

Forskningsnotis

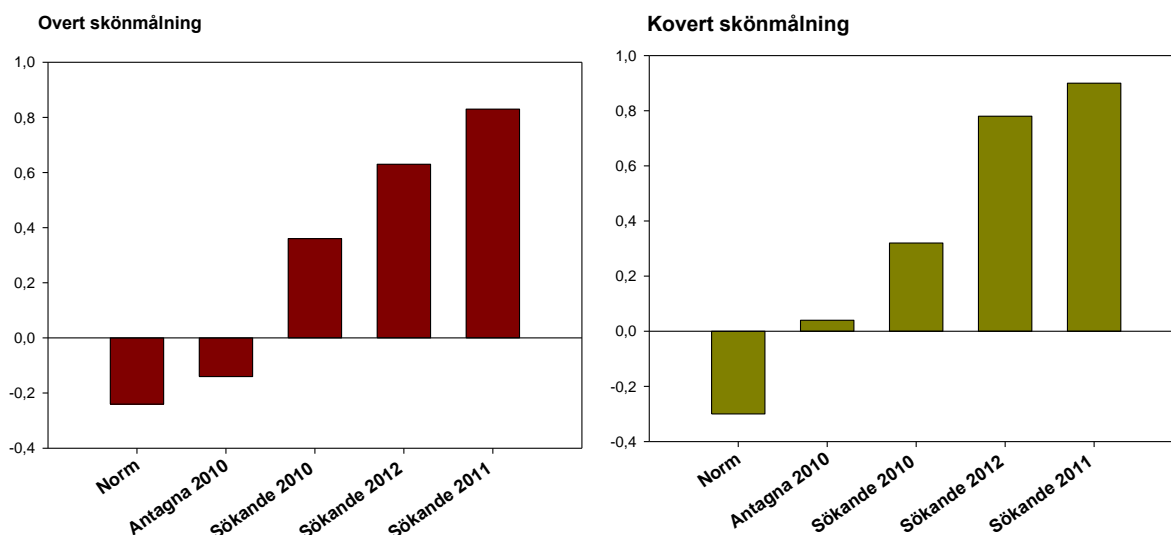
Skönmålning: överjagsstyrka, inställsamhet eller "Need for Approval"?

Lennart Sjöberg
Psykologisk Metod AB
2014-04-06

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 s a s 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 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, 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. 1.



Figur 1. 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 taktisksvar.

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

Datainsamling

De skalor som presenteras i denna forskningsnotis 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

Skalkonstruktioner

Överjagsstyrka

Tanken bakom att inkludera mätning av överjagsstyrka 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

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

moraliska värderingar och de borde kanske 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 en empirisk fråga i vilken mån överjagsstyrka 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. For example, for having extra-marital affairs." Se http://en.wikipedia.org/wiki/Id,_ego_and_super-ego

Efter litteratursökning kunde jag lokalisera en skala för mätning av överjagsstyrka, publicerad 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 överjagsstyrka, men hans personlighetstest har fått mycket starkt kritik². 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.

Överjaget kan illustreras av maktfullkomliga och dominerande politiker som de Gaulle:

² 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.

