

# Den nya utvärderingsvetenskapen

michael scriven

Är utvärderingsforskningen en egen disciplin eller närmare bestämt en »transdisciplin« som statistik och matematik? Kan man tala om att utvärderingen är en modern konst grundad på vetenskap snarare än en renodlad tillämpad vetenskap? I artikeln reflekterar författaren över den nya utvärderingsvetenskapen och dess roll.

## Inledning

En tidigare version av den här artikeln kallades »Utvärdering – en modern konst«, men utvärdering är ju lika mycket vetenskap som konst, och det här temanumret är ett bra tillfälle att betona utvärderingens vetenskapliga sida. Liksom inom socialt arbete är utövarna av utvärdering mycket väl medvetna om i vilken utsträckning tillämpningen kräver en konst grundad på vetenskap snarare än att vara en renodlad tillämpad vetenskap i stil med att bygga broar eller beräkna vilket år Hale-Bopp-kometen kommer att återvända. Det finns tillfällen vid utövandet av dessa konster när man behöver försäkra sig om att man har fast vetenskaplig mark under fötterna; till

---

Michael Scriven is a Professor of Psychology, Claremont Graduate University, California, where he works at the School of Behavioral and Organisational Science.

exempel i inledningsskedet av en ny vetenskap som utvärdering eller i en fullt utvecklad äldre vetenskap som socialt arbete, som då och då behöver ta sig en funderare på och se över sina landvinningar och sin framtid. Därför tänker jag ta det här tillfället i akt att betrakta beskaffenheten av och grunderna för den relativt nya vetenskapen utvärdering. Jag vill börja med att säga att det inte är något nytt med utvärdering – inom och utanför vetenskapsgrenarna – det enda som är nytt är att utvärderingen frigörs för att bilda en egen vetenskap, något som börjar med ett erkännande av dess legitimitet.

## Vetenskapens epistemologiska trygghet

Nuförtiden har det faktiskt blivit inne att angripa legitimiteten hos vetenskapen själv och då i synnerhet föreställningarna om sanning och objektivitet, som är av central

betydelse. Borde det kanske göra oss försiktiga med att försöka få med utvärderingen i vetenskapsgrenarnas skara? Vi får börja med att ta oss an dessa angrepp och den epistemologiska skepsis som de grundar sig på, inte bara för att om utvärdering skall bli betraktad som en vetenskap, så skulle den få samma behandling, utan av ett mera speciellt skäl. Skeptikernas särskilda angreppspunkter – begreppens objektivitet och partiskhet – är i själva verket avgörande inslag inom utvärderingsområdet. Om angreppen är vederhäftiga skulle dessa centrala begrepp inte överleva och utvärderingen skulle gå under tillsammans med dem.

Man träffar på denna skepsis vid diskussioner om teorier i socialt arbete på samma sätt som vid diskussioner om utvärderingsteori. Angreppen görs ofta av dekonstruktionister som säger att de har upptäckt att objektivitet är en myt eller att vetenskap inte är ett sökande efter sanning utan ett maktverktyg som överklassen använder sig av. Det är faktiskt ganska roligt för någon som har tillbringat över tjugofem år av sitt liv som yrkesfilosof att se så många människor med fina akademiska meriter i andra ämnen gå i samma fälla som varenda filosofistuderande gör under sitt första år. (Å andra sidan finns det ganska många yrkesfilosofer som har gått i samma fälla under hela sitt yrkesliv.) Men i vilket fall som helst kan man avfärda skeptikernas angrepp på vetenskap och objektivitet eftersom sist och slutligen det sunda förnuftet inte låter sig förlöjligas. Det kan visserligen vara modernt att driva med vetenskapen, men det finns ingen gränslinje mellan vetenskap och sunt förnuft. Eller som Einstein uttryckte det i *Physics and Reality*: »Hela vetenskapen är

ingenting annat än en förbättring av vanligt tänkande.« Vi har inte för avsikt att ge upp våra välgrundade meningar om möblerna i det här rummet, förekomsten av idoga och lata människor, nyttan av värktabletter och Internets objektiva realitet.

Naturligtvis skiftar det som räknas som vetenskap och sunt förnuft en smula efterhand som vi gör nya upptäckter och upptäcker misstag. Nuförtiden är det förnuftigt att inte dela vattenflaska med någon som har en kraftig förkylning, men förr i tiden skulle någon som vägrade att göra detta ha ansetts irrationell eller vara behäftad med sociala fördomar. Så vetenskapen förändras och sunt förnuft förändras – men nästan alla förändringarna sker stegvis och är noga genomtänkta. Vår fantasi fångas av sporadiska radikala begreppsmässiga förändringar, paradigmskiftena – men Galileo och Einstein gjorde bara mycket små detaljförändringar i data de tolkade. Vi behöver inte vara rädda för att vår tro på att bakterier överför sjukdomar kommer att visa sig vara kapitalistisk hjärntvätt. Även om det i princip är möjligt och inträffar emellanåt – och dessutom har synnerligen stort nyhetsvärde – är sådana händelser så pass sällsynta att vi inte behöver betrakta vår tilltro som otillbörlig. Det finns objektiva sanningar och det finns partiskhet, och de kan båda säkert identifieras och särskiljas utan det naiva antagandet att de alltid skulle skilja sig på samma punkt och att det vi är mest övertygade om aldrig skulle vara fel. Den vurm som förespråkar det motsatta – den aktuella vågen av skepsis – kommer att gå över. Man skulle lika gärna kunna sluta upp att använda räkning när man beräknar sin inkomstskatt eftersom

matematikens grunder fortfarande ifrågasätts. Den typen av diskussioner hör hemma runt filosofernas kaffebord – den är ovidkommande i den riktiga världen.

