

Inriktningar i kvalitativ utvärdering

Ian Shaw

Abstract

Det anses av många som nästan dålig smak att framföra passionerade påståenden om någon metods överlägsenhet. Resonemangen av de flesta konventionella teoretiker är baserade på betraktelsesätt som syftar på att identifiera styrka i olika metoder och som inte uppmuntrar utvärderare att överdriva relevansen och tillämpningen av någon metod i något synsätt gällande utvärdering. Jag ägnar denna artikel åt att utveckla några argument som vänder upp och ner på vissa konventionella argument om urval av metodologi för utvärdering. Jag berör fyra områden där kvalitativ metodologi möjliggör för utvärderare att omarbete huvudaspekter av utvärdering i praxis, nämligen orsaksförståelse, hur metoder väljs ut, utvärdering av professionell praktik och användning av utvärdering.

Inriktningar i Kvalitativ utvärdering

Det anses av många som nästintill dålig smak att framställa passionerade påståenden om någon metods överlägsenhet. Resonemangen av de flesta konventionella teoretiker är baserade på betraktelsesätt som syftar på att identifiera styrka i olika metoder och som inte uppmuntrar utvärderare att överdriva relevansen och tillämpningen av någon metod i något synsätt gällande utvärdering. Till exempel, om utvärderaren arbetar med utvärdering som ansvarskyldighetsperspektiv - som till exempel att mäta resultat eller effektivitet - då utgör det slumpmässiga kontrollerade kliniska försöket (RCT) 'guldstandard, utvärderingsmetodens Rolls Royce' (Chelimsky, 1997: 101). Kvalitativa strategier, som till exempel fallstudiedesign, får inte användas för att angripa spörsmål och problem som bättre kan angripas med kvantitativa metoder. Dock, om perspektivet bakom utvärderingen är att skaffa mer kunskap och förståelse om en viss policy eller ett program, då kan kvalitativ inriktning vara att föredra.

Jag har en viss sympati för denna inställning. Även de bästa argument för att använda utvärderingsmetoder som tillhör en populär paradigm kan motverka sitt eget syfte, och jag tänker inte använda det knappa utrymmet till att förklara resonemanget bakom denna slutsats¹. I detta är jag överens med sociologen David Silverman när han säger: 'det finns ingen principiell grund att välja kvalitativ- eller kvantitativ inriktning. Det beror på vad du är ute efter' (Silverman, 1997b:14). Att demonisera ståndpunkter med vilka vi inte är överens är dåraktigt. Det förleder oss in i sentimentalitet och vidskeplighet. (Shaw, 1999). Jag tycker om berättelsen om Filip av Makedonien som antagligen hade en man anställd vars enda funktion var att slå honom på huvudet när som helst på natten med ett svins urinblåsa som var fäst på en pinne. Syftet var att påminna Filip av Makedonien om att han var en vanlig dödlig. Vi glömmer vår metodologiska dödlighet på egen risk. Det finns som Lather säger 'inga oskyldiga ståndpunkter' (Lather, 1991: 85), och vi behöver 'skydda vårt arbete från våra egna passioner' (Lather, 1986: 77).

Likväl är det så, att när tillämpningen av rätt man på rätt plats används som utvärderingsmått, medför detta förhastade avslutningar av frågor som inte är avslutade och förstärker således status quo, med vissa frågor för evigt kvarlämnade, onåbara för metodologiska utmaningar från andra håll. I praktiken fungerar detta som försvarslinje för traditionella utvärderingsmetoder. Att utvärdera resultat, förstå orsaksmässiga processer, skapa tillförlitliga generaliseringar och så vidare förblir ett område för ett fåtal utvärderare, och 'tjänar bara till att rättfärdiga ett perspektiv på bekostnad av ett annat' (Chelimsky, 1997: 108). Följaktligen blir gränser fastställda vid godtyckliga tidpunkter. Varför skall vi acceptera en bestämd arbetsfördelning som beror på en tillfällighet från sekelskiftet. Detta skulle troligen vidmakthålla onyttiga hierarkier. Om Chelimsky har rätt, när hon drar slutsatsen att RCTs är som Rolls Royce, är kanske fallstudier som Trabant eller som bäst som en liten engelsk Morris Minor.

Jag vill utveckla några sammantagna argument som stöd för att vända upp och ned på några konventionella argument för metodval för utvärdering. Jag skall beröra fyra temata, nämligen orsaksförståelse, metodval, utvärdering av yrkespraxis och användning av utvärdering.

Min allmänna ståndpunkt är att utvärdering är karakteriserad av en klunga av utvärderande ändamål. De möjliggör rimlig och produktiv respons till spörsmål som metodval, utvärderingsteori, utvärderingsetik och befrämjande utvärdering. Utvärdering förstås bäst som utförande av utvärderande forskning än som en åtskild samling av utvärderingsaxiomer eller metod som är skilda från det bredare forskningsarbetet. I detta avseende kan utvärderingsteori och metod till lika stor del tacka erkända utvärderingsteoretiker som de författare och forskare som inte betraktar sig själva som utvärderingsteoretiker men som bidragit med sitt arbete.

Gränslinjen mellan forskning och utvärdering är inte den enda som behöver tas bort. En relevant och noggrann utvärdering kräver skapande av inter-professionell utvärderings teoretisering och strategier, så att utbildning, hälsovård, straffrätt, juridik, arbetsmarknad och så vidare, är ömsesidigt uppmärksammade.

Utvärdering är mer - mycket mer - än programutvärdering. Kvalitativ utvärdering lovar distinkt men sammanhängande perspektiv över policy, program och praktik-utvärdering. Kvalitativ utvärdering ger trovärdiga partiella lösningar till problem av orsaksanalys och resultatutvärdering. Det gör det möjligt för oss att undvika ett onödigt instrumentellt och rationellt synsätt till användningen av utvärderingⁱⁱ.