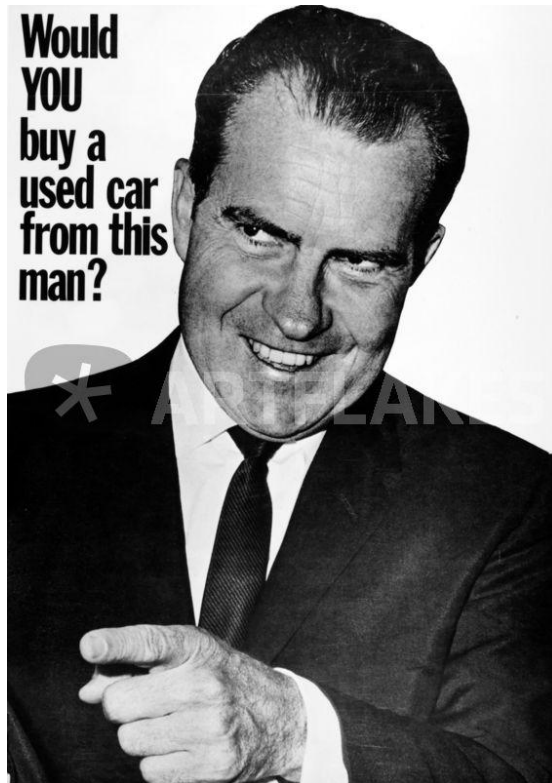


Figur 2. Överjaget personifierat.

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. En klassiker är en bild av Richard Nixon med rubriken "Would you buy a used car from this man"? Se Fig .2.



Figur 3. Klassisk Nixon-affisch som illustrerar inställsamhet

Inställsamhet kan eventuellt användas för att förbättra korrektionen för skönmålning. Som ett exempel har jag analyserat variabeln Beslutsamhet. Overt och kovert skönmålning för klarar 28 % av variansen i den variabeln, lägger man till inställsamhet får man ett signifikant tillskott på 3 %.

Behov att bli uppskattad ("Need for approval")

Nedanstående bild föreställer Karl XIV Johan, med Jakobinmössa. Hans valspråk som svensk kung var *Folkets kärlek, min belöning*, en god sammanfattning av "Need for Approval".



Figur 4. Karl XIV Johan, vars valspråk uttryckte önskan att vara omtyckt.

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. I denna rapport redovisas resultat av att relatera overt skönmålning till NA.

Övergripande analys

De två måtten på skönmålning relaterades till de tre skalorna överjagsstyrka, behov av att vara omtyckt (NA) och inställsamhet. Se Tabell 1.

Tabell 1. Resultatet av regressionsanalys av overt skönmålning som en funktion av överjagsstyrka, behovet av att vara omtyckt samt inställsamhet.

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

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

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

Diskussion

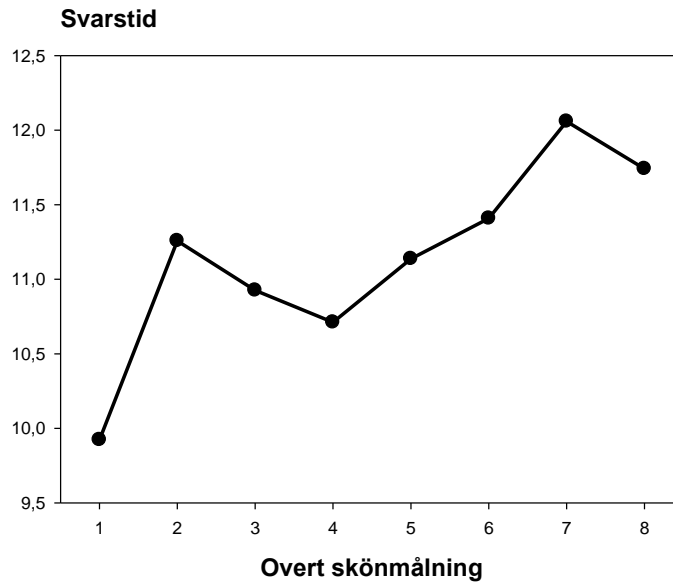
Betrakta lite närmare de typiska påståendena i Marlowe-Crownes skala, t ex "I sometimes think when people have a misfortune they only got what they deserved". Om man svarar "nej" på denna fråga bidrar det positivt till poängen i skalan. Samtidigt är det troligen ett ganska vanligt beteende. En möjlig 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 fråga om ett generellt och tydligt behov av att vara omtyckt*. 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 han eller hon (Higgins, 2003 #10937).

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 överjagsstyrka.

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, 2001 #10931), något som är att förvänta om SD-skolorna 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.



