



PSYKOLOGISK METOD AB



UPP-testet: korrektion för skönmålning

Lennart Sjöberg
Rapport 2009:3

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.

Rapporter utgivna av Psykologisk Metod AB

Sjöberg, L. (2008). Bortom Big Five: Konstruktion och validering av ett personlighetstest. Rapport 2008:1.

Sjöberg, L. (2008). *Big Five Plus*. Manual.

Sjöberg, L., & Möller, K. (2009). Sociala arbetsfunktioner och personlighet. Rapport 2009:1.

Sjöberg, L. (2009). *UPP*-testet: Kriterierelaterad validitet. Rapport 2009:2.

Sjöberg, L. (2009). *UPP*-testet: Korrektion för skönmålning. Rapport 2009:3.

Sammanfattning¹

UPP-testets metodik för att korrigera testvärden för skönmålning undersöktes i fyra delstudier. Testning med *UPP*-testets screeningmodul i skarpt läge jämfördes med data insamlade i ”oskarpt” läge. Skillnaderna var stora mellan grupperna men ca 82 % av dem försvann till följd av korrektionen. Nya mått av potentiellt intresse för mätning av skönmålning togs fram i en särskild delstudie: onyanserat accepterande av påståenden med positivt innehåll och motsatsen, alltså onyanserat avvisande av påståenden med negativt innehåll. Måtten på skönmålning liksom måttet på onyanserat avvisande av negativa påståenden differentierade starkt mellan en grupp om tog hela *UPP*-testet i skarpt läge och normdatagruppen. Fallstudier illustrerade därefter effekterna av korrektion på individuell nivå. Dessa effekter kunde vara mycket stora och de varierade över testvariabler. Slutligen analyserades sambanden mellan testvariabler och 3 mått på datakvalitet som används av *UPP*-testet: jasägartendens, intra-individuell variation samt strukturlikhet. Den senare variabeln är ett mått på hur stor likhet en persons svar har med profilen av medelvärden i en grupp av testade personer. Testvärdena före korrektion hade ganska höga samband med kvalitetsmått vilket tyder på en snedvridande effekt av ett normativt svarsformat. Efter korrektion för skönmålning försvann ca 2/3 av dessa effekter. Slutsats: *UPP*-testets metodik för korrektion för skönmålning ger betydligt mera rättvisande testvärden än de okorrigerade testvärdena. I framtida forskning kan det visa sig möjligt att mäta skönmålning utan att använda särskilda skalor.

Inledning

Skönmålning är ett stort problem vid användning av personlighetstest som bygger på självrapportuppgifter, särskilt vid testningar i skarpt läge. Det är lätt att förstå varför så är fallet. Vid t ex testningar inför en anställning vill många av de testade personerna framstå i så positiv dager som möjligt. De ger en skönmålad bild av sig själva. De försöker göra det på många sätt, också när de besvarar testfrågor. Omfattande forskning har visat att många, troligen alla, självrapporttest påverkas av denna felkälla (Furnham, 1986; Hough, 1998). Hur stora problemen är beror på ett antal detaljer, främst kanske vilket svarsformat som används i testet och om en effektiv metod för korrektion av effekterna av skönmålning används. På den individuella nivån kan effekterna av skönmålning bli enorma.

Självrapport-test är oftast ganska lätta att genomskåda. Frågan är vad resultaten på ett sådant test egentligen är värda, om svaren på testet är starkt korrelerade med skönmålning, och kraftigt påverkas av vilken betydelse testningen förväntas ha.

Det dominerande sättet att hantera problemet tycks vara att använda ipsativa (jämförande) svarsformat. När man använder ett sådant format uppmanas den testade att välja vilket av två eller flera svarsalternativ om passar bäst på honom eller henne (ibland också vilket som passar sämst). Ipsativa format har emellertid allvarliga nackdelar:

¹ Författaren är tacksam till Per Näsman, KTH, för granskning av den statistiska metodiken, samt till Hampus Breistein, Katapult AB, för framtagning av testdata med screeningmodulen i skarpt läge.

- Data kan inte analyseras med standardmetoder för uppgiftsanalys och faktoranalys (Meade, 2004). Detta beror på att skalorna som härleds ur svaren blir matematiskt beroende – alla testade erhåller samma sammanlagda poäng.
- De testade tycker illa om formatet. Det känns onaturligt och obehagligt att välja mellan ”päron och äpplen”.
- Forskning visar att skönmålningstendensen inte elimineras av det ipsativa formatet (Martin, Bowen, & Hunt, 2002), åtminstone inte helt.
- Testresultaten kan inte jämföras *mellan* utan bara *inom* individer. Man kan alltså t ex inte säga om Kalle är mera kreativ än Pelle, bara att Kalle är mera kreativ än han är uthållig.

Motsatsen till ipsativa test är normativa test. *UPP*-testet är ett exempel på ett normativt test. Här besvaras varje testuppgift för sig, t ex avseende hur bra ett visst påstående stämmer in som beskrivning av den testades personlighet, beteende eller känsloliv och tankar. Oftast används en skala i flera steg, en Likert-skala, som medger att den testade kan markera hur väl påståendet stämmer. Flerstegsskalor av denna typ har stora psykometriska fördelar framför tvåstegsskalor av typ stämmer/stämmer inte.

Normativa format har också nackdelar:

- Skönmålning påverkar svaren.
- Även andra snedvridande tendenser kan ha effekt, nämligen jasägartendensen och tendensen att ge alltför lite varierande bedömningar.

Det är en empirisk fråga om det normativa formatet fungerar bra efter korrektion för skönmålning, trots de potentiella nackdelarna. Det är huvudfrågeställningen i denna rapport.

Skalor för mätning av skönmålning

Två skalor för mätning av skönmålning har utvecklats och används i *UPP*-testet. Traditionellt mäter man 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 i viss mån taktiska och tillrättalagda för att ge ett positivt intryck.

Ett viktigt problem är frågan om uppgifterna i en skönmålningsskala går att genomskåda. 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 testuppgifter kan skapa misstänksamhet, och de testade personer som förberett sig genom att läsa på om testmetodik borde ganska lätt kunna identifiera vilka uppgifter det rör sig om.