Jag vill inte bli missförstådd på grund av omfattningen av de påståenden jag försöker stötta. Det skulle inte vara särskilt klokt, och snarare skulle jag då motverka mitt eget syfte om jag skulle påstå att kvalitativ utvärdering skulle vara en rådande utvärdering för alla utvärderingssyften. Snarare hoppas jag att jag kan rubba på den obetingade tron på utvärderande fördelar på vissa former av huvudsakligen kvantitativa utvärderingar - speciellt de som sätter en orimlig tilltro till möjligheter att kontrollera och mäta exakt på oberoende och beroende variabler.

Jag är inte heller övertygad om att alla former och traditioner av kvalitativ metodlära lämpar sig precis och direkt till utvärderingsändamål. Detta skulle bara ersätta en variation av konformism med en annan, genom en tendens att betrakta kvalitativ metodologi på ett orimligt homogent sätt. För att rätta till detta, tror jag att det finns ett behov av att utveckla en dialektisk blandning av metoder inom kvalitativ forskning. Detta skulle ske genom en utveckling av kritiska egenskaper av kunskap för olika kvalitativa metoder. En bra utgångspunkt i detta är en uppsats av McKeganey och hans kolleger, där de diskuterar fördelar och begränsningar av intervju- och observationsmetoder som en del av en studie av professionellt beslutsfattande när folk erbjuds en plats på ett ålderdomshem (McKeganey et al., 1988). Denna inledande analys behöver användas i ett större område av kvalitativa strategier och vara kopplad till kritiska egenskaper av sina associerade kunskapsrön (citerat från Greene och Caracelli, 1997: 12-13). Jag försöker inte ta an uppgiften nu, i betraktande av detta har jag undvikit att ge en övergripande definition av vad som utgör kvalitativa eller kvantitativa studier, men jag har begränsat mig till några illustrativa exempel, huvudsakligen från interpretativ sociologi, etnografi och fallstudier.

Att förstå orsaker

Den konventionella arbetsfördelningen är att kvalitativ undersökning är användbar i framställning av hypoteser/frågor, och beskrivande av förlopp, medan kvantitativa och mer statistiska framställningar behövs för att analysera utfall och verifiera hypoteser. Miles och Huberman summerar den här översikten - från vilken de starkt avviker - som:

Kvalitativa studier lämpar sig bara som utforskande försök för att framställa hypoteser - och ... starka förklaringar, inklusive kausala tillskrivningar kan bara erhållas genom kvantitativa studier.
(Miles och Huberman, 1994: 147)

Detta är mycket begränsande. Varken kvantitativa eller kvalitativa utvärderingar kan lösa problem av orsak och effekt på ett rättframt sätt. Med ledning av nyligen utförda arbeten i realistutvärdering har det skett förändringar i tankesättet gällande orsak och de motsvarande modeller på orsakshypoteser som härrör från det tankesättet. Huvudbegreppet är att det finns underliggande orsaksmekanismer som inte kan skönjas genom bearbetning av och mätningar av det yttre. Följaktligen ...'händelser i sig är inte det definitiva fokus av den vetenskapliga analysen... Verkligheten består inte bara av det vi kan se, men också av de bakomliggande orsaksfaktorer som är inte alltid är skönjbara' (House, 1991: 4). Den bakomliggande verkligheten skapar de aktuella händelser, av vilka vi har empiriska erfarenheter och sinnesintryck.

Detta beskrivs ofta som det generativa begreppet av orsakssammanhang.

När vi beskriver ett resultat generativt kommer vi inte fram med variabler eller korrelat som associerar med varandra; snarare försöker vi beskriva hur sambandet inträffade. Den generativa mekanismen utgör följaktligen resultaten. (Pawson och Tilley, 1997: 408) [kursiv stil i originalet]

De konventionella begrepp av orsakssammanhang som regelmässighet och samband är avfärdade till fördel för orsaksfaktorer som har "tendenser i växelverkan med andra tendenser på så sätt att en iakttagbar händelse kan skapas eller inte skapas alls" (House, 1991: 5). House citerar Manicas och Secord som säger att 'I den klassiska uppfattningen av vetenskap är världen en bestämd serie av tillfälliga händelser; för en realist är det en bestämd serie av verkliga strukturer. Och den skillnaden är enorm'. Följaktligen, istället för att bara dokumentera sekvens och samband av händelser, eftersträvar realisten att förklara händelser.

Även om denna uppfattning inte nödvändigtvis kräver kvalitativ metodologi, lämpar den sig väl till sådana metoder.

Kvalitativa studier är inte utformade att ge definitiva svar på orsaksmässiga frågor...(men) det kan ändå vara en riktig berättigad strävan. (Lofland och Lofland, 1999: 136, 138)

Miles och Huberman är till och med mindre förbehållsamma. De beskriver den konventionella synen som 'felaktig' (Miles och Huberman, 1994: 147), och framhåller att kvalitativ utvärdering är väl lämpad för:

1. att identifiera orsaksmekanismer
2. att studera komplexa lokala nätverk
3. att klarlägga tidsbestämda dimensioner av händelser
4. att ta sig fram och tillbaka mellan olika nivåer av variabler och processer
5. att möjliggöra undersökning och djupare förklaringar i enskilda fall genom analytisk induktion

Orsaksbeskrivningar blir lokala och 'nu-orienterade' (Lofland och Lofland, 1995: 141). Miles och Huberman utvecklar analytiska metoder som inriktar sig på orsaksbestämningar i både enkel- och multipelfall-förklaringar. Till exempel förespråkar de användning av fältundersökning för att kartlägga 'lokala kausala nätverk' som informanter bär med sig för att knyta dessa till utvärderarens egen framträdande orsaksmässiga kartläggning av omgivningen. Sådana kartläggningar börjar från 'orsaksfragment' som leder till länkad uppbyggnad av logiska beviskedjor. Sådana kausala nätverk

är inte sannolika, men specifika och bestämda, baserade på förståelse av händelser över en tidsperiod i den konkreta lokala kontexten - och knuten till en bra föreställning av varje variabel. (Miles och Huberman, 19994:159)ⁱⁱⁱ

Mycket av detta resonemang förutsågs av Cronbach i hans argument beträffande orsaksmodeller. Med avvisande av orsakssamband som händelser som med hög sannolikhet kan förutsägas, utvecklade Cronbach dubbla argument. För det första menade han att orsaker är beroende av lokala interaktioner av grupper av händelser. Mer än en grupp kan räcka, men ingen enskild grupp är nödvändig. För det andra, accepterade han att det ofta är saknade händelser eller förhållande som påverkar resultat av ett givet program, men om detta vet vi för lite. Han var den första utvärderingsteoretikern som framställde en rimlig förklaring av kontextuella faktorer i utvärderingar. Härav drog han slutsatsen att 'efteråt, när undersökaren med sina artificiella inskränkningar har lämnat scenen, skall säkerligen pågående program anpassas till lokala förhållanden' (Cronbach et al., 1980: 217). Dessutom 'programutvärdering är så beroende av sin kontext att reproduktion bara är ett bildligt uttryck' (p.222).