För att ta ett exempel som ligger närmare området samhällsvetenskap och utvärdering, så vet alla mycket väl att man kan förhålla sig ganska objektiv till vissa saker, som till exempel förekomsten av en långvarig diskriminering av kvinnor i moderna västerländska samhällen. Att säga att vi kan förhålla oss objektiva till det är inte det samma som att säga att vi bör vara oengagerade eller oberörda av det. Det säger bara att vi kan vara säkra på det, vi kan veta att det är sant och inte en projicering av våra känslor. Det är helt enkelt ett faktum, på samma sätt som det är ett faktum att Hitler orsakade miljontals oskyldiga människors död. Att vi känner till det visar inte att vi inte kan förhålla oss objektiva till det. Utvärderingar som den här kan vara objektiva – det var bara den positivistiska neurosen som hävdade att det var omöjligt.

En del människor kan inte förhålla sig objektiva till vissa av de här frågorna eftersom de blir för djupt berörda av dem. Det innebär att tillförlitligheten i slutsatserna de har dragit om dem blir lidande och att känslorna i viss mån tar överhanden över deras förnuft. Vi vet att det är mycket svårt för alla att förhålla sig objektiva till vissa saker, som till exempel de goda egenskaperna hos våra egna barn eller karaktären hos den äkta hälft som just har begärt skilsmässa. Vi är ganska bra på att särskilja dessa fall från mera generella fall. Det är helt enkelt tråkigt att hålla på att påminna oss själva om att vi ibland har fel när vi tror att vi är objektiva

och helt enkelt befängda om vi därav drar slutsatsen att det inte finns någon objektivitet. Till och med den till synes mest felfria bil kan gå sönder utan minsta föregående varning. Av det följer inte att alla bilar är opålitliga eller att vi aldrig skulle ha rätt om vi påstår detta om en viss bil.

Det vi måste få från de radikala epistemologerna är specifika exempel på partiskhet som vi hittills skulle ha förbisett. Om de är övertygade om att så kallad objektivitet är rena skämtet borde de kunna hitta några exempel som hittills har undgått oss. I själva verket finns det säkert sådana, och det är troligt att de har varit till skada för någon. Vi borde vara beredda och villiga att börja arbeta med att kontrollera bevisen och sedan identifiera och gottgöra offren och inte ägna oss åt filosofiskt gnäll om den allestädes närvarande omöjligheten att det skulle finnas säker kunskap eller objektivitet. Är den säker så är den säker. Termen som den används inom vetenskapen och det sunda förnuftet innebär inte den absoluta omöjligheten för misstag, bara den extremt låga sannolikheten för misstag, en sannolikhet som är så låg att man i praktiken inte behöver ta hänsyn till den.

Mer än så ger vi alltså inte för det allmänna angreppet på vetenskap och objektivitet. Det är helt enkelt ett filosofiskt misstag och bör inte tas på större allvar än ett angrepp på matematikens grunder – men vi bör alltid vara villiga att lyssna om det dyker upp något specifikt, något av praktisk vikt.

## Utvärderingens historiska osäkerhet

Filosofiska fördomar som exemplet med

radikal skepsis som diskuterades ovan har ofta haft större inverkan på vetenskapen än man insett vid tillfället. Utvärderingens historia är ett sådant fall. Tabubeläggningen mot den infördes av logiska positivisterna – som ansåg att utvärderingsmässiga påståenden inte ens kunde tas som förslag, än mindre kunde vara empiriskt sanna eller falska – och har dröjt sig kvar i fyrtio år efter det att positivismen förlorat alla anhängare av betydelse. Positivism är i dag inget annat än en fågelskrämma som radikala epistemologer använder som måltavla, men positivismens åsikt om utvärdering genomsyrar fortfarande de flesta vetenskaperna. Ta till exempel de naturvetenskapliga läroplansreformernas historia efter Sputnik. Samtliga inledde med att förteckna naturvetenskapens viktigaste logiska begrepp och förfaranden, som de buntade ihop till ett slags föreställning om vetenskaplig metod. Detta bildar den grund på vilken det riktiga, faktiska innehållet sedan har byggts. De vanliga begreppen består av bland annat observation, beskrivning, mätning, klassificering, generalisering, förklaring, hypotesprövning och så vidare. Dessa förteckningar var aldrig särskilt imponerande och förefaller bara bli sämre om man skall döma av en som just införlivats i den nya amerikanska läroplanen för de naturvetenskapliga ämnena. Men de har en sak gemensam: de utesluter det allra viktigaste begreppet i den vetenskapliga repertoaren, det enda som är oundgängligt inom alla vetenskaper. De utesluter utvärdering.

Om man öppnar en vetenskaplig tidskrift inom ett slumpmässigt valt område och stannar vid en slumpmässigt vald artikel

är det troligt att man finner en presentation som börjar med en genomgång av tidigare genomförda arbeten. Denna genomgång är ytterst selektiv och urvalet görs på grundval av kvaliteten på eller betydelsen av tidigare forskning. Detta kommer förmodligen att efterföljas av en diskussion om planen för den forskning som skall redovisas: skäl kommer att anges till varför man ansåg att den valda planen efter förhållandena var optimal. Det är en utvärderingsmässig slutsats grundad på resonemang, det vill säga just det som sades vara en självmotsägelse. Sedan framläggs resultaten av undersökningen, och rönens kvalitet och betydelse kommer att diskuteras – ytterligare två utvärderingsinsatser. Utvärdering är i själva verket den vetenskapliga undersökningens probersten. Det finns vetenskaper som ägnar sig mycket litet åt klassificering och andra som inte gör mycket annat. Det finns vetenskaper som redovisar enstaka händelser, andra som bara går in för generella principer. Men det finns inga vetenskaper som underlåter att utvärdera utformningen eller resultatet av sitt arbete eller andra forskares arbeten. I synnerhet finns det inga vetenskaper som inte kan skilja mellan bra och dåligt vetenskapligt arbete och vid gränfall skilja det bästa arbetet från pseudovetenskap, till exempel astronomi från astrologi – och den skillnaden är helt och hållet grundad på utvärdering.

