

## Två sätt att utvärdera forskning

**Jan Eriksson**

*Statens institut för byggnadsforskning, Gävle*

---

*I artikeln diskuteras två olika sätt att utvärdera forskning. Det handlar inte enbart om olika utvärderingsmetoder, utan också om två olika förhållningssätt till frågorna hur och varför man ska utvärdera forskning. Två exempel får illustrera vad som utvärderas vid den ena eller den andra sortens utvärdering, vad för slags resultat man får och hur utvärderingsresultaten kan användas.*

---

### *Inledning*

Under lång tid har sektorsforskningen i allmänhet och byggforskningen i synnerhet varit föremål för diskussioner. Vad som ifrågasatts är dels om denna forskning<sup>1</sup> haft en kvalitet som är jämförbar med den disciplin-knutna forskningen vid universitet och högskolor, dels om den verkligen varit till någon nytta för den sektor den är tänkt att tjäna. Byggforskningen har haft relativt mycket pengar att arbeta med, vilket sannolikt bidragit till det stora intresse som riktats mot just denna forskning. Under de senaste 10-15 åren har det också ifrågasatts om det, mot bakgrund av den minskande byggnadsverksamheten, verkligen finns samma behov som tidigare av byggforskning. Bland dem som är verksamma inom byggsektorn, dvs byggtreprenörer, materialtillverkare, arkitekter m fl, har det, under minst lika lång tid, funnits ett missnöje med byggforskningen. Man har ansett att det mesta som byggforskningen har åstadkommit varit ointressant och därmed av liten nytta för planering och byggverksamhet. Kritiken har alltså varit av två slag och kommit från två håll. Den har gällt dels den vetenskapliga kvaliteten, dels samhällsnyttan.

Mot denna bakgrund gav bostadsdepartementet, som ansvarar för finansieringen av den statliga byggforskningen, år 1981 i uppdrag åt dåvarande landshövdingen Ingemar Mundebo att göra en "översyn och ut-

värdering" av verksamheten vid byggforskningsrådet och byggforskningssinstituttet. Denna utrednings officiella namn var "Byggforskningsutredningen", men den har sedan kommit att kallas "Mundeboutredningen" (Ds Bo 1982:2). Inom ramen för Mundeboutredningen gjordes en utvärdering. Den genomfördes som en expertgranskning av ett antal forskningsprojekt. Denna utvärdering, som redovisats i en särskild rapport (Bostadsdepartementet 1982) får tjäna som exempel på det ena sättet att utvärdera forskning. I fortsättningen kallar jag den "M-utvärderingen".

Efter M-utvärderingen har byggforskningsrådets vetenskapliga nämnd (BVN) gjort ett antal utvärderingar av olika delar av byggforskningen. De fem första av dessa har behandlats i en studie av Tore Nilstun (Nilstun 1988) och den senaste i raden är den nu aktuella utvärderingen av boendeforskningen vid arkitekturskolorna. Även om BVN efter hand har utvecklat utvärderingsmetoderna på olika sätt så är huvuddragen i BVN-utvärderingarna desamma som i M-utvärderingen. De är alla exempel på vad som brukar kallas "expertbedömningar".

Det minskande byggandet har också använts som motiv för att diskutera dimensioneringen av byggfackutbildningen, däribland utbildningen av arkitekter. I början av 1980-talet gjordes två utredningar om arkitekturskolornas utbildning och forskning (UHÄ-rapport 1981:15, UHÄ-rapport 1983:7). Dessa utgjorde underlag för beslut om nedskärningar både i utbildningen och i forskningsorganisationen.

Byggforskningsrådet (BFR) är vid sidan av universitets- och högskoleämbetet (UHÄ) den största finansiären av arkitektskolornas forskning. Parallellt med den senare av de två UHÄ-utredningarna tog BFR initiativ till en seminarieserie om forskningen vid arkitekturskolorna. Avsikten var ursprungligen att, med utgångspunkt från seminarierna, söka ange den framtida forskningens omfattning, inriktning och organisation. BFR ansåg att detta behövdes som underlag för beslut om vilka nedskärningar som kunde göras.

För att organisera seminarieserien och utarbeta underlag för seminarierna engagerades Örjan Wikforss och jag. Vi skulle arbeta tillsammans med en referensgrupp som, förutom en representant för BFR, bestod av dekanerna vid de tre arkitekturskolorna i Stockholm, Göteborg och Lund. Redan vid första mötet med referensgruppen gavs arbetet formen av en utredning, där dock de seminarier som diskuterats tidigare skulle utgöra ett väsentligt inslag. Vi kom också fram till att det var viktigt att ge en nyanserad bild av forskningen idag, och vilka problem den hade.

Som utredare hade Örjan Wikforss och jag ganska stor frihet att bestämma utredningens uppläggning, och vi svarade också själva för utredningsrapporten (Eriksson och Wikforss 1983). Av det ursprungliga syftet att ange den framtida forskningsinriktningen blev det inte särskilt

mycket. Vi kom fram till att inriktningen i första hand måste beslutas av skolorna själva. I stället kom resultatet framför allt att bli dels en redovisning av vilka olika slags problem som sammanhänger med forskningen och ett försök att kartlägga problemens orsaker, dels en rad förslag till åtgärder för att undanröja dessa problem. I vår rapport sägs inte att det är fråga om en utvärdering av arkitekturskolornas forskning, men i efterhand har jag kommit att betrakta utredningen främst som en sådan.

Vad vi gjorde var alltså en utvärdering av forskningen, även om vi då inte använde beteckningen "utvärdering". Jag tror att detta visar bl a att det inte alltid är så klart vad som menas med "utvärdering" och vad som skiljer utvärderingar från sådana verksamheter som kallas "utredning" och "forskning". Den utvärdering vi gjorde är emellertid av ett helt annat slag än M-utvärderingen och får alltså utgöra exempel på det andra sättet att utvärdera forskning. I fortsättningen kallar jag vår utvärdering för "A-utvärderingen" där A står för "arkitekturskolornas forskning".

I stället för M-utvärderingen kunde jag kanske ha valt BVNs utvärdering av boendeforskningen vid arkitekturskolorna som exempel på det första sättet att utvärdera forskning. Det främsta skälet till att jag inte gjort det är att även om det finns principiella likheter mellan M-utvärderingen och BVN-utvärderingen så finns det också stora skillnader. Mycket av detaljkritiken mot den förra gäller inte den senare och den senare kan kritiseras på andra punkter än den förra. Dessutom anser jag att de två tidiga exempel på utvärderingar som jag valt, med alla sina brister, tydligare framhäver vissa väsentliga skillnader mellan de två sätt jag vill beskriva än om jag valt BVN-utvärderingen som det ena exemplet.

Efter genomgången av de två valda exemplen ska jag emellertid också kort kommentera BVNs utvärdering av boendeforskningen.

Det faktum, att jag själv deltog i den ena utvärderingen kommer givetvis att färga framställningen. Min kännedom om M-utvärderingen grundar sig nästan uteslutande på den skriftliga redovisningen av denna, medan jag vet mycket mer om A-utvärderingen än vad vi redovisat i vår rapport. Även om jag har kritiska synpunkter på både M- och A-utvärderingarna, vill jag redan nu säga att jag är mera kritisk mot M-utvärderingen. Möjligen beror detta på att det är svårt att vara självkritisk. Att båda dessa exempel på utvärderingar brister i sina detaljer är emellertid av mindre betydelse. Det viktiga är att de representerar olika principer. Bristerna kan i många fall avhjälpas.

#### *M-utvärderingens uppläggning*

M-utvärderingen genomfördes som en expertgranskning av 43 avslutade forskningsprojekt inom fem så kallade programområden: samhällsplanering, byggnadsplanering, byggnadsteknik, lokala energikällor och energi-

hushållning. Experterna utgjordes av tretton forskare och tio "praktiker". De arbetade i sex olika grupper och varje grupp granskade mellan fyra och tio projekt.