Kvalitativ utvärdering kan inte lösa problem av orsaksmässiga slutsatser och detsamma gäller kvantitativ utvärdering, men de kan fastställa orsakssammanhang 'som de faktiskt spelas ut i en bestämd omgivning' (Miles och Huberman, 1994:10).

Lofland och Lofland gör den viktiga observationen att orsaksmässiga svar i stort sätt är baserade på passivitets-begrepp av mänsklig natur. Kvalitativ undersökning har alltid undvikit orsaksbeskrivningar, inte på grund av att metoden är svag i det avseendet men på grund av engagemanget i ett aktivist-begrepp av mänskliga natur. Lofland påstår att ett aktivist-begrepp skulle leda till fokus på frågor som inriktar sig på både strukturer och strategier. Detta kommer att omfatta 'tolkning och skildring av

exakt vilken typ av situation deltagarna står inför' (Lofland och Lofland, 1995: 146), och förståelse av de 'oändliga fabricerade' strategier som människor konstruerar för att ta itu med situationen.

Ta till exempel Silvermans arbete i HIV-rådgivning. Han har rätt i att dra slutsatsen att 'det ofta är onödigt att låta våra forskningstema definieras i termer som ... "orsakssammanhang" av "dålig" rådgivning eller "konsekvenser" av "dålig" rådgivning' (Silverman, 1997a: 34) såtillvida att sådana temata reflekterar den syn på sociala problem såsom de uppfattas av professionella eller av samhällsgrupper. Likväl erfordrar detta inte ett övergivande av orsaksmässig undersökning i kvalitativ utvärdering. Forskning i hur de professionella i det oändliga fabricerar utformning av service och strukturer ger löfte om ett bättre sätt att förstå orsakssammanhang.

Som en illustration beskriver Shaw och hans kolleger en fallstudieutvärdering av en landsbygds aktivitetscenter för människor med inlärningsproblem. De observerade och intervjuade projektdeltagare, föräldrar, vårdare, ledningsgrupp-medlemmar, viktiga medarbetare och andra professionella. Projektdokumentation analyserades.

Se figur 1

När de beskrev och förklarade verksamheten på centret verkade de som intervjuades framhålla en eller flera av tre olika modeller på aktivitetscentret. Nämligen 'träning för arbete', 'personlig- och social utveckling' och 'lära för livet'. Dessa fungerade delvis som orsaksmässig kartläggning, innefattande en uppsättning av modellspecifika positioner om projektets ändamål, optimala målgrupper, önskvärda programmonster, anställningskrav, framtida utvecklingsstrategier och rimligt eller önskvärt resultat av projektet^{IV}.

Att välja metoder

Kvalitativ utvärderingsmetod är inte bara väl lämpad för studier av orsak och effekt på lokalnivå, men den berikar också val av metoder relevanta för utvärderingsändamål. Den inbördes relationen mellan kvalitativa- och kvantitativa metoder är inte enbart, och inte ens främst, en fråga om att välja metoder. Det handlar om frågor i figur 2 och är också mycket relevant för politiska spörsmål och ändamål i forskning i socialt arbete, värderingar, delaktighet i forskning, tvärvetenskaplig forskning och forskningsanvändning.

Se figur 2

Kvalitativ undersökning kan kasta ljus över programutfall på ett sätt som är mindre troligt med kvantitativ metodologi. Miller, till exempel, diskuterar hur texter från institutioner som är konstruerade för att förklara tidigare beslut, ofrånkomligen ger en vilseledande framställning av den beslutfattande processens öppenhet och komplexitet (Miller, 1997). Han ger ett vardagligt exempel på utvärderingsforskning gällande muskelträning av magen på ett vårdhem. Utvärderingen byggde på när och hur patienterna hade tömning av tarmen. Programmet bedömdes som framgångsrikt om patienterna använde toalett eller bäcken och som misslyckat för de som forsatte att smutsa ned i sängen. En patient hade smutsat ned i sängen. Emellertid möjliggjorde observationsmetoderna att forskaren kunde iakttaga hur ett

sjukvårdsbiträde ifrågasatte definitionen av detta som "ett misslyckande" eftersom hon menade att patienten visste vad hon gjorde och hade smutsat ned i sin säng som en protest mot att personalen favoriserat en annan patient. Detta illustrerar hur observation av kontexten av textkonstruktion belyser triviala, vardagliga liv. Detta skulle inte ha kommit med i den formella resultatredogörelsen. Textproduktion inom institutioner är "micro-politiskt organiserade" och detta inkluderar även resultatredogörelser gällande texten.

