

Christine Korsgaards moralfilosofi: en kritisk introduktion.

1. Introduktion

Olika upplevelser av *krav* är en viktig del av en människas liv. Människor upplever att de måste vara roliga, duktiga, självständiga, vänliga eller vältränade; att de måste gå till jobbet, borsta tänderna, träffa vänner, skaffa barn, ägna tid åt sina barn och sina föräldrar eller demonstrera mot orättvisor; att de inte får skvallra, inte äta fet mat, inte titta på porr eller kasta skräp på gatan.

I böcker och föreläsningsserier¹ har Harvardfilosofen Christine Korsgaard tagit som sin uppgift att klargöra vilken grund dessa krav har. Hennes kantianska svar är i mycket korta ordalag följande: Endast den som strävar efter att själv bestämma sina handlingar är en moraliskt ansvarig varelse, en varelse som kan uppfatta plikter. Att själv bestämma att man utför en viss handling är vidare detsamma som att uppfylla en uppsättning villkor, och dessa villkor har vittgående

¹ Min redogörelse för Korsgaards uppfattning tar sin utgångspunkt i hennes böcker *The Sources of Normativity* (Cambridge: Cambridge U. P. 1996) – i fortsättning förkortad *SN* – och *Creating the Kingdom of Ends* (Cambridge: Cambridge U. P. 1996) samt *The John Locke Lectures / Hägerströmsföreläsningarna* – i fortsättningen förkortade *HF* – som hölls i Oxford respektive Uppsala år 2002.

implikationer för *vilka* handlingar man själv kan och inte kan bestämma att man utför. Och eftersom man inte kan bestämma att man inte alls handlar – också det är att handla i den relevanta meningen – så kan en moraliskt ansvarig varelse inte själv välja att inte uppfylla villkoren i fråga. Påbudet att uppfylla dessa villkor är i den betydelsen en absolut eller ovillkorlig praktisk norm för den varelse som försöker att själv bestämma de handlingar hon utför, dvs för varje moraliskt ansvarig varelse.

Om Korsgaard lyckats i sina föresatser så har hon undvikit ett par av de svåraste problem som moralfilosofen brottats med. För det första så har hon undvikit det problem som många naturalistiska moralteorier drabbats av, nämligen att moralens krav i dessa teorier inte längre framstår som *måsten*. Om en handlingens egenskap att vara någons plikt till exempel identifieras med dess egenskap att ha bättre totala konsekvenser för alla berörda än vad alternativa handlingar har, så verkar det som om vårt intresse av att göra vår plikt blir beroende av att vi har den allmänna välmågan som mål. I den mån som detta inte är ett viktigt mål för vårt handlande – och den allmänna välmågan är knappast någonting som vårt beslutsfattande normalt fokuserar på – så blir det svårt att förstå varför vi skulle uppfatta vår plikt som något vi *måste* göra. Om kraven i stället har sin grund i normer som varje varelse som försöker bestämma sina handlingar med logisk nödvändighet försöker uppfylla så

blir inte vårt intresse av att handla pliktenligt avhängigt vad vi *råkar* ha för specifika mål för våra handlingar.

För det andra så erbjuder Korsgaard en till synes begriplig naturalistisk *förklaring* till varför vissa krav på vårt handlande verkar vara ofrånkomliga. Andra försök att förklara detta utan att samtidigt förneka kravens verklighet eller giltighet har tenderat att postulera några med intuition greppbara normativa icke-naturliga fakta, utan att förklara hur det kommer sig att dessa fakta kan greppas med sådan intuition, eller varför de framträder som berättigade ofrånkomliga krav. Sådana fakta framstår då som ”olika allt annat vi känner till”, för att använda J. L. Mackies uttryck, och det blir oklart vilken grund vi kan ha för att postulera deras existens.

I det följande ska jag göra två saker. För det första ska jag ge en kort men förhoppningsvis begriplig introduktion till Korsgaards teori. För det andra ska jag kritiskt diskutera hennes argument på några få men centrala punkter. Av utrymmesskäl kommer jag däremot inte att diskutera de många intressanta relationer som Korsgaards teori har till tankar hos hennes huvudsakliga inspirationskällor, Kant, Platon och Aristoteles, eller till tankar hos Nietzsche, Sartre, Rawls och nu aktiva filosofer som Jürgen Habermas, Alan Gewirth, Thomas Nagel och Margaret Gilbert.

2. Normativitet och konstitutiva villkor för handlande

För att förstå Korsgaards teori behöver man förstå hennes kantianska påstående att moraliskt ansvariga varelser är kapabla att själva bestämma sitt handlande. Med detta menas inte blott att sådana varelser på ett intelligent sätt kan välja medel som leder till realiserandet av vissa givna mål. Att själv bestämma sitt handlande är i stället att bestämma huruvida man ska *realisera ett visst mål med ett visst agerande vid en viss tidpunkt på ett visst sätt*. Det intelligenta väljandet av medel tar målet för givet medan bestämmandet av handling väljer både mål och medel.

När vi själva bestämmer att vi ska utföra en viss handling föregås detta av att ett tillstånd tycks oss eftersträvansvärt, och att vår instrumentella intelligens uppfattar ett agerande som ett möjligt medel till att det eftersträvansvärda realiseras. Så långt har vi presenterats med ett handlingsförslag, ett förslag som har formen av en maxim: agera så-och-så under de-och-de omständigheterna för att uppnå det-och-det. Det speciella med varelser som är moraliskt ansvariga är att *de kan reflektera över maximier och strävar efter att själva medvetet bestämma huruvida de faktiskt ska realiseras*.

