



**PSYKOLOGISK METOD AB**



Personlighetstestens validitet  
i arbetslivet  
Lennart Sjöberg

Rapport 2010:8

**Psykologisk Metod L Sjöberg AB arbetar med utveckling och användning av psykologiska test samt undersökningar av attityder och riskuppfattningar, andra psykologiska utredningar och tillämpad forskning.**

**Vår affärsidé är att bedriva arbetet i nära anslutning till den aktuella forskningen inom psykologin.**

Skrifter utgivna av Psykologisk Metod AB

- Sjöberg, L. Bortom Big Five: Konstruktion och validering av ett personlighetstest. Rapport 2008:1.
- Sjöberg, L., & Möller, K. (2009). Sociala arbetsfunktioner och personlighet. Rapport 2009:1.
- Sjöberg, L. (2009). *UPP*-testet: Kriterierelaterad validitet. Rapport 2009:2.
- Sjöberg, L. (2009). *UPP*-testet: Korrektion för skönmålning. Rapport 2009:3.
- Sjöberg, L. (2010). *UPP*-testet: Tredje generationens personlighetstest. Rapport 2010: 1.
- Sjöberg, L. (2010). *UPP*-testet. Manual, reviderad version, februari 2010 .
- Sjöberg, L (2010). *UPP*-testet: Mångfald och jämställdhet gynnas av korrektion för skönmålning. Rapport 2010: 2.
- Sjöberg, L. (2010). A third generation personality test. Rapport 2010:3.
- Sjöberg, L. (2010). Emotionell intelligens och social förmåga hos ungdomar. Rapport 2010:4.
- Sjöberg, L. (2010). Faktorstrukturen hos *UPP*-testet. Rapport 2010:5.
- Sjöberg, L. (2010). *UPP*-testet: Användarhandbok, april 2010.
- Sjöberg, L. (2010). Teknisk manual, april 2010.
- Sjöberg, L. (2010). *UPP*-testet och kundservice: Kriteriestudie. Rapport 2010:6.
- Sjöberg, L. (2010). *UPP/Screen*: Ett screeningtest för personlighet och begåvning. Rapport 2010:7.
- Sjöberg, L. (2010). Personlighetstestens validitet i arbetslivet. Rapport 2010:8

# Innehåll

Sammanfattning.....	1
Inledning.....	1
Historisk exposé.....	1
Obefogad pessimism om test, nyttan av screening.....	3
Femfaktormodellen.....	5
Metodsynpunkter.....	5
Analys av FFM i förhållande till arbetsresultat och andra arbetsrelaterade kriterier.....	7
Allmänna B5-dimensioner.....	7
Specifika dimensioner.....	12
Meta-analys av meta-analys.....	13
Övriga arbetsrelaterade kriterier.....	14
Arbetsstillfredsställelse.....	14
Ledarskap.....	14
Olycksbenägenhet.....	15
Sociala arbetsfunktioner.....	15
Entreprenörskap.....	17
Metodstudier och begreppsvalidering.....	18
Begreppsvalidering mot Assessment Center.....	18
FFM-test jämförda med andra test.....	18
Indexvariabler: exempel.....	19
Diskussion.....	20
Överblick.....	20
Slutsats.....	23
Referenser.....	25

## Sammanfattning

Tidigare forskning om personlighetstest och arbetsresultat gav tämligen svaga resultat. Pessimismen var utbredd, men inte på goda grunder. Det har länge varit känt att test kan ge viktiga tillskott till beslutsfattandet i t ex urvalssammanhang, även om validiteten är relativt blygsam. Från omkring 1990 ökade intresset för frågeställningen i och med introducerandet av femfaktormodellen (FFM) för personlighet. Meta-analyser av Big Five-dimensionernas (B5) användbarhet i arbetslivet sammanfattas. De har publicerats under tiden 1984-2008 och bygger på 100-tals publicerade studier. Korrelationerna visade sig i allmänhet vara låga för alla dimensioner utom noggrannhet, där de tenderade att ligga i intervallet 0.20-0.25, också det ett relativt blygsamt värde. Slutsatsen om B5-dimensionerna som sådana är i enlighet med dessa mycket omfattande undersökningar att de är svagt eller inte alls korrelerade med kriterier i arbetslivet, om de tas var för sig. Multipla korrelationer kan emellertid ge betydligt bättre resultat, även om det är kontroversiellt hur de ska uppskattas på grundval av meta-analyser. Dessa tycks ha gett för låga uppskattningar av sambanden mellan B5-dimensionerna, och därmed en överskattning av de multipla korrelationerna mot kriterier. Alltmera omfattande forskning visar att smala, fokuserade variabler fungerar betydligt bättre än B5-dimensionerna och ger viktiga tillskott till dessa. En del av dessa variabler ingår som underskalor i FFM-test, andra som emotionell intelligens ligger utanför FFM-tänkandet och ingår i UPP-testet. Sådana dimensioner ger viktiga tillskott till intelligens i prognosen av arbetsresultat, och av många andra viktiga aspekter på arbetsrelaterat beteende. Även index som bildas på ett urval av fokuserade personlighetsdimensioner kan ge mycket goda resultat, som närmar sig den prognoskraft som g-faktortest har.

## Inledning

### *Historisk exposé*

Personlighetstest har använts i arbetslivet under många år, och lika lång har diskussionen om deras nytta varit (Giles, Burch & Anderson, 2008). Sedan 1990-talet har denna diskussion dominerats av femfaktormodellen (FFM). Jag ska här sammanfatta vad som är känt om personlighetsdimensioners användbarhet i arbetslivet, med särskild inriktning på FFM och på grundval av publicerade meta-analyser.

Men först en historisk notis av intresse i sammanhanget. Man brukar säga att det första personlighetstestet tillkom 1917: Woodsworth Personal Data Sheet (Gibby & Zickar, 2008). Det var ett självrapporttest. Gibbys bild av den tidiga personlighetstestningen var att den handlade om anpassning ("adjustment"). Woodworth-testet syftade till att predicera vilka soldater som skulle råka ut för "shell shock" under första världskriget. En omfattande verksamhet med personlighetstestning utvecklades under 1920-talet i Woodworths efterföljd och kulminerade i Bernreuters test (BPI), också ett självrapporttest. BPI innehöll skalor på sex egenskaper: neuroticism, självtillräcklighet, dominans, självmedvetenhet, introversion och ensamhet. Några av dessa begrepp finns med i moderna testinstrument, men man saknar noggrannhet i BPI (Mera om noggrannhet senare i denna rapport). En översikt publicerad 1942 tog upp ca 130 publicerade studier av BPI-testet (Super, 1942).

BPI var ett självrapporttest, och det är en teknik som kommit dominera. Meta-analyser tyder på att självrapporter är relativt rättvisande, om de jämförs med andras bedömningar (Connolly, Kavanagh, & Viswesvaran, 2007). Parallellt med BPI utvecklades emellertid helt

annorlunda metoder för personlighetstestning, nämligen projektiva test som TAT (Winter, 1998). Ganska stora förhoppningar knöts till dem, även om de främst var avsedda för kliniskt bruk. Projektiva test används fortfarande i arbetslivet, t ex DMT, men deras validitet är mycket omstridd (Sjöberg, 1981; Sjöberg, Källmén, & Scharnberg, 1998). Olika former av objektiva eller fysiologiska mått har också misslyckats. Simuleringsövningar ("assessment centers") har gett resultat jämförbara med självrapporttest (Collins, et al., 2003; Thornton Iii & Gibbons, 2009), men är dyrbara och det är tveksamt om de är kostnadseffektiva. Termen Assessment Center tycks numera användas mycket liberalt och kan ibland även gälla ett upplägg med två eller flera självrapportinstrument.

Det är svårt att hitta validitetsdata på de tidigaste personlighetstesten. Visserligen nämns begreppet validitet och undersökningar gjordes men resultaten är svåra att tolka eftersom man ofta nöjde sig med att rapportera om sambanden mellan test och kriterium var statistiskt signifikanta eller ej.

Ellis och Conrad publicerade 1948 en översikt av validiteten hos personlighetstest i militära tillämpningar (Ellis & Conrad, 1948). När det gäller arbetsprestation tycks validiteterna ha legat i genomsnitt på ca 0.15 i ett 40-tal redovisade studier<sup>1</sup>. Inga korrelationer för mätfel eller begränsad variation gjordes, trots att kunskaper om dessa problem fanns vid den tidpunkten (Gulliksen, 1950; Thorndike, 1949).

Ett intressant och ambitiöst försök rapporterades av amerikanska armén 1952 (Crook, Burke, & Klieger, 1952). De fann en korrelation på 0.27 med kriteriet "discharged for unfavorable reasons" från armén. De ansåg detta var en alltför låg korrelation, men fann att personlighetstestet i kombination med annan information gav en acceptabel multipel korrelation.

En tidig litteratursammanställning gav ett genomsnitt på 0.25 över 8 yrkesgrupper (Ghiselli & Barthol, 1953). Guion och Gottier (Guion & Gottier, 1965) publicerade 1965 en stort upplagd genomgång av validiteter av personlighetstest i arbetslivet. Deras rapport är något svårläst då de inte ger värdena för icke signifikanta samband. Nivån på de korrelationer som de rapporterar är jämförbar med den som kommit fram i moderna studier. Deras slutsats om användbarheten av personlighetstest var emellertid negativ.

Ghiselli genomförde på 60- och 70-talen ett monumentalt arbete och sammanställde all då tillgänglig valideringsforskning (Ghiselli, 1973a, 1973b), forskning publicerad 1920-71. Sambanden mellan personlighetstest och kriterier låg i intervallet 0.2 – 0.3.

Arbeten av Guion och Gottier samt Ghiselli ledde till en pessimistisk syn på personlighetstesten. Till pessimismen bidrog i hög grad en översikt av personlighetspsykologin som publicerades 1968 av Mischel (Mischel, 1968). Han gjorde en översikt av hur invariant beteendet är över olika situationer och tyckte sig finna att beteendet varierade starkt mellan situationer. Det var situationen som bestämde beteendet, inte personen, eller möjligen interaktionen mellan person och situation. Personlighetstest korrelerade "bara" omkring 0.3 med kriterier, enligt Mischel, som ansåg att de därmed var föga användbara.

---

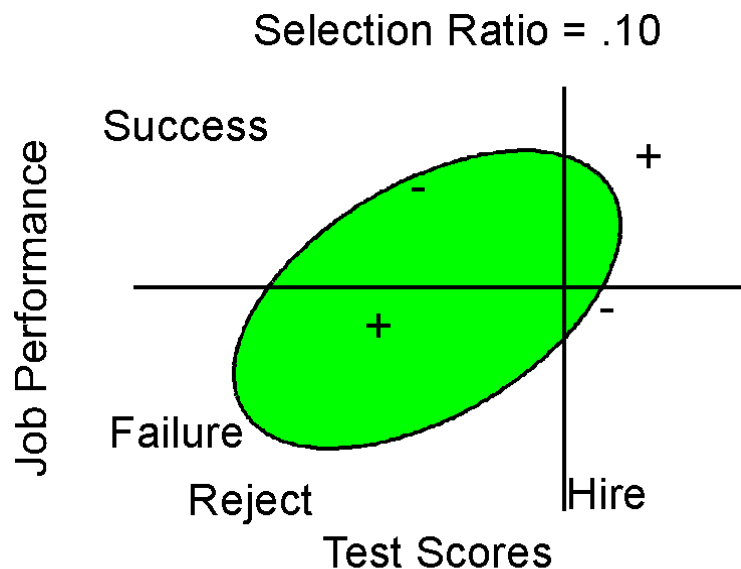
<sup>1</sup> Det är svårt att avgöra var det exakta värdet ligger på grund av att olika studier redovisas på olika sätt: ibland med någon form av korrelation, ibland med signifikansberäkningar.

