

## ERIK CARLSON

### *Kan vi veta vad vi tror?*

---

De flesta av oss accepterar att vi inte kan veta allt om verkligheten. Däremot tycker vi kanske att vi borde kunna ha kunskap, eller åtminstone sann tro, om vad vi själva tror om verkligheten. Detta är tveksamt. Det tycks finnas fall där en rationell agent inte kan känna till sina egna övertygelser.<sup>1</sup> Antag att jag har tillgång till en mycket säker metod att ta reda på dina övertygelser, vilket du känner till. Jag har många gånger demonstrerat min förmåga, med för dig förbluffande säkerhet. Nu berättar jag att jag bestämt mig för att dricka te till frukost i morgon om och endast om du om fem minuter inte tror att jag kommer att dricka te i morgon. Jag kommer alltså att dricka te om du antingen tror att jag inte kommer att göra det, eller saknar övertygelse i frågan.<sup>2</sup> Dessutom ger jag dig chansen att vinna ett pris. Du får om fem minuter skriva ned vad du då tror är det rätta svaret på frågan huruvida jag kommer att dricka te. Jag stoppar lappen med ditt svar i ett kuvert, som jag förseglar och öppnar i din åsyn vid lunchtid i morgon. Har du svarat rätt får du 1000 kronor.

Om du är intresserad av tusenlappen vill du förstås maximera dina chanser att svara rätt. Eftersom du hyser stark tilltro både till min förmåga att känna till dina övertygelser och till min ärlighet, så tror du på vad jag sagt. Detta innebär att du tror att det rätta svaret är

1 Kanske finns det psykologiska mekanismer som gör att vi ibland 'förtränger' övertygelser om obehagliga fakta och på så vis gör dessa övertygelser oåtkomliga för oss själva. Detta är kanhända inte alltid irrationellt. Mitt argument har inget med sådana fall att göra.

2 I fortsättningen kommer jag att använda 'övertygelse' i en vid mening, enligt vilken din övertygelse i frågan om ett påstående  $p$ :s sanning kan bestå i att du varken tror  $p$  eller  $\neg p$ .

beroende av din övertygelse om det rätta svaret. Alltså bör du ta reda på vilken din övertygelse är. Men det kan du inte, vilket vi nu ska se.

Mitt argument förutsätter att du uppfyller följande rationalitetsvillkor: du tror inte samtidigt både ett påstående och dess negation, och om du tror  $p$  och att  $p$  implicerar  $q$ , så tror du också  $q$ . Mer formellt kan dessa villkor uttryckas så här, där 'B' står för 'du tror vid tidpunkten  $t$  att':

$$(1) \neg(Bp \ \& \ B\neg p)$$

$$(2) Bp \ \& \ B(p \supset q) \supset Bq$$

Låt nu  $p$  vara påståendet att jag kommer att dricka te i morgon, och låt  $t$  vara om fem minuter. Då gäller att:

$$(3) Bp \equiv \neg p.$$

Du tror vid  $t$  fortfarande (3). Det vill säga:

$$(4) B(Bp \equiv \neg p).$$

Det finns, antar vi tills vidare, två varandra uteslutande och tillsammans uttömmande möjligheter vad gäller din övertygelse beträffande  $p$ , nämligen  $Bp$  och  $\neg Bp$ . Vi antar först  $BBp$ . Då följer, vilket är lätt att se,  $B\neg p$  via (2) och (4). Med hjälp av (1) följer i sin tur  $\neg Bp$ . Alltså:

$$(5) BBp \supset \neg Bp.$$

Låt oss nu anta  $B\neg Bp$ . Då följer  $Bp$  via (2) och (4). Det gäller således att:

$$(6) B\neg Bp \supset Bp.$$

Med andra ord, om du tror att du tror  $p$ , så tror du inte  $p$ , och om du tror att du inte tror  $p$ , så tror du  $p$ .<sup>3</sup> Genom att berätta för dig att jag tänker basera mitt beslut om att dricka te på din övertygelse om

3 Det finns nog också fall där (5) och (6) är trivialt sanna därför att  $p$  är ett påstående som du inte har någon möjlighet att reflektera över, eller ens formulera. Då tycks det gälla att  $\neg Bp$ ,  $\neg B\neg p$ ,  $\neg BBp$  och  $\neg B\neg Bp$ . (Låt  $t$  ex  $p$  handla om en viss amöba som levde för två miljarder år sedan.) Mitt exempel är intressantare än sådana fall, eftersom det är förenligt med  $Bp$  och med  $B\neg p$ , samt rör påståenden som du har reflekterat över.

mitt tedrickande, så har jag alltså gjort vissa av dina övertygelser oåtkomliga för dig själv. Ditt svar på frågan om jag kommer att dricka te måste alltså bli en ren gissning. Åtminstone gäller detta om du uppfyller de inte särskilt starka rationalitetsvillkoren (1) och (2).