Se på resten av vetenskapsmännens vetenskapliga liv, just de vetenskapsmän som fortfarande inte vill förläna utvärderingen legitimitet som vetenskaplig metod. De undervisar, och de utvärderar varje elev de ger undervisning. Om dessa utvärderingar ifrågasätts försvarar de dem

med hjälp av sakförhållanden och resonemang. Det vill säga, de anser att dessa utvärderingar är objektiva, vilket är något helt annat än att anse att de är ytterst exakta. De granskar arbeten som lämnas in till tidskrifter och ägnar sig därvidlag åt ren utvärdering. De bedömer ansökningar om forskningsmedel – det handlar om utvärdering. De tar sökande till en tjänst på sin institution eller forskningsgrupp under övervägande, vilket är en process helt och hållet grundad på utvärdering. Den egentliga vetenskapliga processen består inte bara till stor och avgörande del av utvärdering, utan varje vetenskaplig disciplins sociala verksamhet hålls samman av ett nät av utvärderingar. I förbigående kan nämnas att alla vetenskapsmän inte bara ägnar sig åt och försvarar utvärdering i sitt privatliv utan också tar råd från sådana som är experter på det, till exempel konsument-tidskrifter eller utredningar om mediciner som görs av statliga organ.

Kan man då inte säga att uteslutningen av utvärdering, tabubeläggningen av utvärdering, som kännetecknar vetenskapens historia under det senaste århundradet i sanning är ett märkvärdigt fenomen? Hur kunde det hända? Denna uteslutning tror jag har mycket djupa rötter. Även om den uppenbarligen i hög grad stöddes och främjades av positivisterna dröjde den kvar så länge efter den offentliga avrättningen av den filosofin att det måste finnas fler förklaringar till den. Jag tvivlar på att doktrinen om värderingsfrihet kunde ta ett sådant grepp och hålla kvar det utan att åstadkomma en genklang i vetenskapsmännens själar. Vad kan den klangen vara för något?

Enligt min åsikt är det ingen händelse att

den religiösa traditionen, i likhet med den vetenskapliga, har samma bisarra, nästan schizofrena inställning till utvärdering. »Dömen icke, så skola I icke bliva dömda« säger Bibeln och ändå är dömandet själva poängen med varenda liknelse och grundtanken hos varenda profet. Är det inte konstigt att den förbjudna frukten i Första mosebok satt på kunskapens träd på gott och ont? Bibeln säger att Gud förbjöd Adam och Eva att äta frukten eftersom den skulle göra dem lika gudarna: gudomen förbehöll sig själv rätten att utvärdera. Och ändå tvekade inte Gud att straffa dem som var olydiga mot honom genom att äta frukten från det trädet trots att de inte kunde veta att det var syndigt att göra det. Under årtusendena som har gått sedan dess har prästerna inom många religioner fortsatt att förbehålla sig rätten till den särskilda kunskapen om vad som är gott och ont. Så den religiöse lekmannen är förbjuden att utvärdera eller att döma, på samma sätt som den vanlige vetenskapsmannen förbjöds att behandla utvärdering som en respektabel kognitiv process. Trots det ägnar sig alla lekmän ständigt åt det, precis som alla vetenskapsmän.

Jag tror att utvärderingen framkallar den här schizofrena responsen eftersom den söker sig ända ned till ångestens och rädslans primitiva källor. I den reviderade standardnomenklaturen för psykiatrisk klassificering finns det numera en åkomma som kallas för utvärderingsångest. Jag tycker att det här bekräftar den ångest jag talar om. Vi är med rätta nervösa inför en sluttentamen. Vi är neurotiskt nervösa när vi hävdar att detta visar att tentamina är olämpliga. När jag hör utbildningsteoretiker hävda att lärare aldrig någonsin borde betyg-

sätta studenternas arbeten, och jag funderar över detta totalt förtryckta förslag, vet jag att jag måste söka djupa orsaker till det och till det allmänna fenomen som det är ett exempel på – de ofta förekommande angreppen på något som helt tydligt är värdefull utvärdering. Jag minns det raseri med vilket många fakultetsmedlemmar mötte de första förslagen om att studenterna skulle tillåtas eller kanske till och med uppmuntras att utvärdera dem. Det var ungefär samma slags reaktion som skulle ha mött en hädelse. »Du skall inga andra gudar hava jämte mig« är det man närmast kommer att tänka på – de nya utvärderarna inkräktar på de gamla reviren. Vetenskaparen, en utvärderare in i ryggmärken inom sitt eget område, blir förbittrad när makten övergår till andra och uppfinner en absurd doktrin för att rättfärdiga tillbakavisandet. Doktrinen är inte absurd bara för att den är rena självmordet, vilket ett ögonblicks allvarlig begrundan av deras eget arbete snart uppenbarar, utan den är absurd på grund av att var och en som följer ens den enklaste form av vetenskaplig metod skulle se att den är falsk.