Argumentet var emellertid traditionellt redan 1968. Hull (Hull, 1928) hade samma pessimistiska syn 40 år tidigare. Ännu längre tillbaka kan Allport och Allport citeras (F. H. Allport & G. W. Allport, 1921). De fann validitet hos mot kamratskattningar som kriterium som låg omkring 0.40. Gordon Allport själv (G. W. Allport, 1921) beskriver ett av deras resultat så här:

”... the scores on sample problems representing the individual's behavior in situations which would require his taking an active or a passive role, correlate *slightly* with personality ratings by associates on this trait ( $r = .40$ )” (Sid. 451, kursiverat här)

### **Obefogad pessimism om test, nyttan av screening**

Är det alldeles för låga korrelationer med kriterier, om de ligger på 0.3 – 0.4? Det beror på vad de ska användas till. Vid urval kan man få bra resultat av tillämpning av ett test med validitet på den nivån, visade Taylor och Russell (Taylor & Russell, 1939) i ett klassiskt arbete. Deras princip kan enkelt illustreras, se Figur 1.



Figur 1. Illustration av urvalskvotens betydelse för resultatet av anställningsbeslut.

I detta exempel anställs 10 % av de sökande. Trots att testet är långt ifrån perfekt kan de som anställs förväntas ge ett bra resultat, betydligt bättre än ett rent slumpmässigt urval. Vid ännu mindre urvalskvot än 10 % blir det förväntade resultatet ännu bättre, även om testet fortfarande är lika ”trubbigt”.

Denna analys är övertygande men missar en viktig aspekt: antalet sökande, oavsett urvalskvoten. Ju flera sökande, desto större är nämligen chansen att et ska finnas en ”stjärna” i gruppen. Om vi är nöjda med att vår ”stjärna” tillhör de 10 % bästa är chansen mycket stor, 96 %, att han eller hon finns i gruppen.

Screening av många sökande till ett jobb är ett bra sätt att hitta högpresterande personer. Varför? Antag att vi är ute efter en verklig elitperson (vem är inte det?), bäst av 100, alltså den bäst presterande procenten. Ju fler vi testar desto större är chansen att det finns minst en

sådan person i den testade gruppen, se Tabell 1 nedan. Tabellen visar också sannolikheten för att det i gruppen av testade finns minst en person i 98:e percentilen och en i 90:e.

Tabell 1. Sannolikheten att få minst en jobbsökande som tillhör 99:e, 98:e eller 90:e percentilen vid testning av 5-75 personer

Antal testade personer	99:e percentilen	98:e percentilen	90:e percentilen
5	0.049	0.096	0.410
10	0.096	0.183	0.651
15	0.140	0.261	0.794
20	0.182	0.332	0.878
25	0.222	0.397	0.928
30	0.260	0.455	0.958
35	0.267	0.507	0.975
40	0.331	0.554	0.985
45	0.364	0.597	0.991
50	0.395	0.636	0.995
55	0.425	0.671	0.997
60	0.453	0.702	0.998
65	0.480	0.731	0.999
70	0.505	0.757	0.999
75	0.529	0.780	1.000

Tabellen visar att testning av 75 personer ger höga sannolikheter för att hitta minst en person med mycket hög kapacitet, och att chansen är mycket god redan vid 25 testade om vi nöjer oss med kapacitet vid 90:e percentilen, även det ett mycket gott resultat. Om vi är ute efter mera extremt hög kompetens, 1 på 100, har vi hyfsade 53 % chans redan vid 75 sökande. I många sammanhang är det realistiskt att vänta sig hundratals sökande, eller fler.

Knepet är alltså att utgå från ett så stort underlag som möjligt. Det är emellertid dyrbart och tidsödande att göra det om man inte använder enkla, valida och billiga screeningtest. Det ger dessutom en konkurrensfördel eftersom screeningstestningens effektivitet inte är känd av flera än ett fåtal. Man kan förstås gå på intuition och "magkänsla" men påverkas då av irrelevanta faktorer som utseende, kroppsspråk, dialekt, klädsel etc, se vår monografi om personkemi (Sjöberg & Tollgerdt-Andersson, 1985). I en studie av anställningsintervjuer (Higgins & Judge, 2004) fann man att det som avgjorde varken var erfarenhet eller förmåga - det var om man tyckte bra eller illa om den jobbsökande, vilket i sin tur berodde på om han eller hon lyckades med "ingratiation". Med den termen menas ett strategiskt försök att göra ett gott intryck, t ex genom smicker, framhålla likhet i värderingar eller allmänt sett att framhålla sina goda egenskaper på ett trovärdigt sätt.

I min egen erfarenhet har jag gjort extremt goda rekryteringar genom screening av 40-60 jobbsökande. Och faktiskt varit en medioker rekryterare i många övriga fall. Det är intressant att fundera över vilka rekryteringsprinciper som kommer till användning för viktiga chefsjobb i samhället. Hur många övervägde man innan ministrar anställdes i Alliansregeringen? (Alla tycks inte ha tillhört eliten, med tanke på vad som sedan hänt). Hur bred är urvalsbasen för VD-jobben och styrelseposterna för de stora företagen? Ju smalare den är desto större är

chansen att missa verkliga superstars ty de tillhör bara i undantagsfall "old boys' network", men han eller hon finns i en grupp om 50 sökande. Om vi är ute efter ännu mera extrem kompetens, den bästa på 100, krävs flera sökande.

## **Femfaktormodellen**

Personlighetsområdet som sådant var så snårigt fram till omkring 1990. Det var svårt att hitta rätt bland alla de hundratals variabler som föreslogs av olika forskare. FFM innebär en dramatisk förändring av bilden. Intresset har sedan 20-30 år inriktats på meta-analyser av femfaktormodellens värde i arbetslivet. De fem faktorerna är

- Emotionell stabilitet
- Extraversion/Utåtvändhet
- Öppenhet
- Vänlighet/Smidighet
- Noggrannhet

FFM med sina fem övergripande dimensioner "Big Five" (B5) (McCrae & Costa, 1987) har kommit att spela en stor roll i arbetspsykologisk testning också i vårt land. Välkända B5-test är OPQ<sup>2</sup>, HPI<sup>3</sup>, MBTI<sup>4</sup> och NEO-PI-R. Flera exempel på utveckling av svenska B5-skalor finns (Bäckman & Carlstedt, 2010; Sjöberg, 2010a). Ett speciellt fall är 16PF, ursprungligen från 1949 och med en ganska snårig uppsättning av 16 personlighetsdimensioner<sup>5</sup>.

I det följande redogör jag för meta-analyser av personlighetstestens, särskilt B5-dimensionernas, validitet mot kriterier i arbetslivet. Jag tar också upp en begreppsvalidering mot Assessment Center-data. I de meta-analyser som citeras nedan är rapporterade validitetskoefficienter vanligen korrigerade för mätfel i kriterierna och begränsad variation i personlighetsvariablerna. Mätfel i kriterierna sänker korrelationerna och detsamma gäller begränsad variation (Gulliksen, 1950). I det senare fallet gjordes korrektionen oftast enligt metodik, Thorndikes Case 2 (se Gulliksen), som gav konservativa uppskakningar av korrelationerna i motsvarande oselegerade grupper. Ett undantag kommer att kommenteras i texten.

Korrektionerna för begränsad variation ger i dessa sammanhang ganska små effekter medan korrektionerna för mätfel i kriterierna kan bli betydande, åtminstone när det gäller subjektiva kriterier (bedömningar), där reliabiliteterna, graden av samstämmighet mellan bedömare, tenderar att ligga om kring 0.5 – 0.6, alltså tämligen låga värden (Viswesvaran, Ones, & Schmidt, 1996).

## **Metodsynpunkter**

I det följande kommer jag att huvudsakligen referera till meta-analyser men i några fall tar jag även upp enskilda undersökningar, eller sammanställningar av forskning med enstaka test. Sammanställningen bygger på linjära korrelationer och man kan undra om icke-linjära

---

<sup>2</sup> OPQ har ganska nyligen befunnits vara ett B5-test (Brown & Bartram, 2009). Det konstruerades ju ursprungligen långt innan B5 hade etablerats; analysen av testets struktur var länge svår att göra på grund av dess ipsativa karaktär.

<sup>3</sup> HPI arbetar med 7 dimensioner i stället för 5, men tycks mäta i grunden samma begrepp som B5.

<sup>4</sup> MBTI täcker bara 4 av 5 dimensionerna. Mått på emotionell stabilitet saknas..

<sup>5</sup> <http://en.wikipedia.org/wiki/16PF>

sambandsanalyser skulle kunna ge starkare resultat. Noggranna analyser tyder på att så inte är fallet i allmänhet (Whetzel, McDaniel, Yost, & Kim, 2010), även om intressanta undantag kan förekomma (Zhang & Bartol, 2010).

Den forskning som diskuteras här är i liten utsträckning svensk. Det är amerikanska studier som dominerar; en del europeiska har också publicerats (Storbritannien, Spanien, Tyskland). Jag har nästan uteslutande använt mig av tidskriftsartiklar. Ytterst få forskare inom området publicerar originalforskning i böcker. Givetvis finns det bra översikter i bokform (Furnham, 2008), men det handlar om läroböcker och de är i allmänhet omkring 5 år efter i förhållande till forskningens utveckling, ibland ännu mera. Databaser över forskning gör det möjligt att relativt snabbt få en överblick över vad som publicerats. Dessutom brukar artiklar i de ledande tidskrifterna innehålla mycket användbara litteraturoversikter.

Forskningen jag diskuterar är kvantitativt orienterad. Personligheten mäts med personlighetstest och olika aspekter på beteende i arbetslivet mäts även de kvantitativt, antingen som ”objektiva” mått eller som självbedömningar eller bedömningar av chefer eller arbetskamrater. Det är, om man så vill, positivistisk forskning och den känns säkert främmande för en del. Men hela frågeställningen – sambandet mellan personlighet och arbetsresultat – är ju från början kvantitativt tänkt och det är svårt att se hur den skulle kunna analyseras på något annat sätt så att resultaten kan summeras överskådligt och jämföras mellan olika branscher och kulturer. Meta-analys är en relativt ny metodik som har gjort sådana sammanfattningar möjliga. Men det krävs förstås också att många forskare har angripit problemställningarna på någorlunda jämförbara sätt. Barrick och Mount (Barrick & Mount, 2003), som var de första som använde meta-analys i stor skala på detta område, har gett en utmärkt historisk överblick. De påpekar att FFM och meta-analys etablerades samtidigt (omkring 1990) och att båda var nödvändiga för det forskningsgenombrott som skulle komma att följa.