Figur 5. 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 taktiska 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 överjagsstyrka.

Referenser

- 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(3), 183-196.
- 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.
- 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. (1986). Response bias, social desirability and dissimulation. *Personality & Individual Differences*, 7(3), 385-400.
- Hough, L. M. (1998). The millennium for personality psychology: New horizons or good ole daze. *Applied Psychology: An International Review*, 47(2), 233-261.
- 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(6), 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(2), 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.
- 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.
- 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.
- Sjöberg, L. (2005). Självskattningstest: Den som ljuger fräckast vinner? (Self report tests: will the liar win?). *Personal och Ledarskap*, *36*(6), 22-25.
- Sjöberg, L. (2008). *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: Korrektion för skönmålning. (The UPP test: Correction for impression management)*. *Forskningsrapport 2009:3*. Stockholm: Psykologisk Metod AB.
- Sjöberg, L. (2012/2013). *Skönmålning i personlighetstest. SSE/EFI Working Paper Series in Business Administration, No. 2012:2*
- 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.
- 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*(3), 281-285.
- 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.
- 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 taktiskskalor³ 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

³ Termen taktiskskala 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 taktiksvar 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 taktiksvar. 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 overt 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 overta taktikskalan. På så sätt får man en möjlighet att mäta taktiksvar som inte kan avslöjas av sofistikerade testtagare. De två taktikskalorna (overt och kovert) har begreppsvaliderats, se Tabell 7.1. Båda korrelerade högt med Crowne-Marloweskalen och högt sinsemellan. Den senare korrelationen var något lägre vilket är rimligt.

Tabell 2. Korrelationer mellan de tre taktikskalorna (N=159), data från Studie 1.

	Crowne-Marlowes taktikskala	Overt taktikskala	Kovert taktikskala
Crowne-Marlowes taktikskala	1,00	0,76	0,73
Overt 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 korrigerar 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 taktiksvarstendens 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 6.