I arbetet med *UPP*-testet har konstruerats en skala utifrån samma principer som Crowne-Marlowe, en overt skala (öppen mätning) och 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 de korrelerade högt med den overta skalan. På så sätt får man en möjlighet att mäta skönmålningstendens som inte kan avslöjas av sofistikerade testtagare. De två skalorna (overt och kovert) har begreppsvaliderats, se Tabell 1. Båda korrelerade hög med Crowne-Marloweskalen och högt sinsemellan. Den senare korrelationen var något lägre.

Tabell 1. Korrelationer mellan de tre taktikskalorna (N=159), data från Studie 1 i den tekniska rapporten.

| | Crowne-Marlowes skala | Overt skala | Kovert kala |
|-----------------------------|-----------------------|-------------|-------------|
| Crowne-Marlowes taktikskala | 1.00 | 0.76 | 0.73 |
| Overt skala | 0.76 | 1.00 | 0.56 |
| Kovert skala | 0.73 | 0.56 | 1.00 |

Den overta skalan har en reliabilitet i normmaterialet (N=1155) på 0.73, den koverta något lägre, 0.67. Korrelationen mellan de två skalorna i normmaterialet var 0.51, vilket motsvarar en korrelation på 0.70 mellan de sanna värdena. Det finns m a o en hög grad av begreppslig konvergens mellan dem. Kontrollstudier har dessutom visat att korrektion enbart med den koverta skalan ger ett mycket likartat resultat jämfört med korrektion med båda skalorna.

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

1. 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".
2. 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 att de är alltför subjektiva (Metod 1) eller ineffektiva (Metod 2). 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.

Båda 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 tendensen att skönmåla kan ju vara mycket varierande beroende på vilken dimension det är som är aktuell.

Den metodik som används i *UPP*-testet utnyttjar effekten av taktikskalorna fullt ut med hjälp av en regressionsmodell och beräkning av residualer, se Fig.1.

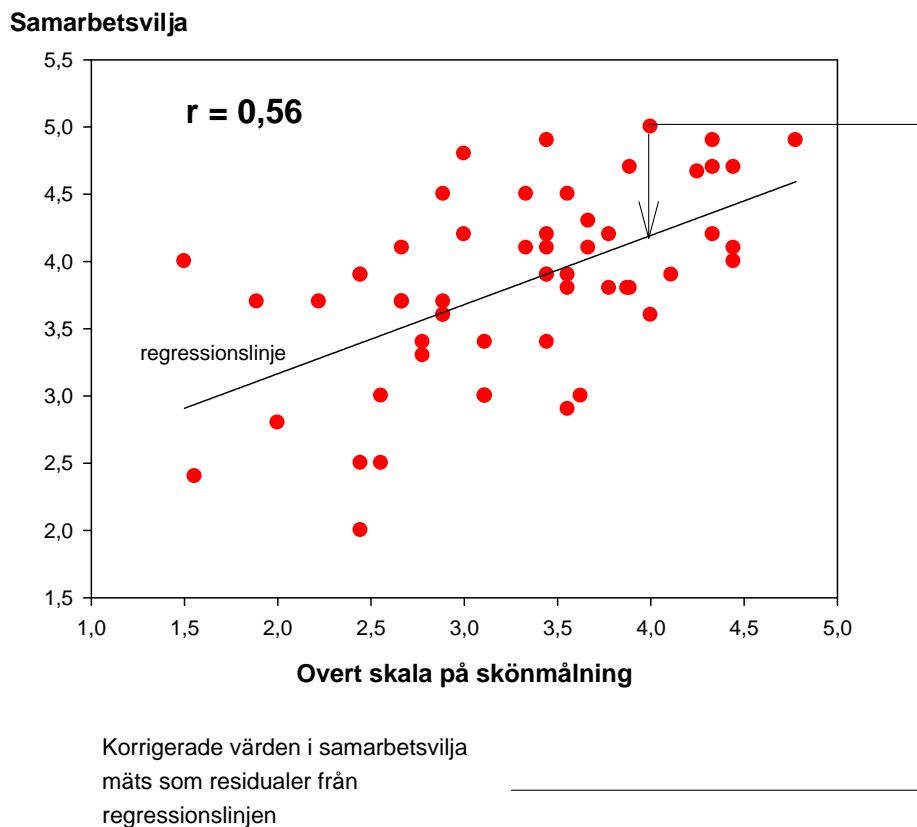
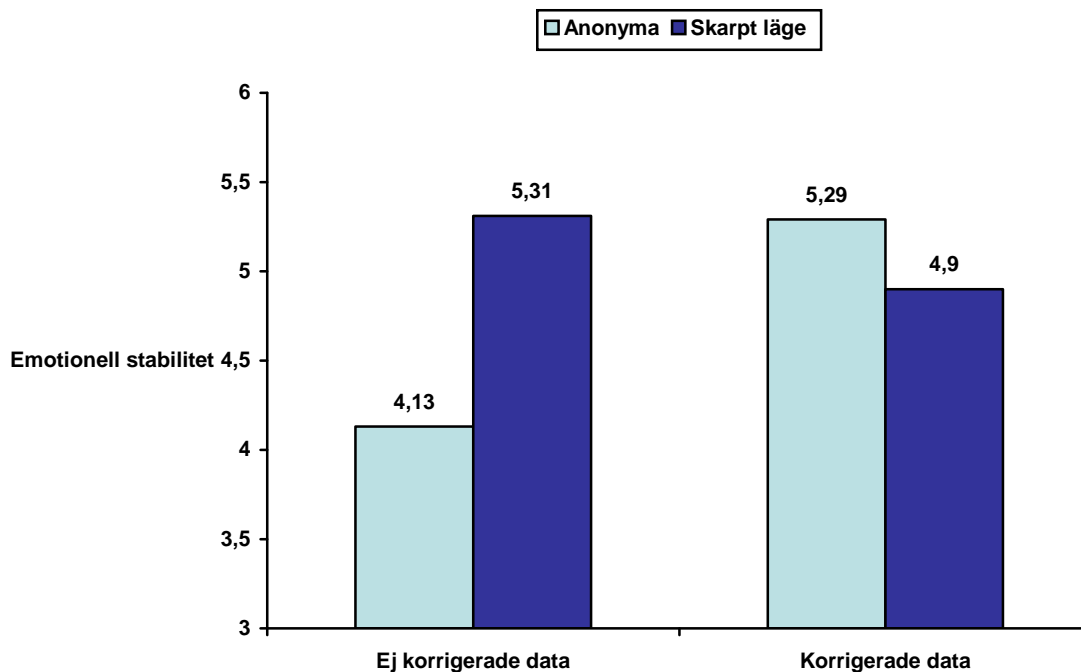


