

En internalistisk teori om moraliska uppfattningar

Ett traditionellt föremål för moralfilosofiska studier är våra moraliska uppfattningar, exempelvis uppfattningar om att det är vår moraliska plikt att utföra vissa handlingar. Många filosofer har tänkt sig att dessa uppfattningar är *föreställningar* om en objektiv moralisk verklighet; föreställningar om att vissa handlingar har vissa objektiva moraliska egenskaper. Låt oss kalla denna tanke för "moralisk kognitivism". Andra har tänkt sig att våra moraliska uppfattningar är en sorts *inställningar* eller attityder till handlingarna i fråga. Vår motivation att handla i enlighet med våra moraliska uppfattningar kommer enligt det här synsättet åtminstone delvis från uppfattningarna själva snarare än från en av uppfattningarna oberoende – extern – önskan att handla rätt. Låt oss kalla denna tanke för "moralisk internalism". Somliga har omfattat både kognitivism och internalism, medan många har uppfattat dem som oförenliga.

En stor del av det filosofiska intresset för internalismen stammar från Humes argument mot att "förnuftet" (dvs vår förmåga att uppfatta hur världen är beskaffad) ensamt kan leda till några *moraliska* slutsatser. Med något andra ord än Humes skulle man kunna säga att argumentet är detta: (1) Förnuftet ger oss enbart en karta över världen. (2) Enbart en karta leder inte till målriktad aktivitet (handling): det behövs mål. (3) Endast inställningar ger oss mål. (4) Våra moraliska uppfattningar leder (av sig själva) till handling. Alltså: (5) Våra moraliska uppfattningar är – åtminstone delvis – inställningar. Till detta lägger Hume premissen att (6) inställningar inte kan vara förnuftiga eller oförnuftiga i betydelsen sanna eller falska. Detta tycks innebära att (7) moraliska uppfattningar inte helt och hållet faller under förnuftets och sanningssökandets domvärjo: vi skulle kunna veta allt som finns att veta om denna värld utan att gå med på någon enda moralisk uppfatt-

ning¹. Med andra ord tycks internalismen vara oförenlig med kognitivismen.

I slutet av denna artikel kommer jag helt kort att återkomma till Humes argument. Min huvudsakliga avsikt är emellertid att kortfattat (och därmed ofullständigt) presentera en moralpsykologisk teori som implicerar att vissa typer av moraliska uppfattningar är inställningar. Tidigare har jag argumenterat för teorin i min avhandling, *Moral Internalism, An Essay in Moral Psychology* (Stockholm: Stockholms Universitet 1998). Eftersom Ingmar Perssons recension av denna bok (Filosofisk tidskrift 4/98, s 50–61) innehöll en del missförstånd skall jag också försöka reda ut några av dessa i ett appendix. Av utrymmesskäl kommer jag att vara mycket sparsam med referenser annat än till Perssons recension: den intresserade hänvisas till avhandlingen.

1. En naturalistisk och funktionalistisk teori

Den teori som jag skall behandla försöker säga något om de *psykologiska tillstånd* i vilka vi befinner oss när vi har olika moraliska uppfattningar, men försöker till exempel inte direkt säga något om *meningen* hos våra moraliska termer eller yttranden. Teorin är empirisk då den ser våra moraliska uppfattningar som en del av den verklighet som är kausalt tillgänglig för våra sinnen. Den är därför på olika sätt öppen för empirisk verifikation och falsifikation, men den är också relativt abstrakt och spekulativ, vilket gör att enstaka observationer av vårt beteende eller vårt inre knappast säger något om dess riktighet.

Utgångspunkten är att det i ett stort antal fall står klart huruvida vi har eller inte har en viss moralisk uppfattning och teorin avses i första hand erbjuda en korrekt och intressant analys av skillnaden mellan dessa "paradigmatiska" fall: vad har tillkommit när en moralisk uppfattning börjat omfattas; vad har försvunnit när den givits upp? Vi känner förstås till gränsfall där det med viss rimlighet men knappast helt utan förbehåll går att tillskriva en person en moralisk uppfattning. Att teorin inte helt och hållet stämmer in på dessa fall kan inte ses som en avgörande invändning, lika lite som våra teorier om vattnets kemiska beståndsdelar falsifieras av att någonting som bara med förbehåll skulle kunna kallas rent vatten avviker från vad som kan förutsägas utifrån dessa teorier.

¹ Se till exempel *A Treatise of Human Nature*, Book III, Part I, Section I.

För en fruktbar analys krävs ett analyspråk som passar analysandum. Jag har funnit det användbart att tala i termer av de *psykologiska funktionerna* hos våra moraliska uppfattningar och hos de psykologiska mekanismer som producerar dessa uppfattningar. I korta ordalag uppfattar jag funktioner hos en typ av tillstånd eller mekanismer som vissa effekter som den typen haft, effekter som förklarar typens fortsatta existens. (Förklaringen kan vara på individualpsykologisk eller genetisk nivå.)

Jag skall använda termerna "föreställning" respektive "inställning" för att göra funktionella klassificeringar av psykologiska tillstånd. Att ett tillstånd är *föreställande* skall jag förstå som att det är en del av en komplex mental struktur som utgör en persons föreställningsvärld, en struktur vars funktion är att tjäna som en inre arsenal av kartor efter vilka personen kan orientera sig i världen och vars ingående element justeras av mekanismer vars funktion det är att åstadkomma någon form av överensstämmelse mellan dessa element och element i världen. Om den del av kartan som utgör en föreställning överensstämmer med faktiska förhållanden i världen kan vi säga att den är sann, annars inte. Jag skall alltså anta att det filosofiskt problematiska uttrycket "överensstämmer med faktiska förhållanden" kan ges en rimlig innebörd som gör detta riktigt. Jag skall vidare anta att våra vanliga uppfattningar om sakers läge i tid och rum, deras form, tyngd, färg är föreställningar, och detsamma gäller uppfattningar om att en stat har ett visst statsskick, att en grupp har en viss lönespridning, eller att en människa har vissa känslor eller tankar.

