

REBECCA SCHWEDER

Förklaring som förståelse – en koherensteori för förklaring

1. Introduktion

Det verkar som förklaringens första och främsta uppgift är att ge förståelse. Att ha förklarat hur en båt flyter är att *förstå* hur detta går till, liksom att ha förklarat hur en addition skall utföras är att *förstå* vad det innebär att addera ett tal till ett annat och få fram ett korrekt resultat. Jag menar att det finns en stark koppling mellan att förklara och förstå, och att denna koppling har förbisetts av de flesta förklaringsteorier.

Trots att många teorier om förklaring nämner förståelse som en viktig komponent, så har förvånansvärt lite skrivits om vilken betydelse förståelse har i förklaringar. Syftet med artikeln är gå igenom och kritisera några teorier om förklaring samt att titta på vad de säger om förståelse. Ingen av de teorier jag granskar har en tillfredsställande syn på kopplingen mellan förklaring och förståelse. Jag menar att en koherensteori för förklaring är det bästa alternativet, och den enda teori som faktiskt kan reda ut förhållandet mellan förklaring och förståelse. Utrymmet tillåter inte någon fullständig översikt över förklaringsteorierna, vilket har gjort att jag begränsat diskussionen till teorier som ligger närmast den teori jag själv förespråkar.

2. Förståelse som nomisk förväntan

Hempels deduktiv-nomologiska (DN) syn på förklaringar är välkänd. Vad säger teorin om förklaring och förståelse? Eftersom en DN-förklaring är ett logiskt argument, så visar förklaringen att fenomenet ifråga var att vänta, givet den eller de lagar och initialvillkor som utgör premisserna (Hempel s 337). Förklaringen ger förståelse därför att den gör det möjligt att förutse fenomenet. I strikt mening behöver fenomenet inte förutses; eftersom det har givit upphov till en för-

klaringssökande varför-fråga, så vet vi ju redan att det har inträffat. DN-modellen gör inte någon skillnad mellan att förklara och förutse, eftersom premisserna kan användas i båda syftena. Att *förstå* ett fenomen är helt enkelt att *veta* att det skulle inträffa, givet vissa lagar och initialvillkor. Det är det deduktiva förhållandet mellan explanans och explanandum som ansvarar för förståelsen i en förklaring. För Hempel är begreppet förståelse en helt och hållet formell angelägenhet. Att förstå är, per definition, att veta vilken den relevanta lagen är, och tillsammans med vilka initialvillkor det som förstås kan härledas deduktivt.

3. Kritik av DN-modellen

Det är vanligt att bara ange en orsak eller ett enda faktum i en förklaring (Scriven 1962). Följande påstående är en vanligt förekommande typ av förklaring, men tycks inte kunna passas in i ett DN-schema: "Dykaren sjönk som en sten därför att hennes torrdräkt läckte in vatten."

Att det finns sådana "enkla" orsaksförklaringar har inte undgått Hempel, som menar att förklaringar som inte anger någon generell lag är *elliptiska*. Den enkla orsaksförklaringen är, menar han, en förkortad variant av en genuin DN-förklaring, och några av premisserna är "underförstådda" (s 415). Frågan är, hur man kan omfatta en formell teori om förklaring och samtidigt hävda att några av premisserna är underförstådda – är det så att orsaksförklaringen logiskt *förutsätter* lagpåståenden om vikt- och volymförhållanden i vatten? Detta verkar orimligt. Ingen meningsanalys i världen kan extrahera lagpåståenden om flytkraft hos föremål i vatten ur påståendet "dräkten läckte in vatten."