I rapporten anges syftet med utvärderingen vara "att bedöma följande tre aspekter i ett urval av projekt:

- Det studerade problemets relevans
  - Kvaliteten i det genomförda arbetet
  - Resultatets relevans/tillgänglighet för olika avnämargrupper"
- (Bostadsdepartementet 1982, sid 9)

Det är inte klart vad dessa aspekter mera precist står för. Anvisningarna för granskningen är mycket kortfattade och ger knappast mer än en antydning om vilka bedömningskriterier som ska användas. Inte heller i slutrapporten preciseras bedömningskriterierna. Tvärtom betonas mångtydigheten hos både "kvalitet" och "relevans" och att experternas egna tolkningar har varit avgörande. Bland annat detta har kritiserats i en uppsats av Tore Nilstun där han diskuterar vilka föreställningar om utvärderingens verkningssätt som föresvävat dem som deltog i utvärderingen (Nilstun 1986).

Expertgrupperna, som var och en bestod av tre till fem personer, arbetade självständigt under några veckor och lämnade skriftliga utlåtanden för varje projekt. Hur grupperna arbetat och vilka avväganden som gjorts får man sluta sig till från dessa utlåtanden. I vissa fall kan man konstatera att man inom en grupp varit oenig om ett projekts värde, och i åtminstone ett fall anklagar en expert en annan i samma grupp för att omfatta en vetenskapssyn som inte är allmänt accepterad.

Förutom forskningsrapporter från projekten granskade experterna också ansökningar om forskningsmedel och remissyttranden över dessa ansökningar.

Experternas utlåtanden presenterades vid ett antal seminarier. Munde-  
boutredningens sekreterare sammanfattade expertutlåtandena och dessa sammanfattningar redovisas i utvärderingsrapporten. För varje projekt har också sekreterarna gjort ytterligare en sammanfattning i form av betygsättning av de nämnda aspekterna. Betygen sattes i en tregradig skala: "bra", "mindre bra" och "ej tillfredsställande".

Det är inte heller riktigt klart hur urvalet av projekt gått till. Framför allt är det inte klart vilket syfte man haft med att göra det urval som gjordes. Enligt utvärderingsrapporten är det fråga om ett slumpurval, som dock styrs av ett antal "kriterier". Urvalet skulle bland annat vara representativt med avseende på de fem nämnda programområdena och med avseende på sex olika typer av forskningsutförande organ. Dessutom skulle det finnas projekt med anslag av olika storlek. I rapporten betonas också att utvärderingen är en "stickprovskontroll", vad som nu menas med detta. Den fråga som kan ställas är dock om de 43 utvalda projekten med dessa utgångs-

punkter över huvud taget kan vara representativa för de nära 1 000 projekt som ingick i populationen, och på vilket sätt de i så fall är representativa?

I utvärderingsrapporten står det: "Urvalet och därmed utvärderingens resultat ger inte tillräckligt underlag för generella utsagor beträffande BFRs samlade FoU-verksamhet" (Bostadsdepartementet 1982, sid 6). Trots detta görs generaliseringar, inte bara om byggforskningsrådets hela verksamhet, utan även om programområdena och andra mindre delar av byggforskningen. Jag ska återkomma till detta.

#### *A-utvärderingens uppläggning*

För att förstå A-utvärderingens uppläggning tror jag det är nödvändigt att klargöra två saker. För det första kom, som jag nämnt, syftet att förskjutas något så att huvudvikten kom att ligga vid en kartläggning av problemen och deras orsaker samt förslag till främst organisatoriska förändringar. Den centrala frågan var dock länge vad forskningen i framtiden skulle handla om, vilket präglade de första seminariernas uppläggning och urvalet av seminariedeltagare.

För det andra bör man ha klart för sig att vi som utredare hade ganska ingående kunskaper om den forskning och forskningsorganisation som skulle utredas. Vi är båda arkitekter och har en omfattande erfarenhet av forskning inom flera områden som finns företrädade vid arkitekturskolorna. Vi var väl förtrogna med undervisningen vid åtminstone en av skolorna (Stockholm) vad gäller både grund- och forskarutbildning. Vi hade också en ganska omfattande kännedom om skolornas forskning och forskningsresultat även utanför våra egna specialområden. Detta innebar att vi redan från början hade en mängd detaljerade uppfattningar. Att våra uppfattningar sedan till en del visade sig felaktiga hade mindre betydelse. Tillsammans med referensgruppen kunde vi i arbetets inledning utarbeta seminarieunderlag som inte varit möjliga att åstadkomma utan vår förhandsuppfattning. Vi gjorde också en genomgång av litteratur om skolornas forskning, allt från tidigare utredningar till debattartiklar. En viktig kunskapskälla var också projektöversikter, studiehandböcker och utredningar som skolorna själva gjort. På grund av ett pressat tidsschema kom dock en del av litteraturgenomgången att ske efter de inledande seminarierna, vilket naturligtvis är en brist.

Vid varje skola hölls två seminarier, ett heldagsseminarium med forskarna och ett halvdagsseminarium med lärarna i grundutbildningen. Inför dessa seminarier sände vi ut ett diskussionsunderlag, där vi dels försökte precisera vilka frågor vi ansåg borde behandlas vid seminarierna, dels framförde våra uppfattningar om forskningen. Syftet med att inbjuda lärare till ett särskilt seminarium var att därigenom få en föreställning om i vilken utsträckning och på vilket sätt forskningen bidrog eller skulle kunna bidra

med kunskaper till grundutbildningen. För att få konkretion i diskussionerna ombads varje skola välja ut tre forskningsprojekt som skolan ansåg vara god forskning. Varje forskarseminarium inleddes med en redovisning av dessa projekt med åtföljande diskussioner. Därefter diskuterades vårt seminarieunderlag.

Efter seminarierna vid skolorna genomfördes ett seminarium med arkitekter som var praktiskt verksamma inom olika delar av byggsektorn eller inom statliga och kommunala myndigheter med ansvar för byggande och planering. Tanken var att dessa skulle representera de viktigaste målgrupperna för A-skolornas forskning. Vid detta seminarium redovisades dels diskussionsunderlaget från seminarierna vid skolorna, dels erfarenheterna från dessa. Vidare ombads deltagarna i förväg att ge exempel på forskningsresultat de haft nytta av, samt ange vad de ansåg att arkitekturskolorna borde forska om. Alla seminarier spelades in på ljudband, och vi gjorde skriftliga referat av diskussionerna vilka finns återgivna i vår slutrapport.

Mot bakgrund av de sju seminarierna och litteraturgenomgången gjorde vi en problembeskrivning samt föreslog en rad åtgärder för att minska eller undanröja dessa problem. Med detta som underlag genomfördes ett åttonde seminarium med några av de tidigare seminariedeltagarna. Vid detta seminarium konstaterades att det ursprungliga syftet, att formulera en framtida önskvärd forskningsinriktning, framstod som mindre intressant jämfört med problemen med forskningens kvalitet, arbetsmiljön, handledningen mm och vad som kunde göras för att lösa dessa problem.

Därefter gjorde vi en preliminär rapport som sändes till alla som deltagit i något seminarium. Avsikten med detta var i första hand att kontrollera att våra seminarierferat på ett rättvisande sätt återgav vad som sagts vid seminarierna.

Under arbetets gång fick vi en mängd information genom brev- och telefonkontakter, genom kontakt med UHÄs pågående utredning och genom diskussioner med vår referensgrupp, vilket kompletterade seminariediskussionerna och litteraturgenomgången.

A-utvärderingens datainsamling har väsentligen skett genom något som skulle kunna kallas strukturerade samtal. Grundstrukturen utgjordes av vårt sätt att ställa frågor och våra påståenden som konfronterades med seminariedeltagarnas uppfattningar. Genom den dialog som seminarieformen möjliggör kunde vi uppnå en viss precision – åtminstone så långt att oenighet på grund av missförstånd kunde undvikas. Det kom givetvis fram olika uppfattningar, men även de uppfattningar som klart avvek från våra finns redovisade i rapporten, så att läsaren kan värdera dessa och argumenten för och emot dem. I slutet av inledningskapitlet i rapporten från A-utvärderingen gör vi följande deklARATION: "Det finns naturligtvis en risk



för att våra personliga uppfattningar och förutfattade meningar i alltför hög grad kommit att påverka utredningsresultatet. Vi har emellertid, genom ovan redovisade metod, försökt minska denna risk dels genom att systematiskt ha försökt lära oss vad andra har att säga, dels genom att redovisa våra egna ståndpunkter så klart att de är möjliga att diskutera och ifrågasätta" (Eriksson och Wikforss 1983, sid 8).

