

Allmän SiS-rapport  
1999:4

# Kausalfrågor och effektfrågor vid utvärdering av behandling

Av Bengt-Åke Armelius och Kerstin Armelius

# **Kausalfrågor och effektfrågor vid utvärdering av behandling**

## **Bengt-Åke Armelius och Kerstin Armelius**

### **Rapport 1999:4**

#### **Förord**

Statens institutionsstyrelse, SiS, inrättades 1993 och svarar för planering, ledning och drift av institutioner för tvångsvård av missbrukare och särskilda ungdomshem. En viktig uppgift är att följa upp och utveckla vården samt att initiera och stödja forskning inom ungdoms- och missbrukarvård. SiS FoU-enhet leder detta arbete. Upp till 30 miljoner kronor får årligen avsättas för forskning, utveckling och utbildning av personal.

FoU-enheten tar initiativ till forskning genom att årligen erbjuda ett ansökningstillfälle för forskning av relevans för SiS inom prioriterade områden som anges i ett särskilt forskningsprogram. Till SiS finns knutet ett vetenskapligt råd som biträder FoU-enheten i vissa principiella frågor samt vid bedömning av forskningsprojekt och större utvecklingsprojekt. FoU-enheten verkar för spridning av resultaten från forsknings- och utvecklingsprojekten genom publicering av kunskapsöversikter och rapporter.

Föreliggande rapport är ett led i utvecklingen av metoder för utvärdering och kvalitetssäkring av verksamheten på de särskilda ungdomshemmen och LVM-hemmen inom SiS. Rapporten är tänkt att tjäna som underlag för diskussion av forsknings- och utvärderingsfrågor när det gäller behandlingsinsatser. Författarna går igenom för syftet tänkbara designar för datainsamling och dataanalys, olika strategier för statistisk kontroll och frågor kring validitet och reliabilitet vid mätning av behandlingseffekter. De diskuterar de svårigheter som är förknippade med utvärdering av den typ av verksamhet som bedrivs inom SiS. Författarna pekar på problemet att isolera effekten av den behandling som klienterna får på SiS institutioner. I rapporten presenteras ett förslag till en modell för kvalitetssäkring av behandlingen inom SiS.

Rapporten är författad av Bengt-Åke Armelius, tillika medlem i SiS Vetenskapliga råd, och Kerstin Armelius, verksam vid Institutionen för Tillämpad psykologi, Umeå universitet, och med 20-årig erfarenhet av forskning om utveckling och utvärdering av psykoterapeutiska vårdprogram inom psykiatrisk vård.

Vera Segraeus FoU-direktör

#### **Innehållsförteckning**

Förord

Innehållsförteckning

Bakgrund

Etiska överväganden

Kausalfrågor och effektfrågor

Designar för datainsamling och dataanalys

Design och metod

Datainsamling

Validitet och design

Intern validitet

Begreppsvaliditet

Extern validitet

Några kvasiexperimentella designar

Ekologisk design

Prepost design med jämförelsegrupp

Institutionscykeldesign

Tidsseriedesign

Dataanalys

Kvalitetssäkring av verksamheter

Kausalfrågor och effektfrågor

Kausalfrågor i verksamheter

Effektfrågor i verksamheter

Utvärdering inom SiS: En modifierad kvalitetssäkringsmodell

Diskussion av utvärderingsfrågor inom SiS

Nivå 1 - Har vården vid SiS någon effekt?

Nivå 2 - en behandlingssideologi/teori inom SiS

Nivå 3 - en specifik behandlingsinsats

Sammanfattning och slutsatser

Referenser

## Bakgrund

Denna skrift är avsedd att tjäna som underlag för diskussion av forsknings- och utvärderingsfrågor gällande behandlingsinsatser. Resonemangen har giltighet för alla verksamheter som sysslar med behandling, men i denna skrift kommer vi att fokusera på behandling inom Statens institutionsstyrelse, SiS. Bakgrunden är att utvärdering och forskning kan ses som en naturlig och nödvändig del av verksamheten inom SiS. Inom vetenskapliga rådet vid SiS har frågan om utvärderingsdesigner diskuterats många gånger och det finns en allmän uppfattning att man borde kunna genomföra utvärderingar som motsvarar högt ställda krav på vetenskaplighet, inte minst mot bakgrund av att det rör sig om tvångsvård som alltså motiverar ett samhällsingripande i syfte att ge den enskilde vård som han inte bitt om.

Inom SiS bedrivs dels LVM-vård för vuxna missbrukare, dels LVU-vård för ungdomar med sociala problem av olika slag. LVM-vården omfattar upp till 6 månader och har som primärt mål att avgifta personen och motivera missbrukaren till fortsatt vård för sitt missbruk. Den fortsatta vården kan ske på SiS institutionen men sker oftast genom socialtjänstens placering hos andra vårdgivare. LVU-vården kan antingen utgöra ett omhändertagande i syfte att utreda ungdomens sociala och psykologiska situation under maximalt 2 månader eller ett erbjudande om behandling, vilket kan omfatta flera års behandlingstid vid SiS-institutionen.

En mycket påtaglig anledning till att dessa frågor har fått aktualitet är att ett behandlingshem (S) har för avsikt att införa en speciell variant av behandling (Kognitiv Beteende Terapi, KBT) och vill ha en utvärdering av dess effekter genomförd av SiS. Det är en behandling med många olika behandlingsinsatser såsom ilskekontroll, social färdighetsträning m m som skall införas för ungdomar med kriminell belastning. Den frågeställning som uppstår är om denna behandlingsform är bättre än andra behandlingsformer för dessa ungdomar.

### Etiska överväganden

Det är väsentligt att etiska överväganden klargörs i samband med forskning på SiS klienter och personal. För det första gäller att forskningen skall underordna sig den kliniska verksamheten. Den kommer i andra hand och får inte skada den kliniska verksamheten. För det andra gäller att, liksom i all forskningsverksamhet, forskningskravet skall vägas mot individskyddskravet. Det betyder att forskningsintresset - att vi/samhället får veta hur det är - vägs mot vars och ens rätt till skydd av sin integritet och att delta i forskning är frivilligt. I de flesta sammanhang kan individskyddskravet sägas väga relativt tungt och forskningskravet får stå tillbaka. Det är dock värt att diskutera denna balans i SiS verksamhet eftersom man arbetar med människor som vårdas under tvång. Det gör kanske att forskningskravet måste väga särskilt tungt. Man kan inte erbjuda tvångsvård vars effekter man inte är absolut säker på. Å andra sidan kan man hävda samma sak för individskyddskravet. Människor som redan är tvångsomhändertagna skall inte behöva utsättas för ytterligare integritetskränkningar därför att man vill ha forskningsinformation. Det är viktigt att det blir en klar policy inom SiS avseende de forskningsetiska huvudfrågorna, men detta är en fråga som faller utanför syftet med denna skrift.

## Kausalfrågor och effektfrågor

Vi anser att det är meningsfullt att skilja mellan två principiellt olika, men relaterade, kunskapsintressen vid utvärderingsforskning. Det första knyts till kausalfrågor och gäller att så klart som möjligt avgöra om en viss faktor (X) orsakar en viss effekt (Y). Detta är den klassiska frågan om orsak och verkan och den kan logiskt besvaras med experimentella och kvasiexperimentella designar där man systematiskt varierar en eller flera oberoende variabler, X, och håller alla andra tänkbara förklaringsvariabler, Z, under kontroll för att se vad som händer med den beroende variabeln, Y. Om man kan arrangera sin datainsamling enligt en sådan design och inga avgörande hot i form av andra förklaringsvariabler finns mot designen så kan man få svar på frågan om X orsakar Y. Det andra kunskapsintresset knyts till effektfrågor och gäller samband mellan effekter (Y) och alla förklaringsvariabler som är av intresse för att förklara variation i Y. I designtermer kan man säga att kausalfrågor fokuserar på att isolera X från Z och på så sätt förklara hur effekten i Y uppkommer, medan effektfrågor fokuserar på att förklara variation i Y med hjälp av relevanta förklaringsvariabler, d v s både X och Z används för att förklara varians i Y.