Om jag förstått Korsgaard rätt så utesluter förmågan att välja huruvida en maxim ska realiseras inte att maximerna uppfattas som en sorts krav för den rationella agenten: tvärtom. Att uppfatta något som ett till en själv riktat handlingsförslag är

just att uppfatta det som lagstiftande för ens handlande *om man inte måste avvisa det*.² En av Korsgaard's kantianska poänger är att förnuftet självt kan kräva att vi avvisar olika sorters krav, och det är i grund och botten i kraft av detta som vi både är förmögna att bestämma över olika krav som ställs och dessutom nödgade att som rationellt beslutande acceptera vissa normer och värden. För att en agent ska bestämma huruvida en *maxim* ska realiseras eller ej måste nämligen ett antal villkor vara uppfyllda, och det är dessa villkor som Korsgaard tänker sig utgör grunden för det *normativa* (*vilka* hon tänker sig att villkoren är ska vi snart komma till). En varelse förmögen till rationella beslut kan i olika mån misslyckas med att uppfylla villkoren – kanske på grund av yttre våld, förvirring, ”viljesvaghet” – och misslyckas i samma mån också med att bestämma sitt handlande, dvs med att vara en reflekterande *agent*. Men hon kan inte själv avgöra att hon helt ska bryta mot villkoren – för att själv avgöra detta måste hon ju uppfylla dem – och på så sätt utgör dessa villkor absoluta praktiska måsten för varje rationell varelse som strävar efter att bestämma sitt

² Det är på grund av detta antagande som Korsgaard säger att *maximer* som överlever reflektion är *intrinsikalt normativa* entiteter. De är *normativa* för att vi inte kan uppfatta dem utan att försöka handla i enlighet med dem, och de är *intrinsikalt normativa* för att reflektionen testas om medlen passar målet, dvs om *maximens interna* struktur är den rätta, dvs om förnuftet inte måste avvisa dem. Se *SN*, s. 108-12.

handlande. Det är just detta som gör villkoren *normativa* i den mening som Korsgaard är intresserad av: vi kan inte bestämma oss för att inte försöka uppnå dem, men vi kan misslyckas med att följa dem.

Inte alla konstitutiva normer är av det slaget. Ett hus är ett föremål som uppfyller vissa för hus konstituerande villkor – det ska skyddas mot väder och vind på vissa sätt. Det går att i olika grad misslyckas med att bygga ett hus, men det går inte att ens försöka bygga ett hus utan att försöka producera något som uppfyller de för hus konstituerande villkoren. Villkoren är därmed också konstitutiva och normativa för husbyggande, men inte absoluta eller rationellt nödvändiga på det sätt som villkoren för rationella beslut är: man kan bestämma att man inte alls ska bygga ett hus.

Hos självmedvetna varelser som strävar efter att vara agenter tar sig denna strävan uttryck i samvetets verkningar: det är genom skuld, skam och ånger som den självmedvetna varelsen känner att hennes identitet som agent är hotad och därigenom blir i stånd att skydda den identiteten, och det är genom bland annat stolthet som hon känner att handlingsförmåga är stärkt.³ Om Korsgaard har rätt förklaras

³ Allmänt så tänker sig Korsgaard att lust och olust av olika slag (inbegripet skamkänslor och ånger) är de sätt på vilka medvetna varelser uppfattar vad som möjliggör respektive hotar deras fysiska och praktiska

alltså den pliktens fenomenologi som är intimt förknippad med samvetet av den pliktkännande varelsens strävan att vara en agent.⁴

Det hypotetiska imperativet

Det första villkor som Korsgaard tänker sig är konstitutivt för bestämmandet av de egna handlingarna är det hypotetiska imperativet: om du vill ett mål ska du välja de nödvändiga medlen därför. Att bestämma att ett mål realiserar är nämligen enligt Korsgaard att vara orsak till målets realiserande, och detta inbegriper att vara orsak till någon orsakskedja⁵ som leder från

→ identitet snarare än det som varelsen uppfattar som intrinsikalt gott respektive ont. Se *SN*, speciellt s. 145-55.

⁴ Se till exempel *SN*, s. 151. Samvetet är inte alls något tema i *HF*, men där utgår Korsgaard helt enkelt från antagandet att vi försöker handla och försöker visa på dess normativa implikationer.

⁵ Om mitt mål är att följa trafikreglerna så måste jag stanna för rött ljus. Att jag stannar för rött ljus är en del av det i kraft av vilket jag följer trafikreglerna, men att säga att mitt stannande för rött ljus *orsakade* att jag följde trafikreglerna kan antyda att dessa två fakta är åtskiljda i tiden eller i rummet, vilket de förstås inte är: mitt regelföljande *bestod i* att jag stannade för rött. "Orsak" ska emellertid förstås i vid bemärkelse där orsaker kan *konstituera* eller vara superveniensbas för sin verkan. Detta innebär också att ett visst agerande kan vara det enda som ska uppnås med just det agerandet så att medel och mål sammanfaller: vi kan önska umgås för umgängets skull, eller dansa för dansandets skull. Och

→

en själv till målet, dvs för att vara orsak till nödvändiga medel till att målet uppnås. Det hypotetiska imperativet är alltså konstitutivt för agenter: man bestämmer inte sitt handlande om man inte själv avgör att maximen i fråga realiserar, och man avgör inte detta om man inte orsakar nödvändiga medel till målet i fråga.

Däremot kan man förstås missta sig på vad som är medel till ens mål. Vi kan misslyckas med att handla i enlighet med principen, men vi kan inte avsiktligt bryta mot den – dvs själva medvetet bestämma att vi bryter mot den – eftersom avsiktliga handlingar alltså förutsätter att vi lyckas uppfylla principen. Principen är alltså normativ i den relevanta meningen.

