

Lars R. Bergman

1995 -05- 0 2

ANPASSNING, BETEENDE OCH PRESTATION

Örebroprojektet

III

Metoder och modeller

David Magnusson - Anders Dunér

Avdelningen för tillämpad psykologi

Psykologiska institutionen

Stockholms universitet

Juni 1967

Avdelningen för tillämpad psykologi
Rådmansgatan 70
Stockholm VA
Tel.: 34 08 60 / 427
Föreståndare: Laborator David Magnusson

Förord

I rapport nr I i den serie där denna rapport ingår har den allmänna målsättningen för ett forskningsprojekt rörande skolelevers anpassning, beteende och prestation redovisats. I föreliggande rapport diskuteras och belyses vissa av de metodologiska problem som hör samman med forskning kring sådana frågeställningar. I rapporten presenteras också de statistiska metoder och modeller med vilka problemen kan angripas.

I diskussionerna av de problem som berörs i rapporten har universitetslektorn i statistik, fil.lic. Erik Leander deltagit.

David Magnusson

INNEHÅLLSFÖRTECKNING

Förord	Sid.
KAP. I. MÅLSÄTTNING	1
KAP. II. GRUNDLÄGGANDE ANGREPPSSÄTT	2
A. Experiment	2
B. Ex post factoundersökningar	3
C. Studier av relationer mellan variabler	3
KAP. III. FORSKNING I ÖPPNA OCH SLUTNA SYSTEM	5
KAP. IV. KARTLÄGGNING AV ANPASSNINGSSYMTOMENS KARAKTÄR	7
A. Stabilitet hos symtöm över tid	7
B. Symtoms prognostiska värde	8
C. Kommentar	8
KAP. V. KORRELATION OCH REGRESSION	10
A. Enkel regressionsanalys	10
B. Korrelations- och regressionskoefficienter	10
C. Extremgruppsmetoden	11
D. Multipel regression och prediktion	12
KAP. VI. RESIDUALER SOM MÅTT PÅ HÄRLEDDA VARIABLER	13
KAP. VII. FÖRSÖK TILL FÖRKLARING AV SAMBAND	20
A. Motivering för den mångsidiga uppläggnen	20
B. Kausalanalys	21
C. Kausalmodeller	22
D. Kontroll av omätbara variabler	25
E. Exempel på kausalmodeller i Örebroprojektet	27
F. Dimensionsanalys	32
KAP. VIII. KLASSIFIKATION	33
A. Motivering för klassifikation	33
B. Latent profilanalys	34
C. Moderatoranalys	34

KAP. I. MÅLSÄTTNING

I en tidigare rapport har presenterats en långtidsplanerad undersökning, vars första syfte är att analysera vilka faktorer, särskilt i skolsituationen, som kan vara av betydelse för elevernas anpassning, beteende och prestation, dels aktuellt i skolsituationen, dels på längre sikt i vidare utbildning och yrkesverksamhet (Magnusson, Dunér och Beckne, 1965). Tyngdpunkten ligger därvid på en analys av anpassningsproblematiken, dels i vad avser bakgrunden till en elevs aktuella anpassningsstatus, dels i vad avser värdet av mått på den aktuella anpassningen som underlag för prognos om anpassning i senare utbildning och yrkesverksamhet.

I den tidigare rapporten definierades skolanpassning (yrkesanpassning) på följande sätt:

"Individens förmåga att finna sig till rätta med de krav, som skolan (yrkeslivet) ställer och den miljö, som skolan (arbetsplatsen) erbjuder." Den anpassning som avses därmed kan komma till uttryck och mätas i tre olika avseenden:

1. Att individen upplever subjektiv tillfredsställelse.
2. Att omgivningen upplever individens beteende som tillfredsställande.
3. Att individen på ett adekvat och tillfredsställande sätt utnyttjar sina begåvningsmässiga förutsättningar.

KAP. II. GRUNDLÄGGANDE ANGREPPSSÄTT

Anpassningsproblemen i skolan kan studeras och har också studerats på en mängd olika sätt. De typer av designs, som närmast kan ifrågakomma, är följande:

- A. Experiment. Det väsentliga i experimentbegreppet är att en förändring i något avseende hos ett stickprov åstadkommes genom aktiv manipulation från experimentledarens sida. I regel genomföres i den utsträckning så är möjligt kontroller av andra variabler som kan influera de variabler som studeras. Ett laboratorieexperiment ger oftast goda möjligheter här till, det försiggår i ett slutet system, d. v. s. det är möjligt att överblicka och kontrollera alla variabler som kan vara relevanta. Ett fältexperiment försiggår däremot i ett öppet system; antalet möjligen relevanta variabler är oöverskådligt och vilka som kontrolleras blir beroende av experimentledarens subjektiva uppfattning om vad som är viktigt. I många fall möter också den svårigheten att kända och väsentliga variabler ej låter sig mätas. I sådana fall kan man bli hänvisad till att i stället mäta andra variabler som i och för sig är ointressanta, men som har ett sådant samband med en o-mätbar variabel att den kan tjänstgöra som indikator på denna. Ytterligare diskussion av denna problematik föres i avsnitt III.

Experimentsituationens grad av slutenhet och därmed följande möjligheter att kontrollera irrelevanta variabler avgör resultatens tolkbarhet. I föreliggande undersökning med ett öppet system är tolkbarheten av experimentresultat tämligen måttlig.

Experiment kring anpassningsproblematiken i öppna system är dessutom förenade med stora praktiska svårigheter och kostnader. Man måste också konstatera, att av två skäl för närvarande i väsentlig utsträckning möjligheter saknas att värdera resultaten av sådana undersökningar, hur väl planerade och genomförda de än är: 1) Vi saknar tillräcklig kännedom om karaktären av de symtom som vanligen brukar betecknas som anpassningssymtom, 2) Vi känner för dåligt vad som ligger bakom

och orsakar dessa symtom. Det måste därför vara en central och grundläggande uppgift att närmare studera just dessa frågeställningar, så att bättre förutsättningar skapas för experimentella studier av anpassningsproblematiken.

Av dessa anledningar kommer rena experiment ej att bli aktuella i de första faserna av projektets arbete. Däremot synes det angeläget att senare med experimentell metodik pröva effekten på anpassning av åtgärder av olika slag, t. ex. pedagogiska, organisatoriska, allmänt miljömässiga.

B. Ex post factoundersökningar. I ett experiment manipulerar experimentledaren vissa förhållanden genom systematisk variation av den oberoende variabeln. I ex post factoundersökningar tar man fasta på utveckling eller förändring, som skett på naturlig väg och försöker att dra slutsatser om bakomliggande faktorer. Vi kan här särskilja två metodiska grepp:

B:1. Studium av individer med specifika karakteristika, t. ex. låg prestationsförmåga eller dålig anpassning i skolan. Data behöver endast föreligga från ett undersökningstillfälle. I analysen försöker man klarlägga orsakssammanhangen t. ex. genom att jämföra en extremgrupp med en kontrollgrupp, och genom att samla in uppgifter om händelser i individens liv som kan antas orsaka extremgruppstillhörighet.

B:2. Undersökning av individer, vilkas status i ett relevant avseende har förändrats mellan två mättillfällen. Händelser under den mellanliggande tiden kartläggs jämte individernas egenskaper och förhållanden i övrigt som underlag för prövning av hypoteser om orsaker till förändringarna.

C. Studier av relationer mellan variabler. Sambanden mellan variabler studeras vanligen genom beräkningar av korrelations- och regressionskoefficienter. Samma kunskap kan också erhållas genom studier av extremgrupper. Vi kan därför särskilja tre situationer.

C:1. Beräkning av korrelationer mellan variabler när mätningarna har utförts vid ett tillfälle. I detta fall är det möjligt att genom regressionsanalys använda vilken eller vilka variabler som helst för att "predicera" det simultant erhållna värdet på

estimation

vilken som helst av de övriga variablerna. I detta fall är det lämpligt att i stället för prediktion tala om estimation. Sambanden mellan två variabler betraktas som symmetriska, d. v. s. man implicerar ej något orsaksförhållande med bestämd riktning. Ur de korrelationsmatriser, som ifrågavarande analyser vanligen utmynnar i, kan man emellertid erhålla värden för prövning av kausalmodeller. En modell för kausalanalys redovisas senare.

Korrelationsmatriserna utgör också grunden för de dimensionsanalytiska metoder man använder för att kartlägga totalstrukturen eller för att gruppera variabler i syfte att få så effektiva dimensioner som möjligt för prediktion eller kausalanalys.

prediktion

C:2. Beräkning av korrelationer mellan en grupp variabler, ofta benämnda prediktorer, och en annan grupp, vanligen benämnda kriterier, som mätts vid olika tidpunkter. Mer allmänt kan här användas beteckningarna oberoende respektive beroende variabler för den variabelgrupp, som mätts vid första respektive andra mättillfället. Detta fall bör betecknas prediktion.

Då prediktionerna används för att studera de individer, som mest avviker från predicerad utveckling, ger metoden utgångspunkten för en ex post factoundersökning av typ B:2.

Då prediktionerna används för att pröva åtgärder för att förhindra en predicerad ogynnsam utveckling, för en viss grupp individer, ger metoden utgångspunkten för ett experiment i ett öppet system.

C:3. Extremgruppsundersökningar. Metoden att utvälja grupper, som i avseende på en variabel är extrema, jämte kontrollgrupper, ger vid linearitet i princip samma upplysningar som korrelationsanalyser. Extremgruppsmetoden är motiverad huvudsakligen då data endast kan insamlas för ett mindre antal individer. Dessutom kan en översiktlig bild av en i visst avseende homogen grupp ge utgångspunkter för hypotesformuleringar. Vid studiet av så komplexa variabler som anpassningssymtom torde det också vara nödvändigt med analyser av individer för att kartlägga den sannolikt stora mängden av alternativa orsakskombinationer.