Om lagpåståendena inte är logiskt förutsatta återstår bara möjligheten att förklaringen relativiseras till en frågeställare och någon som förklarar. Men det uppstår svårigheter med att redogöra för vad det innebär att lagpåståendena är underförstådda. För att använda ett av Hempels egna exempel: Förklaringen till varför smöret smälter är att det lades i en het stekpanna. En person som känner till de enklaste fysikaliska lagarna kan ge förklaringen med ett uttalat eller underförstått citerande av de relevanta lagarna. Men faktum är att nästan vem som helst skulle kunna använda sig av den förklaringen, till och med ett barn utan minsta kännedom om de relevanta lagarna. Hempel blir

oss svaret skyldig vad slags förklaringar enkla orsaksförklaringar egentligen är.

Dessa svårigheter pekar mot att den formella analysen av förklaring måste överges. Scriven är i mitt tycke den som tagit utmaningen på störst allvar. Hans kritik av den formella förklaringsteorin leder fram till en pragmatisk utgångspunkt.

4. Den pragmatiska utgångspunkten

Scriven menar att formalisten gör misstaget att blanda ihop den vetenskapliga formuleringen av en förklaring med förklaringen själv (1975, s 12). Det som intresserar honom är förklaringens essens, och han finner ingen anledning att särskilja vetenskapliga och vardagliga förklaringar. Han utgår från det som är gemensamt för sökandet efter förklaringar i vetenskapliga sammanhang och sökandet efter förklaringar i typiska vardagssituationer. Det är, med andra ord, ingen essentiell skillnad mellan att besvara frågan "Varför dog patienten?" och "Vilken väg är det till festen?" Ett svar på den sistnämnda frågan är "bra, bara i den mån de knyter an till frågarens förkunskaper," och svaret bör vara konsistent med naturlagarna, de lagar som styr språket och förhållandet mellan olika riktningar som norr, söder, öster och väster. Men bortsett från det, "så fungerar vad som helst, beroende på omständigheterna" (s 10). Scriven ser ingen anledning att tro att det skulle förhålla sig annorlunda med vetenskapliga förklaringar.

Det pragmatiska synsättet bygger på insikten att en förklaring alltid är en förklaring *för någon*. Förklaringen söks av någon som iakttagit ett fenomen som för denne verkar gåtfullt eller obegripligt eller överraskande. En förklaring är *svaret på en varför-fråga*, och det är en förklaring om det leder till att frågeställaren inte längre betraktar fenomenet som gåtfullt, obegripligt eller överraskande.

Pragmatismen är en metodologisk utgångspunkt, inte en fullfjädrad teori om förklaring. Pragmatiska teorier skiljer sig åt vad beträffar kriterier för att något skall accepteras som en förklaring och under vilka betingelser en frågeställare upphör att betrakta ett fenomen som gåtfullt och i behov av förklaring.

5. Talaktsteorin för förklaringar

Achinstein (1983) menar att begreppet "förklaring" i första hand skall analyseras som en talakt. Förklaringstalakten är fundamental, och

andra sätt att använda förklaringsbegreppet kan alltid återföras till en sådan talakt (s 22). Det innebär att bland annat att en vetenskaplig teori kan ses som en *produkt* av en sådan talakt, snarare än en förklaring i sig själv (Jag förmodar att Achinstein avser att förklaringen ska ses som en fiktiv eller verklig talakt).

Achinstein ställer upp följande nödvändiga och tillräckliga villkor för att en talakt skall vara en förklaring. S är den som förklarar, u är en utsaga och Q står för en direkt varför-fråga, q för den indirekta formen av samma fråga:

- 1) S yttrar u med avsikten att hans yttrande ska göra q begripligt (för någon åhörare, min anm).
- 2) S tror att u uttrycker en proposition som är ett korrekt svar på Q.
- 3) S yttrar u med avsikten att hans yttrande ska göra q begripligt genom att u frambringar den kunskap som är det korrekta svaret på Q. (s 16ff, min översättning)

Achinsteins tanke är att en förklaring innebär "att man yttrar någonting med avsikten att göra q begripligt på ett särskilt sätt" (s 23). S, den som förklarar, försöker helt enkelt få sin åhörare att *förstå*.