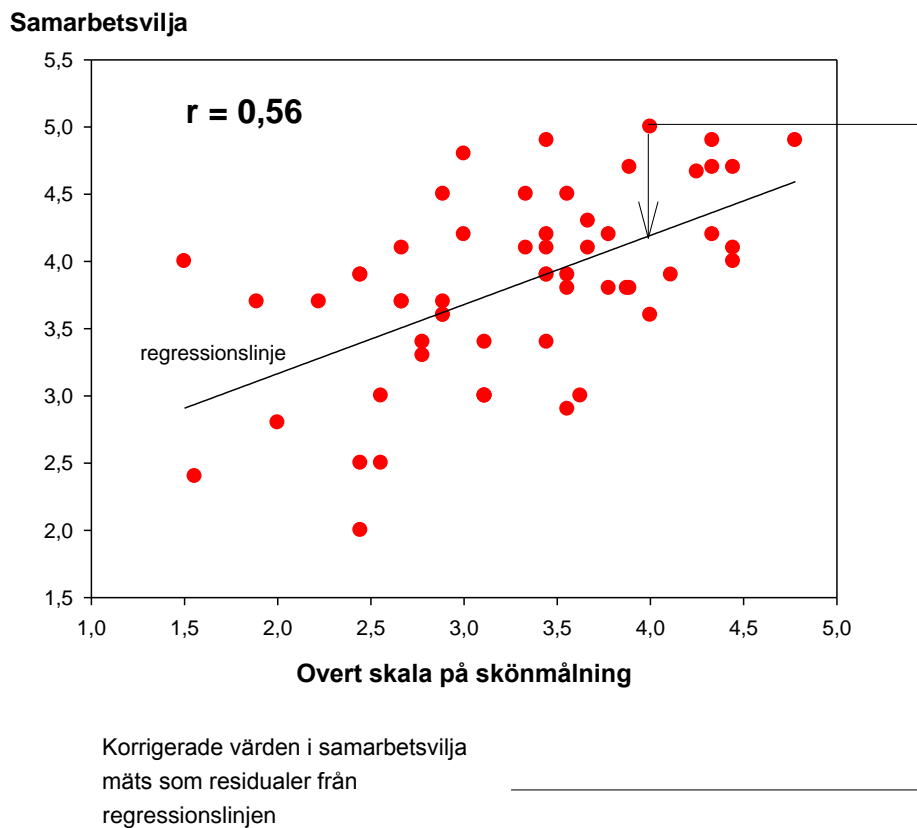
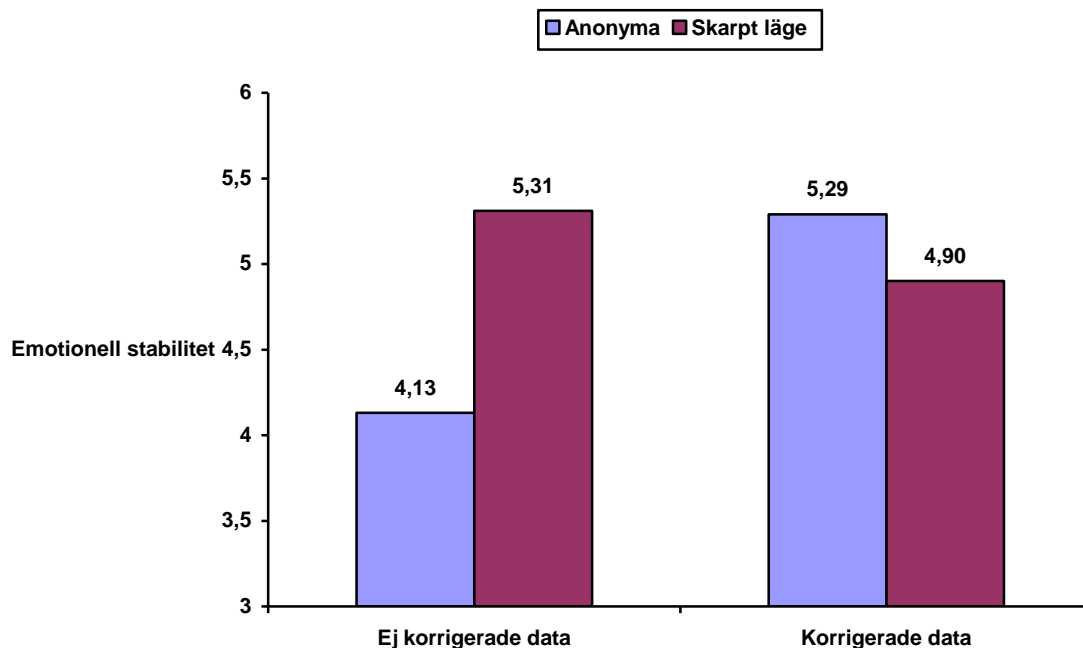


Fig. 6. Illustration av hur korrigerade värden i ett test uppskattas med hjälp av en psykometrisk modell.

Denna metodik utnyttjar taktikskalan maximalt. Effekten av ett starkt incitament att svara taktiskt nästan helt försvinner med denna typ av korrektion (Sjöberg, 2008; 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 Figur 7 som visar att skillnaden i okorrigerade värden starkt reduceras när vi genomfört korrektionen för skönmålning.