KAP. III. FORSKNING I ÖPPNA OCH SLUTNA SYSTEM

En aspekt av avgörande betydelse för bedömning av ett psykologiskt forskningsprojekt, vilket grundläggande angreppssätt som än använts, är om observationerna gjorts i ett öppet eller slutet system. Vissa problem fordrar en miljö, andra problem kan endast lösas genom observationer i en annan. Den problematik som hör samman med val av system skall diskuteras något.

Vi antar att vi önskar fastställa om variabel X orsakar variation i Y , alltså om $X \rightarrow Y$. I ett experiment går vi till väga så, att andra variabler, t. ex. Z på ett eller annat sätt kontrolleras i den mån de anses kunna påverka Y , d. v. s. som man misstänker $Z \rightarrow Y$. Alla variabler kan emellertid ej kontrolleras, någonstans måste en gräns dras, där man gör ett antagande att övriga variabler ej är relevanta.

I ett slutet system antas alla variabler vara kända och mätbara och alltså möjliga att kontrollera. Där är också möjligheterna stora att dra säkra slutsatser om $X \rightarrow Y$. I ett öppet system fungerar många variabler, som kan tänkas inverka på Y . Så är fallet i en undersökning, som går in i människors vanliga livsbetingelser. Många av de variabler man antar kan inverka är mätbara, andra är ej mätbara. I det senare fallet kan man mäta någon annan variabel, som antas vara kausalt förknippad med den relevanta variabeln utan att själv vara relevant. Vanligen betecknas då den mätbara variabeln indikator.

Det slutna systemet approximeras närmast i vissa laboratorieexperiment. Örebroprojektet försiggår i ett öppet system.

Det förhållandet att en undersökning sker i ett öppet system innebär stora metodologiska svårigheter. Emellertid kan man ej bortse från att det ur en synpunkt innebär betydande fördelar jämfört med det slutna systemet. Om man i ett laboratorieexperiment manipulerar X och finner att $X \rightarrow Y$ och därvid konstaterar ett orsakssamband, kan man därvid ej anses ha tillfört vetandet något väsentligt om t. ex. X aldrig varierar i tillämpningssituationer, där Y är relevant. Om problemställningarna i det öppna systemet är mer komplicerade och svåranalyserade, är de också mer meningsfulla.

Beteendet, som det t. ex. manifesterar sig i skolelevers prestationsförmåga eller anpassning, har en komplex orsaksbakgrund. För varje handling vid varje enskilt tillfälle finns flera determinerande faktorer i individens unika miljö. Han har genom sin erfarenhet lärt

sig att uppskatta sannolikheterna för olika utfall av vissa handlingar vid vissa stimuluskonstellationer. Att han ej kan lära sig det säkra utfallet beror på att orsakssammanhangen varierar från tillfälle till tillfälle. Det finns därför heller inte något subjektivt säkert utfall i en bestämd situation, där man blott känner någon eller ett fåtal av orsaksfaktorerna.

Om man isolerar en orsaksfaktor och kontrollerar övriga faktorer med experimentell metodik i ett slutet system kan man studera individens beteende. Detta kan även där komma att karakteriseras av de sannolikheter han lärt in i sin naturliga miljö, men den konstgjorda laboratoriesituationen kan också ge oförutsedda förändringar.

I ett öppet system fungerar individens vanliga sannolikhetsuppfattningar. Vi kan genom ett mångdimensionellt studium lära känna dessa och indirekt även hans effektiva miljö. Vi studerar både individen, hans miljö och interaktionen mellan dem.

Vi kan aldrig vänta oss en fullständig säkerhet i prediktionerna av beteendet, därför att en sådan inte är teoretiskt möjlig. Vi kan däremot lära känna sannolikheterna för en individs utveckling i en bestämd riktning eller hans handling i en bestämd situation.

Vi ansluter oss i dessa synpunkter till Brunswik (1943) och hans efterföljare och till den metodiksyn som utvecklats ur den probabilistiska funktionalismen. Se även Gillis & Schneider (1966). Det måste emellertid betonas att psykologisk forskning är lika nödvändig på alla nivåer av öppen-slutenhet i det studerade systemet. Det måste vara enbart problemet som avgör vilken metodik som skall användas och om en öppen eller sluten miljö skall väljas för observationer.

KAP. IV. KARTLÄGGNING AV ANPASSNINGSSYMTOMENS KARAKTÄR

A. Stabilitet hos symtom över tid

Det är möjligt att vid en viss tidpunkt erhålla en mängd olika mått på anpassningssymtom, där vart och ett av måtten har en hög reliabilitet i traditionell mening. Vanligen nöjer man sig också med sådana mått på reliabilitet som underlag för en bedömning av data som värdefulla mått på anpassning. Det är emellertid nödvändigt för en bedömning av måttens användbarhet för olika syften att också undersöka om symtomen kvarstår under längre tid. Undersökningar av stabilitet bör göras för att dels rangordna symtom efter stabilitet, dels rangordna individer efter stabilitet i störningar.

Stabiliteten i symtom bör provas på olika sätt.

1. a) Korrelation mellan tillfällen för mått på enskilda symtom.
b) Vid ett jämförande studium av korrelationsmatriser från vart och ett av flera tillfällen är det möjligt att konstatera vilka konstellationer av mått som uppvisar en sådan stabilitet i samvariationen att det kan vara meningsfullt att beteckna dem som syndrom och att använda dem för prognos.

2. Genom studiet av individer eller grupper av individer. Den grupp av individer som vid en viss tidpunkt får starka indikationer på missanpassning i symtom eller syndrom, som visat stabilitet över tid, kommer sannolikt att bestå av individer med olika karaktär av missanpassning: Å ena sidan elever med en stadigvarande, på vissa egenskaper hos individen eller konstanta miljöförhållanden beroende missanpassning, å andra sidan elever med akuta störningar, akuta missanpassningssymtom, som beror på att eleven hamnat i situationer som tillfälligt ställer hans förmåga på för svåra prov, men för vilka symtomen försvinner, när den utlösande miljöfaktorn förändras. Denna senare typ av bristande anpassning kan vara vanlig i skolsituationer, där lärare, kamrater, skolarbete, skolväg, bespisning, etc. kan ge anledning till akuta svårigheter. I den mån elever av den senare typen dominerar vid mätningen av ett anpassningssymtom vid ett tillfälle, kommer självfallet korrelationen med mått på samma anpassningssymtom vid ett senare tillfälle att bli låg.

För båda angreppssätten gäller att en kombination av variabler, syndrom, ger större avkastning av bearbetningarna än bearbetning

av enskilda variabler, symtom.

Instabila symtom kan tänkas förekomma även vid stabil missanpassning. Vid det senare av de nämnda angreppssätten är det möjligt att konstatera om enskilda individer har konstant symtombelastning men varierande specifika symtom över tillfällena. Om detta visar sig i mer än enstaka fall är det angeläget att finna eventuella systematiska förskjutningar i symtombilden, sekvenser. Därvid skulle prognosmöjligheterna kunna vara goda även vid instabila symtom.

B. Symtoms prognostiska värde med avseende på senare anpassning i utbildning och arbete

Hela denna frågeställning ligger inbyggd i vad som sagts under A. Här tillkommer emellertid en annan ytterst viktig aspekt, nämligen frågan om vilka beteenden vid en viss tidpunkt, t. ex. i skolan, som faktiskt är indikatorer på en sådan djupgående, stadigvarande störning att den har prognostiskt värde på längre sikt med avseende på anpassning. Av de beteenden, som vanligen betraktas som uttryck för och indikatorer på en bristande anpassning, finns sannolikt vissa som, enskilda eller i kombination med andra, endast utgör en normal individreaktion på en olämplig miljö. Å andra sidan finns det sannolikt beteenden, som betraktas som indikatorer på en god anpassning, men som i kombination med vissa andra beteenden kan utgöra indikatorer på en grundläggande bristande anpassning och störning. Av den anledningen är det även för undersökning av prognosvärdet i symtom och syndrommått nödvändigt att studera dessa som effekter av förhållanden i individens miljö, kausalitetssamband.

C. Kommentar

För ett fullföljande av den målsättning som diskuterats under A och B krävs bl. a. följande:

1. Sampling av symtom. Ju effektivare samplingen av symtom täcker den totala anpassningsbilden, desto effektivare blir resultatet vid kombination av olika variabler.
2. Uppföljning av ett stort antal elever under längre tid.
3. Kartläggning av dessa elevers effektiva miljö, både hem och skola.

Det måste vara angeläget att belysa de frågeställningar som här har berörts. En mängd aktuella skolpolitiska uttalanden antyder före-

ställningar om att lärarens utformning av skolmiljön (skolsituationen), d.v.s. hans pedagogiska och disciplinära åtgärder, är avgörande för barnens anpassning, i varje fall så som den kommer till uttryck i elevernas beteende i skolsituationen. Det är av vikt att belägga i vilken utsträckning sådana föreställningar bygger på riktiga förutsättningar. Först om man vet något om i vilken utsträckning elevers störande beteende är uttryck för grundläggande, stadigvarande störningar hos dem eller för effekter av aktuella faktorer i miljön av övergående karaktär, kan man med större meningsfullhet vidta åtgärder i miljön eller hos individen för att komma till rätta med dessa störningar.

KAP. V. KORRELATION OCH REGRESSION

Korrelations- och regressionsanalys är differentialpsykologins vanligaste metoder. De kommer att spela en central roll vid bearbetningarna av data från screeningundersökningen. Vi skall kortfattat diskutera några hithörande metoder med avseende på speciellt viktiga förhållanden.

A. Enkel regressionsanalys

För analyser av samband är det av central betydelse att känna till ej enbart sambandets styrka utan även dess form. Linearitet är en förutsättning för vissa metoder liksom för tolkning av vanliga korrelationskoefficienter. Sambandens form studeras genom korstabelleringar. Avvikelser från linearitet prövas genom signifikanstestning av skillnader mellan eta- och korrelationskoefficienter. Så har t. ex. variabeln spänning visat sig vara icke-lineärt relaterad till övriga lärarskattningsvariabler. Detta har varit av betydelse för ställningstagandet till fortsatta bearbetningar av dessa data. Sådana resultat kan också vara av psykologiskt intresse.