Fig. 1. Illustration av hur korrigerade värden i ett test uppskattas med hjälp av en psykometrisk modell.

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

Tekniken har 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. 2 som visar hur en skillnad i okorrigerade värden försvann när korrektion för skönmålning infördes.



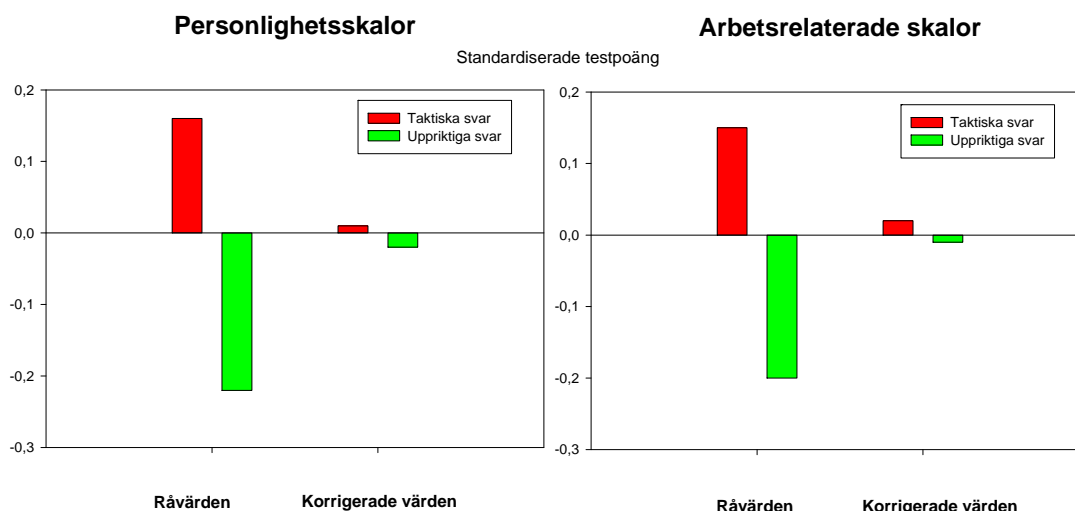
Figur 2. Effekten av korrektion för skönmålning vid jämförelse mellan anonym testning och skarpt läge, emotionell stabilitet.

Fig. 3 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. 3.



Figur 3 Genomsnittliga effekter av korrektion för skönmålning, standardiserade skalor.

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 att gynnas av att man genomför en korrektion (Sjöberg, 2005).

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, men detta är inte alltid sant (Rogosa & Willett, 1983). Analyser av samtliga indexvariabler i testet visade emellertid att det blev en viss sänkning av reliabiliteten efter korrektion 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 troligen 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

Den forskning som sammanfattats ovan ger ett starkt stöd för den använda metodiken för att korrigera för skönmålning. Men saken är mycket viktig och det ansågs angeläget att föra forskningen vidare genom att replikera tidigare resultat och genom en bredare analys av felfaktorerna vid normativt testformat. Den aktuella debatten om ipsativa och normativa format är fortfarande intensiv och det finns anhängare till ipsativa format bland forskare och psykometriker (Brown & Bartram, 2009).

En undersökning av *UPP*-testets mått på datakvalitet aktualiserades av denna debatt om nackdelar med normativa svarsformat. Jasägartendensen föreföll problematisk när

testuppgifterna kunde ha antingen positivt eller negativt innehåll. En särskild studie genomfördes därför av den problematiken.

Syftena med denna rapport är:

- att ytterligare undersöka giltigheten hos korrektionen för skönmålning, denna gång i en fältstudie där en delgrupp bestod av personer som tog testet som en del av en jobbansökan
- att närmare undersöka ”jasägande” i förhållande till testuppgifter med positivt eller negativt innehåll separat
- att visa de konkreta effekterna av korrektionen i några utvalda individuella fall
- att undersöka om denna korrektion även motverkar andra snedvridande svarstendenser i ett normativt test

I rapporten redovisas fyra delstudier:

- Effekten av korrektion för skönmålning vid jämförelse mellan normmaterialet och UPP-testets screeningmodul i skarpt läge. Normmaterialet insamlades i testsituationer som inte var tydligt relaterade till för de testade viktiga beslut.
- Analys av acceptans av positiva påståenden och avvisande av negativa påståenden. Skarpt läge jämfördes med normgruppen.
- Beskrivning av effekter av skönmålning och korrektion på individuell nivå, normmaterialet.
- Effekten av korrektion för skönmålning på vissa andra felfaktorer vid normativ mätning. Denna analys görs på normmaterialet.

Effekten av korrektion för skönmålning vid jämförelse mellan normmaterialet och screeningmodulen i skarpt läge

De data som analyseras i denna delstudie kommer dels från normmaterialet, insamlat i oskarpt läge, dels från data i skarpt läge (jobbsökande) som besvarade screeningmodulen som ett led i en jobbansökan i rekryteringsplattformen Katapult. De variabler som ingår i screeningmodulen är:

- Extraversion
- Uthållighet
- Samarbetsvilja
- Positiv grundattityd
- Kreativitet
- Kovert skönmålningskala

Värdena i dessa testvariabler kommer att jämföras före och efter korrektion för skönmålning. Hypotesen är att det finns skillnader mellan grupperna före korrektionen och att dessa försvinner eller åtminstone starkt minskas genom korrektionen.

