



**Personlighet, värderingar, skönmålning:
Utveckling av UPP-testet (*UPPTM*)**

Lennart Sjöberg

Rapport 2014:1

5 maj 2014

UPP-testet är godkänt av STP (Stiftelsen för Tillämpad Psykologi, Sveriges Psykologförbund) enligt de europeiska EFPA-kriterierna

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: Korrektur 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 korrektur 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). Personlighetsdimensioners validitet i arbetslivet: teorier och empiri. Rapport 2010:8

- Nilheim, H. (2010). Internetplattformen för UPP-testet. Rapport 2010:9.
- Sjöberg, L. (2010). Prognos av riksdagsvalet 2010. Rapport 2010: 10.
- Sjöberg, L. (2010). Skönmålning på UPP hos chefskandidater. Rapport 2010: 11.
- Sjöberg, L. (2011). Ökad testvaliditet genom korrektion för skönmålning. Rapport 2011:1.
- Sjöberg, L. (2011). UPP och UPP/Screen i relation till motivation, personlighet och framgång. Rapport 2011:2.
- Sjöberg, L. (2012). Skönmålning på ett personlighetstest bland sökande och antagna till bildning. Rapport 2012:1.
- Sjöberg, L. (2012). Begåvningsstest vid urval av sökande till Handelshögskolan i Stockholm. Rapport 2012:2
- Sjöberg, L. (2014). Personlighet, värderingar, skönmålning: Utveckling av UPP-testet. Rapport 2014:1.

Sammanfattning	4
Inledning.....	4
Metodutveckling.....	5
Handlingskraft och beslutsamhet	5
Datainsamling.....	5
Resultat.....	5
Handlingskraft.....	5
Beslutsamhet	6
Handlingskraft - beslutsamhet.....	7
Skönmålning.....	8
Diskussion	9
Personlighetens mörka sida.....	9
Inledning.....	9
Datainsamling.....	9
Narcissism	9
Machiavellianism	10
Psykopati	10
Skönmålning.....	11
Slutsats och diskussion.....	11
Skönmålning: moralstyrka, inställsamhet eller "Need for Approval"?	11
Inledning.....	11
Datainsamling.....	13
Resultat.....	13
Moralstyrka	13
Inställsamhet.....	14
Behov att bli uppskattad ("Need for approval").....	14
Övergripande analys.....	14
Diskussion	15
Värderingar.....	17
Inledning.....	17
Datainsamling.....	19
Resultat.....	19
Skönmålning.....	20
Diskussion	20
Personlighet och värderingar.....	21
Validitet: Proxyvalidering	21
Slutsatser	22
Referenser.....	23
Appendix 1. Korrigering för skönmålning.....	27
Appendix 2. Begreppsförklaringar: nya skalor	32
Appendix 3. Proxyvalidering	33
Resultat av proxyvalidering	33

Sammanfattning

Rapporten redogör för utveckling av nya skalor för UPP-testet samt en analys av skönmålning. De nya skalorna valdes bland förslag som kommit fram under den omfattande tillämpningen av *UPPTM*. De var

- beslutsamhet
- handlingskraft
- Machiavellianism
- psykopati (subklinisk, lindrig)
- moralstyrka
- inställsamhet
- värderingar

Upplägget var genomgående att arbeta med nykonstruerade skalor, studera deras reliabilitet samt begreppsvaliditet i förhållande till etablerade standardmätningar hämtade från den internationella forskningen. Alla psykometriska resultat var mycket goda. De skalor som konstruerats för *UPPTM* kan därför förväntas även ha validitet gentemot externa kriterier. Vad gäller värderingar tog vi fram metodik för att mäta 4 olika grundläggande skalor. Dessa var

- hedonism, njutning
- att få respekt
- karriär, makt pengar
- miljö och medkänsla

Begreppsvalideringen här gällde främst sambandet mellan hedonism och psykopati, som var mycket starkt.

Sambanden med skönmålning rapporteras också, de var i vissa fall ganska starka.

Ett proxykriterium bildades genom att kombinera arbetstillfredsställelse, arbetsvilja, förändringsvilja och arbetsintresse. Validitet gentemot ett sådant kriterium har visat sig nära relaterat till externa kriterier av typ arbetsprestation. Vissa av de nya skalorna visade sig ha mycket hög proxyvaliditet, främst handlingskraft, beslutsamhet och värdering av karriär. *UPPTM* kan nu kompletteras med en eller flera av dessa nya skalor.

Inledning

Denna rapport tar upp innehåll av typ "unfinished business", alltså frågeställningar som har dykt under arbetet med att utveckla och tillämpa testet *UPPTM* under en följd av år. Trots att testet nu tillämpats ca 15 000 gånger kommer det ofta upp nya synpunkter som verkar värda att arbeta vidare med, och det är vår målsättning att göra så, och att göra resultaten öppet tillgängliga. Testbranschen styrs ju i hög grad av de utländska leverantörerna som mera sällan sysslar med forskning och utveckling och tydligen inte vill sprida sina forskningsresultat.

Metodutveckling

I denna rapport redovisas utvecklingen av ett antal nya skalor som i många sammanhang i vårt arbete har varit efterfrågade. Den allmänna uppläggningsen har varit att dels konstruera helt nya item, dels begreppsvalidera dessa genom att studera sambandet mellan dem och befintliga, väl kända skalor i den internationella forskningslitteraturen, som vi för detta syfte översatt till svenska.

Alla skalor relateras till skönmålning som är ett stort problem vid all testanvändning i skarpt läge (men även annars, även om problematiken då är mindre akut).

Handlingskraft och beslutsamhet

Handlingskraft och beslutsamhet efterfrågas hos chefer och ledare i många sammanhang, inte minst det militära men naturligtvis inte endast där. De flesta har nog sett de förödande effekterna av låga värden i dessa dimensioner också i det civila arbetslivet och även i vardagslivet. Skalor för att mäta dem bedömdes därför vara av intresse inför en revision och utvidgning av *UPP*-testet.

Jag använde samma strategi här som i andra sammanhang. En etablerad och validerad standardskala hämtades ur den vetenskapliga litteraturen, översattes till svenska och dess reliabilitet i denna nya version kontrollerades. Därefter relaterades den till den nya *UPP*-skalan på samma dimension. På detta sätt erhålls ett mått på begreppsvaliditet. Om denna visar sig vara hög (0.8 eller högre) kan vi ganska lugnt utgå från att *UPP*-skalan mäter samma begrepp som den ursprungliga skalan. Det är viktigt att vara klar över att de fullständiga ursprungliga skalorna var rikt validerade även mot många externa kriterier.

Datainsamling

De skalor som presenteras i denna forskningsrapport bygger på data som samlades in på Internet. 143¹ personer deltog, varav 58 % var kvinnor. Åldern varierade mellan 18 och 74 år, median 40 år. Drygt hälften hade högskoleutbildning.

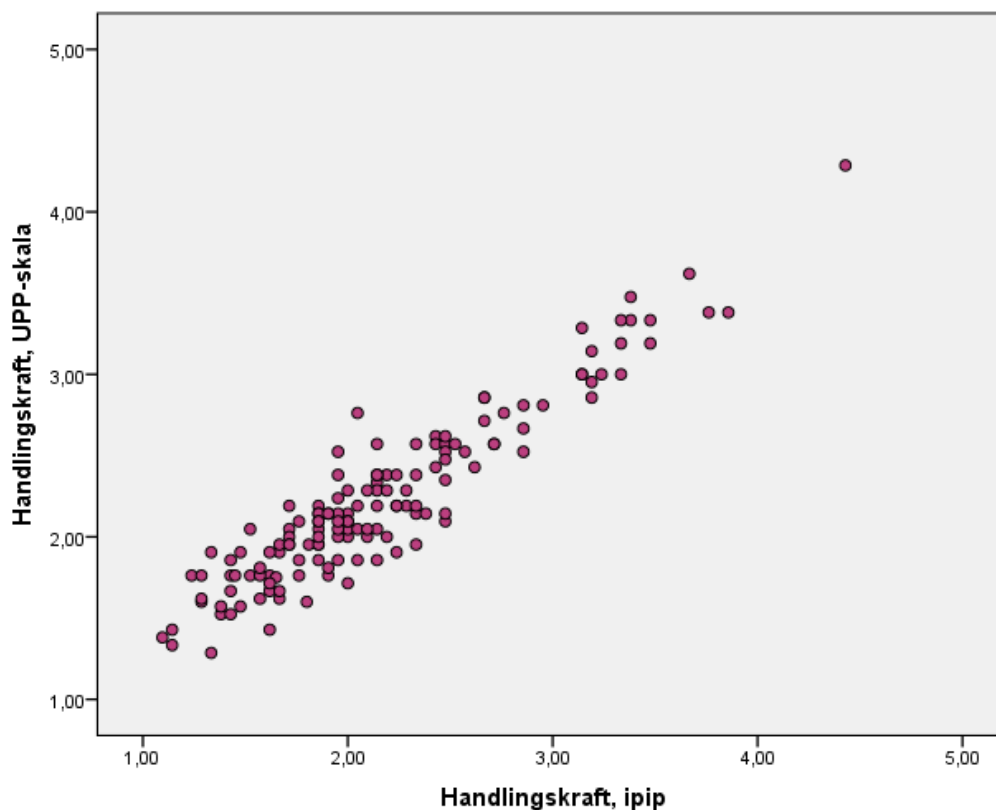
Resultat

Handlingskraft

Valideringsskalan för denna analys hämtades från den internationella sajten www.ipip.org. Här finns väl etablerade skalor för många begrepp inom personlighetspsykologin. För item hämtade från *ipip* var $\alpha = 0.92$, efter vändningar av vissa item. De 21 item som skrevs för *UPP*-testet för att mäta handlingskraft genomgick en komponentanalys. Den första faktorn svarade för 33 % av variansen. För fortsatta analyser vändes 4 av dessa item. Reliabiliteten blev 0.89.

¹ Åtta personer deltog inte i denna fas av undersökningen, som tidsmässigt låg ett par veckor senare.

Två index kunde alltså beräknas: ett för de nya item som konstruerats för *UPP*-testet, och ett för *ipip*-item som avsågs bli använda för begreppsvalidering. Dessa två index korrelerade 0.94. Se Fig. 1.



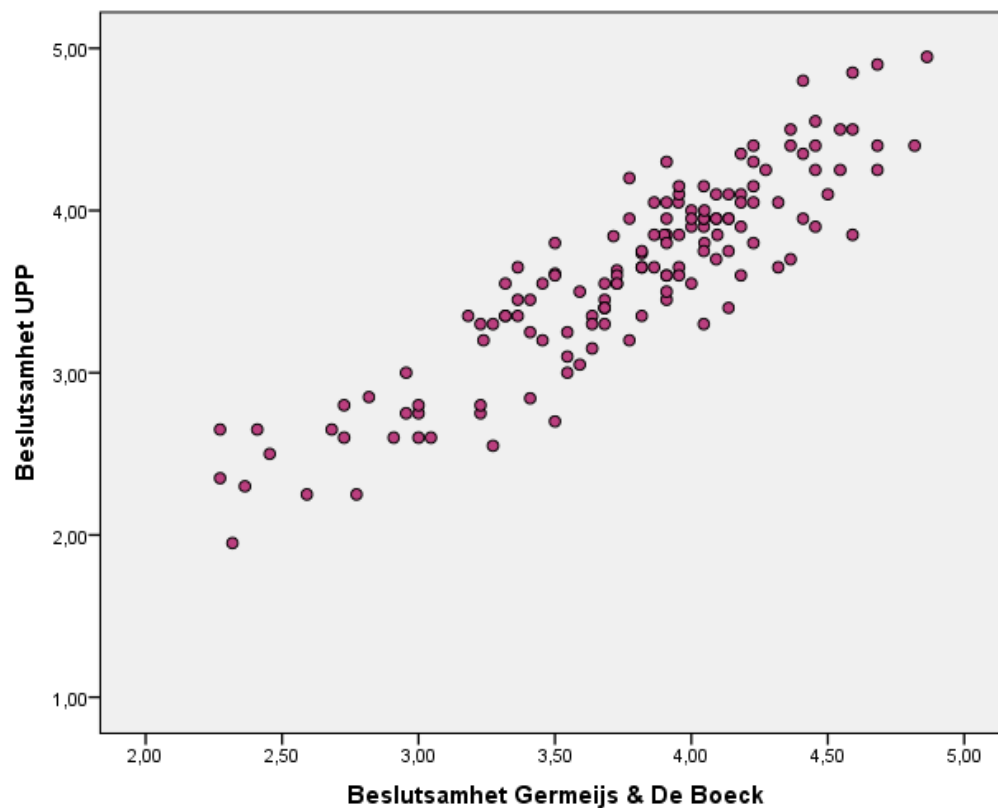
Figur 1. Sambandet mellan två skalor på handlingskraft.