Notera att vi inte uppfattar det hypotetiska imperativet som normativt därför att vi insett att det är "riktigt" och insett att det är tillämplig på vårt eget fall: "om jag ska handla rationellt så måste jag följa det hypotetiska imperativet, och det gör jag bara om jag vill detta medel, alltså måste jag vilja detta medel". Om vi inte redan uppfattat imperativet som tvingande så kan inte tillämpningen av det på vårt eget fall framstå som ett måste. Den praktiska nödvändigheten hos imperativet består alltså i att det

→ Korsgaard skulle säga att målet för att bestämma sina handlingar är att bestämma sina handlingar, dvs att vara en agent, och ingenting utöver detta. Se *HFI*

är konstitutivt för handlande och att vi därför inte själva kan bestämma att vi inte alls ska följa det.

Det kategoriska imperativet

Också det kategoriska imperativet – du ska endast handla så att du kan vilja att maximen för din handling upphöjs till allmän lag – är enligt Korsgaard konstitutivt för bestämmandet av de egna handlingarna. Denna tes bygger på en syn på orsaksbegreppet enligt vilken vi uppfattar händelseförlopp som orsaksförlopp endast om vi uppfattar dem som följande *ceteris paribus*-principer, eller vad Korsgaard kallar ”villkorligt universella principer”.⁶ Sådana principer är inte blott generella – det är inte bara så att de gäller i många fall – men de är inte heller universella – de täcker inte alla fall. I stället så gäller de alla fall där det inte finns skäl till att de inte gäller. En stor sten träffar ett fönster och fönstret går sönder: att det träffades av stenen var en orsak till att det gick sönder. Enligt Korsgaard skulle detta orsaksomdöme vila på uppfattningen att alla fönster som träffas av stora stenar går sönder *om det inte finns något speciellt skäl till att det inte går sönder*, t ex att det är armerat. Enstaka undantag lämnar generella principer oantastade, vederlägger absolut universella principer, men leder till ett tydliggörande av villkorligt universella principer (”alla fönster

⁶”Provisionally universal principles”, se *HFII*.

som träffas av stora stenar går sönder *om de inte är armerade, ceteris paribus*”).

Att bestämma huruvida ett handlingsförslag realiserar eller ej inbegriper att vara en *orsak* till att det realiserar eller inte realiserar. Och om man enligt ovanstående orsaksbegrepp ska kunna uppfatta sig själv som orsak till realiserandet av ett konkret handlingsförslag – ett ”gör så-och-så under de-och-de omständigheterna för att uppnå detta-och-detta” som anger *alla för beslutet relevanta aspekter av situationen* – så måste man identifiera sig med något som i situationer där de relevanta aspekterna råder alltid skulle leda till förslaget realiserande, *ceteris paribus*. I den betydelsen måste alltså ett reflekterande praktiskt förnuft handla endast så att hon kan vilja att maximen för hennes handling upphöjs till *allmän lag*. Det kategoriska imperativet är därför, liksom det hypotetiska, konstitutivt för handlande och för agenter.

Vad det kategoriska imperativet har för konsekvenser är emellertid ännu långt ifrån klart eftersom argumentationen så här långt inte etablerat att den som bestämmer realiserandet av en maxim måste uppfatta sig som något mer än *vad-det-nu-än-är-som-avgör-om-just-den-här-maximen-realiserar*. Det har nämligen inte sagts någonting om varför hon måste uppfatta sig som människa med historia och framtid, eller varför hon måste uppfatta sig som lagstiftande för andra människor. Det kategoriska imperativets *omfång* är alltså ännu en öppen fråga.

Och om det är obestämt så är det kategoriska imperativets implikationer också obestämda, som vi lätt kan se:

Antag först att den reflekterande självbestämmande varelsen måste identifiera sig med – och alltså förstå sig själv som lagstiftande för – varje reflekterande självbestämmande varelse och sig själv tidigare och senare i livet för att alls vara en agent. Då verkar det som om vissa principer är praktiskt uteslutna och som om det kategoriska imperativet är normativt: Handlande enligt maximen *avge falskt löfte för att undkomma en personlig svårighet* verkar till exempel vara uteslutet av just det skäl som Kant anger i *Grundläggning av Sedernas Metafysik*: vi kan inte bestämma att vi ska handla enligt den eftersom vår tilltro till löften kräver att vi normalt *inte* generellt avger falska löften för att undkomma svårigheter.

Om den reflekterande självbestämmande varelsen i stället kan förstå sig som lagstiftande för enbart sig-själv-nu så verkar handlande enligt ovanstående maxim inte alls uteslutet. Den uppenbara frågan är då varför hon måste uppfatta sig som mer än en orsak till realiserandet av handlingsförslaget i fråga: varför hon måste uppfatta sig som människa med historia och framtid, och i ännu högre grad varför hon måste uppfatta sig som lagstiftande för varje rationell varelse. Här krävs det argument, och i avsnitt 3 ska vi se vad Korsgaard har att erbjuda på den punkten.

Ändamål i sig, absoluta ändamål

Om en moraliskt ansvarig varelse själv ska bestämma sig för att eftersträva något – för att hå något som mål för sitt handlande – så måste hon per definition först lyckas uppfylla målet att själv bestämma sitt handlande. Det betyder förstås inte att målet att själv bestämma sitt handlande är ett mål i en maxim som någon *beslutar* sig för att handla efter: i den mån som någon försöker besluta och handla så måste hon (logiskt) redan sträva mot detta mål. Det egna effektiva beslutsfattandet kan därför heller aldrig vara blott ett medel för att uppnå något annat mål, utan är ett ändamål i sig. Och det är därför också ett absolut ändamål, ett som vi inte kan bestämma oss att ge avkall på för att uppnå några andra mål.

Vad denna slutsats har för ytterligare konsekvenser beror emellertid på vilken den reflekterande agenten är, vilket vi kan se på två områden där Korsgaard utgår från tanken om det egna effektiva beslutandets absoluta värde: i en diskussion kring det rationella självmordets möjlighet, och i kritik av den hedonistiska utilitarismen.

