

Lunds Universitet: Filosofiska institutionen: Mats Johansson

Praktisk Filosofi, FPR503 (502:5b), vt04

Jonatan Forsberg, 0704-405029, j@jonatanforsberg.net

2004-06-07

Hemtentamen till kursen "Empati"

1. Diskutera utifrån litteraturen skillnaden mellan

a. empati och sympati:

Det som många människor i vardagspråket pratar om när de använder sig av ordet empati är, skulle jag tro, i själva verket någon form av sympati. Man vill nå någon slags idé om inlevelseförmåga som leder till medkänsla eller att man blir välvillig/tar ställning. Detta eventuella språkbruksfel kan delvis bero på att det länge råddigt brist på en klar definition, vilket i sin tur har att göra med att olika discipliner inom det som kan kallas empatiforskning ställer olika frågor; för en biolog kan en definition gjord av en socialpsykolog eller en kognitionsforskare vara oanvändbar. Att ordet empati har en kort historia (vilket varken själva fenomenet eller dess studium har) och att vardagspråk ofta skiljer sig från det akademiska språket i så mening att det senare ofta försöker vara mer precist spelar antagligen också en viktig roll.

Om man kikar på ordets teoretiska ursprung, estetiken, kan det nog för många verka klarare med en uppdelning mellan empati och sympati, även om de då kanske skulle vilja omdefiniera orden. Man vill, med rätta, troligen inte gå med på att min starka inlevelse i någon av Johan Thomas Lundbyes nationalromantiska målningar av det Själländska landskapet är samma sak som då jag, av inlevelse i en annan individs lidande, slår sönder en främmande persons bilruta för att ta ut en gnyende hund som håller på att förgås av vätskebrist i den gassande solen. Men det är sannolikt att vardagspråket trots det skulle vilja säga att mitt agerande i bägge situationerna är ett utslag av min "empatiska" förmåga eller något liknande. Här ligger det en viktig iakttagelse; empati förutsätter inte sympati, men en rimlig definition av sympati kanske förutsätter empati.

Ett skrämmande exempel är kanske att en riktigt bra torterare kanske är så duktig i sitt yrke just på grund av sin empatiska förmåga, men få skulle kalla honom eller hans yrke sympatiskt. Skillnaden ligger alltså i att om empatibegreppet säger att person A förstår hur det är för individ S att uppleva ett emotionellt tillstånd E, så kräver ett sympatibegrepp utöver detta någon form av omtanke. Empati verkar med nödvändighet inte leda till medlidande eller sympati (ett ord som konnoterar både till något man kan känna för en rörelse eller en individ).

Förhållandet mellan olika definitioner av empati och sympati kommer att diskuteras mycket mer utförligt utifrån litteraturen i senare frågor.

b. empati och känslosmitta:

I min *Psykologisk ordbok* som Carl Ivar Sandström skrev 1948 (A&W 1979) står det under empati:

(ty. Einfühlung) inkännande, känslö – el. upplevelseprojektion
in i ett föremål, särskilt konstföremål

Om vi låter bli att problematisera definitionen ovan, som för övrigt klart hänvisar till ordets historia (se Edith Stein i fråga 2), blir det rimligt att anta att empati förutsätter att man är medveten om sitt emotionella tillstånd som grundat hos en annan individ eller föremål. Känslosmitta däremot verkar fungera mer omedvetet och därmed inte som något man upplever som grundat i något annat. Man kanske kan göra en liknelse mellan känslosmitta och gäspningar som smittas; om någon i ett rum gäspas börjar snart fler i det rummet att gäspa. Varför andra smittas kan ha olika förklaringsmodeller, exempelvis spridningen av feromoner.

Robert M. Gordon skriver om känslosmitta som en primitiv form av sympatiska eller empatiska responser. Det är välkänt från flera empiriska studier att väldigt små spädbarn börjar gråta när de hör att andra barn gråter och att barn tidigt ”smittas” av andras känslouttryck. Detta kan triggas även genom att en vuxen låtsas få vissa ansiktsuttryck eller låtsas gråta, som barnet sedan härmar. Dessa beteenden ändrar inte bara de fysiska uttrycken, utan ändrar även barnets känslotillstånd. Det är rimligt att anta att i alla fall delar av det här sker omedvetet då ett barn blir mer och mer observant och införstådd med omgivningen ju äldre det blir. Något som kommer att refereras till i senare frågor.

Ifall person A känner det emotionella tillståndet E för att individen S känner E, vet eller förstår att individen S känner E och ser sitt E som grundat hos individen S är det empati. Kan man påvisa en tydlig orsaksrelation mellan person A:s upplevelse och individ S:s upplevelse, men utan person A ser sin upplevelse som grundad i individ S:s upplevelse handlar det om känslosmitta.

2. Nilsson ägnar mycket tid åt klassikerna (Spinoza till Scheler). I vilka avseenden kan man säga att klassikerna lägger grunden till det moderna empatibegreppet?

Peter Nilsson gör inledningsvis klart tre idéer som är karaktäristiska för det moderna empatibegreppet: Den första idén är att empati handlar om emotionellt delande. Agenten måste känna en emotion som på något sätt liknar den subjektet har. Den andra idén går ut på att empati kräver att agenten inte bara känner en emotion av en tillfällighet, utan på grund av subjektets tillstånd eller situation, subjektets emotion eller den idé agenten har om subjektets emotion eller situation. Den sista idén är den om att empatiska emotioner liknar den emotion subjektet har, men ändå är något annat. Emotionerna hos agenten och subjektet är inte identiska. Nilsson delar upp de sex klassikerna han skriver om i tre olika kategorier, dessa hänger samman inte bara för att de är samtida utan för att de bryter ny mark. De fyra första personerna, eller de två första kategorierna, var verksamma innan ett empatibegrepp fanns. De sista var verksamma i samma veva som empatibegreppet myntades. Det står mer om Spinozas, Humes och Schopenhauers teorier i fråga 3.

Nilssons tredje idé går alltså ut på att en empatisk emotion inte är ett korresponderande emotionellt tillstånd. Den förra året bortgångna Páll S. Árdal menar att David Hume inte gör någon skillnad mellan att känna en emotion och att vara i ett emotionellt tillstånd (exempelvis stolthet är samma sak som att vara stolt), varför det inte verkar möjligt, enligt Nilssons idé, att det emotionella delande han skriver om är empati. Möjligen är det så att Hume faktiskt skulle gå med på en distinktion mellan en emotion och ett emotionellt tillstånd ställt i ljuset av det Nilsson skriver.

Spinoza menar att kropp och själ är intimt sammanlänkande. Han menar att det är möjligt för någon att anamma en emotion genom att föreställa sig att någon annan har en liknande emotion, och detta trots att denna någon tidigare kanske varit totalt likgiltig inför den andre

personen. Det är detta som kallas tesen om imitation av affekt, en tes som inte med nödvändighet måste ge upphov till emotionellt delande. Spinoza verkar inte göra någon större skillnad på de emotioner som blivit anammade genom en process av imitation och på de emotioner som blivit anammade på något annat sätt. Häri ligger en skillnad mot Nilssons idé om åtskillnad mellan empatiska emotioner och vanliga emotioner.