Ytterligare illustration av relevansen av kvalitativ metodologi för resultatutvärdering kan skönjas genom den förvånansvärda inverkan av Denzins 'tolkande interaktionism' (Denzin, 1989 och 2002; Mohr, 1997). Mohr, till exempel, utsträcker Denzins argument till utvärdering av kliniska resultat i hälsovårdsforskning. Hon påstår att metoden leder oss att ta en överblick över förhållandena mellan personliga problem, erfarenheter, policy, interventioner och institutioner. 'Tolkande interaktionism' tillåter intensiv noggrann undersökning av följder och utfall av olika interventioner' (1997: 284). Det kan:

1. sortera ut på vilka sätt problem är definierade
2. visa hur patienter upplever vården. Vad är det i interventioner som de finner är till hjälp eller ej och under vilka omständigheter
3. identifiera 'sekundära orsaker' t ex kontexter, kultur och den mening patienterna tillför

'Strategiska punkter för interventioner kan identifieras genom att kontrastera och jämföra patienternas otydliga förklaringar, och dessa kan användas för att ändra, att byta eller att förhandla och omförhandla interventioner' (p.284). Det är värdefullt när 'ett resultat inte är lätt märkbart, och ...interventionen är någonting som enbart patienten och inte de professionella kan definiera' (p.285).

Under senare tid har en konstruktiv, om än försiktig, dialog beträffande relativa fördelar och särdrag av kvantitativ och kvalitativ metodologi kommit fram. Det sociala arbetets litteratur ger ett bra exempel. Från den kvantitativa sidan, Reid i USA och Sinclair i Storbritannien, har utvecklat medlande positioner. Reid eftersträvar att 'omdefiniera den konventionella karaktären så att kvalitativ metodologi blir en del därav och inte åtskils'. Han betraktar kvantitativ forskning som stark när det gäller sammankoppling, kontroll, precision och stora datamängder, medan kvalitativ forskning kan skildra hur systemet fungerar, kontextuella faktorer och svårgripbar fenomen samt ge fullständig förklaring. 'Ingendera av metoderna är överlägsen den andra, men varje metod förser forskaren med olika undersökningsverktyg' som kan sättas mot en enhetlig uppsättning av standard (Reid, 1994: 477).

Sinclair lägger till Reids slutsatser i sin diskussion om randomiserade kontrollerade studier (RCTs), när han säger att kvalitativa metoder på många sätt är 'mer anpassade till komplexiteten i den praktikerns värld än den succérika RCT'.

Kvalitativ forskning drar uppmärksamheten till egenskaper i en situation som andra kan ha missat, men som när de upptäcks har stora implikationer för praktiken. Det motverkar en tendens att behandla de maktlösa som varelser med något mindre än normala mänskliga känslor. Det bidrar till ett etiskt försvarbart urval av utfallsmått. Och, i kombination med enkla statistiska beskrivningar, kan det leda till en välinformerad och

genomträngande utvärdering av program inom socialtjänsten. (Sinclair, 2000: 8)

Han vänder upp och ner på vanliga antaganden när han drar slutsatsen att,

Kvantitativ forskning i socialt arbete står inför synnerligen häftiga svårigheter beroende på abstrakta egenskaper i dess variabler, det flytande, sannolika sätt på vilket de hänger samman och den grad till vilken utfallskriterier är omtvistade. (smid 9-10)

Att utvärdera professionell praktik

Kvalitativ metodologi är också mycket handfast när det gäller utvärdering av direkt leverans av service. Huvudpunkten jag vill framlägga är att teori och praktik inom utvärderingsfältet har blivit för mycket påverkade av idéer om att utvärdering är ekvivalent med programutvärdering (t o m ofta som ekvivalent med utfallsutvärdering). Jag tror att vi starkare behöver skilja i vårt tankesätt mellan policy-utvärdering, program, projekt inom program och direkt praktik och leverans av service. Jag är speciellt intresserad av det sistnämnda. Hellre än att se bra praktik som att varande föremål för utvärdering - huruvida det är intern eller extern utvärdering är oväsentlig - tror jag att en god praktik i och av sig själv innefattar utvärdering för och tillsammans med service-användare. Det som påverkar mitt ställningstagande är:

- Reflektiv praktik
- Kvalitativ metodologi
- Empiriska bevis i utvärdering som en dimension av yrkespraktik
- Befrämjande utvärdering och användarledd forskning
- Aktionsforskning

Modellen jag har försökt att utveckla och har testat i vissa praktiker är baserad på en sammansatt bild av bra praktik som kräver utvärderande evidens, utvärderande inläring, och utvärderande rättvisa. Det baseras därför på vad som i Storbritannien alltmer kallas för 'kunskapsbaserad praktik, på lärandeorganisation och reflektiv inlärningslitteratur och på befrämjande utvärderingsmodeller. Den huvudsakliga praktiska infallsvinkel jag har utvecklat är att efterstäva att 'tolka' och 'kolonisera' metoder, speciellt men inte enbart kvalitativa (Shaw, 1996, 1997). Exempel på detta omfattar livshistoria, simulationer, fokusgrupper, berättelsemetoder, kulturöversikt, induktiva lokala logiska modeller och intervjuer av kollegor.

Empiriskt arbete har kastat ljus över hur praktiker efterstämmer att förstå och lösa utvärderingsproblem i sitt vardagliga arbete. Elks och Kirkhart anmodar att 'en alternativ forskningsmodell behövs, en som snarare är undersökande än bekräftande, som bygger en utvärderingsmodell baserad på praktikerns egen beskrivning istället för en påtvingad idealmodell och likvärdighetstester' (Elks och Kirkhart, 1993: 555). Sjutton socialarbetare intervjuades och tillfrågades hur de utvärderade det de sysslade med, och hur de kunde veta om de gjorde ett bra eller

ett dåligt arbete. Praktikerna medgav svårigheter i att veta om de var effektiva. De såg också en oförenlighet mellan rollen som utvärderare och praktiker. Forskarna anser att praktikerna har en implicit modell av praktikutvärdering som de beskriver som 'pragmatisk-professionell modell'. Denna innefattar att lita på intuition och erfarenheter, en internaliserad uppfattning av en ideal praktiker, ett beroende av feedback från kollegor, vänner och familj, och en modell av en ideal klient som alltid innefattar tillväxt och förändringar.

