

Epistemiska trösklar, ett svar på problemet med gemensam kunskap

1. INTRODUKTION

Debatten kring kollektiv intentionalitet handlar om hur vi ska förstå hur två eller flera individer kan agera som en grupp. I debatten finns det fyra etablerade teorier Bratman (1999), Gilbert (1989), Searle (2010) och Tuomela (1995). Gemensamt för alla teorier är att de anser att *gemensam kunskap* (eng. common knowledge) är ett nödvändigt villkor för kollektiv intentionalitet. Även Searle som i tidigare texter (t.ex. Searle 1995) hävdade att gemensam kunskap inte alls är nödvändigt för kollektiv intentionalitet har ändrat sig: ”När jag talar om denna form av kollektiv intentionalitet talar jag om människors och andra djurs förmåga att faktiskt *samarbeta* i sina aktiviteter. Samarbete *implicerar* gemensam kunskap” (Searle 2010, s. 49, författarens betoning).

Gemensam kunskap är en teknisk term i filosofi och logik som beskriver att ett faktum är epistemiskt öppet (allmänt tillgängligt) i en grupp. Traditionellt har gemensam kunskap definierats som *jag vet att du vet att jag vet ... att du vet*. Men det här leder till problemet att det krävs en extremt hög kognitiv förmåga för att kunna ingå i gemensam kunskap. Det här skulle innebära att åtminstone barn inte skulle kunna ingå i en sådan situation, men det finns mycket psykologisk forskning som pekar på att de både kan ingå i situationer av kollektiv intentionalitet och epistemisk öppenhet. Jag kommer i den här artikeln att presentera en psykologiskt rimligt teori om hur barn kan delta i epistemisk öppenhet.

2. GEMENSAM KUNSKAP OCH EPISTEMISK ÖPPENHET

För att illustrera hur ett faktum kan vara epistemiskt öppet kan vi föreställa oss en arbetsplats där alla anställda vet att chefen kommer att komma sent till morgonmötet, men ingen vet att alla andra också vet det. Det här sakförhållandet kallas ömsesidig *kunskap* (eng. mutual knowledge). Däremot om någon ropar ut att de precis har pratat med chefen och att hen hälsar att hen kommer bli sen, så ändras situationen från ömsesidig kunskap till gemensam kunskap att chefen kommer bli sen. Det som har

ändrats är att faktumet att chefen kommer att bli sen är nu öppet och offentligt bland de anställda.

Motivationen bakom gemensam kunskap brukar ligga i att förklara olika typer av gemensam kunskaps problem. Ett sådant problem föreligger t.ex. vid koordinerat handlande.

Vi kan tänka oss ett kärnkraftverk som behöver stängas ner p.g.a. en brand. För att stoppa reaktorn så finns det två personer Olof och Lisa som måste sätta in varsin nyckel i varsitt nyckelhål. Av säkerhetsskäl så är nyckelhålen på olika sidor av reaktorn, dessutom som en säkerhetsåtgärd om bara den ena nyckeln sätts in och inte den andra inom en minut så kommer nyckelhålet att sluta fungera under en timma. P.g.a. eldsvådan så kan Olof och Lisa inte nå varandra och planera när de ska sätta i nycklarna. Som tur är finns det skärmar bredvid nyckelhålen så de kan skicka meddelande till varandra. Olof och Lisa upptäcker dessvärre att skärmarna har blivit trasiga av eldsvådan till den grad att endast 50 procent av deras kommandon/knapptryckningar fungerar. Däremot så lyckas varje meddelande skickas, vilket de inte vet om. Olof skickar till Lisa att de borde sätta in nyckeln vid tidpunkt t. Olof vet inte om Lisa har fått meddelandet vilket Lisa inser, så hon skickar tillbaka ett bekräftelsemeddelande. När Olof får Lisas meddelande inser han att Lisa inte vet om han hade fått bekräftelsemeddelandet och han vet att Lisa inte kommer agera vi om hon inte är säker på att Olof har fått meddelandet. Så Olof skickar en bekräftelse på bekräftelsemeddelandet och så vidare och så vidare. Det här är problemet med gemensam kunskap uttryckt i ett koordineringsexempel.