Antalet analyserade testade personer var 951 i normmaterialet och 1251 i gruppen av jobbsökande.

Tabell 2 visar andelen varians i testvariablerna som förklaras av tendensen till skönmålning.

| Testvariabel | Andel av skönmålning förklarad varians |
|----------------------|--|
| Extraversion | 0.213 |
| Uthållighet | 0.398 |
| Samarbetsvilja | 0.419 |
| Positiv grundattityd | 0.381 |
| Kreativitet | 0.089 |

Denna tabell visar på två viktiga förhållanden:

- Skönmålning svarar för en stor del av variansen i de flesta testvariablerna
- Andelen förklarad varians varierar med ett mycket lågt värde för en testvariabel (kreativitet)

Medelvärden och standardavvikelser för skönmålningsvariabeln framgår av Tabell 3.

| Grupp | Medelvärde | Standardavvikelse |
|----------------|------------|-------------------|
| Normmaterialet | 3.89 | 0.451 |
| Jobbsökande | 4.40 | 0.342 |

Skillnaden mellan de två grupperna i skönmålningstendens var, som framgår av tabellen, mycket stor: $d = 1.04$, $p < 0.0005$. Måttet d avser standardiserade värden. De två delgrupperna hade standardiserats i det sammanslagna materialet.

Skillnader i standardiserade värden är ett vanligt effektmått. Man brukar, i anslutning till Cohen (Cohen, 1988), tala om:

- 0. 20 – 0.49: liten effekt
- 0. 50 – 0.79: moderat effekt
- 0. 80 och större: stor effekt

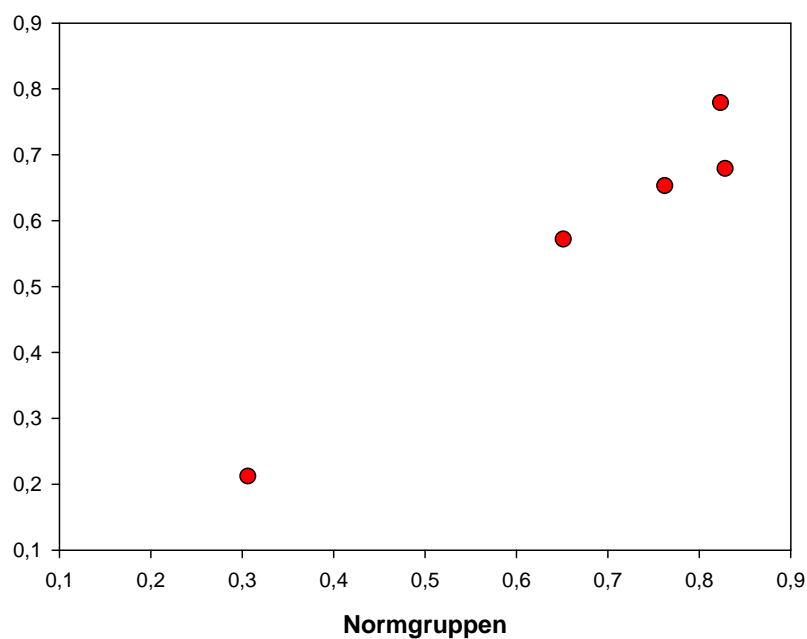
I beteendevetenskaplig forskning är stora effekter ganska sällsynta. Märk att storleken på effekten inte är beroende av stickprovsstorleken vilket är fallet med den vanliga signifikansnivån.

Den fortsatta analysen förutsätter att sambanden mellan testvariabler och måttet på skönmålning är lika starka i båda grupperna. Mera tekniskt är det önskvärt att regressionskoefficienterna är lika (samma lutningskoefficient). Tabell 4 visar korrelationerna, standardavvikelseerna och regressionskoefficienterna i var och en av grupperna.

| Testvariabel | Normgruppen | | | Jobsökande | | | Hela materialet |
|----------------------|---|-------------------|------------------------|---|-------------------|------------------------|------------------------------|
| | r mellan testvariabel och skönmålning | Standardavvikelse | Regressionskoefficient | r mellan testvariabel och skönmålning | Standardavvikelse | Regressionskoefficient | Skillnad kolumn 4 - kolumn 7 |
| Extraversion | 0.393 | 0.747 | 0.652 | 0.353 | 0.554 | 0.571 | 0.081 |
| Uthållighet | 0.565 | 0.609 | 0.763 | 0.505 | 0.442 | 0.652 | 0.111 |
| Samarbetsvilja | 0.589 | 0.630 | 0.824 | 0.596 | 0.447 | 0.778 | 0.046 |
| Positiv grundattityd | 0.561 | 0.665 | 0.829 | 0.516 | 0.450 | 0.678 | 0.151 |
| Kreativitet | 0.245 | 0.564 | 0.307 | 0.173 | 0.419 | 0.211 | 0.096 |

Tabellen visar att regressionskoefficienterna var något lägre i gruppen av jobsökande än i normgruppen. En prickning av dessa koefficienter, se Fig. 4, visar emellertid på mycket likartade egenskaper hos data i de två grupperna

Jobsökande



Figur 4. Regressionskoefficienter för de två grupperna prickade mot varandra.

Tabell 5 visar skillnaderna mellan normgruppen och gruppen i skarpt läge före och efter korrektion för skönmålning.

Tabell 5. Medelvärden i grupperna före och efter korrektion för skönmålning. Alla värden standardiserade till z-form.