Hume och Spinoza är intressanta därför att de bryter en dogm skapad av Thomas Hobbes som går ut på att man behövde ha ett begär (desire) för att få motivation. De menar att vi förvärvar en emotion eller ett mentalt tillstånd från en idé om den andre. Alltså att om agenten känner sorg så är det inte för att dennes önskan på något sätt har frustrerats, utan för att agenten har anammat en idé om subjektets liknande sorg.

Nästa kategori är Adam Smith och Arthur Schopenhauer. De bröt ny mark genom att mena att agenten inte måste känna samma sak som subjektet. Smith gör detta genom att explicit skriva att om agenten föreställer sig vara i subjektets situation, så kommer det att ge upphov till emotioner hos agenten som är annorlunda än de emotioner som agenten skulle ha fått ifall han eller hon faktiskt befann sig i situationen. Smith använder sig av ordet sympati, och menar detta uppstår genom att man föreställer sig vara i den andres situation. I detta föreställande tillstånd kan det ge upphov till emotioner som liknar den andres, men det kan också ge upphov till något helt annat; något jag förväntar mig att den andre ska känna i den situation denna befinner sig i. Agentens sympati är beroende av vilken idé han eller hon har om subjektets situation. Det handlar om en föreställningsakt, och därför kan inte heller agentens önskningar tillfredsställas eller frustreras.

Schopenhauers intressanta ord är medlidande. Med det menar han att man kan känna någon annan persons lidande, fast utan att själv lida. Hur exakt detta går till råder det oklarheter eftersom hans metafysik inte verkar skänka möjligheten till det. Men han är i alla fall intressant då han menar att agentens emotion inte är identiskt med subjektets. En stor skillnad mellan det moderna empatibegreppet och Schopenhauers medlidande, som Nilsson nämner, är att medlidande är något man bara kan känna för någons negativa känslor. Empatiserande däremot, är något man kan göra med både positiva och negativa känslor.

Då den första kategorin, Hume och Spinoza, intresserade sig för hur vi motiveras är den sista kategorin, med Edith Stein och Max Scheler, mer intresserade av hur vi kan ta del av någon

annans känsla, alltså en mer epistemologisk fråga. Vidare formade de empatibegreppet och satte ramarna för det. Stein var den första icke-estet som använde sig av ordet empati och menade att för att känna empati går man igenom tre steg. Det första steget är att agenten på något sätt blir varse subjektets mentala tillstånd. Detta gör agenten enbart genom att titta på externa fakta. Steg två innebär att agenten inte längre tittar på de externa fakta utan istället ser på vad subjektets emotion riktas mot. I steg tre ser agenten återigen på de externa fakta, men mer nyanserat. Nilsson trycker på att det är i det andra steget agenten delar subjektets perspektiv och erfar denne. Det är här agenten får en emotion som liknar subjektets, om än något som inte är helt samma. Stein menar att det man känner inte är något ursprungligt.

Scheler skiljer fyra olika tillstånd från empati. Dessa fyra olika tillstånd är känslomitta, känslagemenskap, emotionell identifikation och medkännande och de är alla genuina känslor. Empati däremot, menar Scheler, är inte en genuin känsla utan handlar enbart om en representation och kräver inte ett faktiskt deltagande i den andres emotionella upplevelse. Scheler menar att agenten och subjektet inte är i korresponderande emotionella tillstånd varför empatiska emotioner skiljer sig från andra. En annan distinktion han gör är den mellan känslotillstånd (exempelvis smärta) och intentionala känslor (exempelvis lidande). Det förra är något som skulle kunna beskrivas som känslor som fungerar som objekt i en varseblivning och det senare skulle vara ett sätt på vilket vi varseblir olika känslotillstånd. En viktig anledning till varför Scheler räknas som en av det moderna empatibegreppets huvudfäder är att han gjorde distinktionen mellan känslomitta och empati, menar Nilsson.

3. Spinoza, Hume och Schopenhauer försökte samtliga förklara hur andras känslor kan påverka oss känslomässigt. Redogör för deras olika uppfattningar:

Baruch Spinoza teori om våra känslor är ungefär något i stil med att det finns tre olika slags primära känslor; glädje, sorg och begär. Alla våra emotioner, menar Spinoza, kan härledas ur de här primära känslorna, men de konstitueras även av våra trosföreställningar, idéer och funderingar. Skillnaden mellan lycka och kärlek, som båda härstammar från glädje, är att vi har olika idéer om dem. Hans teori om hur känslorna påverkar oss är den om tesen av imitation av affekt som går ut på att om vi föreställer oss att en individ, som på ett relevant sätt liknar oss och som inte tidigare omfamnat av någon emotion från vår sida, försätts i en emotion, kommer just vi därigenom att försättas i en liknande emotion. Imitationen kommer

att ske oavsett om imitatören älskar eller inte älskar, hatar eller inte hatar den andra personen. Processen får oss att dela världsuppfattning i en aspekt. Ur det här utläser Peter Nilsson två möjliga problem. Det första är det att imitationen av den andres emotioner kan vara felaktig, att imitationen inte ger upphov till ett tillstånd med en gemensam emotion. Det andra problemet är det om att imitationen kan ha skett korrekt, men att imitatören ändå bara får en partiell avbildning av det den andra individen känner. Med Spinozas resonemang kanske de skulle ha samma primära känsla fast en annan emotion, imitatören känner lycka när den som blir imiterad känner kärlek. Spinoza försöker undvika detta genom att hävda att den emotion vi känner är av samma slag som den vi imiterar, vi kommer att dela objekt för emotionen med den vi imiterar. Den imiterade sorgen eller begäret kommer att vara lika stark som originalet. Nilsson skriver att Spinoza inte är helt konsekvent i sitt resonemang kring hur detta går till och menar att det kanske kan vara kringliggande faktorer eller omständigheter som spelar roll till varför känslan skulle vara, i princip, lika stark hos imitatören som hos den imiterade. Man kanske blir motiverad att hjälpa sig själv ur detta tillstånd, snarare än den man faktiskt imiterar.