Att något är en *inställning* skall jag förstå som att det har som funktion att få element i världen att överensstämma med element hos inställningen. Sådana tillstånd är målgivande för vårt beteende och bestämmer tillsammans med våra föreställningar hur vi handlar. Till våra inställningar hör rimligen önsknings, avsikter och mycket (men inte allt) av det vi kallar emotioner eller känslor. Den teori om moraliska uppfattningar som skall försvaras här säger alltså (bland annat) att (viktiga sorters) moraliska uppfattningar är en slags inställningar. Inställningar inbegriper ett visst mått av motivation att handla i enlighet med dem, och i denna mening försvarar jag en *internalistisk* teori om moralisk motivation: motivation att handla i enlighet med den moraliska uppfattningen är inte beroende av externa inställningar (även om sådana förstås också kan spela in). Vilka moraliska uppfattningar

som är inställningar, och *vilken sorts* inställningar som är moraliska uppfattningar skall vi nu titta närmare på.

2. *Moraliska emotioner*

Låt mig börja med några moralpsykologiska observationer som är svåra att förbigå om man försöker förstå moralen som psykologiskt fenomen:

(1) Vi tenderar att bli indignerade eller arga på personer som utför (eller funderar på att utföra) andra handlingar som vi uppfattar som moraliskt felaktiga, och tenderar att acceptera/gilla de som utför (eller funderar på att utföra) handlingar som vi uppfattar som moraliskt riktiga eller berömvärda.²

(2) Vi tenderar att känna skuld när vi anser att vi själva handlar eller funderar på att handla moraliskt fel, men gillande och självrespekt när vi handlar som vi själva anser rätt eller moraliskt berömvärt.

(3) Vi tenderar att bli indignerade eller arga när vi uppfattar att någon handlar fel men inte känner skuld, och denna upprördhet mildras när den felande uttrycker skuld.

Som jag försöker visa i *Moral Internalism* så det är det rimligt att anta att våra moraliska uppfattningar stammar från emotioner såsom ilska, skuld och praktiskt gillande. Både ur ett naturhistoriskt och ur ett individualpsykologiskt perspektiv kan man se vår kapacitet att forma moraliska uppfattningar som en förfining av vår kapacitet att reagera med sådana emotioner. I nästa avsnitt skall jag skissa denna förfining; i det här avsnittet skall jag säga något om emotionernas psykologiska funktion.

Som jag förstår *emotioner* (som förutom ilska, skuld, och praktiskt gillande inbegriper sådant som glädje, sorg och avsmak) så är de en samling relativt oavsiktliga psykologiska och kroppsliga reaktioner som förbereder oss fysiskt och mentalt för en typ av handlingar genom

2 Vi har nog ingen vardagspsykologisk term för det praktiska gillandet, men fenomenet är bekant. När handlingar som vi uppfattar som moraliskt berömvärda är "riktade" mot oss så tar sig det praktiska gillandet uttryck i tacksamhet; när vi själva gjort något som vi uppfattar som moraliskt berömvärd så känner vi självrespekt; och lite introspektion tror jag ger vid handen en analog emotion i fall där någon annan är föremålet för godheten.

förändring av muskelspänning, andning, och mentalt fokus: ilska förbereder oss för aggressiv handling, företrädesvis gentemot objektet för vår ilska; skuld förbereder oss för viss form av underdånighet. Emotioner är alltså en sorts målgivande tillstånd, precis som önskningar. *Känslan* av skuld, glädje, avsmak, etcetera är känslan av vara i dessa målgivande tillstånd.³ En *önskan* skall jag se som ett tillstånd vars funktion det är att ge upphov till vissa emotiva reaktioner; dessa reaktioner uppfattar jag som de medel genom vilka önskningar påverkar våra handlingar.

Just indignation, skuld, och acceptans – den sorts emotioner som vi förknippar med moraliska uppfattningar – verkar också *signalera* denna beredskap till handling. En del av de förändringar som utgör dessa emotioner är mer eller mindre tydliga förändringar i ansiktsuttryck som man kunde förvänta sig vore obetydliga för vår förmåga att handla: jämför med förändring i andning, hjärttakt och hormonutsöndring, eller muskelspänning i bål eller lemmar. Ilskan liksom skulden signalerar sin egen existens till andra, och detta har viktiga sociala funktioner: den som ilskan är riktad mot får en möjlighet att underordna sig den arge, att böja sig för dennes vilja eller att visa beredskap för sådan underordning: dvs att signalera skuld. Och när det moraliska gillandet av en handling gör sig känt så inbjuder det till fortsatt samarbete eller samlevnad och till fortsatta handlingar av det nämnda slaget.⁴

Följande punkter summerar min uppfattning om några viktiga psykologiska funktioner hos de ovan nämnda moraliska emotionerna:

Indignation/ilska har som funktion att (a) väcka skuld hos andra och att därigenom (b) åstadkomma avhållsamhet när det gäller liknande handlingar genom att åstadkomma eller förstärka en önskan hos den handlande (och andra närvarande) att inte utföra den sortens hand-

3 Viss empirisk evidens för fruktbarheten i detta angreppssätt ges i Antonio Damasio's *Descartes' Error*, G P Putnam's Sons, New York 1994.

