

Kahneman och Lindaproblemet

Bestsellern Daniel Kahneman, *Tänka, snabbt och långsamt* (2013) recenserades i *Filosofisk tidskrift* 2/2014 av Lars Bergström, och jag instämmer i de kritiska reflektioner han ibland gör. Men jag vill komplettera med kritik av Kahnemans berömda Lindaproblem, vilket Kahneman nu sammanfattar sin syn på (kap. 15). Diskussionen kring problemet finns för övrigt kortfattat presenterat i Wikipedia under rubriken ”Conjunction fallacy”, men min typ av kritik finns inte redovisad där.

Se på följande fråga: ”Vilket alternativ är mest sannolikt? (a) Linda är bankkassörska. (b) Linda är bankkassörska och verksam i feministrörelsen.”

Om jag skulle hittat frågan i en bok i sannolikhetslära, så skulle jag omedelbart ha svarat (a). Detta därför att jag vet att sannolikheten för en konjunktion (”Linda är bankkassörska *och* verksam i feministrörelsen”) enligt den matematiska sannolikhetskalkylen alltid är lägre än sannolikheten för någon av konjunkterna tagna för sig (t.ex. ”Linda är bankkassörska”). Men när jag i en offentlig lite interaktiv föreläsning första gången mötte frågan, så svarade jag, i likhet med många andra som kan sin sannolikhetslära, att alternativ (b) är det mest sannolika. Hur kommer detta sig?

Kahneman säger att mitt och andra sannolikhetskunniga personers val av (b) ”strider mot all logik” (s. 179). Han säger också att ”Tyvärr brydde vi oss aldrig om att ta reda på hur den stora minoritet (36 procent) av denna välinformerade grupp resonerade som valde fel alternativ” (s. 180). Den förklaring av mitt förment ologiska svar som jag själv dagen efter föreläsningen kom fram till ser ut som följer.

Till saken hör att i föreläsningen, liksom i Kahnemans experiment, föregicks frågan av följande information: ”Linda är trettioett år, singel, frispråkig och mycket intelligent; har examen i filosofi, och brunn som student för frågor om diskriminering och social rättvisa och deltog även i demonstrationer mot kärnkraft” (s. 177).

Det finns utanför böcker i sannolikhetslära åtminstone två olika rimliga preciseringar av påståendet ”Linda är bankkassörska”:

- (a₁) Linda tillhör mängden bankkassörskor.
 (a₂) Linda är en ganska typisk bankkassörska.

Låt oss nu se på sannolikheterna i de två fallen. Om $(a) = (a_1)$, så är det sannolikare att (a) är fallet än att (b) är det; och till yttermera visso är detta en logisk-matematisk sanning. Frågan som jag ställer är om detta i den givna situationen är den enda logiskt korrekta slutsatsen, och svaret menar jag måste bli nekande.

Om $(a) = (a_2)$, så är det tvärtom sannolikare att (b) är fallet än att (a) är det. Givet informationen om Linda, så det sannolikare att (b) hon är bankkassörska och verksam i feministrörelsen än att (a₂) hon är snarlik stereotypen för en bankkassörska. Man kan inte bedöma om personer drar logiskt korrekta slutsatser förrän man vet hur de uppfattar premisserna. Som jag uppfattade premisserna gjorde jag en helt korrekt logisk slutledning.

Som med all önskvärd tydlighet framgår av Kahnemans bok är jag i sällskap med många personer, varav åtskilliga säkert har mer sannolikhetslära i ryggmärgen än jag har. Något mer allmänt behöver därför sägas.

