

# PERSONLIGHETSTEST I ARBETSLIVET: HISTORIK OCH AKTUELL FORSKNING

Lennart Sjöberg

SSE/EFI Working Paper Series in Business Administration No. 2000:9

## Innehåll

<b>Sammanfattning</b> .....	3
<b>Inledning</b> .....	4
<b>Individfaktorer i allmänhet och personlighet i synnerhet</b> .....	5
<b>Den förkättrade intelligensen</b> .....	9
<b>Felfaktorer i mätning och bedömning av personlighet</b> .....	10
<b>Skillnader mellan länder i fråga om testanvändning</b> .....	12
<b>Hur mäter man personlighet i Sverige?</b> .....	14
MBTI.....	14
Thomassystemet.....	17
OPQ.....	19
16PF.....	19
U-mantestet, grafologi, astrologi.....	20
DMT.....	21
Holismen.....	22
Slutord om personlighetsmätning i Sverige.....	23
<b>Framtiden</b> .....	25
Begåvningsstest.....	25
Emotionell intelligens.....	26
Integritet.....	26
Social önskvärdhet.....	27
Simuleringsövningar.....	28
Slutsatser.....	28
<b>Referenser</b> .....	29

## Sammanfattning

Människor är olika, mycket olika. Individfaktorn har stor betydelse i många sammanhang, inte minst inom näringslivet. Hur ska man då bära sig åt för att göra ett urval och hitta den lämpligaste kandidaten? Under de senaste 10 åren har många intressanta rön gjorts i den psykologiska forskningen och vissa tidigare slutsatser måste nu revideras, bl a om personlighetens roll. Det har visat sig att personlighet är av större betydelse för framgång i yrkeslivet än vad man tidigare trodde, och intelligens - som visserligen fortfarande anses vara viktigast - ses nu i ljuset av nyare forskning som breddat begreppet betydligt. Intervju-metodiken har utvecklats kraftigt och med hjälp av strukturerade intervjuer kan man bidra till prognosförmågan. Även andra typer av data som biografiska data uppmärksammas allt mera. Det finns stabilitet i individernas beteende (vad den nu beror på) och principen "more of the same" står sig som den bästa grunden för prognos, när den kan användas. Psykologiska test utvecklas emellertid alltför långsamt och praktiken ligger decennier efter forskningens ståndpunkt. I ett land som Sverige tycks ännu kvalitetskraven vara alltför låga från testkonsumenternas sida. Nyckelgrupper som ekonomer och socionomer bör få förbättrad utbildning om test - vetenskapligt grundad och kritisk - liksom psykologer.

## Inledning

Vad ligger bakom att somliga företag går bra och andra dåligt? Många faktorer spelar givetvis in. En faktor är individerna, kanske särskilt individerna i nyckelbefattningar som t ex företagsledare. Psykologi handlar till största delen om individer, och är därför relevant för frågan om företagets framgång, även om det är lätt att inse att individen inte är den enda faktorn. Denna artikel handlar om aktuell urvalspsykologi i allmänhet och urval av ledare i synnerhet. Den är en uppföljning av arbeten jag genomförde för ca 10 år sedan [171; 172]<sup>1</sup> - det har varit intressant att notera att 90-talet faktiskt har inneburit nya resultat som leder till delvis anorlunda slutsatser.

Vad är det då hos individerna som bör beaktas vid urval och rekrytering? Frågan är uppenbarligen ytterst viktig - misstag är kostsamma och vanliga [43; 209]. Hogan menade sig ha stöd för påståendet av minst 50 % av amerikanska chefer misslyckades [77]; många exempel på misslyckat militärt ledarskap ges i en läsvärd bok av Dixon [41]. Framgångsrika ledare är självsäkra, dominerande och energiska. De fattar kloka strategiska beslut [137]. De strävar efter makt och ansvar [6]. Men bakom en sådan fasad kan dölja sig arrogans, självgodhet, självupptagenhet och insensitivitet, egenskaper och motiv som kan vara svåra att upptäcka framför allt för högre chefer som fattar beslutet om befordran. Den misslyckade chefen är skicklig på att framstå som "trevlig", konciliant och smickrande uppåt i hierarkin. Hur han eller hon betar sig mot sina underställda medarbetare kommer alltför sent till den högsta ledningens kännedom. Den misslyckade chefen lyssnar sällan på kritik, det är ett av skälen till att han eller hon är misslyckad. Medarbetarna förstår att kritik inte uppskattas och kan leda till negativa konsekvenser för deras del, så de håller tyst. Den misslyckade chefen är ett framgångsrikt filter - informationen om hans eller hennes fungerande når inte uppåt i organisationen. Förrän det är för sent.

Vad betyder personlighet och vad betyder miljö för framgång i yrkeslivet? Jag ska behandla dessa frågor och göra ett avstamp i modern psykologisk forskning. Som vi ska se har mycket hänt under de senaste decennierna inom forskningen - mycket som ännu är okänt för många och mycket som inte påverkat praxis trots att så borde ha skett.

Låt mig emellertid börja med en påminnelse om alternativen till psykologiska test. Denna artikel ger många kritiska kommentarer till test - även till de bästa. En och annan läsare kanske vill dra den slutsatsen att test inte är mycket att ha och att andra metoder - vilka de nu är - vid t ex urval är mycket bättre. Som jag ska visa är det emellertid helt fel att dra den slutsatsen. Test är ofta kostnadseffektiva, och alternativen sämre eller värdelösa. Vissa undantag finns, och de kommer att diskuteras, men då handlar det oftast om att testen kan kompletteras med annan information, inte att de helt bör ersättas. All information om individer är felbemängd, även test, men vi vet speciellt mycket om test och om hur deras fel kan hanteras och mätas. Barnet ska alltså behållas, badvattnet kastas ut.

En intressant historisk parallell är det kinesiska systemet för urval till tjänstemannabefatt-

---

1 . Siffror inom hakparentes hänvisar till den numrerade referenslistan.

ningar [89] . Detta byggde på vad vi skulle kalla verbala test och kunskapsprov. Ytterst få passerade denna eklut (vilket måste ha gjort den effektiv, se nedan om urvalskvotens betydelse), vid en tid när bördens var den viktiga dimensionen för befordran i Europa. Det kinesiska systemet framstår som sympatiskt vid en jämförelse, även om det aldrig tillämpades fullt ut och börd och kontakter fortsatte att spela stor roll för karriären [46] . Kontakter, social bakgrund och skickligt utövad inställsamhet och opportunistik är inte vad vi vill ha som grund för urval och befordran i vårt samhälle. Test är rättvisa - testningen när den är som bäst är ett "sanningens ögonblick".

## Individfaktorer i allmänhet och personlighet i synnerhet

Personlighet kan studeras på många olika sätt. Det som kommit att dominera i arbetspsykologiska tillämpningar bygger på begreppet egenskap (eng. trait). En egenskap är en tendens hos en person att reagera likartad i skilda situationer, t ex att ofta vara sparsam eller noggrann. Om man studerar en grupp personer utifrån en sådan egenskap finner man nästan alltid att de varierar kraftigt. Korrelerar man situationer, eller test, över personer får man en ganska stark korrelation om det verkligen finns en sådan generell egenskap som antagits.

För det sunda förnuftet är nog typbegreppet mera näraliggande än egenskapstänkandet. Typ-psykologiska systemet finns det många av, mest känt är förmodligen det som Jung föreslog [95; 96] . Ganska populära under en begränsad tid var typologier som försökte knyta personlighet till kroppsbyggnad (Kretschmer, Sheldon) - empiriskt visade det sig att de var ohållbara. Psykologiska typologier har emellertid aldrig riktigt lyckats av det skälet att människor inte varierar på det sätt som krävs för typindelning. Det vanligaste är mellanformer, extremerna är ovanliga. Fördelningarna visar inga tendenser till klusterbildning utan det är fråga om jämna övergångar och en gruppindelning måste därför alltid bli mycket godtycklig. Statistiska metoder för klusteranalys finns det en uppsjö av men ingen har lyckats utveckla en metodik som ger något väsentligt värde utöver de enklare former av dimensionsbaserad statistik som är det givna alternativet. Jungs typologi kan lätt översättas i ett dimensionbaserat system. I stället för att tala om typerna extravert<sup>2</sup> och introvert skaffar vi oss mått på graden av inåtvändhet/utåtvändhet.

Egenskap är inte detsamma som attityd, värde och karaktär. Attityd är en mera flexibel tendens att värdera ett begrepp eller ett objekt medan värde syftar på en generellare nivå av värdering. Spranger skilde mellan teoretiska, ekonomiska, sociala, estetiska, religiösa och politiska värdetyper [192] , och idén har visat sig livskraftig [3] , men forskning om generella värden visar i allmänhet att de är *för* generella [173; 186] . När det gäller attityder kan man uppnå höga samband med beteende om de görs tillräckligt specifika [169] . Karaktär är personlighet värderad utifrån moraliska principer snarare än i den deskriptiva mening som är den vanliga i psykologin. I äldre litteratur talade man ofta om "dygd".

---

2 . Troligen bl. a. beroende på Jung har den pseudolatinska varianten extrovert/extroversion blivit vanlig, men numera tycks den vara på väg att ersättas av den korrekta termen extraversion.

Begreppet egenskap kan spåras långt bakåt i psykologins historia. Hippokrates temperamentslära<sup>3</sup> (400-talet f. Kr.) var en typologi men sådana kan ju oftast lätt översättas till egenskaps-system, mera därom senare i denna artikel. Ett hopp på drygt 2000 år för oss till frenologerna, som i början av 1800-talet skapade ett system där många egenskaper enligt dem kunde diagnosticeras med hjälp av att kartlägga "knölar" på skallen [33]. Dessa knölar skulle i sin tur direkt avspegla hjärnans storlek i området i fråga. Eftersom det senare visade sig vara ett felaktigt antagande kom frenologin i vanrykte<sup>4</sup>, men det dröjde minst 100 år innan dess nedgång blev definitiv<sup>5</sup>. Det är emellertid troligt att frenologerna haft stor betydelse genom att de populariserade egenskapstänkandet, och det har mer eller mindre direkt övertagits av eller åtminstone starkt påverkat den moderna psykologin i en tidig fas. Den konkreta tillämpningen på "knölar" var i och för sig humbug, men är intressant som ett exempel på trovärdigheten hos det fysiska, och på en sorts psykologins skuggliv där humbug och viss realism blandas i en salig röra [70].

Bröderna Floyd och Gordon Allport publicerade en intressant artikel redan 1921 [1] och Gordon Allport gav en mera teoretisk analys något senare [2]<sup>6</sup>. I den förstnämnda artikeln redogörs för en undersökning av personlighetstest matchade mot bedömningar av de testade personerna, bedömningar gjorda av deras bekanta. Testen samvarierade i bästa fall 0.3 - 0.4 med bedömningarna vilket författarna tydligen ansåg vara pinsamt lågt. Mera om hur pinsamt det egentligen var följer nedan.

Under flera decennier utvecklades metodiken för att mäta egenskaper med hjälp av frågeformulär, med början 1931 i ett arbete av Bernreuter [15], som avsåg att mäta neuroticism, introversion, dominans (ett begrepp från bröderna Allport), självtillräcklighet, självförtroende och sällskaplighet. Som vi ska se är detta en uppsättning dimensioner som definitivt verkar mycket modern. Många nya mätinstrument konstruerades sedan men metodiken förbättrades bara ytterst långsamt.

Bernreuter konstruerade ett formulär med 125 frågor, som skulle besvaras med Ja, Nej eller ?.

---

<sup>3</sup>. Temperamentsläran brukar tillskrivas Hippokrates. Samma terminologi finns hos andra klassiker som Aristoteles. Den tycks ha byggt på fysiologisk spekulation om "kroppsvätskor" och ytterst på de fyra elementen jord, vatten, eld och luft. Många klassiska texter översatta till engelska kan laddas ner från <http://classics.mit.edu>.

<sup>4</sup>. Den katolska kyrkan och Österrikes regering bannlyste och förbjöd frenologin, som utvecklats av österrikaren Gall, men knappast för att den var empiriskt felaktig, utan för att de inte kunde smälta idén att personligheten kunde lokaliseras till något fysiskt. Frenologerna fick utstå vitt skilda former av kritik, med andra ord. Mesmerismen (hypnosläran) utsattes under ungefär samma epok för en liknande korseld - även i den fanns en kärna av något viktigt och intressant, insvept i rena stolligheter, något som det tog mycket lång tid för etablissemanget att acceptera.

<sup>5</sup>. Försvunnit har den ingalunda. På nätet kan man hitta mycket information om aktuella aktiviteter och förnekande av att kritiken mot frenologin var befogad.

<sup>6</sup>. Dessa och många andra klassiker inom psykologin kan laddas ner från [www.yorku.ca/dept/psych/classics/](http://www.yorku.ca/dept/psych/classics/)

De första frågorna lyder:

1. Does it make you uncomfortable to be "different" or unconventional?
2. Do you day-dream frequently?
3. Do you usually work out things for yourself rather than get someone to show you?

Normdata och uppgifter om reliabilitet ingår, liksom mallar för utvärdering av svaren, och en del valideringsresultat från jämförelse med andra test. Det är inte mycket i Bernreuters instrument som tyder på att det inte skulle vara aktuellt, utom det gulnade papper det är tryckt på (och det betydande detektivarbete som krävdes för att lokalisera och få låna ett exemplar från ett bibliotek i Texas).

Personlighetspsykologin i betydelsen egenskapspsykologi har kämpat i motvind ända sedan 1968 då Walter Mischel publicerade sin bok *Personality and assessment* [126]. Han visade nämligen övertygande att personlighet, mätt med den tidens test, knappast korrelerade högre än 0.3 med kriterier på framgång i yrkeslivet. Senare debatt, mycket omfattande, om Mischels tes har inte kunnat visa att han hade fel, utan tvärtom har forskningen bekräftat den huvudsakliga slutsatsen, se Mabon [102] för ett aktuellt svenskt exempel. Redan 50- och 60-talets personlighetstest kunde uppvisa prediktiva validiteter omkring 0.3 [61] och någon drastisk förbättring har inte skett sedan dess. Som nyss påpekades låg 1921 års värden hos bröderna Allport på samma nivå.

Om kriteriet är specifikt, t ex lön eller förmansbedömning av lämplighet eller prestationsnivå, är personlighet alltså av ganska marginell betydelse. Om kriteriet däremot är en sammanslagning av många aspekter på framgång kan en högre korrelation uppnås, men till priset att man inte riktigt vet vad det är som prediceras. En mycket klar demonstration av denna princip ges i Fishbein och Ajzens demonstration av hur man kan öka prognosstyrkan hos attitydmått [48].

Det är naturligtvis så att en del, kanske en stor del, av en människas beteende bestäms av hennes miljö. Eftersom miljön i väsentliga avseenden tenderar att vara konstant - de flesta lever sina liv inom snäva ramar - blir också beteendet i viss mån predicerbart och ganska konstant. Ganska lite forskning har gjorts om beteendet i naturlig miljö [188]; tidiga försök i detta avseende avstannade och tycktes vara återvändsgränder [11; 25; 168].