I utvärderingsforskning utgörs förklaringsvariablerna i huvudsak av behandlingsvariabler och variabler knutna till klienterna, t ex diagnostiska eller demografiska variabler, eller variabler knutna till själva behandlingsorganisationen eller behandlingsmiljön. Behandlingsvariablerna utgör tänkbara kausala variabler i den meningen att de anger hur orsakerna uppkommer, medan förklaringsvariabler knutna till klienter eller behandlingsorganisation inte kan sägas vara kausala i samma mening. De kan dock mycket väl bidra till att förklara en viss del av variansen i Y.

Det är också värt att notera att det finns en logisk ordning mellan kausalfrågor och effektfrågor. Om man inte har noterat någon effekt, d v s om det inte finns någon varians i Y, så finns inget att förklara. Därför bör man börja med utförliga effektanalyser för att så småningom komma fram till bra kausalfrågor. Möjligen kan man uppskatta vilket stöd olika kausala tolkningar har i befintliga data. Kanske kan man säga att på samma sätt som kausal-samband utgör en delmängd av alla samband så utgör kausalfrågor en delmängd av alla effektfrågor.

## Designar för datainsamling och dataanalys

För att kunna diskutera utvärderingsfrågor är det nödvändigt att tala om design och metod. Eftersom dessa begrepp har många olika innebörder har vi valt att här ge några definitioner som vi vill använda i denna skrift. Valet av såväl design som metod bör idealt styras av den frågeställning som forskaren har, men omvänt gäller också att tillgång till vissa metoder och möjligheterna att genomföra vissa designar skapar förutsättningar för att ställa frågor som tidigare inte gått att besvara på ett vetenskapligt acceptabelt sätt. Ett exempel är tillgången på små persondatorer som revolutionerat forskarvärlden. Vi menar att det inte finns någon enskild design eller metod som generellt kan sägas vara överlägsen, vetenskaplig eller bättre än andra. För oss är vetenskaplighet snarast att betrakta som ett reflekterat och systematiskt förhållningssätt till de problemställningar man tar sig an. Ibland kan man då åstadkomma mer tillförlitliga slutsatser därför att man lyckats samla in och analysera sina data på ett sätt som klargör eller utesluter olika alternativa tolkningar, men ibland kan man inte åstadkomma sådana förutsättningar och då blir kunskapen av begränsat värde.

### Design och metod

Med design menar vi ett sätt att organisera informationen så att det blir möjligt att besvara de frågeställningar som undersökningen har. Designen anger vad som skall få variera och vilken variation som skall kontrolleras på något sätt. En förutsättning för att data som samlas in skall bli meningsfulla är att de kan ställas emot varandra så att intressanta samband eller skillnader kan lyftas fram. Man kan välja att organisera sina data på många olika sätt, men det är bara vissa jämförelser som är av intresse för att ge svar på frågeställningen. På samma sätt kan man resonera när data väl är insamlade, men då gäller det att organisera sina data för att kunna analysera dem. Det finns inte bara ett sätt att analysera data efter det att man samlat in dem och därför är det meningsfullt att skilja mellan om designbegreppet avser en organisation för datainsamling eller en organisation för dataanalys.

Med metod menar vi hur man samlar in eller analyserar data. Datainsamlingsmetoden kan vara test, frågeformulär, intervjuer, observationer, utdrag ur register etc. De kan vara strukturerade utifrån variabler eller ostrukturerade som t ex utskrift av terapeutiska samtal. Vanligtvis tänker vi på olika statistiska metoder när det gäller metoder för dataanalys, men detta förutsätter att data organiserats som variabler. Om data samlas in i ostrukturerad form används oftast en form av kvalitativ metod för att få fram teman eller innebörder. Varje metod kan karaktäriseras med avseende på tillförlitlighet (reliabilitet) och relevans (validitet). I denna skrift

kommer vi inte att ta upp metodfrågor utan koncentrerar oss på designfrågor för datainsamling. Vi kommer inte heller att gå in på design för dataanalys vilket är ett snabbt växande område.

För att besvara kausala frågor är designen för datainsamling mycket central och det krävs att man har kontroll över alla betingelser vid datainsamlingen som kan påverka resultaten. Vad detta innebär skall vi återkomma till under frågan om intern validitet, men låt oss här kalla det för att ha kontroll över datainsamlingen. För att besvara effektfrågor är det nödvändigt att ha tillgång till variationen i alla variabler som kan bidra till att förklara varians i Y. Det innebär att vi i efterhand, med dataanalytiska design, kan undersöka hur variationen i våra variabler hänger ihop. Låt oss kalla detta för statistisk kontroll.

Såväl metod som design kan alltså sägas handla om både insamling och analys av data. Hur vi tänker kring detta beskrivs i den fyrfältstabell som återges i tabell 1.

**Tabell 1.** Design och metod för datainsamling och dataanalys

	<b>Datainsamling</b>	<b>Dataanalys</b>
<b>Design</b>	Kontroll över datainsamlingen	Statistisk kontroll
<b>Metod</b>	Test, frågeformulär, intervjuer, observationer, register	<i>Statistik:</i> medelvärdestest, korrelationer, LISREL, PLS <i>Kvalitativ:</i> diskursanalys, narrativ analys, innehållsanalys, grounded theory

## Datainsamling

När det gäller design för datainsamling utgår vi från Cook och Campbells (1979) klassiska bok om kvasiexperimentella design och analys av fältundersökningar. Vi börjar med en kort sammanfattning av huvuddragen i deras bok.

### Validitet och design

Logiken när man undersöker kausala frågor med experimentell design är att arrangera datainsamlingen så att det enda som kan förklara Y är orsaken X. Idealet kan återges så här:

	Andra förklaringsvariabler		
Experimentgrupp	Z	X	Y
Kontrollgrupp	Z	ej X	Y

Designen förutsätter att man har två grupper som endast skiljer sig i oberoende variabeln definierad som X och som är representativa för den population man vill studera. Det betyder att de skall vara lika i alla andra förklaringsvariabler definierade som Z. En del andra förklaringsvariabler gäller omständigheter före själva införandet av X, andra kan inträffa parallellt med X eller den beroende variabeln, Y. Om alla andra förklaringsvariabler är lika i de två grupperna beror skillnaderna i beroende variabeln, definierad som Y, på att experimentgruppen utsatts för X medan kontrollgruppen inte gjort det. Andra förklaringsvariabler kontrolleras med olika tekniker som konstanthållning, eliminering, matchning och randomisering. Förebilden är ett laboratorium där man isolerar orsaken och har fullständig kontroll över omständigheterna där X administreras och Y mäts. Konstanthållning och eliminering tar bort varians, vilket innebär att man inte kan genomföra en analys av hur samvariationen med Y ser ut. Har man t ex studerat endast kvinnor i en studie så vet man inget om hur betydelsefull variabeln kön är för utfallet. Om man randomiserar en (oändligt) stor mängd försöksobjekt, dvs låter slumpen avgöra vem som hamnar i vilken grupp så har vi en experimentell design. Randomiseringen likställer grupperna med avseende på alla Z som inträffat före X och laboratoriekontrollen likställer dem med avseende på alla Z som inträffar därefter.