En möjlig invändning mot mitt argument är att hävda att det är principiellt omöjligt för mig att ha den kunskap om dina övertygelser som exemplet förutsätter. Detta verkar tveksamt, och även om det skulle vara sant så spelar det egentligen ingen roll för argumentet. (3) är logiskt ekvivalent med följande disjunktion:

$$(7) Bp \ \& \ \neg p \ \vee \ \neg Bp \ \& \ p.$$

Det är obestridligt att det finns åtminstone något påstående  $p$  och någon tidpunkt  $t$ , för vilket (7) är sann. Säkert tror vi alla åtminstone någonting som är falskt, och det är ännu mer uppenbart att vi inte tror allt som är sant. För var och en av oss finns det, med andra ord, många påståenden för vilka (7), och därmed (3), gäller.

En intressantare invändning är att det inte kan vara rationellt av dig att vid  $t$  tro att (3) är sant, för ett givet påstående  $p$ , även om (3) faktiskt är sant. Ett argument för detta kunde bestå i att påpeka att det tycks gälla för var och en av disjunkterna i (7) att det är irrationellt, eller kanske till och med omöjligt, för dig att vid  $t$  tro på den. Men det faktum att du inte rationellt kan tro på någotdera av två påståenden är fullt förenligt med att du rationellt kan tro på deras disjunktion. Om  $p$  är påståendet att antalet gäddor i Mälaren är udda, så har jag ingen rationell grund för att tro vare sig  $p$  eller  $\neg p$ . Däremot är det högst rationellt av mig att tro disjunktionen  $p \vee \neg p$ .

Dessutom finns det realistiska exempel, där det inte verkar irrationellt att vid  $t$  tro (3): Du är medveten om din charm, men du vet samtidigt att din övertygelse om din egen charm tenderar att göra dig mindre charmig. Du är väl förberedd inför din tenta, men vet att du börjar slarva om du medan du skriver är övertygad om att du kommer att klara tentan. Du uppfattar dig själv som en ödmjuk person, men tror att övertygelsen om din egen ödmjukhet är oförenlig med äkta ödmjukhet. Dessa exempel är instanser av (3), om  $p$  är att du charmar din samtalspartner, klarar din tenta, eller är en ödmjuk person. De övertygelser jag tillskrivit dig i de här exemplen tycks inte vara irrationella eller omöjliga. Hur som helst krävs det något starkt argument för att göra troligt att det för *varje* påstående  $p$  är irrationellt

av dig att vid  $t$  tro (3).

En tredje och sista invändning går ut på att förneka att bivalensprincipen, som säger att varje påstående antingen är sant eller falskt, gäller för påståenden om övertygelser. Tro kanske är ett vagt begrepp, på samma sätt som skullighet. Vissa filosofer hävdar att det kan vara så att det varken är sant eller falskt att en viss person är skullig. Påståendet att personen ifråga är skullig kan ha ett tredje sanningsvärde, nämligen 'obestämt'. Låt '?' stå för 'det är obestämt huruvida'. Kanske kan det vara så att  $?Bp$ . I så fall kan du, trots att (4) gäller, vid  $t$  tro  $?Bp$ , utan att denna övertygelse implicerar att det inte gäller att  $?Bp$ .

Också denna invändning är problematisk. För det första är det högst kontroversiellt att förneka bivalensprincipen. För det andra är det enkelt att modifiera mitt exempel, så att det undkommer invändningen. Vi kan helt enkelt stipulera  $t$  ex att  $?Bp \supset p$ , och att du vid  $t$  tror detta. Då gäller, om du uppfyller (2), att  $B?Bp \supset Bp$ . Alltså har du fel om din övertygelse också om du tror  $?Bp$ . Man kan då kontra med att införa en 'andra ordningens' obestämdhet, och hävda att det kan vara obestämt huruvida  $?Bp$ . Det är då lätt att modifiera exemplet för att ta hand om detta, varpå vi kan införa en tredje ordningens obestämdhet, osv. Om invändningen ska hota min tes måste vi alltså acceptera en oändlig hierarki av obestämdheter. För det tredje, och viktigaste, tycks det inte rimligt att anta att det här råder obestämdhet (av någon ordning) huruvida  $Bp$ . (4) innebär att du tror att huruvida  $p$  är sant beror på om  $Bp$  eller  $\neg Bp$  är sant. Den enda evidens du kan ha för  $p$  eller  $\neg p$  är alltså  $Bp$  eller  $\neg Bp$  (eller något påstående som stödjer något av de två sistnämnda påståendena). Men du inser snart att vare sig du vid  $t$  tror  $Bp$  eller  $\neg Bp$ , så har du fel. Du har alltså inte någon evidens för vare sig  $p$  eller  $\neg p$ . Eftersom du inte har någon sådan evidens kan du inte heller tro på  $p$  eller på  $\neg p$ . Du lämnar alltså frågan om  $p$ :s sanning öppen, åtminstone om du är rationell. Detta är ju ett specialfall av  $\neg Bp$ , vilket alltså är din övertygelse. Men detta kan du inte veta, eller ens tro.<sup>4</sup>

4 Tack till Erik Olsson och Krister Bykvist för värdefulla synpunkter på ett tidigare utkast.