"More of the same" är en god princip för att predicera beteende, kanske beklagansvärd, men så är det [128]. Vid urval inom arbetslivet är det en god princip att gå på de prestationer som individer tidigare gjort i liknande jobb, vid arbetsprov eller under provtjänstgöring ("work samples") [147]. Teknikaliteter i sådan mätning diskuteras av Hough [82]. När det gäller ungdomar finns dock sällan den informationen att tillgå, och när det gäller ledande befattningar kan kraven ligga på en sådan nivå att tidigare prestationer och beteende framstår som irrelevanta. Annan information måste sökas, även om biografiska data alltid har sitt intresse.

På 90-talet har emellertid motvinden mot personlighet börjat svänga till medvind [64; 69; 79; 205]. I en stor meta-analys visade Tett et al. på validitetsmått för personlighetstest upp till ca 0.4 för situationer där testen valts på grundval av jobbanalys [205]. 0.4 är ett högst respektabelt värde, i nivå med intelligenstagstest.

Några faktorer som har medverkat till vindkantringen kan nämnas. För det första har det visat sig att intelligenstag, som är av stor betydelse och som jag behandlar i nästa avsnitt, är svåra att förbättra utöver en moderat (men fullt acceptabel) nivå i de flesta sammanhang. För det andra har nyare forskning om personlighet, med början i ett stort projekt inom amerikanska armén [120] och forskning utifrån Big Five (se nedan), t ex vid urval av chefer [76] och även försäljare [75], visat att personlighetsmått ger ett viktigt tillskott utöver vad som kan prediceras utifrån intelligens. Hogan fann prognosvaliditet omkring 0.3 för chefsfunktioner. Hårtill kommer att optimismen i personlighetslägret har ökat betydligt sedan man funnit att alla eller nästan alla av de snåriga egenskapsvariablerna inom området konvergerar mot 4 eller 5 grundläggande dimensioner, de s.k. Big Five [13; 212]. Shafer [163] visade att Big Five, i enlighet med en teoretisk analys av McCrae och Costa [118], hade ganska starka relationer till biodata och självuppfattning. En kritisk analys av några viktiga skalor som bygger på Big Five ges av Widiger och Trull [210]. De pekar bl a på att dessa skalor (inklusive Hogans) visat sig vara känsliga för det tillfälliga stämningläget. Att de dessutom kan vara tämligen genomskinliga och inbjuda till intrycksmanipulation i en urvalstestning är uppenbart, se t ex de resultat som rapporterats för personlighetstestning vid HHS [179] eller en sammanställning av experimentella resultat av Viswesvaran och Ones [213].

Hogan et al. pekar också på behovet av att mäta negativa egenskaper [76] vid chefsurval. Misslyckande som chef kan bero på avsaknad av positiva egenskaper, som mäts av Big Five instrument, men också på förekomst av negativa personlighetsdrag, t ex av typ psykopati. Utvecklingsarbete med att mäta sådana negativa drag pågår [71]. Även Hogan har utvecklat ett sådant instrument, Hogans diagnostiska skala (HDS). Det är troligt att starkt negativa beteenden både är mera sällsynta och mera diagnostiska än positiva beteenden [152].

Ett exempel på negativt personlighetsdrag som troligen är ganska vanligt (vilket psykopati inte är) är perfektionism [155]. Perfektionism leder till låg kvalitet på allt arbete som kräver kreativitet, och till ständiga förseningar, nya deadlines etc. Det är ett av de sex syndrom som förklarar förseningar och uppskjutande (eng. procrastination) enligt en klinisk-psykologisk analys av Sapadin [159]. Ett annat troligen ganska vanligt destruktivt förhållningssätt är uppskjutande av arbetsuppgifter, vilket tjänar funktionen av aggression mot en auktoritetsfigur som kan vara en lärare eller en chef. Man talar i sammanhanget om ”passiv aggression”.

Även äldre idéer om kreativitet och divergent tänkande har överlevt många års misslyckanden och mycket positiva resultat (ca 30 % förklarad varians i ett kriterium som originalitet, mindre men fortfarande intressant i global ledarförmåga) avseende kreativt ledarskap har publicerats [133]. Stabilitet över ganska lång tid hos personlighetsfaktorer kan också nämnas [119]. I det sammanhanget påminner jag om den redan diskuterade principen "more of the same", den lika dystra som giltiga spådom som dömer oss alla att springa livet igenom i ett och samma ekorrhjul.

Ett specialfall av ovanstående diskussion av individ kontra miljö är debatten om personlighet och ledarskap. Den traditionella uppfattningen var länge att det inte fanns något samband [217]. Nyare forskning har emellertid visat att samband finns, låt vara att de inte är särdeles starka utan av den typ som man normalt finner för personlighetsvariabler [78]. Om detta i sin tur betyder att det finns praktiska förutsättningar för att använda sig av personlighetstest eller bedömningar av annan typ, t ex på grundval av intervjuer, beror väsentligen på urvalskvoten. (Urvalskvoten är kvoten mellan antalet slutligt antagna eller anställda och det totala antalet



ursprungliga sökande). Den som arbetar med en bred bas av kandidater (liten urvalskvot) har förutsättningar att lyckas, den som inte gör det riskerar att misslyckas nästan oavsett hur bra test eller intervjuare som han eller hon har. Dessa blir nämligen aldrig så bra att deras precision i individuella fall är extremt hög utan test hamnar i bästa fall på nivån "trubbigt men bra om urvalskvoten är tillräckligt liten". Sedan är det en annan sak vilken typ av ledare man vill ha, se House och Howell [84] för diskussion av olika typer av karismatiskt ledarskap och Conger och Kanungo för ett intressant försök att mäta detsamma [34]. Givetvis är ledarens funktion i slutändan beroende av ett komplext samspel mellan miljö och person [216].

## Den förkättrade intelligensen

Det är en vanlig uppfattning att personlighet är av avgörande betydelse för framgång i yrkeslivet, i synnerhet när det gäller chefsbefattningar. I motsvarande mån tonar man ner intellektuell begåvning, och denna sägs vara garanterad i och med att en person har skaffat en akademisk utbildning, eller anses t o m vara negativ om den är på en "för hög" nivå. Så intelligenta människor är besvärliga, tror många. Och de erinrar sig gärna från skolan att "klassens ljus aldrig blev mer än kamrer". Men vad säger forskningen?

En stor översikt av forskning om urval av personal på grundval av psykologiska test och annan information publicerades för några år sedan [161]. Slutsatsen, på grundval av 100 års forskning, var att intelligens är den variabel som tydligast är relaterad till kriterier på yrkesframgång i de mest skiftande befattningar. Personlighet är också av betydelse och ger ett tillskott till prognosvärdet, men det är mindre viktigt och svarar för ca 25 % av det prognosvärde som intelligens test ger. Dessutom kan intelligens med fördel behandlas mycket generellt. Det finns en g-faktor [190] som återkommer i alla eller nästan alla av de test som vi kallar intelligens test [66; 207]. Det är just denna g-faktor som är bärare av det prognostiska värdet. Detta betyder i sin tur att samma test på begåvning kan användas i många olika urvalssituationer. Specifika begåvningsfaktorer (verbal, spatial, numerisk) kan ibland ha tillkommande värde [67].

Det finns givetvis situationer där personlighet, inte intelligens i traditionell mening, måste ges den avgörande vikten. Jag tänker på urval i flera steg, där man först selegerar med avseende på test eller skolbetyg och sedan gör ett slutligt urval inom gruppen av toppkandidater. Detta är ett mycket effektivt sätt att arbeta, och om man har många sökande i den ursprungliga gruppen av kandidater är man nästan garanterad framgång, eftersom urvalskvoten är liten. Även ganska trubbiga test fungerar bra när urvalskvoten är liten.

I den aktuella debatten har det hävdats att intelligensbegreppet med fördel kan breddas till att omfatta även kreativa och praktiska sidor av det mänskliga fungerandet [196], och detta gäller inte minst chefer inom näringslivet [194]. Det kan också vara fruktbart att knyta an till forskningen om expertfunktioner [44; 45], och relatera till utbildning och individuella skillnader [195]. Ganska populära idéer om intelligens i 7 eller 8 olika avseenden har föreslagits av Gardner [58; 59], men denna teori saknar stringent empiriskt underlag.

Sternberg har föreslagit en teori där intelligens definieras på grundval av framgång i livet i

stort och han specificerar intelligens i tre avseenden: analytisk, kreativ och praktisk [193] . I en serie undersökningar, utförda i många olika länder, har han visat att denna struktur hos intelligensen kan dokumenteras och att test som konstrueras på dess grundval tillför betydande prognosvärde utöver det som kan åstadkommas med traditionella test [196; 198; 199] . Eftersom analytisk (traditionell), kreativ och praktisk intelligens är lågt eller inte alls korrelerade kommer de nya intelligensaspekterna att göra det möjligt att identifiera personer som har möjlighet att bidra med viktiga tillkommande typer av förmåga utöver den traditionella analysförmågan. Företag som förstår detta och utnyttjar det har chansen att vinna en konkurrensfördel. I pedagogiska sammanhang kan Sternbergs teori underlätta identifiering av starka sidor hos elever från vitt skilda kulturer [197] .

En äldre tradition av kritik av intelligensbegreppet representeras av McClelland [116; 170] . McClelland hävdade att intelligens var i stort sett utan relation till kriterier som framgång i yrkeslivet och att man i stället borde gå på konkreta och identifierade former av relevant kompetens. Kritiken sköt över målet i och med att det knappast var eller är kontroversiellt att det finns ett samband mellan intelligens och kriterier av de mest skilda slag, men ingen hävdar att sambandet är perfekt så alla förslag om hur det kan förstärkas är välkomna. Nyligen publicerade McClelland en kortfattad översikt av 25 års försök att realisera programmet [117] . Den idé som har följts har varit att utifrån intervjuer av typ "critical incident" identifiera de avgörande typerna av kompetent beteende hos mycket framgångsrika chefer [21] . Ganska litet finns publicerat om denna ansats av McClelland men den verkar intressant.

Motowidlo [130] har utvecklat den intressanta idén att personlighetsvariabler kan möjliggöra prognos av sociala aspekter på yrkesfunktion, medan erfarenhet och intellektuella faktorer samt arbetsmotivation tycks vara viktiga för prestationen i snäv mening. Validiteten hos personlighetstest tycks vara störst när man studerar sociala kriterier, inte prestation i snäv mening [19] . Så fann t ex McManus och Kelly [121] att fyra av de fem stora personlighetsdimensionerna gav tillskott i validitet för sociala kriterier, och bara en (extraversion) för arbetsprestation i snäv mening. Tillskotten mättes utifrån biodatas prognosförmåga som grund och var ganska betydande.

I en omfattande meta-analys har stabil validitet omkring 0.2 påvisats för biografiska data [17] . Lovande resultat på grundval av biodata vid urval till höga chefspositioner ges i [153] . Biodata ger ett betydande tillskott till prognosförmåga, utöver det man får utifrån intelligens- och personlighetstest [131] ; det kan handla om ca 7 % tillskott.

## **Felfaktorer i mätning och bedömning av personlighet**

Test och intervjuer används för att mäta och bedöma personlighet. Är det svårt eller lätt att göra sådan mätning? Kan man klara det på 10 minuter och med hjälp av en enkel adjektivlista, går det att göra enligt kvällstidningarnas söndagsbilagor, eller krävs det något mer och i så fall vad? Varför inte bara fråga kandidater: Vilken personlighet har du? Beskriv den!

För det första kan man inte ställa en sådan öppen fråga med förhoppning om framgång. Människor skulle närma sig den utifrån olika referensramar och många skulle ha svårt att alls artikulera ett svar. Viktiga dimensioner skulle glömmas bort av somliga, oviktiga skulle få betydande utrymme av andra. Någon form av standardisering behövs tydligen.

Det kan ske genom intervjuer. I Sverige är det populärt att använda "djupintervjuer". Det är oklart vad detta begrepp egentligen står för, men att det är något fint råder allmän enighet om. Ändå är det uppenbart att ostrukturerade samtal av denna typ kan leda i än den ena, än den andra riktningen. Subjektivismen i ostrukturerade intervjuer är enorm, och det har också visat sig att bedömningar på en sådan grund saknar tillförlitlighet [7; 72]. De ger möjligen information om personlighet, men på ett alltför abstrakt och generellt plan [35].

"Personkemi" mellan intervjuaren och kandidaten är en viktig faktor som kan störa processen betydligt [189]. Det första intrycket bör man kanske skaffa sig via telefonsamtal, för att inte störas av utseende och andra ovidkommande faktorer [40]. Just det första intrycket påverkar nämligen speciellt starkt den slutsats man så småningom kommer till. Å andra sidan kan den implicita informationen som förmedlas i intervjun vara ett värdefullt tillskott av information, för den som kan utnyttja den på rätt sätt [27].

Ett första intryck kan vara väl så bra som grundval för en prognos. I en stor meta-analys fann nämligen Ambady och Rosenthal [5] att "thin slices of behavior", d v s ca 5 minuters observation av en främmande person, kan läggas till grund för prognoser med en validitet av ca 0.4! För länge sedan hävdade Gordon Allport just betydelsen av uttrycksfullt beteende (expressive behavior) som informationskälla om en individ [4], kanske hade han faktiskt rätt. Möjligen kan det vara en idé att systematiskt utnyttja denna information vid bedömning på grundval av intervjuer.

Nyare forskning har visat att strukturerade intervjuer kan ge viktiga tillskott i en bedömnings-situation [92; 215]. I en aktuell meta-analys handlar det om ca 17 % -enheters ökning i förklarad varians - en mycket imponerande siffra [37]. Struktur kan uppnås på olika sätt och måhända spelar det mindre roll exakt hur man går till väga [162]. Samma områden måste tas upp med alla sökande, i en viss ordning och med hjälp av frågor formulerade på ett visst sätt. Intervjuarna måste använda vissa i förväg intrimmade bedömningsskalor. Då kan alltså intervjuer fungera bra [85]. Det är emellertid troligt att intervjuer i praktiken ofta görs på ett mindre effektivt sätt, d v s tämligen ostrukturerat och "psykologiserande".

Sammanvägningen av informationen från en intervju bör lämpligen innehålla ett stort mått av "mekanisk" integration av bedömningar av delaspekter [55]. Litteraturen om sk kliniska bedömningar visar ju entydigt att dessa är underlägsna statistisk sammanvägning av data [65].

Personlighetstest är vanligen av typen självbedömning. Den testade får t ex ta ställning till vissa påståenden eller beskrivande adjektiv. Frågan är givetvis hur sanningsenlig han eller hon är, och om man kan lita på svaren [213]. Forskningen visar att det blir förvånansvärt små effekter på validiteten av social önskvärdhet i denna typ av test, även i starkt konkurrensutsatta urvalstestningar [39; 83] och social önskvärdhet verkar inte ha starkt snedvridande effekter [14; 18; 139] i betydelsen att ändra rangordningen i en grupp testade, men givetvis påverkas nivån på svaren [213]. I enskilda fall kan ändå mycket stora effekter uppkomma.