Humphrey och Pease utförde en motsvarande studie i vilken de intervjuade brittiska övervakare beträffande deras uppfattning av skyddstillsynseffektivitet. Övervakare tenderar att koppla bort övervakningsprocessen från övervakningens utfall. Således berättade en person att klienter 'ofta kommer in i mer trassel men jag tror inte att det på något sätt reflekterar på hurvida jag har varit effektiv eller inte'. Det finns faktiskt en utbredd övertygelse om att effektiviteten har med tur att göra. Man kan utföra ett 'strålande' arbete, men om omständigheterna är emot dig kommer de ändå att bryta mot lagen. Följaktligen 'om turen verkar bestämma resultatet, blir skyddstillsynen bara en fråga om att låta en lagförbytare gå fri i samhället och vänta på att turen slår till' (Humphrey och Pease, 1992: 40).

Efterföljande arbeten av Shaw och Shaw visar att socialarbetare verkar ha två olika utvärderingsmodeller i sitt huvud - en formell, 'riktig utvärdering' och en självutvärdering. Formell utvärdering upplevs som oförenlig med det sociala arbetets verklighet och nästan som en motsats till socialarbetarens utvärderande 'kartor' av sin aktuella dagliga utvärdering (Shaw och Shaw, 1997b). Utvärderingsstrategier konstruerades från en 'spelplan', vars framgång betraktades - som i forskning av Humphrey och Pease forskning - som delvis beroende på otillbörlig inverkan av 'ren tur'. Socialarbetare bedömde sin praktik i enlighet med att utfört arbete givit känslomässig belöning; fallet är 'på gång'; interventionen vann en stadig, stegvis förändring; uppgifterna hade utförts utan att ha tillfogat skada via socialvårdens verksamhet; och bekräftande bevis var tillgängliga från andra professionella. De här praktikerna var helt upptagna av orsaker och anledningar till resultat i sitt arbete, hade starka uppfattningar om komplexitet och otydlighet av bevis för det sociala arbetet som utförts, och var medvetna om det ständiga samspelet av att veta och känna (Shaw och Shaw, 1997a). Den känslomässiga betydelsen återspeglar Eriksons anmärkning om klinikern att

beviset inte är fullständig om han inte lyckas att använda sin egen känslomässiga respons under ett kliniskt möte som en beviskälla och som en vägvisare till handling. (Erikson, 1959: 93)

Betoningen hos dessa föregående paragrafer har legat på utvärdering som endemisk och som tagen för given inom professionell praktik, och för de betydande vinster som har erhållits genom att göra det uppenbart som ett steg i att förankra det i själva kärnan av god praktik. Denna förankring skulle ytterligare gynnas om vi undviker förenklade distinktioner mellan intern och extern utvärdering. En nackdel i mycket av praktikforskningen kommer från tendensen att betrakta 'praktik' som skild från teori och följaktligen betraktad som att 'vara teoretisk', som någonting som händer i hjärnan och 'det praktiska' som någonting som härleds till och har 'tillämpade' relationer till de här vägledande idéerna. Omvänt till detta, måste god praktikerutvärdering uppmärksamma att utforska tyst kunskap av olika slag.

Betydelsen av personlig kontakt och att dela praktiska kunskaper mellan praktiker skulle framhävas, och källan till förtroende och misstro mellan socialarbetare skulle klargöras (Collins, 2000). Detta är en stor dagordning, och en som socialarbetare och andra yrkeskategorier bara börjat att ta itu med. En konsekvens av detta är att vi inte kan hålla på med snäva distinktioner mellan praktikerforskning som internutvärdering och 'akademisk' forskning som externutvärdering. I en färsk analys av kvalitativ forskning i socialt arbete, reflekterade många som bidrog i analysen över de här frågorna. Hall och White redogjorde exempelvis för hur de behöll rollerna som både insider och outsider i relation till sina forskningsdeltagare. Hall 'kom in' som en outsider men blev på olika sätt en partiell insider (Hall, 2001). White började som en insider, dock såg hon sig själv genomgå en givande men potentiellt riskabel process av avståndstagande genom vilken hon till en viss grad blev en marginell intern- och externmedlem (White 2001). I samma volym, fokuserar Scourfield sina reflektioner på forskning och praktikrelationen genom sitt betraktande av hur det var att intervjua professionella intervjuexperter i socialt arbete (Scourfield, 2001).

På vilket sätt förväntar vi oss att utvärdering skall vara användbar?

Slutligen, fokus på kvalitativ metodologi, leder tämligen direkt till de allmänna frågor om hur utvärdering kan vara användbar för policy, program, projekt och professionell praktik. På det generella planet skulle utvärdering bedömas som användbar om den bevisligen har bidragit till en eller flera av följande

- bättre policy, service och praktik
- förstärka den morala avsikten/andemeningen i professionell praktik
- befrämjad metodologisk stringens, räckvidd, djup och innovation
- förstärkt känsla av yrkets intellektuella karaktär och belägenhet

Vi tenderar att göra enkla distinktioner mellan forskning som har direkt tillämpning och forskning som är grundläggande och följaktligen där direkta användningsfrågor är mindre relevanta. Detta fungerar inte (precis som den relaterade skarpa distinktionen mellan 'forskning' och 'utvärdering' inte fungerar).

Konventionell kvantitativ forskning på utfall är sammankopplad med förtroende i instrumentell forskningsanvändning. Problemet med detta är att det inte stämmer med bevis på hur utvärdering används, och det missförstår karaktären av den beslutsfattande processen. Den är baserad på en rationalistisk modell. Den rationalistiska modellen av beslutsfattande ser det som en serie diskreta händelser, varje fråga som skall avgöras är klart definierad, och beslut tas av en särskild grupp av aktörer som väljer mellan väl definierade alternativ, efter att de har övervägt bevis om varje alternativs sannolika resultat (Finch, 1986:149-150).

En viktig och tidig gestalt i denna fråga är Carol Weiss, vars arbete på sjuttioalet utforskade hur politiska överväganden spelade in i utvärdering. Hon inriktade sig på tre relaterade frågor. Först beskrev hon den politiska kontexten i vilken utvärderingen var belägen. Fastän hon huvudsakligen var engagerad i utvärdering och policy-forskning på federal nivå, ger hennes empiriska arbete med policy och programpersonal resonans alltigenom utvärderingsteori och praktik. Detta

understryker betydelsen av att vara klar över de som är inblandade i varje undersökning.