Slutsatsen är att det lyckades mycket väl att konstruera en nya skala på handlingskraft.

Beslutsamhet

För denna skalas begreppsvalidering användes en skala på obeslutsamhet av Germeijs och De Boeck (Germeijs & De Boeck, 2002) . Den innehåller 22 item, vilka översattes till svenska. Denna svenska version hade en reliabilitet på 0.92. Germeijs et al. har visat att obeslutsamhet tillför predicerbarhet till externa kriterier utöver Big Five (Germeijs & Verschueren, 2011). Detta är ett av många exempel på att Big Five kan och bör kompletteras som mått på viktiga personlighetsdimensioner; Big Five är alltför allmänt.

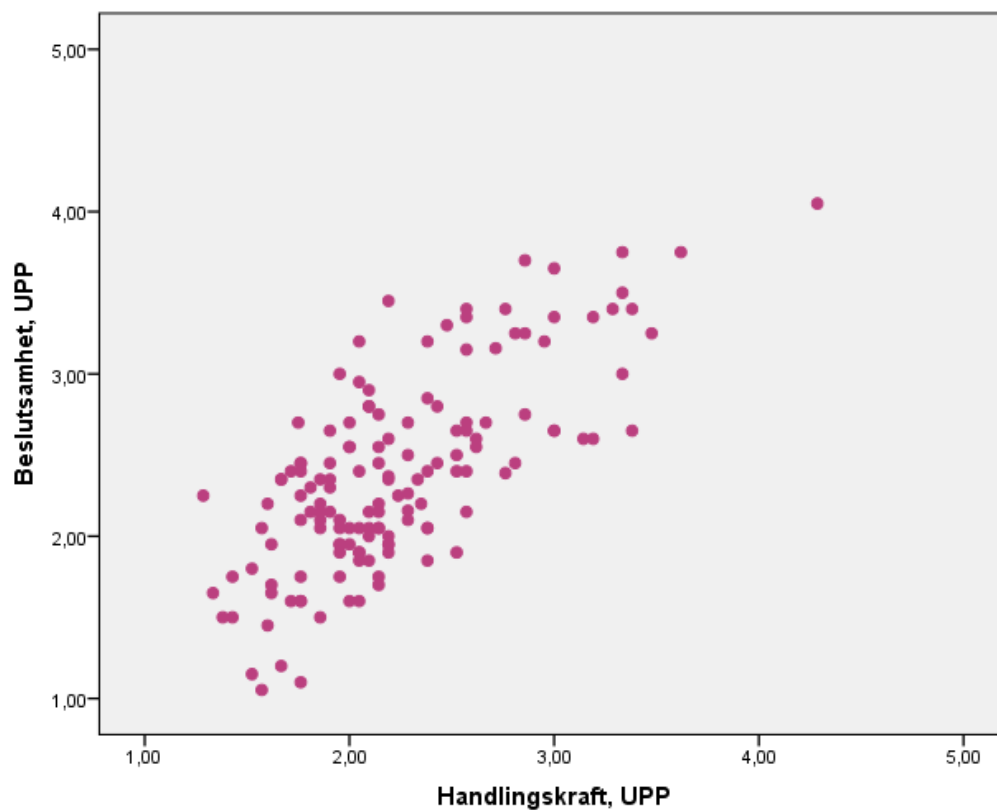
En egen skala för mätning av obeslutsamhet konstruerades, med likartad inriktning, 20 item, reliabilitet = 0.90. I båda skalorna vändes item så att de skulle mäta beslutsamhet, inte motsatsen. Sambandet mellan de två skalorna illustreras i Fig. 2. Korrelationen var = 0.90.



Figur 2. Sambandet mellan två skalor på beslutsamhet.

Handlingskraft - beslutsamhet

De två aktionsinriktade skalorna handlingskraft och beslutsamhet kan förväntas vara relaterade. Så var också fallet, $r = 0.73$. Se Fig. 3. Det är intressant att sambandet ser ut att vara något krökt.



Figur 3. Sambandet mellan handlingskraft och beslutsamhet.

Skönmålning

Sambandet mellan de två nya *UPP*-skalorna och skönmålning framgår av Tabell 1. Skönmålning mättes med två skolor: overt och kovert. Deras innebörd beskrivs i Appendix 1.

Tabell 1. Korrelationer mellan skolor på beslutsamhet och handlingskraft och skolor som mäter skönmålning.

	Skolor som mäter skönmålning	
Ny skala	Overt	Kovert
Beslutsamhet	0.38	0.57
Handlingskraft	0.45	0.55

Det är intressant att se att skalorna är relaterade till skönmålning i varierande grad, och ganska starkt. En korrektion är uppenbarligen påkallad i praktisk användning av dem.

Diskussion

De två nya UPP-skalorna visade sig ha en mycket hög grad av begreppsvaliditet, omkring 0.9. Det är därför ytterst troligt att de mäter samma sak som de kriterieskalor som togs fram ur den internationella forskningen. Sambanden med skönmålning var normala för denna typ av skalor, och i praktisk användning bör de korrigeras. Skalornas innebörd beskrivs i mera detalj i Appendix 2.

Personlighetens mörka sida

Inledning

De flesta personlighetstesten inom arbetspsykologin mäter normalpersonlighet och är inte inriktade på starka avvikelser åt det ena eller andra hållet. Detta har fått många konsulter att söka sig till test som kan ge en kompletterande bild av negativa aspekter på personligheten, som t ex psykopati. Det är emellertid bara ett fåtal test som arbetar med denna typ av dimensioner inom arbetspsykologin.² Det mest kända exemplet är HDS, som emellertid inte validerats särskilt omfattande (Fico, Brady, & Hogan, 2008).

En aktuell sammanställning av forskningen (O'Boyle Jr, Forsyth, Banks, & McDaniel, 2011) tar upp de tre viktigaste aspekterna på den mörka sidan, nämligen

- psykopati
- Machiavellianism
- narcissism

I denna forskningsrapport redogör jag för vårt utveckling av skalor som mäter dessa aspekter. Narcissism finns sedan länge inom UPP-testet, men de två övriga är nykonstruerade.

Datainsamling

De skalor som presenteras i denna forskningsrapport bygger på data som samlades in via Internet. 151 personer deltog, varav 58 % var kvinnor. Åldern varierade mellan 18 och 74 år, median 40 år. Drygt hälften hade högskoleutbildning.

Narcissism

Skalan mäter förekomst av stark och mycket positiv självuppskattning samt självupptagenhet (narcissism). Detta är en dimension som på extrem nivå kan leda till stora problem, inte minst om den finns hos chefer (Board & Fritzon, 2005; McFarlin & Sweeney, 2010; Weiss, 1950)

² I detta läge tycks det vara ganska vanligt att man använder kliniska test, t ex projektiva sådana. Hur bra dessa än är för sina ursprungliga kliniska syften, faller de till föga i en normalpopulation. Kriterierelaterad validitet kan ytterst sällan påvisas, och de har en tendens att "patologisera", dvs överdriva förekomsten av psykiska störningar (Wood, Nezowski, Lilienfeld, & Garb, 2003). En annan ansats är grafologi som inte visat sig ha validitet (Neter & Ben-Shakhar, 1989).

Dimensionen mäts i *UPP* med 8 testfrågor, $\alpha = 0.79$. Den har begreppsvaliderats mot NPI-16 (Ames, Rose, & Anderson, 2006), som är validerat ett mått på narcissism. Reliabiliteten hos den översatta versionen av NPI-16 var 0.64. Korrelationen mellan de två måtten var = 0.63, korrigerat för mätfel = 0.80. *UPP*-testets skala är alltså reliabel och begreppsvaliderad. En person med högt värde i skalan har en mycket hög uppfattning om den egna personen. *UPP*-testets mått på självförtroende korrelerade i denna studie 0.48 med självuppskattning, ett ganska högt värde men samtidigt stöd för att behandla de två dimensionerna som separata. I praktiskt arbete kan extremt höga, eller låga, värden i självuppskattning vara intressanta att följa upp.

Machiavellianism

Sedan ganska lång tid har denna dimension varit av intresse för psykologer (Christie & Geis, 1970). Dimensionen mäter en cynisk och manipulativ attityd i kombination med extrem egoism, därtill inspirerad av Machiavellis klassiska bok "Fursten" (från 1500-talet). Den skala som konstruerades av Christie och Geis är lång och numera delvis föråldrad; en bättre och modernare version med inriktning på företag och organisationer har publicerats av Dahling et al. (Dahling, Whitaker, & Levy, 2009). För min studie av denna skala översatte jag den amerikanska och konstruerade en lika lång (16 item) ny skala på svenska med likartad inriktning. Den översatta skalan hade hög reliabilitet (0.81) vilket även den svenska hade (0.90). De två skalorna korrelerade 0.90. Alltså är den nykonstruerade svenska skalan begreppsvaliderad på hög nivå.

Psykopati

Denna skala mäter s k subklinisk psykopati, alltså en mindre utpräglad form av personlighetsstörning än vad man vanligen menar med psykopati (Babiak & Hare, 2006). För att konstruera och validera en svensk skala översattes den skala som publicerats av Levenson et al. (Levenson, Kiehl, & Fitzpatrick, 1995). Datainsamlingen var densamma som för Machiavellianism.

Levensons skala i svensk översättning hade en reliabilitet på 0.83, den nykonstruerade svenska skalan (18 item) 0.72. De två skalorna på psykopati korrelerade 0.87. Alltså hade den svenska skalan hög begreppsvaliditet.

Skönmålning

De tre skalorna som mäter personlighetens mörka sida relaterades till skönmålning, se Tabell 2.

Tabell 2. Korrelationer mellan skönmålning och de två nya skalorna inom personlighetens mörka sida.

	Psykopati	Machiavellianism
Kovert skala på skönmålning	-0.62**	-0.47**
Overt skala på skönmålning	-0.63**	-0.52**

I praktisk tillämpning korrigeras dessa skalor på samma sätt som övriga för skönmålning. Att korrelationerna är negativa betyder att värdena uppjusteras i den mån som en testad person skönmålat.

Slutsats och diskussion

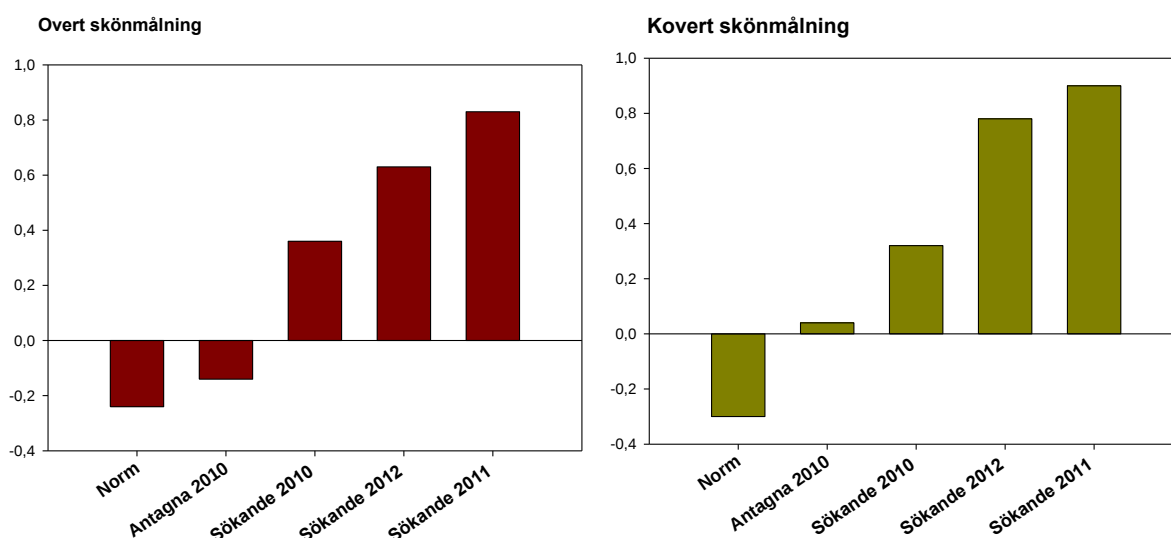
Det lyckades bra att konstruera och begreppsvalidera nya skalor för att mäta psykopati och Machiavellianism. I fortsatt arbete kommer dessa skalor att studeras mera ingående. De läggs till i UPP-testet version 2.3 med eventuella specialkonstruerade undervarianter, med normdata insamlade på sätt som beskrivs ovan. Visserligen är dessa data insamlade i oskarpt läge men i och med korrelationen för skönmålning är de ändå relevanta, och stickprovet är tillräckligt stort (Tett et al., 2009).