#### *M-utvärderingens resultat*

Bedömningen av varje projekt i M-utvärderingen skulle gälla de tre "aspekter" jag citerat tidigare. Redovisningen av experternas bedömningar görs emellertid under fyra rubriker: "Projektplan", "Ämnesrådets relevans", "Genomförande" och "Resultatens relevans/tillgänglighet" (Bostadsdepartementet 1982, sid 6 och 122). Vad som tillkommit är en bedömning av de handlingar som utgjort underlag för ansökan om forskningsmedel. Jag vet inte om detta är tänkt som en fjärde aspekt eller om bedömningen av projektplanen ska betraktas som en del av aspekten "kvaliteten i det genomförda arbetet". Redovisningen tyder dock på att det är fråga om en fjärde aspekt.

Redovisningen under rubriken "Resultatens relevans/tillgänglighet" handlar dessutom helt klart om två olika aspekter. Relevansbedömningen gäller främst om experterna tror att det finns en intresserad publik. Tillgängligheten handlar däremot om resultatens begriplighet (för tänkta målgrupper). Dessa två saker kan bedömas oberoende av varandra, såvida inte resultatet är helt obegripligt och av det skälet blir ointressant. Såvitt jag förstår är det alltså åtminstone fem olika aspekter som redovisas under fyra olika rubriker. I sammanfattningen av utvärderingsresultatet sammanvägs dock relevans/tillgänglighet till ett betyg, vilket innebär att det inte framgår om bristerna består i att resultatet är ointressant eller i att det är svårbegripligt.

Om man vill ge ett grovt sammanfattande betyg är det mindre viktigt om ett forskningsresultat är ointressant eller svårbegripligt. I båda fallen är det fråga om en allvarlig brist. I en grov sammanfattning tycker jag dock att man borde kunna gå ännu längre och nöja sig med ett enda betyg för varje projekt. Denna möjlighet avvisas dock i utvärderingsrapporten: "Någon helhetsbedömning av varje projekt efter en enhetlig skala låter sig inte göras då utvärderingen avser olika aspekter av projekten. Projektplan och relevans kan tex ges positiva omdömen medan genomförandet och presentationen bedöms som svaga" (Bostadsdepartementet 1982, sid 6).

För mig är det uppenbart att de bedömda "aspekterna" hänger ihop på ett sådant sätt att det är möjligt att göra en sammanfattande bedömning av ett projekts värde. Om forskningens genomförande och resultatens relevans/tillgänglighet bedöms som dåliga, kan inte detta kompenseras av att

projektplanen är bra och ämnet mycket relevant. Som helhet måste ett sådant projekt anses vara dåligt. Jag kan heller inte inse varför ett projekt som genomförts så att det motsvarar högt ställda krav på god vetenskap och ger mycket relevanta resultat skulle bedömas som mindre bra därför att projektplanen är dålig.

För att undvika missförstånd vill jag här skjuta in ett förtydligande. Att jag anser det vara möjligt att ge ett sammanfattande omdöme om ett forskningsprojekt innebär inte att det är ointressant att bedöma olika aspekter. Tvärtom menar jag att det är nödvändigt att så noggrant som möjligt ange vilka olika egenskaper hos ett projekt som man beaktar vid bedömningen och vilka kriterier som används för att värdera dessa egenskaper. Den som utsätts för granskning har rätt att få veta på vilka grunder det granskade projektet bedöms som bra eller dåligt.

Jag kan inte här i detalj diskutera experternas granskning, utan bara redovisa några allmänna iakttagelser som också redovisas i utvärderingsrapporten.

Granskningen under rubrikerna "projektplan" och "ämnesområdets relevans" leder till kritik som främst riktas mot byggforskningsrådets handläggning av medelsansökningar och byggforskningsinstitutets styrning av sin forskning. Enligt experterna har man i många fall beviljat medel eller startat projekt på alltför dåliga grunder.

Även då "resultatets relevans/tillgänglighet" diskuteras framförs kritik mot byggforskningsrådet därför att man publicerar nästan allt och inte ställer större krav på redovisningen. Till en del kan alltså expertgranskningen ses som en granskning av de forskningsinitierande och forskningsstyrande organen. Huvuddelen av granskningen gäller dock enskilda forskningsprojekt (och därmed enskilda forskare).

När det gäller experternas bedömning av "kvaliteten i det genomförda arbetet" är det intressant att notera tre saker.

För det första har experterna ibland varit oeniga om hur ett projekt ska bedömas. Om detta sägs följande i utvärderingsrapporten: "Det är naturligt att forskare från olika vetenskapliga discipliner kan ha skilda uppfattningar liksom att man som 'praktiker' ställer andra kvalitetskrav än de som företräder forskarsamhället" (a.a. sid 123). Jag är tveksam till om detta ska anses vara naturligt. En expert kan möjligen sakna den nödvändiga fackkompetensen för att bedöma ett projekt, men i så fall verkar det naturligt att avstå från bedömning, inte att hävda några avvikande vetenskapsideal. Det finns ingen anledning för praktikerna att bedöma den vetenskapliga kvaliteten, såvida de inte, utöver sin praktikerkompetens, även har vetenskaplig kompetens. Enligt min uppfattning kan oenigheten åtminstone delvis förklaras med att experterna varken för varandra eller för sig själva försökt klargöra vilka kriterier man tillämpat vid sin bedömning.



För det andra konstaterades vid expertgranskningen att en stor del av de projekt som granskats inte är "forskningsprojekt i traditionell bemärkelse. Flertalet av projekten har inslag av utredning, litteraturöversikt, utvecklingsarbete, förstudie och experimentbyggande" (a.a. sid 123). Möjligen hade sekretariatet konstaterat detta redan före experterna, men detta förhållande medförde inte något försök till typologisering av "forskningsprojekten" trots att man förutsätter att projekt av olika slag måste bedömas med olika utgångspunkter. I rapporten sägs bara att "utredningsarbete och andra typer av projekt /har/ bedömts utifrån kvalitetskriterier som till stor del valts av respektive expert och expertgrupp" (a.a. sid 123).

Det är visserligen så att byggforskningsrådet även ska stödja bl a utvecklingsarbete, men om forskning framstår som något ovanligt inom något som kallas "byggforskning" är detta anmärkningsvärt. Detta konstaterande leder dock inte till någon kritik mot hur forskningsmedlen används. För övrigt anser jag att även om tex utredning och utvecklingsarbete delvis måste bedömas med andra kriterier än forskning så är det en mängd kvalitetskrav som är gemensamma för de olika verksamheterna.

För det tredje kan man se ett mönster i expertgruppernas bedömning. När det gäller bedömningen av kvaliteten i genomförandet är detta så uppenbart att ordet "mönster" också används i utvärderingsrapporten. Där konstateras "Inom flertalet grupper har 1-2 projekt bedömts vara av hög kvalitet och standard i uppläggning och genomförande, medan 1-2 projekt blivit hårt diskuterade och kritiserade. ...Resterande projekt har med experternas bedömningsgrunder hamnat mellan dessa 'ytterligheter' av kvalitetsmässigt bra och dåliga projekt" (a.a. sid 124). För mig verkar det rimligt att förklara detta bedömningsresultat som ett välkänt socialpsykologiskt fenomen. Man rangordnar det man har och får en topp och en botten.

Jag har tidigare uttalat min skepsis inför möjligheterna att göra generaliseringar med utgångspunkt från de 43 utvalda projekten. I utvärderingsrapporten görs också en del markeringar, bl a den jag citerat tidigare, som måste tolkas så att det inte går att dra några generella slutsatser. Trots detta redovisas experternas bedömningar som om det vore fråga om en stor statistisk undersökning. Resultatet sammanfattas i rapporten så som framgår av tabell 1.