David Lewis var den första att definiera gemensam kunskap och hans definition ses än idag som standarddefinitionen. Han började med att karakterisera gemensam kunskap som en oändlig iteration av jag vet att du vet att jag vet etc. Lewis insåg att det här var ett problem då det inte finns någon människa som kan ha en oändlig kedja av propositioner i huvudet eller ha ett oändligt antal mentala tillstånd samtidigt. För att kunna hantera det här så gör han en skillnad på *faktiska trosföreställningar* och *skäl att tro* (*actual belief* och *reason to believe*). Han definierar gemensam kunskap som följande:

[D]et är *gemensam kunskap* i en population P att X om och endast om något sakförhållande A är fallet, på så vis att:

- (1) Alla i P har skäl att tro att A är fallet.
- (2) A indikerar för alla i P att alla i P har skäl att tro att A är fallet.
- (3) A indikerar för alla i P att X. (Lewis 1969, s. 56, författarens betoning)

Lewis menar att även om denna teori leder till en oändlig regress så är det en regress av implikationer av skäl att tro och inte av faktiska tros-

föreställningar. Men även en oändlig kedja av implikationer kräver en viss nivå av kognitiv kapacitet. Lewis förstår det och fortsätter med att säga:

Resten av vad de behöver är det som de behöver för att bilda högre ordningens förväntningar på det sätt vi behandlar: ömsesidig tillskrivning av några allmänna induktionsstandarder och bakgrundsinformation, rationalitet, ömsesidigt tillskrivning av rationalitet och så vidare. (Lewis 1969 s. 56–57)

För att någon ska kunna delta i gemensam kunskap så behöver åtminstone dessa kriterier uppfyllas. Från ett utvecklingspsykologiskt perspektiv är Lewis teori, och alla andra teorier om gemensam kunskap, väldigt otillfredsställande. Småbarn verkar exempelvis inte uppfylla dessa kriterier, men det finns flera studier som visar på att barn redan vid 1 års ålder förstår vad som är epistemiskt öppet i en grupp och vad som inte är det. Jag ska ge två exempel. I en studie på spädbarn inte mer än tolv månader gamla av Liszkowski, Carpenter och Tomasello från 2008 så undersökte de spädbarns pekande efter objekt när den vuxna letade efter något. De fann att om barnet hade sett att den vuxna inte hade sett objektet ramla ner så pekade de ut objektet för den vuxna, till skillnad från när barnet hade sett att den vuxna hade sett objektet ramla ner. Den andra studien från samma år gjord av Moll, Richter, Carpenter och Tomasello undersökte hur fjortonmånaders barn förstod referens baserad på tidigare erfarenhet med objekten och den vuxna. I studien fick barnet leka med en vuxen och tre leksaker. Den vuxna visade stor entusiasm för den ena leksaken. Efter en stund så frågar den vuxna om hen kunde få en leksak, varpå hen fick den hen hade visat entusiasm över. I kontrollstudien kom en annan vuxen som barnet kände in och frågade om hen kunde få en leksak och då var sannolikheten att det skulle vara den speciella leksaken lika med slumpen. De här två studierna visar på att barn förstår vilka fakta som är öppna mellan dem, men det skulle fortfarande vara orimligt att mena att ett tolv månaders barn skulle uppfylla kraven för gemensam kunskap som nämnts ovan.