Om varje rationell varelse måste ha sitt eget effektiva beslutsfattande som absolut ändamål så verkar det uteslutet att självmord skulle kunna vara en medveten rationell handling. Åtminstone är det vad bland annat Kant hävdar. Men Korsgaard säger någonting som verkar vara i strid med detta: om en persons integritet hotas av fysisk eller psykisk sjukdom så kan hon välja

att bevara sin integritet och skydda det hon finner värdefullt genom att ta sitt liv.⁷ Tanken verkar vara att sjukdomen ändå utesluter att personen fortsätter existera som rationell agent och att den fortsatta existensen hotar saker som personen finner värdefulla: sitt goda namn, sina familjemedlemmars väl. Att ta sitt liv kan då vara det enda personen kan göra för att skydda det hon finner värdefullt. Poängen verkar alltså vara att självmord kan vara rationellt och medvetet valt när den egna fortsatta rationaliteten ändå är utesluten: när det inte är det egna effektiva beslutandet som avslutas, utan bara ens fortsatta existens som djur eller kropp. Vid närmare eftertanke ser man också att det finns andra svagare former av medvetet valda självmord som borde vara rationellt möjliga på samma sätt.⁸ Liksom jag tvivelsutan kan välja att sova och därmed att vara rationellt inaktiv för en stund för att kunna vara mer aktiv senare, så borde det finnas utrymme för den som identifierar sig med ett *samfund* att offra sitt liv för andras väl: samfundet kan fortsätta uppfylla de för självbestämmande konstitutiva

⁷ *SN*, s. 162.

⁸ Motsvarande poänger kan göras angående olika sorters styrelseskick för grupper av människor. Korsgaard diskuterar en del paralleller mellan en människas handlande och politiska styrelseskick, speciellt i HF 5, och en ytterligare parallell är förstås mellan frågan om hur självmord kan vara rationellt valt och frågan om en stat genom demokratiska val legitimt kan avskaffa principen att all makt utgår från folket.

normerna, det är bara en del som görs inaktiv för att det stora hela ska bli mer effektivt. Dvs, det verkande förnuft som agenten då identifierar sig med tas inte av daga, bara en del av dess flerhövdade kropp. Men alla dessa former av självmord förutsätter att personen förstår sig själv på ett visst sätt. För den person som fortfarande identifierar sig med sitt framtida jag och inte identifierar sig med något samfund så verkar självmord uteslutet.

Tanken att vi inte kan bestämma oss för att sätta något annat mål framför uppfyllandet av beslutandets konstitutiva normer ligger också bakom Korsgaards påstående att en rationell agent inte kan vara en helhjärtad hedonistisk konsekventialist: att vara det vore precis som i fallet med vissa sorters självmord att sätta uppnåendet av ett empiriskt mål (frånvaron av lidande, den egna lyckan, den totala lyckan) framför uppfyllandet av de för handling konstitutiva principerna, dvs framför handlandet.⁹

Den här invändningen mot den hedonistiska konsekventialismen eller utilitarismen bygger förstås på det ytterligare antagandet att försöken att uppnå det empiriska målet *kan* komma i konflikt med de konstitutiva principerna för handling, och det är ännu inte riktigt klart. På samma sätt som argumenten så här långt inte visat vilket omfång det kategoriska imperativet måste ha så är det nämligen fortfarande alldeles

⁹ Se *HFV*

öppet *vilken* existens som effektiva beslutsfattare – agenter – vi måste uppfatta som absolut värdefull. Är det existensen som bestämmande orsak / princip relativt det handlingsförslag som tas i beaktande, eller den fortgående existensen som agent i en viss kropp, eller existensen som en verkande rationalitet som kan verka i andra människor likaväl som i en själv? Om det är den fortgående existensen som agent i en viss kropp eller som verkande rationalitet i en själv likväl som i andra så verkar det klart att utilitarismen kan komma i konflikt med uppfyllandet av konstitutiva normer, till exempel om största välmåga realiseras med hjälp av en hallucinatorisk och paralyserande drog eller en dos bedrägeri eller självbedrägeri. Men inga sådana konflikter verkar vara möjliga om det i stället bara är existensen som bestämmande orsak relativt ett konkret handlingsförslag som måste betraktas som ett absolut värde.¹⁰ De normativa

¹⁰ Det finns förstås också andra sätt att se på vad det innebär att vara en helhjärtad utilitarist än det som Korsgaard tar upp. Antag till exempel att vi förstår en utilitarist som någon som accepterar följande generella påbud: *du skall endast handla så att din handling inte har sämre konsekvenser än någon alternativ handling*. Påbudet anger hur man ska välja bland tillgängliga *handlingar*, men vad Korsgaards teori säger är att *vissa beteenden inte fullt ut är handlingar*, och att dessa beteenden därför är uteslutna för moraliskt ansvariga varelser. Det borde alltså inte *kunna* vara någon konflikt mellan accepterandet av Korsgaards teori och ovanstående utilitaristiska påbud, bara mellan hennes teori och

→

konsekvenserna av det effektiva beslutandets absoluta värde är alltså beroende av *vems* effektiva beslutsfattande som är i fråga, precis som det var i frågan om det rationella självmordets möjlighet.