För det andra visade hon begränsningar av vedertagen instrumentell uppfattning om politiskt användning av information, genom sin framställning av användning som upplysning. Tillsammans med sina kolleger intervjuade hon 155 höga tjänstemän inom federala, delstats och lokala inrättningar för psykvård. Tjänstemän och personal använde forskning för att ge information om servicebehov som bevis för vad som fungerar och för att hålla jämna steg inom området. Likväl användes det också som ett ritualistisk överdrag för att rättfärdiga positioner, och för att ge personlig försäkran om att de positioner som hölls var de riktiga. På en bredare begreppsmässig nivå har det hjälpt tjänstemän att få mening i vad det handlar om. Bland alla dessa ändamål, 'var det en källa bland många, och inte alltid stark tillräcklig, för att driva den beslutsfattande processen' (Weiss, 1980: 390). När det gäller direkt forskningsanvändning, 'verkar instrumentell användning vara ovanlig, speciellt när frågorna är komplexa, konsekvenserna är ovissa, och ett otal aktörer är engagerade i den beslutsfattande processen (p. 397).

Användning av forskning reflekterades också i tjänstemäns syn på den beslutsfattande processen. Beslut uppfattades som fragmenterade både vertikalt och horisontellt inom organisationer, och som resultat av en serie gradvisa och formlösa steg. Därför, 'en viktig anledning till varför de inte rapporterar användning av forskning i specifika beslut är att många av dem inte tror att de fattar beslut' (p. 398). Följaktligen är titeln på en av hennes uppsatser - 'kunskap smyger och beslut växer' (Knowledge creep and decision accretion). Detta gav grunden till hennes slutsats att upplysning istället för instrumentell handling representerar den karakteristiska vägen för forskningsanvändning.

Upplysningsmodellen 'ger mycket mer utrymme till kvalitativ forskning, genom betoning på förståelse och begreppsbyggnad, istället för att komma med objektiva fakta' (Finch, 1986: 154).

För det tredje, Weiss genomsyrade användningsmodellerna med en realistisk syn av allmänintresse. Mer är för någonting annat har hon kämpat för en realistisk användningsteori.

Andra har därefter utvecklat sådana realistiska uppfattningar. Till exempel hävdar Chelimsky att utvärdering kan ha en avskräckande funktion. 'Med andra ord, blotta existensen av funktionen, och sannolikheten av en övertygande utvärdering, kan förhindra eller hejda en mängd oönskade metoder från regeringens sida' (Chelimsky, 1997: 105).

Flera försiktiga anmärkningar är behövliga. Först får vi inte bli för upptagna med modeller för forskningsanvändning. Chelimsky tror att 'det ofta är fallet att 'utvärderingar företas utan någon förhoppning om att användas'. Förväntan av icke användning är en karakteristik av några av de bästa utvärderingarna, inklusive 'de som ifrågasätter utbredda allmänna övertygelser i en ideologisk period, eller hotar mäktiga, förankrade intressen under vilken period som helst' (p. 105). Följaktligen,

'det finns bra anledningar till varför utvärderingar kan var skickligt gjorda och också oanvända' (p. 105). Chelimskys kommentarer är förnuftiga och rimliga.

Att rättfärdiga alla utvärderingar för användning på något sätt är en restriktion istället för ett begrepp som ger möjligheter, därför att det tvingar utvärderare i att blir överdriven upptagna med att deras undersökningsresultat är accepterade av användare, och riskerar att förvandla utvärderingar till banala upprepningar av status quo. (Chelimsky, 1997: 106)

För det andra, den upplysningsmodellen får inte antas som universellt lämplig. Till exempel är det mer troligt att praktiker-forskning kommer att försätta efter en mer direkt instrumentell uppfattning (se, till exempel figur 3).

Se figur 3

För det tredje, anammandet av upplysningsmodeller om forskningsanvändning kan lätt leda till en försvarsställning, som uppstår från det faktum att de kan användas för att stödja en stegvis växande syn på social förändring. Steget från en empirisk uppfattning att förändringar i policy och praktik ofta ske genom stegvis växande upplysning till ett underförstått antagande att social förändring ska ske på detta sätt kan vara logiskt omöjligt att stödja, men det är bedrägligt lätt.

För det fjärde och som motvikt till föregående poäng, bör vi undvika att vara orimligt optimistiska om forskningens möjlighet att förändra praktik. Hammersley har kritiserat viss professionellt motiverad forskningssyn av den anledningen att den är baserade på ett mycket snävt begrepp om forskningens relevans och en allför optimistisk tro på forskningens möjlighet att påverka policy och praktik. Han ger två anledningar för sin slutsats 'det finns bra anledning att tro att forskning inte rutinmässigt kan lösa lärarnas problem' (Hammersley, 1993: 430). 'Det finns ingen vetenskaplig metod som garanterar resultat' (p. 430) och lärarnas situationer är olikartade och det är inte troligt att de kommer att bli mottagliga för rutinmässig aktion. Snarare, 'förnuftig praktik kan inte vara detsamma som direkt tillämpning av teoretisk kunskap, men det är en aktivitet som nödvändigtvis involverar bedömning och är hämtad från erfarenhet lika mycket som från ... vetenskaplig kunskap' (p. 430).

För det femte finns det alltid ett viktigt samband mellan forskningsanvändning och etik. Betrakta de etiska och politiska dilemma om hur forskningsmaterial används. De här frågorna är ibland skarpere i kvalitativ forskning. Detta beror delvis på närheten och den resulterande tilltro som kan utvecklas mellan utvärderare och deltagare. I kvantitativ forskning kan det stora avståndet göra att de här frågorna blir mindre aktuella. Risken för svek kan också öka på grund av den typiska användningen av små urval, de resulterande svårigheterna att skydda individen så att denne kan förbli konfidentiell, och betoningen på detaljer om hur folk lever sina liv. Finch beskriver från sin forskning om lekgrupper känslan att 'jag kan eventuellt svika mina informatörer som grupp, men inte som individer' (Finch, 1986: 207). 'Där kvalitativ forskning är riktad på sociala frågor, finns det ett speciellt dilemma i att resultat kan användas på ett visst sätt för att försämra situationen hos målgruppen' (Finch 1985: 117).

