

Hur gick det för Mary?

Olle Häggström¹

Ett av filosofens viktigaste redskap är tankeexperimentet. Ett sådant är inte avsett att vara verklighetstroget. Genom att renodla eller överdriva någon viss aspekt av verkligheten syftar det istället till att hjälpa oss tänka klarare om vad denna aspekt egentligen för med sig. Den kanske mest läste av alla idag verksamma filosofer, amerikanen Daniel Dennett, excellerar i tankeexperiment, och kallar dem intuitionspumpar.

Den australiensiske filosofen Frank Jackson presenterade 1982 följande tankeexperiment. Den briljanta neurobiologen Mary vet allt som överhuvudtaget finns att veta om människokroppens och den mänskliga hjärnans biologi, kemi och fysik, inklusive allt om hur det mänskliga färgseendet fungerar. Exempelvis vet hon – med Jacksons ord – ”exakt vilka våglängdskombinationer från himlen som stimulerar näthinnan och via det centrala nervsystemet ger upphov till den sammandragning av stämband och den utandning som tillsammans resulterar i yttrandet av frasen 'Himlen är blå'”. Men hon har aldrig själv sett några färger, då hon spenderat hela sitt liv i ett svartvitt rum med en svartvit TV-monitor som enda kommunikationslänk med omvärlden. Så en dag får hon komma ut ur sitt rum och uppleva yttervärlden i all dess färgprakt, inklusive ett fång röda rosor. När hon ser dessa rosor, kommer hon då att lära sig något nytt – nämligen hur det känns att se färgen röd – eller visste hon redan allt om den känslan i kraft av sina kunskaper om hjärnan?

Fallet Mary har stötts och blötts av filosofer och andra under hela det kvarstsekel som därefter passerat, dock utan att diskussionen närmast sig någon konsensus. Så till den grad har Marys öde engagerat att hon är huvudperson i inte mindre än två aktuella antologier: *There's Something About Mary: Essays on Phenomenal Consciousness and Frank Jackson's Knowledge Argument* redigerad av Peter Ludlow, Yujin Nagasawa och Daniel Stoljar (MIT Press, 2004) och *Phenomenal Concepts and Phenomenal Knowledge: New Essays on Consciousness and Physicalism* redigerad av Torin Alter och Sven Walter (Oxford University Press, 2007). Titlarna syftar på fenomenella (inte fenomenala!) entiteter, dvs sådana som hör hemma i det mänskliga medvetandet. De två antologierna, här betecknade *TSAM* respektive *PCPK*, ger tillsammans en rik bild av vad som hänt Mary under hennes första 25 år: *TSAM* börjar med Jackson -82 och inkluderar ett urval av de viktigaste texter som följde, medan *PCPK* helt koncentreras på nyskrivet material. Här måste vi nöja oss med sporadiska nedslag.

(Parentetiskt kan infogas att Mary till och med hittat in i skönlitteraturen: David Lodges fiffiga grepp i romanen *Thinks*, att låta huvudpersonen som ger en universitetskurs i *creative writing* uppdra åt sina studenter att avsluta historien om Mary samtidigt som de travestierar valfri profil ur litteraturhistorien, öppnar för underhållande läsning om än måhända inte för någon filosofisk fördjupning.)

Jackson menade i sin ursprungsartikel att det är uppenbart att Mary, då hon till slut får se färger, lär sig något nytt om synupplevelser. Men eftersom hon redan visste allt om fysiken kring dessa, så måste det nya hon lärt sig vara av icke-fysikalisk natur. Det som filosofer kallar *qualia*, dvs karaktären hos våra medvetandes grundläggande sinnesupplevelser, är alltså icke-fysikaliskt, och därmed måste fysikalismen – den metafysiska uppfattning enligt vilken allt som existerar är fysiskt – vara felaktig. Denna och liknande tankegångar har kommit att kallas kunskapsargumentet mot fysikalismen.

Fysikalismen, eller materialismen som i detta sammanhang är en synonym beteckning, växte sig under 1900-talet allt starkare. För 100 år sedan var det svårt att föreställa sig hur atomer och molekyler kunde formiera sig till *levande* varelser utan medverkan av någon mystisk élan vital, något som vi idag stärkta av

¹ Professor i matematisk statistik, Chalmers tekniska högskola.

biologins enorma framsteg inte har några problem med. Förekomsten av *intelligenta* varelser föreföll likaledes gåtfull tills dess att datorrevolutionen visade för oss hur det ur död materia går att bygga maskiner som både kan spela schack och manövrera kryssningsmissiler. Men ännu känner vi oss osäkra på hur förekomsten av *medvetna* varelser går att passa in i ett materialistiskt universum. Hur kan ren materia ge upphov till subjektiva upplevelser? Att en person blickar upp emot himlen och efter en kort stunds neural informationsbearbetning utropar ”Himlen är blå!” kan vi idag utan svårighet acceptera som produkten av rent fysikaliska processer, men att detta förlopp återföljs av att personen i sitt inre förnimmer färgen blå framstår som ett mysterium.