Att kolumnerna inte summerar till 43 beror på att sekreterarna avstått från betygssättning i de fall experterna varit oeniga. Den allmänna slutsats som möjligen kan dras av detta är att en stor del av byggforskningsprojekten, kanske den största delen, inte är så bra, åtminstone vad gäller kvaliteten i genomförandet. I utvärderingsrapporten går man dock betydligt längre och jämför olika programområden och olika forskningsutförande organ. I allmänhet ger också resultatet en bild som överensstämmer med gängse föreställningar om vad som är bra och dålig forskning, utom på en punkt,

	Projektplan	Ämnesrådets relevans	Genomförande	Resultatens relevans
Bra/God	16	36	12	12
Mindre bra/ Mindre god	6	3	12	12
Ej tillfreds- ställande	9	2	16	17
	31	41	40	41

Tabell 1. Utfallet i M-utvärderingen. Antal projekt med olika betyg gällande projektplan, ämnesrådets relevans, kvaliteten i genomförandet och resultatens relevans/tillgänglighet (Bostadsdepartementet 1982, sid 125).

där det också görs ett allvarligt misstag. I utvärderingsrapporten hävdas att "Resultatet av granskningsarbetet visar också att projekt som utförts av traditionella forskningsinstitutioner som regel ligger på en högre kvalitetsnivå än projekt som genomförts av t ex enskilda konsultbolag. ... De etablerade högskoleinstitutionerna har som regel en intern kontroll av kvaliteten ..." (a.a. sid 125). Detta innebär att universiteten, de tekniska högskolorna, nordiska institutet för samhällsplanering och ekonomiska forskningsinstitutet vid handelshögskolan skulle utmärka sig genom projekt av högre kvalitet än de som görs av konsultföretag. Jag kan tänka mig att det i allmänhet förhåller sig så, men i just det urval som granskades förhåller det sig tvärtom. Konsulterna presterar bättre forskning än de etablerade forskningsinstitutionerna. Den slutsats jag citerat ovan är alltså felaktig som en generalisering på grundval av expertgranskningen.

I en strikt mening kan resultatet av M-utvärderingen synas ha blivit tämligen magert, i synnerhet om man inte kan ge några generella omdömen om byggforskningen eller dess olika delar. Sannolikt hade granskningsarbetet ett större värde än så för de experter som deltog och för M-utredningens sekreterare genom att man fick en mängd insikter om byggforskningsrådets handläggning av medelsansökningar, om hur forskarhandledningen fungerade i enskilda fall, om anslagsmottagarnas kompetens, om forskningsmiljöns betydelse mm. Antydningar om sådant kommer ibland fram i experternas yttranden om enskilda projekt. M-utvärderingen har också orsakat en kritisk diskussion om utvärdering av forskning och givit väsentliga erfarenheter inför fortsatt utvärderingsarbete. Jag har emellertid utgått från vad som redovisas i utvärderingsrapporten och därav drar jag slutsatsen att M-utvärderingen var av ringa värde för den aktuella utredningen. De förslag som framläggs i utredningsrapporten skulle, enligt min mening, ha kunnat framläggas på lika goda (eller dåliga) grunder utan M-utvärderingen.

*A-utvärderingens resultat*

Som jag nämnt var A-utvärderingens ursprungliga syfte inte i första hand att utvärdera forskningen vid arkitekturskolorna, utan att ange en önskvärd framtida forskningsinriktning. Det innebar att vi ägnade ganska mycket tid åt att fundera över hur och varför man anger en forskningsinriktning. Jag ska här kort säga något om vad våra funderingar och diskussionerna kring detta resulterade i.

Inom byggforskningen är det vanligt att man gör en indelning av forskningsområden med hänsyn till olika samhällsproblem, dvs t ex energifrågor, ombyggnadsfrågor och förvaltningsfrågor. En sådan indelning kan ha sitt berättigande, men som underlag för att planera den framtida forskningen i en skola eller institution förefaller den mindre lämplig. För att kunna planera forskningen är det rimligare att diskutera inriktningen i termer av forskar- och handledningskompetens, undersökningsmetoder, instrumentell utrustning mm. Detta behöver inte medföra att man utesluter möjligheterna att forska om olika samhällsproblem som t ex energifrågor, men det innebär att man begränsar sig till sådana energifrågor som man har förutsättningar att framgångsrikt behandla.

Vidare hävdade vi att det var viktigt att särskilja åtminstone forskning och utvecklingsarbete. Skälet till detta var främst att vi ansåg att det fanns oklara föreställningar om vad forskning var och vilka resultat man kan förvänta sig av forskning, som i sin tur medförde både dåliga forskningsprojekt och dåliga utvecklingsprojekt. På denna punkt nådde vi inte några större framgångar. Enligt min mening råder det fortfarande stor osäkerhet om vad som är forskning respektive utvecklingsarbete, inte bara vid arkitekturskolorna utan inom byggforskningen i allmänhet.

Som en utgångspunkt för att välja forskningsuppgifter föreslog vi arkitektutbildningens kunskapsbehov. En forskning inriktad på att förbättra det teoretiska underlaget för arkitekturtyrket och den pedagogik som behövs för att lära ut de nödvändiga färdigheterna förefaller naturlig att ägna sig åt inom arkitekturskolorna. Vidare föreslog vi att skolorna gemensamt skulle försöka koncentrera den långsiktiga forskningen till ett fåtal områden så att det skapas en bred bas för kritisk diskussion. Som jag tidigare nämnt avstod vi dock från att föreslå några detaljerade planer för den framtida forskningen.

Vi såg aldrig som vår uppgift att ta reda på hur stor del av forskningsprojekten vid arkitekturskolorna som kunde anses vara dålig eller bra. I stället förutsatte vi att en betydande del av forskningen inte var av så god kvalitet. Om detta gäller 25 % eller 75 % av projekten spelade ingen roll, utan vi antog att kvalitetsproblem var så vanliga att de måste tas på allvar. Detta antagande bekräftades i seminariediskussionerna. Även bland forskarna själva fanns en utbredd uppfattning om att forskningen delvis hade stora brister.

Det kan diskuteras om den uppfattning som kom fram vid seminarierna verkligen kan ses som en bekräftelse av våra antaganden om forskningens kvalitet. Det är lika rimligt att tolka opinionen så att det finns skilda åsikter om vad som är god forskning. Vad den ena forskaren anser vara bra anser den andre vara dåligt, och tvärtom. Därför kan man alltid finna stöd för att en del forskning är dålig, medan en annan del är bra. En sådan invändning mot våra slutsatser om forskningens kvalitet är berättigad, men vi förde också diskussioner om i vilka avseenden forskningen kan anses ha kvalitetsbrister. Resultatet av dessa diskussioner blev att det fanns åtminstone tre typer av brister om vilka det rådde stor enighet. De två första bristerna gällde redovisningen och den tredje det metodiska genomförandet.

Redovisningen av en forskningsuppgift kan vara sådan att det är oklart vilket problem som behandlats, vilka metoder som använts eller vilket resultatet av forskningen är. Redovisningen kan också vara svårförståelig av andra skäl som t ex dålig språkbehandling, oklar disposition eller ovidkommande utläggningar. Att genomförandet är metodiskt dåligt kan ta sig uttryck i att det saknas stöd för de påståenden som framförs, att generaliseringar görs på otillräckliga grunder eller att felaktiga slutsatser dras av de fakta som redovisas. Även om det fanns enighet om att brister av detta slag utmärkte en del forskning, så kan det givetvis finnas olika uppfattningar och helt andra brister. Det viktiga är emellertid att seminariedeltagarna enades om några kriterier för bedömning av forskningens kvalitet och att det fanns ett ganska stort antal projekt av sämre kvalitet enligt just dessa kriterier.

Vid seminarierna diskuterades också forskningens relevans och sambandet mellan kvalitet och relevans. Det var inte alltid så klart vad som menades med "relevans" i dessa diskussioner, men jag kan säga att det handlade dels om huruvida forskningsproblemet eller ämnet för forskningens anknöt till något intressant och aktuellt samhällsproblem, dels om resultaten var intressanta och användbara. När det gäller problemens aktualitet var uppfattningen att det i allmänhet valts intressanta utgångspunkter för projekten. Detta överensstämmer med M-utvärderingens resultat (se tabell 1). Resultatens användbarhet sattes dock i större tvivelsmål. En svårighet här är att det sannolikt finns olika föreställningar om på vilket sätt forskningsresultat kan vara användbara. Jag har i ett annat sammanhang försökt utveckla ett precisare relevanskriterium (Eriksson 1986). Om man alltid förväntar sig rekommendationer, handlingsregler eller metoder, så blir all beskrivande, förklarande eller analyserande forskning oanvändbar. Det enda som godtas blir då utvecklingsarbeten.