David Hume skiljer mellan de han kallar intryck och det han kallar idéer. Intryck är det vi får då vi konfronteras med världen, det refererar inte till något annat. Idéer är en blek kopia av det de faller tillbaka på, intrycken. Hume delar sedan in intryck i primära och sekundära där sensationer är primära, ett exempel är smärta. För att känna smärta behöver jag inte tidigare ha haft något intryck eller idé. Sekundära intryck är de som reflekterar, och då behöver de en korresponderande idé eller tidigare primära intryck att falla tillbaka på. Ett exempel på sekundära intryck är begär (desire) och passioner. En passion som uppstått omedelbart, från till exempel njutning, kallas direkt. Ifall den uppstått genom exempelvis njutning tillsammans med en idé är den indirekt. Men man kan inte dra begreppsliga slutsatser ur passioner eftersom de är intryck, de refererar inte till något annat. Hume kallar vår påverkan av andras känslor för sympati och det är det sätt vilket intryck överförs från en individ till en annan. Detta sker under förutsättning att tre villkor uppfyllts. Det första handlar om att agenten måste kunna sluta sig till att subjektet känner en viss passion. När agenten förstår detta kan en idé om den här passionen uppstå hos denne, och när denna idé övergår i en riktig passion är det första villkoret uppfyllt. Det andra villkoret handlar om att agenten och subjektet är lika. Vari denna likhet finns och till vilken grad det ska vara uppfyllt är oklart. Men Hume vill i alla fall trycka på att alla människor är lika, vilket rimligen borde kunna leda till att man kan sympatisera med alla människor. Det tredje villkoret är det om att agenten och subjektet ska

vara nära varandra i tid och rum. Det är lättare att sympatisera med någon i vår omedelbara närhet. Även här är det naturligtvis oklart vad exakt Hume menar, vidare hur han skulle ha resonerat i en värld med moderna kommunikationer. En intressant fundering som Nilsson har är att det inte klart framgår vilken gradskillnad av likhet som gäller mellan agenten och subjektets passioner, hur lik agentens passion måste vara subjektets för att agenten ska kunna sympatisera. Det verkar också svårt att avgöra om Hume verkligen menade att det för agenten skulle kunna gå ta över både subjektets passion och dennes idé av det passion riktades sig emot.

Arthur Schopenhauer, denne österländskt influerade nitiske ogillare av Hegel, intresserade sig av människans natur för att förstå etikens grundvalar. Han hävdade att de grundläggande drivkrafterna hos människan är egoismen, illviljan och medlidandet. Medlidande innebär att känna en annan persons lidande utan att själv lida. Det är också så att medlidande är det enda som kan motivera någon att utföra en moralisk handling. En handling som utförs för att få bort mitt eget lidande är inte en moralisk handling, utan snarare en egoistisk. Man måste agera för att få bort den andres lidande. Inte heller kan en handling vara motiverad av min önskan att se någon annan lycklig, då tillfredsställer jag nämligen bara min egen önskan. För att det ska finnas några moraliska handlingar alls gör därför Schopenhauer en distinktion mellan direkt och indirekt motivation. Indirekt motivation innebär att ditt lidande motiverar mig först sedan det orsakat smärta i mig. Direkt motivation, vilket behövs för att handlingen ska vara moralisk, innebär att jag endast motiveras av ditt lidande. Jag ska känna ditt lidande som det vore mitt eget, fast utan att jag själv lider. Det jag delar med dig är dina belegg för lidandet, det är inte kvalitativt likt utan numerärt identiskt. Att detta är möjligt har att göra med Schopenhauers metafysiska teori om att alla människor är sammanlänkade, att det inte finns några åtskillnader emellan oss. Nilsson skriver att detta har som konsekvens att Schopenhauer lite grand säger mot sig själv, eftersom om vi är samma enhet borde ditt lidande faktiskt bli mitt. Därför, menar Nilsson, kan man tolka honom som att det faktiskt finns en distinktion mellan dig och mig om vi nu kan känna någon annans lidande utan att själva lida. Fast om man hur som helst tar Schopenhauers metafysik på allvar får det den intressanta konsekvensen att egoismen vilar på en illusion. Ett annat problem med hans teori är att det är svårt att tro att den kan förklara medlidande med hypotetiska individer. Fast å andra sidan kanske det egentligen inte låg i Schopenhauers intresse att förklara sådant.

4. Nilsson lägger stor vikt vid Nancy Snows definition av empati.

a. Vilka fördelar menar Nilsson att Snows definition har framför den som presenterats av Sober och Wilson?

Nancy Snows definition av empati är som följer: [(a) Subjektet känner E. (b) Agenten känner E på grund av att subjektet känner E. (c) Agenten vet eller förstår att subjektet känner E.] Denna definition ställs mot en definition gjord av Elliot Sober och David Sloan Wilson: [(a) Subjektet känner E. (b) Agenten tror att subjektet känner E och detta påverkar (c) agenten att känna E för subjektet.] Skillnaderna Snow är ute efter, vilka också Peter Nilsson har som anledning till att föredra hennes definition framför Sober och Wilsons, är dessa tre: För det första ifrågasätter Snow (och Nilsson) att trosföreställningen är det relevanta och menar att det snarare är vetskapen eller förståelsen. Då utesluter man nämligen den möjligheten då agenten utan goda skäl ändå tror att subjektet känner E. Snows definition hindrar problemet i Sober och Wilsons definition att man kan tro annat än det faktiska, det är skillnad på att tro och veta. Den andra fördelen gäller en annan skillnad och består av två skäl: Genom att agenten noterar att hon själv känner E kan hon börja tro att även subjektet känner E och därefter börja empatisera med denna, samt, det verkar inte finnas några goda skäl till varför man skulle definiera empati som att det måste vara just trosföreställningen som ger upphov till att agenten känner E. Det torde finnas andra möjligheter till att agenten känner E. Ingen av dessa möjligheter finns i Sober och Wilsons definition. Den tredje fördelen Nilsson ser i Snows definition är den att sympati inte följer per automatik när det gäller vissa slags känslor, något det verkar göra hos Sober och Wilson. Detta beror på att de menar lite olika saker med sympati och empati; Sober och Wilson menar att man kan känna både sympati och empati för någon; Snow och Nilsson menar att man känner sympati för någon, men empati med någon. En rimlig definition enligt Snow och Nilsson kan inte omöjliggöra empati utan sympati då subjektet exempelvis känner sorg. Snow menar också, vilket Nilsson kritiserar i nästa fråga, att empati är ett tillstånd där känslan hos agenten är kvalitativt likt det hos subjektet.

b. Vilka problem ser Nilsson med Snows definition?

Nilsson godtar inte Snows definition rakt av, trots att hans mening är att den är bättre än Sober och Wilsons. Den kritik han riktar mot Snow har till att börja med att göra med att den verkar vara fenomenologiskt kontraintuitiv. Om empati kräver kvalitativ likhet borde

empatiserande med någon som är riktigt arg ge upphov till samma ilska hos den som empatiserar. Ibland verkar empatin ge upphov till en lite annorlunda känsla och ofta är det i alla fall så att det är olika styrkeförhållanden på agentens E och subjektets E. Kritik om att Snows definition ej upprätthåller distinktionen mellan empati och sympati går ut på att om E är lika med "tycka synd om sig själv" leder empatiserande, att dela den andres känsla, till sympati. Nilsson menar att empati visserligen kan leda till sympati, men det är inget som är nödvändigt, vilket inte Snows definition möjliggör. Ifall empati betyder att dela någon annans känsla tenderar menar Nilsson att distinktionen mellan empati och känslomitta inte längre upprätthålls. Hos Snow är skillnaden mellan empati och känslomitta att vid den förra vet du att den andre känner det du gör. Nilsson menar att vid känslomitta är det dina riktiga känslor, du är arg eller ledsen på riktigt, men vid empati handlar det inte om samma emotionella tillstånd. Vi delar inte den andres känsla. Ytterligare kritik Nilsson ger av Snows definition är att han tycker hon ibland har för vid definition och ibland för snäv. Man kan uppfylla alla Snows villkor utan att vi skulle vilja kalla det empati, emotionen har nämligen fel ursprung (det krystade exemplet om läkaren som ger medicin som orsakar stress). Där anser han att Snows definition är för vid. Nilsson tycker det är rimligt känna empati för att man vet vilka omständigheter eller i vilken situation någon befinner sig i, inte bara för att man vet vilken emotion den andre befinner sig i. Där anser han Snows definition vara för snäv. Nilsson riktar även kritik mot Snows uppsats, snarare än hennes definition, som går ut på att ifrågasätta om empati med nödvändighet är interpersonell. Ytterligare kritik mot Snow kan vara att hennes definition tar till "vet eller förstår" enbart för att negra Sober och Wilsons definition. Ett problem med vetande kan vara att det verkar mer genomtänkt än vad empati är, som verkar vara något som är mer spontant.