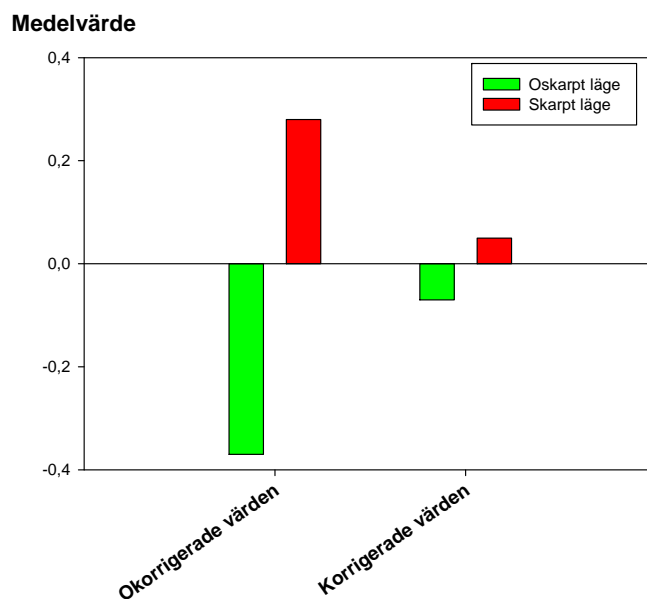
| Testvariabel | Okorrigerade värden | | | Korrigerade värden | | |
|----------------------|---------------------|-------------|---|--------------------|-------------|---|
| | Oskarpt läge | Skarpt läge | <i>p</i> för skillnaden oskarpt – skarpt läge enligt <i>t</i> -test | Oskarpt läge | Skarpt läge | <i>p</i> för skillnaden oskarpt – skarpt läge enligt <i>t</i> -test |
| Extraversion | -0.34 | 0.26 | *** | -0.08 | 0.06 | *** |
| Uthållighet | -0.46 | 0.35 | *** | -0.12 | 0.09 | *** |
| Samarbetsvilja | -0.37 | 0.28 | *** | 0.01 | -0.01 | Ej sign. |
| Positiv grundattityd | -0.41 | 0.31 | *** | -0.05 | 0.04 | * |
| Kreativitet | -0.27 | 0.20 | *** | -0.10 | 0.07 | *** |

*** $p < 0.001$

* $p < 0.05$

Tabellen visar på stora skillnader mellan grupperna före korrektion, och dramatiskt mindre skillnader efter korrektion. Det genomsnittliga spannet mellan oskarpt och skarpt läge var 0.65 före korrektion och 0.12 efter korrektion, alltså en reduktion av 82 % av effekten av skönmålning.

Fig. 5 visar de genomsnittliga värdena före och efter korrektion. Som framgår av Tabell 4 och Fig. 4 fanns det skillnader mellan regressionskoefficienterna mellan de två grupperna. För samarbetsvilja och extraversion var de emellertid små. Samtidigt har korrektionen haft liknande effekter i alla fem fallen varför vi av detta skäl kan anta att variationen i regressionskoefficienter inte allvarligt påverkat resultaten. Möjligen kan man anta att de kvarstående effekterna i viss mån beror på att regressionskoefficienterna inte helt betett sig som de ska för att antagandena ska vara helt uppfyllda. En rimlig slutsats är att effekten av korrektionen underskattats och att den alltså är minst så stor som framgår av Tabell 5. Det kan tilläggas att metodiken torde vara tämligen okänslig för avvikelser från antagandet om lika regressionskoefficienter (Kleinbaum, Kupper, Nizam, & Muller, 2007; Montgomery, 2009).



Figur 5. Genomsnittliga värden i personlighetsvariablerna före och efter korrektion för skönmålning, skarpt och oskarpt läge.

Figuren visar att det fanns en mycket stor skillnad mellan grupperna före korrektionen för skönmålning samt att denna skillnad i stort sett eliminerades av korrektionen för skönmålning.

Slutsatser

Denna delstudie visar:

- Skönmålning var stark och tydlig i gruppen av jobbsökande
- Skönmålningen varierade i styrka för olika testvariabler
- Korrektion för skönmålning enligt den metodik som används i *UPP*-testet eliminerade en mycket stor del av effekten

Detta innebär att vi lyckats replikera tidigare studier av effekten av korrektion för skönmålning, denna gång med två mycket stora stickprov och med data insamlade i skarpt läge från faktiska jobbsökande. Studien gällde screeningmodulens fem personlighetsvariabler men det finns ingen anledning att tro att resultaten skulle ha blivit annorlunda för hela *UPP*-testet. Tidigare studier har visat att effekterna var generella över alla personlighetsvariabler och även de arbetsrelaterade måtten på motivation och attityder.

Skönmålning som onyanserad acceptans/onyanserat avvisande

En annorlunda ansats till mätning av skönmålning är att ta fram den genomsnittliga bedömningen av testuppgifter med positivt innehåll ("Jag har ett lugn och jämt humör") minus den genomsnittliga bedömningen av uppgifter med negativt innehåll ("Jag brukar ge upp ganska snabbt efter ett misslyckande"). På så sätt skapas ett index som mäter hur onyanserade svaren är. Ett högt värde visar att den testade personen tycks reagera huvudsakligen på om testuppgiften säger något positivt eller negativt varefter han eller hon

väljer ett accepterande (för de positiva) eller avvisande (för de negativa). I dessa två index ingick inte de testuppgifter som mätte overt och kovert skönmålning.

Betydelsen av att en person svarar i hög grad ”svart-vitt” kan diskuteras. Eftersom det handlar om positiva respektive negativa självrapporter är det emellertid rimligt att anta att denna form av bristande nyansering är ett uttryck för viljan att skönmåla.