Om vi bara kan se klart på det är det här en scen från ett stort drama i tänkandets historia, och dramat är den långsamma processen att göra vetenskapen till en självrefererande verksamhet. I det här exemplet ser man hur långt vetenskapen har varit från att studera sin egen process innan de lagt fram dogmer om hur deras egen process är beskaffad. Kort sagt visar fallet hur tydligt som helst att vetenskapen är långt ifrån att vara en vetenskaplig verksamhet. Nu talar jag i och för sig i egen sak när det gäller det här angreppet. Jag hävdar att försöket att tabubelägga utvärderingen var

otillbörligt och att vetenskapsgrenen hölls tillbaka under större delen av århundradet på grund av denna blinda fördomsfullhet.

Ibland förefaller det mig som om böjningen av det relevanta verbet, åtminstone i vetenskapsmännens ögon, skulle låta ungefär så här: »Jag utvärderar, du uttrycker en åsikt, och de är helt enkelt känslomässiga.« De ville inte ge utvärderingen legitimitet eftersom de inte ville släppa vargen lös. Så länge den är fastbunden kan vi använda den. Om den släpps loss kan den bita oss. Med tanke på att ett tillbakavisande av utvärderingens legitimitet vore att skära hjärtat ur hela vetenskapen och allt praktiskt liv var det en desperat åtgärd. Med tanke på att vi inte själva har för avsikt att ge upp den, bara förmena den sin legitimitet och därmed andras användning av den, är detta verkligen en ovärdig inställning. Det är en inställning som intellektuella borde skämmas över att ha. Jag kan inte finna någon annan förklaring till en sådan inställning än en rädsla för att själv hamna i händerna på andra, en rädsla för att en granskning skulle avslöja ens egna tillkortakommanden.

I det mänskliga tänkandets historia är tabubeläggningen av utvärderingen onekligen ett av de mest fascinerande, mystiska och på många sätt tragiska teman som finns. Det är tragiskt, inte bara patetiskt, eftersom en av konsekvenserna av tabut, som går raka spåret tillbaka till Max Webers tidiga version av doktrinen om värderingsfrihet, var att samhällsvetenskaperna slutade befatta sig med mänskliga behov. De var stolta över att det inte förekom någon utvärdering i deras arbete och underlät därmed att inrikta sitt arbete mot de ställen där en vetenskaplig inställning

till det sociala sönderfallet skulle ha kunnat leda till stora förbättringar. De letade inte efter behov, de lade inte märke till desperata behov, och i vilket fall som helst ansåg de inte att de hade med det att skaffa. Om ni tvivlar på det här, läs recensionerna av de två första böckerna om fattigdom i USA, som kom ut långt efter det att vi hade passerat halvsekelskiftet. De är väldokumenterade och välskrivna – och mottogs med stort förakt med hänvisning till att det inte kunde finnas någon fattigdom att tala om i USA. Den metodologiska myten hade förblindat vetenskapens iakttagande ögon. När man läser ordväxlingarna mellan företrädarna för socialt arbete och samhällsvetenskap på Internet eller i tidsskrifter kan man än i dag känna spänningen mellan vetenskaparmodellen och hjälparmodellen, en spänning som rätt och slätt består av rök från den positivistiska soptippen.

Tragedin låg i att samhällsvetenskapens fokusering på mänskliga behov, som det skedde inom det medicinska området, i stor utsträckning inom socialt arbete och annorstädes, och som har börjat dyka upp under de senaste årtiondena, kunde ha räddat otaliga liv under de långa mörka tiderna under vårt eget nittonhundratals. Så det här är inte blott och bart ett intellektuellt gruff. Det stod mycket på spel i den här kampen och det är en kamp där de vetenskapliga krafterna skämde ut sig. Men det fanns förstås en och annan framstående person som gick mot strömmen – Gunnar Myrdal, till exempel. Men de kunde inte hejda strömmen i USA och andra engelsktalande länder.

Ni tycker kanske att ni har hört allt det här förut, många gånger, och att poängen bara är kommentaren om de hycklande

vetenskapsmännen som utövade utvärdering men förmenade den dess legitimitet. Faktum är att det bara är en replik man kan glida över. Visst hycklade de, men det finns en mycket allvarigare fråga att ställa: Hade de rätt? När man upptäcker att en predikant är äktenskapsbrytare, lögnare och tjuv så klandrar man hans hyckleri, men man drar inte slutsatsen att han hade fel. Det är en helt annan sak. Och likadant är det i det här fallet. Vi har avslöjat hyckleriet, men vad skall vi göra med de argument som användes för att stödja ståndpunkten? För att uttrycka sig rakt på sak: Skall man utgå från att de hade fel bara för att de var hycklare?

Det är uppenbart att vi inte bara kan utgå från det. Vi måste ha starka argument för det felaktiga i ståndpunkten. Ett av dessa argument finns i ett motsägelsebevis som är underförstått i det jag har sagt: om utvärdering är otillåten eller ovidkommande så är det mesta av vetenskapen och det mesta praktiska beslutsfattandet också otillåtet. Om man tror, vilket förefaller rimligt och som jag har hävdat ovan, att detta är en absurd inställning, så måste vi avfärda doktrinen om värderingsfrihet. Men vi måste också titta närmare på de uttryckliga argument som har framförts för att tabubelägga utvärderingen. Jag skall strax åtminstone antyda varför de var missriktade. Först när de kan vederläggas kan vi känna oss fria från tabu och utvärdering kan försöka uppnå vetenskaplig status.