c. Hur försöker Nilsson undkomma problemen med Snows definition?

Nilsson är intresserad av ett empatibegrepp i den mening där empati är empati i förhållande till de emotionella tillstånd subjektet befinner sig i. Han presenterar sin definition av empati som; [en agent empatiserar med ett subjekt med respekt till ett emotionellt tillstånd av typen E, om och endast om: (a) Subjektet befinner sig i ett emotionellt tillstånd, e, av typen E. (b) Agenten känner endast en emotion, e*, av typen E. (c) Agenten känner endast e* därför att (c1) agenten har blivit utsatt för ett uttryck av e, eller (c2) agenten har hyst en idé om subjektets situation, eller (c3) agenten har hyst en idé av e. (d) Det är inte fallet att agenten

endast känner e* på grund av en faktisk eller föregripande önsketillfredsställelse eller önskefrustration. (e) Agenten vet att subjektet är i det emotionella tillståndet e.] Villkor (a) och (b) betyder två saker. Det handlar dels om att agenten och subjektet känner emotioner av liknande slag, dels om att agentens emotion riktas åt samma objekt som subjektets. Det är alltså exempelvis möjligt för agenten att känna empati med subjektet med respekt till dennes sorg över en viss händelse, inte bara med respekt till subjektets sorg. Villkor (c) handlar om Nilssons kritik av att Snows definition ibland är för vid och ibland för snäv: (c1) betyder att det inte krävs att agenten har en medveten idé eller trosföreställning om subjektets situation för att kunna empatisera. Agenten skulle kunna imitera subjektet för att ge upphov till en empatisk emotion, något som Snows definition inte lämnar utrymme för. (c2) ringar in alla de fall där agenten skaffat sig en idé om subjektets situation. Det som behövs är att agenten är medveten om situationen och att dennes emotion är orsakad av den, men agenten måste inte nödvändigtvis veta eller tro någon om subjektet. (c3) är tillagt för de gånger då agenten har en idé om subjektets emotionella tillstånd. Inte för att agenten blivit varse om subjektets situation eller blivit utsatt för ett uttryck liknande det subjektet haft, utan på grund av den idé agenten har av subjektets emotionella tillstånd. Ett exempel är att agenten har läst om det. Det fjärde villkoret, (d), handlar om skillnaden mellan vanliga emotioner och empatiska emotioner. Empatiska emotioner upplevs inte från ens eget perspektiv och kan därför inte handla om tillfredsställda eller frustrerade önsknings. Nilssons definition är bara rimlig om man enbart intresserar sig för emotioner (vilka i sin tur är ganska komplexa hos honom). Ligger ens huvudintresse i sensationer är hans definition antagligen tämligen svår att tillämpa.

5. Young är kritiskt inställd till "perspektivtagande" inom moralen.

a. Varför?

Iris Marion Young invänder mot perspektivtagande på tre sätt. Hon menar för det första att det leder till att man döljer skillnader, för det andra att det är omöjligt med omvända positioner och för det tredje att perspektivtagande är politiskt suspekt. Hennes första kritik går ut på att om vi väldigt gärna vill se likheter hos varandra omvandlar vi skillnaderna till likheter. Young hänvisar till Lucy Irigaray som menar att ett symmetriskt tankesätt förnekar skillnaderna som finns mellan könen och istället bekräftar den könsmaktsstruktur som ser det kvinnliga könet som kompletterande till det manliga. Irigaray menar att män då ser kvinnor som speglar i vilken de kan finna sig själv. Young använder sig av det här resonemanget hos

alla tänkbara skillnader; kön, klass, etnicitet och så vidare. Man kan inte se symmetri i förhållande till en spegel. Även om det ”inom en struktur” skulle finnas likheter, finns det alltid viktiga skillnader individer emellan; livsstil, erfarenheter och så vidare. Det enda sättet vi behandlar någon annan med moralisk respekt är genom att lyssna på vad denna någon har att berätta. Young menar att vi ska utgå ifrån att vi aldrig kan förstå andra helt och hållet, det bästa vi kan syssla med är dialog och samtal. Att ensamt syssla med perspektivtagande hemma i kammaren ser hon som förkastligt, det handlar snarare om att projicera ett kompletterande perspektiv på den andre än att försöka förstå densamme.

Youngs argument om att det är omöjligt med omvända positioner går ut på att jag är jag inför mig själv på samma sätt som du är jag inför dig själv. Jag är den andre, du, inför dig, på samma sätt som du är den andre, du, inför mig. Detta är något vi inte kan rubba på, jag kan aldrig se ur någon annans perspektiv än mitt eget varför vi inte kan omvända varandras positioner. Våra positioner är asymmetriska och irreversibla och förståelsen av det förebygger missförståndet som heter perspektivtagande.

Det sista argumentet från Young kommer från att perspektivtagande skulle innebära något politiskt suspekt. Hon skriver att eftersom perspektivtagande innebär att någon försöker sätta in sig själv i någons annans perspektiv, med sitt eget perspektiv, missas de hierarkiska skillnader som finns dem emellan. Då människor befinner sig i olika sociala strukturer och är offer för olika stereotyper är det alltid så att risken för falska projektioner är väldigt överhängande. Genom dessa falska projektioner undermineras den undertryckta gruppens självrespekt och den privilegierade gruppens status legitimeras. Young menar att vi måste inse att vi inte kan se ur någon annans perspektiv och att det därför finns saker som vi aldrig kan förstå. Vi måste lyssna på den andra utan att lägga in oss själva.

b. Är hennes kritik rimlig?

Visst är Youngs kritik rimlig på det sättet att den tydliggör saker, den dekonstruerar. Vill man däremot läsa något praktiskt tillämpbart enbart ur hennes teorier, något konstruktivt, tror jag däremot att man får problem. När jag skriver detta menar jag naturligtvis inte att hennes motsats är lösningen, utan snarare möjligen en kompromiss av hennes kunskap och den kunskap hos dem hon kritiserar. Vidare är hennes kritik naturligtvis både rimlig och befogad

på det sättet att hon kritiserar de strukturer som vuxit fram ur ett patriarkalt samhälle och därmed också ett visst förhållningssätt till varandra. Sett ur det här perspektivet är hennes feministiska bidrag är inte bara intressant utan även nödvändigt.