Sambandet mellan detta mått på bristande nyanser och de två måtten på skönmålning i normmaterialet framgår av Tabell 6. Två mått på bristande nyansering har tagits fram: för de positivt och för de negativt formulerade uppgifterna.

| Tabell 6. Korrelationer mellan bristande nyansering och skönmålning. | | | | |
|--|---|--|-------------------|--------------------|
| | Tendens att acceptera positiva påståenden | Tendens att avvisa negativa påståenden | Overt skönmålning | Kovert skönmålning |
| Tendens att acceptera positiva påståenden | 1.00 | 0.40 | 0.41 | 0.34 |
| Tendens att avvisa negativa påståenden | 0.40 | 1.00 | 0.71 | 0.48 |
| Overt skönmålning | 0.41 | 0.71 | 1.00 | 0.51 |
| Kovert skönmålning | 0.34 | 0.48 | 0.51 | 1.00 |

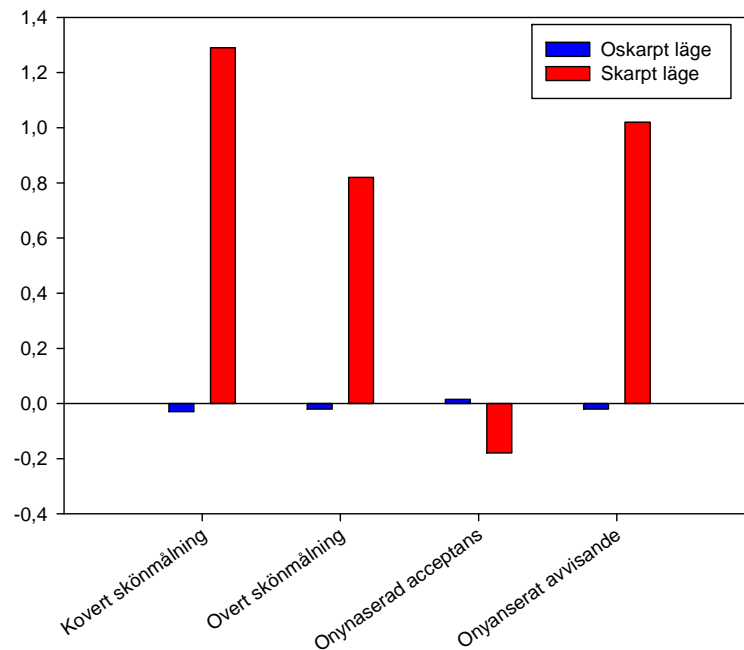
Den multipla korrelationen mellan tendensen att acceptera positiva påståenden och de två måtten på skönmålning av 0.44. Motsvarande värde för tendensen att avvisa negativa påståenden var 0.72.

Tabellen visar att

- Det fanns en tendens att svara onyanserat, ”svart-vitt”, som var gemensam för påståenden med positivt och negativt innehåll
- Denna tendens hade höga korrelationer med måtten på skönmålning
- Särskilt tendensen att avvisa påståenden med negativt innehåll skulle kanske kunna ersätta de två måtten på skönmålning – och tolkas som ett mått på skönmålning, ännu mera ”dolt” än det koverta måttet

Den sistnämnda möjligheten får stöd av en analys av normdatagruppen mot 21 sökande som tagit hela *UPP*-testet i skarpt läge. Medelvärden återges i Fig. 6.

Medelvärde, standardiserade data



Figur 6. Medelvärden av mått på skönmålning, onyanserat acceptande av positiva påståenden samt onyanserat avvisande av negativa påståenden.

Figuren visar tydligt att det fanns stora skillnader i förväntad riktning mellan skarpt och oskarpt läge i de två måtten på skönmålning, men också med avseende på onyanserat avvisande av negativa påståenden. Alla de nämnda skillnaderna var statistiskt signifikanta. För onyanserat acceptande av positiva påståenden fanns däremot ingen signifikant skillnad. Fortsatt forskning om framför allt onyanserat avvisande av negativa påståenden vore av intresse.

Det bör understrykas att de två måtten på onyanserat acceptande respektive avvisande av påståenden inte sammanfaller med jasägartendensen, varken logiskt eller psykologiskt. Jasägande mäts som den genomsnittliga bedömningen av alla påståenden oberoende av om de har positivt eller negativt innehåll.

Individuella effekter av skönmålning

Tre fall har utvalts ur normdatabasen för att illustrera effekterna av korrektion för skönmålning på individuell nivå. De varierade starkt i nivån på skönmålning.

Fall 1: Hög skönmålning

| Skönmålning | |
|---------------|---|
| <i>Overt</i> | 9 |
| <i>Kovert</i> | 9 |

| Personlighetsdimensioner av särskild vikt i arbetslivet | | |
|---|------------------|---------------------|
| | Korrigerat värde | Ej korrigerat värde |
| <i>Samarbetsvilja</i> | 5 | 9 |
| <i>Kreativitet</i> | 4 | 5 |
| <i>Emotionell intelligens</i> | 4 | 6 |
| <i>Social förmåga</i> | 5 | 8 |
| <i>Grad av positiv grundattityd</i> | 5 | 9 |
| <i>Uthållighet</i> | 5 | 8 |
| <i>Självförtroende</i> | 6 | 9 |
| <i>Perfektionism</i> | 8 | 8 |

| Dimensioner enligt femfaktormodellen ("Big Five") | | |
|---|------------------|---------------------|
| | Korrigerat värde | Ej korrigerat värde |
| <i>Utåtvändhet</i> | 5 | 7 |
| <i>Emotionell stabilitet</i> | 5 | 8 |
| <i>Öppenhet</i> | 5 | 7 |
| <i>Vänlighet</i> | 6 | 9 |
| <i>Noggrannhet</i> | 8 | 9 |

| Arbetsmotivation och arbetsattityder | | |
|---|-------------------------|----------------------------|
| | Korrigerat värde | Ej korrigerat värde |
| <i>Arbetsvilja</i> | <i>6</i> | <i>9</i> |
| <i>Arbetsintresse</i> | <i>3</i> | <i>5</i> |
| <i>Arbetsstillfredsställelse</i> | <i>5</i> | <i>7</i> |
| <i>Resultatorientering</i> | <i>7</i> | <i>9</i> |
| Förändringsvilja | <i>7</i> | <i>9</i> |
| <i>Balans arbete – övrigt liv</i> | <i>6</i> | <i>9</i> |

Kommentar: De korrigerade värdena ligger i detta fall i allmänhet betydligt under motsvarande värden före korrektion, upp till 4 stanine-poäng lägre. Samtidigt kan man konstatera att i vissa fall har korrektionen ingen eller en liten effekt. Detta beror på att skönmålningen hade låg korrelation med testvariablerna i just de fallen. Att behovet av att justera testvärdena olika mycket för olika testvariabler är uppenbart.