3. Lagstiftandets omfång och skälens natur

Både i fallet med det effektiva beslutsfattandets absoluta värde och i fallet med det kategoriska imperativet så såg vi att substansen i de normer som ges av för handling konstitutiva principer verkar vara avhängig hur den moraliskt ansvariga varelsen uppfattar lagstiftningens eller den verkande rationalitetens omfång, dvs hur den uppfattar sig själv. Korsgaard försöker mycket riktigt komplettera de kantianska argumenten ovan för att visa att vi som agenter måste tänka oss som lagstiftare för *varje* reflekterande beslutsfattare och måste uppfatta existensen av effektivt reflekterande beslutsfattande *var som helst och när som helst* som absolut värdefullt. Hennes argument bygger på premissen att vi som agenter inte uppfattar → *tillämpningar* av det utilitaristiska påbudet på vad hennes teori klassificerar som icke-handlingar. I modern konsekventialistisk litteratur behandlas agenten för de handlingar som ska bedömas moraliskt ofta som inte mycket mer än en lokal orsak till någon av en mängd för hennes kropp möjliga beteenden vid en viss tidpunkt. Många av dessa beteenden skulle inte vara handlingar utförda av moraliskt ansvariga varelser enligt Korsgaards teori.

skäl som privata eller subjektiva, utan som gemensamma eller objektiva på följande sätt: de talar om hur *jag* ska handla som person utsträckt i tiden och hur *vi* ska handla som en grupp. Det jag måste eftersträva i den mån jag alls försöker handla skall alltså vara att principerna för mina beslut är verksamma inte bara hos mig-nu, utan hos mig som individ vars existens löper över lång tid och hos oss tagna som enhet, och det är för dessa *stora* enheter som jag skulle behöva lagstifta. I det här avsnittet ska jag därför presentera och i viss mån kritisera Korsgaards argument för att vi som agenter uppfattar skäl som gemensamma, och problematisera de slutsatser som Korsgaard tycks vilja dra.

Korsgaards argument skiljer sig från de flesta rationalistiska försök att visa att vi har plikter gentemot andra. Sådana försök utgår normalt från antagandet att var och en har privata (agent-relativa) skäl och försöker visa att vi för att vara rationella måste värdera andra på samma sätt som vi värderar oss själva. Men förutom att dessa argument sällan övertygat – de tycks bara visa att andra måste värdera *sig själva* som vi värderar *oss själva* – så är tågordningen från privata skäl till gemensamma inte uppenbart rimlig. Som Korsgaard ägnar ansenligt utrymme åt att påpeka så verkar det nämligen inte vara någonting

obegripligt med gemensamma (icke agent-relativa) skäl,¹¹ och om skälen dessutom uppfattas som gemensamma *från början* så behövs inget *argument* för att vi måste ta hänsyn till andra. Om vi uppfattar skälen som i första hand skäl för oss att handla på ett visst sätt så uppfattar vi (i första hand) de beslut vi tar på grundval av dessa skäl som bestämmningar av hur *vi* ska handla, dvs det lagstiftande som åsyftas i det kategoriska imperativet har *vårt* handlande som omfång, och det är *vårt* bestämmande av *våra* handlingar som måste uppfattas som absoluta ändamål. Men att skälen är gemensamma verkar i Korsgaards tappning inte bara inbegripa att de gäller för gruppen som helhet utan också att de gäller för var och en: det är genom att var och en inser vad skälen för handling är som var och en kommer att bidra

¹¹ I *The Sources of Normativity* spelar Wittgensteins privatspråksargument stor roll i det här sammanhanget. I en tolkning säger det argumentet att om riktigheten i tillämpningen av ett uttryck inte ligger i annat än att den *tycks* vara riktig så förloras distinktionen mellan riktighet och felaktighet och därmed uttryckets innehåll: det blir tillämpligt på vad som helst och ingenting. Korsgaard vill säga något liknande om skäl: eftersom vi kan ta miste på om någonting är ett skäl för handling eller ej så måste det som avgör om någonting är ett skäl eller inte vara något annat än att det *tycks* oss vara ett skäl. Och om skäl inte är omedelbart och därmed ofelbart givna i ett privat medvetande så skulle andra kunna vara involverade i att avgöra vad som faktiskt är skäl för handling. Se *SN* s. 137-9.

till det gemensamma handlandet. Detta skulle då också innebära att vars och ens bidrag till det gemensamma handlandet är ett mål i sig för oss.

Eftersom Korsgaard tänker sig att plikter stammar ur vår strävan att bestämma våra handlingar så måste analysen av skäl förstås som en analys av ett moment i denna strävan. Om vi från början uppfattar skäl som gemensamma så tyder detta på att vi från början strävar efter ett gemensamt handlande; om vi från början strävar efter ett gemensamt handlande så borde detta omvänt innebära att vi från början uppfattar skäl till handling som skäl för ett gemensamt handlande.

Om vi antar att vi *kan* uppfatta skäl som gemensamma och alltså strävar efter gemensamt lagstiftande är frågan alltså om vi också *från början* gör det. Korsgaards argument för denna uppfattning verkar ha två centrala premisser, som stöds av ytterligare argument:

- (1) Vi uppfattar ibland skäl som gemensamma.
- (2) Vi skulle inte kunna konstruera gemensamma skäl utifrån privata skäl.

Slutsatsen tycks vara att vi från början uppfattar skäl som gemensamma. Men innan vi tittar på vad som följer från premisserna så ska vi se hur Korsgaard försöker underbygga dem.

(1) Argumentet för att vi ibland uppfattar skäl som gemensamma bygger på exempel som det följande: Du behöver träffa mig för att diskutera innehållet i en artikel och föreslår en tid senare samma dag, men jag svarar att jag är väldigt upptagen och föreslår i stället nästa morgon. Du svarar att vi ses då. Vad som *verkar* hända här är att jag tar för givet att ditt behov av att träffas ger oss skäl att träffas, att mina planer för dagen ger oss skäl att inte mötas i dag, och att du uppfattar dem på samma sätt när du blivit upplyst om dem. Och mer allmänt menar Korsgaard att den normala reaktionen på andras önskemål inte är att fråga sig vad dessa har att göra med en själv: vi uppfattar dem inte som skäl först när vi har relaterat dem till egna, privata skäl. I stället så uppfattar vi dem normalt som skäl för vårt eget handlande *om vi inte har skäl för att ignorera dem*. Och det är precis så vi också uppfattar egna önskemål av olika slag: de ger oss skäl att handla, om vi inte har skäl att ignorera dem.