Förståelse spelar en central roll i analysen. Man har *förstått* q bara om man känner till ett korrekt svar till Q, ett svar vilket man *vet* är korrekt (s 23). Achinstein betonar att för att verklig förståelse skall infinna sig räcker det inte med att veta att det finns *något* riktigt svar; inte heller att vara rätt svar "på spåret," (som polisen som, helt korrekt, *tror* att olyckan förorsakades av halka). För att verkligen förstå måste man befinna sig i det epistemiska tillståndet att man *vet* att man är i besittning av det riktiga svaret (s 28).

Jag håller med Achinstein om att förklaring essentiellt involverar förståelse, men jag är skeptisk till hans *beskrivning* av förståelse. Beskrivningen anger ett nödvändigt, men inte tillräckligt, villkor för förståelse. För att visa detta, tar jag till min hjälp ett mycket vardagligt exempel: Låt oss tänka oss en skolelev som räknat färdigt en matematisk uppgift i sin övningsbok. Eleven kontrollerar sitt svar i facit, och upptäcker att han räknat fel. Efter att ha gått igenom facit, *vet* han det korrekta svaret och han *vet* också att det svaret är det korrekta svaret. Men det är inte säkert att han *förstått*. Eleven kan fortfarande vara förbryllad över svaret och okunnig om hur han skall räkna för att

komma fram till det rätta svaret. Eleven har helt enkelt inte de kognitiva resurser som krävs för att han skall *förstå*, i ordets fulla bemärkelse.

Problemet är att Achinsteins beskrivning inte kan skilja kunskap från förståelse, vilket får ödesdigra konsekvenser för hans analys av förklaring. För Achinstein innebär förståelse av x att x är ett svar på en fråga Q , att x utgör kunskap för A , och att A *vet* att x är kunskap. Enligt denna beskrivning är ju x helt enkelt ett stycke internalistiskt rättfärdigad kunskap. Resultatet blir, att den föreslagna analysen kommer att säga något helt annat än vad som avsågs från början. En förklaringstalakt i denna nya mening är, att någon genom att yttra u avser att *förmedla kunskap* om q .

Visserligen innebär alla förklaringstalakter att någon försöker förmedla kunskap, men det är inte fallet att alla försök att förmedla kunskap är förklaringstalakter. Följande ordväxling kan tjäna som exempel:

A (ställer frågan q): – Jag har glömt ditt telefonnummer. Vad var det nu igen?

S (yttrar u med avsikten att förmedla kunskap om q): – Det är . . .
. . . .

A (vet svaret, och vet också att det är det korrekta svaret): – Jo, nu när du säger det, så minns jag.

Alla villkor i "analysen" är uppfyllda, men det tycks inte som S , i verklig mening, *förklarat* något för A . Han har bara bibringat A kunskap.

Slutsatsen blir att även om jag finner Achinsteins talaktsteori om förklaringar tilltalande, så faller analysen på en otillräcklig och inadekvat definition av vad förståelse innebär. Resonemanget pekar på att den pragmatiska utgångspunkten är riktig, men att det krävs en bättre beskrivning av förståelse och dess roll i förklaring. Jag kommer i det som följer att försöka skissera en teori om förklaring som ger utrymme åt förståelse och stöd för våra intuitioner beträffande förståelse.

6. Koherens och förståelse

Nedan skisseras en teori om förståelse som koherens. Koherensteorin har fördelen att den både visar vilken roll förståelse har i förklaring

och vad förståelse innebär. Jag menar att förklaring och förståelse är essentiellt sammanlänkade. Teorin om förklaring som koherens är samtidigt en teori om förklaring som förståelse¹.