3. PROBLEMET MED EN AVKORTAD REGRESS

Av de skäl som ovan nämnts anser många att diskussionen kring gemensam kunskap är en abstrakt filosofisk fråga som inte har bäring på verkligheten. Många forskare har till och med hävdade att vi inte behöver den här långa kedjan av propositioner. Tomasello menar, i sin bok *Origins of Human Communication* från 2008, att vi inte ska se regressen som oändlig, utan snarare som odefinierad. Men denna typ av avkortningsteorier burkar ofta underskatta, eller missförstå motivationen till, problemen med gemensam kunskap. Michael Wilby formulerar det så här:

Problemet med [gemensam] kunskap kommer inte från något förmodat glapp mellan *psykologiska begränsningar* hos mänskliga agenter å ena sidan och en *idealiserad uppfattning* om vad [gemensam] kunskap skulle bestå å den andra, utan snarare från ett glapp mellan *psykologiska begränsningar* hos mänskliga agenter å ena sidan och behovet av en *psykologiskt ändamålsenligt uppfattning* om vad [gemensam] kunskap består i å den andra. (Wilby 2010, s. 86–87, författarens betoning)

För att förstå vad Wilby syftar på här så kan vi återgå till kärnkraftverks-exemplet. Wilby menar att om avkortningsteorierna stämmer så kan vi givet någon nivå k i regressen säga att $k + 1$ är tillräcklig för gemensam kunskap mellan Olof och Lisa. Det här innebär att Lisa som skickar $k+1$ meddelandet kommer sätta in nyckeln vid tidpunkt t utan att ha fått något bekräftande meddelande. Men eftersom vi vet att de bara kan agera tillsammans, så innebär det att Lisa kommer agera oavsett om meddelandet kommer fram eller inte. Men då är $k + 1$ meddelandet överflödigt och det skulle då vara tillräckligt med meddelande k som Olof skickade. Dock leder det här till att hela regressen blir meningslös eftersom meddelande k skulle kunna vara det första meddelandet som Olof skickade och vi skulle inte anse att vi då har någon form av epistemisk öppenhet.

Jag tror dock inte att det är kört för en avkortningsteori och jag kommer senare att lägga fram ett förslag på en sådan teori, efter vi ha sett vilka andra typer av teorier det finns i debatten.

4. ETT RELATIONELLT ALTERNATIV

John Campbell (2005) menar att han har kommit på en lösning av vårt problem, att hitta en teori som både är psykologiskt rimlig och som klarar av problemen med gemensam kunskap som vårt koordinerings-exempel. Campbell pratar inte om gemensam kunskap utan om *delad uppmärksamhet* (eng. joint attention). Men det han säger går att applicera på gemensam kunskap, något Wilby (2010) gör. Campbell skiljer mellan reducerande och relationella analyser av gemensam kunskap. Enligt den reducerande analysen ska vi kunna beskriva de mentala tillstånd Olof är i, utan att implicera att han står i en gemensam kunskapsrelation med Lisa. Lewis teori om gemensam kunskap är enligt denna uppdelning en reducerande teori. En relationell analys, å andra sidan, menar att en analys av Olofs mentala tillstånd implicerar att det finns någon, Lisa, som Olof har gemensam kunskap med. Enligt den relationella analysen är gemensam kunskap ett primitivt fenomen hos medvetandet. Både Wilby och Campbell menar att vi aldrig kommer att kunna lösa ett koordinationsproblem med en reducerande teori. De menar att antingen kommer en sådan teori att vara psykologiskt orimlig eller också kommer den att

skjuta fram problemet ett steg. Exempel på det senare skulle vara, att om vi accepterar Lewis teori som rimlig, så har vi fortfarande en oändlig kedja av implikationer. Man skulle så klart kunna säga att dessa implikationer kan vara implicita hos Olof och Lisa i stället för explicita; om de t.ex. explicit vet att X och explicit vet att om X så Y, så vet de Y implicit. Men då uppstår problemet att de behöver vara säkra på att de i den här situationen får sluta sig till Y och att den andra gör samma sak. Härmed skjuts problemet bara fram ett steg och bekräftelsemeddelandena kommer snarare att handla om de är berättigade till att använda sin implicita kunskap.

I Campbells fall när två individer, Per och Stina, har delad uppmärksamhet, så innebär en relationell analys av Pers mentala tillstånd att Stina är konstituerande för Pers perceptuella medvetande som deltagare (eng. co-attender), utan att vara ett objekt för medvetandet.