Detta förhållande, att det finns olika förväntningar på forskningsresultatens användbarhet, är ett problem. Det återspeglar också olika föreställningar om vad forskning är och vilket syfte forskningen har. För min egen

del är t ex sanning ett avgörande kriterium på om forskning är av god kvalitet. När det gäller tillämpad forskning förefaller det rimligt att på resultaten ställa kraven att de både ska vara sanna och användbara, men sanningen, så långt den är möjlig att finna, är viktigare. Om man ägnar sig åt onyttiga problem kan detta visserligen anses vara slöseri med forskningsmedel, men att producera till synes användbara resultat som grundar sig på dålig forskning och därför möjligen är osanna, bedömer jag vara ett större problem. Tillämpning av felaktiga teorier får svåröverskådligare konsekvenser än resursslöseriet.

Av de diskussioner vi förde vid seminarierna drog vi den slutsatsen att vi på goda grunder kunde hävda att det fanns allvarliga problem vad gäller forskningens kvalitet. Nästa steg i diskussionen var att försöka klargöra orsaker till dessa problem.

Det kom fram en mängd tänkbara orsaker till kvalitetsbristerna. Jag kan inte här i detalj redovisa alla dessa, utan begränsar mig till några av de viktigaste.

Arkitektutbildningen är i första hand en praktisk yrkesutbildning. Det är visserligen så att yrket förutsätter en mängd teoretiska kunskaper om teknik, natur, människor och samhälle, men i undervisningen förmedlas mycket litet kunskaper om vetenskaplig metodik eller ett kritiskt förhållningssätt till det teoretiska lärostoffet. Det innebär att de arkitekter som fortsätter med forskarutbildning har sämre förutsättningar att bedriva forskning än vad som gäller för många andra discipliner där vetenskaplig träning utgör en stor del av grundutbildningen.

Eftersom grundutbildningen vid arkitekturskolorna är inriktad på praktisk yrkesutbildning har professorerna av tradition rekryterats på sina praktiska meriter som arkitekter. Ofta har de ingen forskningserfarenhet alls. Många professorer saknar alltså kompetens både för att bedriva egen forskning och för att handleda forskarstuderande. Trots detta ligger professorernas huvudansvar i att leda forskning och forskarutbildning.

Forskningen vid arkitekturskolorna har i stor utsträckning bedrivits med andra medel än de statsanslag som ges till universitet och högskolor. Många projekt har finansierats av byggforskningsrådet och man kan säga att forskningen inom vissa områden blivit starkt beroende av finansiering utifrån. Det innebär att bl a byggforskningsrådet fått ett stort inflytande över vilken forskning som bedrivs och vem som får pengar. Detta kan ha både för- och nackdelar. Till nackdelarna hör otrygga anställningsförhållanden för många forskare, relativt kortsiktiga forskningsuppdrag och svårigheter att utveckla en självständig forskningsprofil. I den forskning som finansieras externt finns inte mycket utrymme för annat än projektarbete. Tiden för att tänka igenom nya forskningsuppdrag och söka nya anslag blir knapp, liksom tiden för kritisk diskussion. Vid seminarierna



framkom att en fungerande intern kritik i stort sett saknades och att det var svårt att upprätthålla meningsfulla kontakter med omvärlden.

Inom skolorna finns också organisatoriska problem som bidrar till att försämra forskningsmiljön och gör det svårt att finna och behålla kompetenta forskare.

Det får räcka med dessa exempel på problem. Eftersom vi ansåg att vi funnit en mängd orsaker till forskningens kvalitetsproblem - åtminstone tänkbara orsaker - var det relativt lätt att föreslå åtgärder i syfte att minska de negativa effekterna. Bland annat föreslog vi att man skulle inrätta två typer av professurer, en för forskning och en för praktisk färdighetsträning. Vi föreslog också att forskarutbildningen skulle anpassas bättre till de förutsättningar som grundutbildningen ger, men även att grundutbildningen skulle förändras så att den omfattar mera vetenskaplig skolning än den som nu ges.

#### *Jämförande karakteristik av M- och A-utvärderingarna*

Jag har i litteraturen om utvärdering och utvärderingsforskning sökt finna någon taxonomi med vars hjälp man enkelt kan karakterisera dessa två sätt att utvärdera forskning. För mitt ändamål har jag funnit det fruktbarast att använda den problemram som utarbetats av Tore Nilstun (Nilstun 1988, Nilstun 1989, Nilstun och Hermerén 1984).

När det gäller utvärdering av tillämpad forskning och utvecklingsarbete skiljer Nilstun på tre olika nivåer som kan vara föremål för utvärdering: planerings-, genomförande- och användningsnivån. Inom varje nivå kan man urskilja en process och ett resultat. På planeringsnivån är resultatet beslut om forskningsuppgifter, på genomförandenivån är det forskningsresultat och på användningsnivån är det förändringar som följd av forskningsresultatens tillämpning. M- och A-utvärderingarna är båda koncentrerade till genomförandenivån, men tangerar frågor som hör till planerings- och användningsnivån.

M-utvärderingen var i huvudsak en utvärdering av resultaten. Även A-utvärderingen handlade om resultat, men kom också att inriktas på forskningsprocessen och dess förutsättningar. En väsentlig skillnad är dock att M-utvärderingen gällde enskilda forskningsresultat (som tillsammans kan utgöra grund för generaliseringar) medan A-utvärderingen gällde det kollektiva resultatet av arkitekturskolornas forskning.

Så långt kan man säga att utvärderingsobjekten i M- respektive A-utvärderingarna är av i stort sett samma typ. Samma utvärderingsobjekt kan emellertid analyseras med olika utgångspunkter. Nilstun har angivit sex delvis överlappande analystyper som kan användas vid utvärdering: insats-, resultat-, process-, måluppfyllelse-, kontext- och effektivitetsanalys. I tabell 2 återger jag hans karakteristik av dessa analystyper<sup>2</sup> då de



tillämpas på forskning och utvecklingsarbete.

Analystyperna 2 och 5 har han i ett annat sammanhang kallat "effektanalys" respektive "orsaksanalys", vilket kanske förtydligar vad det är fråga om (Nilstun 1984, sid 14 ff). En fullständig utvärdering kräver att alla sex utvärderingsfrågorna besvaras, men av resursskäl måste i praktiken ofta ett urval göras. Resultatanalys, måluppfyllelseanalys och effektivitetsanalys är de analystyper som är primära vid utvärderingar, medan de övriga ses som kompletterande analystyper.

I både M- och A-utvärderingarna ingår måluppfyllelseanalys. Målen är också ungefär desamma om än något oprecist formulerade: forskningen ska vara av god kvalitet och forskningsresultaten ska vara nyttiga för

Analystyper	Frågeställningar
1. Insatsanalys	Vilka FoU-insatser genomfördes i syfte att förverkliga de önskade FoU-resultaten?
2. Resultatanalys	Vilka blev FoU-resultaten - oberoende av om dessa var önskade eller ej?
3. Processanalys	Genom vilka serier av händelser är FoU-insatserna relaterade till FoU-resultaten?
4. Måluppfyllelseanalys	Vilka är FoU-verksamhetens mål och hur förhåller sig resultaten till dessa mål?
5. Kontextanalys	Vilka externa faktorer gynnade, respektive hindrade att önskade FoU-resultat uppnåddes?
6 Effektivitetsanalys	Skulle andra input eller insatser kunnat ge önskade FoU-resultat på ett effektivare sätt?