4 Det här är förstås sådant som etologer studerat hos olika sociala djur. En viktig egenskap hos våra (och vissa primaters) moraliska emotioner verkar vara att de inte enbart gäller beteende gentemot en själv eller avkomman. Detta ger de gemensamma moraliska känslorna större kraft, och låter oss sluta upp gemensamt mot vissa sorters handlingar. (Jfr Humes tal om att vi genom vår medmänsklighet formar ett "party of human-kind" mot obehagligt och skadligt beteende. Se *Enquiry Concerning the Principles of Morals*, Section IX, Part I.)

lingar. (Dessutom förbereder ilskan för mer "fysiskt" aggressivt beteende för händelsen att dess objekt inte reagerar med skuld.)

Skuld har som funktion att (a) återställa möjligheten till samarbete genom att blidka aggression, men också att (b) få den skuldyngde att undvika liknande situationer i framtiden genom att orsaka eller förstärka en önskan hos henne att inte utföra den sortens handlingar.

Praktiskt gillande har som funktion att (a) åstadkomma fortsatta handlingar av det gillade slaget genom att förstärka eller skapa en önskan att utföra den sortens handlingar men också att (b) disponera den gillande att välja den gillade för fortsatt samarbete och omvänt.⁵

Utan dessa emotioner och deras förmåga att anpassa vårt beteende efter varandra skulle vi knappast kunna leva i komplexa samhällen.

Notera att funktionen hos de moraliska emotionerna är att uppmuntra eller motarbeta vissa *sorters* handlingar: emotionernas psykologiska funktion är att lära oss själva och andra visst beteende snarare än annat. När vi säger att exempelvis indignation är riktad mot en viss (föreställd) handling så avser vi väl normalt den (föreställda) handling som givit upphov till indignationen, men i en mening är den emotiva reaktionen (och motsvarande moraliska önskan) alltså riktad mot handlingar av det aktuella *slaget*.

När jag talar om de moraliska emotionernas *funktioner* så avser jag som sagt de effekter som dessa emotioner haft och som förklarar att de har en plats i våra psyken. Att de haft dessa effekter implicerar förstås inte att de *alltid* har dessa effekter. Lika lite kan man från det faktum att emotionernas funktioner har att göra med våra tendenser till samarbete dra slutsatsen att emotiva reaktioner bara uppkommer i sammanhang där (vi tror att) sådant samarbete faktiskt står på spel.

3. *Det moraliska omdömet*

Som jag ser det spelar moraliska emotioner en viktig roll i formandet av moraliska uppfattningar, men det är klart att de två inte *okvalificerat* kan identifieras. Vi hyser moraliska uppfattningar under längre

⁵ Utöver skuld, ilska och praktiskt gillande så finns förstås sociala emotioner som kanske var viktigare i de samhällen som skapade grekiska och isländska sagor än vad de är hos oss idag: skam, förakt, stolthet, beundran.

perioder (dagar, månader, år) medan våra tillstånd av ilska eller skuld ofta är högst tillfälliga och varierar från minut till minut eller från dag till dag. Kanske är en moralisk uppfattning i stället ett psykologiskt tillstånd vars funktion det är att ge upphov till sådana reaktioner, dvs en sorts *moralisk önskan* som verkar genom moraliska emotioner? Ja, men inte heller den identifikationen är oproblematisk då vi tycks kunna vara av uppfattningen att vi handlat rätt men trots det känna skuld, och eftersom det tycks som om vår indignation eller skuld ofta föregår formandet av en moralisk uppfattning.

En jämförelse kan göras med våra vanliga uppfattningar om vår fysiska omgivning. Här spelar våra sinnesintryck en högst väsentlig roll i formandet av trosföreställningar, samtidigt som det är klart att våra trosföreställningar inte kan identifieras med sinnesintryck: ibland misstror vi ju våra sinnen, och många föreställningar bildas genom sammanvägning av en mängd intryck eftersom de gäller förhållanden som inte enkelt kan observeras. Tre delvis sammanlänkade moment av "omdöme" verkar komma in: ett av informationssökning, ett av sammanvägning, och ett av censur. Genom att använda våra sinnen och söka i minnet kan vi ta hänsyn till mängder av information som sedan kan sammanvägas i en föreställning som tillfogas den inre arsenal kartor som vi orienterar oss efter. I den avvägningen kan vi undanhålla vissa intryck, minnen eller uppfattningar därför att de är i konflikt med andra föreställningar eller för att vi anser att de uppkommit under omständigheter som gör dem opålitliga, eller för att de är emotionellt påfrestande eller moraliskt stötande (i de senare kategorierna finner vi exempelvis föreställningar om våra egna förmågor och uppträden och föreställningar om statistiska olikheter mellan folkslag).

Min tanke är att moraliska uppfattningar resulterar ur liknande omdömesakter, fast applicerade på våra moraliska önskningar snarare än på sinnesintryck. Skälet till att vi har och utnyttjar den kapaciteten är densamma i båda fallen: det spelar och har spelat stor roll för oss både vilka föreställningar och vilka moraliska önskningar som vi har. De moraliska önskningarna fyller en central funktion när de möjliggör samarbete genom sin påverkan av våra och andras handlingar och genom att själva påverkas av andras emotiva reaktioner.⁶ Vi behöver

6 De psykologiska, sociala och evolutionära mekanismerna bakom detta har diskuterats i olika sammanhang. Bland filosofiska böcker i ämnet märks Edna

ta hänsyn till relativt komplexa sociala och psykologiska förhållanden för att reagera adekvat – det krävs informationssökning. Vi behöver kunna väga den mångfald av faktorer som kan vara relevanta – det krävs sammanvägning. Och vissa sätt att tänka och vissa sätt att reagera behöver hållas i schack – det krävs censur.