Korrektion för opportunistiska svar tycks inte påverka validiteten hos testen, men kan påverka vilka som föreslås till det sökta jobbet eller den sökta utbildningen, särskilt om urvalskvo-

ten är liten [151]<sup>7</sup>. Andra typer av personlighetstest, som projektiva test, används normalt inte för urval i arbetslivet, men se nedan för svenska tillämpningar.

### **Skillnader mellan länder i fråga om testanvändning**

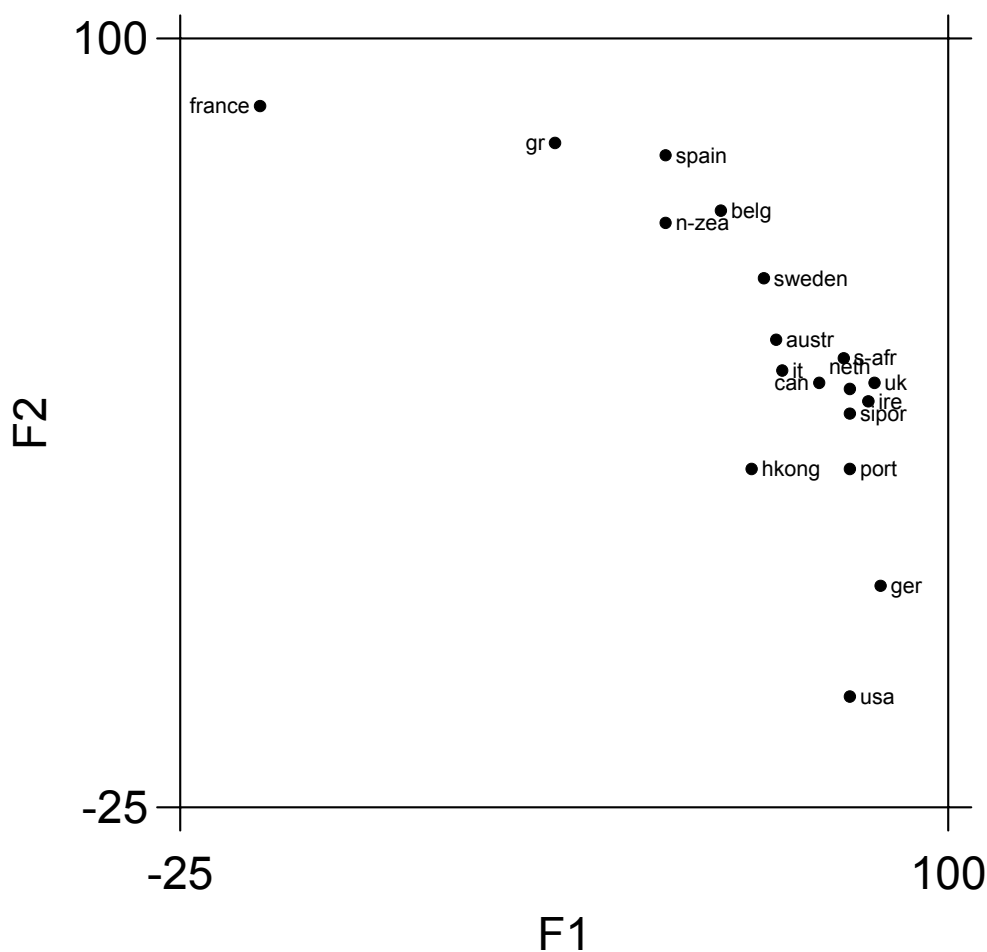
I en artikel [154] analyserades skillnader mellan länder avseende förfarande vid urval, bl a användning av olika sorters test och andra procedurer. Företag i 18 länder svarade på frågor kring vilken praxis de använde sig av. Sverige låg på 8:e plats när det gällde användning av kognitiva test, men på 2:a när det gällde projektiva test och 5:e avseende grafologi (som dock användes sällan i alla länder utom Frankrike). Personlighetsformulär var också populära i Sverige, här låg vi på 3:e plats (USA låg sist). Se Littorin [99; 100] för översikter av svenska förhållanden, som stämmer bra med dessa intryck av den internationella jämförelsen.

Ryan et al. analyserade sina data utifrån kulturbeskrivande dimensioner (undvikande av osäkerhet och maktavstånd) och fann att dessa hade ytterst litet värde för att förklara de stora skillnaderna mellan länderna. Det är emellertid möjligt att göra lite mera av deras data (som återfinns i artikeln) än de själva har rapporterat.

Jag har korrelerat länderna över medelvärden i testanvändning och beräknat värden för de två största komponenterna i en analys av dessa korrelationer. Dessa komponenter förklarade 87 % av variationen mellan länder, och det fanns ingen tydlig tredje komponent. Värdena har prickats i Figur 1.

---

<sup>7</sup>. Ett aktuellt exempel från personlighetstestning vid HHS: korrektion av den sammanlagda testpoängen för värdet i en dimension som mäter tendens att svara opportunistiskt gav små ändringar, mätt i termer av korrelation. Den korrigerade slutpoängen korrelerade nämligen 0.9 med den okorrigerade. Men vid ett urval av 60 personer av total 190 testade (som skulle inbjudas till intervju) visade det sig att 20 personer "bytte plats", d v s 1/3 av dem som inbjöds kom med efter korrektionen. Efter-som validiteten som sådan troligen inte påverkas negativt av korrektion talar enligt min mening mycket för att den är berättigad, speciellt som mätningen av opportunistik i svaren fick stöd av andra analyser.



Figur 1. Positionen av ett antal länder m avs på testanvändning i ett 2-faktorsystem.

Som framgår av figuren fanns en jämn övergång mellan olika typer av filosofi när det gäller test vid rekrytering och urval. USA och Frankrike är de två extrempunkterna. Det senare landet excellerar i grafologi, personlighet, projektiv testning o dyl., medan man i USA sätter "objektiva" metoder i främsta rummet, metoder av typen kognitiva test. Den franska ansatsen kan sägas vara en romansk kulturell ansats, med Spanien, Belgien och Grekland inkluderade i samma grupp (dock ej Italien), men även Nya Zeeland och Sverige tenderar att hamna i den gruppen. Övriga länder har en mera kompromissartad inställning, med undantag för Tyskland som ligger nära USA.

Vad beror dessa skillnader på? Genomslaget för kritik av projektiva test och grafologi är tydligen lågt bland företag och psykologer i Frankrike - och i Sverige finns ett liknande syndrom, ehuru ej lika extremt. Svensk psykologi är ju starkt påverkad av utvecklingen i USA, men tydligen finns det något annorlunda inslag i just den tillämpade psykologin. Vad detta inslag är, är uppenbart: sedan början av 70-talet har den s.k. dynamiska psykologin dominerat psykologutbildning och de flesta psykologers praktiska verksamhet i Sverige. Forskningsresultat om validitet ignoreras helt enkelt. Därför kan man använda grafologi och projektiva test

vid urval - de stämmer bra med de naivpsykologiska uppfattningar man har om personligheten. Det har påvisats att tilltron till vissa projektiva test troligen åtminstone delvis bygger på sådana stereotyper [31; 32]. Projektiva test är trovärdiga därför att de stämmer med allmänkulturell trovärdighet hos dynamisk psykologi och psykiatri, och troligen också med okunighet hos dem som är psykologernas uppdragsgivare.

## Hur mäter man personlighet i Sverige?

Nu kommer jag till den i inledningen utlovade diskussionen av svenska tillämpningar. Nästan inga av de test som används i Sverige för att mäta personlighet är validerade i Sverige, i flera fall inte ens standardiserade, och de bakomliggande teorierna är obsoleta och synnerligen oklara. Några av de viktigaste testen diskuteras i det följande.

### **MBTI**

Myers-Briggs Type Indicator (MBTI) har standardiserats och är en best-seller, men inga svenska valideringsstudier finns dokumenterade i manualen [135]. Internationellt är testet också en storsäljare, med ca 2 miljoner testningar per år över hela världen. Myers-Briggs bygger på Jungs spekulationer som varit av betydelse i litteratur och annan kulturell verksamhet och som är spännande att läsa, men tyvärr är Jungs position inom modern psykologi nästan en chimär [136]. Jungs egen empiri var undermålig och delvis bedräglig; hans rasfördomar mycket pinsamma [191]. MBTI skall enligt den amerikanska manualen inte användas som prognosinstrument vid urval eller för att sammansätta arbetsgrupper utan för att skapa självinsikt, men om dimensionerna inte har etablerad validitet undrar man vad självinsikten kan vara värd.

Jung [95] utgick från två grundläggande dimensioner som vi kan kalla intressets inriktning (extravert - introvert), korsat med 4 "funktioner", nämligen

- (1) Sinnesförmåelse - intuition: söka information via sinnesdata eller via intuitivt tänkande.
- (2) Tanke - känsla: logisk analys gentemot beslut på värderingsgrund<sup>8</sup>.

Enligt Jung kan människor delas in i 8 typer efter vilken av de 2 x 4 kombinationerna som dominerar. Dimensionerna i detta system är oklara, med undantag för extraversion - introversion. Logisk analys är t ex ingen motsättning till värdering. Om beslut ska fattas måste båda till. Det finns inga värdefria beslut. Perceptuell information säger inget om yttervärlden utan att först tolkas och sättas in i ett begreppssystem [24]. Och så vidare. Till detta kommer att antagandet att det skulle finnas generella tendenser till att föredra t ex sinnesdata framför intuition är helt osannolikt och obevisat.

---

<sup>8</sup>. Tydligt används här ordet känsla för att syfta på värdering - en vanlig språklig slapphet som väsentligt ökar förvirringen. Känsla är ett ord som bör syfta på emotioner, och dessa är i sin tur något helt annat än värderingar.

Myers och Briggs gjorde saken värre genom att föreslå ytterligare en "funktion", nämligen

(3) Bedömning - perception: preferens för sensorisk eller perceptuell information å ena sidan, tanke eller känsla (värdering) å den andra, vid hantering av den yttre världen.

Förslaget är illa anknutet till Jungs teori [10]. Myers och Briggs menade emellertid att man kan härleda 16 typer utifrån

Extraversion - introversion  
Sinnesförmåelse - intuition  
Tanke - känsla  
Bedömning - perception

MBTI mäter dessa 4 dimensioner reliabelt<sup>9</sup> [29], vilket är ett nödvändigt men inte tillräckligt villkor för att testet ska vara användbart. Däremot finns inte tydligt stöd för att faktorstrukturen i testet sammanfaller med de antagna dimensionerna [158]. MBTI resulterar i en indelning av individerna i 16 typer, men internationell erfarenhet av typologin är negativ [8; 16; 22].

Wiggins har gjort en analys av MBTI som återfinns i MMY (Mental Measurements Yearbook) vol. 10, 1989 [211]. Han ställer upp 4 krav på testet för att det ska kunna anses som validerat och samtidigt understödja Jungs bakomliggande teori (i Myers och Briggs hemkokade tappning). Dessa krav är:

bimodala fördelningar  
diskontinuiteter i regressioner  
typindelning som är stabil över tid  
predicerade skillnader mellan kriteriegrupper, t ex olika yrkesgrupper

De tre förstnämnda kraven handlar alla om hur data ska se ut om det faktiskt finns en tendens att människor bildar klart åtskilda grupper eller typer. Enligt Wiggins är endast den sistnämnda kravet uppfyllt i de studier som gjorts. Wiggins analys är milt formulerad men i sak förödande för testet. Om det inte ger resultat som understödjer teorin, annat än i mycket begränsad omfattning - kan då slutsatsen bli annat än negativ? Det är inte heller brist på forskning som är förklaringen till situationen, det finns mycket omfattande forskning om detta test och manualen är på flera hundra sidor.

Men låt oss se på vad testet ska användas till. Indelningen i de 16 typerna ska användas för att öka den testades självkännet. Det tycks emellertid inte finnas någon forskning som visar att testet ifråga verkligen skapar självinsikt utöver den *känsla* av sådan som ges av det mesta oavsett verkligt värde, inklusive horoskop. Det är välbekant att en känsla av självinsikt uppkommer ganska lätt, något som länge utnyttjats av astrologer. Ett roande exempel beskrivs av Sändh [201]. Ett horoskop ställdes på honom och stämde till ca 60%. Samma horoskop stämde lika bra på andra personer. Det finns allmängiltiga och/eller smickrande påståenden som alla har en tendens att anse som riktiga beskrivningar av just dem. Strategin bakom att övertyga främmande människor att man vet allt eller det mesta av betydelse om dem är just

---

9. Testet är med andra ord tämligen fritt från rena slumpfel.

den.

Självinnsikt och självkunskap är i själva verket subtila och svårfångade begrepp [150]. Det finns en hel del som talar för att fullständig kunskap om hur andra människor uppfattar en inte är bra för den mentala hälsan, ett visst mått av önsketänkande är önskvärt [202]. Positiva upplysningar är givetvis lättare att acceptera än negativa. Hur ett testresultat ska kunna bidra till acceptering av nya tankar om den egna personen, och därigenom vara till stöd och hjälp för den testade, är en stor och svår fråga. Å andra sidan kanske testresultat kan hjälpa till att vidmakthålla och skapa myter om den egna personen, myter som kan vara destruktiva eller konstruktiva. Minne och egenbedömning av prestation är ofta kraftigt felbemängda i optimistisk riktning [124].

En typisk reaktion på ett MBTI-utlåtande är förmodligen att "Det ligger en hel del i det där". Bortsett från de positiva värderingarna, som givetvis är smickrande, tycks det stämma i de flesta avseendena. Men vid närmare eftertanke inser man att det helt enkelt är en sammanfattning av de val som gjorts i testet. Testet beskriver dig på samma sätt som du själv beskriver dig! Det är därför osannolikt att testet skulle ge ny information eller en helt annorlunda självbild. Testet befäster fördomar, och beskriver dem på ett smickrande sätt. Kanske ligger där nyckeln till dess enorma framgång.

Beteendet ändras emellertid från en situation till en annan [127]. Det är en överförenklad syn på den egna personen om man ser sig som ständigt utåtvänd, aggressiv eller noggrann. Det är bättre att lära sig inse att man varierar [123]. Test av typ MBTI tar inte hänsyn till det behovet, utan erbjuder – i bästa fall – en kraftigt förenklad bild av den testade.

MBTI är ett märkligt test på flera sätt. Det består av 84 uppgifter. Av dessa är 46 frågor där man ska välja mellan två svarsalternativ för att meddela hur man "vanligtvis" brukar känna eller handla. De övriga 38 uppgifterna består av ordpar. Uppgiften är att markera "vilket ord som tilltalar dig mest" m av på betydelse. Vad betyder nu detta? I uppgift 64, t ex, ska valet ske mellan "instämma" och "diskutera". Det förstnämnda kan ju vara trevligt om andra gör det - hur instruktionen tolkas torde vara högst varierande. Vilket är mest tilltalande av orden tänka och känna? Är det ett meningsfullt val? Och så vidare.