Eftersom denna situation sällan eller aldrig föreligger i samhällsvetenskaper är grundtanken hos Cook och Campbell att diskutera design för orsaksstudier när experimentella design är omöjliga. Dessa design kallas kvasiexperimentella eftersom logiken som används bygger på experimentella design. För att värdera olika undersökningsdesign använder Cook och Campbell sig av tre olika validitetsbegrepp: intern validitet anger vilken kontroll man har av andra möjliga orsaker, begreppsvaliditet anger hur väl man definierat X och Y, och extern validitet avser hur generaliserbar kunskapen är. Dessa validitetsbegrepp är centrala för att diskutera värdet

av designer för såväl datainsamling som dataanalys och därför ska vi gå lite närmare in på innebörden av de olika validitetsbegreppen.

### Intern validitet

De variabler eller faktorer som utgör andra förklaringsvariabler kallas hot mot intern validitet. Hot mot intern validitet är alltså alla Z som kan finnas med i en undersökning och sammanblandas med X. På nästa sida ges några exempel på de viktigaste hoten mot intern validitet som Cook och Campbell diskuterar.

Z = Mognad - man tror att X ger effekt på Y men det är en effekt av mognad eller naturliga förlopp t ex att ungdomar blir mer lugna i vissa åldrar.

Z = Historia - det inträffar något annat samtidigt med X som ger effekter i Y, t ex att man ger en avsevärd bidragsförbättring till klienterna samtidigt som man genomför en studie av effekter av behandlingen X.

Z = Observationseffekter och försöksledareffekter - olika socialpsykologiska fenomen som t ex.

uppmärksamhetseffekter d v s Y återspeglar förbättringar beroende på att man är med i en undersökning, eller effekter av engagemang och förväntningar hos personal och forskare, s k "allegiance".

Z = Statistiska effekter - regression mot medelvärdet, d v s beroende på reliabilitetsbrister i Y kommer en förnyad mätning av Y alltid att uppvisa en regression mot medelvärdet så att de som var bäst före blir sämre och de som var sämst blir bättre. Tak- och golfeffekter hör också hit.

Z = Selektion - personerna är ovanliga eller olika redan från början, t ex om bara de som var motiverade för behandling tillfrågades om de ville vara med.

Z = Bortfall - t.ex. om många personer slutar, flyttar, dör etc och endast de som mår bra av behandlingen blir kvar.

Om man endast har en grupp av personer att studera i en enkel före-efter design så blir nästan alla ovan uppräknade variabler aktuella som hot mot den interna validiteten i kausalfrågan huruvida X orsakat Y. Utan tillgång till jämförelsematerial, och om det inte råkar finnas en stor variation i Z, blir det också svårt att statistiskt kontrollera variansen i Y för inflytanden av Z så även effektfrågor blir i det närmaste omöjliga att undersöka på ett tillfredställande sätt. En jämförelsegrupp är således av stort värde för att förbättra designen även om man inte kan randomisera.

Att randomisera är ett effektivt sätt att jämställa grupperna innan X inträffar. Förutsättningen är naturligtvis att grupperna är tillräckligt stora för att slumpens lagar skall kunna antas verka. Denna kontrollåtgärd får ibland status som det viktigaste och avgörande och det kan därför vara skäl att fråga sig vilka hot mot intern validitet som kan kontrolleras med randomisering av personer till grupper. Randomisering kontrollerar för selektion och mognad och vissa statistiska effekter medan historia, observations-/försöksledareffekter och bortfall inte kontrolleras utan måste kontrolleras på annat sätt t ex med konstanthållande eller balansering. Däremot återstår alla de hot som har att göra med övriga Z (historia, bortfall, observations-/försöksledareffekter, statistiska effekter). En slutsats blir därför att randomisering av personer har ett begränsat värde och detta gäller särskilt i longitudinella studier. I en del sammanhang kan finnas ett motstånd mot randomisering, vilket då blir ett eget hot mot den interna validiteten t ex därför att de inblandade inte låter alla tänkbara klienter medverka i studien eller ordnar med kompletterande behandling.

I tabell 2 kan vi se vilka hot som en randomisering av personer tar hand om och vilka som måste hanteras oavsett om man randomiserar eller ej.

### Tabell 2. Hot (Z) som inträffar före och parallellt med X och Y

Hot före X	Hot parallellt med X och Y
Selektion, mognad, vissa statistiska effekter	Historia, observationseffekter, bortfall, statistiska effekter

Cook och Campbell argumenterar för att man även i de fall man inte kan randomisera kan komma långt i sina orsakstolkningar med olika kvasiexperimentella designer. Man kan med olika slags tekniker i viss grad kontrollera många hot. Tex. kan selektionsfaktorn och mognad kontrolleras genom att matcha personer. Om man inte kan kontrollera (jämställa) övriga Z-variabler kan man kanske kartlägga och sedan relatera dem till X och Y för att se om de utgör alternativa orsaker.

## Begreppsvaliditet

Vi har hittills talat om orsaksbegreppet i anslutning till kausalfrågor i allmänna ordalag. Det är ett svårt begrepp med många filosofiska rötter som vi inte kan reda ut här. En liten nyansering är emellertid nödvändig. Med begreppet orsaksmekanism brukar vi avse en beskrivning av hur en händelse kan orsaka en annan händelse på ett sätt som är begripligt för oss mot bakgrund av vår allmänna förståelse av hur saker och ting hänger ihop eller mot bakgrund av en mer genomarbetad teori. I bästa fall är det bara en variabel som är orsaken till att Y inträffar, men det kan också röra sig om en serie händelser eller ett orsaksmönster där man inte är riktigt klar över varje enskild variabels betydelse. Hur pass precis man är i sin ambition att isolera exakt vad som åstadkommit en viss effekt är avhängigt av dels hur kunskapsläget är inom området och vilket intresse man själv har som forskare.

Om X är komplex eller sammansatt av flera olika komponenter ger inte en kontrollerad design svar på frågan vad som orsakat en förändring i Y. I den typ av verksamhet som SiS sysslar med är behandlingen (X) oftast komplex och svår att definiera vilket försvårar studier av kausala mekanismer. Det hindrar emellertid inte att man kan ställa sig frågan om en viss behandling eller till och med en hel verksamhet kan ha en viss effekt eller inte. Vi har då glidit över till en effektfråga där de kausala mekanismerna inte längre är huvudsaken utan intresset knyts till att förklara behandlingseffekter med att man deltagit i en komplex verksamhet. Inte ens den mest väl designade studie med randomisering och kontroll av alla betingelser kan ge ett bättre svar på kausalfrågor än vad validiteten i X medger. Om vi inte vet vilka verksamma ingredienser behandlingen består av får vi inte veta särskilt mycket ens i en väl genomförd kontrollerad studie. Det kan då vara klokt att försöka definiera X bättre innan man lägger ner alltför mycket möda på att genomföra kostsamma och krävande kontrollerade studier.

En mycket väsentlig faktor för att försäkra sig om hög validitet i begreppet behandling är att förvissa sig om att behandling faktiskt har ägt rum. Det kan liknas vid att i läkemedelsstudier förvissa sig om att patienterna faktiskt tagit sin medicin och inte nöja sig med att räkna hur många tabletter man delat ut! Andra viktiga frågor att besvara för att få hög begreppsvaliditet är hur mycket behandling av olika slag klienterna har deltagit i och med vilket engagemang, etc? En annan viktig fråga gäller karaktären och kvaliteten på den behandling som givits. Hur kan den beskrivas? Har man genomfört behandlingen så som manualen föreskriver?