Utifrån ovanstående föreslår jag följande analys av några enkla typer av moraliska uppfattningar:

(1) Uppfattningar som går ut på att *en handling är moraliskt felaktig* (att den inte får utföras, moraliskt sett) är önskningsar som ej förkastats i omdömesakt och vars funktion är att framkalla moraliskt ogillande gentemot den som utför handlingen (eller funderar på detta): indignation gentemot andra, skuld gentemot en själv.

(2) Uppfattningar som går ut på att *det är moraliskt fel att inte utföra en handling* (att den är agentens moraliska plikt) är önskningsar som ej förkastats i omdömesakt och vars funktion är att framkalla moraliskt ogillande gentemot den som inte utför handlingen (eller funderar på detta).

(3) Uppfattningar som går ut på att *en handling är moraliskt tillåten eller acceptabel* är tillstånd som ej förkastats i omdömesakt och vars funktion är att avhålla från moraliskt ogillande både gentemot den som utför handlingen (eller funderar på detta) och mot den som inte utför handlingen (eller funderar på detta).⁷

(4) Uppfattningar som går ut på att *en handling är moraliskt berömvärd* är önskningsar som ej förkastats i omdömesakt och vars funktion

Ullman-Margalits *The Emergence of Norms* (Oxford: Oxford U P 1977) och Allan Gibbard's *Wise Choices, Apt Feelings* (Cambridge Mass: Harvard U P 1990).

7 På samma sätt som det är viktigt att ha en viss emotiv reaktion i ett visst sammanhang kan det förstås också vara viktigt att *inte* ha en viss emotiv reaktion i ett annat sammanhang, och en sådan frånvaro av emotion kan förstås också vara resultatet av våra bedömningar. Användande av omdöme i både moraliska och icke-moraliska spörsmål kan förstås också sluta utan bestämt avgörande, vilket resulterar i fortsatt osäkerhet/öppenhet som måste skiljas från ett bestämt/slutet avhållande från vissa omdömen (handlingen är varken fel eller plikt, hans mustach var varken vildvuxen eller välansad).

är att framkalla praktiskt gillande gentemot den som utför handlingen.

4. *Emotionell konflikt, apati, och omoral*

Ovanstående analyser torde kunna förklara den täta koppling mellan moraliska uppfattningar och moraliska emotioner som diskuterades i avsnitt 2. Men de erbjuder också förståelse av ett antal fenomen som har uppfattats som problem för internalistiska teorier, som det faktum att våra moraliska emotioner ibland "motsäger" våra moraliska uppfattningar, att vi ibland inte känner någon som helst motivation att handla i enlighet med våra moraliska uppfattningar, och att vi kan välja att göra något som vi anser vara moraliskt fel:

(1) Om teorin är riktig så är det förstås inget konstigt med att vi reagerar emotivt i strid med vårt omdöme: det är inte konstigare än våra sinnen kan ge oss andra besked än det vi kommer fram till efter att ha tillämpat våra omdömesförmågor. Korsa pek- och långfingret på ena handen och dra dem över näsryggen: det kommer att kännas som om du har två näsor, men du tror inte därför att du har två näsor. Fokusera på att ditt handlande väckt ilska hos personer vars gillande betyder mycket för dig och du kan komma att känna skuld trots att motsvarande moraliska önskan inte överlevt eller kommer att överleva dina omdömesfrågor.

(2) Utmattade, deprimerade, gravt berusade eller sömndruckna kan vi ibland "inse" att det är vår plikt att göra det ena eller det andra utan att därför uppfatta någon särskild vilja att handla på det sättet. Vår internalistiska modell låter oss säga att våra tidigare moraliska omdömen/önskningar är tillgängliga i vårt minne under sådana omständigheter, men att vi är emotionellt ur spel: de mentala mekanismer som normalt tar oss från önsknigen (vars funktion är att ge upphov till emotiva reaktioner på vissa föreställningar) till de emotiva reaktionerna (vars funktion det är att påverka handling i enlighet med önsknigen) är (tillfälligt) utslagna.

(3) Det framstår också som begripligt att vi kan välja att handla i strid med våra moraliska uppfattningar. Även om vi i centrala fall av moralisk deliberation försöker besluta oss för hur vi skall *handla* så gäller den *moraliska* deliberationen primärt den *moraliska* inställningen till den handling som bedöms. Att komma fram till att något är moraliskt fel men ändå besluta sig för att göra det är att forma ett moraliskt ogillande av handlingen, men ändå forma avsikten att utföra

den utifrån ett icke-moraliskt gillande av densamma. (Moraliskt gillande är som bekant inte det enda gillande som vi är kapabla.)⁸

5. *Problem med de klassiska argumenten för internalism*

Traditionella argument för internalism hänvisar oftast till att våra moraliska uppfattningar normalt följer våra önskningar eller emotionella reaktioner och att vi normalt förväntar oss detta. Utifrån internalismen kan vi förklara varför vi tar brist på motivation som tecken på frånvaro av motsvarande moraliska uppfattning, varför moraliska uppfattningar och uttryck för moraliska uppfattningar kan ha en praktisk funktion – att de kan tala om för oss vad vi skall göra – och varför vi förväntar oss att goda, viljestarka personer skall vara motiverade i enlighet med sina moraliska uppfattningar. Men en närmare analys visar att externalister kan förklara dessa och liknande fenomen utifrån till synes rimliga antaganden om att nästan alla har någon önskan att handla rätt eller är disponerade att forma önskningar i enlighet med sina moraliska uppfattningar (som i sig själva inte är moraliska önskningar utan enbart en sorts föreställningar).⁹

En annan sorts argument för internalismen utgår från någon form av icke-kognitivism och försöker visa att moraliska uppfattningar inte kan förstås som föreställningar om en moralisk verklighet. Detta skulle ge de föregående argumenten ny styrka. När man funderar på hur moraliska uppfattningar skall förstås om de *inte* är föreställningar så är nämligen de kopplingar till motivation och moraliska emotioner som påvisats svåra att ignorera: varför då inte helt enkelt anta att moraliska uppfattningar är moraliska önskningar?