Låt mig börja med att beskriva vad förklaring innebär enligt koherensteorin. För att begripa detta, måste vi titta på *dhöraren* i förklaringstalakten, eller kanske mer exakt *frågeställaren*. Frågeställaren är den som observerat det som är i behov av förklaring, och som formulerat en varför-fråga som förklaringen är ett svar på. Följande kontextuella faktorer måste tas i beaktande: (i) Den förklaringsökande varför-frågan "varför p?" (ii) det kognitiva tillståndet C hos frågeställaren. C inkluderar varför-frågan "varför p?" (iii) Svaret A på varför-frågan, och slutligen (iv) det kognitiva tillståndet C+A som inträder då frågeställaren fått ett svar på sin varför-fråga.

Koherensteorin för förklaring säger helt enkelt att för att A, svaret, ska kunna sägas vara en förklaring, så måste det nya kognitiva tillståndet C+A ha en *högre grad av koherens än C* (Schurz, s 4). Det är ett nödvändigt villkor för A, om A skall accepteras som en förklaring, att A gör C mer koherent. För att detta ska vara möjligt, krävs det att det "finns plats" för A i C eller att C kan assimilera A.

Koherensegenskapen har framförallt diskuterats i teorier för rättfärdigande av kunskap. Dessvärre har det visat sig svårt att ge en exakt beskrivning av den (se Bonjour, s 92ff). Allmänt kan man säga att koherens är en egenskap som tillfaller ett helt system av trosföreställningar. Det är *systemet* som sådant som uppvisar en större eller mindre grad av koherens, inte trosföreställningarna själva. I ett koherent system ger trosföreställningarna varandra stöd, och ju starkare detta stöd är, desto mer koherent är systemet. En diskussion av koherensteorin för förklaring i termer av motexempel kan vara belysande.

1) Antag att det kunde finnas en enda förklaring som hade större koherentistiskt förklaringsvärde än någon annan förklaring vi kunde tänka oss. Vi kan i princip föreställa oss ett kunskapssystem sådant att uppfattningen att "Gud är orsak till allt som händer" skulle öka dess

¹ Denna teori är huvudsakligen inspirerad av Schurz' artikel "Explanation as unification". Artikelnen behandlar i första hand unifikation av vetenskapliga teorier. I artikeln har jag hittat formuleringar som pekar mot en koherensteori liknande den jag försöker beskriva.

koherens, och därmed accepteras som en förklaring. Intuitivt verkar det önskvärt att avfärda den sortens påståenden – de är inte acceptabla som förklaringar.

Jag kan se två lösningar på denna invändning. Det första lösningsalternativet innebär att påståendet ”Gud är orsak till allt som händer” till en början accepteras som en förklaring. Det är viktigt att betona att ett kunskapssystem inte är immunt mot revision. Det försiggår en ständig växelverkan mellan uppfattningarna i kunskapssystemet och omvärlden, där gamla uppfattningar i kunskapssystemet förkastas till förmån för nya erfarenheter. Så småningom måste föreställningen att Gud är orsak till allt förkastas i ljuset av nya erfarenheter. En sådan erfarenhet skulle kunna vara upptäckten att människan själv orsakar och intervenserar i olika händelseförlopp i världen.

Den andra möjligheten är att kunskapssystemet C redan är relativt sofistikerat och innehåller antaganden om orsaksförhållanden i världen. Kanske C är grundat på en materialistisk världsuppfattning. För att en trosföreställning ska ingå i ett koherent system gäller att den logiskt medför och själv kan härledas ur andra trosföreställningar i systemet. Men uppfattningen att Gud är orsak till allt kommer i det aktuella fallet att dels strida mot övriga uppfattningar i systemet, dels vara en trosföreställning som inte kan härledas ur systemets övriga trosföreställningar. Svaret kan inte *assimileras* av det aktuella kunskapssystemet C.

2) Koherensteorin för förklaring tillåter att två rivaliserande svar båda accepteras som förklaring på ett fenomen trots att svaren utesluter varandra.