Men ingen av de här protesterna minskar det övergripande värdet av försiktigt utvecklande av vad vi menar när vi säger att forskning är användbar. Till exempel, vardagliga diskussioner av bevisbaserad praxis grundas ofta på en missuppfattning när det gäller detta. Praktik är inte och kan inte bli 'baserad' på bevis på ett direkt och problemfritt sätt som många av dess anhängare föreställer sig.

I denna uppsats har jag argumenterat för att utvärderingsfältet har mycket att vinna genom att distansera sig från ett synsätt av prematur konsensus till utvärderingsmetodologi. Jag har berört fyra områden där kvalitativ metodologi tillåter utvärderarna att stöpa om centrala aspekter på utvärderingens praktik, dvs kausal förståelse, val av metodologi, utvärdering av professionell praktik och utvärderingsanvändning. Jag kunde lika lätt ha utvecklat samma argument för andra domäner - kvalitetsstandard i utvärdering, utvärderingens etik och styrning, utvärderingens sammanställning, utvärderingens relation till huvudströmningar i socialvetenskapen, tvärvetenskaplighet och så vidare. Inom alla dessa områden är jag övertygad av Paul Feyerabend's provokativa avsidesreplik att 'det inte är själva lösningen av pusslet som är anledningen till att vetandet växer, utan det är den aktiva växelverkan av åtskilliga ihärdiga åsikter' (citerat från Trend, 1979:84).

Referenser

- Atkinson, Paul, Delamont, Sara, Coffey, Amanda, Lofland, John and Lofland, Lyn (2000) *Handbook of Ethnography* London: Sage Publications.
- Chelimsky, Elinor (1997) 'Thoughts for a new Evaluation Society', in *Evaluation*, 3 (1): 97-118.
- Collins, Harry (2000) *Tacit Knowledge, Trust and the Q of Sapphire* Cardiff University School of Social Sciences. Working Paper.
<<http://www.cf.ac.uk/socsi/>>
- Cronbach, Lee, Ambron, S., Dornbusch, S., Hess, R., Hornik, R., Phillips, D., Walker, D., and Weiner, S. (1980) *Toward Reform of Program Evaluation*. San Francisco: Jossey-Bass.
- Denzin, Norman (1989) *Interpretive Interactionism*, Englewood Cliffs, NJ: Prentice Hall.
- Denzin, Norman (2002) 'Social work in the seventh moment' in *Qualitative Social Work: Research and Practice* 1 (1): 25-38.
- Eisner, Elliot (1991) *The Enlightened Eye: Qualitative Inquiry and the Enhancement of Educational Practice*. New York: Macmillan.
- Elks, Marion and Kirkhart, K. (1993) 'Evaluating effectiveness from the practitioner's perspective' in *Social Work*, 38 (5): 554-563.
- Erikson, Erik (1959) 'The nature of clinical inference' in Lerner, D. (ed) *Evidence and Inference*. Illinois: Free Press.
- Finch, Janet (1986) *Research and Policy: the Uses of Qualitative Methods in Social and Educational Research*. London: Falmer Press.
- Geertz, Clifford (1973) *The Interpretation of Cultures*. New York: Basic Books.
- Hall, Tom (2001) 'Caught not taught: Ethnographic research at a young people's accommodation project' in Shaw, I and Gould, N. *Qualitative Research in Social Work* London: Sage Publications.
- Hammersley, Martyn, (1993) 'On the teacher as researcher', in *Educational Action Research*, 1 (3): 425-445.
- House, Ernest (1991) 'Realism in research', in *Educational Researcher*, 20 (6) 2-9.
- Humphrey, Chris and Pease, Ken (1992) 'Effectiveness measurement in the Probation Service: a view from the troops' in *Howard Journal*, 31 (2): 31-52.
- Greene, Jennifer and Caracelli, Valerie (1997) *Advances in Mixed Method Evaluation: the Challenge and benefits of Integrating Diverse Paradigms*. New Directions For Evaluation, No 74, San Francisco: Jossey-Bass.
- Lather, Patti (1986) 'Issues of validity in openly ideological research', in *Interchange*, 17 (4): 63-84.
- Lather, Patti (1991) *Getting Smart: Feminist Research and Pedagogy with/in the Postmodern*. New York: Routledge.
- Lofland, John and Lofland, Lynn (1995) *Analysing Social Settings*. Belmont: Wadsworth.