Som jag ser det så saknas emellertid övertygande argument för icke-kognitivismen, åtminstone om man inte kan anta att internalismen är riktig (vilket vi förstår inte kan i det här sammanhanget). Traditionella hänvisningar till att det råder stor oenighet i moraliska frågor eller till att vi saknar okontroversiella metoder för att avgöra moraliska

8 Det är vanligt att argumentera mot internalismen med hänvisning till att vi kan *tänka oss* möjligheten av en person som anser att en handling är fel utan att för den skull ogilla handlingen. Men min ambition är att förstå faktiska, paradigmatiska fall av moraliska uppfattningar, inte i första hand vår fantasikapacitet.

9 Jag argumenterar utförligt för detta i kapitel 3 av *Moral Internalism*.

tvistemål verkar till exempel också drabba stora vetenskapliga domäner.

Hur skall man då ge stöd åt en internalistisk teori? Mitt argument i *Moral Internalism* är i korta ordalag följande: Även externalistiska teorier – som alltså säger att moraliska uppfattningar är en sorts föreställningar – måste anta att paradigmatiska fall av moraliska uppfattningar utöver dessa föreställningar ackompanjeras av moraliska önskningar. Om man då kan visa att det räcker med att anta sådana moraliska önskningar för att förklara allt som postulerandet av moraliska föreställningar skulle förklarat så verkar den externalistiska teorin omotiverat komplicerad. Därmed skulle det internalistiska antagandet att moraliska uppfattningar är moraliska önskningar bjuda den enklaste av två i övrigt likvärda förklaringar av de fenomen som vi associerar med moraliska uppfattningar. (Kognitivismen skulle fortfarande kunna vara riktig om Hume har fel i att den är oförenlig med internalismen. Se vidare avsnitt 7.)

6. Inkonsistens, följd och komplexa moraliska uppfattningar

Frågan är om det går att visa att det internalistiska antagandet räcker. Även om analyserna i avsnitten 2 till 4 ovan är riktiga så utelämnar de viktiga beståndsdelar av vårt moraliska tänkande. En stor sådan del är nog våra värderingar av andra företeelser än handlingar. Men det som oftare uppfattas som problematiskt för internalistiska teorier, och det jag skall ägna mig åt i detta avsnitt, är att förklara att vårt moraliska tänkande har en "logisk" struktur:

(1) Vi har intuitioner om att vissa moraliska uppfattningar står i "logiska" relationer av följd eller inkonsistens med andra moraliska uppfattningar. Traditionellt så förklarar man följd och inkonsistens i termer av sanning och falskhet, men det är oklart hur det kan tillämpas på inställningar.

(2) Förutom den sortens enkla uppfattningar som analyserats i avsnitt 3 så har vi komplexa moraliska uppfattningar som på olika sätt verkar stå i relationer av följd och inkonsistens med dessa. Ta exempelvis:

- (A) Om det är fel att ljuga så är det fel att få sin lillebror att ljuga.
- (B) Det är fel att ljuga.
- (C) Det är fel att få sin lillebror att ljuga.

Låt oss kalla de uppfattningar som konventionellt uttrycks av dessa satser för (A*), (B*) och (C*). Det tycks då som om (A*) är *inkonsistent* med *konjunktionen* av (B*) och *negationen* av (C*). Dessutom tycks (C*) *följa* av konjunktionen av (A*) och (B*).¹⁰

För att förklara dessa fenomen behöver vi en idé om vilken psykologisk funktion som tillkommer våra intuitioner om inkonsistens och följd mellan olika uppfattningar, och vilken funktion som tillkommer komplexa moraliska uppfattningar. Mina förslag – som är avsedda att vara giltiga för både moraliskt och icke-moraliskt tänkande – är följande:

Intuitioner om inkonsistens är tillstånd vars funktion är att hindra oss från att acceptera kombinationer av uppfattningar.

Intuitioner om följd är tillstånd vars funktion är att få oss att acceptera vissa uppfattningar givet att vi accepterar vissa andra uppfattningar.

Negationer av uppfattningar är tillstånd vars funktion det är att få oss att inte acceptera det negerade tillståndet.

Villkorsuppfattningar – uppfattningar som normalt uttrycks med satser av typen ”Om det snöar (försats) så är solen i moln (eftersats)”; ”Om någon gjort sin plikt (försats) så är det fel att klandra denna någon (eftersats)” – är tillstånd vars funktion det är att leda till acceptans av det tillstånd som uttrycks av eftersatsen givet acceptans av det tillstånd som uttrycks av försatsen.