En liknande invändning är standard när det gäller koherensteorin för rättfärdigande av kunskap. Där gäller invändningen vilket av två lika koherenta system som är rättfärdigat. Vad gäller koherensteorin för kunskap, så kan den försvaras med att man tillåter intryck från omvärlden, att det finns en ständigt pågående växelverkan mellan systemet och världen. Ett kunskapssystem revideras och förbättras i ljuset av nya erfarenheter. Så småningom kommer de nya erfarenheterna att göra att ett kunskapssystem snarare än ett annat överlever. Jag tänker mig att man kan försvara koherensteorin för förklaring på ett liknande sätt. Vid en viss given tidpunkt kan det förefalla som två konkurrerande svar ger lika stor koherens till kunskapssystemet. Genom att tillåta nya intryck från omvärlden och jämföra kunskapssystemens koherens

över tid, så kommer det system som uppvisar störst koherens i *det långa loppet* att vara det som accepteras.

3) Antag att A helt enkelt innebär att ett kunskapssystem *förminskas* genom att man utelämnar flera trosföreställningar. Detta skulle kunna öka C's koherens om de trosföreställningar som utelämnades var problematiska eller kunskapssystemet C inte hade en hög grad av koherens från början. Det vill säga, att just själva utelämnandet av trosföreställningarna måste accepteras som en förklaring.

Det verkar under alla omständigheter vara en rimlig epistemisk strategi att göra sig av med trosföreställningar som minskar koherensen i ett kunskapssystem. Jag tror dock inte att förfarandet behöver accepteras som en förklaring. Det kan inte beredas plats i det pragmatiska ramverket av en talakt, och dessutom, vilken förklaringsökande varför-fråga skulle utelämnandet av en trosföreställning kunna vara svaret på?

Det som framförallt är tilltalande med koherensteorin för förklaring är att den kan visa hur förklaring och förståelse hänger ihop. När subjektet iakttar ett fenomen som han upplever är i behov av förklaring, befinner han sig i ett tillstånd av kognitiv dissonans. Om svaret A är en förklaring, så upphävs denna dissonans. Upplevelsen av att förstå inträffar när A assimileras i C och A "faller på plats".

Jag finner Achinsteins talaktsteori för förklaringar tilltalande också av det skälet att en talakt är sådan att den kan lyckas eller misslyckas. Koherensteorin för förklaring visar *varför* en förklaringstalakt lyckas, men också hur den kan misslyckas. Historien med eleven som inte "förstod" svaret kan begripas som en misslyckad förklaringstalakt. Det som hände var helt enkelt att eleven saknade de kognitiva resurser som krävdes för att A, det svar som stod angivet i facit, skulle kunna assimileras

7. Korrigering av ett kunskapssystem

Den koherentistiska utgångspunkten är att en förklaring gör det som ska förklaras mer bekant genom att öka koherensen hos systemet av trosföreställningar och därigenom ge upphov till förståelse. Men ibland måste ett system av trosföreställningar korrigeras därför att vi omfattar en eller flera felaktiga trosföreställningar. Antag att jag hyser trosföreställningen att om min torrdräkt läcker in vatten, så kommer jag att sjunka till botten. En dag när jag dyker med min torrdräkt

upptäcker jag till min förvåning att den läcker, men att jag trots detta inte sjunker. Någon trosföreställning i mitt kunskapssystem är felaktig och måste korrigeras. Frågan är om en förklaring, i koherentistisk bemärkelse, kan vara korrigerande.

Situationen som uppstått kan beskrivas på två sätt. Det första innebär, att jag tillfälligtvis glömt bort de lagar som styr flytkraften hos föremål i vatten. Egentligen känner jag till dem, men har dem inte aktuella. Någon kan "förklara" varför jag inte sjönk genom att påminna mig om dessa lagar. Närmare reflektion visar, att min trosföreställning att man sjunker om torrdräkten tar in vatten inte var välgrundad. Detta fall är inte av så stort intresse, eftersom det inte är fråga om någon verklig korrigerande av kunskapssystemet.