Skönmålning: moralstyrka, inställsamhet eller "Need for Approval"?

Inledning

Skönmålning är ett stort problem vid alla former av självrapporttest. I UPP-testet hanteras det med hjälp av två särskilda skalor, overt och kovert mätning. Den overta är så som grunden bygger på den klassiska skalan av Crowne och Marlowe (Crowne & Marlowe, 1960) som mäter social önskvärdhet ("social desirability"). Den innehåller item som handlar om att man förnekar att man utfört vanliga men ej önskvärda handlingar, eller påstår att man utför ovanliga men önskvärda handlingar. UPP-testets overta skala är uppbyggd enligt denna filosofi men inte en direkt översättning. Den korrelerar högt med en översatt version av Crowne-Marlowe. Den koverta skalan i UPP är uppbyggd av "normala" item från självrapportskalor, som valdes för att de var särskilt högt korrelerade med Crowne-Marlowe. Den skalan var som väntat högt korrelerad med den overta skalan. För en närmare beskrivning av vår metodik för att korrigera för skönmålning, se Appendix 1.

Denna rapport behandlar det psykologiska innehållet i den overta skalan. Är den ett uttryck för att man försöker svara "taktiskt" och därigenom beroende av vad som är viktigt för stunden, eller är den ett uttryck för en mera generell personlighetsegenskap, och i så fall vilket? Vår tidigare forskning har visat att det finns starka tecken som tyder på att det handlar om taktiska svar som är viktiga för stunden (Sjöberg & Wolgers, 2012), och därmed åtminstone inte enbart en fråga om "personlighet". I en studie av sökande till officersutbildning analyserades 5 olika grupper som skilde sig åt med avseende på hur "skarp" situationen var, dvs. om de testade personerna trodde att testresultatet skulle påverka deras chanser att bli antagna i olika hög grad, eller definitivt inte alls - den senare gruppen bestod av testade personer som redan var antagna. Även data från en normgrupp som tagit testet i helt oskarpt läge ingick i studien. Resultaten framgår av Fig .4.



Figur 4. Overt och kovert skönmålning i fem grupper av testade. X-axeln: ökande grad av skarpt läge vid testning, Y-axeln: standardiserad skala.

Sökande 2011 testades under det skarpaste läget, normgruppen och redan antagna 2010 i oskarpa lägen. Sökande 2010 och 2012 testades i mellanliggande lägen vad gäller skarphet. Det kan noteras att spännvidden i medelvärde mellan grupperna är ca 1 standardavvikelse, en mycket stor skillnad enligt konventionerna för att mäta skillnader (Cohen, 1988). Antagandet att måtten på skönmålning enbart mäter "personlighet" får inget stöd av dessa data eftersom skönmålningen varierade mellan betingelser som den förväntades göra om den var ett mått på taktiksvär.

I denna rapport går jag vidare med frågan genom att inkludera relevanta mått på två personlighetsvariabler, nämligen moralstyrka och behov av att vara omtyckt ("need for approval", förkortas NA). Dessutom konstruerades ett nytt mått på taktiksvär, nämligen inställsamhet. Syftet var att undersöka i vilken mån dess tre skalor korrelerade och kunde förklara overt skönmålning.

Datainsamling

De skalor som presenteras i denna forskningsrapport bygger på data som samlades in på Internet. 143³ personer deltog, varav 58 % var kvinnor. Åldern varierade mellan 18 och 74 år, median 40 år. Drygt hälften hade högskoleutbildning.

Resultat

Moralstyrka

Tanken bakom att inkludera mätning av moralstyrka i designen var att de item som ingår i det overta måttet handlar om att man förnekar att man handlat "omoraliskt", eller att man påstår att man handlat "moraliskt". Det finns givetvis människor som är starkt styrda av moraliska värderingar och de borde kanske därför få ett högt värde i skalor av typ Crowne-Marlowes skala för att mäta social önskvärdhet. Å andra sidan är det rimligt att tro att svaren på ett personlighetstest i ganska hög grad är taktiska svar, särskilt om det handlar om att ett skarpt läge. Det är därför en empirisk fråga i vilken mån moralstyrka avspeglas i skalorna för skönmålning.

Enligt Wikipedia: "The super-ego aims for perfection. It forms the organized part of the personality structure, mainly but not entirely unconscious, that includes the individual's [ego ideals](#), spiritual goals, and the psychic agency (commonly called "conscience") that criticizes and prohibits his or her drives, fantasies, feelings, and actions. "The Super-ego can be thought of as a type of conscience that punishes misbehavior with feelings of guilt." Se http://en.wikipedia.org/wiki/Id,_ego_and_super-ego

Efter litteratursökning kunde jag lokalisera en skala för mätning av moralstyrka, publicerad i en doktorsavhandling av Cosgro (Cosgro, 1991). Den översattes till svenska och underkastades faktor- och itemanalys. Av de ursprungliga 27 item som skalan innehöll, behölls 12, som hade en reliabilitet på 0.90. Dessa var de dominerande item som föll ut i första faktor.

Cattells 16PF innehåller också en skala som avses mäta moralstyrka, men hans personlighetstest har fått mycket starkt kritik⁴. De faktorer som han och hans efterföljare menar sig ha etablerat har inte kunnat replikeras (Cooper, 2002). Cooper skriver

" The problem is that no one apart from Cattell and his colleagues can find anything approaching the 16 factors in the 16 PF." (Sid. 108).

Jag höll mig därför till Cosgros skala som bygger på hans doktorsavhandling och som har goda psykometriska egenskaper och grundligt begreppsvaliderats.

³ Åtta personer deltog inte i denna fas av undersökningen, som tidsmässigt låg ett par veckor senare.

⁴ Märkligt nog används detta test fortfarande ganska ofta i Sverige, inte minst inom offentlig sektor, trots att kraven på vetenskapligt stöd där är mycket stora. Ett av de första svenska försöken att validera det var Ramfalks doktorsavhandling (Ramfalk, 1957); det misslyckades fullständigt.

Inställsamhet

Inställsamhet kunde tyckas vara en uppenbart viktig och intressant dimension för mätning i samband med testningar i skarpa lägen, men någon etablerad skala för att mäta detta begrepp kunde inte lokaliseras. Av detta skäl konstruerades en skala med 8 item, reliabilitet = 0.70.

Alla som arbetar med anställningsintervjuer torde ha blivit uppmärksammade på att människor gör allt de kan för att framstå på ett sätt som de tror kan gynna dem. Detsamma gäller i många andra sammanhang, t ex bland säljare eller hos vissa (många?) politiker.

Behov att bli uppskattad ("Need for approval")

Crowne och Marlowe hade enligt min mening gjort ett viktigt pionjärbete med sin skala för mätning av social önskvärdhet men i sin egen forskning spårade de snabbt in på idén att det handlar om ett mera permanent personlighetsdrag, som de kallade "Need for Approval" och inte ett efter tillfället anpassat taktiskt beteende (Crowne & Marlowe, 1964). Denna tes tycks ha vunnit vitt spridd trovärdighet, men utan god saklig grund. Man har försökt validera den mot direkt mått på behovet att vara uppskattad, t ex med hjälp av GABS – skalan (Bernard, 1998; Lindner, Kirkby, Wertheim, & Birch, 1999), som har validerats i skilda sammanhang. I vår studie användes 7 översatta GABS-item som hade en reliabilitet på 0.84.

En aktuell doktorsavhandling av Barger (Barger, 2012) ger en mycket användbar översikt av mätning av NA, inklusive fullständiga förteckningar av de item som ingår i skalorna. Förutom Marlowe-Crownes skala studerade han DAS-skalan (Weissman & Beck, 1978) samt GABS – skalan (Bernard, 1998; Lindner et al., 1999). GABS och DAS korrelerade 0.63 i Bargers studie, men Marlowe-Crownes skala korrelerade bara -0.08 och -0.15 med respektive DAS och GABS. Det verkar alltså inte som om Marlowe-Crownes skala mäter den postulerade personlighetsdimensionen. Det är tydligt att DAS är mycket konsekvent inriktad just på behovet att vara omtyckt medan Marlowe-Crownes skala bygger på en samling av olika handlingar som är vanliga, men anses icke önskvärda, eller ovanliga, men anses önskvärda. Det är i den meningen ett dolt mått på NA, om den nu alls kan anses vara ett sådant mått. I denna rapport redovisas resultat av att relatera overt skönmålning till NA.

Övergripande analys

De två måtten på skönmålning relaterades till de tre skalorna moralstyrka, behov av att vara omtyckt (NA) och inställsamhet. Se Tabell 3.

Tabell 3. Resultatet av regressionsanalys av overt skönmålning som en funktion av moralstyrka, behovet av att vara omtyckt samt inställsamhet.

Oberoende variabel	Overt skönmålning
Moralstyrka	0.231
Behov av att vara omtyckt	-0.022
Inställsamhet	0.280
Förklarad varians	0.167

Korrelationen mellan moralstyrka och inställsamhet var 0.46. Resultaten visar att behovet av att vara omtyckt har ytterst lite med skönmålning att göra medan både moralstyrka och inställsamhet tycks vara av stor betydelse.

I en regressionsanalys av moralstyrka fann jag att 36 % av variansen i denna variabel kunde förklaras av overt och kovert skönmålning.

Diskussion

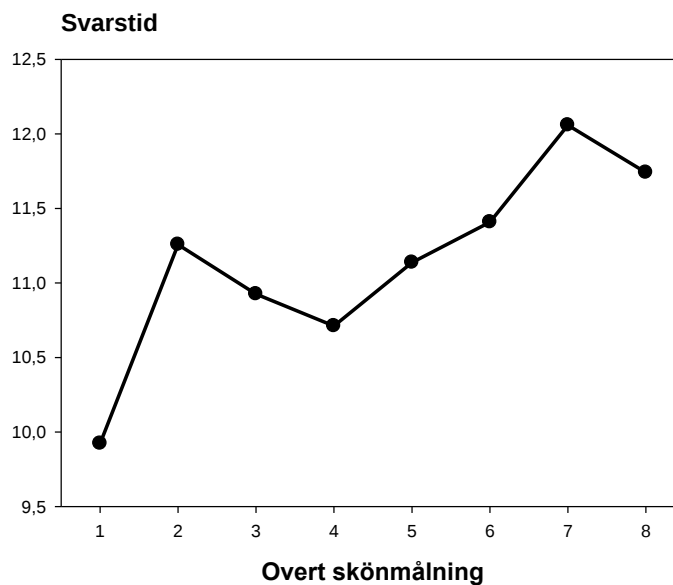
Betrakta lite närmare de typiska påståendena i Marlowe-Crownes skala, t ex "I sometimes think when people have a misfortune they only got what they deserved". Om man svarar "nej" på denna fråga bidrar det positivt till poängen i skalan. Samtidigt är det troligen ett ganska vanligt beteende. En möjlig tolkning av varför någon svarar nej är att han eller hon vill framstå som en sympatisk person, och att detta är ett uttryck för en generell sådan tendens. En annan tolkning är önsknigen att framstå som en sympatisk person *i testningen*. *I inget fall är det påvisat att det är fråga om ett generellt och tydligt behov av att vara omtyckt. Det är enbart en spekulativ tolkning.* Det är en tolkning som går mycket längre, och som enligt vad vi sett inte får stöd när man relaterar Marlowe-Crownes skala till övriga skalor som direkt mäter behovet av att vara omtyckt. Marlowe-Crownes skala som mått på NA är ett exempel på s k "subtle items", dvs item som inte tydligt ger uttryck för det begrepp som ska mätas, utan där det krävs en mer eller mindre långsökt tolkning. Forskningen om sådana item har inte gett positiva resultat (Weed, Ben-Porath, & Butcher, 1990).