## Extern validitet

Cook och Campbell tar upp generalisering under rubriken extern validitet för att markera att frågan gäller hur användbar kunskapen är i andra sammanhang än de man studerat. Den viktigaste frågan för extern validitet gäller hur man valt ut sina undersökningspersoner. Om man inte studerar en totalpopulation måste man med hjälp av särskilda urvalsmetoder (både randomiserade och systematiska) åstadkomma ett representativt urval av personer som man studerar. I denna skrift utgår vi från att urvalet redan genomförts enligt någon systematisk metod. En annan viktig faktor när det gäller extern validitet är hur pass stort bortfallet varit. Det finns flera stora internationella behandlingsstudier som inte kunnat bidra med användbara kunskaper därför att bortfallet varit allt för stort. Ett exempel är den s k Bostonstudien där man i en randomiserad design skulle svara på frågan om dynamisk psykoterapi var bra för schizofrena patienter eller ej. Man började med 180 patienter men slutade med drygt 50 efter två år. Bortfallet var således mycket stort och gjorde såväl tolkningar av gruppskillnader som generaliseringar till verkliga kliniska förhållanden närmast omöjliga.

Man kan kortfattat säga att den externa validiteten beror på hur pass representativa de individer och betingelser man studerat är. Man kan tänka på undersökningen som ett stickprov ur en population. Stickprovet gäller såväl personer som ingår i undersökningen som värdena på X, Y och Z. Ju mer representativa undersökningsbetingelserna är för den situation eller grupp man vill generalisera till, desto större är den externa validiteten. Av detta följer att ju mer man begränsar och kontrollerar den naturliga variationen, desto lägre är i allmänhet generaliserbarheten. Studier designade för kausalfrågor, med hög grad av kontroll och begränsning av variation i form av konstanthållning av t.ex. kön, diagnos, behandlingssinnehåll, terapeutkompetens m m, har lägre extern validitet än studier där man studerar variationen i variabler, vilket är fullt möjligt och önskvärt vid designer upplagda för studier av effektfrågor.

En intressant debatt förs inom psykoterapiforskningen där man diskuterar värdet av välkontrollerade jämförelser, s k "efficacy"-studier, mellan t ex dynamisk psykoterapi och kognitiv psykoterapi. Man renodlar betingelserna så långt som möjligt genom att randomisera patienter till en selekterad grupp av terapeuter som tränats att följa en manual via flera videospelade fall. Alla patienter får lika många timmar och lika mycket handledning ges till alla terapeuter. Samma mätmetoder används i alla betingelser. Det enda som skall skilja är vilken terapeutisk metod som används. Resultaten kan tolkas som uttryck för en specifik behandlingseffekt. Värdet av dessa studier ligger i att man kan se hur den maximala effekten ser ut. Praktiskt är det dock begränsat därför att resultaten inte är direkt generaliserbara till de förhållanden som råder i verkligheten. Kritikerna av "efficacy"-studier menar att

det kanske inte alls går att replikera dessa terapier i den kliniska vardagen där terapeuterna har olika kompetens, terapitiderna varierar etc och förordar därför s k "efficiency"-studier, d v s studier av vilken effekt en viss behandlingsmodell har i en verklig klinisk miljö. Hur mycket av de maximala effekterna kan man förvänta sig att uppnå i en klinisk miljö där man inte har kontroll över särskilt många betingelser utan låter dem variera på ett representativt sätt? Det är inte säkert att det ger samma resultat som "efficacy"-studier.

Låt oss sammanfatta genomgången av Cook och Campbells validitetsbegrepp med hjälp av tabell 3.

**Tabell 3.** Sammanfattning av Cook och Campbells validitetsbegrepp

	Validitetstyp		
	Intern	Begreppslig	Extern
Hög	Experimentell design	En teoretiskt definierad X	Representativ design
Låg	En grupp utan jämförelser	Komplex eller många X	Laboratoriedesign

Syftet med validitetsgenomgången var att fördjupa frågan om design för utvärdering. Slutsatserna blir:

1. Vi kan åstadkomma varierande grad av kontroll över undersökningsbetingelserna, X och Z genom att ha en bra design för datainsamling. Med bra design menar vi att vi har kontroll över variationen i andra förklaringsvariabler genom att randomisera, balansera och konstanthålla variationen eller att registrera hur variationen ser ut för att så småningom kunna relatera den till X och Y.
2. Definitionen av X är avgörande för vilken möjlighet vi har att besvara kausalfrågor. Ju mer avgränsad och meningsfull X är, desto säkrare kausala slutsatser kan vi dra. Vid effektfrågor är det nödvändigt att beskriva de olika komponenterna i X så utförligt som möjligt för varje individ för att kunna relatera dem till Y.
3. Värdet av undersökningen är ofta beroende av att vi kan generalisera till andra situationer än den som undersökts. Ibland dominerar kontrollkraven medan generaliserbarheten får komma i andra hand, men ibland kan generaliserbarheten vara viktigare och man får kompromissa med kontrollkraven.

### Några kvasiexperimentella design

Låt oss titta på några kvasiexperimentella design som kan vara av intresse för SiS.

#### Ekologisk design

Låt oss kalla den genomsnittliga behandlingen vid SiS för Treatment As Usual (TAU). Det är kanske ett olyckligt begrepp, men det är användbart för designdiskussionen. Bakom detta stora medelvärde ligger en fördelning av enheter med olika behandlingseffekt. Det intressanta är att vi kan få kunskap om denna fördelning genom att SiS har en begränsad mängd institutioner där vi skulle kunna genomföra en gemensam mätning av effekter (Y) av behandlingsinsatserna. Man skulle kunna säga att vi har definierat SiS domän eller population och vi har fått kunskap om variationen i SiS verksamhet. Det är mot denna helhetsbild som nya behandlingsinsatser i första hand skall jämföras, inte mot något enstaka alternativ.

G1:	O1	X	O2
-----			
ALLA:	O1	TAU	O2

O1 står för observationstillfälle 1 o s v och X för behandling. Naturligtvis kan bilden nyanseras genom att skilja mellan män och kvinnor, unga och gamla, korta och långa vårdtider etc i takt med att man får in kunskap om hela populationen. Detta är i själva verket en variant av den gamla testpsykologins normtänkande men nu applicerat på prognosnormer om vad som skall hända med en viss grupp klienter som får en viss behandling. Liksom i de flesta kvasiexperimentella design spelar Selektionsfaktorn en stor roll. Vi vet inte om klienterna är olika från början, men förmätningen O1 kan användas för att få en uppfattning om så är fallet eller ej. Många andra hot mot kausaltolkningen att X orsakar Y finns också, vilket gör att man måste undersöka om de är aktuella i varje specifikt fall. Styrkan i denna design ligger främst i den externa validiteten.

#### Prepost design med jämförelsegrupp

Detta är en enkel före-efter design där man mäter samma personer före och efter behandling vid två eller flera enheter. Jämförelsen kan ske dels för varje grupp, dels som en differens före-efter mellan grupper. Även här kan vi se att Selektionsfaktorn utgör ett allvarligt hot



G1:	O1	X1	O2
-----			
G2:	O1	X2	O2

### Institutionscykeldesign

För institutioner som har återkommande grupper eller kohorter kan man utveckla en återkommande institutionscykeldesign t ex på följande sätt:

G1: O1	X	O2	O3
-----			
G2:		O1	X
-----			
G3:		O1	X
			O2 osv

En närmare analys visar att G1 och G2 kan ses som en "prepost test med bara post-test" design. Där kan man direkt få en hel del svar genom att jämföra G1:O2 och G2:O1 med varandra. Om den skillnaden är lika stor som skillnaden mellan G1:O1 och G1:O2 talar det för att man har en behandlingseffekt och inte en selektionseffekt. Är skillnaden mellan G1:O2 och G2:O1 däremot inte är lika stor som skillnaden mellan G1:O1 och G1:O2 talar det för att en selektionseffekt kan finnas. På liknande sätt kan man slå ihop alla O1 för att få ett ingångsvärde och alla O2 för att få ett utgångsvärde och alla O3 för att få ett uppföljningsvärde. Man kan också jämföra dessa med O1, O2 och O3 vid en viss tidpunkt för att se om det ger samma uppskattning av Y. Detta är ett exempel på en kvasiexperimentell design som har möjlighet att kontrollera många hot mot den interna validiteten. Detta är en design som skulle kunna användas inom många organisationer med återkommande behandlingscykler.