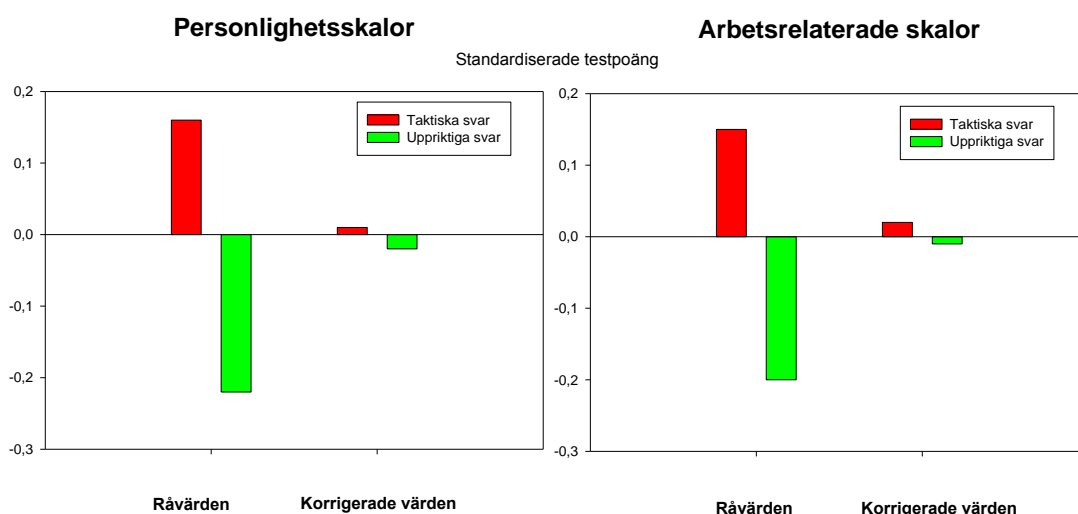


Figur 7. Effekten av korrektion för skönmålning vid jämförelse mellan anonym testning och skarpt läge, emotionell stabilitet. Medelvärdet=5, standardavvikelsen=2.

Fig.8 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:

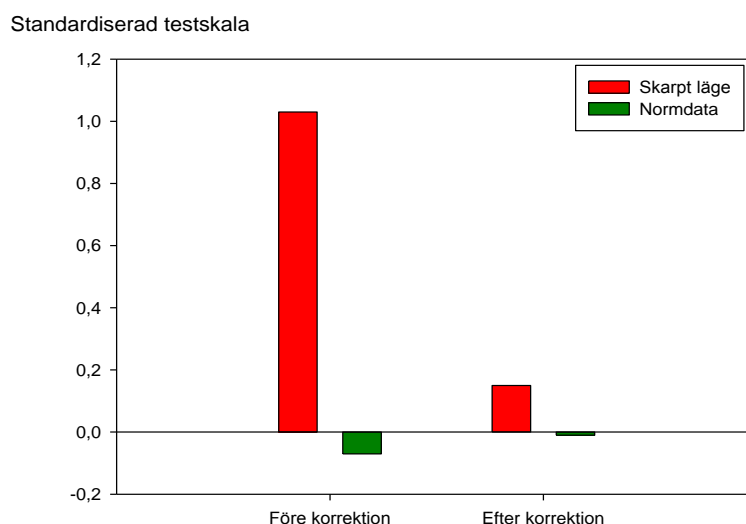
”Tänk på ett jobb som du mycket gärna skulle vilja ha. Tänk dig nu att du sökt just det jobbet och att den här testningen är ett *mycket viktigt* led i anställningsförfarandet. Svara på testuppgifterna så att framstår som en person som är precis sådan som man söker till jobbet. Det kanske innebär att 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. 8.



Figur 8. Genomsnittliga effekter av korrektion för skönmålning, standardiserade skalor.

Delar av *UPP*-testet (screening- modulen) används i rekryteringsplattformen Katapult⁴. De som där tar testet gör det i skarpt läge, dvs. i samband med att de söker ett jobb. Se Fig. 9 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 korrigerar på det sätt som görs i *UPP*-testet.



Figur 9. 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).

⁴ www.katapult.com

Appendix 2. Begreppsförklaringar: nya skalor

Begreppsförklaringar <i>UPP</i> -testet: Beslutsamhet och värderingar	
Testskala	Mycket högt värde (>7) utmärks av de flesta eller alla av nedanstående beteenden:
<i>Överjagsstyrka</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