Sätt det jag har sagt i motsats till den åsikt som nästan är en del av den intellektuella kulturen i dag, nämligen åsikten att den värderingsfria inställningen avslöjades som humbug av krafterna inom den politiska liberalismen och radikalismen.

De och alla andra som blivit övertalade av dem anser att den radikala vänstern var riddaren i skinande rustning som dräpte den värderingsfria draken. Det de gjorde var i själva verket fullständigt ovidkommande vad doktrinen om värderingsfrihet beträffar. Det de trodde sig göra var att visa att samhällsvetenskaperna var ett exempel på kejsarens nya kläder, eftersom just de samhällsvetenskapare som avfärdade utvärderingens berättigande inom vetenskapen i själva verket tog ställning i en rad frågor, ställningar som hade hämtats från deras yttre politiska och personliga värderingar. Den typen av argument visar precis samma totala okunnighet om frågan som ledande vetenskapsmän själva uppvisade. Doktrinen om värderingsfrihet hävdade aldrig någonsin att vetenskapsmän inte hade både politiska och personliga värderingar eller att vetenskapsmännen inte skulle använda dem för att bestämma vilken vetenskap de skulle ägna sig åt eller om de skulle ägna sig åt vetenskap. Doktrinen om värderingsfrihet var något helt annat. Det var en doktrin om att utvärderingsmässiga påståenden inte hade någon vetenskaplig status, det vill säga inte kunde bekräftas med hänvisning till logik eller vetenskapliga metoder. Det hävdades inte att de inte hade någon inverkan på vetenskapsmännen eller ens att de inte skulle ha någon inverkan på vetenskapsmännens beslut om huruvida de skulle ägna sig åt vetenskap eller vilken vetenskap de skulle ägna sig åt eller hur de skulle använda sig av den. Radikalerna dräpte i sanning en drake – den var en mytisk fiende.

Bland de argument som framfördes ovan om hyckleriet i den vetenskapliga inställningen finns det däremot ingen hän-

visning alls till de värderingar som driver fram sådana beslut. Det var alltid klart och tydligt och förnekades aldrig för ett ögonblick av någon företrädare av värderingsfri vetenskap att det är yttre värderingar som driver fram dessa personliga beslut av vetenskapsmännen. Det som framförs i argumentet ovan är att hyckleriet härrörde från den faktiska användningen av utvärdering i själva hjärtat – i teorin och praktiken – i den logiska kärnan – hos varje vetenskap. Slår man ihop denna genomgripande och uppenbara roll med avfärdandet av dess berättigande får man hyckleri som grundats på självbedrägeri. Det som den radikala vänstern visade var att en doktrin som aldrig någon var anhängare till var oförenlig med en praxis som alla anslöt sig till. Det är inte någon särskilt upphetsande slutsats.

Det enda seriösa sättet att vederlägga doktrinen om värderingsfrihet är att ta itu med en undersökning av vetenskapens struktur, inte dess sociala omgivning, och visa att vetenskap i allt väsentligt är utvärderande. Och det var det jag gjorde, litet löst, när jag genomförde tankeexperimentet att slå upp en vetenskaplig tidsskrift och noggrannare betrakta de kognitiva processer som innehållet innefattar. Det ger oss de motsägelsebevis vi behöver. Det ger oss också slutsatsen om hyckleri, om man inser att vilka vetenskapsmän som helst hade kunnat göra det vi diskuterade och upptäcka att det de gjorde var oförenligt med det som vetenskapsmän säger om vetenskapens beskaffenhet. Om man fastställer slutsatsen om hyckleri på detta sätt är den välgrundad snarare än ett angrepp på en fågelskrämma.

Men det innebär att vi sitter där med klara argument för att det är omöjligt att



vetenskapligt rättfärdiga värderande slutsatser, vilket naturligtvis tyder på att det är omöjligt att ha en utvärderingsvetenskap.

### Argumenten för att utvärdering är en omöjlighet

Det finns två sådana som är nära besläktade, och de är mycket enkla. Det första var argumentet att vetenskapen bara ägnar sig åt, eller bara borde ägna sig åt, att beskriva världen som den är, det vill säga bara ägna sig åt beskrivningar och inte tala om hur den borde vara, det vill säga rekommendationer. Den andra är en version av Humes argument, att man utifrån deskriptiva premisser – vetenskapliga data eller observationer – inte kan dra giltiga utvärderingsmässiga slutsatser. Det första argumentet är uppenbart felaktigt om man bara bryr sig om att titta närmare på bevisen. Vetenskapen ägnar sig faktiskt åt hur världen ser ut, men det kräver att man också ägnar sig åt hur den kan beskrivas på bästa sätt, med de bästa förklaringarna, de bästa sätten att undersöka den och med de bästa verktygen för sådana undersökningar. Med andra ord måste vetenskapen vara utvärderande för att kunna vara deskriptiv.

När man väl fått klart för sig att vetenskapens logik kräver användning av utvärderingens logik, legitimerar det erkännandet av det sätt på vilket tillämpad vetenskap alltid och vederbörligen utfärdar rekommendationer om hur saker bör göras, till exempel hur temperaturen i en järnsmältugn skall kontrolleras, hur en bro bör byggas, hur en stor damm skall konstrueras för att komma upp i den bestämda kapaciteten, etc. Så vetenskapen har inte bara med instrumen-

tell utvärdering att göra utan även med tillämpad utvärdering. Faktum är ju att när konsumenttidskrifter rekommenderar den bästa datorn för vissa ändamål ägnar de sig åt produktutvärdering, ett slags tillämpad vetenskap. Och dessa tillämpade matematiska och byggnadstekniska exempel skiljer sig inte från exempel på samhällsbyggande, till exempel tillämpad psykologi eller tillämpad sociologi, utom när det gäller parametrarnas komplexitet. Det är med andra ord ingen principiell skillnad, inget brott mot logiska begränsningar.