Erfarenheterna av meta-analys inom området personlighet och arbetsresultat är nu ganska omfattande; ca 20 års intensivt arbete med analyser av bokstavligen hundratals publicerade studier har gett stabila resultat. Forskningsområdet är kumulativt. Vi vet mera nu än vi gjorde 1990. Praktiska tillämpningar inom området personlighetstestning utvecklas i anslutning till forskningen. Ett exempel är emotionell intelligens som blev ett mycket populärt tema på 1990-talet, mycket tack vare Golemans böcker (Goleman, 1995, 1998) och den enorma uppmärksamhet de fick i media. På svaga grunder gjorde Goleman mycket starka anspråk på att i emotionell intelligens ha funnit en suveränt betydelsefull personlighetsdimension. Media och konsultvärlden kopplade genast på. Det gjorde emellertid även forskarna. Omfattande empirisk forskning har nu visat att det finns en kärna av sanning i idén om vikten av emotionell intelligens, låt vara att den inte alls ligger på den nivå som Goleman och andra tänkte sig för 15 år sedan (Joseph & Newman, 2010). Tillämpningar inom testområdet har med framgång utnyttjat den omfattande forskningen på området (Sjöberg, 2001b, 2010c).

Kvantitativt orienterad beteendeforskning gör det alltså möjligt för oss att veta *mera*, inte bara att veta *något annat*. Det är en verksamhet som varken domineras av filosofisk spekulering eller inåtvänt teoribygge. Poängen är att en stor internationell forskargrupp samverkar med likartade metoder och frågeställningar på att sådant sätt att man kan bygga på varandras resultat och, som sagt, nå fram till bättre kunskaper och bättre praktiska metoder. Det är ”sanningens ögonblick” när man frågar om ett forskningsområde: Vet vi mera idag än vi gjorde för 20 år sedan? Kan vi göra bättre prognoser?

# Analyser av FFM i förhållande till arbetsresultat och andra arbetsrelaterade kriterier

## Allmänna B5-dimensioner

Den första publicerade meta-analysen kom redan 1984 (Schmitt, Gooding, Noe, & Kirsch, 1984). Den ger inga detaljer när det gäller personlighet, bara en övergripande siffra på 0.15. Efter 7 år kom nästa analys (Barrick & Mount, 1991). Den är fortfarande speciellt intressant eftersom den rapporterar data på en mera specifik nivå än senare meta-analyser, nämligen för olika yrkesgrupper samt för både subjektiva och objektiva kriterier. De övergripande resultaten framgår av Tabell 1.

Tabell 1. Genomsnittliga korrelationer mellan B5-dimensioner och arbetsresultat (Barrick & Mount, 1991)

Dimension	Korrelation
Emotionell stabilitet	0.08
Extraversion/Utåtvändhet	0.13
Öppenhet	0.11
Vänlighet/Smidighet	0.07
Noggrannhet	0.22

Resultaten bygger på 162 studerade stickprov, totalt 23994 respondenter; data från 1950-talet och framåt t o m 80-talet. Värdena var snarlika för objektiva och subjektiva kriterier samt för olika yrkesgrupper. Denna artikel är bland de mest citerade inom området och den anslöt tonen för senare arbeten. Ofta sägs att den skapade optimism bland personlighetsforskare med intresse för tillämpning inom arbetslivet, men det är lite svårt att se varför. Korrelationerna är ju låga och på samma nivå som tidigare översikter.

Samma år (1991) publicerade Tett, Jackson och Rothstein en meta-analys av personlighet och arbetsresultat (Tett, Jackson, & Rothstein, 1991). De byggde på 97 stickprov och totalt 13 521 respondenter. Deras sammanfattande slutsats var att det fanns ett samband på 0.24 mellan personlighetsskalor och arbetsresultat. Tabell 2 ger mera detaljerade resultat.

Tabell 2. Genomsnittliga korrelationer mellan B5-dimensioner och arbetsresultat (Tett, et al., 1991).

Dimension	Korrelation
Emotionell stabilitet	0.19
Extraversion/Utåtvändhet	0.13
Öppenhet	0.24
Vänlighet/Smidighet	0.28
Noggrannhet	0.16

Resultaten av denna meta-analys skiljer sig i viss mån från de flesta övriga. Det beror troligen på att de var speciellt inriktade på studier där personlighetsskalor och kriterier var matchade genom att de hade likartat innehåll, styrda av att just dessa samband var förväntade. De fick också högre korrelationer än Barrick och Mount, med undantag för noggrannhet.

Donovan och Hurtz publicerade år 2000 en meta-analys av B5-dimensionerna och arbetsresultat (Hurtz & Donovan, 2000), se Tabell 3. Data bygger på 35-45 korrelationer och 5000 – 8000 respondenter.

Tabell 3. Genomsnittliga korrelationer mellan B5-dimensioner och arbetsresultat (Hurtz & Donovan, 2000)

Dimension	Korrelation
Emotionell stabilitet	0.20
Extraversion/Utåtvändhet	0.13
Öppenhet	0.11
Vänlighet/Smidighet	0.09
Noggrannhet	0.06

Denna analys avviker från huvudfåran genom att noggrannhet får ett så lågt värde. En europeisk meta-analys har publicerats av Salgado (Salgado, 1997). Se Tabell 4 för resultatet som bygger på 18-32 studier, 2700-3900 respondenter. Detta är den enda studien som bygger uteslutande på europeiska data.

Tabell 4. Genomsnittliga korrelationer mellan B5-dimensioner och arbetsresultat (Salgado, 1997)

Dimension	Korrelation
Emotionell stabilitet	0.13
Extraversion/Utåtvändhet	0.08
Öppenhet	0.06
Vänlighet/Smidighet	0.01
Noggrannhet	0.15

Tabell 5 ger genomsnittliga validiteter för OPQ-testet (Brown & Bartram, 2009). Dessa värden grundas visserligen bara på ett stickprov, N=853, men de är ändå intressanta. B5-dimensionerna härleddes ur de 32 underskalorna. Kriterierna var 16 dimensioner som mätte olika aspekter på arbetsresultat. De skattades både av chefer, och av de testade själva.

**Tabell 5. Validitetskoefficienter (medelvärden) hos B5-dimensioner härledda ur OPQ-testet, N=853, chefsbedömningar och självbedömningar (Brown & Bartram, 2009).**

Dimension	Chefsbedömningar	Självbedömningar
Emotionell stabilitet	0.06	0.14
Extraversion/utåtvändhet	0.15	0.15
Öppenhet	0.05	0.11
Vänlighet/smidighet	0.00	0.03
Noggrannhet	0.01	0.18

Värdena är uppenbart låga för denna typ av härledda variabler i OPQ-testet. Bartram rapporterade också låga samband mellan arbetsprestation och B5-variabler härledda från OPQ (Bartram, 2005).

Vinchur et al. (Vinchur, Schippmann, Switzer, & Roth, 1998) har publicerat en meta-analys av sambanden mellan framgång som säljare och olika psykologiska variabler, bl. a B5-dimensionerna. Se Tabell 6.

Tabell 6. Genomsnittliga korrelationer mellan B5-dimensioner och framgång i säljarbete (Vinchur, et al., 1998).

Dimension	Korrelation
Emotionell stabilitet	0.12
Extraversion/Utåtvändhet	0.22
Öppenhet	0.06
Vänlighet/Smidighet	-0.03
Noggrannhet	0.31

Dessa resultat bygger på 12-18 studier, 818 – 2629 testade personer. Intressant nog fick de bättre resultat med mera fokuserade variabler, nämligen prestationsinriktning (medelvärde av korrelationer = 0.41) och intresse (medelvärde av korrelationer = 0.50). (Detta är variabler som ingår i *UPP*-testet).

När det gäller *OPQ*-testet och prognos av säljframgång finns en sammanställning av 9 studier (SHL, 1999-2007). För säljarbete används 14 härledda skalor från detta test, men eftersom informationen i grunden är av B5-typ är det intressant att återge resultaten här. Se Tabell 7. Namnen på variablerna har inte översatts.

Tabell 7. Genomsnittliga korrelationer mellan säljresultat (olika kriterier) och *OPQ*-skalor (SHL, 1999-2007).

Skala	Genomsnittlig validitet
Sales confidence	0.16
Sales drive	0.13
Sales resilience	-0.02
Adaptable	-0.04
Listening	-0.08
Embracing change	0.06
Developing a game plan	0.04
Making contact	0.11
Building desire	0.07
Creating options	0.08
Presenting	0.11
Closing the sale	0.18
Satisfying the customer	0.01
Managing and growing	0.14
Genomsnitt, totalt	0.07

Det är intressant att notera dessa resultat eftersom *OPQ* har en mycket stor andel av den svenska testmarknaden. Det har också ett visst intresse att endast två av de nio studierna hade ett  $N > 100$ , den gräns som kan utläsas ur EFPA:s kriterier för testgranskning. Studier med mindre stickprov sägs vara ”inadequate”. Se vidare Sjöberg (Sjöberg, 2010c).

En av de bästa meta-analyserna av B5-validitet och arbets- och utbildningsresultat publicerades år 2008 av Schmidt et al. (Schmidt, Shaffer, & Oh, 2008). De har gjort ett ambitiöst försök att hitta all relevant valideringsforskning och dessutom korrigerat resultaten för mätfel i kriterierna och begränsad variation (ny metodik) i B5-dimensionerna. Deras

värden torde därför vara de mest rättvisande som publicerats. Särskilt intressant är deras nyskapande metodik för att korrigera för begränsad variation - tidigare använda metoder gav för små effekter (Schmidt, Oh, & Le, 2006). Medianvärden av validitetskoefficienterna (korrelationer) återfinns i Tabell 8.

**Tabell 8. Validitetskoefficienter (medianer) hos B5-dimensioner samt intelligens (g-faktorn) enligt meta-analys av Schmidt et al. (Schmidt, et al., 2008).**

Dimension	Arbetsprestation	Utbildningsresultat
Emotionell stabilitet	0.13	0.16
Extraversion/utåtvändhet	0.09	0.17
Öppenhet	0.04	0.24
Vänlighet/smidighet	0.08	0.12
Noggrannhet	0.21	0.26
Intelligens <sup>6</sup>	0.62	0.67

Det är alltså bara noggrannhet som ligger över 0.20 när det gäller arbetsprestation. Medianen för alla skalorna ligger på 0.09 för prestation och 0.17 för utbildningsresultat. Intelligens har uppenbarligen en mycket högre validitet än B5-dimensionerna.

Trapmann et al. har publicerat en mycket ambitiös och noggrann meta-analys av B5-dimensionerna i relation till studieresultat i akademisk utbildning (Trapmann, Hell, Hirn, & Schuler, 2007). För studieresultat kan deras resultat sammanfattas i Tabell 9.

**Tabell 9. Genomsnittliga korrelationer mellan B5-dimensioner och studieresultat (Trapmann, et al., 2007).**

Dimension	Korrelation
Emotionell stabilitet	0.04
Extraversion/Utåtvändhet	0.01
Öppenhet	0.08
Vänlighet/Smidighet	0.04
Noggrannhet	0.22

Data bygger på 34-59 studier och totalt 11 000 – 15 000 testade personer. Kortfattat kan sägas att resultaten är mycket tydliga: endast noggrannhet tycks ha något prognosvärde i detta sammanhang, och även i det fallet är tillskottet i prognosvärde utöver kognitiv förmåga mycket blygsamt (mindre än 2 %). Trapmann et al. anser att noggrannhet har ett visst värde vid urval till högre utbildning men påpekar också att det är marginellt i förhållande till kognitiva mått. De pekar också på risken med skönmålning och skriver:

“The disadvantage of using self-reports is that they are amenable to impression management, that is, coachable and fakable. Goldberg (2001) suggested that this is a problem especially in university admission, which has very high stakes for applicants”. (Sid. 147).