Fall 2: medelhög skönmålning

| Skönmålning | |
|--------------------|----------|
| <i>Overt</i> | <i>5</i> |
| <i>Kovert</i> | <i>5</i> |

| Personlighetsdimensioner av särskild vikt i arbetslivet | | |
|--|-------------------------|----------------------------|
| | Korrigerat värde | Ej korrigerat värde |
| <i>Samarbetsvilja</i> | <i>6</i> | <i>6</i> |
| <i>Kreativitet</i> | <i>3</i> | <i>3</i> |
| <i>Emotionell intelligens</i> | <i>4</i> | <i>4</i> |
| <i>Social förmåga</i> | <i>4</i> | <i>4</i> |
| <i>Grad av positiv grundattityd</i> | <i>3</i> | <i>3</i> |
| <i>Uthållighet</i> | <i>6</i> | <i>6</i> |
| <i>Självförtroende</i> | <i>5</i> | <i>5</i> |
| <i>Perfektionism</i> | <i>5</i> | <i>5</i> |

| Dimensioner enligt femfaktormodellen ("Big Five") | | |
|--|-------------------------|----------------------------|
| | Korrigerat värde | Ej korrigerat värde |
| <i>Utåtvändhet</i> | 5 | 5 |
| <i>Emotionell stabilitet</i> | 6 | 6 |
| <i>Öppenhet</i> | 5 | 5 |
| <i>Vänlighet</i> | 4 | 4 |
| <i>Noggrannhet</i> | 7 | 7 |

| Arbetsmotivation och arbetsattityder | | |
|---|-------------------------|----------------------------|
| | Korrigerat värde | Ej korrigerat värde |
| <i>Arbetsvilja</i> | 2 | 3 |
| <i>Arbetsintresse</i> | 3 | 3 |
| <i>Arbetsstillfredsställelse</i> | 3 | 3 |
| <i>Resultatorientering</i> | 5 | 5 |
| <i>Förändringsvilja</i> | 4 | 4 |
| <i>Balans arbete – övrigt liv</i> | 8 | 8 |

Kommentar: i detta fall finns nästan inga skillnader mellan korrigerade och okorrigerade testvärden, vilket ju är vad man skulle vänta sig med tanke på att skönmålningen låg på medelvärdet 5.

Fall 3: Låg skönmåling

| Skönmåling | |
|---------------|----------|
| <i>Overt</i> | <i>1</i> |
| <i>Kovert</i> | <i>1</i> |

| Personlighetsdimensioner av särskild vikt i arbetslivet | | |
|---|------------------|---------------------|
| | Korrigerat värde | Ej korrigerat värde |
| <i>Samarbetsvilja</i> | <i>5</i> | <i>1</i> |
| <i>Kreativitet</i> | <i>6</i> | <i>4</i> |
| <i>Emotionell intelligens</i> | <i>7</i> | <i>3</i> |
| <i>Social förmåga</i> | <i>4</i> | <i>2</i> |
| <i>Grad av positiv grundattityd</i> | <i>7</i> | <i>2</i> |
| <i>Uthållighet</i> | <i>7</i> | <i>2</i> |
| <i>Självförtroende</i> | <i>4</i> | <i>1</i> |
| <i>Perfektionism</i> | <i>6</i> | <i>6</i> |

| Dimensioner enligt femfaktormodellen ("Big Five") | | |
|---|------------------|---------------------|
| | Korrigerat värde | Ej korrigerat värde |
| <i>Utåtvändhet</i> | <i>4</i> | <i>1</i> |
| <i>Emotionell stabilitet</i> | <i>5</i> | <i>1</i> |
| <i>Öppenhet</i> | <i>4</i> | <i>1</i> |
| <i>Vänlighet</i> | <i>7</i> | <i>2</i> |
| <i>Noggrannhet</i> | <i>3</i> | <i>1</i> |

| Arbetsmotivation och arbetsattityder | | |
|---|-------------------------|----------------------------|
| | Korrigerat värde | Ej korrigerat värde |
| <i>Arbetsvilja</i> | <i>2</i> | <i>1</i> |
| <i>Arbetsintresse</i> | <i>5</i> | <i>2</i> |
| <i>Arbetsstillfredsställelse</i> | <i>2</i> | <i>1</i> |
| <i>Resultatorientering</i> | <i>5</i> | <i>2</i> |
| <i>Förändringsvilja</i> | <i>6</i> | <i>3</i> |
| <i>Balans arbete – övrigt liv</i> | <i>4</i> | <i>2</i> |

Kommentar: De korrigerade värdena ökade betydligt i detta fall, beroende på de låga värdena i skönmålning. De okorrigerade värdena gav en helt annan bild av personen än de korrigerade. Kanske har personer med denna låga nivå på skönmålning en speciell motivation att inte framställa sig själva i en alltför positiv dager. Det vore intressant att studera den saken närmare.

Det kan tyckas märkligt att vissa personer får sina resultat uppgraderade i den utsträckning som visas i denna delstudie. Det är emellertid konsekvensen av den använda metodiken, och ingenting i data tyder på att det finns icke-linjära samband eller interaktioner som skulle påkalla en annan typ av analys för dem som har låga värden i skönmålning, än vad som görs för dem med höga värden.

Korrektion för skönmålning och effekten av datakvalitet

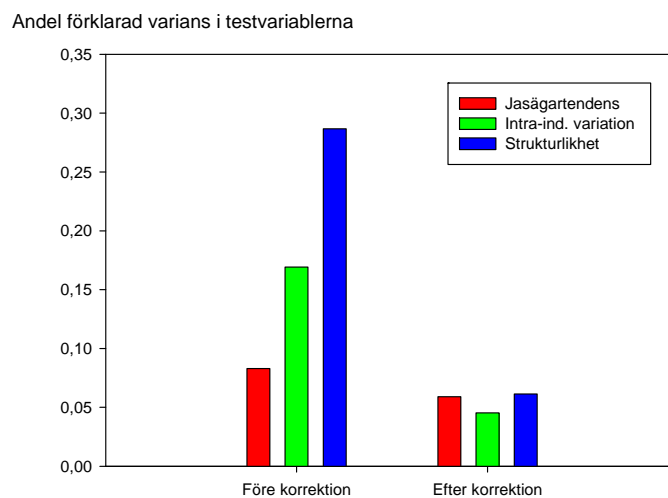
Som nämndes i inledningen har det hävdats att normativa svarsformat kan medföra vissa besvärliga felkällor: skönmålning, samt att personer varierar i tendens att hålla med om påståenden oavsett innehåll (jasägartendens) och vad gäller variationen i svar. Somliga utnyttjar skalan fullt ut, andra har en tendens att använda bara några få svars-kategorier.