MBTI är starkt laddat i Big Five [54]. En analys av data insamlade vid personlighetstestning av sökande till HHS visar att Big Five förklarade i genomsnitt ca 40% av variansen i de fyra grundläggande MBTI indikatorerna, alltså större delen av den varians som ej är rent mätfel [179]. MBTI omfattar dessutom inte emotionell stabilitet som dimension och ger ipsativa<sup>10</sup> mått varför dess värde för många psykologiska tillämpningar ej är dokumenterat.

Den omfattande forskningen om MBTI har visat på många samband. Detta är inte förvånande eftersom det huvudsakligen mäter samma grunddimensioner som andra självrapporttest.

---

<sup>10</sup>. De bedömningar som görs av den testade är av jämförande typ, dvs AA stämmer bättre på mig än B@, men inte av absolut typ, dvs AA stämmer mycket bra på mig@.



## Thomassystemet

Ett annat system som är framgångsrikt i Sverige, är det s k Thomassystemet<sup>11</sup>, som är delvis inspirerat av Hippokrates och även av Jung, och som går tillbaka på en bok från 1928 av amerikanen Marston [106]. Marston var en energisk skriftställare och populärpsykolog [26], nu nästan helt bortglömd i akademisk psykologi men känd på 20-talet. Ett centralt tema både i hans vetenskapliga produktion och i hans många populära skrifter och serien Wonder Woman är dominans/underkastelse<sup>12</sup> - serien tycks ha nått en stor popularitet bland läsare med sado-masochistiska böjelser. Marston var också lite av "crackpot" och befordrade gamla ingrodda fördomar om temperamentsskillnader mellan blondiner, brunetter och rödhåriga - vid ett tillfälle demonstrerade han dem praktiskt på en teaterscen i New York och med hjälp av sin "lögn-detektor"<sup>13</sup>, till publikens bestörtning.

I Thomassystemet sägs personligheten vara möjlig att beskriva i 4 dimensioner, nämligen (med de engelska termerna eftersom översättning bara leder till ännu större förvirring):

Dominance  
Inductance (eller Influence)  
Steadiness  
Compliance

Thomas-systemet är ett alltså ett DISC-system (efter initialerna i de fyra dimensionernas engelska beteckningar) och av dessa finns ett 50-tal som lett till olika test [103]. Thomassystemets enkla ansats är att mäta dessa faktorer med hjälp av bedömningar av adjektiv, men det tycks vara tveksamt om det lyckas göra detta på ett psykometriskt acceptabelt sätt [141].

"Motsvarigheterna" till de fyra dimensionerna i Thomassystemet i Hippokrates temperamentslära är:

Kolerisk  
Sangvinisk  
Flegmatisk  
Melankolisk

---

<sup>11</sup>. Namnet Thomassystemet härrör från en annan psykolog, Hendrickson, som konstruerade den enkla adjektivlista som används i "systemet". Hendrickson tycks inte ha publicerat något om detta utvecklingsarbete men en företagsintern manual föreligger [73].

<sup>12</sup>. Dominans (eng. ascendancy) var en vanlig dimension i 20-talpsykologin och finns med bland bröderna Allports personlighetsdimensioner i 1921 års tappning. Eftersom maktsträvan är ytterst politiskt inkorrekt i Sverige och nästan ingen vill medge att han eller hon vill ha makt och vill dominera är det onekligen tänkvärt att ett system som bygger på dominans som grundläggande dimension kunnat bli så populärt.

<sup>13</sup>. Marston var en av upphovsmännen till lögn-detektorn. I hans variant var den ineffektiv, men långt senare har forskning lett till lovande varianter av idén [104].

Hippokrates syftar tydligen på tendenser att reagera emotionellt, men DISC-systemen är baserade på vissa typer av social interaktion (dominans etc). Nu kan man förstås påstå att emotioner har mycket att göra med sociala relationer och typer av samspel med andra, men det är inte detsamma som att säga att emotioner och social interaktion är samma sak. De är inte samma sak. Förvirring blir resultatet om man påstår motsatsen.

Thomas-schemat förefaller egendomligt vid en första anblick men systemet är egentligen enkelt. Det handlar om att påverka eller påverkas, i en fientligt eller vänligt inställd miljö. Man sluter sig till en persons läge i denna fyrfältare (påverka eller påverkas är den ena dimensionen, fientlig eller vänlig miljö den andra) utifrån hans eller hennes val av självbeskrivande adjektiv. Det tar bara ca 10 minuter att besvara det enkla formuläret och datorn matar sedan fram en personbeskrivning.

Det hela kan ha varit en respektabel ansats år 1928, men en del forskning har hunnit utföras sedan dess. Personligheten handlar faktiskt inte bara om att påverka eller påverkas. Dominans/underkastelse är måhända en viktig personlighetsdimension men inte den enda. Faktorstukturen i formuläret har för ö ej verifierats till fullo [74]. Irvine [90] har sammanställt en del teknisk information men ger inga konkreta upplysningar om validering, se Cooks [36] kritik. Ett av de större problemen med systemet är dess ipsativa karaktär [93]. Om det är en nackdel eller fördel är tveksamt och frågan om ipsativa mätningar är onekligen svår att utreda [110]. Givetvis leder ipsativ mätning till att man tappat information om absolutnivån [103], och även till en del statistiska problem vid dataanalys.

Irvine [90] ger en del sparsmakade upplysningar om prediktiv validitet hos Thomassystemet: opublicerade amerikanska data ger validiteter upp till ca 0.3, men ibland nära noll. Det antyds att data föreligger från tre grupper. Brittiska data ser ut på ett liknande sätt i en grupp som Irvine själv analyserat (ej närmare beskriven grupp bestående av 132 testade personer).

En senare publikation [91] ger prognostisk validering i tre grupper, för enskilda skalor ligger validiteten på den för personlighetsvariabler typiska 0.3. Kombination av skalorna kan eventuellt ge en ökning till 0.4, men det är oklart om detta är ett korsvaliderat värde (troligen inte). Det är värt att notera att ännu i denna artikel från 1994 fanns tydligen inte flera valideringsresultat att hänvisa till, två opublicerade studentuppsatser anförs i en tydlig brist på vetenskapliga publikationer att hänvisa till.

Ska Thomas-systemet, i likhet med MBTI, egentligen ses som ett sätt att öka den testades självinsikt? Cook [36] nämner inte ens detta syfte, inte heller testets försvarare Irvine [90; 91]. Det är märkligt att personlighetstest utvecklats för och i bästa fall validerats för ett syfte (prognos, urval) men sedan används för ett annat, utan forskningsstöd.

Och det är i sanning märkligt att Thomassystemet, föråldrat och överförenklat som det är, med framgång kan säljas till svenska företag som ett sätt att mäta personligheten eller som en grund för organisationsutveckling. Det säger mera om skickligheten hos försäljarna än om något annat. Validiteter som publicerats av Irvine (och som citerats av honom) ligger i underkanten av vad som rutinmässigt presteras av 100-tals andra, mera moderna instrument - som dessutom ger en betydligt mera nyanserad beskrivning av personligheten.

Sambandet mellan de 4 systemvariablerna och kriterier tycks dessutom vara helt okänt under

svenska förhållanden. Det troliga är att systemet har en viss validitet på den för personlighetstest normala nivån, vilket lämnar mycket utrymme för förbättringar.

## **OPQ**

OPQ är ett test som framtagits och marknadsförs av en brittisk konsultfirma vid namn SHL. Enligt deras brittiska hemsida är OPQ det vanligaste personlighetstestet i industriell tillämpning i UK och det används i flera än 20 länder. Det mäter enligt reklamen, bland mycket annat, också emotionell intelligens. Det konstruerades 1984 och utvärderas i 31 dimensioner.

Forskningsdatabasen PsycLit listar omkring 25 publicerade artiklar om testet. Några av de viktigaste slutsatserna ska nämnas. Matthews och Stanton [111] fann att det inte gick att utvärdera så mycket som 31 skalor på grundval av den begränsade uppsättningen item, omkring 200. De vanliga 5 faktorerna i personlighetsmätning räckte bra till, möjligen med tillägg av ännu en. Det finns knappast stöd för att OPQ mäter något utöver dessa. Robertson och Kinder [148] fann att kriterievaliditeten låg omkring 0,2 i genomsnitt, vilket är lågt för ett personlighetstest. Vidare fann Martin, Bowen och Hunt [107] att testet i en normativ variant påverkades starkt av skönmålning av svaren.

I Sverige används detta test i tämligen stor utsträckning. Den svenska versionen förefaller inte ha varit föremål för någon validering eller annan forskning.

## **16PF**

Detta är ett test som ursprungligen konstruerades 1949 av Cattell. Det är ett konventionellt självrapporttest som kan genomgå ganska snabbt. Kortversionen (se nedan) klaras normalt av på ca 20 minuter. Det har publicerats ganska mycket forskning om testet, främst om kliniska tillämpningar. Namnet står för att det mäter 16 personlighetsfaktorer. Det är en komplicerad bild som tonar fram, och senare har ju forskningen funnit att fem grundläggande faktorer är tillräckliga. De 16 faktorerna i Cattells test kan återföras på de fem grundläggande [28]. För övrigt lider testet av samma svagheter som andra självrapportskalor. Det påverkas starkt av om de testade är motiverade att ge en skönmålning av sig själva [53], något som är vanligt i urvalssammanhang när de testade är starkt motiverade att lyckas väl.

Buros Mental Measurements Yearbook, numera Buros Institute, publicerar auktoritativa utvärderingar av test. 16PF värderades för några år sedan. Nyligen kom en bedömning av kortversionen 16PF Select.

5:e upplagan av testet kommenterades så här:

*”The 16PF Fifth Edition also has a limited range of predictive value and must be used cautiously in situations involving occupational selection and appraisal of specific qualities.*

*The empirical validation efforts concerning these interpretive reports are neither mentioned nor described in the administrator's manual; thus, users must consider the possibility that these reports may be based on content interpretation and face validity rather than substantial empirical validation efforts.*

*Overall, the identified improvements appear to have strengthened the utility of this instrument; however, the lack of sufficient information in the form of an available technical manual currently is a significant limitation of the 16PF Fifth Edition. Complete evaluation of this measure and its improved psychometric properties will be necessary following availability of this critical documentation.”*

Utvärderingen av den förkortade versionen som speciellt avsetts för urval gjordes av två bedömare som skrev var sitt utlåtande. Ett av dem tar inte upp validitetsfrågan (märkligt nog) varför jag bortser från det utlåtandet. Den andre, Paul Arbisi<sup>14</sup>, skriver:

*”The most troubling issue with the 16PF Select is the lack of construct and predictive validity evidence. Virtually no data are presented in the manual that demonstrate the relationship between 16PF Select dimensions and relevant job performance parameters. “*

Det finns alltså anledning till betydande skepsis mot detta test, som funnits sedan 1949 och vars producenter fortfarande inte kan redovisa övertygande validitetsdata. Forskningen tyder också på att det inte ger något tillskott utöver de grundläggande fem personlighetsdimensionerna som kan mätas bättre med andra test. Det har för mycket länge sedan passerats av utvecklingen inom testforskningen.

Om test av detta slag ska användas i starkt motiverande urvalstestningar är det ett minimikrav att man korrigerar för skönmålning. Sådana korrekationer kan göras, ett exempel inom området emotionell intelligens har publicerats [185] .

## **U-mantestet, grafologi, astrologi**

Samma kritik som gäller Thomassystemet verkar gälla en annan bästsäljare i Sverige, Scientologikyrkans s.k. U-mantest, som består av 200 självrapportfrågor av traditionellt snitt. Enligt hemsidan <http://www.u-man.se/> har man 93 % träffsäkerhet vid rekrytering med hjälp av detta test. Det anges ej hur man kommit fram till denna siffra. Över 3000 svenska företag påstås ha använt sig av testet. Forskning pågår enligt uppgift i Norge, men inga resultat har dokumenterats.

En ren stollighet, som dock inte saknar anhängare, är numerologin. Följande beskrivning kan läsas på hemsidan [www.ditthoroskop.nu](http://www.ditthoroskop.nu):

*”Det finns flera olika sätt att tolka bokstäver och siffror, men det är namnet och*

---

14 . Docent (assistant professor) i psykologi vid University of Minnesota.

*födelsedatumet som har störst betydelse för en människas livsöde. Siffrorna 1-9 har alla olika kosmiska vibrationshastigheter. Hastigheten på vibrationen förutspår en människas egenskaper och framtid, och också platser och händelser”.*

Grafologi och astrologi skall inte heller förglömmas i sammanhanget. Grafologi är ett snäpp mera trovärdigt än astrologi, men omfattande forskning ger inget stöd åt dess användning inom näringslivet [42] , och aktuell forskning är inte speciellt uppmuntrande [68] . Inte desto mindre kan vi läsa på Internet<sup>15</sup> att

*”Grafologin får fram människans attityd genom att kartlägga den sanna personligheten, människans kapacitet, värderingarna, temperamentet och inte minst samspelet mellan dessa. Den visar också om man har en teknisk begåvning eller en social begåvning (ledarförmåga). Tillämpningsområden är många”.*

Vad gäller astrologi har forskning visat att det helt saknas stöd för dess påståenden [165] , vilket inte hindrar företrädare för denna industri att fortsätta att försvara den, på sätt som f ö är mycket lika dem som används av psykologer som försvarar olika varianter av projektiva test som Rorschach [98] . Det förekommer att man använder sig av astrologi vid anställning av chefer, åtminstone i Storbritannien [149] .

## **Rorschach**

Det finns också andra exempel på test vars validitet internationellt inte ens av förespråkarna anses lämpliga för arbetspsykologiska tillämpningar, men som i Sverige tillämpas utan vidare prut. Jag tänker främst på Rorschach, se en auktoritativ förespråkares nyktra syn på när testet är lämpligt [200] . Detta test från 1923, några år äldre än Marstons spekulationer, är en favorit bland kliniska psykologer världen över och det är möjligt att det bidrar positivt i en del kliniska sammanhang. Men när det gäller att göra specifika prognoser av vikt i näringslivet har det inte kunnat påvisas att testet är användbart. Detta hindrar inte att det tycks ha trovärdighet nog i vissa sammanhang, just för sådana tillämpningar, i vårt land. Samma konsult som använder Rorschach är f ö också känd för att använda grafologi.

Debatten om Rorschach har lett till mycket negativa konklusioner om detta test [183] . I en omfattande genomgång av testets historia och av den aktuella forskningen påvisas ganska obarmhärtigt att nästan ingen validitet har kunnat dokumenteras, med undantag av några rent triviala och praktiskt ointressanta samband [214] .