Det kan tilläggas att Marlowe-Crownes skala i en meta-analys (Ones, Viswesvaran, & Reiss, 1996) befanns ha mycket svaga samband med arbetsresultat, omkring 0. Det skulle knappast ha varit fallet om den skalan mäter personlighet. Andra studier har visat tydliga negativa effekter av skönmålning på arbetsresultat (Peterson & Griffith, 2006). Kortsiktiga positiva bedömningar kan en anställd uppnå i jobbet genom olika taktiska sätt att bete sig men ganska snart genomskådas i de flesta fall han eller hon (Higgins, Judge, & Ferris, 2003).

Slutsatsen blir att Marlowe-Crownes skala inte mäter personlighet i betydelsen NA, utan - som ursprungligen avsett - tendensen att svara socialt önskvärt i en testsituation, men också personlighet i betydelsen moralstyrka.

Detta hindrar inte att en tendens att svara taktiskt kan återfinns i flera situationer, t ex i olika test eller vid olika tidpunkter. Man har funnit för gruppen av testpersoner med höga SD-värden att testvaliditeten är låg (White, Young, & Rumsey, 2001), något som är att förvänta om SD-skalorna faktiskt mäter snedvridning av testresultaten.

Det finns samband mellan Marlowe-Crownes skala och svarstid (Sjöberg, 2012/2013), vilket också stämmer med antagandet att Marlowe-Crownes skala mäter skönmålning. Se Fig. 5.



Figur5. Sambandet mellan genomsnittlig svarstid i sekunder (per item) och overt skönmålning i 8 grupper av testade personer

Slutsatsen av de analyser som presenterats här är ett overt skönmålning i hög grad, men inte enbart är ett uttryck för takiska svar i testsituationen. Det finns en hel del kvar att förklara i variansen i overt skönmålning, och även om idén att overt skönmålning är ett uttryck för NA inte får stöd här eller i andra kritiska studier tyder resultaten på att personlighet kan spela in, inte bara den tillfälliga taktiska attityden. Fortsatt forskning får visa vilka implikationerna bör vara för praktisk testverksamhet. Om man använder både overt och kovert skönmålning korregerar man dock för en mycket stor del av effekten av moralstyrka.

Värderingar

Inledning

Det är en vanlig uppfattning att värderingar, eller värdegrund, är av stor betydelse i arbetslivet. Det är ett rimligt antagande, även om det finns ganska lite forskning till stöd för det. Om man studerar förekommande personlighetstest ser man emellertid ganska sällan att de täcker av viktiga värderingar.

För en tid sedan skapade jag ett formulär för att mäta värderingar. Skälet till att jag ville skapa ett helt eget formulär var att de befintliga systemen för att mäta värderingar tycktes mig ofullständiga och ganska godtyckliga. De låg för långt från arbetslivets verklighet. Formuläret prövades vid Handelshögskolan i Riga. När jag tog upp det på nytt var det bl a för att pröva ett nytt svarsformat. (Det tidigare var ipsativt). Jag ville också utnyttja de kunskaper och den metodik som jag utvecklat i samband med arbetet med UPP-testet. Alltså valde jag en normativt svarsformat, och de skalor som fanns att tillgå för att mäta skönmålning (overt och kovert skala).

Detta är inte platsen för en utförlig diskussion om värderingar; jag ger bara ett par exempel. En klassiker inom området är Rokeach (Rokeach, 1973, 1979), vars värden omfattar:

1. **True Friendship**
2. **Mature Love**
3. **Self-Respect**
4. **Happiness**
5. **Inner Harmony**
6. **Equality**
7. **Freedom**
8. **Pleasure**
9. **Social Recognition**
10. **Wisdom**
11. **Salvation**
12. **Family Security**
13. **National Security**
14. **A Sense of Accomplishment**
15. **A World of Beauty**
16. **A World at Peace**
17. **A Comfortable Life**
18. **An Exciting Life**

Listan är tankeväckande. Åtskilliga av värdena är subtila och existentiella. Inget fel i det, men det praktiska värdet är inte övertygande. Märk också att vissa av värdena gäller samhället eller kulturen (t ex jämlikhet och frihet) och är mycket allmänna. Vem är inte för jämlikhet och frihet i vår kultur? Jag valde att arbeta med mindre subtila värderingar som gäller personen som gör bedömningarna.

En ambitiös vidareutveckling har föreslagits av Schwartz (Schwartz, 1992), och den tycks fin vara dominerande och ligger till grund för stora internationella jämförelser. Många andra har spekulerat om värden, men få har uppnått en klar och lätt-tillgänglig struktur. Det är viktigt i praktiskt arbete att undvika komplexa strukturer vars betydelse för arbetslivet är svårt att fastställa. Schwartz system har prövats mot specifika attityder med nedslående resultat (Sjöberg, 2008a); inga tydliga samband alls.

Syftet med mitt arbete har varit att uppnå just en sådan enkel struktur. Data insamlade med det tidigare formuläret gav en enkel 2-dimensionell struktur, se Fig.6.

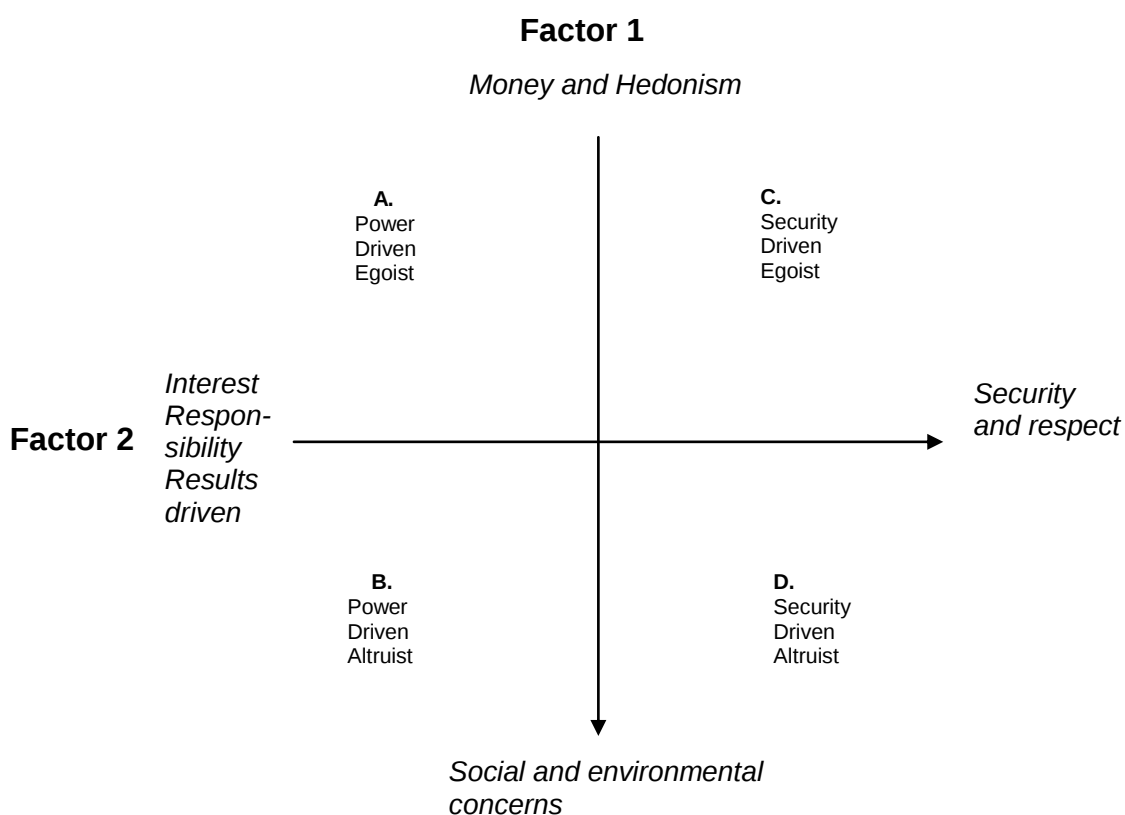


Fig. 6. Ursprunglig värdestruktur.

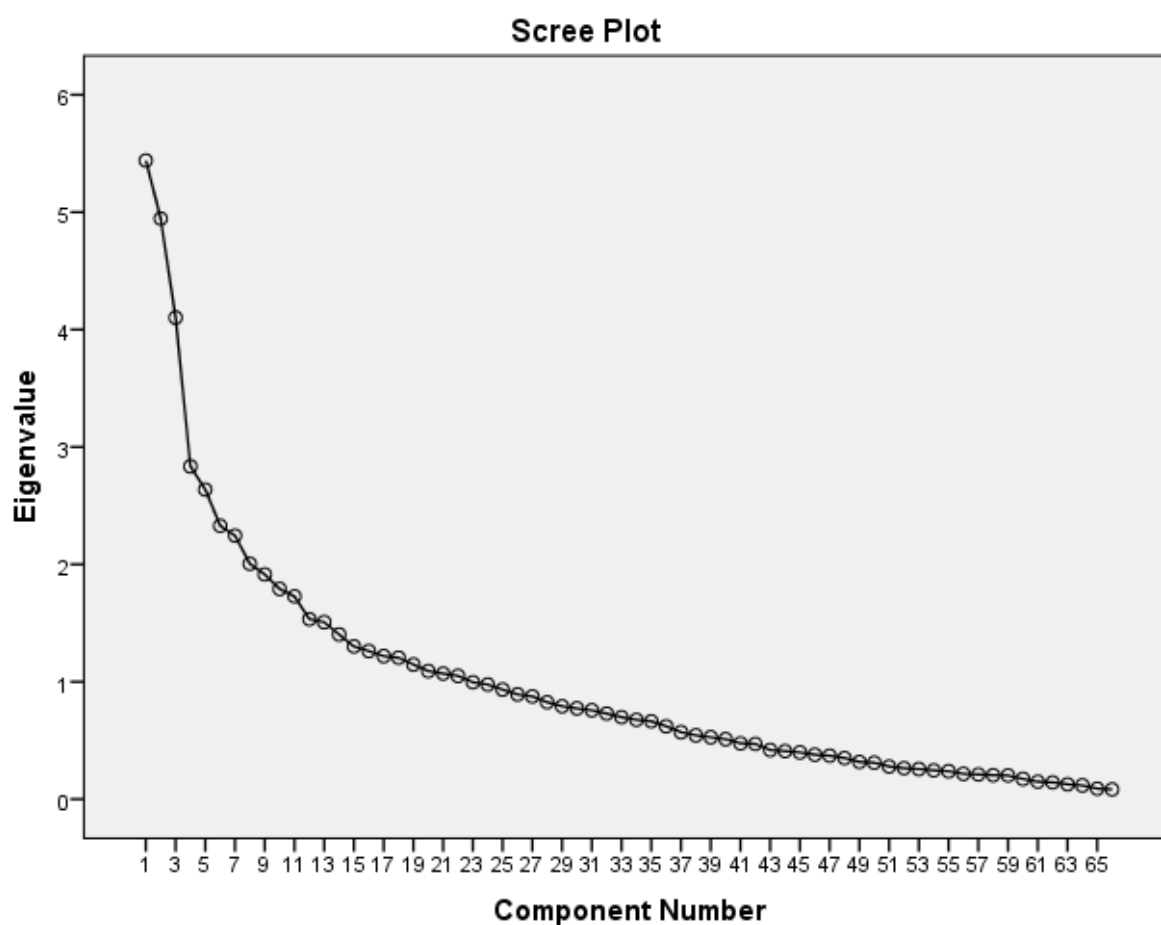
Systemet tillåter fyra övergripande dimensioner, inte ca 10-20 som annars är vanliga. Sextio-sex påståenden formulerades utifrån det ursprungliga värdeformuläret, och gavs till ett antal testpersoner, se nästa avsnitt. Skalorna för skönmålning fanns också med. Problemet med skönmålning är givetvis lika akut här som i t ex vanliga item för självskattning av personlighet, men möjligen är sambanden svagare om testpersonerna inte uppfattar hur betydelsefulla dessa item är, eller vad som är önskvärda svar.