Låt oss ta ett exempel på hur det skulle kunna se ut för ett utbildningsprogram som återkommer regelbundet. Den första gruppen (G1) mäts vid intagningen (G1:O1) och efter avslutad utbildning (G1:O2). Samtidigt tas nästa elevgrupp (G2) in och mäts vid intagningen (G2:O1). Redan här kan man göra två jämförelser. Skillnaden G1:O2-G1:O1 återspeglar den eventuella utbildningseffekten, men kan t ex också bero på att det råkar vara en speciellt begåvad elevgrupp (selektionsfaktorn). Detta kontrolleras genom att jämföra intagningsvärdena mellan de två grupperna (är G1:O1 lika med G2:O1?). I en institution med många återkommande grupper kan man effektivt eliminera selektionsfaktorn som hot mot kausaltolkningen att det finns en effekt av utbildningen.

### Tidsseriedesign

Här upprepar man mätningen av samma grupp vid flera tillfällen och ser hur förändringar i Y samvarierar med X.

O1	O2	O3	O4	O5	O6	O7
			X			

Om det visar sig att Y ökar vid O5 så har man visst stöd för att behandlingen haft effekt. Naturligtvis måste man analysera alternativa tolkningar (Z) som vanligt. Om man kan låta X återkomma finns ännu fler möjligheter att se om X har någon rimlig kausal förklaringskraft.

### Dataanalys

Att man kan tala om designer och metoder för dataanalys kanske de flesta inte har tänkt på. Vid planeringen av utvärderingsprojekt lägger man ofta ner den mesta energin på att välja designer och metoder för datainsamling medan design och metod för dataanalys inte alls får motsvarande uppmärksamhet. Designer och metoder för dataanalys är en fråga om hur man organiserar sina data efter de är insamlade för att få svar på olika frågeställningar. Med kunskaper om designer och metoder för dataanalys kan man ofta få ut betydligt mer av de data som samlats in än vad man tänkte på när man planerade sin datainsamling. Det sker en snabb utveckling inom detta område i takt med att små datorer blir allt mer kraftfulla och tillgängliga och det finns för närvarande ett stort antal olika metoder d v s dataprogram att välja mellan. Det är därför ett krävande och stort område att

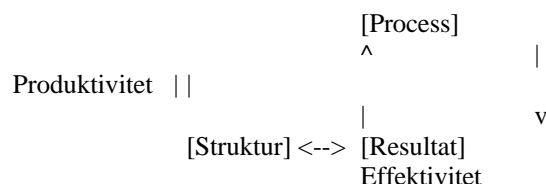
sätta sig in i för behandlingsforskare. Vi kommer inte att ge någon utförlig beskrivning av området utan endast peka på de möjligheter som finns.

Gemensamt för de olika designers och metoder som vi avser är att man vill förklara varians i Y som funktion av andra variabler. En grundläggande distinktion kan göras mellan designers där man arbetar enbart med de observerade variablerna och designers där man arbetar med tänkbara strukturer i latent variabler som förklaring till variation i de manifesta variablerna. I studier av kausala frågor där data samlas in med välkontrollerade designers brukar valet av analysmetod vara givet eftersom kausaliteten ingår i datainsamlingsdesignen. I regel analyseras medelvärdeskillnader med variansanalytiska modeller och metoder. I studier av effektfrågor blir möjligheterna att välja lämpliga analysdesigners och metoder större. Det gäller här att känna till vilka designers och metoder som är möjliga och hur man ska organisera sina data för att genomföra analyserna.

Till exempel kan man vilja förklara samband mellan en uppsättning variabler med hjälp av faktoranalys. Det kan också vara intressant att försöka förklara varians i en variabel utifrån flera andra variabler med hjälp av regressionsanalys eller variansanalys. Man kan vilja testa olika kausala modeller med mer sofistikerade multivariata metoder där man arbetar med latent variabler som antas orsaka varandra på ett systematiskt sätt. Dessa kan sedan knytas till manifesta variabler och på det sättet kan man pröva olika kausalmodeller mot befintliga data. Den mest kända metoden för kausalanalys baserad på samvariation mellan variabler är LISREL som bygger på strukturmodeller. Ett annat syfte med multivariata sambandsdesigners är att få en bild av hur strukturen i data ser ut med hjälp av modeller som bygger på principalkomponentanalyser. Med dessa modeller kan man gruppera både individer och variabler. I klinisk forskning är det ofta meningsfullt att identifiera individer med likartade profiler på diagnosdata eller demografiska data för att senare kunna kontrastera dem mot behandlingsutfall. Med longitudinella data kan tidsserieanalyser eller stiganalyser användas för att analysera strukturen mellan mätningar vid olika tidpunkter och få fram tänkbara kausalstrukturer i ett sådant material.

## Kvalitetssäkring av verksamheter

Kan man ha någon nytta av tankarna kring validitet i designers när det gäller SiS verksamhet? Det gäller ju inte en enskild variabel utan ofta en hel verksamhet vid ett eller flera behandlingshem med många X, Y och Z-variabler inblandade. I detta avsnitt vill vi vidga perspektivet något genom att anknyta till Donabedians (1966) klassiska modell för kvalitetssäkring som också presenteras i SPRI-rapport nr 230. Modellen är framtagen för att utvärdera kvalitet i sjukvården och återges i figur 1 nedan. Modellen skiljer mellan tre olika aspekter av kvalitet: struktur, process och resultat.



Figur 1. Donabedians modell för kvalitetssäkring

Struktur är själva organisationen och de resurser som finns i form av personal, kunskaper, utrustning, lokaler, material etc. Process är den verksamhet som utförs i organisationen såsom mediciner, samtal, terapi, vidareutbildning, planerings- och uppföljningsarbete etc. Resultat är den förändring i patienternas hälsa som kan tillskrivas strukturen och processen. En stor del av kvalitetssäkringsarbetet handlar om att upprätta standarder och kriterier för de olika boxarna. Man skall t ex ha en viss bemanning (ex läkare) med viss kompetens (ex kirurg) på en enhet för att klara av de uppgifter man ålagts (ex operationer) eller en viss typ av händelser (ex diagnostik) måste inträffa för att nödvändiga processer (ex matchning diagnos-behandling) skall kunna äga rum.

Av större intresse för oss är att man med denna modell kan definiera olika effektivitetsmått. Produktivitet hänför sig till relationen mellan struktur och process eller hur de förutsättningar man har utnyttjas i processen, t ex antal terapisaftal för psykoterapeuter av olika kompetens, behandlingstider per anställd, planeringsmöten, ronder etc per avdelning. Alla processer sätts i relation till olika aspekter av strukturen. Effektivitet hänför sig till relationen mellan struktur och resultat d v s vad man har åstadkommit med de resurser man har, t ex mängd symptom minskning, drogfrihet, sjukskrivning etc i relation till antal anställda eller någon annan strukturaspekt. Som

exempel på modellen kan vi tänka oss att en klient med missbruks-problem kommer till ett behandlingshem med dess struktur i form av lokaler, utrustning och personal. Därinne vidtar processen i form av utredningar och behandling. Huvudresultatet är förändring i patientens missbruksproblem. Produktivitet är frågan om hur många samtal, utredningsinsatser, samvarotimmar med personal, rapporter och journalanteckningar etc man genomfört vid behandlingshemmet per resursenhet, t ex krona. Effektivitet är hur mycket missbruksproblemen minskat per insatt resurs.