## **DMT**

Ännu ett exempel på ett projektivt test som används för urval av chefer är det svenska DMT, Defence Mechanism Test. Testet är utvecklat för urval av stridspiloter och inte validerat för urval av t ex bankdirektörer. Detta hindrar inte att det används, med hänvisning till forskning på helt andra områden (Flygvapnet, kliniska tillämpningar). Denna forskning är dessutom av

---

15 . Se <http://www.iqgrafologcenter.se/>

låg kvalitet och högst suspekt [187] . Resultaten har inte kunnat replikeras [108] .

## **Holismen**

”Personligheten är ett integrerat helt” är ett vanligt påstående, vars sanningshalt beror av vad man menar med personlighet och vad man menar med integrerat helt. Men låt oss inte hänga upp oss på terminologin. Det finns något intressant bakom påståendet, nämligen *holismen*.

På 30-talet formulerade Kaila [97] förföriskt elegant holismens credo:

*”...trots att personligheten består av flera skikt och företer inre söndring är den en holistisk enhet; dess skikt och de strimmor, som genomkorsa dem, äro icke oberoende av varandra. Personlighetens enhetlighet uppenbaras i varje steg som människan tager, i varje åtbörd, i varje min, i varje ord, i var gärning - även då de äro sinsemellan motstridiga. Var människa följer samma lag för sitt personliga liv, hon tecknar samma kurva för sin karaktär, då hon går eller springer, då hon talar, skriver eller gör åtbörder och då hon älskar eller tänker. ” (Sid. 31).*

Märk att *allt tycks påverka allt*, och därför är även de ytligaste tecken i djupaste mening avslöjande. Kaila citerar Schopenhauer: "Der Mensch ist so, wie er assieht", t o m utseendet är avslöjande!

Ett typiskt exempel på holistiskt resonemang finner vi tidigt hos Husén [88] . Han säger att betydelsen hos ett visst personlighetsdrag hos en person måste tolkas i sammanhanget med alla hans eller hennes övriga egenskaper. De bildar hierarkier, säger Husén, och vilka slutsatser vi ska dra beror alltså så hela den bild vi får. Delarna påverkas av helheten, med andra ord. Detta låter bra, men är omöjligt att systematiskt belägga.

Holismen är fortfarande vid liv [51; 142] , detta trots att omfattande forskning har visat att typbegreppet är ofruktbart (se diskussionen ovan om MBTI). Men det finns även andra skäl till varför holismen tycks vara en återvändsgränd. De anspråk som många psykologer har på att kunna göra integrerade bedömningar som går utöver det man kan sluta sig till från test och annat kvantifierbart underlag, har gång på gång visat sig vara ogrundade [65; 122] . Ju mera information som psykologen får, desto säkrare blir han eller hon på sin sak, men när informationen ökar utöver en ganska låg nivå blir bedömningarna i själva verket allt godtyckligare och osäkrare [105] .

Holismen har ett nära samband med den humanistiska psykologin (Maslow, Rogers m fl). I den formen har den fått ett stort genomslag i arbetslivet i vårt land. Maslows motivationsteori, många gånger funnen ohållbar i empirisk forskning [175] lärs ut på högskolor som om den vore sann. Den på 70-talet populära sensiträningen har övervintrat och finns nu utbredd inom olika former av konsultverksamhet. Det finns inget vetenskapligt stöd för denna verksamhet i positiv mening, men väl en hel del allvarliga nackdelar [177; 182] . Risken för betydande psykisk skada är uppenbar, liksom möjligheterna för ”terapeuterna” att profitera ekonomiskt och emotionellt [125] .

## **Slutord om personlighetsmätning i Sverige**

Man kan konstatera att gamla projektiva test som föreföll stendöda redan på 50-talet ingalunda saknar nutida förespråkare och att negativa forskningsresultat absolut inte är ett hinder för att ett test används i arbetslivet och genererar goda inkomster för konsulterna. Och kanske är det just detta som är det viktiga för dem, inte att det som de gör har ett dokumenterat värde. Till allt detta kommer att den psykometriska kompetensen bland psykologer och andra praktiker är sviktande [134].

Inget hindrar givetvis någon från att helt frankt påstå att ett test eller en bedömningsmetod "fungerar bra", trots avsaknad av forskning till stöd för påståendet. Det är därför viktigt för användarna att utveckla sin kompetens så de lär sig att ställa höga krav på de bedömningsinstrument som de använder. Testförläggare och konsulter tycks i åtskilliga fall enbart lägga ekonomiska aspekter på sin verksamhet. Kan något säljas, så sälj det! Utbildning av och information till uppdragsgivarna är därför av strategisk betydelse för området; problemen är följaktligen likartade i USA där många (troligen ca 80 % av företagen) inte använder sig av forskningsresultaten, förmodligen på grund av okunskap [204]. En survey i Storbritannien gav ett liknande resultat: ca 70 % av företagen använde sig inte av test vid chefsurval, men väl av intervjuer av tveksam validitet [149]. Detta innebär i sin tur att de företag som använder sig av forskningsresultat inom test och urval kan få betydande konkurrensfördelar, eftersom det finns god anledning att tro att de därigenom ökar sin potential för ekonomisk framgång [203]. Mabon har påvisat betydande ekonomiska vinster vid användning av personlighetstest [102]. Integritetstest har troligen lett till miljonvinster i form av minskade stölder av anställda, hos företag som använt dem [94], för att ta ett exempel.

Den praktiska användningen av test utgår i teorin från att testen ska ingå som en del av underlaget för ett beslut som fattas om en person, t ex att han eller hon ska gå vidare till intervju eller erbjudas ett visst jobb eller en utbildningsplats. Det är naturligt att man utvärderar testets kvalitet utifrån kriterier som utbildningsresultat eller arbetsprestation, och det är detta som litteraturen på området till största delen handlar om. Testanvändning är emellertid i praktiken inte bara detta, utan mycket annat dessutom [178]. Test, särskilt personlighetstest, används även för att:

- legitimera organisationens teknisk/vetenskapliga nivå
- legitimera ett beslut om anställning, d v s personen testas för att man ska kunna säga att han eller hon också blivit testad med personlighetstestet X
- bidra till förbättrad självinsikt hos den testade
- skapa underlag för intervju eller utvecklingssamtal genom att intervjuaren bättre ska förstå den intervjuades personlighet
- skapa motivation för en intervju och trovärdighet, så att känsliga ämnen kan tas upp

- ge ett underlag för associationer hos parterna i en intervju, så att intressanta och oväntade teman kan aktualiseras
- signalera till den testade att man är engagerad i honom/henne som person

Självinsikt är troligen en viktig målsättning och det är intressant att fundera över hur den kan åstadkommas med test - ytterst lite forskning tycks finnas om den saken. Detsamma gäller de flesta av de övriga tillkommande funktionerna hos personlighetstest. Det har länge varit känt att testade personer har en stark tendens att tro på utlåtanden de får om sina resultat på testet [50], i synnerhet om det handlar om positiva, allmänt sanna eller oklart formulerade påståenden – den så kallade Barnumeffekten<sup>16</sup>. Det märkliga är att sådana ganska klichéartade påståenden ofta tycks upplevas som specifika, och relevanta, just för en själv. De tror också i hög grad på påståenden som är falska [52], bara de inte avviker alltför mycket från deras självuppfattning. Trankell demonstrerade Barnumeffekten på grafologiska diagnoser [206]. Hans undersökning är ju en klassiker i Sverige men måhända har den förlett en del läsare till att tro att den bara gäller grafologiska utlåtanden. Så är inte fallet.

Vad ska man se upp med om man är kund till testare? Det viktigaste är troligen att undvika att gå i fällan "Vad var det jag sa" eller "Jag är en god psykolog". Vi har alla en förutfattad mening om vår förmåga att bedöma andra människor. Konsulten smickrar oss genom att sätta i våra händer ett instrument som artikulerar vissa common-sense begrepp, och vi tror på honom för att vi redan hade denna uppfattning (plus en massa andra) även om vi inte uttryckt den riktigt så. Försäkringar om metodens förträfflighet accepteras sedan, och vetenskaplig dokumentation efterfrågas inte - eller den finns men är svår att tillgodogöra sig för den som inte studerat statistik och psykometrik. Text är det ganska vanligt att konsulten pekar på att hans eller hennes metod uppnått vissa "statistiskt signifikanta" samband, och det krävs då av kunden att han eller hon förstår vad signifikans är och vilka slutsatser den tillåter [176]. Dessa är mycket begränsade men ganska få torde vara medvetna om det. Konsulten klarar sig därför att de kritiska frågorna aldrig blir ställda, och för att människors spontana uppfattningar om andras personlighet är (1) att den har en allt överskuggande betydelse, och (2) att de är skarpa iakttagare av sina medmänniskor. Alla är nöjda, oavsett kvaliteten på använd metodik. Skengiltighet (face validity) hos ett test är det intryck som det ger av att vara tillförlitligt och giltigt, ett intryck som ofta säger ytterst lite om den faktiska tillförlitligheten, som bara kan fastställas genom systematisk forskning om testet [38].

Ännu 40 efter Cronbachs varningar om skengiltighet tycks test accepteras okritiskt bara det "verkar bra". Högskolans utbildning av psykologer, men också av andra nyckelgrupper som ekonomer och socionomer, borde innehålla mycket mera omfattande och mycket mera vetenskaplig och kritisk utbildning om test och personbedömning.

---

16 . Man brukar tillskriva P. T. Barnum det cyniska påståendet "There is a sucker born every minute". Även om det inte är historiskt korrekt att det var Barnum som gjorde uttalandet så är det tankeväckande. Godtrogenhet är en vanlig egenskap, utnyttjad av skojare i alla tider. Att människor låter sig luras till att tro på sådant som horoskop är ett mera specifikt fenomen, som först beskrevs av Forer [50]. Forer presenterade ett godtyckligt valt horoskop för en grupp personer. De flesta tyckte att det stämde mycket bra in på just dem. Hans experiment har upprepats många gånger, med samma resultat.



## Framtiden

Denna artikel har handlat om prognos av ledarskap och chefslämplighet. Hur blir det i framtiden? Framtidens ledarskap vet vi inget om, vi kan bara gissa. De ganska stereotypa uppfattningar som ständigt hörs i denna fråga kan diskuteras [174]. Men hur det än blir med värderingar och vilka personlighetsdimensioner som kommer att efterfrågas så torde de rent metodiska problem vid mätning och observation förbli ungefär desamma. Låt oss se på några av de mera lovande ansatserna på området.

## Begåvningsstest

Under de senaste decennierna har mycken möda ägnats åt att förbättra begåvningsstesten genom att knyta an till den moderna kognitiva psykologin. Hittills har resultaten av denna forskning och utveckling varit magra. Ungefär samma typ av test används idag som för flera decennier sedan, och med ungefär samma grad av framgång. Denna grad av framgång är som ovan påpekats modest men viktig och kostnadseffektiva testapplikationer kan konstrueras på nuvarande testmetodiks grund.

Schmidt [160] summerade läget på följande sätt:

*"After over 50 years of research, including some very major research efforts during the 1980s, it is now evident that refinements in the measurement of abilities and aptitudes are unlikely to contribute nontrivial increments to validity beyond that which is produced by good measures of general mental ability. The areas of personality, biographical data, physical abilities, and perhaps interests are considerably more promising in that respect." (p.348-349)*

Utvecklingen har alltså gått långsamt framåt, ibland inte alls. Sternberg skriver:

*"Whereas the ideas generating most technologies have moved forward by leaps and bounds over the last few decades, the testing industry, like its tests, has remained largely (although not entirely) static...Thus, whereas old computers and VCRs and old telephones go rapidly out of date, old tests never seem to die, except for updating of norms and cosmetic changes". [196].*

Det påpekades ovan att i Sverige använda test som MBTI, Thomas-systemet och Rorschach har sina rötter i 20-talet. Vilken annan verksamhet i samhället med anspråk på vetenskaplig status är en direkt tillämpning av teori och metod som är 75 år gammal?

Testen har alltså inte utvecklats speciellt snabbt. Det har visat sig svårt att komma längre än bröderna Allport gjorde 1921. Ett exempel är engelska flygvapnets testbatteri för urval av piloter under första världskriget; det fungerade ungefär så bra som moderna test [86]. Inom just detta område visar sig följande: testen bli sämre och sämre allt eftersom tekniken utvecklas [86].

Sternberg förklarar situationen med okunskap både bland testare och deras kunder, och eta-

bländat av utvärdering av t ex skolor som bygger på vissa standardtest. Problematiken är inte exakt densamma i Sverige, men stora likheter finns onekligen. Troligen är ett av problemen den fast grundmurade tron som många har att deras erfarenhet är en god grund för insiktsfull bedömning av människor. Erfarenheten har ju skapats i arbete bl a med vissa traditionella test. Forskningen om erfarenhet och psykologbedömning visar att det är en felaktig föreställning att den skulle vara entydigt och starkt positiv [56; 57] .

Det verkar alltså som om prediktioner på en ganska global nivå av förmåga till anpassning och prestation kan göras med ganska enkla medel, och det är svårt att nå väsentligt längre. Några andra uppslag till förnyelse ska emellertid nämnas. En persons resultat på ett begåvningsstest beror givetvis på hur mycket han eller hon anstränger sig. Kan man mäta testmotivation? Frågan kan tyckas uppenbar men ytterst lite forskning finns om den och den borde uppmärksammas mera. I flervalstest borde t ex stereotypa sekvenser av svar kunna tas som tecken på låg motivation [181] . Modern datorteknik kan möjliggöra en sådan utveckling.

## ***Emotionell intelligens***

Ett begrepp som fått mycket stor uppmärksamhet är emotionell (och social) intelligens [62; 63; 180; 184] . Emotionell intelligens är ett begrepp som bygger på forskning kring förmågan att hantera emotioner, t ex att identifiera emotioner hos andra eller i musik och konst [112; 114] . Detta är en intressant variabel och det tycks finnas vissa samband mellan olika aspekter på emotionell och social förmåga [179; 185] . Men det finns också ett suddigt sätt att hantera denna term, och man kan kalla lite av varje för emotionell intelligens och saluföra begreppet som en ny beteckning på samma gamla personlighetsvariabler som varit i omlopp i ett århundrade. Bar-Ons test av emotionell intelligens är av denna typ [12] .

Det finns två typer av test på emotionell intelligens: prestation och självrapport. Prestationstest bygger på de ursprungliga idéerna av Mayer, Salovey och Caruso [30; 113] . Det innehåller uppgifter av typen att identifiera emotioner i ansiktsuttryck. Testet, som kallas MSCEIT, har acceptabel reliabilitet och har gett en del intressanta resultat i forskningen [115] . Den andra typen av test på emotionell intelligens exemplifieras av Bar-On testet; det finns många liknande ansatser. Prestations- och självrapportskalorna på emotionell intelligens ger olika resultat. Självrapportskalorna är känsliga för de testades önsknings att ge ett gott intryck, medan prestationsskalorna inte är det [185] .