Det är detta mysterium – varför upplever vi qualia och varför har vi medvetanden? – som en annan australiensisk filosof, David Chalmers, döpt till *the Hard Problem*, för att betona att det ur filosofisk synvinkel är av en helt annan division än de många ”enklare” problem som rör specifika detaljer i hur hjärnan bearbetar information: hur signalerna från två ögon kombineras till djupseende, hur vi prioriterar mellan olika sinnesintryck som samtidigt pockar på uppmärksamhet, etc, etc.

Ett sätt att formulera *the Hard Problem* är i termer av zombier – filosofers beteckning på varelser som fysikaliskt är identiska med oss själva ned till minsta elementarpartikel men som inte har någon inre värld av förnimmelser och medvetande. Varför är vi inte zombier?

Om vi köper Jacksons kunskapsargument så måste vi förkasta fysikalismen. Men vad för slags världsbild skall vi anamma i dess ställe? Huvudalternativet är den dualism enligt vilken å ena sidan fysik, och å andra sidan medvetande, samexisterar som metafysiskt separata kategorier. Inom naturvetenskapen är dualism idag mer eller mindre ett skällsord, medan filosofer diskuterar begreppet med större allvar.

Antag nu att vi med stöd i kunskapsargumentet godtar dualismens uppdelning i det fysikaliska och det fenomenella. Det är då förhållandevis okontroversiellt att acceptera att den fysikaliska världen påverkar den fenomenella: de fysikaliska fotoner som strömmar mot min likaledes fysikaliska näthinna bidrar till min fenomenella upplevelse av färg. Svårare att bestämma sig för är om det föreligger någon kausal påverkan i motsatt riktning: har det fenomenella någon inverkan på det fysikaliska? Om inte, så kallas medvetandet ett epifenomen.

Jackson tar i sin ursprungsartikel om Mary ställning för den epifenomenalistiska uppfattningen att våra medvetna förnimmelser – qualia – inte utövar någon påverkan på det fysikaliska. Hans motivering består huvudsakligen i att visa att de vanligaste argumenten emot denna uppfattning inte håller, och vi står då kvar med argumentet om att qualia inte utgör epifenomen så är fysiken inte kausalt sluten (dvs den påverkas av sådant som inte hör till fysiken), vilket strider mot idén om en fulländad fysik. Om det finns icke-fysikaliska faktorer som påverkar fysikaliska så borde väl dessa faktorer inkorporeras i fysiken?

Här invänder ännu en filosof, amerikanen David Lewis, i en läsvärd och inflytelserik uppsats från 1988 som återfinns i *TSAM*, att epifenomenalismen har en del bisarra konsekvenser: när Mary ropar av hänförelse över att äntligen få uppleva känslan av att se rött, så orsakas utropet på intet vis av hennes upplevelse. Vi kan på samma sätt konkludera att det faktum att filosofer lägger så mycket tid på att diskutera medvetandet inte alls orsakas av att vi har medvetanden.

Lewis avvisar Jacksons kunskapsargument med stöd i en argumentation som går ut på att särskilja två slags kunskap: faktakunskap och färdigheter. Det är bara om Mary vid åsynen av de röda rosorna tillägnar sig ny faktakunskap som kunskapsargumentet håller, men Lewis menar att hon endast lär sig färdigheterna att föreställa sig, identifiera och återkalla vissa upplevelser. I denna brutala sammanfattning låter Lewis' resonemang mest som demagogi, men nyckeln till hans argument ligger i hans noggranna dissekering av begreppet faktakunskap, som det dock skulle föra för långt att gå in på här.

Jämfört med Lewis så är ovan nämnde Daniel Dennetts grepp för att avvisa kunskapsargumentet mer radikalt – han förnekar helt enkelt att Mary lär sig något alls av sina nya färgupplevelser. I sin uppmärksammade bok *Consciousness Explained* från 1991 ägnar han Mary ett avsnitt som också återges i *TSAM*. För det första påpekar han att premissen att Mary vet *allt* om hjärnans kemi och fysik måste tas absolut bokstavligt för att kunskapsargumentet skall fungera. För det andra hävdar han att den intuitiva slutsatsen att Mary naturligtvis lär sig något nytt bygger på vår oförmåga att riktigt föreställa oss tankeexperimentet fullt ut, och leker med detta i ett eget slut på Jacksons historia. Marys vänner beslutar sig för att skoja med henne genom att förbereda en ljusblå banan som det första hon får se av yttervärlden, och hon utropar glatt ”Ha! Skämtare där! Bananer är gula, inte blå!”, för att därefter förklara för sina förbluffade vänner att hon ju faktiskt redan visste *allt* om färgseendets neurologi och mer specifikt om vad för konsekvenser gula respektive blå objekt skulle ha på hennes nervsystem, och därmed vad för slags tankar hon skulle få – tankar hon omedelbart kunde känna igen när de dök upp vid åsynen av den blå bananen.