Det första och kanske mest direkta problemet med Youngs kritik är att den inte verkar ge möjlighet att ta hänsyn till intressen som vi vanligen gärna ser som moraliskt viktiga intressen. Människor med vissa handikapp, små barn och inga djur kan falla inom den kategorin vi kan föra samtal med. Om man då inte får syssla med perspektivtagande eller anta vissa basala likheter blir det svårt att veta hur man ska behandla dem. Det gäller egentligen inte bara de grupperna: Jag kan själv ibland vara väldigt stressad eller upprörd av någon annan anledning, då är det ibland bättre att enbart anta något än att fråga. En kram kan göra mer än tusen ord.

Young verkar mena att moralisk respekt innebär att man tar hänsyn till den andre personens speciella säregenskaper, och detta till en sån grad att de är de som styr relationen. Detta i sig själv är naturligtvis ett problem, men det blir ännu större när man betraktar fenomenet samhälle. Ett samhälle bygger på generaliseringar, antaganden och snabb interaktion, något som Young inte skulle ha gillat alls. En teori som inte är tillämbbar annat än möjligen i mycket små samhällen är kanske inte speciellt praktiskt intressant.

Ett problem med att prata om att någon kan hävda att den andre ändå inte kan förstå är att det ger upphov till en slags definitionsrätt. Detta kan i sin tur faktiskt vara ett väldigt farligt maktmedel för att få sin vilja igenom eller för att trycka ner någon annan. Det ifrågasätter dessutom starka rättviseintuitioner kvinnor och män världen över faktiskt har. Vidare är det problematiskt med tolkningsföreträde då man inte alltid kan säga vem det är som är mest privilegierad, och eftersom Young själv trycker på starka individuella skillnader borde hon vara införstådd med att det ofta kan vara väldigt svårt att avgöra.

Young beskriver alltid väldigt komplicerade perspektivtaganden, något som inte riktigt klargör hur hon ställer sig till den uppsjö av mycket mindre komplicerade perspektivtaganden. Vidare är det så att samtal faktiskt bygger på att man hänvisar till något gemensamt, språket. Man kan inte samtala utan att göra tolkningar. En konsekvens av hennes resonemang är möjligen att vi helt slutar upp med att vara språkbrukande sociala varelser, en annan är att hon inser att vi faktiskt kan och måste ha gemensamma referenser.

Hur hjälper man någon, med Youngs teori, som inte vill bli hjälpt eller förstådd? Om talaren alltid har tolkningsföreträde, kan lyssnaren då tillföra något alls? Kan vi förstå något?

Ytterligare kritik som är intressant för Youngs ståndpunkt hittas i nästa fråga (6), och då främst under kritiken mot omsorgsetiken.

6. Antag att man vill försöka länka samman empati och moral.

a. Hur kan man då gå till väga?

b. Vilka problem kan man stöta på?

Man kan, precis som Alvin I. Goldman gör i sista delen av sin text, börja med att dela upp moralen i en deskriptiv och en preskriptiv del. Den deskriptiva delen delar han sedan upp i ytterligare två delar, en som beskriver och förklarar de olika moralkodex som finns i olika kulturer och ytterligare en som handlar om hur mycket människor faktiskt följer dessa kodex. Den senare fokuserar på motivationella faktorer, vad som möjliggör eller hindrar människor från att agera utifrån sin moral. I alla dessa delar av etik är empiriska studier av empati relevanta, menar Goldman.

Historiskt sett, skriver Goldman, spelade Arthur Schopenhauer en viktig roll för införandet av empati i deskriptiv etik. Han hävdade det grundläggande fenomenet för etik är medlidande, vilket han karaktäriserade som ett deltagande i någon annans lidande. Det är inte så att Schopenhauer menade att medlidande är den övervägande motivationella faktorn i vårt liv, men han menade att medlidande är roten till moraliska principer och anledningen till att vi följer dessa principer. Inte heller är det så att medlidande är den enda eller den direkta anledningen till, formuleringen av och eftergivenheten till, olika moralkodex, tvärtom verkar medlidande agera indirekt.

Den första starka kritiken mot det här sättet att resonera kom 1982 med Carol Gilligans *In a Different Voice* (HUP). I den skriver hon att kvinnor har en moralisk orientering som fokuserar på associering och omsorg, snarare än abstrakta moralteorier om rättigheter och rättvisa. Hennes tes har fått stor kritik därför att den har svårt att överensstämma med empiriska studier som inte hittar någon statistik som talar för att kvinnor skulle vara mer

benägna att syssla med hennes slags etik. Men om man bortser från hennes resonemang kring könsskillnader ligger det i hennes teser en stark och viktig kritik mot det sätt att resonera kring moraliska rättigheter som varit rådande. Det är också det senare omsorgsetiker tagit tillvara på, nämligen att det är mer relevant att prata om moral som något som härleds ur empati, och där empatiserande är framträdande. Moralen, på det sätt den rådande diskussionen förts, är kanske felaktigt grundad. Möjligen, menar vissa, har det här mycket mer långtgående konsekvenser än just kring moralen som isolerat fenomen.

Om vi istället vänder blicken mot den preskriptiva etiken, är det, i ljuset av kritiken från omsorgsetikerna, ett rimligt krav att en tillämpbar moralteori bör vara förenlig med hur vi faktiskt fungerar. De som inte vill anpassa gamla moralteorier till denna nya kunskap gör klart, som jag nämnde i stycket ovan, att skapa en helt ny, omsorgsbaserad, etik. Sett ur den här synvinkeln blir våra moraliska skyldigheter grundade i vår empatiska förmåga. Ett problem med det här är naturligtvis att empati tenderar att vara partisk eller fördomsfull, vi bryr oss mer om våra nära och kära. För att ge ett klassiskt exempel så kanske de som är mest behövande inte skulle få något alls; om någon, av en eller annan anledning, är svår att empatisera med skulle den individen inte ligga i moralens strålkastarljus. Annan kritik jag brukar rikta mot ”ren omsorgsetik”, vilken är relevant mot all partikularism, är att den helt enkelt har svårt att ta itu med det som är kanske det viktigaste i frågan om etik, nämligen hur man löser moraliska konflikter. För att vi ska klara av det verkar det behövas vissa grundläggande axiom om rätt och fel och till det ett mer förnuftsmässigt resonemang om moralisk ärlighet eller konsistens. För att kunna lösa konflikter behöver moralen vara mer som en karta, och en karta kan inte säga att det är rätt att både gå åt norr och åt söder. Vad som är helt klart är att till det här kommer en frågeställning om vad som i så fall faktiskt är moral och vår individuella frihet i förhållande till den. Det hela är kanske inte så blodigt som det verkade vid första ögonblicket.