Dock verkar den relationella förklaringsmodellen inte fungera, eller rättare sagt den förklarar inte det den ska förklara. Relationella teorier, som Campbell och Wilbys, menar sig förklara hur *vi* kan ha en specifik kunskap *ihop*. Men varifrån kommer det här *vi ihop*? De säger att *vi* är ett primärt och primitivt fenomen. Men vad betyder det egentligen? Det verkar betyda att Olof är berättigad att säga att han och Lisa konstituerar ett *vi ihop*, eller tillsammans vet att de ska agera vid tidpunkt *t*, på samma grunder som han är berättigad att säga att det finns en sten framför honom om han ser en sten. I det senare fallet så skulle vi ju säga att Olof är berättigad att tro att det är en sten framför honom om det verkar vara en sten där. Det skulle innebära att Olof anser att han och Lisa konstituerar ett *vi* endast om det verkar som om Lisa gör det också. Visserligen är det viktigt, till och med ett nödvändigt villkor som vi kommer se nedan, att något verkar vara öppet för att det ska vara öppet. Men det kan inte vara hela historien, att det verkar vara så gör att det är så. I exemplet med arbetsplatsen så räcker det inte med att säga att det verkar som om alla vet att chefen kommer bli sen, det behövs något mer. Den relationella analysen verkar inte kunna förklara det vi vill att den ska förklara. Jag kommer nu försöka återuppliva både den reducerande analysen och den avkortande versionen av en sådan analys för att kunna få fram en tillfredställande teori om gemensam kunskap.

5. ÖPPENHET OCH EPISTEMISKA TRÖSKLAR

Till att börja med tar vi det vi lärde oss från den relationella analysen, att något verkar vara fallet, som väsentligt. Om vi pratar om epistemisk öppenhet så verkar det vara nödvändigt att det verkar vara så att det är öppet, för att något ska vara öppet. Om vi lägger till det till vår formu-

lering av problemet så får vi; om det är ömsesidig kunskap att chefen är sen och det *verkar* som faktumet att chefen är sen är öppet för alla, så är det gemensam kunskap bland arbetarna att chefen kommer bli sen. I det här fallet, om vi lägger till att det verkar vara öppet till faktumet att alla redan visste det, så verkar vi vara berättigade till att sluta oss till att det är öppet. Det intressanta är att något kan verka publikt eller öppet innan alla faktiskt vet om det, däremot måste alla veta om det för att det i sin tur ska bli gemensam kunskap. Vi kan tänka oss en situation där Stina säger till Per att han verkar vara arg. Det kan till och med vara så att Per inte vet om att han är arg och först genom att Stina pekar ut det upptäcker att han är arg. Det är först då Pers ilska blir gemensam kunskap mellan Stina och Per. Andra tillfällen kan det vara så att Per upptäcker att han verkar arg men helt enkelt inte är det varpå det inte blir gemensam kunskap.

Av det här kan vi definiera vad jag kallar för *semi-öppet*: B är semi-öppen för A om intention $X =_{df}$ det verkar/ser ut för A som att B är öppen med sin intention att X.

Öppen med sin intention innebär helt enkelt att man ger en publik indikation för något som nämnts ovan. Med det här kan vi beskriva fenomenet att man kan *ställa in* sin uppmärksamhet efter någon annans uppmärksamhet. Vi kan säga att om Lisa är medveten om att Olof är medveten om X och Lisa är semi-öppen med att vara medveten om detta för Olof, så tror Olof att deras uppmärksamhet är *inställd* efter X. Om samma sak gäller för Lisa så får vi en situation där både Olof och Lisa tror att de har korrigerat sin uppmärksamhet efter varandra till att de båda är inställda på att uppmärksamma X. Det här leder till att X verkar vara öppet för dem båda.