Datainsamling

De skalor som presenteras i denna forskningsrapport bygger på data som samlades in via Internet. 151 personer deltog, varav 58 % var kvinnor. Åldern varierade mellan 18 och 74 år, median 40 år. Drygt hälften hade högskoleutbildning.

Resultat

En komponentanalys visade tydligt på fyra övergripande dimensioner. Se Fig 7 för ett skree-test⁵.



Figur 7. Scree test för de 66 värdepåståendena.

De fyra dominerande komponenterna förklarade 26% av variansen och transformerades till enkel struktur med Varimax. (Oblik transformation gav samma resultat, eftersom axlarna var i stort sett ortogonala). Faktorerna var tydligt möjliga att tolka:

⁵ Används för att uppskatta hur många komponenter som bör behållas.

I. Karriär, pengar, makt. (15 item, $\alpha = 0.78$).

II. Miljö, medkänsla. (9 item, $\alpha = 0.80$).

III. Njutning, hedonism. (8 item, $\alpha = 0.64$)

IV. Respekt från andra. (5 item, $\alpha = 0.60$)

Det normativa formuläret gav alltså en resultat som kunde tolkas tydligt, och mätas med hög tillförlitlighet.

Skönmålning

Sambandet mellan värdeskalorna och skönmålning framgår av Tabell 4.

Tabell 4. Korrelationer mellan värde skalor och skalor som mäter skönmålning.

	Skalor som mäter skönmålning	
Värdedimension	Overt	Kovert
Karriär, pengar, makt	0.06	0.27
Miljö, medkänsla	0.15	0.04
Njutning, hedonism	-0.40	-0.52
Respekt från andra	-0.24	-0.28

Det är intressant att se att skalorna är relaterade till skönmålning i mycket varierande grad. En korrektion är uppenbarligen påkallad i praktisk användning av dem.

Diskussion

Den erhållna strukturen var 4-dimensionell, inte som den tidigare 2-dimensionell. Detta beror troligen på att data insamlats med ett normativt format, inte det ipsativa som användes tidigare. Med ipsativa format skapas artificiellt förhöjd korrelation mellan item och därmed ofta en enklare struktur.

Alla 4 dimensionerna kunde mätas med tillräcklig reliabilitet. De ger en ganska enkel och överblickbar bild av värderingar med inriktning på individernas liv och målsättningar, med förbigående av subtila existentiella aspekter och samhällsmål av generellt slag.

Personlighet och värderingar

Ytterligare validering av vissa av skalorna är möjlig genom att beräkna korrelationer mellan olika typer av skalor. Se Tabell 5 för resultatet.

Tabell 5. Korrelationer mellan värderingar och personlighetsskalor

	Inställsamhet	Moralstyrka	Beslutssamhet	Handlingskraft	Machiavellianism	Lindrig psykopati
Inriktad på karriär	-0.18*	-0.37**	0.53**	0.64**	0.27**	0.11
Engagerad i miljö, har medkänsla	-0.22**	0.03	-0.01	0.11	-0.30**	-0.12
Lägger vikt vid njutning	0.34**	0.35**	-0.24**	-0.34**	0.51**	0.50**
Behov av att mötas med respekt	0.21*	0.41**	-0.42**	-0.30**	0.10	0.12

* Korrelation signifikant på nivån 0.05 (2-tailed).

** Korrelation signifikant på nivån 0.01 (2-tailed).

Det är intressant att se att psykopati och Machiavellianism båda har höga samband med värdering av njutning, som väntat. Detta innebär ännu en validering av dessa skalor.

Validitet: Proxyvalidering

Ett proxykriterium bildades genom att kombinera arbetstillfredsställelse, arbetsvilja, förändringsvilja och arbetsintresse. Validitet gentemot ett sådant kriterium har visat sig nära relaterat till externa kriterier av typ arbetsprestation (Sjöberg, 2013). Närmare information om proxyvalidering ges i Bilaga 3. De nya skalorna korrelerades med proxykriteriet i en grupp av 68 personer som tagit det fulla UPP-testet förutom de nya skalorna⁶. Tabell 6 visar skalornas validitet.

⁶ . Alla var erbjudna att ta det full UPP-testet; 68 gjorde det. En datamängd som innehöll svaren på UPP samt de nya skalorna bildades för att analysera de nya skalornas validitet.

Tabell 6. Skalornas validitet mot förväntad arbetsprestation (proxy)

Skala	Validitet
Inriktad på karriär	0.46***
Engagerad i miljö, har medkänsla	-0.23
Lägger vikt vid njutning	0.38**
Behov av att mötas med respekt	0.26*
Inställsamhet	0.28*
Beslutsamhet	0.57***
Handlingskraft	0.71***
Machiavellianism	0.02
Psykopati	0.26*

* Korrelation signifikant på nivån 0.05 (2-tailed).

** Korrelation signifikant på nivån 0.01 (2-tailed)

*** Korrelation signifikant på nivån 0.001 (2-tailed)

Tabellen visar att det fanns hög validitet framför allt hos handlingskraft, beslutsamhet samt värdering av karriär. Dessa värden var av betydande storlek, främst handlingskraft.

Slutsatser

Den valda strategin har fungerat bra. Flera nya skalor kunde konstrueras och valideras. Begreppsvalidering mot kända och etablerade internationella skalor ligger på mycket hög nivå, vilket innebär att man kan vänta sig att även samband med externa kriterier av olika slag kommer att vara höga.

Validering mot ett proxykriterium gav mycket höga värden för handlingskraft, beslutsamhet och värdering av karriär.

Referenser

- Ames, D. R., Rose, P., & Anderson, C. P. (2006). The NPI-16 as a short measure of narcissism. *Journal of Research in Personality*, 40, 440-450.
- Babiak, P., & Hare, R. D. (2006). *Snakes in suits. When psychopaths go to work*. New York: Harper.
- Barger, V. A. (2012). *Individual differences in need for approval: Measurement and marketing implications*. ProQuest Information & Learning, US.
- Bernard, M. E. (1998). Validation of the General Attitude and Belief Scale. *Journal of Rational-Emotive and Cognitive-Behavior Therapy*, 16, 183-196.
- Board, B. J., & Fritzon, K. (2005). Disordered personalities at work. *Psychology Crime & Law*, 11, 17.
- Christie, R., & Geis, F. L. (1970). *Studies in Machiavellianism*. New York: Academic Press.
- Cohen, J. (1988). *Statistical power analysis for behavioral sciences (2nd ed.)*. Hillsdale, NJ: Erlbaum.
- Cooper, C. (2002). *Individual differences. 2nd edition*. London: Hodder EducationW.
- Cosgro, M. A. (1991). *The development of a Superego Assessment Scale*. Fresno CA: California School of Professional Psychology.
- Crowne, D. P., & Marlowe, D. (1960). A new scale of social desirability independent of psychopathology. *Journal of Consulting and Clinical Psychology*, 24, 349-354.
- Crowne, D. P., & Marlowe, D. (1964). *The approval motive*. New York: Wiley.
- Dahling, J. J., Whitaker, B. G., & Levy, P. E. (2009). The development and validation of a new Machiavellianism scale. *Journal of Management*, 35, 219-257.
- De Colli, D. (2011). *Ett nytt svenskt arbetspsykologiskt test och arbetsprestation inom polisen – samtidig validitet*.
- 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.
- Fico, J. M., Brady, R., & Hogan, R. (2008). Identifying potential derailing behaviours: Hogan Development Survey. In J. Passmore (Ed.), *Psychometrics in coaching: Using psychological and psychometric tools for development*. (pp. 171-188): London, England: Kogan Page.
- Furnham, A. (1986). Response bias, social desirability and dissimulation. *Personality & Individual Differences*, 7, 385-400.

- Germeijs, V., & De Boeck, P. (2002). A measurement scale for indecisiveness and its relationship to career indecision and other types of indecision. *European Journal of Psychological Assessment*, 18, 113-122.
- Germeijs, V., & Verschueren, K. (2011). Indecisiveness and Big Five personality factors: Relationship and specificity. *Personality and Individual Differences*, 50, 1023-1028.
- Higgins, C. A., Judge, T. A., & Ferris, G. R. (2003). Influence tactics and work outcomes: A meta-analysis. *Journal of Organizational Behavior*, 24, 89-106.
- Hough, L. M. (1998). The millennium for personality psychology: New horizons or good ole daze. *Applied Psychology: An International Review*, 47, 233-261.
- Johnson, J. W. (2000). A heuristic method for estimating the relative weight of predictor variables in multiple regression. *Multivariate Behavioral Research*, 35, 1-19.
- Johnson, J. W. (2004). Factors Affecting Relative Weights: The Influence of Sampling and Measurement Error. *Organizational Research Methods*, 7, 283-299.
- Johnson, J. W., & LeBreton, J. M. (2004). History and Use of Relative Importance Indices in Organizational Research. *Organizational Research Methods*, 7, 238-257.
- 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, 376-407.
- Levenson, M. R., Kiehl, K. A., & Fitzpatrick, C. M. (1995). Assessing psychopathic attributes in a noninstitutionalized population. *Journal of Personality and Social Psychology*, 68, 151-158.
- Lindner, H., Kirkby, R., Wertheim, E., & Birch, P. (1999). A brief assessment of irrational thinking: The shortened General Attitude and Belief Scale. *Cognitive Therapy and Research*, 23, 651-663.
- Lönnqvist, J.-E., Paunonen, S., Tuulio-Henriksson, A., Lönnqvist, J., & Verkasalo, M. (2007). Substance and style in socially desirable responding. *Journal of Personality*, 75, 291-322.
- 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.
- McFarlin, D. B., & Sweeney, P. D. (2010). The corporate reflecting pool: Antecedents and consequences of narcissism in executives. In B. Schyns & T. Hansbrough (Eds.), *When leadership goes wrong: Destructive leadership, mistakes, and ethical failures*. (pp. 247-283): Greenwich, CT, US: IAP Information Age Publishing.
- Meade, A. W. (2004). Psychometric problems and issues involved with creating and using ipsative measures for selection. *Journal of Occupational and Organizational Psychology*, 77, 531-552.

- Neter, E., & Ben-Shakhar, G. (1989). The predictive validity of graphological inferences: A meta-analytic approach. *Personality and Individual Differences*, 10, 737-745.
- O'Boyle Jr, E. H., Forsyth, D. R., Banks, G. C., & McDaniel, M. A. (2011). A meta-analysis of the dark triad and work behavior: A social exchange perspective. *Journal of Applied Psychology*, 97, 557-579.
- 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.
- 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.
- Paulhus, D. L., & Reid, D. (1991). Enhancement and denial in socially desirable responding. *Journal of Personality and Social Psychology*, 60, 307-317.
- Peterson, M. H., & Griffith, R. L. (2006). Faking and job performance. In R. L. Griffith & M. H. Peterson (Eds.), *A closer examination of applicant faking behavior* (pp. 233-261). Greenwich, CT: IAP.
- Ramfalk, C. W. (1957). *A study of the selection of personnel for higher positions in industrial organizations*. Stockholm: Almqvist & Wiksell.
- Rokeach, M. (1973). *The nature of human values*. New York: Free Press.
- Rokeach, M. (1979). *Understanding human values*. New York: Free Press.
- Schwartz, S. H. (1992). Universals in the content and structure of values: theoretical advances and empirical tests in 20 countries. *Advances in Experimental Social Psychology*, 25, 1-65.
- Sjöberg, L. (2005). Självskattningstest: Den som ljugar fräckast vinner? (Self report tests: will the liar win?). *Personel och Ledarskap*, 36, 22-25.
- Sjöberg, L. (2008a). *Attityd till slutförvar av använt kärnbränsle: Struktur och orsaker. (Attitudes toward the final repository for spent nuclear power: Structure and causes)* (Research Report No. R-08-119). Stockholm: SKB. Svensk Kärnbränslehantering AB.
- Sjöberg, L. (2008b). *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. (2009). *UPP-testet: Korrektur för skönmålning. (The UPP test: Correction for impression management)*. *Forskningsrapport 2009:3*. Stockholm: Psykologisk Metod AB.
- Sjöberg, L. (2010). *UPP-testet och kundservice: Kriteriestudie. Forskningsrapport 2010:6*. Stockholm: Psykologisk Metod AB.
- Sjöberg, L. (2012/2013). *Skönmålning i personlighetstest. SSE/EFI Working Paper Series in Business Administration, No. 2012:2*