## ***Integritet***

Kommer personlighetstestningen att bjuda på ännu mera radikala innovationer? Konvergens mot Big Five har redan lett till att det finns ett antal test som mäter dessa dimensioner [210] . Åtminstone en av dem, nämligen Noggrannhet/Pålitlighet, har befunnits ha validitet i arbetslivet [13] . (Ej att förväxla med Perfektionism [155] som troligen har negativ validitet för många typer av kvalificerade uppgifter ehuru det kan vara en viktig sak i andra sammanhang). Mätning av integritet kan ske på två sätt: direkt genom frågor av uppenbart innehåll

relaterade till moral och ärlighet, eller indirekt [156] . Det senare är av tveksam relevans, det förra kan genomskådas och manipuleras av den testade [23; 87] .

De två typerna av integritetstest tycks inte samvariera nämnvärt [81] , och de direkta har visat sig ha betydligt högre validitet [157] . Tveksamhet råder alltså bland experterna om denna typ av test, även om vissa positiva resultat dokumenterats och den senaste utvecklingen varit positiv för testen [208] . Ashton [9] fann att specifika mått hade bättre prognostisk validitet än de breda Big Five när det gällde sambandet med ärlighet. Intressant är att den enda bakgrundsfaktorn som samvarierade med integritet var kön, i en studie av Ones och Viswesvaran [138] . Kvinnor framstod som hederligare än män. I en studie av missbruk visade sig ett integritetstest användbart för att upptäcka en del anställda med sådana påproblem, men långt oiffrån alla [109] .

## **Social önskvärdhet**

Ny forskning behövs för att närmare studera social önskvärdhet och tendens att svara "ja" på alla frågor. Tidsmätningar på mikronivå, alltså för varje fråga, är numera möjliga och kan underlätta utvecklingen av en sådan metodik [80; 164] . Responstid tycks vara användbar som ett sätt att begreppsligt validera traditionella lögnskalor [80] . Återigen är det datorn som erbjuder nya möjligheter. Datormetodik kan också användas för att skraddarsy en testsession för varje testad person, så att tiden minimeras och tillförlitligheten maximeras vid given kostnad i form av tid [101] . Datorbaserade test tycks ge resultat som är mycket jämförbara med traditionellt administrerade test [47; 146] .

Nya idéer om mätning av tendensen att ge socialt önskvärda svar har publicerats [143] . Han skiljer mellan självbedräglig skönmålning och intrycksmanipulation och ser den senare dimensionen som den bästa "lögnskalan" [145] . Skönmålning av den egna personen befanns vara en destruktiv tendens [144] . Stora svårigheter är förenade med personlighetstestning i situationer där de sökande är intensivt angelägna om ett ge ett gott intryck för att få chansen att t ex komma in på en eftertraktad utbildning, i synnerhet om man använder ett normativt snarare än ett ipsativt format<sup>17</sup> [20] . Ett citat:

*"Faking has indeed been a current concern in research on personality measurement [213] . However, research on the issue has supported an optimistic conclusion: the effects on validity are small even in a selection situation. Extensive research on the question has shown that social desirability does not appear to influence scores to the extent of jeopardizing their validity [14; 18; 139] , and that faking more generally seems to be a somewhat marginal problem [39] , as does coaching on similar and related tests [140] ." [179]*

Frågan är om emellertid om detta är hela sanningen. Det visar sig för det första att om testpoäng korrigeras för intrycksmanipulation (mätt med sk lögnskalor) kan rangordningen mellan de testade bli väsentligt förändrad. Om t ex de 15% som har högst poäng beaktas kan ca 1/3 av dessa bytas ut om man går från okorrigerad till korrigerad poäng. I individuella fall kan effekten av korrigerad bli dramatisk. Det är kanske sant att validiteten hos testen inte förbätt-

---

17 . Samtidigt är ju ipsativa format olämpliga eftersom de inte tillåter jämförelser mellan individer.

ras av korrektion, men den tycks inte heller försämrats. Då är det närmast en etisk fråga om korrektion ska göras eller ej och enligt min mening talar det mesta för att korrektion är berättigad, om lögnskalorna har tillräckligt trovärdighet. Dessa skalor får visst stöd i resultat med aktiv simulering av intryck, d v s de testade får instruktionen att försöka framställa sig själva i så positiv dager som möjligt. De resultat som erhållits med en sådan instruktion i en reell urvalssituation av stor betydelse är stöd för lögnskalornas användbarhet (och ger uppslag till nya sätt att mäta intrycksmanipulation) [185]. Personlighetsdimensionerna kunde till ca 30% förklaras av intrycksmanipulation, vilket är mycket, men långt ifrån hela den sanna variansen. Deras användbarhet i denna typ av situation kan inte ifrågasättas utifrån dessa resultat, men givetvis behövs mycket mera forskning om "faking".

## **Simuleringsövningar**

En radikalt annorlunda idé än traditionella test är att använda simuleringsövningar, s k assessment center [60]. Här försöker man bygga en social miljö med uppgifter av en typ som liknar dem som kommer att ingå i det framtida arbetet. Psykologer observerar kandidaterna i sådan miljö under lång tid, av storleksordningen flera dagar. Liknande tankegångar tillämpades av den tyska krigsmakten före och under andra världskriget, men systematiska och statistiska utvärderingar gjordes troligen ej, i vilket fall publicerades de inte [166; 167], och idén föll i vanrykte. Den moderna forskningen visar emellertid på en del goda resultat av denna metodik men den är dyrbar och knappast bättre än goda test. Dessutom finns en hel del kritik av de psykometriska egenskaperna hos bedömningar av simuleringsövningar [49].

## **Slutsatser**

Personlighet har betydelse för framgång i ledarskap och andra sammanhang, och en flerstegsmodell med ett första filter baserat på begåvningsstest som följs av personlighetstest och/eller strukturerade intervjuer är en mycket kostnadseffektiv metod för urval. Förutsättningen är att ganska många kandidater finns till befattningen. Om så ej är fallet torde det inte hjälpa ens med de mest effektiva metoder som står till buds för personbedömning och det enda som återstår är att gå på biografiska data och principen "more of the same", alltså att noga studera vederbörandes tidigare prestationer i liknande jobb, om sådan information finns att tillgå.

Samtidigt är det efter så mycket intensiv forskning allt tydligare att individfaktorer på psykologisk nivå bara mäter en del av det som är viktigt för framgång och prestation. Vad kan svara för resten - i själva verket för huvudparten av variationen? Det vet vi inte, men troligen är situationsfaktorer av stor betydelse [132]. Det finns stabila individuella skillnader i kategorier av situationer [129], men om situationerna är psykologiskt olika uppkommer en variation i beteendet. Det kan också handla om psykologiska faktorer som vi ännu ej känner till, om fysiologiska faktorer eller om sociala skeenden som interagerar med det individbundna på ett intrikat och tills vidare okänt sätt. Psykologerna saknar inte arbetsuppgifter.

## Referenser

- [1]. Allport, F. H., & Allport, G. W. (1921). Personality traits: Their classification and measurement. *Journal of Abnormal and Social Psychology*, 16, 6-40.
- [2]. Allport, G. (1927). Concepts of trait and personality. *Psychological Bulletin*, 24, 284-293.
- [3]. Allport, G. W., & Vernon, P. E. (1931). A test for personal values. *Journal of Abnormal and Social Psychology*, 25, 368-372.
- [4]. Allport, G. W., & Vernon, P. E. (1933). *Studies in expressive movement*. New York: Haffner.
- [5]. Ambady, N., & Rosenthal, R. (1992). Thin slices of expressive behavior as predictors of interpersonal consequences: A meta-analysis. *Psychological Bulletin*, 111, 256-274.
- [6]. Andersen, J. A. (1999). Are power motivated leaders more effective? A review of McClelland's theory. In M. A. Rahim & R. T. Golembiewski (Eds.), *Current topics in management* (pp. 41-59). Stamford, CT: Jai Press.
- [7]. Anderson, N. R. (1992). Eight decades of employment interview research: a retrospective meta-review and prospective commentary. *Eur. Work Organ. Psychol.*, 2, 1-32.
- [8]. Arnau, R. C., Green, B. A., Rosen, D. H., Gleaves, D. H., & Melancon, J. G. (2003). Are Jungian preferences really categorical?: An empirical investigation using taxometric analysis. *Personality & Individual Differences*, 34, 233-251.
- [9]. Ashton, M. C. (1998). Personality and job performance: The importance of narrow traits. *Journal of Organizational Behavior*, 19, 289-303.
- [10]. Barbuto, J. E., Jr. (1997). A critique of the Myers-Briggs Type Indicator and its operationalization of Carl Jung's psychological types. *Psychological Reports*, 80, 611-625.
- [11]. Barker, R. G. (1968). *Ecological psychology*. Stanford: Stanford University Press.
- [12]. Bar-On, R. (1997). *The emotional quotient inventory (EQ-i): Technical manual*. Toronto: Multi-Health Systems.
- [13]. Barrick, M. R., & Mount, M. K. (1991). The Big Five personality dimensions and job performance: A meta-analysis. *Personnel Psychology*, 44, 1-26.
- [14]. Barrick, M. R., & Mount, M. K. (1996). Effects of impression management and self-deception on the predictive validity of personality constructs. *Journal of Applied Psychology*, 81, 261-272.
- [15]. Bernreuter, R. G. (1931). *The personality inventory*. Stanford, CA: Stanford University Press.
- [16]. Bess, T. L., & Harvey, R. J. (2002). Bimodal score distributions and the Myers-Briggs Type Indicator: Fact or artifact? *Journal of Personality Assessment*, 78, 176-186.
- [17]. Bliesener, T. (1996). Methodological moderators in validating biographical data in personnel selection. *Journal of Occupational and Organizational Psychology*, 69, 107-120.
- [18]. Borman, W. C., Hanson, M. A., & Hedge, J. W. (1997). Personnel selection. *Annual Review of Psychology*, 48, 299-337.
- [19]. Borman, W. C., & Motowidlo, S. J. (1997). Task performance and contextual performance: The meaning

- for personnel selection research. *Human Performance*, 10, 99-109.
- [20]. Bowen, C.-C., Martin, B. A., & Hunt, S. T. (2002). A comparison of ipsative and normative approaches for ability to control faking in personality questionnaires. *International Journal of Organizational Analysis*, 10, 250-259.
- [21]. Boyatzis, R. E. (1982). *The competent manager: A model for effective performance*. New York: Wiley.
- [22]. Boyle, G. J. (1995). Myers-Briggs Type Indicator (MBTI): Some psychometric limitations. *Australian Psychologist*, 30, 71-74.
- [23]. Brown, R. D., & Cothorn, C. M. (2002). Individual differences in faking integrity tests. *Psychological Reports*, 91, 691-702.
- [24]. Bruner, J. S. (1957). On perceptual readiness. *Psychological Review*, 64, 123-152.
- [25]. Brunswik, E. (1956). *Perception and the representative design of psychological experiments* (2nd Ed. ed.). Berkeley: University of California Press.
- [26]. Bunn, G. C. (1997). The lie detector, Wonder Woman and liberty: the life and work of William Moulton Marston. *History of the Human Sciences*, 10, 91-119.
- [27]. Burnett, J. R., & Motowidlo, S. J. (1998). Relations between different sources of information in the structured selection interview. *Personnel Psychology*, 51, 963-983.
- [28]. Byravan, A., & Ramanaiah, N. V. (1995). Structure of the 16 PF fifth edition from the perspective of the five-factor model. *Psychological Reports*, 76, 555-560.
- [29]. Capraro, R. M., & Capraro, M. M. (2002). Myers-Briggs Type Indicator score reliability across studies: A meta-analytic reliability generalization study. *Educational & Psychological Measurement*, 62, 590-602.
- [30]. Caruso, D. R., Mayer, J. D., & Salovey, P. (2002). Relation of an ability measure of emotional intelligence to personality. *Journal of Personality Assessment*, 79, 306-320.
- [31]. Chapman, L. J., & Chapman, J. P. (1967). Genesis of popular but erroneous psychodiagnostic observations. *Journal of Abnormal Psychology*, 73, 193-204.
- [32]. Chapman, L. J., & Chapman, J. P. (1969). Illusory correlation as an obstacle to the use of valid psychodiagnostic signs. *Journal of Abnormal Psychology*, 74, 271-280.
- [33]. Colbert, C. (1998). *A measure of perfection: Phrenology and the fine arts in America*.
- [34]. Conger, J. A., & Kanungo, R. N. (1994). Charismatic leadership in organizations: Perceived behavioral attributes and their measurement. *Journal of Organizational Behavior*, 15, 439-452.
- [35]. Conway, J. M., & Peneno, G. M. (1999). Comparing structured interview question types: Construct validity and applicant reactions. *Journal of Business and Psychology*, 13, 485-506.
- [36]. Cook, M. (1992). An evaluation of the DISC/Personal Profile Analysis. *Selection and Development Review*, 8, 3-6.
- [37]. Cortina, J. M., Goldstein, N. B., Payne, S. C., Davison, H. K., & Gilliland, S. W. (2000). The incremental validity of interview scores over and above cognitive ability and conscientiousness scores. *Personnel Psychology*, 53, 325-351.
- [38]. Cronbach, L. J. (1960). *Essentials of psychological testing*. New York: Harper & Row.
- [39]. Cunningham, M. R., Wong, D. T., & Barbee, A. P. (1994). Self-presentation dynamics on overt integrity

- tests: Experimental studies of the Reid Report. *Journal of Applied Psychology*, 79, 643-658.
- [40]. Dipboye, R. L. (1992). *Selection interviews: Process perspectives*. Cincinnati, OH: South Western.
- [41]. Dixon, N. F. (1976). *On the psychology of military incompetence*. London: Jonathan Cape.
- [42]. Driver, R. W., Buckley, M. R., & Frink, D. D. (1996). Should we write off graphology? *International Journal of Selection and Assessment*, 4, 78-86.
- [43]. Drucker, P. F. (1985). How to make people decisions. *Harvard Business Review*.
- [44]. Ericsson, K. A., & Charness, N. (1994). Expert performance. Its structure and acquisition. *American Psychologist*, 49, 725-747.
- [45]. Ericsson, K. A., & Smith, J. (Eds.). (1991). *Toward a general theory of expertise*. Cambridge: Cambridge University Press.
- [46]. Fairbanks, J. K., & Goldman, M. (1998). *China: A new history*. Cambridge MA: Belknap Press of Harvard University Press.
- [47]. Finger, M. S., & Ones, D. S. (1999). Psychometric equivalence of the computer and booklet forms of the MMPI. A meta-analysis. *Psychological Assessment*, 11, 58-66.
- [48]. Fishbein, M., & Ajzen, I. (1975). *Belief, attitude, intention, and behavior: An introduction to theory and research*. Reading, MA: Addison-Wesley.
- [49]. Fleenor, J. W. (1996). Construct and development assessment centers: Further troubling empirical findings. *Journal of Business and Psychology*, 10, 319-335.
- [50]. Forer, B. R. (1949). The fallacy of personal validation: a classroom demonstration of gullibility. *Journal of Abnormal & Social Psychology*, 44, 118-123.
- [51]. Friedman, H., & MacDonald, D. A. (2002). Using transpersonal tests in humanistic psychological assessment. *Humanistic Psychologist*, 30, 223-236.
- [52]. Furnham, A. (1989). Personality and the acceptance of diagnostic feedback. *Personality & Individual Differences*, 10, 1121-1133.
- [53]. Furnham, A., & 711-716, V. (1990). The fakeability of the 16-PF, Myers-Briggs and FIRO-B personality measures. *Personality & Individual Differences*, 11, 711-716.
- [54]. Furnham, A., Moutafi, J., & Crump, J. (2003). The relationship between the revised NEO-Personality Inventory and the Myers-Briggs Type Indicator. *Social Behavior & Personality*, 31, 577-584.
- [55]. Ganzach, Y., Kluger, A. N., & Klagman, N. (2000). Making decisions from an interview: Expert measurement and mechanical combination. *Personnel Psychology*, 53, 1-20.
- [56]. Garb, H. N. (1986). Clinical judgment, clinical training, and professional experience. *Psychological Bulletin*, 105, 387-396.
- [57]. Garb, H. N. (1998). *Studying the clinician: Judgment research and psychological assessment*. Washington, DC: APA.
- [58]. Gardner, H. (1983). *Frames of mind: The theory of multiple intelligences*. New York: Basic Books.
- [59]. Gardner, H. (1993). *Multiple intelligences: The theory in practice*. New York: Basic Books.
- [60]. Gaugler, B. B., Rosenthal, D. B., Thornton, G. C., III, & Bentson, C. (1987). Meta-analysis of assessment