Mänsklig interaktion kan se annorlunda ut förstås: det skulle kunna vara så att du bara uppfattar mina skäl som faktorer som kausalt bestämmer mitt beteende, och därmed som faktorer som påverkar vad du kan få mig att göra. I sådana fall kan du försöka modifiera dessa faktorer så att jag ska bli mer användbar för dina syften. Men vi känner till skillnaden mellan dessa två fall: skillnaden mellan en interaktion där vi delar skäl och ett där vi försöker få de andra att göra något. Alltså uppfattar vi

åtminstone ibland skäl som gemensamma, vilket är detsamma som att vi uppfattar oss som utgörande *en* integrerad agent vars beteende ska koordineras genom ett gemensamt beslut.¹²

(2) Om vi åtminstone ibland behandlar skäl som gemensamma, varför kan vi inte göra detta av privata skäl? Svaret är helt enkelt det att om jag behandlar vad du uppfattar som skäl som om de också vore mina *därför att detta skulle gynna mina privata mål* så skulle jag ändå i grunden förstå detta till synes respektfulla beteende som avhängigt av att det gynnade mina privata mål. Skälen skulle därmed inte uppfattas som gemensamma, bara sammanfallande.¹³

Korsgaard's slutsats är alltså att vi *från början* behandlar skäl om gemensamma. Men det är tveksamt om Korsgaard's argument visar det hon tycks tänka sig att det visar, även om premisserna skulle vara riktiga. Från att vårt behandlande av skäl som gemensamma inte kan vara medvetet och rationellt valt utifrån privata skäl så följer på sin höjd att vi i de fall där vi

¹² Se *SN* s. 140-3 samt *HF* VI. Liknande argument skulle kunna framföras vad gäller vår interaktion med oss själva över tiden.

¹³ En möjlighet som Korsgaard inte nämner men som diskuterats ingående bland annat av David Gauthier i *Morals by Agreement* (Oxford: Clarendon 1986) är att *nu* välja att *senare* komma att behandla andras skäl som sina egna. Men notera att man i ett sådant val inte såsom beslutsfattande skulle kunna identifiera sig med sitt senare jag: den beslutsprincip som man följer nu kommer man ju att förkasta senare.

faktiskt behandlar skäl som gemensamma inte valt att göra detta på grundval av privata skäl. Det följer inte att vi *alltid* från början uppfattar skäl som gemensamma. En möjlighet är nämligen att skälens omfång väljs från fall till fall, men inte medvetet av oss som förnuftiga agenter, utan av våra instinkter eller vårt instrumentella tänkande. Resultatet skulle då vara att en agent lagstiftar för dem hon faktiskt identifierar sig med och alltså respekterar, men att medlemskap i den gruppen avgörs av utomrationella faktorer. Och det resultatet är förenligt med att vissa radikalt avvikande individer – vissa sociopater, till exempel – *alltid* handlar utifrån privata skäl: Korsgaard ger oss nämligen ingen anledning att tro att privata – strikt individrelaterade – skäl är obegripliga.

Denna mer modesta slutsats är kanske nog för att lösa det problem som Korsgaard i första hand försöker lösa: att visa *hur* det kommer sig att vi har oavvisliga plikter gentemot andra och oss själva, handlingar som vi inte kan välja bort om vi alls vill handla. Däremot skulle den vara för svag för att stödja någonting annat som hon vill säga, nämligen att *den praktiska reflektionens egna normer* är sådana att om man följer dem så kommer man att värdesätta *allas* självbestämmande på det sätt som man värdesätter sitt eget självbestämmande när man försöker bestämma sitt handlande.¹⁴ Det må vara så att vi *kan* komma att

¹⁴ *SN*, s. 254-8

delar andras skäl och därmed se deras rationella självbestämmande som lika värdefullt som det egna. Men det är en annan fråga om vi bryter mot några för det praktiska förnuftet konstitutiva normer om vi aldrig kommer att uppfatta deras skäl som våra. Det har inte Korsgaard argument visat: att skäl är *tillgängliga* för gemensamt lagstiftande innebär inte att de är gemensamma för alla rationella varelser. Och en stunds reflektion över interaktionens förutsättningar understryker detta ytterligare.¹⁵

4. Interaktionens förutsättningar och lagstiftandets omfång

Utöver det hypotetiska och det kategoriska imperativet, menar Korsgaard, borde det finnas en princip som reglerar målkonflikter av intra- och interpersonellt slag och är konstitutiv för komplexa agenter såsom vanliga människor eller mänskliga samfund. Om olika mål ska kunna uppfyllas med en och samma kropp eller av en och samma grupp människor – en och samma

¹⁵ Motsvarande kvalifikation måste hållas i åtanke vad gäller Korsgaards mycket problematiska diskussion om våra plikter gentemot djur. (Se *SN* kap. 4, speciellt s. 153.) Vi kanske ofta uppfattar djurs intressen som krav på samma sätt som vi uppfattar andra människors intressen som krav, men det innebär inte att vi som agenter *måste* uppfatta saken på det sättet.

samhällskropp – måste nämligen någon form av samordning eller prioritering till.¹⁶

Korsgaard har tyvärr inget riktigt förslag på hur en princip för detta skulle vara beskaffad, och varken det kategoriska eller det hypotetiska imperativet verkar ge mycket till vägledning eftersom de inte säger något om *hur* olika ändamål ska vägas eller aggregeras.¹⁷ Och detta är ingen trivial brist: även om Korsgaard inte uppmärksammar det så slår denna samordningens nödvändighet för effektivt handlande åt två håll. Det är nämligen så att jag knappast kan lagstifta effektivt för en kropp – min mänskliga lekamen eller en samhällskropp – om inte dess temporala och spatiala delar styrs av *en och samma* prioriteringsprincip, eller åtminstone av principer som sammanfaller med avseende på den sortens mål som kommer i konflikt i det aktuella fallet. Och om jag inte kan lagstifta för en

¹⁶ Se *HF II*.