- Sjöberg, L. (2013). *UPP-testet. Teknisk manual*. Stockholm: Psykologisk Metod AB.
- Sjöberg, L., Bergman, D., Lornudd, C., & Sandahl, C. (2011). *Sambandet mellan ett personlighetstest och 360-graders bedömningar av chefer i hälso- och sjukvården*. Stockholm: Karolinska Institutet, Institutionen för lärande, informatik, management och etik (LIME).
- Sjöberg, L., & Engelberg, E. (2005). Measuring and validating emotional intelligence as performance or self-report. In K. B. S. Kumar (Ed.), *Emotional intelligence. Research insights* (pp. 97-125). Hyderabad, India: ICFAI University Press.
- Sjöberg, L., & Wolgers, G. (2012). Personlighetstestning vid antagning till officersutbildning. En jämförelse mellan CTI och UPP-testet i olika grader av "skarpt läge". *Ledarskapscentrum, Försvarshögskolan. Forskningsrapport F:39*.
- Tett, R. P., Fitzke, J. R., Wadlington, P. L., Davies, S. A., Anderson, M. G., & Foster, J. (2009). The use of personality test norms in work settings: Effects of sample size and relevance. *Journal of Occupational and Organizational Psychology*, 82, 639-659.
- Weed, N. C., Ben-Porath, Y. S., & Butcher, J. N. (1990). Failure of Wiener and Harmon Minnesota Multiphasic Personality Inventory (MMPI) subtle scales as personality descriptors and as validity indicators. *Psychological Assessment: A Journal of Consulting and Clinical Psychology*, 2, 281-285.
- Weiss, E. (1950). *Principles of psychodynamics*: Oxford, England: Grune & Stratton.
- Weissman, A. N., & Beck, A. (1978). Development and validation of the dysfunctional attitudes scale: A preliminary investigation, *Proceedings of the meeting of the American Educational Research Association*. Toronto, Canada.
- White, L. A., Young, M. C., & Rumsey. (2001). Assessment of background and life experiences (ABLE) implementation issues and related research. In J. P. Campbell & D. J. Knapps (Eds.), *Exploring the limits in personnel selection and classification* (pp. 526-558). Mahway, NJ: Erlbaum.
- Wood, J. M., Nezworski, M. T., Lilienfeld, S. O., & Garb, H. N. (2003). *What's wrong with the Rorschach?: Science confronts the controversial inkblot test*. San Francisco, CA, US: Jossey-Bass.
- Vrij, A. (2000). *Detecting lies and deceit*. New York: Wiley.

Appendix 1. Korrigering för skönmålning

Personlighetstest används i stor omfattning vid anställningar i Sverige. Det handlar nästan alltid om självrapporttest där den testade får ta ställning till ett antal påståenden och ange i vilken grad de stämmer in på honom eller henne, eller om de håller med om påståendena. Test av den typen är oftast lätta att genomskåda. Omfattande forskning har visat att många, troligen alla, självrapporttest lider av denna svaghet (Furnham, 1986; Hough, 1998). Frågan är vad resultaten på ett sådant test egentligen är värda, om svaren på testet är starkt korrelerade med skönmålning, eller kraftigt påverkas av vilken betydelse testningen förväntas ha.

Forskningen visar emellertid att testen trots allt har ett visst prognosvärde för många olika typer av befattningar (Ones et al., 1996). Det låter ju bra att testens validitet finns kvar trots skönmålade svar, men det finns en hake i resonemanget. Man tar nämligen i praktiken inte alls, eller bara i mycket begränsad omfattning, *individuell* hänsyn till tendensen att svara taktiskt. På den individuella nivån kan effekterna av skönmålning bli enorma.

Det dominerande sättet att hantera detta problem tycks vara att använda ipsativa (jämförande) svarsformat. Dessa leder emellertid till stora och olösta statistiska och psykometriska problem med analys av strukturen i testuppgifterna och bestämning av reliabilitet (Meade, 2004). Empirisk forskning visar dessutom att de ipsativa formaten inte effektivt motverkar tendensen till skönmålning (Martin, Bowen, & Hunt, 2002). I en bilaga till den tekniska rapporten ges en mera omfattande diskussion av ipsativa testformat (Bilaga 3).

Två taktikskalor⁷ utvecklas och används i denna rapport. Traditionellt mäter man taktisksvar, eller skönmålning, med hjälp av en skala av den typ som Crowne och Marlowe publicerade för ganska länge sedan (Crowne & Marlowe, 1960), ofta kallad en skala för mätning av social önskvärdhet (social desirability). Uppgifterna i den skalan bygger på idén att de allra flesta någon gång betar sig på tvären mot gällande normsystem och att den som inte medger detta om sig själv antagligen svarar taktiskt. Ett exempel är:

”Jag har aldrig kastat skräp på gatan.”

Om en person svarar taktiskt på sådana uppgifter är det rimligt att tro att även övriga svar på testet är taktiska och tillrättalagda för att ge ett positivt intryck.

Paulhus har fört vidare diskussionen om skönmålning på ett intressant sätt (Paulhus, 1991; Paulhus & Reid, 1991). Han använde samma typ av uppgifter som Crowne och Marlowe men också en ny typ som avsåg att mäta självbedrägeri. Dessa uppgifter handlar om psykologiska tillkortakommanden som då och då inträffar för alla – den som förnekar dem ljuger för sig själv. Ett exempel är:

”Det är svårt för mig att stänga ute en störande tanke.”

Lönnqvist et al. (Lönnqvist, Paunonen, Tuulio-Henriksson, Lönnqvist, & Verkasalo, 2007) fann att Paulhus-skalorna åtminstone till en del tycktes mäta metodartefakter på det avsedda

⁷ Termen taktikskala föredras här framför lögnskala – det handlar ju om att man svarar taktiskt vilket inte behöver beskrivas med den nedsättande termen lögn. En annan vanlig beteckning är social önskvärdhet.

sättet. Man kan diskutera om inte samma sorts taktiksvär ligger bakom nekande svar på sådana uppgifter som på de mera traditionella uppgifterna som handlar om normbrott. Hursomhelst korrelerar de ganska högt med den traditionella typen av taktiksvär. I en urvalssituation (N=282, sökande till Handelshögskolan) användes Crowne-Marlowes skala och Paulhus båda skalor. Hans skala som byggde på frågor om normbrott korrelerade 0,84 med Crowne-Marlowe, medan skalan på "självbedrägeri" korrelerade 0,54 – också det ett högt värde.

Ett viktigt problem är frågan om uppgifterna i en taktikskala går att genomsåda. Tag som exempel ovanstående uppgift om att kasta skräp på gatan. Det är absurt att påstå att man aldrig gör det – åtminstone för de allra flesta. Sådana uppgifter kan därför skapa misstänksamhet, och de testade personer som förberett sig genom att läsa på om testmetodik kan kanske identifiera vilka uppgifter det rör sig om och räkna ut hur de ska besvaras för att nå ett maximalt positivt resultat på testet. (Detta är inte så lätt att göra). .

I arbetet med *UPP*-testet har konstruerats en egen skala på samma principer som Crowne-Marlowe och Paulhus, en övert taktikskala (öppen mätning), men också en kovert skala (dold mätning). Den senare består av uppgifter av samma typ som i den övriga delen av testet, men de är särskilt utvalda för att korrelerade högt med den överta taktikskalan. På så sätt får man en möjlighet att mäta taktiksvär som inte kan avslöjas av sofistikerade testtagare. De två taktikskalorna (övert och kovert) har begreppsvaliderats, se Tabell 5. Båda korrelerade högt med Crowne-Marloweskalen och högt sinsemellan. Den senare korrelationen var något lägre vilket är rimligt.

Tabell 5. Korrelationer mellan de tre taktikskalorna (N=159), data från Studie 1.

	Crowne-Marlowes taktikskala	Övert taktikskala	Kovert taktikskala
Crowne-Marlowes taktikskala	1.00	0.76	0.73
Övert taktikskala	0.76	1.00	0.56
Kovert taktikskala	0.73	0.56	1.00

Hur korrektionen för skönmålning görs är av stor betydelse. Det finns tre vanliga metoder, alla otillfredsställande:

1. Ipsativt format används, ibland med försök att matcha valalternativ i social önskvärdhet.
2. Psykologen som tolkar testresultatet gör en intuitiv bedömning av resultatets trovärdighet på grundval av värdet på en "lögnskala" och sin egen bedömning av den testades "ärlighet".
3. Man använder sig av ett gränsvärde. De 25 % av de testade som har högst värde på taktikskalan sällas bort och kommer inte ifråga för antagning eller anställning.

Problemen med dessa metoder är dels att forskning visat att de inte har den avsedda effekten (Metod 1), de är alltför subjektiva (Metod 2), eller ineffektiva (Metod 3). Att avslöja lögnare utifrån subjektiva intryck är något de flesta tror att de kan göra – men de kan det inte (Vrij,

2000). Metoden med gränsvärde rensar bara bort en mindre del av variansen som beror på skönmålning. De som rankas högst på testet tenderar att vara bluffmakare även efter det att den metoden har använts.

Alla metoderna har den gemensamma nackdelen att de inte korregerar testvariablerna specifikt utan de försöker göra en global korrektion. Den metodik som använts i *UPP*-testet innebär att korrektionen görs för varje testvariabel separat. Genomslaget för taktiksvartstendens kan ju vara mycket varierande beroende på vilken dimension det är som är aktuell. Den metodik som använts här utnyttjar effekten av taktikskalorna fullt ut med hjälp av en regressionsmodell och beräkning av residualer, se Fig 8.