Tabell 2. Sammanställning av de sex utvärderingsanalyserna med motsvarande utvärderingsfrågor för genomförandenivån i det sektoriella FoU-systemet (Efter Nilstun 1988, sid 22 ff och Nilstun 1989).

byggsektorn. För A-utvärderingen gäller detta åtminstone den forskning som finansieras av byggforskningsrådet. I båda fallen är det dock ett problem att kriterierna för att avgöra om dessa mål uppfylls är obefintliga eller oklara. En måluppfyllelseanalys på dessa grunder förutsätter konsensus om vad som är bra och nyttigt. Enligt min mening är dock detta ett mindre problem i A-utvärderingen där inte några enskilda projekt eller forskare utpekas. Där finns också en antydning till några mer precisa bedömningskriterier.

Varken i M- eller A-utvärderingarna förekom någon egentlig effektivitetsanalys. Man kan naturligtvis säga att i de fall då målen inte uppfylls är forskningen ineffektiv, men det är knappast detta som åsyftas med effektivitetsanalys. Det är svårt att göra en effektivitetsanalys av forskning. Idealt skulle man då behöva två likadana forskningsresultat som tillkommit på olika sätt och krävt olika resursinsatser. Detta är nog sällsynt om det alls förekommer. Vad som måste göras är i stället en rimlighetsbedömning av resursinsatsernas storlek i förhållande till forskningsresultatet. Någon sådan allmän bedömning förekom inte i M- och A-utvärderingarna.

Även när det gäller resultatanalysen kan man fråga sig om den är möjlig när det gäller forskning. Kan man få andra resultat än sådana som antingen uppfyller målen eller inte gör det? Här kan man dock tänka sig att ett forskningsproblem startas med syftet att behandla ett visst problem men att resultatet handlar om ett helt annat problem. Att konstatera detta är inte ett resultat av måluppfyllelseanalys om målen gäller forskningens kvalitet och nytta.

Vid M-utvärderingen konstaterades att man i flera projekt avvikit från projektplanerna. Hur detta värderats tycks dock ha varierat beroende på hur projekten bedömts med avseende på forskningens kvalitet och nytta. Vid M-utvärderingen konstaterades också att många resultat inte var forskningsresultat utan något annat som man kallade utredning, utvecklingsarbete m m. Följden av detta blev dock att målen vad gäller kvaliteten i genomförandet ändrades (på ett oklart sätt) med hänsyn till detta. Detta innebar att förutsättningarna för måluppfyllelseanalysen varierade mellan de granskade projekten.

A-utvärderingen utgick från att inte allt som kallades "forskningresultat" var forskningsresultat, att den viktiga skiljelinjen gick mellan forskning och utvecklingsarbete samt att resultaten av dessa verksamheter sannolikt borde bedömas på olika sätt. Om detta kan kallas resultatanalys, så utgjorde den en förutsättning för A-utvärderingen.

I både M- och A-utvärderingen saknades insatsanalys.

M- och A-utvärderingarna berörde också båda användningsnivån vad gäller informationen om och utnyttjandet av forskningsresultat. Detta gäller dock i mindre grad M-utvärderingen.

När det slutligen gäller process- och kontextanalysen finns det väsentliga skillnader mellan de två exemplen på utvärderingar.

I A-utvärderingen finns det åtminstone en uttalad vilja att tillämpa såväl process- som kontextanalys. En viktig del av denna utvärdering var att söka orsakerna till att resultaten inte alltid blev så bra. Det gällde då att finna både interna och externa faktorer (i förhållande till arkitekturskolornas forskning) som kunde tänkas orsaka kvalitetsbristerna. Det gällde då t ex också byggforskningsrådets (minimala) stöd till mera långsiktig forsk-

ning, UHÄs stöd till forskning (med hot om nedskärning) och reglerna för tjänstetillsättning mm.

I M-utvärderingen saknas både process- och kontextanalys. Visserligen berörs sådana frågor som handledning, forskningsmiljö och kompetens i enstaka projektgranskningar men någon allmän diskussion om vad sådana förhållanden betytt för de granskade projekten saknas. Ett undantag utgör möjligen försöken att jämföra forskning vid universitet och högskolor med konsulter forskning och förklara kvalitetsskillnader med hänsyn till dessa skilda miljöer. Som jag nämnt tidigare misslyckas dock detta förklaringsförsök: man förklarar något som M-utvärderingen inte visar. Genom att M-utvärderingen omfattade en granskning av projektplaner och annat beslutsunderlag som använts av byggforskningsrådet och byggforskningsinstitutet kan man säga att den till någon del kom att omfatta kontextanalys. Resultatet av denna blev en del kritiska anmärkningar mot byggforskningsrådet och byggforskningsinstitutet. Detta berör emellertid mera en utvärdering på planeringsnivån.

Med dessa utgångspunkter vill jag sammanfatta den viktigaste skillnaden mellan M- och A-utvärderingarna på följande sätt: A-utvärderingen är en fullständigare utvärdering än M-utvärderingen genom att den försöker förklara de brister som konstaterats.

Vad denna genomgång med utgångspunkt från Nilstuns problemram inte belyser är de grundläggande metodologiska skillnaderna mellan M- och A-utvärderingarna. Jag skall kommentera dessa i nästa avsnitt.

För att göra rättvisa åt M-utvärderingen vill jag här slutligen betona att den ingick som en del i en större utredning. Inom Mundeboutredningen fördes det samtal med bland andra administratörer vid byggforskningsrådet och byggforskningsinstitutet. Dessa samtal kan kanske delvis ses som kompletterande process- eller kontextanalyser. Hur de genomförts och vad de resulterat i är emellertid dåligt dokumenterat.

#### *Några kommentarer om typskillnader*

Det vore kanske möjligt att göra en utvärdering av dessa exempel på utvärderingar med hjälp av Nilstuns kategoriseringar och begreppsapparat, men både M- och A-utvärderingarna har så stora brister att en sådan fullständig metautvärdering knappast skulle tillföra särskilt mycket utöver vad jag redan sagt. Vidare har jag valt M- och A-utvärderingarna som exempel på två principer som båda är möjliga att utveckla och förbättra. När det gäller den grundprincip som tillämpades i M-utvärderingen, dvs expertgranskning av genomförda projekt, så har denna använts senare vid utvärderingar som gjorts av byggforskningsrådets vetenskapliga nämnd. I dessa har metodiken förfinats och analysen utvidgats i jämförelse med M-utvärderingen. Det är också ganska lätt att nu föreslå förbättringar av den

grundprincip som A-utvärderingen byggde på. En mer översiktlig jämförande utvärdering vill jag därför här begränsa till några principiella skillnader mellan utvärderingar av M-typ och av A-typ. (Med Nilstuns termer kommer jag att inrikta mig på process- och effektivitetsanalys.)

En viktig principskillnad gäller utvärderingsmetodiken. De data som utgör grunden för utvärderingar av M-typ fås genom att ett antal experter granskar ett antal avslutade (eller pågående) projekt. Samma experter bedömer också värdet av varje projekt. Experterna kan sägas utgöra det mätinstrument med vilket man avgör om ett projekt är bra eller dåligt, åtminstone om det saknas i förväg preciserade bedömningskriterier. För att resultatet av en sådan utvärdering ska vara riktigt krävs att ett annat urval av experter skulle komma fram till samma resultat. Detta är dock inte helt säkert. Även om man kan anta att det bland experter råder en viss konsensus om på vilka grunder forskningsresultat i allmänhet bör bedömas, är det troligt att man skulle göra olika bedömningar av enskilda projekt, vilket också M-utvärderingen visade. För att undvika detta bör det finnas någorlunda entydiga bedömningskriterier.

I utvärderingar av A-typ genereras data genom dialog mellan utvärderare och det forskarkollektiv som utvärderas. Även bedömningen av värdet görs i en sådan dialog och i en dialog med dem som kan tänkas använda forskningsresultaten. Att utvärderarna är väl insatta och har vissa förhandsuppfattningar är knappast någon nackdel, om de är beredda att överge sina uppfattningar då nya data kommer fram i dialogen. Mätinstrumentet i detta fall (när det gäller forskningens kvalitet) kan sägas vara dialogen och deltagarna i denna dialog. Utvärderingsresultatets riktighet beror här främst på om det skulle bli detsamma med andra utvärderare. Det är möjligt att valet av utvärderare spelar stor roll för utvärderingsresultatet. Även här skulle precisare bedömningskriterier sannolikt förbättra testbarheten.