B. Korrelations- och regressionskoefficienter

Vid linearitet erhålls upplysning om sambandets styrka genom beräkning av korrelationskoefficienter. I kvadrerad form (r^2 , determinationskoefficienten) anger de den andel av varians som är gemensam för två variabler.

Regressionskoefficienter anger å andra sidan storleken av den effekt en förändring av en enhet i en oberoende variabel har på en beroende variabel. Dessa två koefficienter svarar alltså på två olika frågor, som båda är väsentliga.

Förändring och substitution. I experimentella undersökningar åstadkommes förändring i en oberoende variabel (i fortsättningen betecknad X) genom manipulationer. I differentialpsykologiska sammanhang kan det gälla t. ex. behandlingar av individer. Man studerar de effekter dessa manipulationer har på en beroende variabel (Y) under det att övriga variabler kontrolleras.

I icke-experimentella undersökningar kan en annan typ av "förändring" studeras. Den innebär att individer med visst X-värde

ersätts med andra individer med annat X-värde. Vi skall här använda termen substitution för denna typ av "förändring". Man kan notera en väsentlig skillnad mellan egentlig förändring och substitution. Antag att man studerar sambandet $X \rightarrow Y$. Vid förändring av X-värdet hos samma individer kommer förändringar av Y-värden att kunna tolkas i termer av X under förutsättning av situationskontroll. Om vid förändring av Y däremot X ej förändras, blir en tolkning av orsaksriktning möjlig. Vid substitution däremot kommer utbyte av individer med visst Y-värde mot annan grupp med annat Y-värde att också medföra en förändring av X även om det gällande sambandet är $X \rightarrow Y$. Tolkning av orsaksriktning är då omöjlig.

C. Extremgruppsmetoden

Samband mellan variabler kan studeras genom att jämföra kontrasterande grupper. I de fall möjligheter finns att regressioner är icke-lineära, tas även en mittgrupp eller en slumpmässigt utvald genomsnittsgrupp med. Denna metod kommer till användning för analys av sådana variabler som det ej varit möjligt eller lämpligt att mäta på totalgruppen utan där individuella metoder som t. ex. intervjuer är nödvändiga. Av kostnadsskäl måste då stickprovets storlek starkt begränsas. I det läget är det effektivare att undersöka extremgrupper än ett litet slumpmässigt utvalt stickprov. Som exempel kan nämnas en intensivundersökning, där för små stickprov screeningundersökningens data kompletterats med ett antal uppgifter om hemförhållanden, barnets inställning i vissa avseenden och lärares uppfattning om eleverna, vilka erhållits genom intervjuer med föräldrar, elever och lärare.

I regel är vinsterna av att ur ett större material välja ut extremgrupper för analys av samband mellan urvalsvariabel och andra variabler ur statistisk synvinkel begränsade. Emellertid kan ett studium av grupper av individer, som utmärks av extrem utpräglingsgrad i ett avseende, ge en lättillgänglig bild av denna grupp och dess egenskaper, som kan vara av värde för vidare teoretiserande och för hypoteskonstruktion. Vi har med utgångspunkt från screeningundersökningens data studerat ett antal sådana grupper, t. ex. en grupp som ej trivs i skolan, en grupp med psykosomatiska symtom, en med uttalad ångslan för skolsituationen och en med föräldrar med speciella normsystem.

D. Multipel regression och prediktion

Korrelations- och regressionskoefficienter kan användas vid estimation av individuella värden i en variabel från kännedom om en eller flera simultant mätta variabler. De kan också användas så att från kännedom om en eller flera variabelvärden en prediktion kan göras om resultatet av en senare mätning av en annan variabel. Detta sker genom att en multipel regressionsekvation bestämmes, där för varje variabel en vikt har bestämts så, att den totala prediktionen blir maximalt god. Inom projektet kommer intresset att knyta sig till variablerna prestation, anpassning i skolan, anpassning i samhället m. fl. som predicerade eller estimerade variabler. Vi kan härvid först inom screeningundersökningen med estimation studera hur simultana mätningar struktureras.

Då vi ej känner något om orsakssamband måste vi t. v. betrakta alla samband som symmetriska. Vi kan alltså från andra variabler estimerat t. ex. vilka elever som har goda studieresultat eller god anpassning men också vilka som har hög intelligens.

Genom prediktion får vi svar på en annan typ av frågor, nämligen sådana som avser vem som kommer att prestera högt i en viss situation, vem som kommer att bli missanpassad etc. Även om det tekniskt vore möjligt att även vid prediktion använda vilka variabler som helst som prediktorer respektive kriterium, d. v. s. oberoende respektive beroende variabler, bör intresset här knytas till relevanta frågeställningar som förutsätter antaganden om orsak och verkan, d. v. s. riktade kausalsamband, t. ex. $X \rightarrow Y$ (variabel X orsakar Y men ej tvärtom).

Man måste då konstatera, att de i multipla regressionsekvationer ingående, matematiskt beräknade vikterna ej säger något om orsaker till variationer i den predicerade variabeln Y. Det är t. o. m. tänkbart att ingen direkt orsakande faktor ingår i ekvationen. Det som kunnat mätas och som kommer med i ekvationen kan vara enbart effekter av gemensamma bakomliggande variabler som antingen undgått vår uppmärksamhet eller som t. v. är omätbara. I detta senare fall kan vi i bästa fall mer eller mindre medvetet ha valt indikatorer på dessa orsakande variabler och gjort mätningar på dessa.

Den vikt med vilken en viss variabel ingår i ekvationen säger heller inget om dess betydelse som orsaksfaktor. Vikternas storlek bestäms av många praktiskt svåranalyserade faktorer, de ger en maximal prediktion till Y, men de säger inget om orsak och verkan.

KAP. VI. RESIDUALER SOM MÅTT PÅ HÄRLEDDA VARIABLER

Med två givna variabler som utgångspunkt är det möjligt att härleda en tredje, psykologiskt intressant variabel genom att individuellt beräkna avvikelser från det värde i variabeln Y som prediceras från variabeln X. Residualen, i fortsättningen betecknad E, i regressionen Y på X kan därefter studeras med avseende på samband med andra variabler i systemet.

Detta tillvägagångssätt har hittills kommit till användning vid mätning av över- underprestation (Thorndike, 1963). Y-variabel är då prestation, X-variabel intelligens. Denna approach har använts för att explorativt studera övriga faktorer som har inverkan på prestation och skolframgång (se t. ex. Magnusson, 1964).

Inom Örebroprojektet kommer studiet av över- underprestation att vara en av de viktigare uppgifterna. Metodiken kommer dessutom att användas vid mätning av icke-kognitiva variabler varvid dels intelligens, dels specifika snabbhetsfaktorer kommer att fungera som prediktorer (X-variabler).

Genom denna metodik komprimeras data och ger ett mått på en psykologiskt väsentlig variabel, vilken kan användas vid vidare analyser, t. ex. i kausalmodeller och i faktoranalyser. Användningen av metodiken är emellertid begränsad till ett första explorativt undersökningsstadium, då man generellt vill studera vilka variabler som samvarierar och eventuellt orsakar över- underprestation. I fortsättningen skall här användas beteckningen (E) för detta mått på generell över- underprestation, medan vi i det följande skall visa att det på ett senare stadium av en undersökning blir önskvärt att i stället studera en aspekt på över- underprestation, som är specifik för varje ytterligare variabel i systemet (Z-variabel), som samvarierar med över- underprestation.

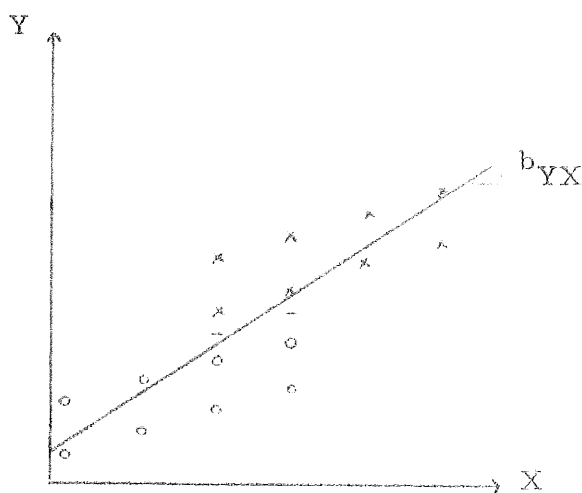
Om variabeln E (generell över- underprestation) relateras till en Z-variabel, t. ex. motivation, kan följande nackdelar noteras.

(a) Även om Y (prestation) helt kan förklaras av X (intelligens) och Z (motivation) följer inte att E (över- underprestation) låter sig förklaras av motivation.

(b) Sambandet mellan E (över- underprestation) och Z (motivation) är ej annat än i ett specialfall det samband som skulle råda mellan Y (prestation) och Z (motivation) för individer med samma X-värde (intelligens) även om en linjär modell gäller.

Situationen illustreras i figur 1. Variabeln Z (motivation) antas för överskådlighetens skull endast kunna anta två värden, 0 eller 1. I figurerna markeras individer med $Z=0$ med ringar och de med $Z=1$ med kryss. Medelskillnaden i prestation mellan individer med $Z=0$ och $Z=1$ är tre enheter för individer med samma intelligens. Individer med $Z=1$ har högre intelligens och prestation än individer med $Z=0$. I figur 1 är regressionslinjen som representeras av koefficienten b_{YX} dragen. Om för varje individ avvikelser från denna linje (E-värdet) mätes och prickas mot Z (motivation) erhålles figur 2. I denna figur är regressionslinjen b_{EZ} , som representerar styrkan på sambandet mellan över- underprestation och motivation inritad.