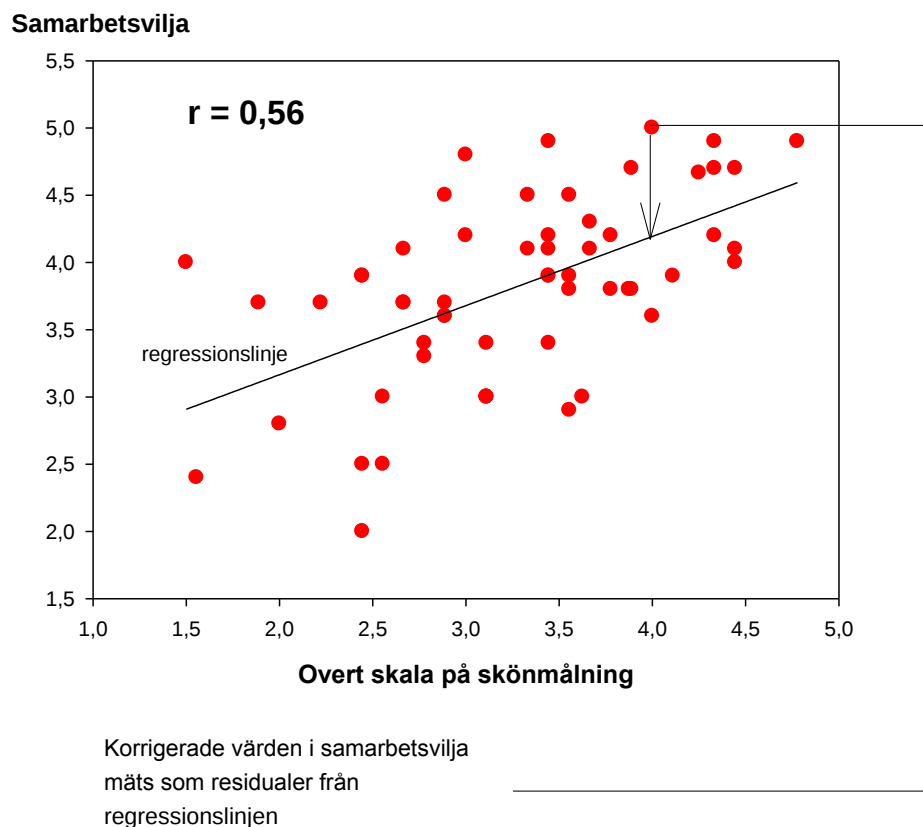
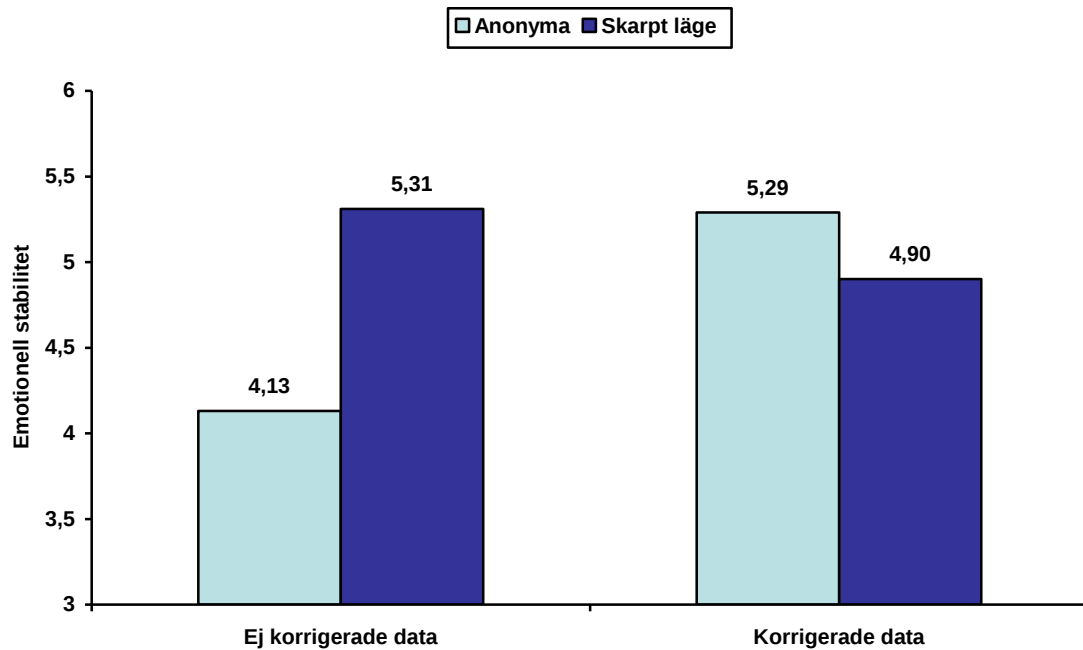


Fig. 8. Illustration av hur korrigerade värden i ett test uppskattas med hjälp av en psykometrisk modell.

Denna metodik utnyttjar taktikskalan maximalt. Effekten av ett starkt incitament att svara taktiskt nästan helt försvinner med denna typ av korrektion (Sjöberg, 2008b; Sjöberg & Engelberg, 2005). Skönmålningen har alltså inte längre någon effekt när korrektionen har gjorts.

Den här tekniken har fungerat utmärkt och validerats i en studie av sökande till HHS, vars värden jämfördes med studerande på skolan som tog samma test men i medvetande om att deras testvärden inte skulle ha någon för dem viktig betydelse (Engelberg & Sjöberg, 2005). Se Fig. 9 som visar att skillnaden i okorrigerade värden starkt reduceras när vi genomfört korrektionen för skönmålning.

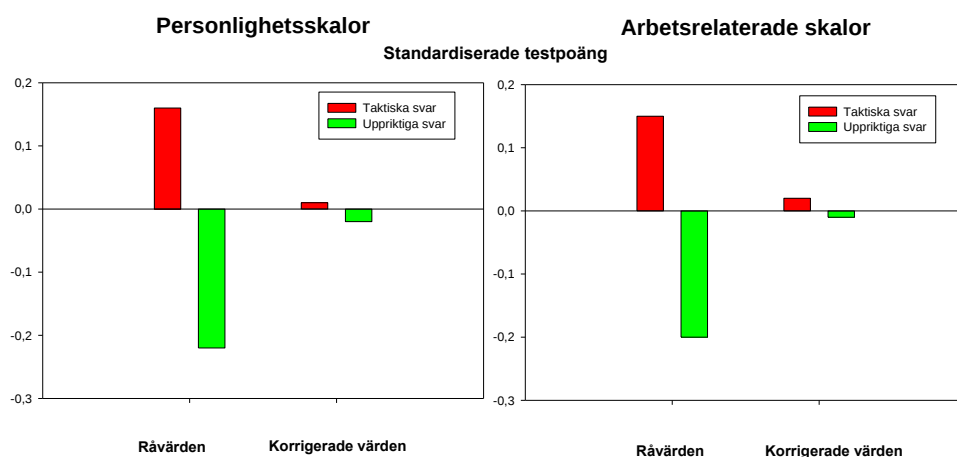


Figur 9. Effekten av korrektion för skönmålning vid jämförelse mellan anonym testning och skarpt läge, emotionell stabilitet. Medelvärde=5, standardavvikelsen=2.

Fig.10 visar resultatet av ett experiment som beskrivs närmare i den tekniska rapporten. Somliga av de testade personerna uppmanades att skönmåla så mycket de kunde, andra uppmanades att ge ärliga svar. De som skulle bluffa uppmanades så här:

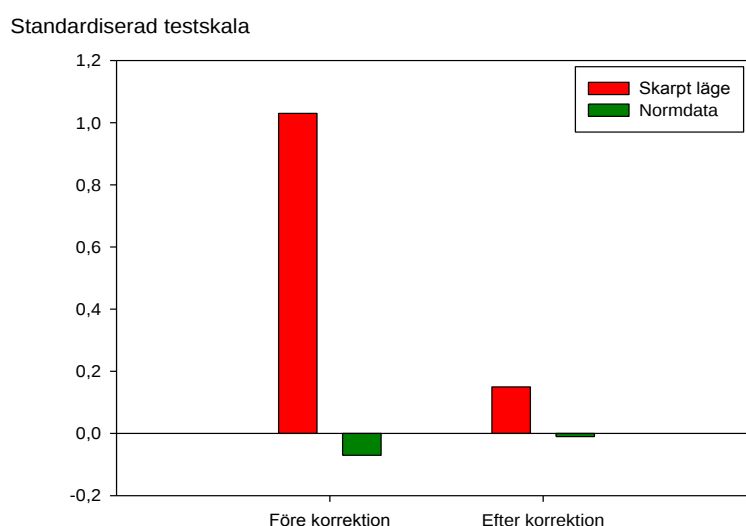
”Tänk på ett jobb som du mycket gärna skulle vilja ha. Tänk dig nu att du sökt just det jobbet och att den här testningen är ett *mycket viktigt* led i anställningsförfarandet. Svara på testuppgifterna så att framstår som en person som är precis sådan som man söker till jobbet. Det kanske innebär att du bluffar en del, men det är just det som vi har för avsikt att studera i den här undersökningen. Känn dig därför fri att svara taktiskt!”

Skillnaden var stor före korrektion men försvann till 90-95 % efter korrektion, se Fig. 9.



Figur 10. Genomsnittliga effekter av korrektion för skönmålning, standardiserade skalor.

Delar av *UPP*-testet (screening-modulen) används i rekryteringsplattformen Katapult⁸. De som där tar testet gör det i skarpt läge, dvs. i samband med att de söker ett jobb. Se Fig. 11 för skillnader mellan testade i skarpt läge och normdatabasen (ej skarpt läge), före och efter korrektion för skönmålning. Figuren gäller utåtvändhet men liknande resultat erhöles för andra testvariabler. Ca 90 % av effekten av skönmålning kunde elimineras. Detta resultat betyder också att normdatabasen är relevant även vid arbete i skarpt läge, om man nämligen korregerar på det sätt som görs i *UPP*-testet.



Figur 11. Utåtvändhet i skarpt läge, före och efter korrektion för skönmålning.

En korrektion av denna typ har stor praktisk betydelse. På individnivå kan effekterna bli mycket stora. Personer som hamnat i tätgruppen efter testning kan förpassas till botten av fördelningen, om de har skönmålat starkt. Kvinnor tenderar vid chefsrekrytering att gynnas av att man genomför en korrektion (Sjöberg, 2005), och att missgynnas om man inte gör det.

Enligt testteorin bör de korrigerade testvärdena ha lägre reliabilitet än de ursprungliga. Detta beror på att de har formen av differensvärden. Det har ofta hävdats att differensvärden har en låg reliabilitet. Analyser av samtliga indexvariabler i testet visade som väntat att det blev en viss sänkning av reliabiliteten efter korrektionen för skönmålning, från ett genomsnitt på 0,77 till 0,72. Å andra sidan minskade också felspridningen, dvs. precisionen i diagnostiska bedömningar av enskilda personer. Denna minskning uppskattades till, i genomsnitt, 14 %. För praktiska syften vid användningen av testet finns de alltså två skäl att korrigera för skönmålning:

- Testvärdena ger efter korrektion en sannare bild av testpersonen eftersom eventuella försök att svara taktiskt i stort sett inte påverkar dem
- Testvärdena är mindre påverkade av slumpfel i mätningarna

UPP-testet använder alltid korrektion för skönmålning enligt den metod som beskrivs i denna manual, den tekniska rapporten och i (Sjöberg, 2009).

⁸ www.katapult.com

Appendix 2. Begreppsförklaringar: nya skalor

Begreppsförklaringar för nya skalor i UPP-testet	
Testskala	Mycket högt värde (>7) utmärks av de flesta eller alla av nedanstående beteenden:
<i>Beslutsamhet</i>	Fattar snabba och genomtänkta beslut
<i>Handlingskraft</i>	Genomför med kraft de åtgärder som krävs i en viss situation
<i>Självuppskattning</i>	Mycket hög självuppskattning, på gränsen till "självförälskelse", kan innebära svårigheter i en chefsposition: vill inte ta kritik, negligerar andras behov
<i>Machiavellianism</i>	En cynisk och manipulativ attityd, utpräglad egoist
<i>Psykopati</i>	Manipulativ, kortsiktig, obehärskad
<i>Moralstyrka</i>	Styrd i hög grad av sina moraliska övertygelser
<i>Inställsamhet</i>	Letar efter sätt att bli omtyckt i situationer av vikt, som en anställningssituation, beter sig taktiskt
<i>Behov av att bli uppskattad</i>	Vill mycket gärna vara omtyckt av personer som hon uppfattar som viktiga för henne
<i>Värde: Karriär</i>	Betonar vikten av att göra karriär, få makt, tjäna mycket pengar
<i>Värde: Miljö</i>	Betonar miljön, medkänsla med dem som har det svårt
<i>Värde: Hedonism, njutning</i>	Vill uppnå njutningar
<i>Värde: Respekt</i>	Vill få andras respekt

Appendix 3. Proxyvalidering

En enkel metod för validering är att använda proxykriterier, alltså data som kan fungera som ersättning för oberoende bedömningar av arbetsprestationer och som kan samlas in i samband med testningen. Det kan handla om mått på arbetsmotivation eller balans mellan arbete och övrigt liv. Det är väl känt från mycket omfattande forskning att sådana data är starkt relaterade till arbetsresultat. Se t ex Judge m fl (Judge, Thoresen, Bono, & Patton, 2001) för en metaanalys av arbetstillfredsställelse och arbetsresultat.

Resultat av proxyvalidering

Omfattande valideringsarbete genomfördes mot arbetsmotivation och övriga arbetsrelaterade variabler. Här redovisas endast resultat från normmaterialet men analyser av övriga datamängder gav liknande resultat. Multipla korrelationer för vart och ett av de sex kriterierna mot FFM-modellen Big Five och mot *UPP*-testet framgår av Tabell 7. I *UPP*-testet ingick också 8 tillkommande variabler – uthållighet, samarbetsvilja etc.

Tabell 7. Validering mot 6 kriterievariabler, förklarad varians (kvadrerade multipla korrelationer).

Kriterievariabel	Big Five	<i>UPP</i> -testet
Förändringsvilja	0.137	0.274
Arbetstillfredsställelse	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

I samtliga fall var ökningen av validiteten högt signifikant och av dramatisk storlek. Förklarad varians på 0.317 motsvarar en korrelation på 0.56. Märk att det var balansdimensionen som drog ner värdet för *UPP*-testet i viss mån. Övriga kriterier kunde förklaras på högre nivå.