Det andra argumentet, att man inte kan dra utvärderingsmässiga slutsatser av faktiska premisser, verkade ganska förnuftigt på den tiden när Hume framförde det, när slutsatser på det hela taget betydde deduktiva slutsatser. Men nuförtiden förstår man härledningens beskaffenhet bättre. Man förstår att slutledningen är en särskild typ av härledning, en som gör sina slutsatser logiskt nödvändiga utifrån de givna premisserna, men man förstår också att det finns två andra typer av härledning som bara gör sina slutsatser troliga: statistisk härledning, den mest erkända, och beviskraftig härledning, den typ som styr de flesta juridiska härledningar och sådana som grundar sig på sunt förnuft. (Ett specialfall av beviskraftig härledning är luddig härledning, som är mycket mer berömd än sitt upphov.) Beviskraftig härledning är härledning av slutsatser som vid första påseendet är troliga, med de förutsättningar som givits. Detta innebär inte att man kan räkna ut en siffermässig sannolikhet för slutsatserna, vilket är möjligt för statistisk härledning, utan att de flesta bevisen talar för dem, som man brukar säga, eller att de troligtvis

är sanna, under i övrigt lika förhållanden, som man ofta säger. Det som ofta talar för den beviskraftiga härledningen är att den förefaller ge de bästa förklaringarna till premisserna och att dessa premisser förefaller vara av ett sådant slag som det på det hela taget alltid finns förklaringar till (med andra ord är determinism tillämplig).

Beviskraftig härledning är just den typen som leder oss från premisser om behov och sätt att tillfredsställa dessa behov och tillgängliga resurser, till slutsatser om vad som bör göras, det vill säga slutsatserna om programutvärderingar. Det är också den typen som leder oss från kunskap om den täckning som en högskolekurs har plus information om bakgrundkunskaperna hos studenterna som läser kursen, tiden som är tillgänglig för kursen och de angivna nivåerna, till slutsatser om betyg för deras specialarbeten, det vill säga slutsatser grundade på prestationsutvärderingar. Och samma sak gäller personalutvärderingar. Här har vi med andra ord utvärderingens grundläggande logik, underförstådd i den praktiska utvärderingsprocessen inom de många delområden som utvecklade sina egna procedurer på grund av det praktiska behovet av att göra det, ogenerade av vetenskapens underlåtenhet att erkänna deras berättigande. Och det humeska argumentet, i likhet med radikalernas angrepp på doktrinen om värderingsfrihet, får därmed betraktas enbart som ett tjusigt angrepp på en ståndpunkt som ingen intar.

Bergen av invändningar mot det berättigade i en utvärderingsvetenskap kan därför bestigas, och det är dags för oss att inrätta vårt basläger. Vad är egentligen utvärdering? Är det numera en disciplin snarare än

en rad underförstådda tillämpningar, och har den disciplinen bidragit med något som främjar vetenskapens framsteg?

## Utvärderingens beskaffenhet och logiska grunder

Som vi har sett är utvärdering inte en svår-tillgänglig process begränsad till vetenskap eller etik. Småbarn lär sig mycket tidigt att i sin omgivning skilja mellan det de gillar och det de ogillar. Det rör sig här inte om utvärdering utan om att lägga grunden till utvärdering. Efterhand som de blir äldre lär de sig att förknippa orsakerna till att resultaten blivit behagliga med själva välbehaget, och har därmed börjat identifiera tillstånd eller saker som inte omedelbart står till buds men som eventuellt, och ofta indirekt, är värdefulla för dem. Fortsatt utbildning och mognad utsträcker omfattningen av de överväganden som innefattas i utvärdering till andra värden än välbehag och ger en färdighet i utvärdering vad avser endast dessa andra värden. På så sätt lär vi oss att utvärdera olika slags saker i enlighet med normen som gäller för just den saken, inklusive konstakning, motorecyklar och hyresfastigheter.

Och så är vi där till slut och har blivit förtroliga med det naturliga och enkla utövandet av vägning och bedömning av styrkan och stabiliteten i de många överväganden som måste göras i samband med fastställandet av förtjänster och värden, det vill säga själva utvärderingsprocessen. Det är faktiskt en av de allra vanligaste och viktigaste kognitiva processer som vi är i stånd att utföra. Men det är en som framkallar blandade känslor i oss eftersom

vi växte upp under sådana förhållanden att våra handlingar och vi själva ofta utvärderades, inte sällan till vår nackdel, och ofta med bestraffningar förknippade med dessa utvärderingar. Så vi slits mellan tillbakavisande och godtagande, mellan ogillande och erkännande av nyttan av den. Till och med inom de vetenskapsgrenar där motsättningarna ofta inte är lösta och resultaten motsägelsefulla.

Fundera på hur många gånger någon som undervisar i en introduktionskurs i samhällsvetenskap har sagt: »Det är inte lämpligt att vetenskapsmän gör värderingar« utan att de tänker på att de just yttrat en självmotsägelse. Det här är inget område där vetenskapsmän är immuna mot logiska misstag av enklaste slag.