Det finns andra lösningar på problemet än att anamma en ren omsorgsetisk moralföreställning, som faller i en partikularistisk grop. Ett exempel på en empati, eller snarare sympati-grundad teori som är ganska universalistisk är den som John Rawls skisserade på i *A Theory of Justice* (HUP 1971). Det handlar om att vi måste använda vår förmåga till förnuft (tillsammans med vår delikata förmåga till simulering, se nästa fråga) för att ta reda på vilka moraliska teorier eller rättvisepprinciper vi skulle önska vara rådande,

oberoende av oss själva. Bakom en slöja av okunnighet, menar Rawls, kan vi ta reda på vad som är rätt och fel. Detta gör vi med hjälp av vår förmåga till empatiserande.

Man hade även kunnat tänka sig att olika moralteorier som kan tänkas vara abstrakta och rationella, exempelvis den från Immanuel Kant, skulle kunna anpassas till den kritik som kommit från omsorgsetikerna. (Möjligen är det så att vi faktiskt läser Kant fel och att han inte är speciellt orubblig i förhållande till hur vi faktiskt fungerar som empatiserande varelser, i alla fall låter jag den frågeställningen vara öppen, oavsett om han vänder sig i sin grav eller ej. Puh!) Andra moralteorier än den från Kant som brukar anklagas för att vara orimliga utifrån den mänskliga naturen är den rena konsekventialismen, eller dess kanske främsta förgrening; utilitarismen. En lösning på det här, som för vissa känns ad hoc, är den som går ut på att man istället för att maximera lycka eller preferenstillfredsställelse ska maximera de attityder som leder till det samma, alltså någon form av rättighetsutilitarism eller omsorgsutilitarism. Det som möjligen är ytterligare en annan väg är den som Peter Singer verkar ta. Han har skrivit på senare tid att han förstår att många tycker att han kräver för mycket när han säger att vi ska ge en väldigt stor del av vår förhållandevis stora rikedom till tredje världen (vilket han skrev om i *Famine, Affluence, and Morality*, Philosophy & Public Affairs 1972), och skriver därför att ett kanske rimligare bistånd är det utifrån en viss slags princip om exempelvis trettio procent av sin månadsinkomst. I Singers fall kanske det handlar om någon slags anpassning av sin teori i ljuset av den kritik som omsorgsetikerna kom med. En annan sådan anpassning kan möjligen vara den som Robert Nozick skrev om i *The Examined Life* (Touchstone Books 1990). Hur som helst, vilket Goldman avslutar sin text med, kan moralteoretiserande dra nytta av kognitionsforskningen samtidigt som kognitionsforskningen kan dra nytta av de observationer och antydningar filosofer har gjort på området.

7. Goldman, Gordon och Walton utgår samtliga från simuleringsteorin.

a. Vad är simuleringsteorin?

I boken *Hur Homo blev sapiens – om tänkandets evolution* skriver Peter Gärdenfors, i kapitel 3.1, om det han kallar skinnerska varelser. De kan skaffa sig kunskap genom trial-and-error, det vill säga att de provar och ser vad som händer. Han skriver vidare att detta är tidskrävande eftersom varelsen måste prova många beteenden innan det hittar ett som lyckas, vidare är det så att trial-and-error ofta är direkt farliga: ”En hjort som försöker hoppa över en ravin för att

ta en genväg löper en stor risk att inte få göra ett försök till.” Därför klarar sig de som han kallar popperska varelser mycket bättre, de kan nämligen representera världen i en inre värld. I denna inre värld kan varelsen föreställa sig konsekvenserna av olika handlingar. Varelsen simulerar konsekvenserna av olika beteenden och reagerar därefter.

Simuleringsteorin är något som modern kognitionsforskning använder sig av för att förklara hur vi tolkar och agerar i världen. Den sätts ibland som motpol till en behavioristisk betingningsmodell. Den är rimlig av flera anledningar, en är att den ofta stämmer bra överens med de empiriska data vi har av hur vi fungerar; det är exempelvis så att det verkar ta lika lång tid att leta igenom en fysisk karta som en likadan karta vi memorerat. Vidare är det så att syncentrum i hjärnan faktiskt aktiveras då jag visualiserar något och min hjärna skickar liknande impulser när jag iakttar någon utföra något som ifall jag själv skulle ha gjort samma sak. Det sistnämnda dock utan att jag med nödvändighet faktiskt själv utför det. Simuleringarna kan alltså kopplas av, de kan kopplas, som vissa i nästa fråga (7 b) uttrycker det, off line. På många sätt spelar simuleringar en viktig roll i, i alla fall, människors mentala liv. Men det är rimligt att anta att många andra djur i mer eller mindre utsträckning har förmåga till simulering.

I vårt fall är simuleringsteorin intressant som kontrahent till teoriteorin. Enligt den senare förstår vi våra empatiska tillstånd på samma sätt som vi förstår allt annat i världen, genom att forma teorier. Problemet med den teorin är att den har problem att förklara hur vi faktiskt upplever tillstånden, något inte simuleringsteorin har. Simuleringsteorin menar att vi inte teoretiserar, vi simulerar de empatiska tillstånden och vi drar nytta av de processer vi försöker att simulera. Vår egen hjärna är den perfekta plattformen för att simulera empati, vi har nämligen all den kunskap vi behöver för att själva klara oss i vardagen. Det enda vi behöver göra är att inhämta kunskap om den vi empatiserar med och se vad som skiljer.

b. Diskutera simuleringsteorins för- och nackdelar utifrån litteraturen?

Kendall Walton intresserar sig för hur vi föreställer oss och bygger upp fantasivärldar. Han har skrivit en bok som heter *Mimesis as Make-Believe* (HUP 1990) i vilken han verkar vilja förklara hur vi gör detta, och som bokens titel förklarar är det främst konst- och litteraturvärlden han har intresserat sig för. Hans frågeställning, på det sätt som det är

intressant för vårt ämne, handlar om vad som skiljer, om det är något som skiljer, mellan att empatisera med någon i en påhittad värld och någon i den ”riktiga” världen. Walton har tidigare kallat sin teori inbillningsteorin (make-believe theory), men skriver att det nästan är läskigt så mycket den relativt färska teorin om mental simulering (simuleringsteorin) sammanfaller med hans.

En kort bit in i texten berättar han en kort historia om hur du och en kompanjon grottvandrar, han berättar om hur gångarna blir trängre och trängre efter hand som ni fortsätter längre och längre in i den mörka grottan. När man läser texten föreställer man sig hela scenariot, man simulerar det. Detta kan ge konsekvenser för hur man agerar och känner sig i verkliga livet, inte bara nu utan även i framtiden. Det är inte bara så att ens händer kanske blir svettiga och hjärtat dunkar fortare, nästa gång man befinner sig i en trång hiss kanske hjärtat dunkar till en extra gång. Trots att man vet att det är en påhittad värld och man fantiserar är det fortfarande så att man simulerar det med hjälp av sig själv. Walton skriver att det inte kan råda tvivel om att det faktiskt rör sig om riktiga attityder, i det här fallet kanske riktig spänning och/eller klaustrofobi.