Om vi för ett ögonblick betraktar enbart frågan om forskningens kvalitet, så vill jag påstå att resultaten av M-utvärderingen inte är säkrare än resultaten av A-utvärderingen. Det har inte bara att göra med brister i M-utvärderingens metodik vad gäller generaliserbarheten. Den skenbara exaktheten är kanske inte alls intressant. Det räcker med att konstatera att en (inte obetydlig) del av forskningen har allvarliga kvalitetsbrister, så som gjordes i A-utvärderingen. Något mer konstaterades inte i M-utvärderingen (om man inte är intresserad av enskilda projekt eller forskare).

I de senare utvärderingar av M-typ som genomförts finns det ansatser till att precisera bedömningskriterierna, att utvidga analysen till forskningsmiljön och att föra en dialog med de forskare som arbetat med de utvärderade projekten (Nilstun 1988, sid 65 ff). Vidare väljs begränsade forskningsområden inom vilka samtliga projekt under en viss tidsperiod utvärderas. Detta är givetvis förbättringar i förhållande till M-utvärde-

ringen, men i den mån man konstaterar kvalitetsbrister saknas det i stor utsträckning förklaringar till dessa.

Det är uppenbart att ett viktigt syfte med både M-utvärderingen och dess efterföljare varit att utgöra beslutsunderlag på planeringsnivån, dvs i detta fall beslut om åtgärder som byggforskningsrådet bör vidta. Mot den bakgrunden är det anmärkningsvärt att så litet intresse ägnats åt såväl insats- som kontextanalys.

M-utvärderingen var avsedd som underlag för Mundeboutredningens förslag men tycks ha givit mycket liten ledning. Även de senare utvärderingarna av M-typ avslutas med rekommendationer om åtgärder. Det är emellertid ofta svårt att förstå på vilket sätt dessa rekommendationer är relaterade till projektutvärderingarna. Rekommendationerna tycks i stor utsträckning grunda sig på oredovisade föreställningar om vad som orsakat brister och förtjänster hos de granskade projekten. Detta utesluter inte att rekommendationerna kan vara både välgrundade och värdefulla, men det innebär att man kan ifrågasätta nödvändigheten av projektutvärderingarna.

I en utvärdering av A-typ blir förslagen till åtgärder, enligt min mening, klarare relaterade till utvärderingen. Vid en dialog mellan utvärderare och utvärderade blir intresset större av att förklara brister hos forskningsresultat på andra sätt än enbart som en följd av forskarnas bristande kompetens.

Den metodik som användes vid M-utvärderingen har mycket stora likheter med den granskning av personer som görs vid tillsättning av akademiska tjänster och med den granskning som vetenskapliga tidskrifter ibland utsätter manuskript för. Det finns också vissa likheter med opposition vid seminariebehandling och disputation, om man kan tänka sig sådant med respondenten frånvarande. Mitt intryck är att dessa granskningsmetoder överfördes ganska okritiskt till ett sammanhang där de inte självklart passade. I senare utvärderingar av M-typ har det dock givits större utrymme för de utvärderade forskarna att försvara sig.

A-utvärderingens metodik har hämtat sina förebilder från andra områden än forskningens, utom vad gäller den kritiska dialogen.

Även om vi då inte var medvetna om det, tycks det finnas stora likheter med vad som görs vid s.k. organisationsutveckling. Det finns också likheter med traditionell utredningsmetodik och de metoder som ibland används vid verksamhetsrevision.

Jag ska också säga något om effektiviteten i de olika sätten att utvärdera forskning.

Utvärderingsobjektet för M-utvärderingen var väsentligt större än A-utvärderingens objekt, vilket torde motivera en större resursinsats. I senare utvärderingar av M-typ är utvärderingsobjekten väsentligt mindre än A-utvärderingens objekt, vilket då torde motivera mindre resursinsatser. Jag känner inte till kostnaderna för M-utvärderingen. Däremot finns det vissa



uppgifter om att de fem första BVN-utvärderingarna kostat mellan 190000 och 300000 kronor per utvärdering (Nilstun 1988, sid 75 ff).

Om vi bortser från kostnaderna för att ett stort antal forskare och andra personer ställer upp för diskussioner eller vid seminarier, så motsvarade resursinsatserna för A-utvärderingen ungefär fem manmånader. Det hade nog krävts något mera tid i förberedelseskedet för att undvika de största bristerna, kanske ett par månader. Expertbedömningar av M-typ, även om det bara gäller ett tiotal projekt, tycks dock kräva minst lika stora resursinsatser. Om resultaten av de senare är av större värde än det resultat som kan uppnås vid en utvärdering av A-typ vill jag överlåta åt läsaren att avgöra.

Utvärderingar av M-typ kan säkert utvecklas så att de ger en mycket detaljerad information om brister hos enskilda forskningsprojekt, men frågan är om sådana detaljkunskaper är särskilt intressanta som underlag för åtgärder på från anslagsgivande och forskningsplanerande organ. Med utvärderingar av A-typ får man inte veta mycket om enskilda projekt, men däremot något om frekventa brister och något generellt intressant om bristernas orsaker, vilket bör vara viktigast att veta på planeringsnivån.

Eftersom det gått några år sedan M- och A-utvärderingarna genomfördes bör det i princip vara möjligt att säga något om vilka faktiska konsekvenser de fått, dvs en utvärdering av utvärderingarna på användningsnivån. Detta är en forskningsuppgift för sig som säkert inte är särskilt lätt. Något skulle ändå kunna sägas om vilka förändringar som skett och på vilket sätt de kan tänkas vara relaterade till utvärderingarna. Någon sådan analys har jag dock inte haft för avsikt att göra här.

Mot bakgrund av vad jag sagt kan man spetsa till saken och fråga: Behövs det över huvud taget några utvärderingar av M-typ? Och vad ska de då användas till?

Utvärderingar av M- och A-typ kan kanske komplettera varandra på något bra sätt. I senare utvärderingar av M-typ har dialogen fått ett allt större utrymme. De kan därför ses som kombinationer av M- och A-ansatser. Utvärderingen av boendeforskningen är ett tydligt exempel på detta. Men frågan blir då: Hur stor vikt ska läggas vid den ena eller andra ansatsen för att uppnå det bästa resultatet?

#### *Om utvärderingen av boendeforskningen*

Utvärderingen av boendeforskningen (Properties of Housing – Conditions of Living 1988) avviker i flera avseenden från både M-utvärderingen och från senare utvärderingar av M-typ. Det förefaller också vara den största utvärdering som BVN har genomfört.

En väsentlig skillnad jämfört med tidigare utvärderingar är att expertbedömningen föregicks av en särskild studie av forskningen vid arkitekturskolorna (Kärholm & Wilhelmsen 1988). Denna kan, åtminstone delvis,



ses som en utvärdering av A-typ. Förutom skriftligt material och de båda författarnas egna erfarenheter av forskning vid arkitekturskolorna baserar sig denna studie på samtal och intervjuer med företrädare för denna forskning. Den berör alla tre nivåerna i FoU-systemet men tyngdpunkten ligger liksom i A-utvärderingen på genomförandenivån. Inom denna handlar det främst om process-, måluppfyllelse- och kontextanalyser och i någon mån resultatanalys.

I denna studie görs många insiktsfulla konstateranden om arkitekturforskningens förutsättningar, om olika problem och hinder för forskningens bedrivande och om forskningens förhållande till arkitektyrket och dess krav. Det hade förmodligen varit möjligt att dra mer långtgående slutsatser om vilka åtgärder som skulle ha behövt vidtas för att förbättra forskningen vid arkitekturskolorna. Kanske skulle detta ha krävt en något utvidgad studie och intensivare dialoger med både forskare och forskningsanvändare, men en sådan utvidgning skulle ha varit mycket måttlig jämfört med expertutvärderingen av boendeforskningen. Frågan är om inte en sådan utvidgning hade kunnat leda till i stort sett samma slutsatser och rekommendationer som expertutvärderingen?

