en men inte båda av dessa Göte juniors, och att Göte vill att hans barn ska skriva poesi. Anta också, mycket orealistiskt, att inga andra preferenser är berörda än Götarnas olika poesipreferenser.

Resultatet blir alltså att oavsett vilken modal partiskhet vi antar att den tidsneutrala tillfredsställelsepreferentialismen omfattar så får vi agentrelativitet.

6. Relativiserat värde och konsekventialism

Vi har funnit att varje rimlig form av objektpreferentialism och vissa rimliga former av tillfredsställelsepreferentialism, två tidspartiska och två modalt partiska, alla är agentrelativa i den meningen att de ger olika värderingar av samma sakförhållande till olika agenter. Utifrån en agent sett är ett sakförhållande p bättre än ett annat sakförhållande q , men utifrån en annan agent sett är q bättre än p . Men innebär inte detta att dessa preferentialismer implicerar en motsägelse? Det är väl omöjligt att p är bättre än q och att q är bättre än p . Nog för att vi ofta är oense om vilka strukturvillkor som gäller för relationen "bättre än" men så mycket kan vi väl ena oss om att den måste vara asymmetrisk?

Det enkla svaret är att vi undviker motsägelsen genom att göra relativiseringen till agentpositioner explicit i värderingen. Med andra ord, enligt dessa preferentialismer är den korrekta formen för en "bättre-än"-värdering inte den tvåställiga " p är bättre än q " utan den treställiga " p är bättre än q utifrån agentpositionen A ". När vi sedan som preferensutilitarister rör oss från axiologin till det normativa, så är det dessa agentrelativa värderingar som ska bestämma den normativa statusen hos handlingarna. Varje agent bör utföra den handling som leder till det utfall som maximerar värdet relativt den agentposition vid vilken agenten befinner sig.

Hur går dessa värderingar ihop med konsekventialism? Inte alls, om vi utgår från Schefflers definition av konsekventialism som jag citerade i inledningen. Enligt Scheffler är konsekventialismer agentneutrala, men det är ju helt klart att de preferensutilitarismer vi gått igenom inte är agentneutrala.

Notera också att dessa preferentialistiska värderingar inte går ihop med konsekventialism tolkad som en teori enligt vilken den normativa statusen hos en handling bestäms enbart av det *intrinsikala* värdet hos dess utfall (och de alternativa handlingarnas utfall). Även om det är oklart hur man mera exakt ska förstå intrinsikalt värde, så torde man kunna säga att det intrinsikala värdet inte kan variera över tidpunkter inom en och samma värld. Det verkar också rimligt att lägga till att det intrinsikala värdet inte kan variera över världar. Nu är det så, om vi ska vara lite petiga, att det värde som de beskrivna preferentialismerna omfattar inte varierar över tidpunkter eller världar, eftersom relativiseringen inte rör tidpunkter eller världar utan rör agentposi-

tioner. Men om det nu överhuvudtaget är möjligt att relativisera intrinsikalt värde till agentpositioner – något som moralfilosofer i Moores anda skulle sätta sig emot – så verkar det rimligt att här liksom vad gäller tidpunkter och världar kräva invarians. Med andra ord, om ett värde varierar över agentpositioner, så kan inte detta värde vara intrinsikalt.

Enligt de preferensutilitarismer vi gått igenom så bestäms den normativa statusen hos en handling av ett värde som varierar över agentpositioner. Men om ett värde varierar över agentpositioner så kan det inte kan vara intrinsikalt. Alltså får vi resultatet att dessa teorier *inte* är former av konsekventialism! Ett något överraskande resultat eftersom så många moderna utilitarister, utan att förneka att de är konsekventialister, bekänner sig till någon av dessa preferensutilitarismer. Var ligger felet? Hos dessa utilitarister, som vi kanske bör kalla ”deontologer”, om vi förstår en deontolog som en icke-konsekventialist? Eller hos de föreslagna definitionerna av konsekventialism som båda utgår från att värdet är invariant?

Jag tror att vi måste hålla fast vid att intrinsikalt värde inte kan variera över agenter. Även om vi kan vara oense om många strukturvillkor för intrinsikalt värde och om hur den riktiga definitionen lyder, så borde väl vi enas om att *ett* kännetecken för intrinsikalt värde, det värde något har i kraft av sina intrinsikala egenskaper, är att det inte kan variera över agentpositioner. Att ett sakförhållande ligger i framtiden, nuet eller det förflutna, är *extrinsikala* egenskaper hos sakförhållandet, egenskaper det har i kraft av att de står i relation till någon tänkt tidpunkt. Likaså är det att ett sakförhållande innehåller eller konstitueras av tillfredsställelsen av någon kontingent preferens en extrinsikal egenskap hos sakförhållandet som det har i kraft av att stå i en viss relation till någon agent. Det är därför de föreslagna definitionerna av konsekventialism som vi bör förkasta. En definition som är för snäv utifrån våra förteoretiska intuitioner bör vi inte hålla fast vid.

Hur ska vi då skilja på deontologi och konsekventialism, något som de flesta moralfilosofer förutsätter är möjligt? Kanske visar svårigheten att finna en godtagbar distinktion mellan konsekventialism och deontologi att det egentligen är lönlöst att leta efter formella eller strukturella olikheter med vilka man skulle kunna definiera traditionerna. Gemensamt för definitionsförsöken är ju att man försöker finna skillnader som inte ligger direkt på det substantiella planet. Man hävdar att det

är *genom* sina strukturella egenskaper som konsekventialismen utsluter, och deontologin kan omfatta, vissa substantiella värderingar. Kanske ska man i stället se konflikten mellan konsekventialister och deontologer som en konflikt mellan rent substantiella värderingar. Kort sagt: konsekventialismen och deontologin kan stöpas i samma form. Oenigheten rör inte formen utan vad som ska läggas i formen. (Jag vill tacka Erik Carlson och de andra deltagarna på filosofidagarna i Umeå, 1995, för värdefulla synpunkter på en tidigare version av denna text.)

Litteratur

NAGEL, T, *The View from Nowhere*, Oxford, 1986.

PETTIT, P, (red), *Consequentialism*, Aldershot, 1993.

SCHEFFLER, S, *The Rejection of Consequentialism*, Oxford, 1987.