<sup>6</sup> Onekligen är det intressant att se vilken överlägsen validitet som allmänbegåvning har. För en meta-analys av kognitiva testvariabler och arbetsprestation, se Bertua, Anderson och Salgado (Bertua, Anderson, & Salgado, 2005). Deras genomsnittliga validiteter ligger på 0.5 – 0.6.

Det är lätt att hålla med om detta, men ”high stakes” torde föreligga i många andra sammanhang än vid ansökningar till högre utbildning. Jag har behandlat frågan om skönmålning på annat ställe, där jag visar hur man framgångsrikt kan korrigera för denna felkälla (Sjöberg, 2009).

O’Connor och Paunonen har också genomfört en meta-analys av B5-dimensionerna och studieresultat på högskolenivå (O’Connor & Paunonen, 2007). Det är ett viktigt arbete som också går in på betydelsen av specifika dimensioner. Se Tabell 10 för resultaten.

Tabell 10. Genomsnittliga korrelationer mellan B5-dimensioner och studieresultat (O’Connor & Paunonen, 2007).

Dimension	Korrelation
Emotionell stabilitet	-0.03
Extraversion/Utåtvändhet	-0.05
Öppenhet	0.06
Vänlighet/Smidighet	0.06
Noggrannhet	0.24

Dessa resultat bygger på 19-23 studier och 4490 – 5878 personer. Som synes stämmer de bra med Trappman et al. Resultatorientering och pålitlighet var viktiga underskalor som bidrog utöver B5-dimensionerna.

Hogan och Holland har gjort ett viktigt och originellt bidrag till debatten om testvaliditet (Hogan & Holland, 2003). Deras argument är att den mest rättvisande uppskattningen av validitet erhålls om man matchar testdimensionerna med kriteriernas innehåll, i efterföljd till den tidiga meta-analysen av Tett et al., som jag redogjorde för ovan. Detta är ett exempel på en princip som ofta gäller i psykologiska data, nämligen att variablers interkorrelationer samvarierar med deras likhet (Sjöberg, 1980). Hogan och Hollands meta-analys, som bygger på data från testningar med Hogans personlighetstest HPI, tyder på att principen fungerar även i detta sammanhang. I nästan samtliga fall ger deras meta-analys den största korrelationen för matchande testdimensioner och kriterier, se Tabell 11 för fallet med matchande variabler och kriterier. (Övriga korrelationer var oftast lägre, i många fall betydligt lägre). De har en lite annorlunda strukturering av FFM och annorlunda terminologi än den vanliga och därför har jag inte översatt deras termer.

Tabell 11. Genomsnittliga korrelationer mellan personlighetsdimensioner<sup>7</sup> och matchande kriterier (Hogan & Holland, 2003).

Dimension	Korrelation
Adjustment	0.43
Ambition	0.35
Likeability	0.34
Prudence	0.36
Intellectance	0.34
School success	0.25

Det är först med denna meta-analys som det finns solid grund för optimism om personlighet och arbetsresultat.

### **Specifika dimensioner**

Varför fungerar B5-dimensionerna relativt dåligt? Ett skäl är troligen att de är för allmänna. De fokuserar inte på viktiga dimensioner i arbetslivet. Testvariabler som är specifikt inriktade på arbetslivet fungerar bättre än mera allmänna personlighetsskalor (Bowling & Burns, 2010). Ett annat skäl är B5-dimensionernas kognitiva, beskrivande, karaktär. Emotionella funktioner är viktigare än de kognitiva (Sjöberg, 2008, 2010).

Allt flera forskare inriktar sig numera på smalare aspekter än B5-dimensionerna, även om de ofta tycks begränsa sig till underskalor eller ”fasetter” i FFM (Ashton, 1998; Bergner, Neubauer, & Kreuzthaler, 2010; Lounsbury, Sundstrom, Loveland, & Gibson, 2002; Paunonen, Haddock, Forsterling, & Keinonen, 2003). En aktuell studie var inriktad på ledarskap (Bergner, et al., 2010).

Dudley et al. har gjort ett pionjärarbete med sin meta-analys av validiteten hos specifika underskalor till B5-dimensionen noggrannhet (Dudley, Orvis, Lebiecki, & Cortina, 2006). De fann att vissa underskalor i många sammanhang hade högre validitet än det övergripande måttet och att de gav tydliga tillskott till prognosförmåga med avseende på arbetsresultat. Deras resultat är relativt komplexa men man kan sammanfattningsvis säga att pålitlighet (dependability) och prestationsinriktning (achievement) är de underskalor som är mest lovande.

Tabell 12 ger en översikt av proxyvalideringar<sup>8</sup> (Sjöberg, 2010b) av *UPP*-testet (normmaterialet, ca 1000 testade personer); se Sjöberg (Sjöberg, 2008a). B5-dimensionerna jämfördes med hela testet som mäter B5 + 8 fokuserade variabler. Alla variabler hade samma vikt i analyserna.

<sup>7</sup> Det är oklart hur ”school success” kan anses vara en personlighetsvariabel.

<sup>8</sup> Validering mot variabler som man vet är relaterade till direkta mått på arbetsresultat, t ex arbetsvilja. Proxyvalidering fungerar som ersättare för explicit validering mot arbetsresultat (Sjöberg, 2010b); är mycket enklare, snabbare och billigare. Vid användning av *UPP*-testet ingår mätning av proxykriterier som standard.

Tabell 12. Validering mot 6 proxykriterier, förklarad varians (kvadrerade multipla korrelationer).

Kriterievariabel	B5	B5 + 8 fokuserade variabler
Förändringsvilja	0.137	0.274
Arbetsstillfredsställelse	0.008	0.454
Arbetsvilja	0.010	0.475
Resultatorientering	0.106	0.212
Arbetsintresse	0.054	0.389
Balans arbete – övrigt liv	0.045	0.097
Medelvärde	0.041	0.317

Förklarad varians = 0.317 motsvarar en multipel korrelation på 0.56, medan B5 ligger på det typiska  $R = 0.20$ . Slutsatsen är att man bör satsa på test som går utöver femfaktormodellen.

### **Meta-analys av meta-analys**

Barrick, Mount och Judge publicerade år 2001 en översikt av de meta-analys av sambandet mellan B5-dimensioner och arbetsresultat som då fanns tillgängliga (Barrick, Mount, & Judge, 2001). För övergripande analys av arbetsresultat mot B5-dimensioner rapporterade de följande genomsnittliga validiteter, Tabell 13.

Tabell 13. Genomsnittliga korrelationer mellan personlighetsdimensioner och arbetsresultat (Barrick, et al., 2001). Baseras på 5 meta-analys (4 för öppenhet).

Dimension	Korrelation
Emotionell stabilitet	0.12
Extraversion/Utåtvändhet	0.12
Öppenhet	0.05
Vänlighet/Smidighet	0.10
Noggrannhet	0.23

Detta kan tjäna som en sammanfattning av alla meta-analys av B5-dimensioner mot arbetsresultat. Vi ser att denna typ av analys ger en tydlig validitet enbart för noggrannhet och inte ens denna är särskilt stor. Det är 20 års forskning som ligger bakom siffrorna och hundratals studier.

## Övriga arbetsrelaterade kriterier

### Arbetstillfredsställelse

En meta-analys av sambandet mellan B5-dimensioner och arbetstillfredsställelse gav det resultat som framgår av Tabell 14 (Judge, Heller, & Mount, 2002).

Tabell 14. Genomsnittliga korrelationer mellan B5-dimensioner och arbetstillfredsställelse (Judge, Heller, et al., 2002).

Dimension	Korrelation
Emotionell stabilitet	-0.29
Extraversion/Utåtvändhet	0.25
Öppenhet	0.02
Vänlighet/Smidighet	0.17
Noggrannhet	0.26

Man kan ju säga att detta är ganska höga värden i sammanhanget. En multipel korrelation hamnar på 0.41<sup>9</sup>. Det bör emellertid noteras att arbetstillfredsställelse är en relativt svag prediktor av arbetsresultat (Judge, Thoresen, Bono, & Patton, 2001), med en korrelation på omkring 0.3.

### Ledarskap

Ledarskap och personlighet är ett stort forskningsområde. Tidigare ofta citerade påståenden om att ledarskap var i hög grad situationsspecifikt (Stogdill, 1948) har fått vika för en mera nyanserad syn där personligheten spelar stor roll (Hogan, Hogan, & Kaiser, 2011; Zaccaro, 2007). En meta-analys av B5-dimensioner mot ledarskap publicerades av Judge et al. (2002). Se Tabell 15.

Tabell 15. Genomsnittliga korrelationer mellan B5-dimensioner och ledarskap (Judge, Bono, Ilies, & Gerhardt, 2002).

Dimension	Korrelation med karriär	Korrelation med effektivitet
Emotionell stabilitet	-0.24	-0.22
Extraversion/Utåtvändhet	0.33	0.24
Öppenhet	0.24	0.24
Vänlighet/Smidighet	0.03	0.21
Noggrannhet	0.33	0.16
Multipel korrelation	0.53	0.39

<sup>9</sup> Se diskussionsavsnittet gällande multipla korrelationer.

Även om enskilda värden i tabellen är modesta kan man sluta sig till att B5 ger en bra nivå på den multipla korrelationen, särskilt prognos av vem som kommer att göra karriär som ledare. Se emellertid den avslutande diskussionen vad gäller uppskattningen av multipla korrelationer på grundval av meta-analyser; de är troligen kraftigt överskattade.

## Olycksbenägenhet

En meta-analys av Clarke och Robertson ger sambanden mellan olycksbenägenhet och B5-dimensioner, se Tabell 16.

Tabell 16. Genomsnittliga korrelationer mellan B5-dimensioner och olycksbenägenhet (Clarke & Robertson, 2005)

Dimension	Korrelation
Emotionell stabilitet	-0.21
Extraversion/Utåtvändhet	0.16
Öppenhet	0.32
Vänlighet/Smidighet	-0.26
Noggrannhet	-0.27

Antalet studier varierade mellan 11 och 30, antalet respondenter mellan 1147 och 6048. Analys av data på trafikolyckor och arbetsolyckor gav liknande resultat.

## Sociala arbetsfunktioner

Det finns en omfattande litteratur om sociala arbetsfunktioner, ”Organizational Citizen Behavior” (OCB), som är en viktig aspekt på fungerande i arbetslivet. OCB används ofta som ett kompletterande kriterium och faller ut som en tydlig faktor i tillräckligt omfattande uppsättningar av kriterievariabler, se t ex Sjöberg (Sjöberg, 2010b)

Borman et al. har publicerat en meta-analys av olika variabler, däribland några av B5-dimensionerna, i relation till OCB-mått (Borman, Penner, Allen, & Motowidlo, 2001), se Tabell 17. Endast namnen på B5-dimensionerna har översatts. B5-resultaten bygger på 8-12 studier, 1554-2378 personer. Resultaten bygger på OCB-mått som ej innehåller självskattningar av sådant beteende.

Tabell 17. Genomsnittliga korrelationer mellan B5-dimensioner och OCB-beteende (Borman, et al., 2001)

Dimension	Korrelation
Extraversion/Utåtvändhet	0.06
Vänlighet/Smidighet	0.13
Noggrannhet	0.19
Positive affectivity	0.16
Negative affectivity	-0.12
Locus of Control	0.12
Collectivism	0.04
Other oriented empathy	0.17
Helpfulness	0.15

Sambanden är som synes lika låga som för arbetsresultat, eller t o m lägre. Dalal et al. dokumenterade samband mellan OCB och negativt beteende på arbetsplatsen (Dalal, 2005). Berry, Ones och Sackett (Berry, Ones, & Sackett, 2007) gjorde en mera omfattande meta-analys av B5-dimensioner och negativt beteende på arbetsplatsen. Två typer av negativt beteende har studerats: mot arbetskamrater (ID) och mot företaget (OD). Se Tabell 18.