- center validity. *Journal of Applied Psychology*, 72, 493-511.
- [61]. Ghiselli, E. E. (1966). *The validity of occupational aptitude tests*. New York: Wiley.
- [62]. Goleman, D. (1995). *Emotional intelligence*. New York: Bantam Books.
- [63]. Goleman, D. (1998). *Working with emotional intelligence*. New York: Bantam Books.
- [64]. Goodstein, L. D., & Lanyon, R. I. (1999). Applications of personality assessment to the workplace: A review. *Journal of Business and Psychology*, 13, 291-322.
- [65]. Grove, W. M., & Meehl, P. E. (1996). Comparative efficiency of informal (subjective, impressionistic) and formal (mechanical, algorithmic) prediction procedures: The clinical-statistical controversy. *Psychology, Public Policy, and Law*, 2, 293-323.
- [66]. Gustafsson, J.-E. (1984). A unifying model for the structure of intellectual abilities. *Intelligence*, 8, 179-203.
- [67]. Gustafsson, J.-E., & Balke, G. (1993). General and specific abilities as predictors of school achievement. *Multivariate Behavioral Research*, 28, 407-434.
- [68]. Guthke, J., Beckmann, J. F., & Schmidt, G. (2002). Ist an der Graphologie doch etwas dran? Untersuchungen zur Uebereinstimmung von Graphologenuurteil und psychometrischen Pereoenlichkeitstests./Is graphology of diagnostic value at all? On the relationship between graphological judgement and psychometric te sts. *Zeitschrift fur Personalpsychologie*, 1, 171-176.
- [69]. Hampson, S. (1999). State of the art: Personality. *Psychologist*, 12, 284-288.
- [70]. Hardy, T., & Evans, G. L. (1983). *Psychology's occult doubles: Psychology and the problem of pseudo-science*: Nelson-Hall.
- [71]. Harkness, A. R., & McNulty, J. L. (1994). The Personality Psychopathology Five (PSY-5): Issue from the pages of a diagnostic manual instead of a dictionary. In S. Strack & M. Lorr (Eds.), *Differentiating normal and abnormal personality* (pp. 291-315). New York: Springer.
- [72]. Harris, M. M. (1989). Reconsidering the employment interview: a review of recent literature and suggestions for future research. *Personnel Psychology*, 40, 691-726.
- [73]. Hendrickson, T. M. (1983). *Personal profile analysis: a technical manual*. Marlow: Thomas International Systems (Europe) Ltd.
- [74]. Henkel, T. G., & Wilmoth, J. N. (1992). Factor analysis of the Personal Profile System. *Journal of Experimental Education*, 60, 271-280.
- [75]. Hogan, J. C., Hogan, R., & Gregory, S. (1992). Validation of a sales representative selection inventory. *Journal of Business and Psychology*, 7, 161-171.
- [76]. Hogan, J. C., Hogan, R., & Murtha, T. (1992). Validation of a personality measure of managerial performance. *Journal of Business and Psychology*, 7, 225-237.
- [77]. Hogan, R. (1994). Trouble at the top. Causes and consequences of managerial incompetence. *Consulting Psychology Journal*, 46, 9-15.
- [78]. Hogan, R., Curphy, G. J., & Hogan, J. (1994). What we know about leadership: Effectiveness and personality. *American Psychologist*, 49, 493-504.
- [79]. Hogan, R., Hogan, R., & Roberts, B. W. (1996). Personality measurement and employment decisions. *American Psychologist*, 51, 469-477.



- [80]. Holden, R. R. (1998). Detecting fakers on a personnel test: Response latencies versus a standard validity scale. *Journal of Social Behavior and Personality*, 13, 387-398.
- [81]. Hough, J., & Brinkmeyer, K. (1997). Bridging the gap between overt and personality-based integrity tests. *Personnel Psychology*, 50, 587-599.
- [82]. Hough, L. M. (1984). Development and evaluation of the "accomplishment record" method of selecting and promoting professionals. *Journal of Applied Psychology*, 69, 135-146.
- [83]. Hough, L. M., Eaton, N. K., Dunnette, M. D., Kamp, J. D., & McCloy, R. A. (1990). Criterion-related validities of personality constructs and the effect of response distortion on those validities. *Journal of Applied Psychology*, 75, 581-595.
- [84]. House, R. J., & Howell, J. M. (1992). Personality and charismatic leadership. *Leadership Quarterly*, 3, 81-108.
- [85]. Hufcutt, A. I., & Arthur, W. (1994). Hunter and Hunter (1984) revisited: interview validity for entry-level jobs. *Journal of Applied Psychology*, 79, 184-190.
- [86]. Hunter, D. R., & Burke, E. F. (1995). *Handbook of pilot selection*. Aldershot: Avebury Aviation.
- [87]. Hurtz, G. M., & Alliger, G. M. (2002). Influence of coaching on integrity test performance and unlikely virtues scale scores. *Human Performance*, 15, 255-273.
- [88]. Husén, T. (1946). Psykologiska undersökningar av flygspiranter. *Tidskrift i Militär Hälsovård*, 71, 51-98.
- [89]. Ichisada, M. (1981). *China's examination hell: the civil service examinations of imperial China*. New Haven, CN: Yale University Press.
- [90]. Irvine, S. H. (1989). *Personal profile analysis technical handbook. A contemporary frame of reference for research and development*. Ormskirk: T. Lyster.
- [91]. Irvine, S. H. (1994). Valid and more valid? Keys to understanding personal appraisal practices at work. *Current Psychology: Developmental, Learning, Personality, Social*, 13, 27-59.
- [92]. Jelf, G. S. (1999). A narrative review of post-1989 employment interview research. *Journal of Business and Psychology*, 14, 25-58.
- [93]. Johnson, C. E., Wood, R., & Blinkhorn, S. F. (1988). Spuriouser and spuriouser: the use of ipsative personality tests. *Journal of Occupational Psychology*, 61, 153-162.
- [94]. Jones, J. W., & Terris, W. (1991). Integrity testing för personnel selection: An overview. *Forensic Reports*, 4, 117-140.
- [95]. Jung, C. G. (1950). *Psychologische Typen*, 8. Aufl. Zürich: Rascher.
- [96]. Jung, C. G. (1971). *Psychological types*. Princeton NJ: Princeton University Press.
- [97]. Kaila, E. (1935). *Personligheters psykologi*. Stockholm: Natur och Kultur.
- [98]. Kelly, I. W. (1999). "Debunking the debunkers". A response to an astrologer's debunking of the skeptics. *Skeptical Inquirer*, 23, 37-41.
- [99]. Littorin, P. (1994). *Talangjakten*. Stockholm: Ekerlid.
- [100]. Littorin, P. (1994). *Tester: hjälpmedel vid rekrytering, outplacement, projektarbete och utvecklingssamtal*. Stockholm: Ekerlid.

- [101]. Luecht, R. M. (1998). Computer-assisted test assembly using optimization heuristics. *Applied Psychological Measurement*, 22, 224-236.
- [102]. Mabon, H. (1998). Utility aspects of personality and performance. *Human Performance*, 11, 289-304.
- [103]. Mabon, H. (2001). *Arbetspsykologisk testning. Om urvalsmetoder i arbetslivet*. Stockholm: Psykologiförlaget.
- [104]. MacLaren, V. V. (2001). A qualitative review of the Guilty Knowledge Test. *Journal of Applied Psychology*, 86, 674-683.
- [105]. Magnusson, D. (1967). *The generality of behavioral data III*. Stockholm: Applied Psychology Unit, Psychological Laboratories, University of Stockholm.
- [106]. Marston, W. M. (1989/1928). *Emotions of normal people*. Ormskirk, Lancs.: Thomas Lyster.
- [107]. Martin, B. A., Bowen, C.-C., & Hunt, S. T. (2002). How effective are people at faking on personality questionnaires? *Personality & Individual Differences*, 32, 247-256.
- [108]. Martinussen, M., & Torjussen, T. (1993). Does the Defense Mechanism Test (DMT) predict pilot performance only in Scandinavia? In R. D. Jensen & D. Neumeister (Eds.), *Proceedings of the Seventh International Symposium on Aviation Psychology* (pp. 398-403). Columbus: Ohio State University.
- [109]. Mastrangelo, P. M., & Jolton, J. A. (2001). Predicting on-the-job substance abuse with a written integrity test. *Employee Responsibilities & Rights Journal*, 13, 95-106.
- [110]. Matthews, G., & Oddy, K. (1997). Ipsative and normative scales in adjectival measurement of personality: Problems of bias and discrepancy. *International Journal of Selection and Assessment*, 5, 169-182.
- [111]. Matthews, G., & Stanton, N. (1994). Item and scale factor analyses of the Occupational Personality Questionnaire. *Personality & Individual Differences*, 16, 733-743.
- [112]. Mayer, J. D., Caruso, D., & Salovey, P. (1999). Emotional intelligence meets traditional standards for an intelligence. *Intelligence*, 27, 267-298.
- [113]. Mayer, J. D., DiPaolo, M. T., & Salovey, P. (1990). Perceiving affective content in ambiguous visual stimuli: A component of emotional intelligence. *Journal of Personality Assessment*, 54, 772-781.
- [114]. Mayer, J. D., & Geher, G. (1996). Emotional intelligence and the identification of emotion. *Intelligence*, 22, 89-113.
- [115]. Mayer, J. D., Salovey, P., Caruso, D. R., & Sitarenios, G. (2003). Measuring emotional intelligence with the MSCEIT V2.0. *Emotion*, 3, 97-105.
- [116]. McClelland, D. C. (1973). Testing for competence rather than for "intelligence". *American Psychologist*, 28, 1-14.
- [117]. McClelland, D. C. (1998). Identifying competencies with behavioral-event interviews. *Psychological Science*, 9, 331-339.
- [118]. McCrae, R. R., & Costa, P. T. (1996). Towards a new generation of personality theories: theoretical contexts for the Five Factor Model. In J. Wiggins (Ed.), *The Five Factor Model of personality: theoretical perspectives* (pp. 51-87). New York: Guilford Press.
- [119]. McCrae, R. R., & Costa, P. T., Jr. (1987). Validation of the five-factor model of personality across instruments and observers. *Journal of Personality and Social Psychology*, 52, 81-90.

- [120]. McHenry, J. J., Hough, L. M., Toquam, J. L., Hanson, M. A., & Ashworth, S. (1990). Project A validity results: The relationship between predictor and validity domains. *Personnel Psychology*, *43*, 335-354.
- [121]. McManus, M. A., & Kelly, M. L. (1999). Personality measures and biodata: Evidence regarding their incremental predictive value in the life insurance industry. *Personnel Psychology*, *52*, 137-148.
- [122]. Meehl, P. E. (1954). *Clinical versus statistical prediction: A theoretical analysis and a review of the evidence*. Minneapolis: University of Minnesota Press.
- [123]. Mendoza-Denton, R., Ayduk, O., Mischel, W., Shoda, Y., & Testa, A. (2001). Person \* Situation interactionism in self-encoding (I am . . . when . . .): Implications for affect regulation and social information processing. *Journal of Personality & Social Psychology*, *80*, 533-544.
- [124]. Metcalfe, J. (1998). Cognitive optimism: Self-deception or memory-based processing heuristics? *Personality & Social Psychology Review*, *2*, 100-110.
- [125]. Milton, J. (2002). *The road to Malpsychia. Humanistic psychology and our discontents*. San Francisco: Encounter Books.
- [126]. Mischel, W. (1968). *Personality and assessment*. New York: Wiley.
- [127]. Mischel, W. (2004). Toward an integrative science of the person. *Annual Review of Psychology*, *55*, 1-22.
- [128]. Mischel, W., & Peake, P. K. (1982). Beyond deja vu in the search for cross-situational consistency. *Psychological Review*, *89*, 730-755.
- [129]. Mischel, W., & Shoda, Y. (1995). A cognitive-affective system theory of personality: Reconceptualizing situations, dispositions, dynamics, and invariance in personality structure. *Psychological Review*, *102*, 246-268.
- [130]. Motowidlo, S. J., Borman, W. C., & Schmit, M. J. (1997). A theory of individual differences in task and contextual performance. *Human Performance*, *10*, 71-83.
- [131]. Mount, M. K., Witt, L. A., & Barrick, M. R. (2000). Incremental validity of empirically keyed biodata scales over GMA and the five factor personality constructs. *Personnel Psychology*, *53*, 299-323.
- [132]. Mumford, M. D., Connelly, M. S., Helton, W. B., Strange, J. M., & Osburn, H. K. (2001). On the construct validity of integrity tests: Individual and situational factors as predictors of test performance. *International Journal of Selection & Assessment*, *9*, 240-257.
- [133]. Mumford, M. D., Marks, M. A., Connelly, M. S., Zaccaro, S. J., & Johnson, J. F. (1998). Domain-based scoring of divergent thinking tests: Validation evidence in an occupational sample. *Creativity Research Journal*, *11*, 151-163.
- [134]. Mårdberg, B. (1996). Improper use of tests in Sweden. *International Journal of Selection and Assessment*, *4*, 154-158.
- [135]. Mårdberg, B., Nienmaa, P., Hillström, B., & Carlstedt, B. (1994). *Myers-Briggs Type Indicator. Manual*. Stockholm: Psykologiförlaget.
- [136]. Noll, R. (1994). *The Jung cult. Origins of a charismatic movement*. Princeton, NJ: Princeton University Press.
- [137]. Nutt, P. C. (1986). Decision style and strategic decisions of top executives. *Technical Forecasting and Social Change*, *30*, 39-62.
- [138]. Ones, D. S., & Viswesvaran, C. (1998). Gender, age, and race differences on overt integrity tests. Results across four large-scale job applicant data sets. *Journal of Applied Psychology*, *83*, 35-42.