Figur 1. Regressionen Y (skolprestation) på X (intelligens) för en grupp individer med varierande Z (t.ex. motivation) för avläsning av E "prestation för sin intelligens".



x individ med $Z=1$

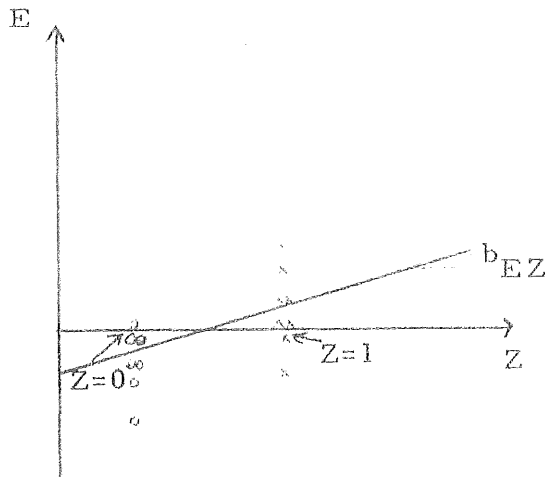
o individ med $Z=0$

$r_{XZ} > 0$

$r_{YZ} > 0$

— betingat medeltal

Figur 2. Individuella E-värden för individer med olika Z-värden.



Använda beteckningar: se fig. 1.

Vid ett studium av figur 1 finner vi

(a) att de individer med $Z=0$ som är mest intelligenta underpresterar mest,

(b) att de individer med $Z=1$, som är mest intelligenta överpresterar minst.

Man kan konstatera att E (under- överprestation) samvarierar med X (intelligens). X -variabeln är alltså icke konstanthållen vid residualmåttmetoden, så som den tillämpats av Thorndike m.fl.

Detta förhållande kan också visas på följande sätt.

Vi antar modellen

$$Y = b_{YX \cdot Z} X + b_{YZ \cdot X} Z + e$$

Sambandet mellan E och Z kan då skrivas

$$E = b_{EZ} Z, \text{ där}$$

$$b_{EZ} = b_{YZ} - b_{YX} b_{XZ} = b_{YZ \cdot X} (1 - r_{XZ}^2)$$

Den första av dessa likheter erhålles vid

$$b_{EZ} = \frac{\sum(Y - b_{YX}X)Z}{\sum Z^2} = b_{YZ} - b_{YX} b_{XZ}$$

Den andra får man genom att omforma normalekvationerna

$$\sum YX = b_{YX \cdot Z} \sum X^2 + b_{YZ \cdot X} \sum XZ$$

$$\sum YZ = b_{YX \cdot Z} \sum XZ + b_{YZ \cdot X} \sum Z^2$$

till

$$b_{YX} = b_{YX \cdot Z} + b_{YZ \cdot X} b_{ZX}$$

$$b_{YZ} = b_{YX \cdot Z} b_{XZ} + b_{YZ \cdot X}$$

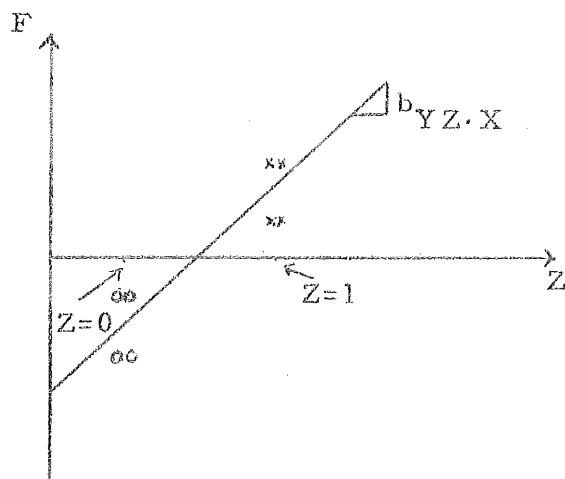
och därefter multiplicera den första av de två sista ekvationerna med b_{XZ} och subrahiera den från den senare av de två.

Sambandet mellan E (över- underprestation) och Z (motivation) beror alltså inte endast på hur prestationen samvarierar med intelligens och motivation utan också på hur X (intelligens) och Z (motivation) samvarierar.

Den tidigare använda metoden för beräkning av mått på över- underprestation innebär konstanthållande av intelligensvariabeln endast under den speciella omständigheten att intelligens ej samvarierar med Z-variabeln ifråga.

I stället för måttet E, avvikelser från regressionslinjen, beräknas i figur 3 avvikelser från betingade medeltal i observationerna från figur 1. Dessa senare avvikelser betecknas i fortsättningen F. Detta innebär att X (intelligens) hålls konstant. Betingade Y-medeltal är i figur 1 markerade för de två X-värden, som är representerade i båda grupperna $Z=0$ och $Z=1$ med tecknet —. I figur 3 har den regressionslinje dragits som kan betecknas $b_{YZ \cdot X}$. Dess lutning är uppenbarligen brantare än regressionslinjen i figur 2. Sambandet mellan det mått på över- underprestation, som vi infört i figur 3 (F) och variabeln Z (motivation) är starkare än sambandet mellan E och Z.

Figur 3. Individuella F-värden (avvikelser från betingade medeltal) för individer med olika Z-värden.



Använda beteckningar: se fig. 1.

Det sist nämnda angreppssättet är en tillämpning av vanlig enkel partialteknik. Vi har använt oss av $b_{YZ \cdot X}$, d. v. s. konstanthållit X (intelligens). Emellertid skall vi också arbeta med individuella mått på F, här benämnd specifik över- underprestation. Tekniken är att simultant anpassa modellen

$$Y = b_{YX \cdot Z} \cdot X + b_{YZ \cdot X} \cdot Z$$

och låta

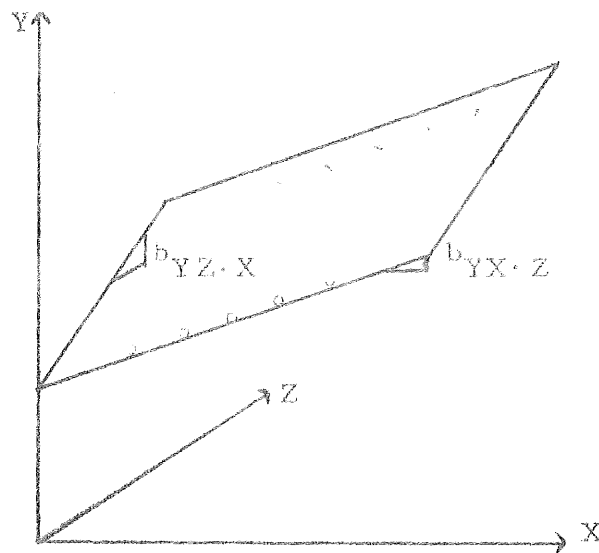
$$F = Y - b_{YX \cdot Z} X.$$

Då man därefter vill studera sambandet mellan F och Z kan man endast tillämpa förhållandet att $b_{FZ} = b_{YZ \cdot X}$.

Det är av intresse att b_{EZ} och b_{FZ} alltid har samma tecken och att $|b_{EZ}| < |b_{FZ}|$. Det kan också visas att om b_{EZ} visar sig vara signifikant skild från noll på en viss signifikansnivå, gäller detta å forteriori också b_{FZ} .

Den här redovisade metodiken för beräkning av specifik över- underprestation kan betecknas som en multipel ansats. I figur 4 visas ett exempel med samma beteckningar som tidigare använts. Fem individer har motivationen $Z=0$, fem andra har $Z=1$. Relationerna är i övrigt likartade exemplet i figur 1. Den multipla ansatsen innebär att ett plan anpassas, och ett sådant har inritats i figur 4. I detta har linearitet förutsatts. I vissa variabler inom Örebroprojektet har det visat sig att vissa relationer icke har linearitet. Den multipla ansats som har föreslagits här kan anpassas även till sådana situationer. I ett fall har t. ex. sociometrisk data varit Z-variabel. Den "mjuka" modellen, där hänsyn tas till krökta regressioner har där givit approximativt samma resultat som den "hårda" modellen, lineärmodellen.

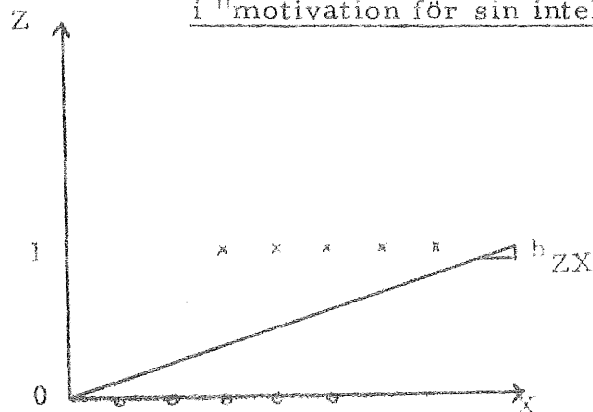
Figur 4. Ett lineärt plan anpassat till värden i X, Y och Z för en grupp om 10 individer.



Använda beteckningar: se fig. 1.