### **Kausalfrågor och effektfrågor**

Med denna modell som grund vill vi nu knyta an till kausal- och effektfrågan i utvärdering av behandling. I den klassiska utvärderingsmodellen ovan behandlas inte frågan om kausalitet - begreppet finns inte med i modellen. Vi vill därför utvidga modellen genom att föra in kausal-frågor och effektfrågor i kvalitetssäkring av verksamheter.

### **Kausalfrågor i verksamheter**

Låt oss börja med att diskutera möjligheterna att genomföra en design för att svara på en kausalfråga inom en verksamhet som SiS. Det första man slås av är att komplexiteten är mycket stor. I Struktur-boxen har vi många variabler att ta hänsyn till. Varje skillnad mellan olika behandlingshem eller enheter med avseende på bemanning, lokaler, ideologi, kunskaper m m ger en mängd andra förklaringsvariabler som potentiella hot i en kausalfråga. På samma sätt finns många möjliga skillnader i Process-boxen i form av behandlingsutbud, tider för behandling, miljöterapeutiska inslag, resor, samråd, personalutveckling, information, föräldrakontakter, myndighetskontakter m m. En kausalfråga kräver att en eller möjligen några få processvariabler avgränsas till X medan alla andra betraktas som andra förklaringsvariabler och kontrolleras på något sätt. Slutligen har vi en stor mängd tänkbara variabler i Resultat-boxen. Det rör sig i princip om variabler relaterade till klienternas/personalens/anhörigas /socialtjänstens beteende/upplevelser/känslor/förmågor inne på behandlingsenheten/vid utskrivning/vid uppföljning. Frågeställningen avgör vilka variabler som är av intresse och data kan samlas in med olika metoder som vi antytt inledningsvis.

För kausalfrågor gäller regeln att ju mindre variation vi har att ta hänsyn till, desto bättre är våra förutsättningar att besvara frågan på ett tillfredställande sätt. Grundprincipen är att kontrollera variation för att få likhet i alla avseenden utom X. Därför är kausalfrågor svåra att besvara i komplexa verksamheter där mycket varierar. De lämpar sig bäst för laboratoriemiljöer.

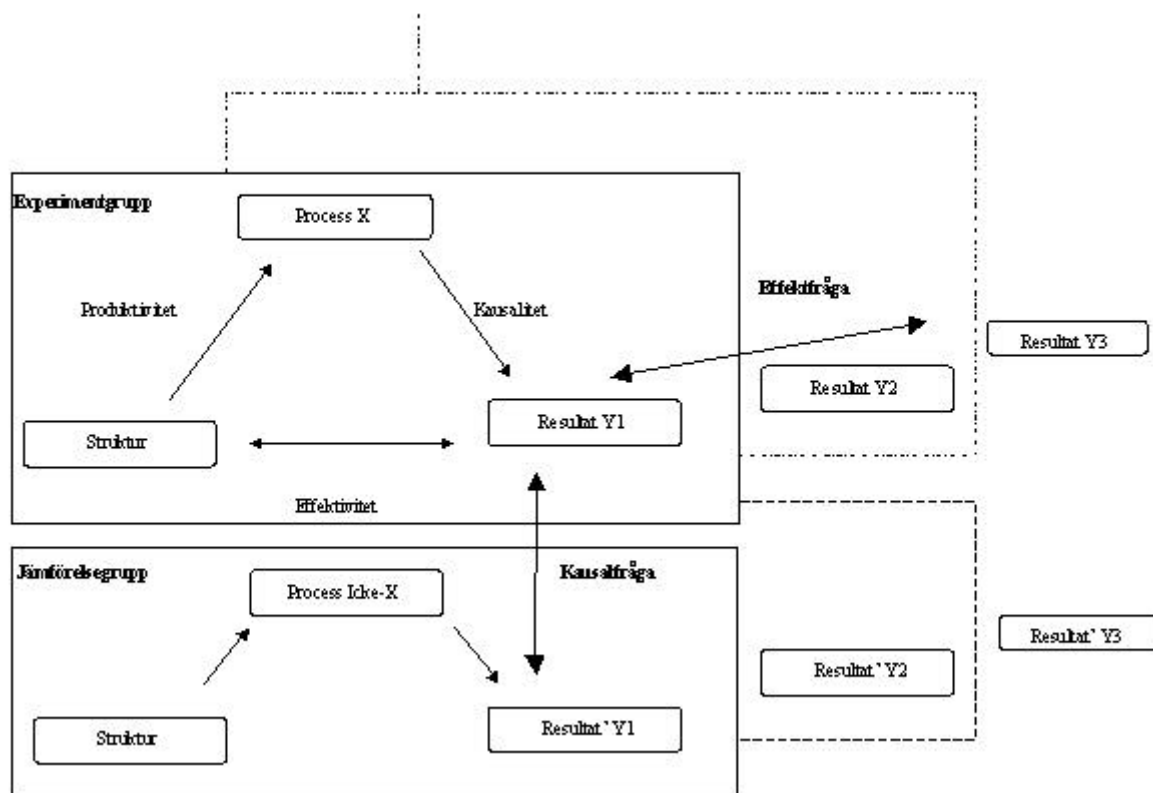
### **Effektfrågor i verksamheter**

I effektfrågor vill vi förklara varians i Y med hjälp av X. Det betyder att vi vill ha variation i såväl X som Y eftersom vi annars inte kan undersöka om de samvarierar. I en verksamhet kan många variabler fungera som X i den meningen att de samvarierar med Y. Det gäller alltså inte bara sådana processvariabler som mer direkt kan tänkas orsaka effekten i Y utan även sådana som normalt ingår i Struktur-boxen. Det kan vara av intresse att se om t ex personalens attityder i vårdideologiska frågor samvarierar med Y eller om lokalernas storlek har någon samvariation med Y. Med lämpliga statistiska metoder kan vi också kontrollera för variation i en eller flera andra förklaringsvariabler för att se om det finns indirekta länkar mellan Y och olika variabler. Med effektfrågor är det därför viktigt att så noggrant som möjligt registrera ett antal centrala förklaringsvariabler för att kunna relatera varians i Y till varians i dessa variabler.

Vi kan se att effektfrågor lämpar sig bra för komplexa verksamheter. Förutsättningen är naturligtvis att vi har tillgång till variabler som återspeglar verksamhetens olika delar så att vi kan vi bygga modeller där delarna sätts i relation till varandra. Idealet för en effektfråga är att ha tillgång till all varians i verksamheten. Det är också en fördel att ha med många behandlingshem eller varianter på utformning av verksamheten. Då kan man få variation i många avseenden och man kan undersöka om den är systematiskt relaterad till målsättningen med verksamheten, Y.

### **Utvärdering inom SiS: En modifierad kvalitetssäkringsmodell**

Låt oss avsluta diskussionen om designfrågor i verksamheter med hjälp av en lite mer utförlig figur över kvalitetssäkring baserad på Donabedians modell med tanke på SiS verksamhet



Figur 2. Modell över kausalfrågor och effektfrågor i utvärdering inom SiS.

I Figur 2 har vi återgett modellen inom ramen för en verksamhet, t ex ett behandlingshem inom SiS. Här har vi kallat relationen mellan Process och Resultat för Kausalitet eftersom det är något i processen som orsakar förändringen. Antag att man inom verksamheten vill veta om X orsakar effekten Y, d v s man ställer en kausalfråga. Man kan tänka sig att t ex undersöka betydelsen av att lägga till ett program för träning av ilskekontroll till ett redan befintligt behandlingsprogram inom en enhet. Ett sätt att få svar på den frågan är att skapa en experimentgrupp och en kontrollgrupp inom behandlingshemmet. Det enda som skiljer de två grupperna är X. Om det är möjligt kan vi randomisera klienterna till de två grupperna för att få kontroll över deras ingångsvärden. Vi får med denna uppläggning en god kontroll över många variabler, men vi måste fortfarande hålla uppsikt över vad som händer under tiden X pågår så att inga nya hot mot validiteten i designen uppstår.