Expertutvärderingen av boendeforskningen har koncentrerats till genomförandenivån och gäller främst resultat- och måluppfyllelseanalyser. I någon mån görs också (på ett allmänt plan) process- och kontextanalyser. Utvärderingsrapporten ger en ganska detaljerad beskrivning av tio "forskningsprogram" (forskargrupper) i termer av problem, angreppssätt, undersökningsmetoder och forskningsresultat. Beskrivningarna baseras på forskningsdokumentation (i allmänhet på svenska med engelsk sammanfattning), svar på en enkät och ett seminarium med varje forskargrupp. Dessutom har man samtalat med byggforskningsrådets tjänstemän. Med utgångspunkt från detta görs en del kritiska kommentarer i anslutning till varje forskningsprogram.

I rapporten sägs att man ser utvärderingen som en process. Därmed menas bl a att man vill sätta in forskningen och forskarna i ett större socialt sammanhang och föra en dialog med forskarna. Det är dock svårt att avgöra i vilken utsträckning man faktiskt levt upp till dessa ambitioner, men därigenom markerar man ett visst avståndstagande till utvärderingar av renodlad M-typ till förmån för A-typen. Själv uppfattar jag dock utvärderingen av boendeforskningen främst vara av M-typ. Men jag skall inte här gå in i någon detaljgranskning.

Utvärderingsrapporten är mycket bra om man vill ha en beskrivning av boendeforskningen vid arkitekturskolorna. Ansatsen att försöka relatera den svenska forskningen till en internationell standard är lovvärd. Rekommendationerna till byggforskningsrådet är också bra, även om de inte är särskilt originella och förmodligen kunde ha gjorts oberoende av expertut-

värderingen. Det problem jag ser är emellertid att den bild av svensk boendeforskning som ges kanske är alltför positiv.

En sak är att jag, med min kännedom om åtminstone vissa forskningsprogram och projekt, inte alltid delar experternas positiva uppfattning. Det kan man bortse ifrån. En annan sak är att experterna inte har haft möjlighet att göra vissa bedömningar.

Eftersom experterna i allmänhet har varit hänvisade till korta referat i översättning har man inte kunnat detaljgranska forskningsdokumentationen. Det innebär att bl a en del metodfrågor måste bli obesvarade. Har det t ex funnits tillräcklig evidens för de slutsatser som dragits i forskningsrapporterna? Är de data som använts eller insamlats tillförlitliga? Hur ser analysen av det empiriska materialet ut?

Även om de slutsatser som dras i en forskningsrapport är riktiga (eller överensstämmer med vad några experter tror är riktigt), så är det inte säkert att de är väl underbyggda (med några allmänt vedertagna kriterier på vetenskaplig underbyggnad).

Även om man (med utgångspunkt från sammanfattningar) finner att forskningsresultat är både nya och relevanta (även i ett internationellt perspektiv) är det inte säkert att de baseras på någon bra forskning.

Detta är ett dilemma för expertbedömningar med utländska experter så länge den mesta forskningsdokumentationen sker på svenska. Man kan inte både avgöra om forskningen är av god kvalitet (i alla avseenden) och samtidigt upprätthålla det internationella perspektivet (som tycks förutsätta internationella experter). Därmed kan man inte heller säkert säga något om vad som orsakar eventuella problem och vilka åtgärder som skulle behöva vidtas för att undanröja dessa.

När det gäller utvärderingen av boendeforskningen överdriver jag möjligen de svårigheter experterna kan ha haft. I vissa fall tycks man ha genomskådat de korta beskrivningarna och funnit intressanta teori- och metodproblem. Det är emellertid en mycket försiktig kritik som framförs. Och det är nog bra. Det finns säkert en insikt hos utvärderarna om att deras uppgift inte i första hand är att peka ut och kritisera enskilda projekt eller personer. Den grundläggande kvalitetskontrollen är en uppgift för forskarkollektiven själva. Genom en utvecklad intern kritik med seminarier och kvalificerad handledning bör det finnas en garanti för att elementära krav på vetenskaplighet vad gäller t ex begriplighet, testbarhet och trovärdighet upprätthålls. Om inte detta fungerar är det en brist i forskarmiljön. Felet behöver alltså inte vara den enskilde forskarens.

Utvärderingen av boendeforskningen vid arkitekturskolorna är givetvis något positivt för boendeforskarna. Tvärtom vad boendeforskningens kritiker brukar hävda har ett antal ansedda utländska forskare funnit att man i stort sett sysslar med vettiga problem och att man kommit ganska

långt jämfört med många andra länder. Vad man inte tycks ha upptäckt är dock att de olika forskargrupperna sinsemellan är mycket olika vad gäller problemuppfattning och vetenskapsideal. I grunden finns en mycket kritisk inställning från den ena gruppen mot den andra.

Om experterna ägnat mera tid åt den ansats de själva kallat "utvärderingen som process" i stället för att granska enskilda program, projekt, grupper och personer så hade man kanske fått en annan och delvis mera nyanserad bild av boendeforskningen vid arkitekturskolorna. Med detta menar jag att dialogen kunde ha utvidgats med både fler deltagare och till att i samtalen med de utvärderade grupperna omfatta även annan forskning än den egna. Utvärderingen hade då närmast sig en utvärdering av A-typ. De allmänna kommentarerna och rekommendationerna hade kanske blivit desamma, men kanske också intressantare och bättre underbyggda. Och jag tycker inte att det är helt oberättigat att upprepa frågan: Behövs det över huvud taget några utvärderingar av M-typ?

#### Noter

1. Det är ofta viktigt att betona skillnaden mellan forskning, utvecklingsarbete, utredning mm. Om man vill ha någon samlingsbeteckning för dessa olika verksamheter kan man som bl.a Tore Nilstun envist hålla fast vid "FoU" i stället för "forskning". Av främst stilistiska skäl har jag emellertid här använt "forskning" som samlingsbeteckning. Detta överensstämmer också med språkbruket i de utvärderingar jag diskuterar.
2. I en tidigare version av min uppsats citerade jag en opublicerad uppsats av Tore Nilstun. Denna finns något nedkortad och bearbetad som artikel i detta nummer av *Arkitekturforskning*. De olika analystyperna finns också redovisade i Nilstun 1988. Jag har dock bibehållit de ursprungligen citerade formuleringarna av analystyperna eftersom de passar mina syften bäst.

#### Referenser

- Bostadsdepartementet 1982, *Utvärdering av byggforskningsprojekt: Delrapport från byggforskningsutredningen*.
- Ds Bo 1982:2, *Byggforskningen - en översyn och utvärdering: Slutrapport från byggforskningsutredningen*.
- Eriksson, J och Wikfors, Ö 1983, *Arkitekturforskning*. Byggforskningsrådet G7:1983.
- Eriksson, J 1986, "Byggforskningens samhällsrelevans" i B Sollbe (red), *Vishets frukter. Uppsatser om sektorsforskning och forskningskvalitet*. Statens institut för byggnadsforskning.
- Kärrholm, G och Wilhelmsen, A M 1988 *Forskning vid arkitekturskolorna*. BVN Skriftserie 1988:2.
- Nilstun, T 1986, "Mundebo och Fosolifen" i B Sollbe (red), *Vishets frukter. Uppsatser om sektorsforskning och forskningskvalitet*. Statens institut för byggnadsforskning.

- Nilstun, T 1988, *Expertbedömningar. Om teori, ideal och verklighet vid utvärdering av sektors FoU*. BVN Skriftserie 1988:1.
- Nilstun, T 1989, "Utvärdering av sektors FoU" i *Arkitekturforskning*, Vol 2, nr 3, 1989.
- Nilstun, T och Hermeren, G 1984, *Utvärderingsforskning och rättsliga reformer: Analys av orsaker och effekter*. Studentlitteratur.
- Properties of Housing - Conditions of Living. An International Evaluation Study*.  
BVN Skriftserie 1988:5.
- UHÄ-rapport 1981:15, *Arkitektutbildning - dimensionering och lokalisering*.
- UHÄ-rapport 1983:7, *Arkitektur: Utbildning och forskning*.
- (Byggforskningsrådets vetenskapliga nämnds utvärderingar har publicerats i BVN Skriftserie 1985:1, 1985:2, 1985:3, 1985:4, 1985:5, 1988:3, 1988:4 och 1988:5.)