¹⁷ Är det inte uppenbart att det handlingsalternativ som är *bäst på det hela taget* (bland de alternativ som i övrigt uppfyller för handling konstitutiva normer) som ska realiseras, dvs det alternativ som maximerar värde? Problemet är att ge substantiellt innehåll åt tal om vad som är ”bäst på det hela taget”. Att säga att vi ska maximera *antalet* uppnådda mål verkar orimligt eftersom mål kan vara av så olika vikt, och att säga att det är summan av de tillfredsställda målens vikt som ska maximeras lämnar helt öppet vilken vikt de olika empiriska målen ska ges: det är just det som en prioriteringsprincip behöver svara på.

kropp så är det svårt att se hur det lagstiftande som det kategoriska imperativet påbjuder kan omfatta den kroppen, och hur jag ska kunna praktiskt identifiera mig med den. Problemet är omfattande eftersom det verkar klart att människor faktiskt har olika prioriteringsprinciper, beroende på läggning och olika sorters historiska och personliga band. Alltså kan vi inte med klar blick praktiskt identifiera oss med varje annan moraliskt ansvarig varelse. Vi kan förstås anpassa oss till en annans prioriteringsprincip, men vi kan inte anpassa oss till två andra människor som har substantiellt olika prioriteringsprinciper. Kontentan är förstås att människor (sociopater möjligen undantagna) bara kan vara den sortens varelse som de strävar efter att vara i rätt sociala miljö: endast där deras egna och andras prioriteringsprinciper sammanfaller.

5. Misslyckanden och absoluta handlingsnormer

En teori som säger att moraliska måsten är konstitutiva för handling verkar utesluta omoraliskt handlande: om ett beteende är omoraliskt så tycks det inte vara något handlande, och den som beter sig omoraliskt tycks inte vara någon agent. Detta vore en synnerligen kontraintuitiv konsekvens, speciellt om människor (som det verkade i förra avsnittet) bara lyckas bestämma sitt handlande under väldigt speciella omständigheter. Korsgaard ägnar följaktligen en del utrymme åt att poängtera att det finns *grader* av misslyckande vad det gäller

sådana måsten och motsvarande grader av handlande.¹⁸ Så länge en människa kan tillskrivas en strävan att konstituera sig själv som agent – en strävan att följa de för handling konstitutiva normerna – så kan man tala om att hon utför en *misslyckad* handling eller är en *misslyckad* agent i den mån som hon inte lever upp till normerna. Men jag tror att vi inom ramarna för Korsgaards filosofi kan och bör säga något mer om vissa fall av vad som intuitivt verkar vara medvetet, avsiktligt omoraliskt *handlande*. Enligt vad jag sagt ovan så är det helt möjligt för någon att medvetet och avsiktligt handlar mot moralens normer, nämligen om hon i beslutsögonblicket identifierar sig med en *liten* praktisk enhet: om hon ser sig själv som ensam och bara lagstiftar för sig själv, och kanske bara för sig själv för stunden. Det den medvetet omoraliska människan ändå uppfattar som moralens krav och som hon handlar emot skulle då kunna vara en strävan som hon som människa – men inte som liten *agent* – har att vara en social varelse, dvs en större praktisk enhet. Som människa misslyckas hon med att vara den sorts agent som hon försöker vara, men som liten agent så *lyckas* hon fullt ut bestämma sitt eget handlande.

Förutom att Korsgaards teori om normativitet verkar lämna utrymme för det här sättet att uppfatta avsiktlig omoral så tror jag att det är fenomenologiskt träffande. Tolkningen kan

¹⁸ HFV

emellertid tyckas vara i strid med Korsgaards påstående att var och en som uppfattar någonting som värdefullt – som försöker bestämma sig som orsak till att någonting uppnås eller bevaras – också måste uppfatta moraliska plikter:¹⁹ en agent som uppfattar sig som *praktiskt ensam* skulle med rätta kunna uppfatta sig som den som bestämmer över sina handlingar utan att behöva uppfatta sig som lagstiftande över andra, och utan att respektera andra. Men jag är inte så säker på att Korsgaard behöver uppfatta detta som ett problem. Hon skulle nämligen kunna säga att hon vänder sig till *sociala* varelser, nämligen till nästan alla människor och i synnerhet till de tvivlare som uppfattat moralens krav som krav (vilket innebär att de strävar efter att vara sociala agenter) men undrar vad de har för grund: ambitionen är kanske att visa att en viss sorts kategorisk praktisk nödvändighet är möjlig för en viss sorts varelse, inte att den gäller varje agent. För dessa varelser är moralens bud kategoriska måsten: *som sådana varelser* kan de inte helt lyckas handla utan att följa buden och inte heller medvetet välja att ignorera dem.²⁰

¹⁹ Se *SN*, speciellt s. 163.

²⁰ Denna tolkning verkar synnerligen rimlig i ljuset av Korsgaards karakterisering av vad hon kallar ”den normativa frågan”, vilket är den fråga som hon hoppas att hennes teori ska besvara: den normativa frågan ställs nämligen *av någon som uppfattat moralens krav* (men undrar om de eventuellt är grundlösa och går att ignorera). Se *SN*, kapitel 1.

Det kan tyckas som om denna svaga tolkning av Korsgaards teori är i strid med vad hon säger i slutet av *SN*, nämligen att det är förnuftets egna normer som gör att vi har moraliska plikter. Kort sammanfattat verkar hennes resonemang där vara följande:

Det reflekterande praktiska förnuftets funktion är att ifrågasätta olika handlingsprinciper för att se om de verkligen måste accepteras. Men den enda handlingsprincip som ovillkorligen måste accepteras är den att välja sin egen lag. Och om jag identifierar mig med den principen så måste jag eftersträva dess realisering överallt: skäl är gemensamma. Alltså måste jag högakta och identifiera mig med vars och ens förnuft varhelst det förekommer, dvs lagstifta för varje varelse med reflekterande praktiskt förnuft.²¹

Ingenstans i resonemanget verkar det förutsättas att jag har någon strävan att vara social varelse, vilket jag precis föreslagit att Korsgaard tar för givet. Men antagandet att skäl är gemensamma inbegrep just detta antagande: att uppfatta skäl som gemensamma är just att sträva efter gemensamt lagstiftande. Så den föreslagna tolkningen kan fortfarande vara riktig.