Beroende på olika faktorer agerar man annorlunda när man simulerar, sådana faktorer kan naturligtvis röra en själv som person, vilken situation man befinner sig i och hur man får fantasivärlden presenterad för sig. Klart är i alla fall att man agerar annorlunda än man hade gjort om det hade skett i verkliga livet. Walton skriver att ens psykologiska mekanismer är ”off line” och att de oftast inte orsakar samma beteenden som de brukar. När man sitter i soffan och läser om grottvandrarerna är man fullt medveten om det, trots att man simulerar det hela. En man som heter Richard Moran hade här argumenterat mot Waltons teori och sagt att den saknar förklarlighet i den riktiga världen. Moran menar att uttryck för genuina attityder är något annat än de uttryck vi får av fantasivärldar. Walton svarar med att ge ett exempel om att man läser en text om kattungar som blir torterade. Att man bara läser texten spelar antagligen ingen roll på ens moraliska karaktär i det här fallet, men om man skulle komma på sig själv att tycka historien var lustig hade man anledning till oro.

Walton håller med Moran om att våra emotionella upplevelser ofta är något som inte pågår här och nu, något man kan rycka på axlarna åt. Men han tycker inte att man måste uppfatta fiktion som något distinkt annorlunda än ”verkligheten”, vilket Moran verkar implicera. Som stöd för sin tes hänvisar Walton till att både fiktion och ”verkligheten” är simuleringar och

han uttrycker även att detta är en stark fördel med simuleringsteorin, att den tillåter en att se fiktion som en kontinuitet med resten av livet. En annan viktig fördel är att den ger en fenomenologiskt rimligare förklaring av hur vi agerar utifrån något vi föreställer oss; om man tänker tillbaka på då man på håret klarade sig undan att bli överkörd av en lastbil blir man rädd, man måste inte föreställa sig först att man nästan blev överkörd och sedan föreställa sig rädsla.

Men, vilket Walton skriver i slutet, simuleringsteorin räcker inte. Den inbillningsvärld som han pratar om har även en nödvändig dimension av att man faktiskt föreställer sig, att man fantiserar. Den inbillade delaktighet vi har i en fantasivärld består dels i en mental simulering, man simulerar exempelvis att man blir jagad av ett slemmigt monster och får känslor därur. Men vi reagerar inte med orden ”jäklar, vilket hemskt påhittat slemmigt monster”, utan snarare bara ”jäklar, vilket hemskt slemmigt monster”. Vi måste alltså, utöver simuleringen, föreställa oss fantasivärlden.

Robert M. Gordon skriver med utgångspunkt från 1700-talsfilosofen David Humes beskrivning av hur vi sympatiserar. Hume förklarar att våra medvetanden är speglingar av varandras och att de reflekterar varandras känslor, lustar och åsikter. Gordon vill förklara, i ljuset av färsk empirisk forskning att denna spegling är mycket mer genomträngande än vad Hume föreställde sig. Han förklarar att det Hume skriver om mer verkar handla om någon slags protoempati i form av icke-kognitiv känslomitta, motormimik eller reaktiva skrik som utslag av empatisk stress. Gordon exemplifierar; en mamma ler på grund av att hon är glad över att bli befordrad, vid denna åsyn börjar hennes spädbarn att le och känna glädje. Denna glädje är inte glädje *för* något speciellt, i alla fall inte på grund av att barnets moder blivit befordrad. Gordon vill gå från smitta till förutsägelse.

I fråga om inlevelse gör Gordon en uppdelning mellan varm och kall metodologi, där den senare handlar om att man först och främst engagerar intellektet och inte drar nytta av sin egen kapacitet till motivation, känsla och praktiskt resonerande. Den kalla metodologin kan falla in under det som man brukar kalla teoriteorin. Men han berättar att det är ganska tydligt att vi ofta använder oss av den varma metodologin, det vill säga simuleringsteorin, då vi till exempel skulle vilja förklara vårt eget beteende eller någon annans beteende. Vidare skriver han om hur många skådespelare pratar om att förvandla sig själv till den personen de spelar.

Man använder sig av den kunskapen man inhämtat om den personen man spelar, men genom att använda sig av sig själv. Gordon skriver att man ändrar plats på sin egocentriska karta så att någon annan karaktär är mer centrerad, detta kan betyda att man ändrar på den personlighet man egentligen har. Andra värderingar och normer kan råda, men verkligheten kan vara densamma. Man låtsas eller föreställer sig bara att man är i sin roll. Vid inlevelse simulerar man någon annan, oavsett om man är skådespelare eller ej.

Som argument för simuleringsteorin nämner han först rimligheten. Gordon skriver att det är orimligt att anta att man använder sig av olika kalkyler och uträkningar mellan olika kategorier såsom trosuppfattningar (beliefs), känslor, begär (desires) och ageranden för att räkna ut hur en person beter sig i en viss situation. Vid en inlevelsesituation, såsom på scen, skulle skådespelaren bli helt låst. Det andra argumentet handlar om att det är mycket mer ekonomiskt. Gordon ber en tänka på en repertoarskådespelare som måste byta mellan olika roller från en dag till en annan och förklarar att det är mer rimligt att skådespelaren har en mall över vad som måste förändras för att han eller hon mentalt ska "bli" rollen. Det snarare än att skådespelaren har en hel uppsättning med korta och långsiktiga mentala tillstånd, karaktärens alla trosuppfattningar, begär, och så vidare för alla de olika rollerna han eller hon spelar. Gordon gör en analogi med en vanlig metod att komprimera videosignaler som går ut på att istället för att lagra alla de olika bilderna som visas på skärmen lagras man bara informationen om vad som förändras mellan de olika bilderna. Den största skillnaden mellan analogin och verkligheten är att när det gäller videosignalerna så måste den första bilden lagras i sin helhet, vilket inte skådespelaren behöver göra eftersom att han eller hon simulerar utifrån sig själv.

Ytterligare ett argument som Gordon framför är att simuleringsteorin, inom ramen för vardagspsykologin, mycket enklare kan förklara småbarns oförmåga att kunna redogöra för att samma fakta inte finns tillgängliga för andra som för dem själva. Istället för att komma med, som Gordon uttrycker det, ad hoc-hypoteser om att barnen inte har några trosuppfattningar kan simuleringsteorin förklara det genom att det tar ett tag innan man utvecklat förmågan att simulera andra och skilja det från sig själv. Han menar vidare att teoriteorin-förespråkare naturligtvis kan komma med liknande ad hoc-konstruktioner för att nå samma konsekvenser som simuleringsteorin. Men hans poäng är att simuleringsteorin inte kan leda till något annat, och att det talar till dess fördel.