- [139]. Ones, D. S., Viswesvaran, C., & Reiss, A. D. (1996). Role of social desirability in personality testing for personnel selection: The red herring. *Journal of Applied Psychology, 81*, 660-679.
- [140]. Palmer, D. R., & Busciglio, H. H. (1996). Coaching on the ASVAB: Analysis of posttest questionnaire responses. *Military Psychology, 8*, 267-278.
- [141]. Paltiel, L. (1986). Self-appraisal personality batteries. *Guidance & Assessment Review, 2*, 3-7.
- [142]. Parangimalil, B. J. (1990). *Toward integral holism in psychology*. New Dehli, India: Inter-India Publications.
- [143]. Paulhus, D. L. (1991). Measurement and control of response bias. In J. P. Robinson & P. R. Shaver & L. S. Wrightsman (Eds.), *Measures of personality and social-psychological attitudes* (pp. 17-59). San Diego, CA: Academic Press.
- [144]. Paulhus, D. L. (1998). Interpersonal and intrapsychic adaptiveness and trait self-enhancement: A mixed blessing? *Journal of Personality and Social Psychology, 74*, 1197-1208.
- [145]. Paulhus, D. L., Bruce, N. N., & Trapnell, P. D. (1995). Effects of self-presentation strategies on personality profiles and their structure. *Personality and Social Psychology Bulletin, 21*, 100-108.
- [146]. Potosky, D., & Bobko, P. (1997). Computer versus paper-and-pencil administration mode and response distortion in noncognitive selection tests. *Journal of Applied Psychology, 82*, 293-299.
- [147]. Robertson, I. T., & Kandola, R. S. (1982). Work sample tests: Validity, adverse impact and applicant reaction. *Journal of Occupational Psychology, 55*, 171-183.
- [148]. Robertson, I. T., & Kinder, A. (1993). Personality and job competences: The criterion-related validity of some personality variables. *Journal of Occupational & Organizational Psychology, 66*, 225-244.
- [149]. Robertson, I. T., & Makin, P. J. (1986). Management selection in Britain: A survey and critique. *Journal of Occupational Psychology, 59*, 45-57.
- [150]. Robins, R. W., & John, O. P. (1997). The quest for self-insight. Theory and research on accuracy and bias in self-perception. In R. Hogan & J. Johnson & S. Briggs (Eds.), *Handbook of personality psychology* (pp. 649-679). San Diego, CA: Academic Press.
- [151]. Rosse, J. G., Stecher, M. D., Miller, J. L., & Levin, R. A. (1998). The impact of response distortion on preemployment personality testing and hiring decisions. *Journal of Applied Psychology, 83*, 634-644.
- [152]. Rothbart, M., & Park, B. (1986). On the confirmability and disconfirmability of trait concepts. *Journal of Personality and Social Psychology, 50*, 131-142.
- [153]. Russell, C. J. (1990). Selecting top corporate leaders: An example of biographical information. *Journal of Management, 16*, 73-86.
- [154]. Ryan, A. M., McFarland, L., Baron, H., & Page, R. (1999). An international look at selection practices: Nation and culture as explanations for variability in practice. *Personnel Psychology, 52*, 359-391.
- [155]. Saboonchi, F. (2000). *Perfectionism. Conceptual, emotional, psychopathological, and health-related implications*. Stockholm: Department of Psychology, University of Stockholm.
- [156]. Sackett, P. R., Burris, L. R., & Callahan, C. (1989). Integrity testing for personnel selection: An update. *Personnel Psychology, 42*, 491-529.
- [157]. Sackett, P. R., & Wanek, J. E. (1996). New developments in the use of measures of honesty, integrity, conscientiousness, dependability, trustworthiness, and reliability for personnel selection. *Personnel*

*Psychology*, 49, 787-829.

- [158]. Saggino, A., Cooper, C., & Kline, P. (2001). A confirmatory factor analysis of the Myers-Briggs Type Indicator. *Personality & Individual Differences*, 30, 3-9.
- [159]. Sapadin, L. (1999). *Beat procrastination and make the grade. The six styles of procrastination and how students can overcome them*. Harmondsworth, Middlesex: Penguin.
- [160]. Schmidt, F. L. (1994). The future of personnel selection in the U.S. Army. In M. G. Rumsey & C. B. Walker & J. H. Harris (Eds.), *Personnel selection and classification* (pp. 333-350). Hillsdale, NJ: Erlbaum.
- [161]. Schmidt, F. L., & Hunter, J. E. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. *Psychological Bulletin*, 124, 262-274.
- [162]. Schmidt, F. L., & Rader, M. (1999). Exploring the boundary conditions for interview validity: Meta-analytic findings for a new interview type. *Personnel Psychology*, 52, 445-464.
- [163]. Shafer, A. B. (2000). Relation of the Big Five to Biodata and aspect of the self. *Personality and Individual Differences*, 28, 1017-1035.
- [164]. Siem, F. M. (1996). The use of reponse latencies to enhance self-report personality measures. *Military Psychology*, 8, 15-29.
- [165]. Silverman, B. I. (1971). Studies of astrology. *Journal of Psychology*, 77, 141-149.
- [166]. Simoneit, M. (1940). *Vom Werden der deutschen Wehrmachtpsychologie: Ein geschichtlicher Ruckblick* (Wehrpsychologische Mitteilungen 1-60).
- [167]. Simoneit, M. (1944). *Grundriss der charakterologischen Diagnostik*. Leipzig: Teubner.
- [168]. Sjöberg, L. (1971). The new functionalism. *Scandinavian Journal of Psychology*, 12, 29-52.
- [169]. Sjöberg, L. (1982). Attitude-behavior correlation, social desirability and perceived diagnostic value. *British Journal of Social Psychology*, 21, 288-292.
- [170]. Sjöberg, L. (1982). *Upplevelse och prestation. Några utvecklingslinjer inom teoretisk psykologi. (Experience and achievement. Some lines of development in theoretical psychology)*. Lund: Doxa.
- [171]. Sjöberg, L. (1987). *Test vid urval av sökande till Handelshögskolan i Stockholm. (The use of tests for the selection of students applying to the Stockholm School of Economics)* (Forskningsrapport 2-87). Stockholm: Psykologisk Metod AB.
- [172]. Sjöberg, L. (1991). Urvalstest för stress- och chefsbefattningar. Var står forskningen idag? (Selection tests for stressful and management jobs. Current research).
- [173]. Sjöberg, L. (1998). World views, political attitudes and risk perception. *Risk - Health, Safety and Environment*, 9, 137-152.
- [174]. Sjöberg, L. (1999). Framtidens ledarskap - myt eller verklighet? (Future leadership - myth or reality?). In E. Lemaire (Ed.), *Anno 1998* (pp. 114-126). Malmö: Akademiförlaget Corona AB.
- [175]. Sjöberg, L. (1999). Maslowmyten. (The Maslow myth). *Folkvett*, 6-16.
- [176]. Sjöberg, L. (1999). Psykologisk forskningsmetodik och praktik: tre haverier. (Methods and practice of psychological research: three disasters). *VEST: Journal for Science and Technology Studies*, 12, 5-25.

- [177]. Sjöberg, L. (2000). Exploatering av ångestkapitalet. (Exploiting the anxiety capital). *Folkvett*.
- [178]. Sjöberg, L. (2000). Lättare sagt än gjort att hitta radikalt nya urvalsmetoder. *FOA Tidningen*, 38, 14-15.
- [179]. Sjöberg, L. (2001). Emotional intelligence: A psychometric analysis. *European Psychologist*, 6, 79-95.
- [180]. Sjöberg, L. (2001, 17 Augusti). Emotionell intelligens ingen framgångsgaranti. (Emotional intelligence no guarantee of success). *Svenska Dagbladet*, pp. 5.
- [181]. Sjöberg, L. (2003). Good and not-so-good ideas in psychological research. A tutorial in idea assessment and generation. *VEST: Journal for Science and Technology Studies*, 16, 33-68.
- [182]. Sjöberg, L. (2003). Humanistisk psykologi - den mörka sidan (The dark side of humanistic psychology). *Folkvett*, 21, 6-15.
- [183]. Sjöberg, L. (2003, 14 September). Rättssäkerheten hotas av en bläckplump. *Svenska Dagbladet*, pp. Kultur 9.
- [184]. Sjöberg, L., & Engelberg, E. (2003). *Emotionell intelligens - teori och empiri i ett psykologiskt perspektiv. (Emotional intelligence - theory and empirical research in a psychological perspective)* (SSE/EFI Working Paper Series in Business Administration 2003:6). Stockholm: Stockholm School of Economics.
- [185]. Sjöberg, L., & Engelberg, E. (2004). *Measuring and validating emotional intelligence as performance or self-report* (SSE/EFI Working Paper Series in Business Administration 2004:3). Stockholm: Stockholm School of Economics.
- [186]. Sjöberg, L., & Engelberg, E. (in press). Lifestyles, risk perception and consumer behavior. *International Review of Sociology*.
- [187]. Sjöberg, L., Källmén, H., & Scharnberg, M. (1998). Selection for stressful jobs: Is the Defence Mechanism Test the solution? In R. Flin & M. Salas & M. Strub & L. Martin (Eds.), *Decision making under stress: Emerging themes and applications* (pp. 252-260). Aldershot: Ashgate.
- [188]. Sjöberg, L., & Magneberg, R. (1990). Action and emotion in everyday life. *Scandinavian Journal of Psychology*, 31, 9-27.
- [189]. Sjöberg, L., & Tollgerdt-Andersson, I. (1985). *Vad är personkemi? Socialpsykologisk forskning om attraktivitet. (What is person chemistry? Social psychological research on attractiveness)*. Stockholm: Scandinavian Executive Search.
- [190]. Spearman, C. E. (1904). General intelligence, objectively determined and measured. *American Journal of Psychology*, 15, 201-293.
- [191]. Spillmann, B. (1998). Die Wirklichkeit des Schattens: Kritische Überlegungen zu C. G. Jungs Haltung während des Nationalsozialismus und zur Analytischen Psychologie. (The reality of the Shadow: Critical reflections on Jung's attitudes during the time of the German National Socialism and on Jungian Psychology.). *Analytische Psychologie*, 29, 272-295.
- [192]. Spranger, E. (1925). *Lebensformen : geisteswissenschaftliche Psychologie und Ethik der Persönlichkeit*. Halle: Niemeyer.
- [193]. Sternberg, R. J. (1985). *Beyond IQ: A triarchic theory of human intelligence*. New York: Cambridge University Press.
- [194]. Sternberg, R. J. (1997). Managerial intelligence: Why IQ isn't enough. *Journal of Management*, 23, 475-493.

- [195]. Sternberg, R. J. (1998). Metacognition, abilities, and developing expertise: What makes an expert student? *Instructional Science*, 26, 127-140.
- [196]. Sternberg, R. J. (1999). The theory of successful intelligence. *Review of General Psychology*, 3, 292-316.
- [197]. Sternberg, R. J. (1999). A triarchic approach to the understanding and assessment of intelligence in multicultural populations. *Journal of School Psychology*, 37, 145-159.
- [198]. Sternberg, R. J., Forsythe, G. B., Hedlund, J., Horvath, J. A., Wagner, R. K., Williams, W. M., Snook, S. A., & Grigorenko, E. L. (2000). *Practical intelligence in everyday life*. New York: Cambridge University Press.
- [199]. Sternberg, R. J., Wagner, R. K., Williams, W. M., & Horvath, J. A. (1995). Testing common sense. *American Psychologist*, 50, 912-927.
- [200]. Stricker, G., & Gold, J. R. (1999). The Rorschach: Toward a nomothetically based, idiographically applicable configurational model. *Psychological Assessment*, 11, 240-250.
- [201]. Sändh, B. (1999). Heltäckande horoskop. *Folkvett*, 40-43.
- [202]. Taylor, S. E., & Brown, J. D. (1988). Illusion and well-being: A social psychological perspective on mental health. *Psychological Bulletin*, 103, 193-210.
- [203]. Terpstra, D. E., & Rozell, E. J. (1993). The relationship of staffing practices to organizational level measures of performance. *Personnel Psychology*, 46, 27-48.
- [204]. Terpstra, D. E., & Rozell, E. J. (1997). Why some potentially effective staffing practices are seldom used. *Public Personnel Management*, 26, 483-495.
- [205]. Tett, R. P., Jackson, D. N., & Rothstein, M. (1991). Personality measures as predictors of job performance: A meta-analytic review. *Personnel Psychology*, 44, 703-742.
- [206]. Trankell, A. (1961). *Magi och förnuft i människobedömning*. Stockholm: Bonnier.
- [207]. Undheim, J. O., & Gustafsson, J.-E. (1987). The hierarchical organisation of cognitive abilities: Restoring general intelligence through the use of linear structural relations (LISREL). *Multivariate Behavioral Research*, 22, 149-171.
- [208]. Wanek, J. E. (1999). Integrity and honesty testing: What do we know? How do we use it? *International Journal of Selection and Assessment*, 7, 183-195.
- [209]. White, R. P., & DeVries, D. L. (1990). Making the wrong choice: Failure in the selection of senior-level managers. *Issues & Observations*, 10, 1-6.
- [210]. Widiger, T. A., & Trull, T. J. (1997). Assessment of the five-factor model of personality. *Journal of Personality Assessment*, 68, 228-250.
- [211]. Wiggins, J. (1989). Review of the Myers-Briggs Type Indicator. *Mental Measurements Yearbook*, 10, 537-538.
- [212]. Wiggins, J. S., & Trapnell, P. D. (1997). Personality structure: The return of the Big Five. In R. Hogan & J. Johnson & S. Briggs (Eds.), *Handbook of personality psychology* (pp. 737-766). San Diego, CA: Academic Press.
- [213]. Visweswaran, C., & Ones, D. S. (1999). Meta-analysis of fakability estimates: Implications for personality measurement. *Educational and Psychological Measurement*, 59, 197-210.
- [214]. Wood, J. M., Nezworski, M. T., Lilienfeld, S. O., & Garb, H. N. (2003). *What's wrong with the Ror-*

*schach?* San Francisco, CA: Jossey-Bass.

- [215]. Wright, P. M., Lichtenfels, P. A., & Pursell, E. (1989). The structured interview: additional studies and a meta-analysis. *Journal of Occupational Psychology*, 62, 191-199.
- [216]. Yammarino, F. J., & Bass, B. M. (1991). Person and situation views of leadership: A multiple levels of analysis approach. *Leadership Quarterly*, 2, 121-139.
- [217]. Yukl, G., & Van Fleet, D. D. (1992). Theory and research on leadership in organizations. In M. D. Dunnette & L. M. Hough (Eds.), *Handbook of industrial and organizational psychology* (Vol. 3, pp. 147-198). Palo Alto, CA: Consulting Psychologists Press.