Konjunktiva uppfattningar – uppfattningar som normalt uttrycks med satser av typen ”Det snöar (konjunkt 1) och skoveln står i skjulet (konjunkt 2)”; ”Sokrates är grek (konjunkt 1) och om Sokrates är grek

10 Argument som går ut på att moralens logik kräver kognitivism utgår ofta från Peter Geachs ”Assertion” i *The Philosophical Review*, Vol 74, 1965, s 449–65. Ibland uppfattas internalismens problem vad gäller moralens logik som ett problem att *rättfärdiga* våra moraliska slutledningar. Men det problemet kan inte gälla den sortens *deskriptiva* internalism som försvaras här: ambitionen är förklaring och förståelse av de psykologiska aktiviteter och tillstånd som utgör vårt moraliska tänkande. Om internalismen erbjuder en sann beskrivning av detta tänkande men underminerar vår moraliska praxis så är detta ett problem för oss eller vår praxis, men det motsäger inte alls vad internalistiska teorier påstår.

så är han vis (konjunkt 2)” – är tillstånd vars funktion det är att hålla de tillstånd som uttrycks av konjunkterna tillsammans för slutledning så att villkorsuppfattningar kan fylla sin funktion och så att inkonsistens kan upptäckas.

Tillämpat på ovanstående exempel så kan analysen kortfattat se ut som följer: Intuitionen om att (C*) följer av konjunktionen av (A*) och (B*) har som funktion att få en att acceptera (C*) givet att man accepterar konjunktionen av (A*) och (B*). Eftersom acceptans av konjunktionen av (A*) och (B*) har som funktion att få en acceptera (C*) så kommer tänkandet i konflikt om det samtidigt accepterar negationen av (C*): därav intuitionen att konjunktionen av (A*) och (B*) är inkonsistent med negationen av (C*).

Hur kommer det sig då att vi har tillstånd av ovanstående slag: vilken vital funktion kan de ha fyllt? Om vi antar att moraliska uppfattningar är föreställningar rörande element i den moraliska verkligheten så ger sig följande (skissartade) förklaring. Det finns vissa komplexa sätt på vilka den moraliska verkligheten inte kan vara eller sannolikt inte är och vissa komplexa sätt på vilka den måste vara eller sannolikt är: i den mån som det är viktigt för oss att ha korrekta föreställningar om denna verklighet så är det viktigt att vårt tänkande på ett systematiskt sätt avspeglar dessa förhållanden. En sådan avspeglning vore omöjlig om vi inte negerade vissa föreställningar och hade villkorsuppfattningar vars funktion det är att ta oss från (negationen av) vissa föreställningar till (negationen av) andra föreställningar.¹¹

Vad jag hävdar är emellertid att en internalist kan förklara dessa fenomen utan hänvisning till någon kognitiv funktion hos våra moraliska uppfattningar – det räcker med antagandet att de är resultatet av praktiskt omdöme rörande moraliska emotioner. Artikelns omfång tillåter nu inte några detaljerade redogörelser, men grundtanken är i princip och i (väldigt) korta ordalag följande:

(1) De moraliska uppfattningarnas funktion är att påverka handling så

11 En fyllig kognitivistisk-externalistisk förklaring av detta måste rimligen säga något om denna moraliska verklighet: varför det är viktigt för oss att ha korrekta föreställningar om den, och varför den måste vara på vissa sätt samtidigt som den inte kan vara på andra sätt.

att vissa mål uppnås.

(2) Moraliska önskningar kan endast påverka handling och uppnå målen om de är förbundna med andra moraliska önskningar och med icke-moraliska föreställningar på vissa systematiska sätt snarare än andra. Detta förutsätter att vissa önskningar negeras och att vi har villkorsuppfattningar vars funktion det är att ta oss från föreställningar till (negationer av) moraliska önskningar och från (negationer av) moraliska önskningar till föreställningar eller (negationer av) andra moraliska önskningar.

Exakt vilka sådana systematiska förbindelser som frammanas hos en given individ kommer förstås variera, på samma sätt som de exakta systematiska förbindelser som föreligger mellan olika föreställningar varierar. (I andra sammanhang utvecklar jag ovanstående förklaringar och tittar närmare på några sådana förbindelser.)¹²

Om denna förklaring av "logiska" intuitioner och komplexa moraliska uppfattningar är adekvat så tycks vår internalism därmed erbjuda en enklare (men på andra sätt likvärdig) teori om våra moraliska uppfattningar än vad kombinationen av externalism och kognitivism utlovar.

7. *Internalism och kognitivism*

Låt mig avslutningsvis kort kommentera Humes argument från internalism till icke-kognitivism. Hume tänker sig att moraliska uppfattningar består av "passions" och att dessa inte kan vara sanna eller falska eftersom de inte är representationer av verkligheten. Men han förklarar inte varför passioner inte kan vara representationer eller "kopior" och lämnar därmed frågan om internalismen verkligen är oförenlig med kognitivism öppen. (Ingmar Persson menar sig ha ett argument för denna oförenlighet. I appendix nedan visar jag varför det inte duger.)

Jag har antagit att något är en föreställning i kraft av att vara en del av en uppsättning kartor vars element justeras av mekanismer vars funktion det är att få kartan att överensstämma eller "korrespondera" med verkligheten. Idén om en sådan korrespondens är en idé om att det finns vissa regler för hur kartan (inte) skall se ut givet att världen

¹² Se kapitel 4 av *Moral Internalism*, men framför allt "Morals, Reasoning, and Logic: The Emotivist Account" (manuskript).

ser ut på ett visst sätt: låt oss kalla dessa för ”korrespondensregler”. Vi kan tänka oss flera källor till sådana regler, men jag nöjer mig med att nämna några få som förfäktats under detta århundrade: (1) Jag har ovan (och i enlighet med modern teleofunktionalism) tänkt mig regler baserade på biologisk funktion. De justerande mekanismerna har reproducerats för att de *de facto* följt vissa regler: det är deras biologiska funktion att följa dessa regler. (2) Man kan också (i positivistisk anda) fokusera på regler baserade på vad individen omfattar som villkor för att acceptera eller ge upp tillståndet i fråga. (3) Andra är mer intresserade av regler baserade på vad ”språksamfundet” uppfattar som riktigt acceptera eller uppgivande av tillståndet (Wittgenstein à la Kripke, Crispin Wrights disciplinerade syntaxism). (4) Ytterligare andra (Donald Davidson, exempelvis) har intresserat sig för ”transcendentala” regler som följer ur principer för tillskrivande av mentala tillstånd. Frågan gäller väl närmast om vissa av dessa regler är mer intressanta än andra (och denna fråga är kanske framför allt en fråga om vilken poängen är med att följa sådana regler).