Tabell 18. Genomsnittliga korrelationer mellan B5-dimensioner och negativt beteende av två typer (Berry, Ones, et al., 2007) .

Dimension	OD – mot företaget	ID – mot arbetskamrater
Emotionell stabilitet	-0.23	-0.24
Extraversion/Utåtvändhet	-0.09	0.02
Öppenhet	-0.04	-0.09
Vänlighet/Smidighet	-0.32	-0.46
Noggrannhet	-0.42	-0.23

Det är intressant att se att värdena är av betydande storlek, jämfört med de meta-analyser som var inriktade på arbetsresultat och OCB. Dimensionen vänlighet som sällan gett viktig information kommer här fram som definitivt viktig.

Integritetstest syftar till att predicera negativt beteende. Det finns två slag av sådana test: overta och koverta. I det första fallet handlar det om frågor som direkt tar upp brott mot normer och lagar. I det andra fallet handlar det om personlighetsdimensioner som visat sig samvariera med sådana brott mot normer, nämligen emotionell stabilitet, vänlighet/smidighet och noggrannhet. Validiteterna i förhållande till negativt beteende har i meta-analyser visat sig ligga på 0.39 (overta test) och 0.29 (koverta test) (Berry, Sackett, & Wiemann, 2007; Ones, Viswesvaran, & Schmidt, 1993).

En ”state-of-the-art” artikel ska också nämnas (Bolton, Becker, & Barber, 2010). Bolton et al. studerade kontra-produktivt (CWB) beteende i relation till B5. Deras definition av CWB är intressant:

Counterproductive work behaviors were measured with the 33-item Counterproductive Work Behavior Checklist (CWB-C; Spector et al., 2006). The measure consists of five subscales: abuse (e.g., "been nasty or rude to a client or customer"), production deviance (e.g., "purposely did your work incorrectly"), sabotage (e.g., "purposely damaged a piece of equipment or property"), theft (e.g., "stolen something belonging to your employer"), and withdrawal (e.g., "came to work late without permission"). These subscales are also classified into organizationally-directed (i.e., production deviance, sabotage, theft, and withdrawal dimensions) or interpersonally-directed behaviors (i.e., consisting of mostly the same items as the abuse dimension). Lastly, the subscales also allow for the computation of a composite CWB score. Responses are chosen on a five-point frequency scale ranging from 1 (*never*) to 5 (*everyday*).

(Sid. 538).

De olika måtten på CWB hade höga inbördes samband och slogs ihop till ett index. Detta index korrelerades med B5-dimensionerna med följande resultat:

- Emotionell stabilitet: -0.27
- Extraversion/Utåtvändhet: -0.14
- Öppenhet: 0.02
- Vänlighet/Smidighet -0.28
- Noggrannhet: -0.28

Multipel regressionsanalys visade att 13 % av variansen i CWB-index kunde förklaras av B5-dimensioner.

## Entreprenörskap

En meta-analys av Zhou och Seibert sammanställde undersökningar som jämförde B5-värden för entreprenörer och chefer i allmänhet (Zhao & Seibert, 2006) . Deras resultat gavs som effektmått (*d*-värden) som kan transformeras till korrelationer. Detta har gjorts, se Tabell 19.

Tabell 19. Genomsnittliga korrelationer mellan personlighetsdimensioner och entreprenörskap (Zhao & Seibert, 2006) .

Dimension	Punktbiserial korrelation
Emotionell stabilitet	-0.16
Extraversion/Utåtvändhet	0.10
Öppenhet	0.15
Vänlighet/Smidighet	-0.07
Noggrannhet	0.19

Storleken på sambanden är troligen något underskattad eftersom den beroende variabeln är en dikotomi. I övrigt är de jämförbara med andra resultat som presenterats här, kanske något lägre.

## Metodstudier och begreppsvalidering

### Begreppsvalidering mot Assessment Center

Assessment Center (AC) är en annorlunda ansats till personlighetsmätning, som har fått ett visst genomslag även i vårt land. Forskningen är relativt omfattande och validiteten respektabel. En intressant fråga är om B5-skalar samvarierar med AC-bedömning. Se sammanställning av en meta-analys av Collins et al. (Collins, et al., 2003) i Tabell 20.

Tabell 20. Genomsnittliga korrelationer mellan personlighetsdimensioner, intelligens och övergripande AC-bedömning (Collins, et al., 2003).

Dimension	Korrelation
Emotionell stabilitet	0.34
Extraversion/Utåtvändhet	0.47
Öppenhet	0.23
Vänlighet/Smidighet	0.16
Noggrannhet (inga data)	-
Intelligens (GMA)	0.65

Den multipla korrelationen mellan AC-bedömningen och B5+GMA var 0.81. Den är troligen en underskattning eftersom data på noggrannhet inte var tillgängliga för analysen. Kostnader för AC som citeras av Collins et al. låg på 1700 – 4000 dollar per person; data från 1986. Om AC används för urval är det diskutabelt om metodiken är kostnadseffektiv. Dessa data är emellertid också ett stöd för både B5 och AC i den meningen att de ger likartade resultat.

### FFM-test jämförda med andra test

I en meta-analys av Salgado jämfördes validiteter för test som explicit konstruerats för att mäta B5-dimensioner och andra test där vissa skalar kunde kategoriseras som tillhörande FFM (Salgado, 2003). Resultaten framgår av Tabell 21.

Tabell 21. Genomsnittliga korrelationer mellan B5-dimensioner och arbetsresultat för två typer av personlighetstest (Salgado, 2003).

Dimension	FFM-test	Övriga test
Emotionell stabilitet	0.16	0.05
Extraversion/Utåtvändhet	0.04	0.05
Öppenhet	0.05	0.05
Vänlighet/Smidighet	0.08	0.08
Noggrannhet	0.28	0.18

Det verkar alltså som om FFM-test hade något högre validitet än övriga test.

## ***Indexvariabler: exempel***

Den forskning som behandlats ovan har nästan enbart tagit upp sambanden mellan enskilda test variabler och kriterier. Bilden som man får är att sambanden oftast (men inte alltid) är svaga och att personlighetstest är av tveksamt värde. Detta är emellertid missvisande. Om man i stället analyserar sammanställningar av flera testvariabler till index kan man få en helt annan bild, som bättre beskriver det praktiska värdet av testen.

Skälet till att vi har så lite data på indexvariabler och multipla korrelationer är troligen att så stora krav ställs på data underlaget i sådana analyser. I nästa avsnitt diskuterar jag kritiken av de multipla regressionsanalyser som gjorts i meta-analyser.

Det exempel som ges här har beskrivits i detalj på annat ställe (Sjöberg, 2010b). Anställda i kundservice i ett finansbolag testades med *UPP*-testet. Deras arbetsprestationer bedömdes av cheferna i ett stort antal kriteriedimensioner. Kriterierna kunde reduceras till 3 aspekter: övergripande intryck (värde för företaget), kärnuppgifter (effektivitet) och sociala funktioner (OCB). Av testets 19 variabler togs 6 ut som speciellt användbara. Se Tabell 22.

Tabell 22. Korrelationer mellan testvariabler och tre kriterieindex.

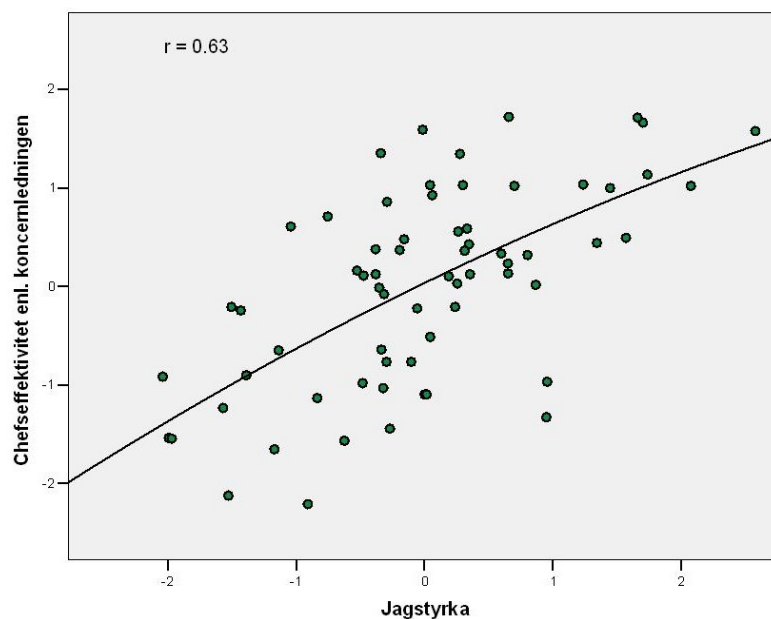
	Värde för företaget ("halo")	Effektivitet (kärnuppgifter)	Social funktion (OCB)
Samtliga testvariabler med lika vikter, perfektionism omvänt	0.34**	0.28*	0.20
Sex bästa testvariablerna, lika vikter	0.50**	0.39**	0.29*

\* $p < 0.05$

\*\*  $p < 0.01$

Det övergripande intrycket kan eventuellt likställas med "halo". Forskningen om halo är omfattande (Jackson & Furnham, 2001; Chockalingam Viswesvaran, Schmidt, & Ones, 2005). Det handlar om att cheferna bedömde att en anställd var till nytta för företaget, var en sådan person som man var angelägen om att anställa och behålla. Om ett personlighetstest kan bidra till att välja ut sådana personer är det naturligtvis mycket värdefullt för företaget; det är ingen felkälla i bedömningarna som det handlar om utan ett användbart mått på arbetsresultat (Jackson & Furnham, 2001).

Generella dimensioner av typ jagstyrka (Sjöberg, 2005) fungerar bra mot övergripande kriteriemått som chefsbedömningar av "värde för företaget". Se Figur 2.



**Figur 2. Samband mellan jagstyrka och arbetsresultat, varuhuschefer bedömda av koncernledningen.**

## Diskussion

### Överblick

Bilden av B5-validiteten mot arbetsresultat i allmänhet är entydig. Endast noggrannhet har en tydlig korrelation med kriterierna, men även den är ganska låg, omkring 0.20 – 0.25, och inte generellt giltig, t ex inte i polisarbete (Barrett, Miguel, Hurd, Lueke, & Tan, 2003; Varela, Boccaccini, Scogin, Stump, & Caputo, 2004). Dessa slutsatser bygger på mycket omfattande forskning: 100-tals publicerade undersökningar och många tusen respondenter. Bara ett fåtal analyser har publicerats av om nivån 0.20 - 0.25 är tillräcklig för att ge ett väsentligt tillskott av prognosförmågan utöver *g*-faktorn, vars stora validitet är väl känd. Det tycks vara mycket tveksamt om ett väsentligt tillskott ges. Hogans och Hollands analyser av matchande kriterier och personlighetsdimensioner gav en mera positiv bild men förutsätter givetvis att det är de matchande kriterierna som är av praktiskt intresse. Sammansatta variabler, t ex baserade på regressionsanalys, ger ofta betydligt bättre validitet än enskilda B5-skalar. Validiteter omkring 0.4 – 0.5 är inte ovanliga. Sådana anpassade och sammansatta skalor är vanliga i praktiskt arbete och därför är dessa validiteter relevanta, mera relevanta än de enskilda skalorna. Dessutom kan sägas att vissa av meta-analyserna, som gällde andra aspekter än arbetsresultat, har gett mera lovande resultat.