För att betrakta den mer positiva sidan av vår uppgift, låt oss se om vi kan ställa upp utvärderingsvetenskapens grundläggande grundsatser. Vi måste börja med att skilja mellan två slags vetenskap, som vi kan kalla grunddiscipliner och transdiscipliner. Gränsen mellan dem är inte särskilt tydlig, men den generella skillnaden är att grunddiscipliner i första hand är inriktade på empiriska undersökningar av fysiska, sociala, beteendemässiga eller psykologiska skeenden, medan transdiscipliner är: 1) analytiska snarare än empiriska, 2) verktygsdiscipliner som underlättar de undersökningar och utredningar som genomförs av grunddisciplinerna snarare än att utföra arbetet själv, 3) också självständiga studieområden, vilket ger dem rätt att kallas egna discipliner. De främsta transdisciplinerna är nog statistik, sannolikhetsteori, mätvetenskap, logik och utvärdering. Varje disciplin har ett eget studiefält, i vissa fall

en hel akademisk institution, men varje fält är formellt och inte empiriskt. Varje transdisciplin är också ett viktigt verktyg inom ett halvduzin grundfält. Och var och en får mycket av sin utvecklingskraft från försöken att lösa problem som uppstår i dessa andra discipliner. Experimentell design är ett annat exempel, liksom beslutsteori, och möjligen etik.

Det finns en viktig skillnad i transdisciplinernas historiska utveckling. Logik och utvärdering utvecklades till en sofistikerad – men för det mesta underförstådd – nivå samtidigt som grunddisciplinerna mögnade. Logiken, som började underförstått, kom fram ganska tidigt som en egen disciplin för två tusen år sedan. Statistik och sannolikhetsteori kom senare, som biprodukter av vissa vetenskapliga och matematiska verksamheter, men utvecklade sedan snabbt sina egna teorem och esoteriska begrepp liksom otaliga värdefulla tjänster åt grunddisciplinerna, beträffande allt från kvantmekanik till demografi. Utvärdering och mätvetenskap är fortfarande i hög grad underförstådda i vanligt vetenskapsutövande och har först på senare tid börjat utvecklas självständigt, och då för att återföra grunddisciplinerna med nyttiga verktyg. Mätvetenskapen gjorde förstås det här tidigare än utvärderingen. Experimentell design var litet före dessa två grupper. Den delade sig i och med Fishers och andras arbeten, och har bevisat sitt värde, men det är klart att mycket av experimentell design inom de fysiska vetenskaperna har utvecklats på egen hand och inte haft någon större glädje av den självständiga disciplinens arbete. Beslutsteori är också ett mellanliggande fall. Vad den här typen av skillnader

beträffar, vill jag hävda att vi nu kan peka på ett flertal ytterst viktiga bidrag från utvärderingens sida till vanlig vetenskapsutövning, varav jag nämner några nedan. Detta är en synnerligen viktig milstolpe eftersom det är mycket svårt att avfärda ett ämne som har blivit eller bidragit med ett viktigt verktyg till ens egen och andra grunddiscipliner.

Om man har i åtanke det sätt på vilket S. S. Stevens inrättade det mätvetenskapliga fältet genom att först identifiera kärnbegreppen och sedan utveckla någon teori om dem, kommer man att se likheter med det sätt på vilket jag inrättat utvärderingen. I det följande gör jag en allmän översikt av området.

1. Utvärdering är en process där enheters värde, förtjänst eller betydelse bestäms; utvärderingar blir resultatet av denna process. Utvärdering kan vara extern eller intern eller en blandning härav, och den kan vara kvantitativ eller kvalitativ eller en blandning härav. Den är starkt men inte alltid skarpt avgränsad från förklaring.
2. Det finns bara fyra grundläggande utvärderingsgrunder, varav den ena – en kognat – måste uppträda i alla utvärderingsmässiga slutsatser (eller ges av de slutsatserna utifrån sammanhanget). Utvärderingsgrunderna är gradering, rangordning, poängsättning och fördelning. Var och en av dessa måste understödjas av ett lämpligt och separat utredningsförfarande. Det finns bara två samband – poängsättning medför rangordning, och fördelning medför, men är inte begränsad till, en blandning av gradering och rangordning. Rekommendationer är en del av den tillämpliga utvärderingsvokabulären och består av kontextberoende slutsatser baserade på någon kombination av de grundläggande utvärderingsgrunderna.
3. Utvärdering är en självständig disciplin som bestäms av 1 och 2, och den utvecklar sina egna modeller, teorier och förfaranden. Dessa täcker frågor som att anföra logiken i de grundläggande utvärderingsgrunderna, kartlägga sambandet mellan utvärdering och förklaring, planera, beskriva, klassificera, göra förutsägelser och ge rekommendationer. Vidare att skapa och fastställa metoder för interna synteser (att integrera delutvärderingar i en övergripande utvärdering) och externa synteser (att integrera utvärderingar som gjorts av flera utvärderare i en övergripande utvärdering – som också är ett slags metaanalys). Utvärdering använder också, som verktyg, många metoder och tekniker från andra discipliner.
4. Utvärdering är också en transdisciplin, en disciplin som förser andra discipliner med verktyg. Vissa av dessa verktyg har utvecklats som en viktig del av disciplinerna efterhand som de kommit fram ur sina förvetenskapliga ursprung. Således är intradisciplinär utvärdering – utvärdering av hypoteser, data, resultat, utredningsmässiga förfaranden (experimentell design, till exempel),

instrument, tidigare arbeten – en viktig del av alla andra discipliner, mycket viktigare än de tilläggskomponenter som härletts ur de flesta transdisciplinerna. Det är i själva verket det nyckelverktyg som skiljer dessa discipliner från irrationell eftergivenhet och ren spekulat. Arbetet inom den självständiga disciplinen utvärdering kan nu påvisa många sätt på vilket intradisciplinär utvärdering kraftigt kan förbättras (till exempel utvärdering av förslag, kandidater, studerande, vetenskapliga instrument, programvaror och maskinvaror och genom att förbättra den interna syntesen).