För att se hur det här kan hjälpa oss få gemensam kunskap så behöver vi analysera koordineringsexemplet igen. En intressant sak med exemplet är att innan Olof skickar första meddelandet, så har varken Olof eller Lisa någon anledning att agera vid tidpunkt t . Vi kan beskriva det här som att deras individuella subjektiva sannolikhet för att agera vid tidpunkt t utan någon evidens eller skäl att tro så, är lika med noll. Vi kan formellt skriva ut det som att den subjektiva sannolikheten $P(\text{agera vid tidpunkt } t) = 0$. Den enda evidens som kan ändra deras subjektiva sannolikhet är, om det verkar som den andra kommer agera vid tidpunkt t . Om exempelvis Olof anser sig tillräckligt berättigad att anse det, så kommer han att agera, det samma gäller för Lisa.

När Olof skickar sitt första meddelande, vi kan kalla det M_1 , så får vi $P_{\text{Olof}}(\text{agera vid tidpunkt } t | M_1) = x_1$, där $0 < x_1 \leq 1$. P_{Olof} betyder Olofs subjektiva sannolikhet. Även om Olof inte vet om meddelandet kommer fram, så kommer själva akten av att skicka meddelandet att öka hans subjektiva sannolikhet dvs. $0 < x_1 \leq 1$. x_1 ligger här emellan p.g.a. att Olof är

osäker på om hans intention att agera vid tidpunkt t är öppen eller inte. Men hans subjektiva sannolikhet är högre nu, eftersom det *kan* vara så att meddelandet kommit fram.

När Lisa får meddelandet så händer samma sak för henne, dvs. hennes subjektiva sannolikhet för att agera vid tidpunkten t ökar. När hon skickar iväg sin bekräftelse, kalla det M_2 , så kommer även denna handling också att öka sannolikheten. Vi får då att $P_{\text{Lisa}}(\text{agera vid tidpunkt } t | M_1 \& M_2) = x_2$, där $0 < x_2 \leq 1$ och $x_1 < x_2$.

Något väldigt intressant händer när Olof får bekräftelsemeddelandet, nämligen, att genom själva meddelandet att Lisa vill agera vid tidpunkten t så ökar hans subjektiva sannolikhet, men även hans tidigare syn på M_1 ändras då han nu vet att det faktiskt kom fram. Vi får nu att $P_{\text{Olof}}(\text{agera vid tidpunkt } t | M_1 \& M_2) = x_2$, där $0 < x_2 \leq 1$ samt $x_1 < x_2$ och så vidare.

Givet premisserna för exemplet kommer Olof och Lisa aldrig att nå visshet dvs. det händer aldrig att $P(\text{agera vid tidpunkt } t) = 1$. Jag tror dock inte att det är nödvändigt. I stället vill jag postulera att alla agenter har en *subjektiv tröskel*, som bedömer om något är epistemiskt öppet eller inte. I det här exemplet så var det just att något verkade öppet, som var det enda som kunde ändra deras subjektiva sannolikheter och om tröskeln blir nådd, så kommer de att agera. Den brännande frågan är egentligen om de kommer att sätta in nycklarna när tidpunkten t kommer? Om de gör det, gör de det inte de då p.g.a. att det att agera vid tidpunkten t verkade öppet för dem?

Vi vet att efter ett par meddelandeväxlingar så vet båda att den andre vet om den andras planer att agera vid tidpunkt t och båda vet att den andra har bekräftat detta. I termer av semi-öppenhet: det verkar öppet för båda två att den andra är öppen med sin intention att agera vid tidpunkt t och att denne är medveten om att den andra är öppen med samma intention. Om vi utgår från att det finns en subjektiv tröskel, låt oss säga 0,7 för Olof och 0,9 för Lisa, så gäller att om de har nått dessa trösklar genom en ökning av sina sannolikheter baserad på sin kommunikation, så verkar det rimligt, baserat på deras standard för att agera, att de just agerar.

Om vi accepterar existensen av en subjektiv tröskel så kommer vi få en version av både den reducerande analysen och den avkortade teorin att fungera. Den här typen av teori är dessutom kompatibel med sannolikhetsteorier om hur hjärnan hanterar och interagerar information, t.ex. fri energi-principen (Friston, 2009), som argumenterar för att hjärnan försöker minimera överraskningar. Med en sådan läsning så kommer den subjektiva upplevelsen av gemensam kunskap vara till för att skapa förväntningar på vad andra kommer att göra. Om Olof och Lisa skulle fortsätta att skicka ett par meddelanden efter att deras tröskelnivåer har nåtts så skulle det vara för att minimera möjligheten att bli överraskade.