Antag nu att vi vill veta om hela behandlingsinsatsen vid enheten har någon effekt på Y utan att vilja precisera exakt vad det är i X som orsakar Y. För att få svar på en sådan fråga bör man ha en longitudinell design där man studerar förändringar hos klienterna över tid. Det bör vara en skillnad mellan ingångsvärdena och utgångsvärdena och helst bör effekterna sitta kvar när klienterna återkommer till sin vardagliga situation för att vi skall känna oss säkra på att behandlingen är effektiv. Det är oftast nödvändigt att ha upprepade mätningar för att svara på effektfrågor om behandling varför vi har ritat flera bilder bakom varandra i figuren. Det finns alltid en Struktur-box där vi gör mätningarna i Resultat-boxen, men vid uppföljningar är det inte längre samma struktur som omger klienten. I vårt exempel frågar vi oss om förändringarna i Y står sig oavsett hur strukturen ser ut, d v s om de kan generaliseras till den struktur där klienten befinner sig just då.

Låt oss anta att Resultat Y1 är värdet på Y vid inskrivning, Resultat Y2 är värdet vid utskrivning från behandlingshemmet och Resultat Y3 är värdet vid en uppföljning ett år efter utskrivning. Y kan vara ett mått på missbruk, kriminalitet, social förmåga, självkänsla eller någon annan relevant beroende variabel. Ju mer vi vet om vad som inträffat mellan mätillfällena (Z), desto bättre förutsättningar har vi att undersöka sådana variablers eventuella värde som andra förklaringsvariabler. Om vi finner att det sker en förbättring från Y1 till Y2 och att den står sig till Y3 har vi goda skäl att tro att behandlingen har en positiv effekt som inte är strukturbunden, vilket är ett idealt utfall. Om resultatet inte står sig till Y3 vet vi att det rör sig om en begränsad effekt och att det sannolikt finns andra förklaringsvariabler utanför behandlingshemmet som vi inte har kunskap om. Om vi inte får en förbättring till Y2 behöver vi inte förklara något alls, även om uteblivna behandlingseffekter naturligtvis kan vara ett resultat av olika Z-variabler.

Om vi antar att vi har tillgång till en eller flera jämförelsegrupper ökar variationen i alla variabler. Det förbättrar möjligheterna att studera samband mellan Z, X och Y, speciellt som man då kan göra longitudinella jämförelser också av förloppen inom varje enhet. Det är alltså bra med flera jämförelsegrupper även för effektfrågor, men det är inte absolut nödvändigt under förutsättning att det finns tillräcklig variation inom den studerade gruppen. För att få kontroll över selektionsfaktorn och mognad är det idealiskt att randomisera klienterna till de olika enheterna. Vi får då en rimlig garanti för att selektionsfaktorn och mognadsfaktorn är lika mellan enheterna, vilket är en bra förutsättning för att knyta varians i Y till den globala X-variabeln. Man måste dock som vanligt registrera alla andra tänkbara förklaringsvariabler som kan inträffa parallellt med X och Y för att skapa förutsättningar för statistisk kontroll. Problemen med randomisering ligger inte i designens logik utan i de praktiska och etiska förutsättningarna för att genomföra en randomisering. Kan man motivera klienter och socialtjänstemän att gå med på förfarandet? Är det etiskt att inte erbjuda det man tror är den bästa behandlingen till vissa klienter? Kan man skicka klienter till enheter som ligger långt bort från hemorten?

### **Diskussion av utvärderingsfrågor inom SiS**

Avslutningsvis skall vi diskutera utvärderingsfrågor i anslutning till SiS verksamhet. Som framgått av resonemangen kring kvalitetssäkringsmodellen menar vi att det finns många möjligheter till såväl kausalfrågor som effektfrågor inom SiS egna institutioner. Den logiska ordningen bör vara att arbeta med effektfrågor i första hand och kausalfrågor i de fall man har konstaterat att en effekt finns eller med stor sannolikhet kommer att finnas. Man bör också ha bra begreppsvaliditet och extern validitet i X innan man genomför en studie avsedd att besvara en kausalfråga.

En design som förefaller värd att utveckla mer systematiskt för SiS är den som kallas för Institutionscykeldesign. Villkoren för en sådan design är gynnsamma genom att det är ett relativt likartat och väl avgränsat behandlingsprogram som upprepas över tid för olika personer. Vid SiS görs också systematiska intagnings- och utskrivningsintervjuer med DOK och ADAD som kan användas. Även om klienterna inte tas in gruppvis utan kontinuerligt kan man bilda grupper i efterhand som jämförs, t ex alla som klienter som tas in under första kvartalet blir grupp 1 och O1 och O2 definieras individuellt utifrån datum för intagning och utskrivning. Även om man med denna design kan få kontroll över selektionsfaktorn måste man försöka uppskatta de andra hoten mot validitet som finns när det rör sig om en longitudinell design, som t ex stora förändringar i behandlingsinnehåll eller stort eller selektivt bortfall.

För att strukturera diskussionen kommer vi att utgå från en modell som anger 3 olika nivåer (som exempel) av komplexitet i X som kan bli föremål för utvärdering:

Nivå 1 är hela SiS som behandlingsinstitution

Nivå 2 är en behandlingssideologi eller behandlingsteori som används inom SiS

Nivå 3 är en specifik behandlingsinsats vid ett behandlingshem

### **Nivå 1 - Har vården vid SiS någon effekt?**

Den viktigaste frågan att försöka besvara är om behandlingen vid SiS har någon effekt på klienternas framtida beteende. Det är i grunden en kausalfråga där X är behandling vid SiS respektive ej behandling vid SiS. Låt oss granska denna fråga ur de tre validitetsperspektiven. Den begreppsliga validiteten, som innebär att X definieras på ett tillfredställande sätt riskerar att bli mycket låg eftersom komplexiteten i behandlingen inom SiS är stor. För att åstadkomma en god intern validitet måste man ha kontroll över undersökningsbetingelserna i form av andra förklaringsvariabler, Z. Detta kan åstadkommas genom någon kvasiexperimentell design, t ex en återkommande institutionscykeldesign som inte kräver jämförelsebetingelser utanför SiS, eller genom en randomiserad design där flera socialbyråer placerar sina klienter antingen till SiS eller till ett annat behandlingsalternativ. Det kräver ett samarbete med institutioner utanför SiS som också skall ha kontroll över sådana förklaringsvariabler som ligger utanför SiS. Den externa validiteten kräver att det är ett representativt urval av klienter och SiS-enheter som studeras. Om man endast har valt klienter som placerats hos SiS kan man naturligtvis endast uttala sig om sådana som kommer till SiS och inte om socialtjänstens klienter i allmänhet. Risken för en systematisk selektion av svåra klienter till SiS gör designen som arbetar med interna jämförelser inom SiS olämpliga. Om man däremot kan samarbeta med socialtjänsten så att man inkluderar även klienter som inte hamnar hos SiS ökar möjligheterna att generalisera till olika klientgrupper liksom till jämförelse med andra behandlingsalternativ inom socialtjänsten. Som framgår av detta resonemang skulle ett samarbete med socialtjänsten som innebär att man får tillgång till en jämförelsegrupp utanför SiS öka såväl den interna som den externa validiteten, men inte påverka den begreppsliga validiteten.