En internalist som vill argumentera för en form av kognitivism behöver visa att moraliska uppfattningar i den internalistiska tappningen (essensiellt) är föremål för någon form av intressanta korrespondensregler. Sådana argument har också framförts. Allan Gibbard tänker sig till exempel att en form av internalism som liknar den här diskuterade ger vid handen att moraliska uppfattningar justeras enligt regler av typ (1)¹³ medan John McDowell framhåller existensen av regler av typ (3).¹⁴ Det är också klart att vi till mans håller oss med regler av typ (2), vilket internalister som Richard Hare och Charles Stevenson ofta påpekat, men sådana regler kan tyckas göra moralisk sanning till något outhärdligt subjektivt. Givet mångfalden av möjliga system av korrespondensregler bör vår fråga nog inte vara om moraliska önskningar alls lyder under korrespondensregler, utan om det finns några särskilt *intressanta* regelsystem som ger en *intressant* form av kognitivism. Men den uppgiften finns det inte rum att ta sig an här.

13 *Wise Choices, Apt Feelings*, Oxford U P 1990, s 117–20. Gibbard poängterar dock att dessa regler inte behöver gå hand i hand med regler av typ (2) eller (3).

14 ”Non-Cognitivism and Rule-Following” i *Wittgenstein: To Follow a Rule*, redigerad av S Holzman och C Leich, Routledge and Kegan Paul 1981.

Appendix: Ingmar Persson om moralisk internalism

I sin recension av min avhandling framför Ingmar Persson vad han uppfattar som vägande kritik, dels mot avhandlingens upplägg, och dels mot dess slutsatser. Men jag tror inte att dessa invändningar håller streck, vilket jag skall försöka visa på några punkter. (Jag kommer inte att beröra invändningar som rör min definition av *motivation* eftersom presentationen ovan inte berört detta närmare.)

(1) Persson förordar att internalismen formuleras så att den är neutral mellan kognitivism och icke-kognitivism, och menar att jag misslyckas med detta. Enligt Persson innebär detta i sin tur att jag sammanför element som borde diskuterats separat. Stämmer det? Ovan har jag förstås antagit att min internalistiska teori kan vara förenlig med kognitivism: frågan var om den är förenlig med någon intressant version därav. Här hävdar Persson att den förfäktade formen av internalism är "logiskt eller i djupaste mening" oförenlig med kognitivismen (FT 4/98, s 51). Varför då? Jo, tycks han resonera, det att något "essensiellt" är en föreställning kan inte garantera att det också är en inställning. Alltså är inte den kognitivistiska tesen att moraliska uppfattningar essensiellt är föreställningar förenlig med internalismen som ju också säger att moraliska uppfattningar garanterat – essensiellt – är inställningar.

Felslutet i ett sådant resonemang är uppenbart: för att en viss sorts tillstånd essensiellt skall ha två funktioner behöver inte den ena funktionen garantera förekomsten av den andra. Tänk på förhoppningar och avsikter. När vi har en förhoppning föreställer vi oss – essensiellt – att det vi hoppas på kan inträffa men vi har också – essensiellt – en viss inställning till denna möjlighet. Inte desto mindre kan vi ha föreställningen utan inställningen och omvänt: den ena garanterar inte den andra, men det faktum att vi hoppas garanterar båda. När vi avser någonting så har vi – essensiellt – en positiv inställning till att det avsedda blir fallet, och samtidigt föreställer vi oss – essensiellt – att det kommer att bli fallet: den ena funktionen garanterar inte den andra, men det faktum att vi har en avsikt garanterar båda. Dessa exempel kanske kan ifrågasättas, men poängen är att vissa psykologiska tillstånd essensiellt kan ha flera funktioner utan att dessa funktioner essensiellt är förknippade (annat än som delar i ett sådant psykologiskt tillstånd). Detta visar förstås inte att internalismen oproblematiskt är förenlig med kognitivismen, bara att de två inte är "logiskt" oförenliga

på det sätt som Persson tänkt sig. Och det lämnar öppet för den sorts diskussion som jag skissade i förra avsnittet.

Möjligen är Persson förvirrad av min diskussion i kapitel 4 av *Moral Internalism*. Jag försöker där visa att internalistiska analyser gör det begripligt varför moralisk diskurs uppvisar drag som vi associerar med paradigmatisk faktuell diskurs utan *explicita* hänvisningar till att moraliska uppfattningar är föreställningar. Men att *explicita* hänvisningar till kognitivistiska antaganden inte behövs i en förklaring innebär inte att de element som ingår i förklaringen inte till slut ger kognitivism vid handen (som jag antydde i avsnitt 7). Vad det innebär är att om våra moraliska uppfattningar är föreställningar så kan dessa föreställningars funktion och existens förklaras utifrån internalismen.

Men kanske har Persson ytterligare premisser i tankarna. Den form av kognitivism som är förenlig med internalismen tycks också vara förenlig med att man kan veta allt om världen utan att ha någon bestämd moralisk uppfattning. I linje med vad Persson påpekat så garanterar ju inte det att man har ett tillstånd som lyder under de för moraliska åsikter relevanta korrespondensreglerna och som uppfyller dessa regler på ett reliabelt sätt att man också har den moraliska önskan som enligt internalismen gör att man har en viss moralisk uppfattning. Kanske kan man argumentera för att detta är nog för att kognitivismen i fråga skall vara ointressant, men det är en fråga som debatteras livligt (gällande McDowell's kombination av kognitivism och internalism, exempelvis), och därför ingenting som jag velat ta för givet.