Ganska lite arbete har ägnats åt att försöka förstå *varför* vissa personlighetsvariabler korrelerar med arbetsresultat, och *när*. I en studie fann man emellertid att målsättningar och planering förmedlade sambandet mellan noggrannhet och arbetsresultat (Barrick, Mount, & Strauss, 1993): i en annan att autonomi i arbetssituationen medförde starkare samband mellan resultat och personlighet (Barrick & Mount, 1993).

FFM har tydligen så starkt fascinerat forskarna att de ägnat sig nästan uteslutande åt att försöka hitta validitet hos B5-dimensionerna. Det har varit ett ganska ofruktbart arbete, enligt min mening, men det har pågått i decennier och 100-tals studier har publicerats. Det är möjligt att FFM blivit så allmänt accepterat att forskarna sett chansen att få sina arbeten publicerade maximeras genom att ansluta sig till denna huvudfåra. Den oerhört omfattande forskningen om B5-dimensionerna torde inte ha haft sin grund i realistiska förhoppningar om deras validitet utan i modellens dominans inom nästan all modern personlighetspsykologi (Matthews, Deary, & Whiteman, 2003).

Ger personlighetstest något utöver begåvning? Begåvning är starkt relaterad till arbetsresultat i de allra flesta jobb (Schmidt & Hunter, 1998). I en grupp som är selegerad med avseende på begåvning finns emellertid fortfarande en stor variation, som kan vara relaterad till personlighet. Särskilt för komplexa jobb är det väl känt att variationen mellan individer i arbetsresultat är mycket stor (Hunter, Schmidt, & Judiesch, 1990). Det var denna omständighet som låg till grund för mitt personlighetsprojekt vid HHS, där jag utvecklade mina metoder för att mäta emotionell intelligens och en del andra personlighetsvariabler (Engelberg & Sjöberg, 2005; Sjöberg, 2001a, 2001b, 2008b). En viktig översikt (Roberts, Kuncel, Shiner, Caspi, & Goldberg, 2007) har visat att personlighet är nästan lika viktig som begåvning. Allt detta understryker att sökandet efter effektiva personlighetstest måste fortsätta. De övergripande B5-dimensionerna räcker inte. Smalare underskalor till B5 ger ett visst, men begränsat, tillskott.

Hur ska man väga samman personlighet och begåvning? En enkel och rättfram lösning är att använda en kompensatorisk modell, dvs. summera de två typerna av information med vissa vikter, som kan vara beräknade statistiskt mot kriterier eller ansatta a priori. Det är emellertid inte säkert att man vill acceptera en extremt intelligent chef som är personlighetsmässigt klart olämplig. Det kan man hantera genom att ställa upp miminikrav i båda avseendena och sedan tillämpa en kompensatorisk modell på de som uppfyller båda kraven. Ett närmare studium av data kan ge uppslag till andra och bättre modeller i varje enskilt fall.

Multipla korrelationer mellan B5-dimensioner har rapporterats i vissa studier och de kan, som påpekats ovan, ge uppmuntrande resultat. Svårigheten är emellertid att uppskatta interkorrelationerna mellan B5-dimensionerna utifrån meta-analyserna; kritik har riktats mot sådana analyser och man har hävdats att de uppskattade interkorrelationerna varit alltför låga, vilket gett överskattningar av de multipla korrelationerna. Morgeson et al. (Morgeson, et al., 2007a) skriver:

“The meta regressions cited in these papers rely on an unrealistically low estimate of the intercorrelations among the Big Five. The estimated population intercorrelations among the Big Five are typically taken from the unpublished dissertation of Ones and are described as ranging from .00 to .27 (Ones, Viswesvaran & Reiss, 1996) with an average intercorrelation of only .15. There is clear evidence that the Big Five factors are not as orthogonal as suggested by these low intercorrelations (Block, 1995; Funder, 2001). For example, the revised NEO Personality Inventory (Costa & McCrae, 1992) reports domain-scale intercorrelations as high as .53 in self-ratings. The intercorrelations among the five factors appear even stronger in applicant samples (McManus & Kelly, 1999).” (Sid. 6-7).

Data från användning av *UPP*-testet kastar ett visst ljus över denna frågeställning.

Korrelationer beräknades mellan B5-skalorna i testet i fyra grupper av testade personer, före och efter korrektion för skönmålning. De vägledande värdena från Ones som Morgeson et al. åberopar låg alltså i genomsnitt så lågt som 0.15. I vårt fall var värdena 0.27 och 0.22 för oskarpt resp. skarp läge, betydligt högre men inte riktigt i nivå med de värden som Morgeson et al. tar fram. Intressant nog var interkorrelationerna mellan B5-skalorna efter korrektion för skönmålning mycket låga: 0.07 och 0.09 för oskarpt resp. skarp läge. Våra data stödjer alltså

den kritik som Morgeson et al. riktar mot beräkning av multipla korrelationer på grundval av meta-analyser, och de stödjer också korrektionen för skönmålning som ger en förbättrad chans till ökad validitet i multipel regressionsanalys på grund av de mycket låga interkorrelationerna. Det kan tilläggas att det var endast få negativa korrelationer i dessa analyser, den starkaste var -0.17. De fyra korrelationsserierna var inbördes högt korrelerade. Däremot finner vi inte den väntade skillnaden mellan skarpt ("applicants") och oskarpt läge. Nivån på interkorrelationerna mellan B5-skalor tycks vara ungefär densamma i dessa två fall. Skälet till de låga värdena som Ones kommit fram till tycks vara oklart, men så mycket är tydligt att de är underskattningar, och detta betyder ju att de multipla korrelationer som uppskattats med deras hjälp är för höga. En noggrann analys av skillnaden mellan testningar i skarp och oskarpt läge tyder inte på att korrelationerna mellan testvariabler sinsemellan eller med kriterier är mycket olika (Bradley, 2007).

En senare meta-analys av Mount et al. (Mount, Barrick, Scullen, & Rounds, 2005), som bygger på 4000 testade personer, gav ett genomsnittligt värde av interkorrelationerna på 0.29. Den mest aktuella och omfattande meta-analysen av interkorrelationer publicerades 2010 (van der Linden, te Nijenhuis, & Bakker, 2010) och bygger på 212 studier och sammanlagt 144 117 testade personer (!). Den genomsnittliga korrelationen mellan B5-dimensionerna var 0.31. Våra resultat var alltså mycket lika värdena hos Mount et al.. samt van der Linden et al.

Det har ett visst intresse att kriterier för testgranskning som formulerats av EFPA<sup>10</sup> (Lindley, Bartram, & Kennedy, 2008), och som används i Sverige, bygger på medianberäkning av korrelationer mellan enskilda skalor och kriterier, inte på multipla korrelationer. Om man räknar på det sättet blir alla FFM-test – och flesta andra personlighetstest - "underkända". Medianen av kriteriekorrelationer mot arbetsprestation hamnar normalt en bra bit under 0.2. Det är emellertid meningslöst att beräkna medianer av korrelationerna mellan testskalor och kriterier. Ett test utvärderas ju normalt i ett antal skalor. Mot arbetsresultat kan en del vara valida och mot OCB andra. Hur bra testet är avgörs också av interkorrelationerna mellan testskalorna. Medianen av korrelationerna säger ingenting.

Fokuserade och "smala" testvariabler är den mest lovande utvecklingsinriktningen. Stödet för fokuserade dimensioner exemplifieras väl av forskningen om *UPP*-testet. Testet mäter B5-dimensionerna, *men framför allt dimensioner mera tydligt fokuserade på arbetsfunktioner samt kriteriemått av typ arbetsvilja och balans arbete/övrigt liv* (Sjöberg, 2010a). Dessa kriteriemått har använts som ställföreträdande kriterier (proxykriterier) (Sjöberg, 2010b) och gjort det möjligt att göra omfattande valideringar av testet. Dessa valideringar har visat

1. Mycket låg validitet för B5-dimensioner, i överensstämmelse med de meta-analyser som citerats ovan
2. Hög validitet för flera av de fokuserade dimensioner som *UPP*-testet mäter, med mycket stora tillskott utöver den validitet som B5-dimensionerna ger
3. Hög validitet hos sammansatta skalor – bäst är givetvis när komponenterna valts och viktats för deras relevans för en speciell tillämpning.

---

<sup>10</sup> European Federation of Psychologists' Associations

## Slutsats

Personlighet och arbetsresultat – i vid mening – är som vi sett ett livaktigt forskningsområde, internationellt sett. Det är bokstavligen hundratals studier som kan åberopas som stöd för de slutsatser som dragits i denna rapport. Bilden man får av alla dessa meta-analyser är tydlig: B5-dimensionerna som sådana är svagt relaterade till arbetsresultat. Det är tydligt att nya idéer behövs. Den mest lovande är användningen av fokuserade, ”smala” personlighetsdimensioner. Dessa har fått vara i skymundan under lång tid, när forskningen dominerats av de övergripande B5-dimensionerna. Det är realistiskt att vänta sig att sammansatta variabler som konstruerats för speciella syften fungerar mycket bättre än B5-dimensionerna. I praktiskt arbete är det sådana fokuserade ”framgångsprofiler” som är av intresse.

Många svaga samband mellan B5-dimensioner och arbetsresultat har publicerats. Men det är fel att på grundval av dessa resultat helt avfärda personlighetstestning i arbetslivet, på det sätt som görs av Morgeson et al. (Morgeson, et al., 2007a, 2007b). Om man använder ett personlighetstest på ett förnuftigt sätt, anpassat till behoven i varje särskilt fall, är dess effektivitet på en helt annan nivå än den som beskrivs i de grova övergripande meta-analyserna – se Tabell 22 för ett exempel; många andra kunde ges. Psykometrisk teknik utvecklad för *UPP*-testet gör det möjligt att estimeras ”lämplighet” och ”risk” för varje specificerat syfte, t ex anställning i en viss befattning, utifrån företagets egen bedömning (Sjöberg, 2010c). Validering har lyckats väl, både mot övergripande chefsbedömning och mot proxykriterier, som i sin tur ger resultat liknande dem man erhåller med externa kriterier. Personlighetstest ger viktiga tillskott till prognos i arbetslivet inte bara av arbetsresultat utan också av olycksbenägenhet, OCB, arbetstillfredsställelse, ledarskap mm. Det är en kostnadseffektiv metod att samla in viktig information som står sig mycket väl i konkurrens med andra metoder som ”djupintervjuer” och AC/simuleringsövningar.

Ones och Viswesvaran fann validiteter omkring 0.4 för ett index sammansatt av noggrannhet, emotionell stabilitet och vänlighet (Ones & Viswesvaran, 2001a, 2001b). B5-dimensionerna har en korrelationsstruktur som tyder på en övergripande gemensam faktor (van der Linden, et al., 2010), denna kan man tolka som jagstyrka men GFP (”general personality factor”) enligt van der Linden m fl. bygger på alldeles för få variabler för att kunna jämföras med jagstyrka i den mening som begreppet definieras i *UPP*-testet. Inte desto mindre skriver de:

”In Big Five terms, high GFP-individuals are described as open-minded, hard-working, sociable, friendly and emotionally stable”. (Sid. 316).