Det andra fallet är ett genuint fall av korrigerande. Jag är, per hypotes, okunnig om de lagar som styr flytkraften hos föremål i vatten. Frågan är om någon kan sägas ha *förklarat* för mig varför jag inte sjunker genom att nämna de relevanta lagarna. Mitt ursprungliga kunskapssystem C är litet, men relativt koherent. Det innefattar trosföreställningen att lätta föremål flyter, och att tyngre föremål sjunker. En båt flyter så länge den inte tar in vatten, men sjunker annars. Det verkar rimligt att tro att om en torrdräkt tar in vatten så sjunker dykaren på samma sätt. Det faktum att jag *inte* sjönk när min torrdräkt tog in vatten gör att jag måste korrigera min uppfattning om föremåls flytkraft. Det verkar som detta måste ske i två steg.

(i) I ljuset av min nya erfarenhet, så måste jag acceptera att min tro att alla flytande föremål som tar in vatten sjunker är felaktig. Denna uppfattning förkastas från kunskapssystemet C.

(ii) Det nya, mindre kunskapssystemet C' inkluderar den nya erfarenheten att somliga föremål kan flyta trots att de är vattenfyllda. Givet C' kommer de korrekta lagarna om föremåls flytkraft i vatten att ha en "kohererande" effekt, och fungera som en förklaring. C'+A är mer koherent än C'.

Exemplet visar att det är möjligt att göra mindre korrigeringar i ett kunskapssystem, åtminstone så länge det gäller empiriska erfarenheter. Men det är inte fallet att alla trosföreställningar i ett kunskapssystem är reviderbara:

En natt drömmer jag att jag kör och krockar med min bil. Dagen

därpå inträffar just det jag drömde. Givet att jag inte tror på möjligheten att "se" in i framtiden, hur ska jag förklara att jag drömde om olyckan innan den inträffat? Om drömmar står i någon kausal relation till verkligheten alls, verkar det vara till sådant som inträffat dagen före. Personer vi mött, saker vi gjort och sagt har en tendens att dyka upp i våra drömmar. Vanligtvis tror vi att en orsak *föregår* sin verkan i tiden, och av denna anledning verkar det inte möjligt att förklara drömmen med hänvisning till olyckan.

Koherensteorin för förklaring kräver att det finns ett kunskapssystem C som förmår assimilera svaret A. I ett kunskapssystem ingår inte bara empiriska uppfattningar, utan också sådana föreställningar som *konstituerar* våra erfarenheter, som kausalitet, tidens riktning och ett erfalande "jag". Det förefaller mig att en förklaring, i koherentistisk bemärkelse, inte kan utmana eller bestrida dessa uppfattningar. Givet att vårt kunskapssystem är normalt beskaffat skulle vi aldrig kunna *förstå*, i verklig mening, hur olyckan orsakade drömmen. Den förklaringen kan helt enkelt inte assimileras av ett kunskapssystem som grundar sig på en ordinär kausalitetsuppfattning.

Koherensteorin visar, att förklaring och förståelse är sammanlänkade via kunskapssystemets förmåga att assimilera den föreslagna förklaringen. Kopplingen mellan förklaring och förståelse är essentiell. Det är min övertygelse att en noggrann undersökning av kunskapssystemets struktur också kommer att visa vilka acceptansvillkor som gäller för förklaringar.

Litteratur

- ACHINSTEIN, P, 1983, *The nature of explanation*, Oxford University Press.
- BONJOUR, L, 1985. *The structure of empirical knowledge*, Harvard University Press.
- HEMPEL, C G, 1965. Aspects of scientific explanation. [I] *Aspects of scientific explanation and other essays in the philosophy of science*, The Free Press.
- SCHURZ, G, 1999, Explanation as unification, *Synthese*, 120:1.
- SCRIVEN, M, 1962, Explanations, predictions and laws, [I] Feigl och Maxwell (red), *Minnesota studies in the philosophy of science*, vol III, University of Minnesota Press.
- SCRIVEN, M, 1975, Causation as explanation, *Nous*, 9.