I kognitionsvetenskap och delar av lingvistik finns det många forskare som skriver under på den s.k. prototyp teorin. De mest namnkunniga tidiga företrädarna är Eleanor Rosch och George Lakoff. Peter Gärdenfors, för att nämna en svensk, har arbetat med att väva in antagandet om prototyper i en vidare teoretisk ram där prototyperna representeras av punkter i vad han kallar begreppsrum; se hans *Conceptual Spaces* (2000) och *The Geometry of Meaning* (2014). Enligt prototyp teorin förstår vi en mängd begrepp inte via en implicit hänvisning till en uppsättning nödvändiga och/eller tillräckliga villkor för begreppets applicering, utan primärt som implicit hänvisande till en prototyp.

Jag har långt före jag stötte på Lindaproblemet trott att det ligger mycket i prototyp teorin, vilket kanske t.o.m. var det som gjorde att jag så snabbt fann förklaringen till mitt skenbart felaktiga svar. Enligt teorin uppfattar vi normalt begreppet ”bankkassörska” som hänvisande till ”typisk bankkassörska”, vilket jag ju gjorde. Slutsats: Kahneman borde ha tänkt mer på hur vardagsspråket kan fungera innan han började beskylla sannolikhetskunniga personer för att tänka helt ologiskt.

Märkligt nog finns i Kahnemans bok inte en enda hänvisning till prototyp teorin. Detta trots att han drar in begreppet om prototyper i sina resonemang, och t.o.m. har indexerat ordet prototyp. Specialiseringen i forskarsamhället får ibland lite konstiga konsekvenser.

Som jag redan citerat, så skriver Kahneman i anslutning till Lindaproblemet att han inte brydde sig om att ta reda på hur de som enligt honom

tänkte ologiskt egentligen tänkte. Men det låter lite underligt, eftersom hela hans bok är ett försök att identifiera och klassificera olika mekanismer bakom mänskligt vardagligt feltänkande. Och det är faktiskt lätt att rekonstruera hur han ser på det förmenta misstag jag och många andra med mig gjort. Han säger för övrigt i förbifarten senare i boken att hans förklaringsmodell kan appliceras på Lindaproblemet (s. 429). Såvitt jag förstår menar han så här.

För att på ett enkelt sätt förstå många av våra felslut kan vi betrakta vårt intellekt *som om* det består av två åtskilda system. Det första, System 1, fungerar snabbt och okritiskt; ofta fungerar det bra, men förhållandevis ofta också fel. System 2 är metodiskt reflekterande, kritiskt-logiskt och kräver lite mera tid. Det kan träda in och rädda situationen när System 1 gått vilse, men detta sker långt ifrån alltid. En av människans egenheter är att det kan byta ut ett svårt problem, vilket egentligen kräver System 2, mot ett enklare problem som System 1 snabbt klarar av; och sedan ta svaret på det enklare problemet som ett svar på det svårare problemet. System 1 arbetar med olika typer av förenklingar. Det ”föreställer sig uppsättningar med hjälp av genomsnitt, normer och *prototyper* [min *emfas*], inte med hjälp av summor” (s. 429). Tillämpas förklaringsmodellen på mig och Lindaproblemet blir det så här.

Jag förstod först att (a) ”Linda är en bankkassörska” borde tolkas som (a_1) ”Linda tillhör mängden bankkassörskor”. Men för att förenkla det givna problemet, så tolkade jag snabbt om (a) till (a_2) ”Linda är en typisk bankkassörska”, och gav därför det enligt Kahneman felaktiga svaret att (b) är sannolikare än (a). Jag måste säga att detta i mina öron låter högst otroligt. Hade jag tolkat (a) som (a_1) hade jag precis lika snabbt och oreflekterat kunnat ge mitt svar som när jag tolkade (a) som (a_2). Kahneman är med sitt tal om prototyper snubblande nära den rätta förklaringen, men genom sin oproblematiska syn på språkförståelse ändå mycket långt ifrån.

Jag upprepar därför: Kahneman (och likasinnade) bör tänka lite mer på hur vardagsspråket kan fungera, innan de drar slutsatsen att vissa av deras experimentpersoner tänker på ett sätt som ”strider mot all logik”.