Det låter ju onekligen som en beskrivning av jagstarka personer. GFP är också positivt relaterat till emotionell intelligens (Veselka, Schermer, Petrides, Cherkas, et al., 2009). Det är intressant att se att GFP tycks ha en mycket stor genetisk komponent (Veselka, Schermer, Petrides, & Vernon, 2009), 50 % enligt Rushton et al. (Rushton, et al., 2009).

Som framgått av redogörelsen för meta-analyser av samband mellan B5-dimensioner och arbetsrelaterade variabler finns det åtskilliga aspekter där personlighetstest är av värde; de ska inte diskuteras enbart i relation till arbetsresultat. Att vara närvarande på jobbet är en aspekt, att inte byta till annan anställning är en annan. Närvaro kunde till 18 % för klaras av B5-dimensioner i en studie, som kontrollerade för bakgrundsvariabler (Judge, Martocchio, &

Thoresen, 1997). Barrick och Mount fann samband mellan avsikt att söka sig till annan anställning och emotionell stabilitet och noggrannhet (Barrick & Mount, 1996).

Forskarna tycks sedan 1990 delvis ha letat på fel ställen när de ville dokumentera personlighetsdimensionernas prognosvärde i arbetslivet. Många studier har emellertid bidragit till att en mera positiv bild successivt har växt fram. Varken den historiska pessimismen eller den moderna som bygger enbart på B5-dimensioner mot arbetsresultat är motiverade. Personlighetsdimensioner har betydelse nästan i nivå med *g*-faktorn.

## Referenser

- Allport, F. H., & Allport, G. W. (1921). Personality traits: Their classification and measurement. *Journal of Abnormal and Social Psychology*, 16, 6-40.
- Allport, G. W. (1921). Personality and character. [doi:10.1037/h0066265]. *Psychological Bulletin*, 18(9), 441-455.
- Ashton, M. C. (1998). Personality and job performance: The importance of narrow traits. [doi:10.1002/(SICI)1099-1379(199805)19:3<289::AID-JOB841>3.0.CO;2-C]. *Journal of Organizational Behavior*, 19(3), 289-303.
- Barrett, G. V., Miguel, R. F., Hurd, J. M., Lueke, S. B., & Tan, J. A. (2003). Practical Issues in the Use of Personality Tests in Police Selection. *Public Personnel Management*, 32(4), 497-517.
- Barrick, M. R., & Mount, M. K. (1991). The Big Five personality dimensions and job performance: A meta-analysis. *Personnel Psychology*, 44, 1-26.
- Barrick, M. R., & Mount, M. K. (1993). Autonomy as a moderator of the relationships between the Big Five personality dimensions and job performance. [doi:10.1037/0021-9010.78.1.111]. *Journal of Applied Psychology*, 78(1), 111-118.
- 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(3), 261-272.
- Barrick, M. R., & Mount, M. K. (2003). Impact of meta-analysis methods on understanding personality-performance relations *Validity generalization: A critical review*. (pp. 197-221): Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.
- Barrick, M. R., Mount, M. K., & Judge, T. A. (2001). Personality and performance at the beginning of the new millennium: What do we know and where do we go next? [doi:10.1111/1468-2389.00160]. *International Journal of Selection and Assessment*, 9(1-2), 9-30.
- Barrick, M. R., Mount, M. K., & Strauss, J. P. (1993). Conscientiousness and performance of sales representatives: Test of the mediating effects of goal setting. [doi:10.1037/0021-9010.78.5.715]. *Journal of Applied Psychology*, 78(5), 715-722.
- Bartram, D. (2005). The Great Eight Competencies: A Criterion-Centric Approach to Validation. [doi:10.1037/0021-9010.90.6.1185]. *Journal of Applied Psychology*, 90(6), 1185-1203.
- Bergner, S., Neubauer, A. C., & Kreuzthaler, A. (2010). Broad and narrow personality traits for predicting managerial success. [doi:10.1080/13594320902819728]. *European Journal of Work and Organizational Psychology*, 19(2), 177-199.
- Berry, C. M., Ones, D. S., & Sackett, P. R. (2007). Interpersonal deviance, organizational deviance, and their common correlates: A review and meta-analysis. [doi:10.1037/0021-9010.92.2.410]. *Journal of Applied Psychology*, 92(2), 410-424.
- Berry, C. M., Sackett, P. R., & Wiemann, S. (2007). A review of recent developments in integrity test research. [doi:10.1111/j.1744-6570.2007.00074.x]. *Personnel Psychology*, 60(2), 271-301.
- Bertua, C., Anderson, N., & Salgado, J. F. (2005). The predictive validity of cognitive ability tests: A UK meta-analysis. [doi:10.1348/096317905X26994]. *Journal of Occupational and Organizational Psychology*, 78(3), 387-409.
- Bolton, L. R., Becker, L. K., & Barber, L. K. (2010). Big Five trait predictors of differential counterproductive work behavior dimensions. [doi:10.1016/j.paid.2010.03.047]. *Personality and Individual Differences*, 49(5), 537-541.

- Borman, W. C., Penner, L. A., Allen, T. D., & Motowidlo, S. J. (2001). Personality predictors of citizenship performance. [doi:10.1111/1468-2389.00163]. *International Journal of Selection and Assessment*, 9(1-2), 52-69.
- Bowling, N. A., & Burns, G. N. (2010). A comparison of work-specific and general personality measures as predictors of work and non-work criteria. [doi:10.1016/j.paid.2010.03.009]. *Personality and Individual Differences*, 49(2), 95-101.
- Bradley, K. M. (2007). *Personality test validation research: Present-employee and job applicant samples*. US: ProQuest Information & Learning.
- Brown, A., & Bartram, D. (2009). *Development and psychometric properties of the OPQ32r. Supplement to the OPQ 32 technical manual*: SHL.
- Bäckman, C., & Carlstedt, B. (2010). A construct validation of a profession-focused personality questionnaire (PQ) versus the FFPI and the SIMP. [doi:10.1027/1015-5759/a000019]. *European Journal of Psychological Assessment*, 26(2), 136-142.
- Clarke, S., & Robertson, I. T. (2005). A meta-analytic review of the Big Five personality factors and accident involvement in occupational and non-occupational settings. [doi:10.1348/096317905X26183]. *Journal of Occupational and Organizational Psychology*, 78(3), 355-376.
- Collins, J. M., Schmidt, F. L., Sanchez-Ku, M., Thomas, L., McDaniel, M. A., & Le, H. (2003). Can basic individual differences shed light on the construct meaning of assessment center evaluations? [doi:10.1111/1468-2389.00223]. *International Journal of Selection and Assessment*, 11(1), 17-29.
- Connolly, J. J., Kavanagh, E. J., & Viswesvaran, C. (2007). The Convergent Validity between Self and Observer Ratings of Personality: A meta-analytic review. [doi:10.1111/j.1468-2389.2007.00371.x]. *International Journal of Selection and Assessment*, 15(1), 110-117.
- Crook, G. H., Burke, L. K., & Klieger, W. A. (1952). *The Army Personality Inventory. PRS Report 865*.
- Dalal, R. S. (2005). A Meta-Analysis of the Relationship Between Organizational Citizenship Behavior and Counterproductive Work Behavior. [doi:10.1037/0021-9010.90.6.1241]. *Journal of Applied Psychology*, 90(6), 1241-1255.
- Dudley, N. M., Orvis, K. A., Lebiecki, J. E., & Cortina, J. M. (2006). A meta-analytic investigation of conscientiousness in the prediction of job performance: Examining the inter-correlations and the incremental validity of narrow traits. [doi:10.1037/0021-9010.91.1.40]. *Journal of Applied Psychology*, 91(1), 40-57.
- Ellis, A., & Conrad, H. S. (1948). The validity of personality inventories in military practice. [doi:10.1037/h0056021]. *Psychological Bulletin*, 45(5), 385-426.
- Engelberg, E., & Sjöberg, L. (2005). Emotional intelligence and interpersonal skills. In R. D. Roberts & R. Schulze (Eds.), *International handbook of emotional intelligence* (pp. 289-308). Cambridge MA: Hogrefe.
- Furnham, A. (2008). *Personality and intelligence at work. Exploring and explaining individual differences at work*. Hove, East Sussex: Routledge.
- Ghiselli, E. E. (1973a). The validity of aptitude tests in personnel selection. *Personnel Psychology*, 26(4), 461-477.
- Ghiselli, E. E. (1973b). *The Validity of Occupational Aptitude Tests*. New York: Wiley.
- Ghiselli, E. E., & Barthol, R. P. (1953). The validity of personality inventories in the selection of employees. [doi:10.1037/h0059438]. *Journal of Applied Psychology*, 37(1), 18-20.
- Gibby, R. E., & Zickar, M. J. (2008). A history of the early days of personality testing in American industry: An obsession with adjustment. [doi:10.1037/a0013041]. *History of Psychology*, 11(3), 164-184.

- Giles St. J. Burch, & Anderson, N. (2008). Personality as a Predictor of Work-Related behavior and Performance: Recent Advances and Directions for Future Research. *International Review of Industrial and Organizational Psychology* 23, 261-305.
- Goleman, D. (1995). *Emotional intelligence*. New York: Bantam Books.
- Goleman, D. (1998). *Working with emotional intelligence*. New York: Bantam Books.
- Guion, R. M., & Gottier, R. F. (1965). Validity of personality measures in personnel selection. *Personnel Psychology*, 18, 135-164.
- Gulliksen, H. (1950). *Theory of mental tests*. New York: Wiley.
- Hogan, J., Hogan, R., & Kaiser, R. B. (2011). Management derailment *APA handbook of industrial and organizational psychology, Vol 3: Maintaining, expanding, and contracting the organization*. (pp. 555-575): Washington, DC, US: American Psychological Association.
- Hogan, J., & Holland, B. (2003). Using theory to evaluate personality and job-performance relations: A socioanalytic perspective. *Journal of Applied Psychology*, 88(1), 100-112.
- Hull, C. L. (1928). *Aptitude testing*. London: Harrap.
- Hunter, J. E., Schmidt, F. L., & Judiesch, M. K. (1990). Individual differences in output variability as a function of job complexity. *Journal of Applied Psychology*, 75, 28-42.
- Hurtz, G. M., & Donovan, J. J. (2000). Personality and job performance: The Big Five revisited. [doi:10.1037/0021-9010.85.6.869]. *Journal of Applied Psychology*, 85(6), 869-879.
- Jackson, C. J., & Furnham, A. (2001). Appraisal ratings, halo, and selection: A study using sales staff. [doi:10.1027//1015-5759.17.1.17]. *European Journal of Psychological Assessment*, 17(1), 17-24.
- Joseph, D. L., & Newman, D. A. (2010). Emotional intelligence: An integrative meta-analysis and cascading model. [doi:10.1037/a0017286]. *Journal of Applied Psychology*, 95(1), 54-78.
- Judge, T. A., Bono, J. E., Ilies, R., & Gerhardt, M. W. (2002). Personality and leadership: A qualitative and quantitative review. *Journal of Applied Psychology*, 87(4), 765-780.
- Judge, T. A., Heller, D., & Mount, M. K. (2002). Five-factor model of personality and job satisfaction: A meta-analysis. *Journal of Applied Psychology*, 87(3), 530-541.
- Judge, T. A., Martocchio, J. J., & Thoresen, C. J. (1997). Five-factor model of personality and employee absence. [doi:10.1037/0021-9010.82.5.745]. *Journal of Applied Psychology*, 82(5), 745-755.
- Judge, T. A., Thoresen, C. J., Bono, J. E., & Patton, G. K. (2001). The job satisfaction-job performance relationship: A qualitative and quantitative review. *Psychological Bulletin*, 127(3), 376-407.
- Lindley, P., Bartram, D., & Kennedy, N. (2008). *EFPA review model for the description and evaluation of psychological tests. Test review form and notes for reviewers. Version 3.42*: European Federation of Psychological Associations.
- Lounsbury, J. W., Sundstrom, E., Loveland, J. L., & Gibson, L. W. (2002). Broad versus narrow personality traits in predicting academic performance of adolescents. [doi:10.1016/j.lindif.2003.08.001]. *Learning and Individual Differences*, 14(1), 65-75.
- Matthews, G., Deary, I. J., & Whiteman, M. C. (2003). *Personality traits, 2nd ed*. Cambridge: Cambridge University Press.
- 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.
- Mischel, W. (1968). *Personality and assessment*. New York: Wiley.