5. Utvärdering har massor av namngivna praktiska områden, där kompetent utvärdering har pågått i många år, ibland i århundraden eller årtusenden. Bland dessa finns produkt-, program-, prestations-, personal-, förslags- och policyutvärderingar, liksom institutionell utvärdering, läroplansutvärdering, litteraturkritik, kvalitetssäkring inom industrin, etiska bedömningar, juridiska granskningar, metautvärdering (utvärdering av utvärdering), liksom de intradisciplinära utvärderingarna som nämndes ovan. Den grundläggande logiken i utvärdering är precis samma inom vart och ett av dessa områden, och – vilket tål att understrykas – många av de professionella metoder som utvecklats inom ett av dem fungerar inom många av de andra. I dagens läge finns det gott om utrymme för förbättringar av utvär-

deringen inom alla dessa områden. Medan en del av det helt enkelt kan åstadkommas genom överflyttning av metoder eller begrepp som redan slagit rot inom andra områden (till exempel skillnaden mellan primära (kriteriegrundande) indikatorer och – vanligen otillåtna – sekundära (växelverkande) indikatorer) återstår mycket att utveckla från de grundläggande satserna (som framgår av punkt 3 ovan) och genom att tillämpa utvärderingens grundläggande logik, som är – liksom inom juridiken och det mesta av vårt praktiska liv – beviskraftig logik (det vill säga härledning av slutsatser som gjorts vid första anblicken), inte slutledning eller statistisk härledning.

6. Utvärdering är en nyckelprocess inom alla meningsfulla aktiviteter i vardagslivet, inklusive konstnärliga, fritidsmässiga och reflekterande. Inom många av dessa områden, liksom inom många av de tillämpade fälten, är de nuvarande nivåerna på utvärderingskvaliteten mycket begränsade och priset man får betala för undermåliga utvärderingar är mycket högt. Det beror delvis på dålig täckning i skolan inom relevanta ämnesområden (till exempel försäkring, investering, dubbla yrkesutbildningar för dagens arbetsmarknad, strategier för beslutsfattande), men också delvis på dålig täckning i vilken som helst av läroplanerna för K-8 (från förskolan till och med åttonde studieåret) avseende de grundläggande principerna för eller utövan-

det av utvärdering (till exempel skillnaden mellan god utvärdering och ren åsikt eller beskrivning, skillnaden mellan engagemang och bias, metoder för biaskontroll, värdet av och metoder för kritisk utvärdering och självutvärdering, färdighet i att utvärdera produktutvärderingar). Med tanke på detta och situationen som beskrevs under punkt 5 ovan, är det kanske dags för professionen att överväga en rejäl insats för att öka mängden utvärdering som finns i vanliga läroplaner, inte minst de som gäller naturvetenskap, samhällsvetenskap och yrkesförberedelser.

7. Färdighet i utvärdering är mycket värdefullt inom många angränsande aktiviteter som inte är rent utvärderande, till exempel planering, målförtydligande, diagnosticering, idékläckning, ledning, förordande, förklarande, problemsökning, undervisning och yrkesutbildning. Även om det är lämpligt och viktigt för utvärderare att företa och delta i sådana aktiviteter bör det stå klart att det i sig inte är någon ersättning för att skaffa sig eller tillämpa utvärderingsfärdigheter (eftersom det inte rör sig om alternativa sätt att gripa sig an utvärdering), och det är troligt att de ligger utvärderaren i fatet vid utvärdering av samma utvärderingsobjekt vid ett senare tillfälle.
8. De flesta utvärderingstillämpningar påverkar människor i grunden eftersom de står i samband med kvaliteten på deras arbete eller deras värde. Alla de angränsade aktiviteter som finns

uppräknade under punkt 7 påverkar människor. Det är därför viktigt att utvärdering används med ansvar och viktigt att det står klart att många av dess tillämpningar och angränsade aktiviteter kräver mellanpersonliga färdigheter som ligger långt från de färdigheter som krävs för att utveckla logiken i disciplinens utredningar.

Dessa överväganden antyder att betydande förbättringar inom yrkesmässig utvärdering har möjlighet att kunna komma till avsevärd nytta för det mesta av vad människor tar sig för. Jag tror att vårt arbete inom utvärdering, på samma sätt som inom socialt arbete, bör låta sig inspireras av den möjligheten. Vi borde ständigt vara på alerten, inte bara inför möjligheten att utvidga resultaten från ett utvärderingsområde till ett annat, utan också inför chansen att nå ut från den akademiska världens vanliga områden till sådana som alla medborgare har intresse i. Utvärdering är inte längre bara en samling olika aktiviteter, må vara indelade i discipliner. I likhet med den medicinska vetenskapen är den nu inte bara en akademisk disciplin utan även något mycket mer än en akademisk disciplin.

## Noter

1. Vissa avsnitt av den här artikeln lades fram för Australian Evaluation Societys årsmöte 1996 på Nya Zeeland i augusti förra året. Kommentarer och kritik är mycket välkomna: P. O. Box 69, Point Reyes, CA94956, [scriven@aol.com](mailto:scriven@aol.com), fax 415 663-1913. En version av artikeln har tidigare publicerats i *Scandinavian Journal of Social Welfare*, volym 7, nummer 2, 1998.