Men simuleringsteorin leder även till att, med Richard Hares ord, interpersonella konflikter kan bli intrapersonella. Gordon skriver att de beslut vi fattar när vi simulerar någon annan måste vara avskurna från vårt eget beslutsfattande. Gränserna mellan ens egna mentala liv och representationerna av någon annans blir tunnare och då borde benägenheten till förvirring vara större. Att upprätthålla gränserna är därför viktigt menar Gordon. För även om många moralteorier, allt från kristendomens gyllene regel till möjligen kravet på universaliserbarhet, kräver att man kan sätta sig i någon annans situation måste vi ändå hålla oss borta från att helt identifiera oss med den andre. Om någon ber om ett råd måste vi helt enkelt ha viss distans för att kunna komma med ett bra råd. "Blir" vi den andra personen torde vi vara lika oförmögna som den rådfrågandes, och därmed förlorat förmågan att kunna hjälpa vederbörande med ett råd. Att vi befinner oss i en fördelaktig situation gör att vi har know-how eller oberoende omdöme.

Gordon diskuterar den huvudkritik som Adam Smith framförde mot Humes sympatibegrepp, om hur man bedömer någons uppförande. Men han menar att han klarar sig undan den genom att han gjort distinktionen mellan att föreställa sig vara i någons situation och att föreställa sig vara denna någon i dess situation. Genom denna senare distinktionen kan vi bedöma någons uppförande genom att först föreställa oss vara i den situationen och sedan att vara denna någon i dess situation, om vi agerar lika i båda situationerna godkänner vi uppförandet, om vi inte agerar lika ogillar vi uppförandet. Fast Gordon ser minst ett problem med det här resonemanget, och det är att vi ibland ogillar vårt eget uppförande, ibland till och med vårt nuvarande uppförande. För att lösa det tar han till Smith, som menar att när vi bemödar oss med att rannsaka vårt eget uppförande delar man upp sig själv som om man blev två personer, både bedömaren och den som ska bedömas. Här blir det krångligt, men det är också här Gordons kanske viktigaste moraliska implikation ligger; när man simulerar en annan person än jag själv, kan jag vara i den rollen i sympati med mig själv, men hur kan detta ske när jag, mig själv, med mina egna känslor och min egen motivation, är åskådaren? Det kan bara ske om själva processen att bli en åskådare, och genom att komma att se mig själv som en annan, förändrar mig. Gordon skriver att den han tror att den processen faktiskt förändrar oss, och att det även var Smiths åsikt. Han avslutar: "But the topic is a complex one: observing myself trying to argue the case in a paragraph or two, I see a lot of hand waving. So I will not try." Vilket skulle bevisas ☺.

Alvin I. Goldman argumenterar mot det som brukar kallas teoriteorin, argumenterar för det som brukar kallas simuleringsteorin (han kallar den simuleringsteorin) och kommer fram till att det antagligen är en mix av de båda som stämmer. När han gjort detta konstaterar han att man då kan argumentera för en svag hypotes om att simuleringar bara gäller ibland, till förmån för teoriteorin, eller för en stark hypotes där man ser simuleringar som grundläggande och något som man sysslar med tidigt i livet som bas för hur man förstår andras medvetanden. Goldman skriver att han attraheras av den starkare hypotesen av simuleringsteorin.

Tidigt i texten tar Goldman upp ett exempel som även Gordon nämnde, nämligen det från Daniel Kahnemans och Amos Tverskys experiment från 1982. Det går ut på att ett antal testpersoner föreställer sig att Mr Crane och Mr Tees bägge skulle ha lyft från flygplatsen med olika flyg som avgick samtidigt. De reser tillsammans i samma limousin ut till flygplatsen, men de fastnade i trafiken och ankom därför 30 minuter efter att deras flyg skulle ha lyft. Mr Crane fick reda på att hans flyg hade lyft i tid, medan Mr Tees fick berättat att hans flyg hade varit försenat och hade därför avgått för bara några minuter sedan. Man ska därefter berätta vem av de bägge som är mest upprörd. Det visar sig då att nittiosex procent av de som Kahneman och Tversky frågade berättade att Mr Tees skulle vara mer upprörd. Goldman gör det tydligt att teoriteorin skulle förklara det här genom att säga att alla testpersonerna höll inne med några relevanta generaliseringar utifrån vilka de kunde härleda relativa grader av upprördhet, någonting han inte tycker verkar rimligt. Inte heller tycker han det verkar rimligt att som vissa försvarare av teoriteorin antyder, nämligen den att vi lär oss dessa vardagspsykologiska generaliseringar ”i vår mammas knä” eller ”genom modersmjölken”. Detta därför att, som Goldman uttrycker det, få av oss har mammor eller andra vuxna, som yttrar vardagspsykologiska plattityder. Filosofer är i princip de enda som har försökt att explicit formulera dessa plattityder, och få barn har kontakt med filosofer.

Ett problem för simuleringsteorin är det om Anita Hill och Clarence Thomas. I och med att få män kunde sätta sig in i att hon inte kände att hon hade möjligheten att säga ifrån eller säga upp sig själv betyder det inte att det finns en garanti för verklighetstroga simuleringar. I det nämnda fallet är det helt enkelt inte så att de sociala strukturerna är de samma för män som för kvinnor. Här ligger skillnaden mellan de bägge teorierna, simuleringsteorin kräver inte att man förstår hur ens beslutfattande (eller andra psykologiska mekanismer) fungerar, det räcker med att man har förmågan till beslutfattande och att man är kapabel att mata den med lämpligt input och tolka dess output. Detta, skriver Goldman, är ett mycket mindre åtagande

än det från teoriteorin. Varför han skriver att inget hindrar en från att ha en mix av de båda teorierna som hypotes, snarare är det oundvikligt. För visst finns det visa enkla beteendemönster vi kan härleda slutsatser från, precis som vi kan härleda generaliseringar från tidigare simuleringar.

Eftersom vi faktiskt kan syssla med kontrafaktiska hypotetiska resonemang klargör Goldman att en rimlig förklaringsmodell är minst den svaga hypotesen, det vill säga att simuleringar åtminstone spelar någon roll. Vad han tycker talar för den starka hypotesen är återigen det som Gordon var inne på, nämligen det om barnens utveckling och förmåga att kunna skilja sig själv från andra i exempelvis interpersonella tolkningar. I korthet, vilket inte gör resonemanget rättvisa, skriver Goldman att teoriteoriförespråkarna måste prata om en radikal växling mellan olika teorier någonstans mellan en tre- och fyraåring; från en icke-representerande till en representerande begreppsförmåga rörande exempelvis trosuppfattningar. Goldmans förklaringsmodell är den att det snarare rör sig om att treåringar är sämre på att isolera sina egna faktiska trosföreställningar så att de ”släpps in” i simuleringarna. Detta verkar vara ett rimligt antagande, uttrycker han, eftersom det faktiskt är så att treåringarnas standardmisstag är att de hänvisar till sina egna trosföreställningar, något som fyraåringar, generellt sett, inte gör. Det är också så att mycket tyder på att treåringar är kapabla att förstå andras liknande misstag, något som inte teoriteoriförespråkarna verkar gilla. Goldmans slutsats är att simuleringsteorin, mycket bättre än teoriteorin, förklarar de empiriska data som finns till hands och därför är bättre på att uppskatta de grundläggande processer av interpersonella tolkningar. Därför försvarar Goldman den starka hypotesen, simuleringar är grundläggande i vår förståelse av andra.