- Morgeson, F. P., Campion, M. A., Dipboye, R. L., Hollenbeck, J. R., Murphy, K., & Schmitt, N. (2007a). Are we getting fooled again? Coming to terms with limitations in the use of personality tests for personnel selection. *Personnel Psychology*, *60*(4), 1029-1049.
- Morgeson, F. P., Campion, M. A., Dipboye, R. L., Hollenbeck, J. R., Murphy, K., & Schmitt, N. (2007b). Reconsidering the use of personality tests in personnel selection contexts. *Personnel Psychology*, *60*(3), 683-729.
- Mount, M. K., Barrick, M. R., Scullen, S. M., & Rounds, J. (2005). Higher-order dimensions of the big five personality traits and the big six vocational interest types. *Personnel Psychology*, *58*(2), 447-478.
- O'Connor, M. C., & Paunonen, S. V. (2007). Big Five personality predictors of post-secondary academic performance. [doi:10.1016/j.paid.2007.03.017]. *Personality and Individual Differences*, *43*(5), 971-990.
- Ones, D. S., & Viswesvaran, C. (2001a). Integrity tests and other criterion-focused occupational personality scales (COPS) used in personnel selection. [doi:10.1111/1468-2389.00161]. *International Journal of Selection and Assessment*, *9*(1-2), 31-39.
- Ones, D. S., & Viswesvaran, C. (2001b). Personality at work: Criterion-focused occupational personality scales used in personnel selection *Personality psychology in the workplace*. (pp. 63-92): Washington, DC, US: American Psychological Association.
- Ones, D. S., Viswesvaran, C., & Schmidt, F. L. (1993). Comprehensive meta-analysis of integrity test validities: findings and implications for personnel selection and theories of job performance. *Journal of Applied Psychology Monograph*, *78*, 679-703.
- Paunonen, S. V., Haddock, G., Forsterling, F., & Keinonen, M. (2003). Broad versus narrow personality measures and the prediction of behaviour across cultures. *European Journal of Personality*, *17*(6), 413-433.
- Roberts, B. W., Kuncel, N. R., Shiner, R., Caspi, A., & Goldberg, L. R. (2007). The power of personality: The comparative validity of personality traits, socioeconomic status, and cognitive ability for predicting important life outcomes. *Perspectives on Psychological Science*, *2*, 313-345.
- Rushton, J. P., Bons, T. A., Ando, J., Hur, Y.-M., Irwing, P., Vernon, P. A., et al. (2009). A general factor of personality from multitrait-multimethod data and cross-national twins. [doi:10.1375/twin.12.4.356]. *Twin Research and Human Genetics*, *12*(4), 356-365.
- Salgado, J. F. (1997). The five factor model of personality and job performance in the European Community. [doi:10.1037/0021-9010.82.1.30]. *Journal of Applied Psychology*, *82*(1), 30-43.
- Salgado, J. F. (2003). Predicting job performance using FFM and non-FFM personality measures. [doi:10.1348/096317903769647201]. *Journal of Occupational and Organizational Psychology*, *76*(3), 323-346.
- 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.
- Schmidt, F. L., Oh, I.-S., & Le, H. (2006). Increasing the accuracy of corrections for range restriction: Implications for selection procedure validities and other research results. *Personnel Psychology*, *59*(2), 281-305.
- Schmidt, F. L., Shaffer, J. A., & Oh, I.-S. (2008). Increased accuracy for range restriction corrections: Implications for the role of personality and general mental ability in job and training performance. [doi:10.1111/j.1744-6570.2008.00132.x]. *Personnel Psychology*, *61*(4), 827-868.

- Schmitt, N., Gooding, R. Z., Noe, R. A., & Kirsch, M. (1984). Metaanalyses of validity studies published between 1964 and 1982 and the investigation of study characteristics. [doi:10.1111/j.1744-6570.1984.tb00519.x]. *Personnel Psychology*, 37(3), 407-422.
- SHL. (1999-2007). *OPQ32. SHL sales report. Technical manual*: SHL Group Ltd.
- Sjöberg, L. (1980). Similarity and correlation. In E.-D. Lantermann & H. Feger (Eds.), *Similarity and choice* (pp. 70-87). Bern: Huber.
- Sjöberg, L. (1981). Värdet av DMT vid urval av flygförare. *Nordisk Psykologi*, 33, 241-248.
- Sjöberg, L. (2001a). *Emotional intelligence measured in a highly competitive testing situation* (SSE/EFI Working Paper Series in Business Administration No. 2001:13). Stockholm: Stockholm School of Economics.
- Sjöberg, L. (2001b). Emotional intelligence: A psychometric analysis. *European Psychologist*, 6, 79-95.
- Sjöberg, L. (2005). *Ledarskap och personlighet: Testning av jagstyrka. (Leadership and personality: Testing ego-strength)* (SSE/EFI Working Paper Series in Business Administration No. 2005:7). Stockholm: Stockholm School of Economics.
- Sjöberg, L. (2008a). *Bortom Big Five: Konstruktion och validering av ett personlighetstest. (Beyond Big Five: Construction and validation of a personality test)* (SSE/EFI Working Paper Series in Business Administration No. 2008:7). Stockholm: Stockholm School of Economics.
- Sjöberg, L. (2008b). Emotional intelligence and life adjustment. In J. C. Cassady & M. A. Eissa (Eds.), *Emotional Intelligence: Perspectives on Educational & Positive Psychology* (pp. 169-183). New York: Peter Lang Publishing.
- Sjöberg, L. (2009). *UPP-testet: Korrektion för skönmålning. (The UPP test: Correction for impression management)*. *Forskningsrapport 2009:3*. Stockholm: Psykologisk Metod AB.
- Sjöberg, L. (2010a). *A third generation personality test* (SSE/EFI Working Paper Series in Business Administration No. 2010:3). Stockholm: Stockholm School of Economics.
- Sjöberg, L. (2010b). *UPP-testet och kundservice: Kriteriestudie. Forskningsrapport 2010:6*. Stockholm: Psykologisk Metod AB.
- Sjöberg, L. (2010c). *UPP-testet. Teknisk manual*. Stockholm: Psykologisk Metod AB.
- 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.
- 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.
- Stogdill, R. M. (1948). Personal factors associated with leadership; a survey of the literature. *Journal of Psychology: Interdisciplinary and Applied*, 25, 35-71.
- Super, D. E. (1942). The Bernreuter Personality Inventory: a review of research. [doi:10.1037/h0058418]. *Psychological Bulletin*, 39(2), 94-125.
- Taylor, H. C., & Russell, J. T. (1939). The relationship of validity coefficients to the practical effectiveness of tests in selection: Discussion and tables. *Journal of Applied Psychology*, 23, 565-578.
- Tett, R. P., Jackson, D. N., & Rothstein, M. (1991). Personality measures as predictors of job performance: A meta-analytic review. *Personnel Psychology*, 44(4), 703-742.
- Thorndike, R. L. (1949). *Personnel selection; test and measurement techniques*: Oxford, England: Wiley.
- Thornton Iii, G. C., & Gibbons, A. M. (2009). Validity of assessment centers for personnel selection. *Human Resource Management Review*, 19(3), 169-187.

- Trapmann, S., Hell, B., Hirn, J.-O. W., & Schuler, H. (2007). Meta-analysis of the relationship between the Big Five and academic success at university. [doi:10.1027/0044-3409.215.2.132]. *Zeitschrift für Psychologie/Journal of Psychology*, 215(2), 132-151.
- van der Linden, D., te Nijenhuis, J., & Bakker, A. B. (2010). The General Factor of Personality: A meta-analysis of Big Five intercorrelations and a criterion-related validity study. [doi:10.1016/j.jrp.2010.03.003]. *Journal of Research in Personality*, 44(3), 315-327.
- Varela, J. G., Boccaccini, M. T., Scogin, F., Stump, J., & Caputo, A. (2004). Personality Testing in Law Enforcement Employment Settings: A Meta-Analytic Review. [doi:10.1177/0093854804268746]. *Criminal Justice and Behavior*, 31(6), 649-675.
- Veselka, L., Schermer, J. A., Petrides, K. V., Cherkas, L. F., Spector, T. D., & Vernon, P. A. (2009). General factor of personality: Evidence from the HEXACO model and a measure of trait emotional intelligence. [doi:10.1375/twin.12.5.420]. *Twin Research and Human Genetics*, 12(5), 420-424.
- Veselka, L., Schermer, J. A., Petrides, K. V., & Vernon, P. A. (2009). Evidence for a heritable general factor of personality in two studies. [doi:10.1375/twin.12.3.254]. *Twin Research and Human Genetics*, 12(3), 254-260.
- Whetzel, D. L., McDaniel, M. A., Yost, A. P., & Kim, N. (2010). Linearity of Personality-Performance Relationships: A large-scale examination. *International Journal of Selection and Assessment*, 18(3), 310-320.
- Vinchur, A. J., Schippmann, J. S., Switzer, F. S., III, & Roth, P. L. (1998). A meta-analytic review of predictors of job performance for salespeople. [doi:10.1037/0021-9010.83.4.586]. *Journal of Applied Psychology*, 83(4), 586-597.
- Winter, D. G. (1998). "Toward a science of personality psychology": David McClelland's development of empirically derived TAT measures. [doi:10.1037/1093-4510.1.2.130]. *History of Psychology*, 1(2), 130-153.
- Viswesvaran, C., Ones, D. S., & Schmidt, F. L. (1996). Comparative analysis of the reliability of job performance ratings. *Journal of Applied Psychology*, 81(5), 557-574.
- Viswesvaran, C., Schmidt, F. L., & Ones, D. S. (2005). Is There a General Factor in Ratings of Job Performance? A Meta-Analytic Framework for Disentangling Substantive and Error Influences. [doi:10.1037/0021-9010.90.1.108]. *Journal of Applied Psychology*, 90(1), 108-131.
- Zaccaro, S. J. (2007). Trait-Based Perspectives of Leadership. *American Psychologist*, 62(1), 6-16.
- Zhang, X., & Bartol, K. M. (2010). The influence of creative process engagement on employee creative performance and overall job performance: A curvilinear assessment. [doi:10.1037/a0020173]. *Journal of Applied Psychology*, No Pagination Specified.
- Zhao, H., & Seibert, S. E. (2006). The Big Five personality dimensions and entrepreneurial status: A meta-analytical review. [doi:10.1037/0021-9010.91.2.259]. *Journal of Applied Psychology*, 91(2), 259-271.