- McKeganey, Neil, MacPherson, I. And Hunter, D. (1988) 'How "they" decide: exploring professional decision making' in *Research, Policy and Planning*, 6 (1): 15-19.
- Miles, Matthew and Huberman, A, (1994) *Qualitative Data Analysis: an Expanded Sourcebook*. Thousand Oaks: Sage.
- Miller, Gale (1997) 'Contextualizing Texts: Studying Organisational Texts' in Miller G and Dingwall R (eds) *Context and Method in Qualitative Research* London: Sage.
- Mohr, Wanda K. (1997) 'Interpretive Interactionism: Denzin's Potential Contribution to Intervention and Outcomes Research' in *Qualitative Health Research* 7 (2): 270-286.
- Pawson, Ray and Tilley, Nick (1997b) 'An Introduction to scientific realist evaluation', in Chelimsky, E. and Shadish, W. (eds) *Evaluation for the 21st Century*. Thousand Oaks: Sage.
- Reid, William J. (1994) 'Reframing the epistemological debate', in Sherman, E. and Reid, W. (eds), *Qualitative Research in Social Work*. New York: Columbia University Press.
- Scourfield, Jonathan (2001) 'Interviewing interviewers and knowing about knowledge' in Shaw, I and Gould, N. *Qualitative Research in Social Work* London: Sage Publications.
- Shaw, Ian (1999) 'Seeing the trees for the wood: the politics of evaluating in practice', in Broad, B. (ed), *The Politics of Research and Evaluation*. Birmingham (England): Venture Press.
- Shaw, Ian (1996) *Evaluating in Practice*, Aldershot (England): Ashgate.
- Shaw, Ian (1997) *Be Your Own Evaluator: a Guide to Reflective and Enabling Evaluating*. Wrexham (Wales): Prospects Publishing.
- Shaw, Ian and Shaw, Alison (1997a) 'Keeping social work honest: evaluating as profession and practice', in *British Journal of Social Work*, 27 (6): 847-869.
- Shaw, Ian and Shaw, Alison (1997b) 'Game plans, buzzes and sheer luck: doing well in social work', in *Social Work Research*, 21 (2): 69-79.
- Silverman, David (1997a) *Discourses of Counselling: HIV Counselling as Social Interaction*. London: Sage.
- Silverman, David (1997b) 'The logics of qualitative research' in Miller, G. and Dingwall, R (eds) *Context and Method in Qualitative Research*. London: Sage.
- Sinclair, Ian (2000) 'Methods and measurement in evaluative social work', paper from ESRC seminar series 'Theorising social work research'. National Institute for Social Work web site <http://www.nisw.org.uk/tswr/sinclair.html>
- Stake, Robert and Trumbull, D. (1982) 'Naturalistic generalizations' in *Review Journal of Philosophy and Social Science*, 7, 1-12.
- Weiss, Carol (1980) 'Knowledge creep and decision accretion', in *Knowledge, Creation, Diffusion, Utilisation*, 1 (3): 381-404.
- White, Sue (2001) 'Auto-ethnography as reflexive inquiry: the research act as self surveillance' in Shaw, I and Gould, N. *Qualitative Research in Social Work* London: Sage Publications.

Figure 1 Stakeholder models of a Rural Activity centre

	Training for Work	Personal/social growth	Education for life
Aims	Credible work skills for independent/sheltered work	Personal and social growth	Alternative occupation to enhance the quality of life
Target group	Demonstrable ability to benefit; younger	Wide range of age and ability	Wide ability range; younger
Programme	Time limited stay; skill learning; assessment and review; contracts; move-one facility; integration into work	Open stay period; small project; small-group activities; counselling; liaison with carers and social work agencies	Loosely held time limits; the best learning context; interest-led contracts; community based activities and outside links; craft work and home making skills
Staffing	Education and special needs employment skills; plus volunteers	Social and group work qualifications; plus expert consultants	Education and social work qualifications; plus volunteers
Outcome	Regular throughput; work placements; normalization of work patterns; skill learning	No clear distinction between programme and outcome	Wide range of social skills; integration into community networks; change of attitudes on the part of outside community members

Figure 2 Qualitative and Quantitative Methodology

Single cases or comparison.
 Cause and meaning.
 Context as against distance.
 Homogeneity and heterogeneity.
 Validity and the criteria of quality in social work research.
 The relationship of researcher and researched.
 Measurement.

Figure 3 Evaluation models and information use

<i>Model of use</i>	<i>evaluation focus</i>	<i>Evaluation base</i>	<i>Discipline links</i>	<i>Time span for action</i>
<i>Enlightenment</i>	Policy evaluation Programme development Advocacy evaluation	Higher education	Stronger discipline links Commitment to theorising	Longer term
<i>Instrumental</i>	Project evaluation; programme feasibility studies; practitioner evaluation	Agency sponsored ; 'Insider' evaluation ; self evaluation	Limited social science theorising	Immediate 'applications'

¹ I have attempted this in some detail in Shaw (1999) *Qualitative Evaluation* London: Sage. Eg Chapter 3.

² This paper has been written primarily to an evaluation audience. I have corresponding misgivings about the ways in which a majority of mainstream qualitative research academics are silent on the evaluation relevance of their field. For a recent example of the dog that does not bark, see Atkinson, Delamont, Coffey, Lofland and Lofland (2000).

³ An interesting connection can be drawn between Miles and Huberman's assessment and current work, mainly in American evaluation writing, on logic models. A qualitative methodology stance on logic models suggests the potential value of inductive 'logic models'. I see the existing USA approach as a way of combining a consensual stakeholder input with outcome analysis. The drawback of this approach is that it under-emphasizes the likely persistence of variant

informal logic models. Inductive, informal logic models offer a use of 'logic' in a sense not far from ideas of frames of meaning as used in Anthony Giddens and others.

IV Rethinking approaches to understanding causal processes leads naturally to rethinking the ways in which it is possible to generalize from one program and its evaluation to another program. Qualitative researchers and evaluators have developed partial answers to this question through ideas about vicarious experience (Stake and Trumbull. 1982), transfer (Eisner, 1990), 'thick description' (Geertz, 1973) and analytic generalisation.

ⁱ I have attempted this in some detail in Shaw (1999) *Qualitative Evaluation* London: Sage. Eg Chapter 3.

ⁱⁱ This paper has been written primarily to an evaluation audience. I have corresponding misgivings about the ways in which a majority of mainstream qualitative research academics are silent on the evaluation relevance of their field. For a recent example of the dog that does not bark, see Atkinson, Delamont, Coffey, Lofland and Lofland (2000).

ⁱⁱⁱ An interesting connection can be drawn between Miles and Huberman's assessment and current work, mainly in American evaluation writing, on logic models. A qualitative methodology stance on logic models suggests the potential value of inductive 'logic models'. I see the existing USA approach as a way of combining a consensual stakeholder input with outcome analysis. The drawback of this approach is that it under-emphasizes the likely persistence of variant informal logic models. Inductive, informal logic models offer a use of 'logic' in a sense not far from ideas of frames of meaning as used in Anthony Giddens and others.

^{iv} Rethinking approaches to understanding causal processes leads naturally to rethinking the ways in which it is possible to generalize from one program and its evaluation to another program. Qualitative researchers and evaluators have developed partial answers to this question through ideas about vicarious experience (Stake and Trumbull. 1982), transfer (Eisner, 1990), 'thick description' (Geertz, 1973) and analytic generalisation.