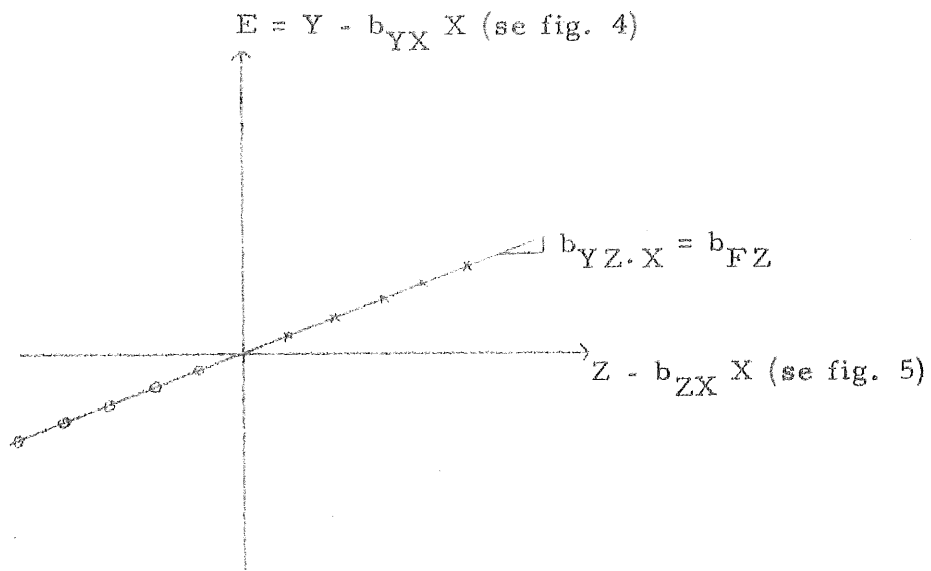
Det kan slutligen påvisas att regressionskoefficienten $b_{FZ} = b_{YZ \cdot X}$ kan uppfattas som den icke-partiella regressionskoefficienten för sambandet mellan $(Y - b_{YX}X)$ och $(Z - b_{ZX}X)$, d. v. s. sambandet mellan "prestation för sin intelligens" och "motivation för sin intelligens" i vårt exempel. Vi utgår från exemplet i figur 4. Vi prickar samma material i figur 5 där X prickats mot Z. Regressionslinjen b_{ZX} är markerad. I figur 6 har slutligen avsatts på X-axeln $Z - b_{ZX}X$ (som kan avläsas i figur 5) och på Y-axeln $Y - b_{YX}X$ (som kan avläsas från figur 4 om regressionslinjen med lutningen b_{YX} inritas). Regressionslinjen i figur 6 kommer att representera den partiella regressionskoefficienten $b_{YZ \cdot X}$ och korrelationskoefficienten kommer att vara den partiella korrelationskoefficienten $R_{YZ \cdot X}$.

Figur 5. Regressionen Z på X för avläsning av individuella värden i "motivation för sin intelligens".



Använda beteckningar: se fig. 1.

Figur 6. Sambandet mellan "prestation för sin intelligens" och "motivation för sin intelligens".



KAP. VII. FÖRSÖK TILL FÖRKLARING AV SAMBAND

A. Motivering för den mångsidiga uppläggnings

I tidigare anpassningsundersökningar har man i övervägande antalet fall arbetat med mått på ett symptom i taget, varvid data insamlats vid ett tillfälle eller inom en kort tidrymd. Instrumenten har växlat mellan olika anpassningsundersökningar. Man har därför erhållit mått på olika anpassningsvariabler, och resultaten från olika undersökningar blir inte jämförbara med varandra. De kan inte utan svårigheter ställas mot varandra för att öka vår kunskap om anpassningsmåttens karaktär och innebörd.

Genom sambandsanalys av data från olika instrument insamlade vid ett tillfälle eller inom en kort tidrymd kan man sammanföra symptomen i grupper med höga samband, syndrom. Resultaten av en sådan analys medger dessutom en bättre och säkrare tolkning av innebörden av enskilda mått på anpassning.

Ju mera varierad till sin karaktär den information är, som införes i en sådan sambandsanalys, desto säkrare blir tolkningen av innebörden i den syndrombild som erhålles. I Örebroprojektet finns eller kommer att finnas unika möjligheter till sådana tolkningar dels genom att rena anpassningssymtom täcks med mått av mycket olika karaktär och från olika källor, dels genom att systematiska uppföljningar genomföres, dels genom att det utöver mått på anpassning också finns en mångfald data av betydelse för tolkningen av innebörden i de rena anpassningsmått, nämligen a) andra personlighetsmått (intressen, attityder, värderingar, prestation, begåvning), b) fysiologiska mått och c) information om elevernas hem och familjebakgrund.

Sambandsanalyser av data från en tidpunkt eller en begränsad tidrymd av det slag som just berörts kan ge viss information om de erhållna syndromens karaktär och om deras betydelse i den totala anpassningsbilden, tillfälligt och på längre sikt. Kunskapen ökar emellertid väsentligt om syndrombilden kan studeras med data från olika tidpunkter genom a) jämförelse mellan korrelationsmatriser från olika tillfällen, vilket bl. a. ger information om i vilken utsträckning syndrombilden är densamma vid de två tillfällena, och b) studiet av samband mellan tillfällen för enskilda syndrom, vilket bl. a. ger information om i vilken utsträckning syndrombilden är den samma för individer mellan tillfällena.

Sambandsanalyser av det slag som just berörts, av data från ett

tillfälle eller från flera tillfällen, medger också studium av kausala samband.

B. Kausalanalys

Uppgiften att studera kausalitet i icke-experimentella situationer är mycket svår. Det finns dock metodik vid bearbetningen av data, som ger möjligheter till vissa slutsatser om kausalitet även i sådana fall.

En skillnad mellan olika slag av studier ligger i storleken av den risk man måste ta, då man drar slutsatser. Den mest väsentliga skillnaden i möjligheterna till säkra slutsatser ligger ej i om undersökningen är experimentell eller icke utan i om den försiggår i ett öppet eller i ett slutet system. Problematiken har närmare diskuterats i kapitel III. I Örebroprojektets öppna system kommer möjligheterna till kausalanalys att bli beroende av i vilken utsträckning relevanta variabler kan täckas av mätningar eller kontrolleras på annat sätt.

Oberoende och beroende variabler betecknas liksom tidigare, X respektive Y. De variabler som därutöver kan vara intressanta vid kausalitetsstudiet betecknas Z. De senare kan vara av två huvudtyper.

- a) Variabler, som påverkar Y men ej X: $X \longrightarrow Y$
 $\quad\quad\quad Z \nearrow$
- b) Variabler, som påverkar både X och Y: $X \longrightarrow Y$
 $\quad\quad\quad \nwarrow Z \nearrow$

I de två system, som representerar dessa fall betecknar pilarna orsaksriktning.

I fall a) finner vi Y predicerbar från X, då $b_{YX} > 0$. Vidare anger r^2 hur stor del av variansen i Y som är gemensam med X och orsakas av X. Övrig varians i Y kan hänföras till dels Z-variabler, dels mätfel. Om inverkan av Z kan elimineras med partialteknik kommer b_{YX} att bli oförändrad, medan r_{XY}^2 ökar till följd av variansminskning i Y. Detta förhållande framgår av följande formler för den partiella korrelations- respektive regressionskoefficienten.

$$|r_{XY \cdot Z}| = \frac{|r_{XY} - r_{XZ} r_{YZ}|}{\sqrt{1 - r_{XZ}^2} \sqrt{1 - r_{YZ}^2}} = \frac{|r_{XY}|}{\sqrt{1 - r_{YZ}^2}} \geq r_{XY}$$

$$b_{YX \cdot Z} = \frac{b_{YX} - b_{YZ} b_{ZX}}{1 - b_{ZX} b_{XZ}} = b_{YX}$$

då $r_{XZ} = b_{ZX} = 0$.

Utomliggande variabler av typ a) som utan vår kännedom påverkar den beroende men ej den oberoende variabeln inverkar alltså icke på regressionskoefficienten men däremot på korrelationskoefficienten.

I fall b) kommer vi om $r_{XY} \neq 0$ ej att kunna dra slutsatsen, att $X \longrightarrow Y$. Sambandet kan helt vara orsakat av den tredje variabelns (Z) inflytande på både X och Y. Kontroll av Z genom beräkning av $r_{XY \cdot Z}$ och $b_{YX \cdot Z}$ visar om sambandet är helt eller delvis skenbart. Om både $r_{XY \cdot Z}$ och $b_{YX \cdot Z}$ sjunker till 0 är hypotesen $X \longrightarrow Y$ ej verifierad.

C. Kausalmodeller

Ett hjälpmedel vid kausalanalys i icke-experimentella undersökningar är kausalmodellen.

Metodiken innebär att man konstruerar en modell av verkligheten i termer av variabler och kausalriktningar (Simon, 1957). Den uppställs som ett pilsystem, se t.ex. figur 7 (sid. 23). Modellen av verkligheten är här med samma nödvändighet som i andra sammanhang, där man använder modelltänkande, en förenkling. Den är förenad med explicita antaganden om vissa kända variablers funktion likaväl som om det inflytande okända variabler kan ha. Från modellen kan dras slutsatser eller uppställas hypoteser om korrelationer, regressioner eller vikt-koefficienter i multipla regressionsekvationer, vilka kan prövas. Om modellen ej ger prediktioner, som stämmer med empiriska fakta, kan den modifieras genom ändrade kausalitetsantaganden och/eller tillförande av nya variabler.

I allmänhet kan man inte lösa en uppsättning ekvationer, i vilken varje variabel är beroende av de återstående. Därför begränsar man sig i det kausala tänkandet till s. k. rekursiva system av ekvationer, som kan handhas med minsta kvadratmetoden. Där antas icke-reciproka samband, d. v. s. om X_i orsakar X_j så kan inte samtidigt X_j orsaka X_i . Man kan då skriva ekvationer, där X_1 inte beror på någon av övriga förekommande variabler, X_2 beror på endast X_1 , X_3 på X_1 och X_2 o. s. v.

Ett rekursivt ekvationssystem kan således skrivas:

$$X_1 = e_1$$

$$X_2 = b_{21} \cdot X_1 + e_2$$

$$X_3 = b_{31} \cdot X_1 + b_{32} \cdot X_2 + e_3$$

.

.

.

$$X_k = b_{k1} \cdot X_1 + b_{k2} \cdot X_2 + \dots \dots \dots b_{k,k-1} \cdot X_{k-1} + e_k$$

Ett sådant system av ekvationer kan anknytas till en kausalmodell och göra det möjligt att uppställa prövbara hypoteser, med vilkas hjälp modellen kan verifieras.

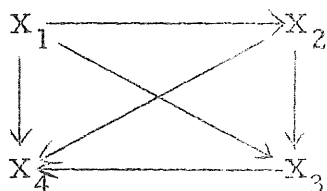
För att testa modellen måste vi anta att feltermerna är oberoende. Vidare måste vi, för att kunna ställa upp hypoteser, som går att pröva, utforma modellen så att någon regressionskoefficient sätts $= 0$.