(2) Persson vill att internalismen skall formuleras så att den är förenlig med både "icke-kognitivistiska" varianter (dit han alltså förhastat räknar min internalism) och dels en påstått kognitivistisk variant. Den internalism som jag förfäktar är en icke-normativ, deskriptiv moralpsykologisk teori som gäller funktionen hos de psykologiska tillstånd som är våra moraliska uppfattningar. Som Persson mycket riktigt påpekar används etiketten "moralisk internalism" också på en annan tes vars ungefärliga innebörd är att en person endast kan vara moraliskt förpliktad att utföra en handling som hon är (potentiellt) motiverad att utföra, eller omvänt att hon garanterat är motiverad att utföra en handling som hon är förpliktad att utföra. (Hobbes och Kant är kanske de historiskt mest kända företrädarna för uppfattningar av detta slag.) Till skillnad från min psykologiska tes så

är denna andra internalism inte alls en teori om moraliska uppfattningars koppling till motivationen *hos den som har uppfattningarna*. Snarare tycks den vara en normativ tes (eller i alla fall en tes som är logiskt oförenlig med vissa normativa teser) om pliktens förutsättningar och som sådan neutral mellan kognitivism och icke-kognitivism i den mån som icke-kognitivismen tillåter normativa teser av detta slag (Kant ansåg exempelvis inte att moraliska uppfattningar var *föreställningar*).

Tyvärr har även filosofer som klart och tydligt sett skillnaden mellan dessa teser diskuterat dem i utan tydlig åtskillnad (se exempelvis David O Brinks *Moral Realism*, Cambridge: Cambridge U P, kap 3). Ingmar Persson tycks gilla denna ordning och menar också att bägge dessa teorier stammar från en och samma "vaga internalistiska kärnintuition" (FT 4/98, s 51–2). Men även om detta är en riktig genealogi (vilket jag betvivlar när det gäller min egen tilltro till denna artikels internalism) så finns det inget som hindrar oss från att gå från vaghet till klarhet och särskilja de olika teser som hämtar kraft ur denna intuition, eller från att behandla dem i skilda sammanhang – vilket jag härmed rekommenderar.

(3) Ytterligare en terminologisk invändning är att jag inte – som annars varit gängse – definierar "externalism" som blotta negationen av internalism utan låter den stå för ett relativt fylligt alternativ till min internalistiska teori (FT 4/98, s 52–4). Bakgrunden till detta är att bedömningen av teorier om psykologiska mekanismer oftast innebär att man undersöker vilken av konkurrerande teorier som bäst förklarar tillgänglig evidens: eftersom negationen av en fyllig internalistisk teori är alltför obestämd för att förklara någon evidens så lämpar den sig inte för den sortens bedömningar. Som jag nu definierar externalismen så implicerar den en form av kognitivism, men detta är – *pace* Persson – inte den bärande skillnaden mellan internalism och externalism: internalismen bör tills annat visats uppfattas som förenlig med sådan kognitivism. Den bärande skillnaden är i stället att externalismen uppfattar moralisk motivation som helt avhängig önskningar som ligger utanför våra moraliska uppfattningar. (Se *Moral Internalism*, s 39.)

(4) Ingmar Persson argumenterar för att komplexa moraliska uppfattningar inte *bara* kan förstås som dispositioner att acceptera vissa kombinationer av enkla moraliska uppfattningar men förkasta andra.

Det håller jag helt med om, och något annat har jag inte heller hävdad. Vad fattas? Den enda restpost som Persson nämner är ett *medvetet accepterande* av denna disposition (FT 4/98, s 57). Det kan jag förstås gladeligen gå med på: det tillstånd som utgör den komplexa uppfattningen och är ansvarigt för dispositionen i fråga skall liksom andra uppfattningar ha uppkommit i en omdömesakt, och detta erbjuder rimligen den sorts medvetna tillgänglighet som Persson efterlyser. Viktigare är emellertid att detta inte specifikt är ett krav på den fullständiga analysen av komplexa *praktiska* uppfattningar: det gäller i precis lika hög grad analysen av komplexa *föreställningar*. Tyvärr var jag för oklar på den punkten i *Moral Internalism*.

(5) En passage i *Moral Internalism* (s 83) skiljer mellan moraliska uppfattningar som tagit allt relevant i beaktande och moraliska uppfattningar som bedömer en handling ur en begränsad synvinkel. Man kan till exempel uppfatta en handling som moraliskt riktig om man ser till att den räddade en persons sociala anseende men samtidigt anse den vara moraliskt felaktig, *allt relevant taget i beaktande*. Ingmar Persson tror att en moralisk uppfattning "allt relevant taget i beaktande" måste innebära en avsikt att handla i enlighet med detta (givet att uppfattningen formas i en beslutssituation): detta skulle innebära problem med att förklara hur vi kan handla mot våra moraliska uppfattningar (FT 4/98, s 60). Som jag påstod i *Moral Internalism* så är de fall där formandet av ett moraliskt omdöme "allt relevant betraktat" innebär formandet av en avsikt kanske de mest *uppenbara*, men "allt relevant" skall förstås som "allt relevant för formandet av den moraliska attityden", vilket alltså inte behöver vara allt som vi uppfattar som relevant för *beslutsfattande*. Vi har, som sagt, andra önsknings- och emotioner än de moraliska.