Hur skulle det se ut om vi gjorde om frågan till en effektfråga? Liksom vid alla effektfrågor vill vi då maximera variationen för alla förklaringsvariabler och samtidigt notera vilket värde som gäller för varje betingelse. Om vi vill öka den begreppsliga validiteten skall vi alltså ha indikatorer på alla tänkbara aspekter av behandling (X) som finns vid SiS. Om vi har enheter eller personer utanför SiS med i designen måste relevanta data om behandlingsinnehållet samlas in även för dessa. Om man nöjer sig med en globalt definierad X för att få veta om det överhuvudtaget finns någon effekt att förklara ökar möjligheterna att genomföra studien. Då kan vi koncentrera oss på att samla information om tänkbara alternativa förklaringsvariabler och välja en design för dataanalys som tillåter analys av hur de samvarierar med X och Y, vilket är ett sätt att stärka den interna validiteten. För att tillgodose den externa validiteten måste jämförelser göras med klienter utanför SiS i enlighet med resonemanget ovan. En sådan studie skulle förmodligen kunna utnyttja registerdata av olika slag och såväl X-, Y- som Z-variabler.

## **Nivå 2 - en behandlingsideologi/-teori inom SiS**

Låt oss ta exemplet med behandlingshemmet S där man vill utvärdera effekterna av kognitiv beteendeterapi (KBT). Eftersom tanken är att låta hela verksamheten vid S genomsyras av KBT kan man säga att det handlar om en effekt- eller kausalfråga avseende en hel behandlingsteori med tillhörande tekniker. Det är då viktigt att inte sammanblanda denna fråga med utvärderingen av S, som är ett enskilt behandlingshem. Det kommer att vara en oupplöslig sammanblandning av både struktur och process vid S och KBT som behandlingsteori om det endast är S som använder sig av KBT. Samma sak gäller för Icke-X. För att tillgodose kraven på intern validitet måste vi således se till att det blir en variation i strukturvariabler både inom X och Icke-X för att kunna utvärdera hela behandlingsteorin. I processvariablerna ingår många icke-specifika effekter som har att göra med personalens engagemang och tilltro till sin behandlingsteori och många generella effekter som finns även vid behandlingshem som inte arbetar enligt KBT. Det är därför nödvändigt att ha med minst två behandlingshem med var och en av X-betingelserna.

Kan detta utformas som en kausalfråga? Problemen är mycket likartade dem som gällde ovan för nivå 1 och komplexiteten i X måste betraktas som alltför stor för att man skall kunna definiera X på ett meningsfullt sätt. Det rimliga är att betrakta en utvärdering av en behandlingsideologi som en effektfråga och försöka förklara varians i Y med hjälp av goda beskrivningar av såväl X som Z. Det är emellertid inte alls nödvändigt att involvera enheter utanför SiS för att tillgodose krav på vare sig intern validitet eller extern validitet. Om det går att randomisera klienter inom SiS till enheter som har KBT respektive inte har KBT är det en effektiv metod att få kontroll över selektions- och mognadsfaktorn när man skall förklara effekter i Y. Kanske kan man randomisera efter det att klienterna kommit till SiS och ännu inte placerats vid någon enhet. Om det inte går att randomisera kan man ändå använda en kvasiexperimentell design för att få kontroll över olika hot mot den interna validiteten. Ju fler enheter och individer som ingår i studien, desto större är variationen och möjligheterna att också få en statistisk kontroll av variationen även om kontrollen vid själva datainsamlingen är begränsad. Med de interna dokumentationssystemen ADAD och DOK finns goda möjligheter att göra intressanta designers för dataanalys som kan hjälpa till att få en överblick över såväl struktur- som processvariabler som har systematiska samband med Y.

## **Nivå 3 - en specifik behandlingsinsats**

Vi har redan använt införandet av ilskekontroll som ett tillägg till annan behandling vid S som ett exempel på hur en kausalfråga skulle kunna läggas upp. Här finns de flesta villkoren uppfyllda för att försöka testa kausaleffekten av ilskekontroll i behandlingsprogrammet. Det finns en dokumenterad effekt av denna specifika insats i forskningslitteraturen och man kan säga att begreppet är väldefinierat såväl teoretiskt som operationellt. Den externa validiteten i förhållande till klienter inom SiS kan också sägas vara god i den meningen att många klienter har problem med att kontrollera sin aggressivitet. Om man vill öka den externa validiteten med avseende på behandlingsinnehållets effektivitet kan man pröva att lägga till ilskekontroll vid någon behandlingsenhet som inte arbetar med KBT.

## **Sammanfattning och slutsatser**

Vi har i denna skrift försökt gå igenom villkoren för att genomföra utvärderingar av behandlingsverksamheten inom SiS utifrån principer för experimentell design och kvalitetssäkring. Genomgången visar att det är svårt att få den grad av kontroll som behövs över alla tänkbara konkurrerande förklaringsvariabler för att dra slutsatsen att det är behandlingen vid SiS som har orsakat de observerade effekterna. Därför kan man inte tala om behandlingsinsatserna som orsaken till effekterna i strikt logisk mening. Det är emellertid mycket viktigt att man undersöker effekterna av SiS behandlingsverksamhet. Detta är en lättare uppgift eftersom vi då inte behöver kontrollera att alla tänkbara alternativa förklaringsvariabler är utjämnade eller likställda mellan

undersökningsgrupperna i samband med datainsamlingen. En effektfråga kräver endast att vi kan genomföra en statistisk kontroll av variabelernas betydelse som förklaringsvariabler. Det betyder att vi måste registrera värdet på varje variabel för varje klient för att kunna beräkna hur den samvarierar med utfallet av behandlingen. Denna typ av analys kan också inkludera kontroll av variation i andra variabler i syfte att renodla effekterna. Syftena med kausalfrågor och effektfrågor är således mycket likartade, men genomförandet skiljer sig ibland avsevärt.

När den verksamhet som skall studeras är mycket komplex eller sammansatt av många komponenter som samverkar för att åstadkomma en effekt är det inte möjligt att få svar på kausalfrågor. Däremot är det möjligt att utforma effektfrågor som idealiskt inkluderar mycket information om behandlingsinsatserna, men det är också möjligt att utelämna hela frågan om vad som ingår i behandlingen och endast fokusera på att beräkna storleken på behandlingseffekten, även om det inte är att rekommendera eftersom man får veta så lite i förhållande till arbetsinsatsen.

Begreppet kontroll spelar en central roll i både kausalfrågor och effektfrågor. Vid kausalfrågor gäller det att ha kontroll över villkoren för datainsamling. Ibland kan man åstadkomma detta genom att ta bort eller begränsa variation i variabelerna precis som i ett laboratorium. Detta är i regel en dålig strategi vid utvärderingar därför att det också förstör möjligheten att statistiskt kontrollera betydelsen av den variabel man hindrat att variera. Andra kontrollmetoder påverkar inte variationen och kan därför vara lämpliga både för kausalfrågor och effektfrågor. Hit hör metoder som randomisering, matchning och balansering. Vid effektfrågor vill vi kontrollera samvariation mellan variabler på olika sätt. Detta kan göras med datorprogram som tillåter analys av både manifesta och latent strukturer i data. Kunskaper inom detta fält är centrala för att utforma bra effektdesigner.

Vår genomgång visar att det är viktigt att ha ett gemensamt system för att samla in data från SiS olika enheter. Härvidlag utgör de gemensamma dokumentationssystemen en mycket bra grund att stå på. Huruvida de bör kompletteras eller modifieras kommer att visa sig när den insamlade informationen börjar analyseras i syfte att besvara relevanta frågor.

### *Referenser*

Cook, T. D. & Campbell, D. T.

Quasi-experimentation. Design and analysis issues for field settings. Chicago: Rand McNally College Publishing Company, 1979.

Se även föregångaren:

Campbell, D. T. & Stanley, J. C. Experimental and quasi-experimental designs for research. Chicago,: Rand McNally, 1966.

Donabedian, A. Evaluating the quality of medical care. Milbank Memorial Fund Quarterly 1966: 44 s 166-206.

Kvalitetssäkring. Att mäta, värdera och utveckla sjukvårdens kvalitet. Stockholm: SPRI-rapport 230. 1987.