Exempel på kausalmodeller av enklaste slag har redan förekommit i samband med definitionen av två slags störande variabler. De antaganden, som modellen måste kompletteras med innebär bl. a. att andra variabler än X , Y och Z ej har någon betydelse i sammanhanget. Av de nämnda enkla modellerna framgår två tänkbara prövningar, användningen av partialkorrelationsteknik och jämförelser mellan tendenser i korrelation och regression vid förändrade betingelser.

För att modelltänkandet skall bli fruktbart måste antalet i modellen ingående variabler liksom antalet hypotetiska nollsamband vara så stort att ett väsentligt antal hypoteser kan uppställas och prövas. Vid ett litet antal hypoteser kommer alternativa förklaringar till erhållna resultat att kvarstå.

En kausalmodell enligt ett rekursivt ekvationssystem med fyra variabler presenteras i figur 7.

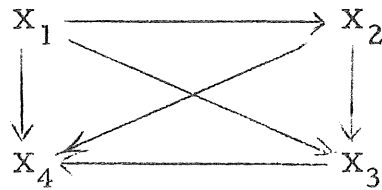
Figur 7. Kausalmodell med fyra variabler.



Om man i denna modell sätter en viss partialregressionskoefficient $= 0$ innebär detta en hypotes att det inte finns någon direkt orsakslänk mellan de två variablerna. Det kan emellertid finnas indirekta länkar, som ger koefficienter som är $\neq 0$ för korrelationer mellan variablerna i fråga.

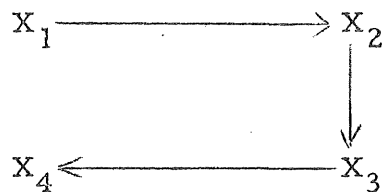
I figur 8 antas att $r_{13.2} = 0$. Genom inflytande från variabel X_2 kommer emellertid r_{13} att prediceras att bli $\neq 0$.

Figur 8. Kausalmodell med möjlighet till en nollhypotes.



För varje partiell regressionskoefficient, som sättes = 0 erhåller man ett förhållande som data måste satisfiera för att modellen skall vara konsistent med data. Om t. ex. tre partiella regressionskoefficienter sätts = 0 som i figur 9 blir antalet obekanta tre färre än antalet ekvationer. Vi får tre förhållanden som data måste passa in på och kan uppställa tre prediktionsekvationer för testning av modellen.

Figur 9. Enkel orsakskedja.



Modellen i figur 9 ger de tre ekvationerna

$$r_{13 \cdot 2} = 0$$

$$r_{24 \cdot 3} = 0$$

$$r_{14 \cdot 2} = 0 \text{ eller } r_{14 \cdot 3} = 0$$

Av bl. a. den vanliga formeln för partialkorrelation framgår att om

$$r_{13 \cdot 2} = 0$$

$$\text{så är } r_{13} = r_{12} \cdot r_{23}$$

och att om

$$r_{14 \cdot 3} = 0$$

$$\text{så är } r_{14} = r_{13} \cdot r_{34}$$

Alltså kan vid modellen i figur 9, vilken representerar en enkel orsakskedja, korrelationen mellan första och sista variabeln i kedjan

skrivas

$$r_{14} = r_{12} \cdot r_{23} \cdot r_{34}$$

Generellt kan i ett sådant fall skrivas

$$r_{1n} = r_{12} \cdot r_{23} \cdot r_{34} \cdot \dots \cdot r_{(n-1)n}$$

Då två variabler i en modell ej är sammanbundna med orsakspil testas modellen genom att en partialkorrelation sätts lika med noll. Därvid kontrolleras variabler som i orsakskedjan ligger mellan eller före, men ej variabler som ligger efter, d. v. s. beroende variabler. För det fall som illustreras i figur 9 gäller sålunda att $r_{13.2} = 0$, medan en kontroll av även X_4 skulle kunna ge en felaktig tolkning. Denna partialkorrelation

$$r_{13.24} = \frac{r_{13.2} - r_{14.2} \cdot r_{34.2}}{\sqrt{1 - r_{14.2}^2} \sqrt{1 - r_{34.2}^2}}$$

kommer ej att försvinna samtidigt med $r_{13.2}$ om modellen är riktig i avseende på pilarna både mellan X_1 och X_4 och mellan X_3 och X_4 .

I större modeller kan man börja med hypoteser, som berör variabler, som antas vara före de andra i orsakskedjorna. Då man möter betydande diskrepanser mellan prediktion och data får man göra modifieringar i modellen innan man prövar den vidare.

D. Kontroll av omätbara variabler

I en fältundersökning i ett öppet system föreligger stor risk att man ej identifierat och mätt alla variabler med tänkbart inflytande. En metod att pröva en modell med okända variabler har föreslagits av Blalock (1964). Vi tänker oss att variabeln Z i den förut diskuterade modellen är okänd eller otillgänglig för mätning (se sid. 21). Tidigare nämnda prediktionsekvationer kan då ej uppställas, då de förutsätter kännedom om t. ex. r_{XZ} och r_{YZ} .

En ny mätbar variabel definieras, som har så högt samband som möjligt med en oberoende variabel X men mycket lågt eller inget samband alls med Z. Om denna nya variabel benämnes W erhålles modellen $W \rightarrow X \xrightarrow{Z} Y$. Z står för alla tänkbara omätbara och okända variabler, som kan tänkas ha orsakat att en observerad korrelation $r_{XY} \neq 0$. I detta läge kan prediceras att $r_{XY.W} > r_{XY}$, d. v. s. att en utpartialisering av W medför en höjning av korrelationen, eftersom Z's inflytande ökar proportionellt sett, då en del med Z oberoende varians subtraheras från X. Samtidigt kan förväntas att regressions-

koefficienten ej påverkas av partialisering. Då detta utfall ej är unikt för den skisserade situationen behöver vi en alternativ prövning av modellen. Denna kan ske genom Blalock's metod, som innebär att försöksgruppen grupperas efter variabel W. Korrelationskoefficienter beräknas mellan grupperas medeltal i stället för som tidigare individvärden. Mellan gruppssvarianser i W kommer att i det närmaste bli lika stor som totalvariansen i W. Variabler som är högt korrelerade med W kommer att få måttlig sänkning i varians medan en okorrelerad variabel får en varians i grupperade data, som minskar väsentligt och på ett sätt som kan beräknas.

Enligt antagandena samvarierar X med W, medan Z-variabeln antas vara väsentligt lägre eller ej alls korrelerad med W. Inflytande på r_{XY} från Z kommer att minska och om modellen stämmer, d. v. s. om bl. a. $r_{XY.Z} = 0$ kommer korrelationer och regressioner på grupperade data mellan variablerna X och Y att sjunka, och om antaganden visar god överensstämmelse med verkligheten och om grupperna är tillräckligt stora blir $r_{XY} = 0$.

Då man på detta sätt förändrar enheten för analysen från enskilda personer till grupperas medeltal påverkar alltså detta mängden variation i alla variabler inklusive okända sådana, som kan påverka den beroende variabeln. Påverkan är givetvis starkare ju större enheter (grupper) man arbetar med. Effekten på variationen i varje variabel är emellertid en funktion av hur dessa enheter bildas. Om individerna slumpas tillsammans till enheter blir effekten densamma på alla variabler och ingen systematisk inverkan kan förväntas på korrelationer och regressioner. Sker sammanslagning helt utifrån variabel X kommer variationen i denna variabel att bli praktiskt taget oförändrad, variabler som är korrelerade med X får något minskad variation, mindre minskning ju högre korrelationen är. Korrelationer mellan variabler i ett system där X ingår kommer att förändras på ett systematiskt sätt. Dessa effekter är sammanställda i tabell 1. Det är denna systematik, som kan utnyttjas för att kontrollera okända variabler i ett öppet system. Blalock har genomfört beräkningar på artificiella data, som synes visa, att metoden kan vara effektiv. Emellertid föreligger risken att grupperingsvariabeln W är korrelerad med den okända och störande variabeln. Blalocks räkneexempel visar att metoden ger större avvikelser från verkliga förhållanden först då denna icke önskade korrelation r_{WZ} är ganska hög (i exemplet 0.59) och högre än r_{WX} .

Tabell 1. Effekter på r och b av gruppering efter oberoende respektive beroende variabel då $X \rightarrow Y$.

Grupperingsmetod	Mått	Förändring med större enheter
Randomisering	r_{XY}	invariant
	b_{YX}	invariant
Gruppering efter X	r_{XY}	ökar
	b_{YX}	invariant
Gruppering efter Y	r_{XY}	ökar
	b_{YX}	ökar

Eftersom det sällan existerar kunskap om hur de störande variablerna opererar, rekommenderas forskaren att gruppera efter ett antal distinkta variabler, som orsakar den oberoende variabeln, alltså att flerfaldiga W .

Bland de svårigheter som möter i vårt fall där vi har ett helt öppet system märks reliabilitetsproblem. Här skall noteras att då partialkorrelationer används i prediktionsekvationer måste reliabilitetskorrektur vara utförd på kontrollvariablerna (i våra exempel Z) om partialkorrelationen i förekommande fall teoretiskt skall kunna förväntas närma sig 0. I annat fall blir data svårtolkade. Bland de kontrollvariabler som används i föreliggande undersökning märks bakgrundsvariabler. Dessas reliabilitet är ofta okänd. I regel står variabler endast som indikatorer på psykologiskt verk samma variabler varvid bristande identitet mellan åsyftad variabel och indikator fungerar som mätfel. Bakgrundsdata måste av den anledningen transformeras med varsamhet så att reliabiliteten ej onödigtvis minskas. Kategorisering får t. ex. ej ske till alltför få kategorier.