Den här tröskelteorin möjliggör självklart att Olof och Lisa har olika tröskelnivåer, som i exemplet ovan. Vi kan tänka oss att Olofs når sin tröskel men inte Lisa, vilket innebär att Olof agerar men inte Lisa. Det här är såklart något som kan hända och jag anser att en teori om gemensam kunskap även behöver kunna förklara varför det kan bli fel.

Det här är som sagt en reducerande analys av gemensam kunskap och en avkortningsteori, då jag menar att efter ett par nivåer så kommer tröskeln att vara nådd för dem båda. Teorin faller därför inte för Wilbys kritik av avkortningsteorier, då den kräver att det blir några iterationer, men att det är subjektivt för vardera agenten hur många som behövs givet kontexten. Dessutom så klarar den här teorin av att vara både psykologiskt rimlig och uttrycka gemensam kunskap i psykologiskt ändamålsenliga termer, utan att det blir ett glapp. Dessutom är frågan om det finns en tröskel en rent empirisk fråga som kan leda till flertalet intressanta studier.

6. SLUTSATS

Vi började med att se att alla etablerade teorier inom kollektiv intentionalitetsdebatten anser att gemensam kunskap är nödvändigt för att två eller flera personer ska kunna samarbeta. Vi såg sedan att som gemensam kunskap brukar definieras, så blir teorierna antingen psykologiskt orimliga eller så skjuter de bara upp problemet ett steg. Mitt förslag på hur vi ska analysera gemensam kunskap tycks hantera dessa problem samtidigt som det lyckas förklara den epistemiska öppenheten som de relationella teorierna smyger in i sin analys. Teorin om subjektiva trösklar leder dessutom till en del intressanta empiriska frågor, t.ex. om det finns en sådan tröskel, hur den utvecklas hos barn och hur tröskelnivåerna skiljer sig hos introverta och extroverta individer?¹

REFERENSER

- Bratman, M. 1999. *Face of Intention: Selected Essays on Intention and Agency*. Cambridge: Cambridge University Press.
- Campbell, J. 2005. "Joint Attention and Common Knowledge". I *Joint Attention: Communication and Other Minds*, utg. N. Eilan, C. Hoerl, T. McCormack och J. Roessler, s. 287–97. Oxford: Oxford University Press.
- Friston, K. 2009. "The Free-Energy Principle: A Rough Guide to the Brain?", *Trends in Cognitive Neuroscience* 13, nr 7, s. 293–301.
- Gilbert, M. 1989. *On Social Facts*. London: Routledge.

¹Tack till William Hedly Thompson för litteraturtips och kommentarer på en tidigare versioner av uppsatsen samt till Niklas Rydberg som har stått för översättning av citaten.

- Lewis, D. 1969. *Convention: A Philosophical Study*. Cambridge: Harvard University Press.
- Liszkowski, U., M. Carpenter och M. Tomasello. 2008. "Twelve-Month-Olds Communicate Helpfully and Appropriately for Knowledgeable and Ignorant Partners". *Cognition* 108, s. 732-39.
- Moll, H., N. Richter, M. Carpenter och M. Tomasello. 2008. "Fourteen-Month-Olds Know What 'We' Have Shared in a Special Way". *Infancy* 13, s. 90-101.
- Searle, J. 1995. *Konstruktionen av den sociala verkligheten*. Göteborg: Daidalos.
- Searle, J. 2010. *Making The Social World*. Oxford: Oxford University Press.
- Tomasello, M. 2008. *Origins of Human Communication*. Cambridge: The MIT Press.
- Tuomela, R. 1995. *The Importance of Us: A Philosophical Study of Basic Social Notions*. Stanford: Stanford University Press.
- Wilby, M. 2010. "The Simplicity of Mutual Knowledge". *Philosophical Explorations* 13, nr 2, s. 83-100.