För Big Five var den genomsnittliga korrelationen med kriterierna endast 0.20, ett ganska normalt värde för personlighetstest. En ökning av validiteten från 0.20 till 0.56 ger mycket stora personalekonomiska vinster om testet används vid urval.

Dessa värden på validiteten påverkas ej av skönmålning eller andra felfaktorer eftersom dessa rensats bort ur både beroende och oberoende variabler. Om man inte gör denna korrektion blir validiteten, i termer av korrelation mellan test och kriterium, förhöjd med ca 0.10.

De traditionella beta-vikterna i multipel regressionsanalys kan vara vilseledande eftersom de är starkt beroende av kontexten av olika förklarande eller oberoende variabler. En variabel av betydelse kan få en alltför låg beta-vikt om den korrelerar med en annan oberoende variabel. Johnsons metod är ett försök till lösning av detta problem och uppskattar varje oberoende variabls bidrag till förklaringen av den beroende variabeln, både det unika bidraget och

bidraget tillsammans med andra oberoende variabler. Tabell 8 ger relativa vikter beräknade enligt Johnsons metod (Johnson, 2000, 2004; Johnson & LeBreton, 2004).

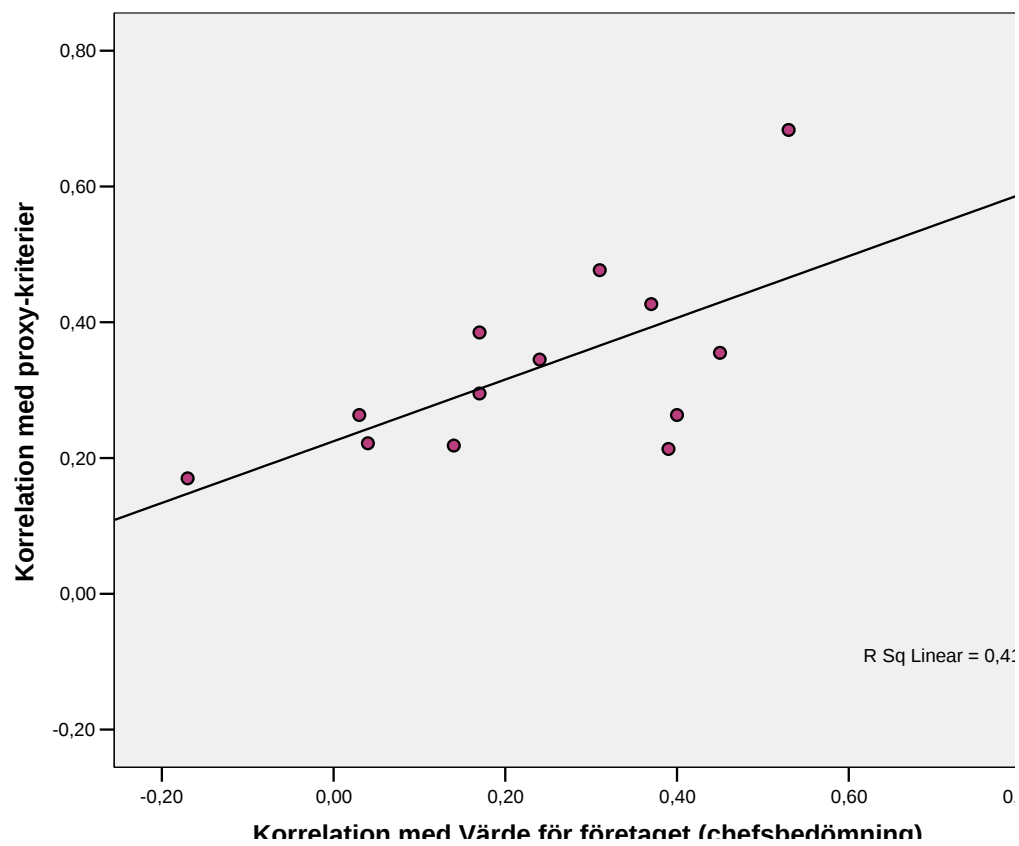
Tabell 8. Relativa vikter för testvariablerna vid förklaring av kriterierna. Varje vikt ger den andel av det totala förklaringsvärdet som testvariabeln svarar för. De två största värdena i varje kolumn anges i fetstil.

	Föränd-ringsvilja	Arbets-tillfreds-ställelse	Arbets-vilja	Balans	Arbets-intresse	Resultat-orientering
Följsamhet	1,3	1,7	1,9	5,9	1,3	1,6
EI	0,9	0,3	3,9	4,6	3,3	2,4
Öppenhet	12,3	0,5	3,8	1,7	2,4	2,3
Emotionell stabilitet	2,8	0,3	2,2	2,3	2,8	9,6
Extraversion	3,7	25,9	6,0	1,0	1,5	3,7
Perfektionism	20,0	0,9	8,9	2,9	3,8	12,0
Kreativitet	12,5	0,3	9,8	64,9	10,4	25,3
Uthållighet	1,5	0,2	0,5	5,1	66,6	1,7
Noggrannhet	2,7	2,2	9,0	1,4	0,9	1,0
Samarbetsvilja	3,5	65,6	28,3	7,1	0,5	18,9
Positiv attityd	24,9	1,4	2,9	1,3	5,4	10,6
Självförtroende	13,2	0,5	,6	1,3	0,9	10,7
Social förmåga	0,7	0,2	0,1	0,5	0,3	0,2

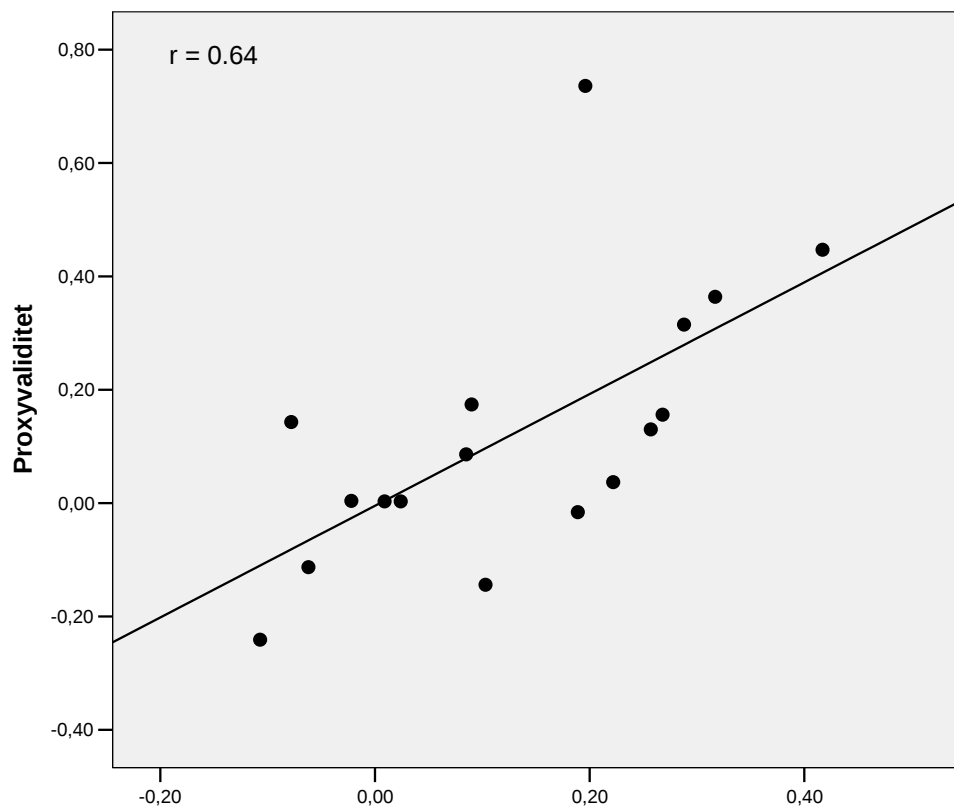
En övergripande bedömning visar att de viktigaste personlighetsdimensionerna var:

1. Samarbetsvilja
2. Kreativitet
3. Extraversion
4. Uthållighet
5. Positiv attityd
6. Perfektionism

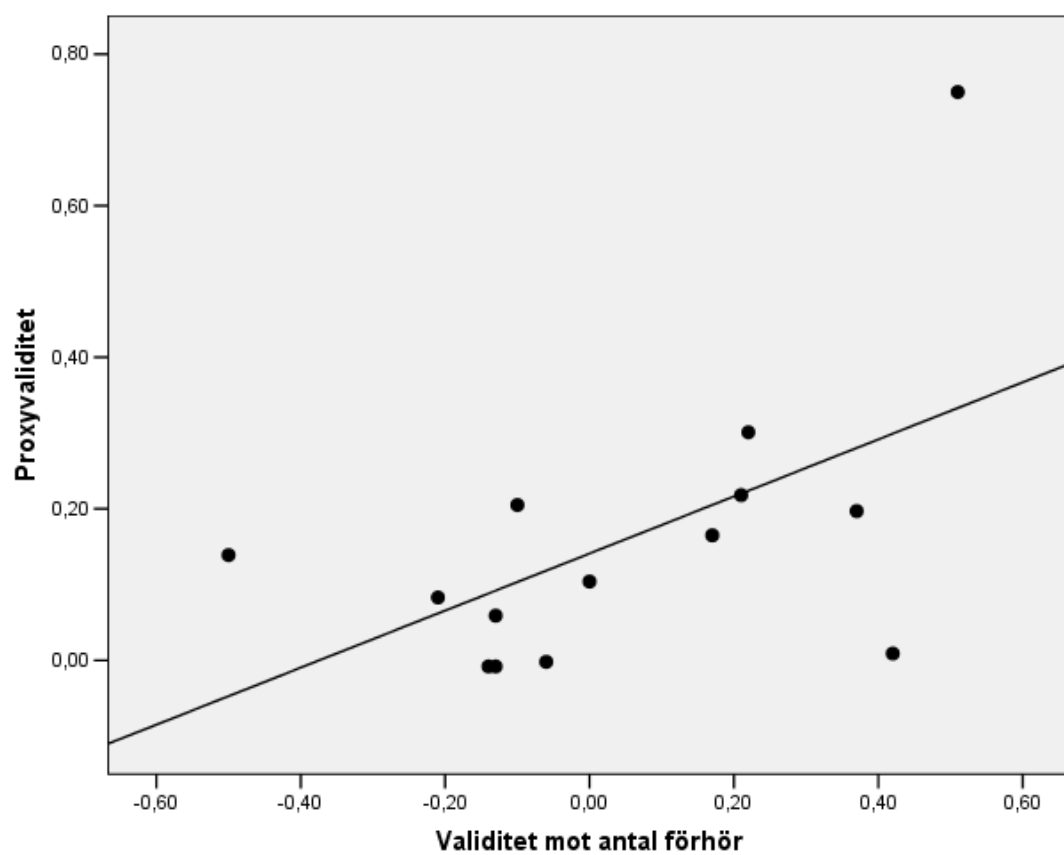
Resultaten avseende validering torde vara mycket stabila eftersom stickprovet var stort. Det är intressant att av de sex viktigaste variablerna var det endast en, extraversion, som ingår i FFM. I båda de fall där extraversion var bland de två bästa förklarande variablerna (arbetstillfredsställelse och arbetsvilja) var det en variabel bland dem som tillkom utöver FFM som var ännu bättre (samarbetsvilja). Figuren 12-14 visar samband mellan proxyvaliditeter och validiteter mot externa kriterier i 3 studier. Det framgår att proxyvaliditeter är informativa när det gäller de externa validiteter som kan förväntas när sådana data finns att tillgå.



Figur 12. Proxyvaliditeter och validiteter mot externt kriterium (kundtjänst vid ett finansbolag). (Sjöberg, 2010)



Figur 13. Proxyvaliditeter och validiteter mot externt kriterium (mellanchefer i sjukvården). (Sjöberg, Bergman, Lornudd, & Sandahl, 2011)



Figur 14. Proxyvaliditeter och validiteter mot externt kriterium (anställda vid polismyndigheter). (De Colli, 2011)