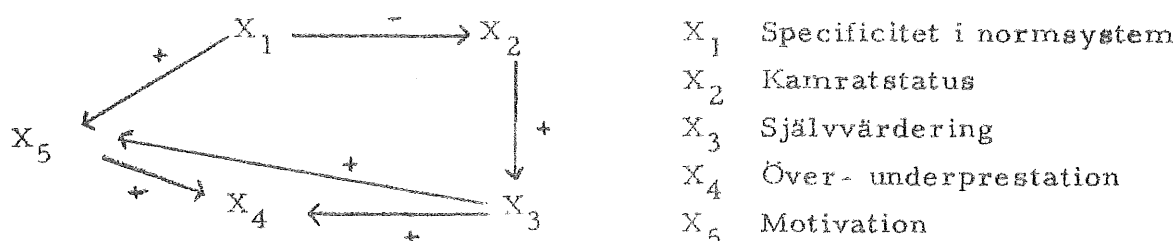
E. Exempel på kausalmodeller i Örebroprojektet

Inom Örebroprojektet kommer intresset att knytas till prestation och anpassning som beroende variabler. Variabler av olika slag, som kan ha orsakssamband med de beroende variablerna skall prövas. Därvid måste små starkt förenklade system konstrueras, prövas och modifieras. På grund av den stora komplexiteten i systemet måste vi utgå från att i egentlig mening goda prediktioner ej kan göras i undersökningens nuvarande fas. Vår avsikt är att pröva den nämnda metodiken med kausalmodeller för att få hypoteser till kommande undersökningar, bl. a. uppföljningar som kan planläggas med bestämda *à priori* modeller som utgångspunkt.

Till den nämnda explorativa kausalforskningen skall också läggas vår möjlighet att utnyttja uppföljningsundersökningar och däri inbyggda tidsförskjutningar mellan oberoende och beroende variabler för utformningen av andra typer av kausalmodeller. (Jfr ex post facto-undersökningar med tidsförskjutning på förändrade elever, B:2).

Arbetssättet kan exemplifieras i anknytning till variabler inom Örebroprojektet. I figur 10 visas en preliminär modell, där X_1 - specificitet i normsystem (avvikelse från allmänt gällande normsystem för beteendet) i barnets hem - antas orsaka förändringar i kamratstatus. Kamraternas uppfattning av barnet antas vidare orsaka ändrad självvärdering, vilket i sin tur inverkar på skolbeteenden, t. ex. över- underprestation. Modellen blir likformig med den som allmänt har diskuterats i figur 9 och kan testas på samma sätt. Om därvid $b_{13.2}$ mot hypotesen ej försvinner (blir 0) kan det bli naturligt att ändra modellen så, att en pil införes från X_1 till X_3 , d. v. s. normsystemet kan hypotetiskt antas inverka direkt på självvärdering. Låt oss i stället anta att modellen stämmer i alla avseenden utom att $b_{14.2}$ ej försvinner utan har ett markerat positivt värde liksom att $|r_{12} \cdot r_{23} \cdot r_{34}| < |r_{14}|$. Då det förefaller orimligt att ett normsystem i hemmet direkt kan inverka på barnets prestation i skolan behöver vi komplettera modellen med en ny variabel. Vi erhåller modellen i figur 10 om vi tillägger motivation (X_5) med där redovisade orsaksriktningar, som torde vara rimligare att anta. Denna modell provas med bl. a. undersökning av om $b_{14.52}$ går mot noll och om $|r_{15} \cdot r_{54}| > |r_{14}|$. I det fortsatta arbetet kan man tillföra ytterligare variabler, t. ex. utbildningsgrupp som orsak till motivation, eller tillägga orsakspilar inom systemet, t. ex. att självvärdering positivt påverkar motivation.

Figur 10. Kausalmodell.



Variabler av den typ som står i centrum i det valda exemplet, normsystem, kommer att mätas i ökad omfattning i följande etapper av projektet. Det gäller där hemmets och lärarens centrala roll som uppfostrare och orsaksfaktorer till anpassning.

Det kan för många kausalsamband vara en alltför grov förenkling att anta att en och samma kausalmodell fungerar för alla individer och för alla grupper. Man kan också förvänta betydelsefulla interaktionseffekter mellan variabler, som ingår i samma kausalmodell. Angreppssätten blir här parallella till dem som i andra avsnitt skisseras för andra problem, studium av undergrupper, moderatoranalys m.m.

Här skall endast i grova drag ges ett exempel på hur en modell, som tar hänsyn till denna problematik skulle kunna se ut.

De värderingar som gäller i hemmet har enligt denna modell i många avseenden en avgörande betydelse för barnets egna värderingar, för barnets anpassning och anpassningsmöjligheter och prestationer på olika områden. Vi skall i detta exempel ta hänsyn till två värderingsobjekt, som kan ha olika betydelse i ett hem. Föräldrarnas utbildningsnivå torde ha betydelse för vilket värdesystem som utvecklas, varför X_1 - utbildningsnivå - satts först i orsakskedjan i modellen.

Det ena värderingsobjektet rör intellektuell verksamhet, studier och liknande. Det andra rör fysiska prestationer, sådant som i skolan bäst svarar mot gymnastikämnet.

Mot varje värderingsobjekt svarar alltså vissa skolprestationer, som kan tas som beroende variabler i modellen.

De olika värderingsobjekten tänkes nu manifesteras sig i intresse för barnets verksamhet på de olika områdena, i klander och beröm respektive likgiltighet.

Intresset för att barnet skall vara duktigt inom ett område kan nu ta sig olika uttryck. Beröm och klander kan utdelas för prestation i relation till förutsättningar, i modellen intelligens respektive fysiska förutsättningar. Å andra sidan kan kraven vara absoluta och ej ta hänsyn till förutsättningarna. Vi har alltså dels fallet med, dels fallet utan interaktion mellan kravnivå och förutsättningar. Detta kan införas i modellen som två variabler, som båda torde vara mätbara: Intressets styrka respektive kravnivåns anpassning till barnets förutsättningar.

I figur 11 a betraktas den grupp, där vi möter intresse för intellektuell verksamhet men ej för gymnastik och idrott. Barnets prestationer belönas i studier men ej i idrott, och detta intresse från hem-

mets sida inverkar på barnets upplevelse av sig själv och dess säkerhet och trygghet. Det avgörande för effekten är interaktionen mellan intresse och förutsättningar. Vid högt utvecklad interaktion utvecklas barnet till realistisk uppfattning om sig självt och till säkerhet och trygghet. Vid absoluta krav däremot kommer endast de barn som faktiskt har de bästa anlagen att kunna svara mot kraven, de övriga får utveckla försvar mot för höga krav. Detta kan yttra sig i skolleda, överambition eller andra mindre önskvärda reaktioner, d. v. s. i tecken på missanpassning.

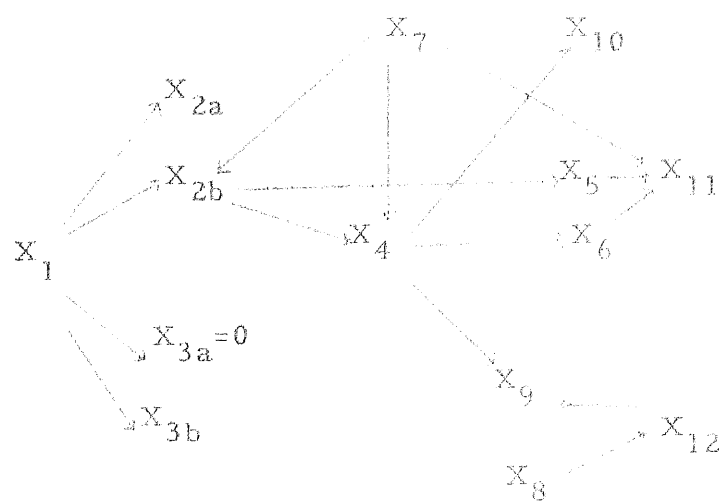
För barnen i denna grupp inverkar hemmet inte alls på de idrottsliga prestationerna. Dessa har emellertid betydelse för kamratstatus och det blir anlagen som blir orsaksfaktorer jämte självvärdering, som utvecklas under inverkan av bl. a. hemkraven.

I figur 11 b visas situationen för en grupp barn med motsatta värderingar i hemmet. Det framgår av pilsystemet hur självvärderingen här kommer att grundläggas utifrån hur barnet svarar mot hemmets förväntningar på idrottsliga prestationer medan studieprestationerna och motivationen för sådan verksamhet byggs upp utifrån barnets intellektuella förutsättningar.

På analogt sätt kan man konstruera modeller för de två övriga fallen, då i hemmet båda typerna av verksamhet uppmuntras och där förväntningar på barnets prestationer inom båda områdena är starka respektive då inget intresse knyter sig till något av de nämnda områdena för barnets verksamhet.

Vi måste vidare förutsätta att samma modell ej kan komma att gälla för alla olika åldrar. I de givna exemplen kan vi t. ex. förvänta att de orsaker till ett visst beteende som ligger i hemmets värderingar kommer att avta i betydelse för en del av de införda variablerna. Kamratstatus kommer i stället att öka i betydelse och ett antal orsaks-pilar bör förmodligen inritas från denna variabel mot anpassning, självvärdering, motivation etc. då vi arbetar med högre årskurser. Kamratgruppernas värderingar ersätter hemmets.

Figur 11 a. Kausalmodell.



- X₁ Föräldrarnas utbildningsnivå
- X₂ a: Hemmets intresse: studier
b: Kravnivåns anpassning till barnets förutsättningar (studier)
- X₃ a: Hemmets intresse: idrott etc.
b: Kravnivåns anpassning till barnets förutsättningar (idrott)

X₄ Säkerhet och självvärdering

X₅ Motivation

X₆ Spänning

X₇ Intelligens

X₈ Anlag för idrott etc.