Dennetts behandling av fallet Mary kom att väcka en del ytterligare diskussion. Då han kände att flera av kritikerna missförstod honom (och om man t.ex. läser Howard Robinsons bidrag till *TSAM* kan man rentav få intryck av att vissa missförstånd var avsiktliga) valde han att återkomma mer utförligt till fallet i *PCPK*. Med stor analytisk skärpa kopplad med polemisk lustfylldhet fingranskar han här sina meningsmotståndares argument för att Mary faktiskt lär sig något nytt, och finner att de i samtliga fall förlitar sig på den avskalade intuitionen: ”För bövelen, hon kommer att bli förvånad!”. Ingen kan artikulera varför, klagar Dennett, eller svara på följande: ”Om det nu går bra att förklara upplevelsen av att se en triangel med några dussin ord, eller den av månsken över Paris i maj månad med några tusen [...], hur kan vi då vara så säkra på att det inte går att förklara hur det är att se rött, eller blått, med några miljoner eller miljarder ord?”.

Här kan man försvara Jackson och de andra med att varje filosofiskt argument måste ta avstamp i obevisade intuitiva antaganden (om inte annat så åtminstone *modus ponens* och andra logiska grundaxiom), och att vad Dennett här kritiserar är att vägen från intuitivt antagande till slutsats är mycket kort. Ett längre resonemang för att ta sig från antagande till slutsats må vara en större intellektuell prestation, men borgar det för en säkrare slutsats? Knappast. Vad Dennett behöver göra är att underminera kunskapsargumentets premisser genom att argumentera för dess motsats. Och det är detta han gör när han i sitt *PCPK*-bidrag går vidare och introducerar en turboversion av Jacksons intuitions pump: RoboMary.

RoboMary befinner sig i samma situation som sin kusin Mary, med den huvudsakliga skillnaden att hon är gjord av kisel och plåt istället för av kött och blod. Detta har vissa fördelar för tankeexperimentet, bl.a. då vi helt enkelt kan byta ut RoboMarys färg-TV-kameror till synorgan mot svartvita, och på så vis slippa en lång rad av de praktiska svårigheterna i fallet Mary, som t.ex. att hon på något vis måste hindras från att se rött genom att blunda och vända sig mot en lampa. (Någon kanske invänder att fallet RoboMary inte kan säga något om oss människor utan bara om robotar, en invändning som Dennett enkelt avfärdar med att den som kan modern biologi vet att vi människor i princip inte är något annat än robotar, uppbyggda av robotar, uppbyggda av robotar.) Den största fördelen med RoboMary är emellertid att det blir mycket lättare att specificera hur hon använder sig av sina fysikaliska kunskaper för att förutse hur hon kommer att uppleva färgen röd, och Dennett förklarar hur hon helt enkelt med hjälp av dessa kunskaper kan gå in och manipulera sitt eget syncentrum för att försätta det i ett fysikaliskt tillstånd som svarar mot att se rött. ”Fusk”, kanske någon nu ropar, vilket Dennett besvarar dels genom att påpeka att vad RoboMary här gör svarar mot vad vi själva gör när vi föreställer oss saker, och dels genom att förklara ett (lite omständligare) sätt på vilket RoboMary kan lösa uppgiften även om vi gått in och ”skrivskyddat” hennes syncentrum.

Debatten fortsätter. Men kommer vi någonsin att fullt ut förstå sambandet mellan hjärna och

medvetande? Frank Jackson avslutar sin artikel från -82 med ödmjuka reflektioner kring detta: Det är fantastiskt hur mycket av världen som den människohjärna som evolutionen frambringat tycks vara i stånd att begripa, men det finns knappast någon garanti för att den skall vara i stånd att begripa allt. Och kanske hör medvetandets plats i eller jämte det fysikaliska universum till det som vi faktiskt inte har kapacitet att förstå. Liknande tankegångar har framförts av en rad andra tänkare, inklusive kognitionsforskaren Steven Pinker och filosofen Colin McGinn.

Jackson fortsatte under ett antal år att försvara kunskapsargumentet, men gjorde vid mitten av 90-talet vad som i inledningen till *TSAP* beskrivs som ”en av de mest remarkabla helomvändningarna i nutida filosofi”: han förkastade kunskapsargumentet och anslöt sig till fysikalismen. I ett bidrag till *PCPK* förklarar han sin nuvarande position, där han i sitt försvar för fysikalismen tar stöd i den så kallade representationalism som betonar att vi inte har någon direkt upplevelse av världen utan bara av våra idéer och tolkningar av vad som finns i världen.

Den uppsats i *PCPK* som med sin mer jordnära diskussion av mänskligt färgseende sticker ut mest mot de övriga är den norske (men i Sverige uppväxte) psykologen Knut Nordbys bidrag. Nordby, som gick bort 2005 kort efter att han skickat in sin text till *PCPK*, kan beskrivas som ett slags verklighetens Mary: själv totalt färgblind forskade han kring det mänskliga synsinnet, bland annat i samband med utveckling av videotelefoni på Telenor. Men till skillnad mot Mary slapp Nordby aldrig ut ur sitt svartvita rum.