²¹ Se framför allt sidorna 254-8.

Alldeles bortsett från detta måste Korsgaards argument emellertid kvalificeras på det sätt som vi gjorde i förra avsnittet. För att kunna samarbeta måste vi styras av liknande prioriteringsprinciper, och eftersom människor faktiskt har olika principer så innebär det att *även* om jag identifierar mig med det kategoriska imperativet så *kan* jag inte identifiera mig med varje moraliskt ansvarig människa. Detta må vara ett misslyckande för mig som människa, men inte ett som jag rimligen kan hållas ansvarig för.

6. Några avslutande anmärkningar

Enligt Korsgaards teori stammar de förpliktelser som tillkommer moraliskt ansvariga varelser ur dessa varelsers strävan att vara en viss sorts agenter. En stor del av hennes författarskap ägnas just åt en karakterisering eller analys av vad det är att vara en agent – att bestämma sitt handlande – men detta är knappast något försök att utröna *meningen* hos ett visst uttryck ("moraliskt ansvarig varelse", "bestämma sitt handlande") eller att med övertalningsdefinitioner få folk att acceptera en viss normativ ståndpunkt. Snarare är det ett försök att ange vad det är för egenskaper hos moraliskt ansvariga varelser som gör att de uppfattar olika sorters plikter och krav. I detta tar hon för givet att vi har ett grepp om vilka varelser som är moraliskt ansvariga utan att vi för den skull vet exakt vad det är hos dessa varelser som gör att de uppfattar olika sorters krav och plikter: detta är

den fråga som hon försöker besvara. Korsgaards svar är alltså att det just är deras strävan att bestämma sina handlingar på det i teorin karakteriserade sättet. Tanken är att det svaret både förklarar den pliktens fenomenologi som vi är bekanta med och visar att denna fenomenologi inte bygger på någon illusion: de krav som vi uppfattar som ovillkorliga plikter är verkligen rationellt ofrånkomliga.

Så här långt har jag inte ifrågasatt Korsgaards karakterisering av innehållet i vår strävan att vara en viss sorts agenter, bara försökt komplettera den och pekat på vad Korsgaards argument *inte* visar, även om vi i stort accepterar hennes premisser. Men några korta reflektioner över ämnet kan vara på sin plats.

Det är förstås i allmänhet problematiskt att tillskriva varelser omedvetna strävanden, och speciellt problematiskt blir det om det eftersträvade målet i många fall är långt ifrån uppfyllt: hur vet man då att strävan räcker hela vägen fram, att varelsen inte bara strävar att nå en bit i den riktningen? Och precis så verkar det vara med det mål som Korsgaard (implicit) postulerar, nämligen målet att vara en *maximalt social* självbestämmande varelse. Det kanske är så att social strävan visar sig i någon form av missnöje och otillfredsställelse (till exempel genom skuld känslor, ånger eller irritation) med alla former av icke-samarbete. Men det är också ett faktum att vi kan vara missnöjda med många former av samarbete och detta

hänger rimligen ihop med en strävan vi har att inte göra våld på våra egna prioriteringsprinciper, dvs med vad vi finner viktigt: vi tenderar att identifiera oss med likar och med människor tillsammans med vilka vi uppnår saker vi finner viktiga. I ljuset av detta tycks följande karakterisering av människors strävan att vara agenter vara minst lika rimlig som den Korsgaard lutar sig mot:

Människor strävar efter att som reflekterande agenter identifiera sig med andra varelser (och sig själva tidigare eller senare) i den mån som en sådan identifikation bäst gynnar sådant som de finner viktigt.²²

Givet att en människa såsom reflekterande agent identifierat sig med en viss grupp och en viss utsträckning av sin biologiska existens så skulle denna sammansatta agents effektiva handlande vara ett mål i sig för den reflekterande agenten, ett mål som hon som reflekterande agent inte kan underordna andra mål. *Givet* den identifikationen kan hon alltså ha ovillkorliga plikter på det sätt som Korsgaard tänker sig. Men om hennes

²² Här skiljer jag på människan och den reflekterande agenten: människan är den biologiska varelse som under en stor del av sin levnad till ganska stora delar styrs med rationella val, dvs som en reflekterande agent; den reflekterande agenten är den verkande princip som styr ställningstagandet till maximer som människan blir medveten om.

egna eller andra gruppmedlemmars mål förändras så kommer människan att sträva efter en annan identitet som reflekterande agent, en identitet som sedan kommer att implicera andra förpliktelser. Och i den mån hon är medveten om att människor faktiskt har olika prioriteringsprinciper så kommer gruppen i fråga aldrig vara mänskligheten som helhet. Upplysningsmoralen skulle *inte* vara sann: ja, *ingen* specifik moral skulle vara sann för alla människor eller alla självbestämmande varelser, eftersom människor strävar efter olika identiteter såsom reflekterande agenter och eftersom den reflekterande agentens plikter bestäms av hennes identitet. För på den sista punkten tror jag trots allt att Korsgaard har helt rätt: det finns en sorts ovillkorlig praktisk nödvändighet som stammar ur vår identitet som reflekterande agenter.²³

²³ Jag har fått användbara synpunkter på tidigare versioner av denna artikel av ett stort antal personer, bland annat flera studenter på C-, D- och forskarnivå. Speciellt värdefulla och genomarbetade var kommentarer från Niklas Möller, Lars-Erik Hjertström, Åsa Carlsson, Caj Strandberg, Johan Brännmark, Jan Österberg och Bo Peterson