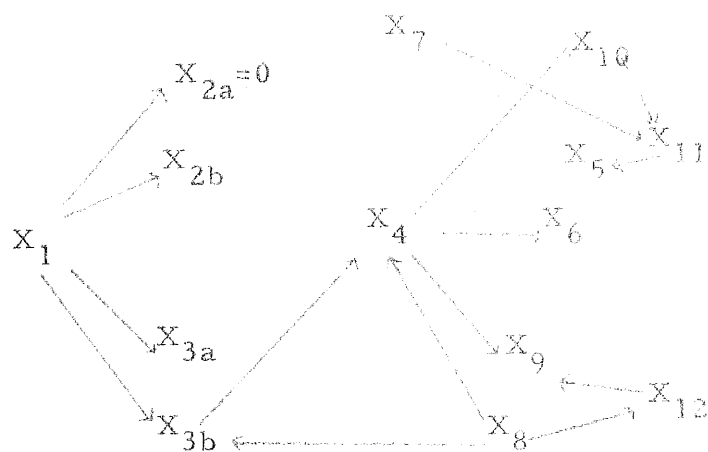
X₉ Kamratstatus

X₁₀ Anpassning

X₁₁ Över- underprestation

X₁₂ Gymnastikbetyg

Figur 11 b. Kausalmodell.



F. Dimensionsanalys

Inte minst vid modellkonstruktion kommer antalet variabler snabbt att växa till oöverskådlighet. Det är nödvändigt att arbeta med så få och så effektiva variabler som möjligt. Även för prediktionsändamål är det önskvärt att använda variabler täcker av relevanta områden. Ett frågeformulär eller ett testbatteri lämnar ej avsedd information om data ej kan sammanställas till ett mindre antal centrala skalor. Härvid använder vi oss av dimensionsanalytiska metoder.

Komponentanalys med följande orthogonal transformation enligt varimaxkriteriet har använts för konstruktion av indexskalor på elevenkäter. Metoden medger en enkel bestämning av transformerade individvärden på de matematiskt definierade indexskalorna (principalaxlarna).

Kanonisk faktoranalys har använts för att strukturera materialet och ge en modell, där latent variabler förklarar kovariansen mellan de manifesta variabler, där vi förfogar över data. Denna struktur används för teoretiserande hypoteskonstruktion och kan i vissa fall betraktas som en kausalmodell, som kan provas som en sådan, nu eller i senare faser av undersökningen. Vidare ger strukturen i materialet möjligheter till effektivisering av mätinstrument i uppföljningar.

KAP. VIII. KLASSIFIKATION

A. Motivering för klassifikation

Med de dimensionsanalytiska metoderna studeras materialets totalstruktur. Men analysen av enskilda symtom, sammanförandet av symtom till syndrom, samt studiet av innebörden och betydelsen i de så erhållna syndromen är icke den slutliga målsättningen utan endast ett led i arbetet att skapa bättre förutsättningar att uppnå det som är huvudmålsättningen, studiet av individer. Eftersom undersökningen syftar till att om möjligt belysa generellt giltiga drag i anpassningsproblematiken är det knappast meningsfullt att syssla med en individ i taget. Individerna bör grupperas i syndromgrupper. Vi söker därvid klasser, som med utnyttjande av all tillgänglig information innehåller individer som liknar varandra så mycket som möjligt och är så olika individerna i andra klasser som möjligt. En jämförelse av individer eller grupper av individer som uttagits på en enda variabel, vilket är det sätt som i övervägande antalet anpassningsundersökningar har använts, kan förutses ge resultat som är mycket svårtolkbara om de inte blir direkt missvisande. Detta har sin grund i det förhållande att det sannolikt är interaktionen mellan olika variabler och det sätt på vilket dessa kommer till uttryck i symtom, som har avgörande betydelse för i vilken utsträckning beteendet kan betecknas som anpassat eller ej. Den metodik som använts för grupperingen av individer bör därför uppfylla villkoren dels att utnyttja informationen från olika fördelningar av data på en gång, dels att utnyttja interaktionen mellan olika variabler. Traditionella metoder byggda på regressionsanalys uppfyller endast i undantagsfall och i begränsad utsträckning dessa villkor. Meningsfulla latent variabler kan ej definieras utan att man tagit i beaktande de enskilda individernas variabelprofiler.

Definitionen av latent variabler betyder icke att klasser av individer har påvisats. För att nå detta syfte då två eller tre variabler studeras kan grafiska metoder användas. Man kan då finna anhopningar i den tredimensionella rymden av punkter (= individuella profiler) som meningsfullt kan medge en indelning av rymden, så att individer inom en del hänförs till en klass. I de fall då mer än tre latent variabler studeras kan i stället för en dimensionsanalytisk metod användas en objektinriktad metod, för kontinuerliga skalor latent profilanalys eller vid kända klasser, diskriminantanalys.

B. Latent profilanalys

Latent profilanalys (LPA) är en utveckling av Lazarsfelds (1959) latent strukturanalys, som i föreliggande undersökning används enligt metodik och program utarbetade av Mårdberg (1966).

Denna metod går icke via dimensionsanalys utan innebär i stället punktskattningar av klasser (anhopningar av punkter) i den latent rum, som modellen bygger på. Basmaterialens variabler behöver inte uppfylla några fordringar på normalfördelning. Vid parameterskattningarna utnyttjas hela materialet på en gång, vilket innebär maximal effektivitet. LPA har använts på föreliggande lärarskattningar, där utifrån sju manifesta skalaer en meningsfull indelning i tre klasser har kunnat utföras. Vidare används modellen på data som erhållits med semantisk differentialteknik och intresseformulär.

Studiet av individegenskaper har i tidigare undersökningar i huvudsak koncentrerats kring genomsnittliga förändringar i måtten, utvecklingskurvor. Detta har gällt både anpassningsmått, andra personlighetsmått och fysiologiska mått. Sådana genomsnittsmått kan inte ge annat än referensramar, men är givetvis ändå nödvändiga. Det är emellertid alldeles tydligt att den individuella utvecklingen i flertalet fall fortskrider på ett annorlunda sätt än det som illustreras av en genomsnittlig utvecklingskurva. Det är därför angeläget att studera relativa förändringar hos de grupper och klasser av individer som kan bildas med den metodik som har utnyttjats i det föregående. Den information som ligger i skillnader i relativa förändringar kan nämligen också utnyttjas för klassificering av individer i meningsfulla kategorier. Här avses i första hand de förändringar som erhålles för mått på anpassning. När klasser har karakteriserats med avseende på dessa mått kan relativa förändringar relateras till a) andra mått (personlighetsmått, fysiologiska mått, bakgrundsvariabler) och b) händelser. Ett exempel: Vad skiljer den grupp som gått från normal eller god anpassning till dålig anpassning över en viss tidsperiod från den grupp, som kvarstår i god anpassning, med avseende på dels a) andra mått vid tidpunkten 1 och dels b) händelser under mellanliggande period för individerna i respektive grupper, t. ex. lärarbyte, flyttning från ett bostadsområde till ett annat, skilsmässa i hemmet, nära kamraters avflyttning, etc.

C. Moderatoranalys

Den struktur, som vi försöker kartlägga, kommer att visa variationer mellan olika grupper och klasser. Vi skiljer i alla bearbetningar

på könen, mycket ofta studeras utbildnings- eller socialgrupper var för sig. Med andra tekniker erhållna klasser, kommer att på liknande sätt särhållas i olika bearbetningar. Grunden för gruppering eller klassifikation benämnes här moderatorvariabler. Analyser med särhållande av undergrupper är en form av moderatoranalys.

Då uppdelning önskas i undergrupper efter flera indelningsgrunder kommer vi trots den stora försöksgruppen att slutligen arbeta med alltför små undergrupper för att resultaten skall kunna tolkas. Detta utgör en begränsning av möjligheterna till simultan moderatoranalys enligt flera indelningsgrunder.

I viss mån kan denna nackdel undanröjas genom tillämpning av kovariansanalys eller en metod, utvecklad av Saunders (1956). Vid kovariansanalys kan hela materialet användas för skattningar av undergruppers parametrar. Metoden innebär en korrigering för differenser mellan grupper på olika nivå i fråga om sambands styrka. Enligt Saunders "moderator multiple regression" kan samma korrigering göras på en kontinuerlig moderatorvariabel.

LITTERATURFÖRTECKNING

- Blalock, H.M. Causal inferences in nonexperimental research.
Univer. of North Carolina 1964.
- Kollas!* Brunswik, E. Organismic achievement and environmental probability.
Psychol. Rev., 1943, 50, sid. 255-272.
- Gillis, J. - Schneider, C. The historical preconditions of representative design, i K.R. Hammond, The psychology of Egon Brunswik. New York 1966.
- Lazarsfeld, P.F. The logic and mathematical foundation of latent structure analysis. The interpretation and computation of some latent structure, i Stouffer (Ed.), Studies in social psychology in World War II, Vol. 4: Measurement and prediction. Princeton 1959.
- Magnusson, D. Anpassning och skolprestation. Stockholm 1964.
Stencil.
- Magnusson, D. - Dunér, A. - Beckne, R. Anpassning, beteende och prestation - Örebroprojektet, Rapport nr I, Planläggning. Stockholm 1965. Stencil.
- Mårdberg, B. Statistiska klassificeringsmetoder med specifika tillämpningar inom arbetspsykologi. Stockholm 1966.
Stencil.
- Saunders, D.R. Moderator variables in prediction. Ed. Psychol. Meas. 1956, 16, sid. 209-222.
- Simon, H.A. Models of man. New York 1957.
- Thorndike, R. The concepts of over- and underachievement.
New York 1963.

FÖRTECKNING ÖVER INOM PROJEKTET
TIDIGARE UTGIVNA RAPPORTER

Huvudrapporter

I. Planläggning

David Magnusson - Anders Dunér - Rolf Beckne
(Mars 1965)

II. Sammanställning av tidigare forskning

David Magnusson - Rolf Beckne
(Mars 1967)

Delstudier

1. En analys av data från elevenkäter

Rolf Beckne
(December 1966)

2. Yrkesenkät, intresseschema och yrkesvalslärares bedömning
i årskurs 8

Dorothy Been - Göran Zetterblom
(Maj 1967)