Det är sant att ipsativa format undviker problemen med jasägande och intra-individuell variation. Men hur är det med ett normativt format *med data korrigerade för skönmålning*? Frågan ska här analyseras med hjälp av data från normgruppen (N=953).

Korrelationer beräknades mellan vart och ett av de tre måtten på datakvalitet och alla testvariabler (19 stycken). Kvalitetsmått var:

- Jasägartendens
- Intra-individuell variation
- Strukturlikhet

Alla korrelationer kvadrerades för att få ett mått på hur stor del av variansen i testvariablerna som förklarades av kvalitetsmått. De genomsnittliga kvadrerade korrelationerna illustreras i Fig. 7.



Figur 6. Genomsnittlig andel av variansen i testvariablerna som förklaras av kvalitetsmåten jasägartendens, intra-individuell variation och strukturlikhet.

Figuren visar en mycket tydlig skillnad mellan data före och efter korrektion för skönmålning. Okorrigerade data kunde till ganska betydande del förklaras utifrån kvalitetsmått. Så kunde till exempel ett högt värde i testvariablerna bero på att en person tenderat att välja olika svars kategorier och inte klustrat sina svar till några få kategorier. Det är just den sortens resultat som anhängare av ipsativa testformat brukar anföra som kritik av normativa format. Men data efter korrektion för skönmålning förklaras bara till ca 6 % av kvalitetsmåten. Detta är en reduktion med ca 2/3 av effekten av kvalitetsvariationen.

Diskussion och slutsatser

Om någon ljugar – kan vi ”korrigera” vad han säger och få fram sanningen? Det påstås ibland att det är omöjligt att göra en sådan korrektion. Men visst kan det göras, under vissa omständigheter. Antag att en person alltid säger sig vara 10 år äldre än vad han är. Då korrigerar vi hans uppgift om ålder och får enkelt fram sanningen. Antag att någon säger:

”Innan jag röstar sätter jag mig noga in i alla partiernas program”.

Det är troligen få om ens några som gör det. De som gör detta påstående är därför troligen ute efter att göra ett gott intryck. De skönmålar bilden av sig själva. Ju mera de håller med om påståendet, och ju flera liknande påståenden de håller med om, desto mera angelägna tycks de vara att skönmåla bilden av sig själva. Vi mäter graden av skönmålning med hjälp av våra skalor. Om en testvariabel är korrelerad med en skönmålningsskala är den troligen i viss mån ett mått på just skönmålning och inte på den personlighetsvariabel vi ville mäta. Om sambandet är linjärt kan vi enkelt korrigera på följande sätt:

Utifrån måttet på skönmålning beräknar vi det förväntade värdet i personlighetsvariabeln med hjälp av en statistisk modell (regressionsanalys). Det verkliga, erhållna, värdet skiljer sig i allmänhet från det utifrån modellen förväntade. Skillnaden mellan det erhållna och förväntade värdet är en *residual*. Detta är måttet på personlighetsvariabeln rensat från skönmålning. Om metoden fungerar eller inte måste undersökas empiriskt genom experimentella studier eller undersökningar av testning i skarpt läge jämfört med testning där inga stora värden står på spel. Ett sådant forskningsprogram beskrivs i föreliggande rapport. Resultaten visar att ca 90 – 85 % av effekterna av skönmålning har eliminerats. Det är inte 100 % men det är bra nära.

De tre fallstudierna visar att effekterna av korrektion för skönmålning kan vara dramatiska på individuell nivå. Detta gäller både för höga och låga värden i skönmålning, men effekterna blir givetvis motsatta. Det är nödvändigt att ta ställning till vilken bild av en testad person som man bör sätta tilltro till: den före eller den efter korrektion. Eftersom korrektionen har validerats både experimentellt och i en jämförelse mellan skarpt och ej skarpt läge finns det goda skäl att tro på de korrigerade värdena snarare än de okorrigerade.

Andra normativa personlighetstest använder sig oftast av okorrigerade värden. Det betyder att de regelmässigt levererar resultat som i hög grad styrs av skönmålning, om det handlar om skarpa lägen. De ipsativa testen är troligen mindre påverkade av skönmålning men de har i stället andra och stora problem, särskilt om de används vid urval och liknande beslutssituationer som befordran. Vid ipsativt format kan man ju inte jämföra personer med avseende på nivån i testvärdena, bara avseende mönstret av värden.

Vad gäller datakvalitet har det visat sig här att det *finns* samband mellan kvalitetsmåten och testvariablerna, *speciellt före korrektion för skönmålning*. Detta är ett problem för det normativa format som *UPP*-testet arbetar med. Det är emellertid intressant att det största sambandet mellan testvariabler och kvalitetsmått fanns för måttet strukturlikhet som troligen mäter sådant som slumpmässiga svar, idiosynkratisk förståelse av testuppgifterna o dyl. Den problematiken finns i lika hög grad för ipsativa format.

Efter korrektion för skönmålning kvarstod endast ca 6 % av testvariation som relaterad till kvalitetsmåten. Korrektionen hade lyckats eliminera ca 2/3 av effekterna av kvalitetsmåten. Jasägartendens, intra-individuell variation och slumpmässiga svar hade alltså en underordnad roll för testvariablerna efter korrektion.

Denna rapport har visat att:

- Skönmålning förekommer och har stora effekter men dessa kan till mycket stor del elimineras med hjälp av den metodik som används i *UPP*-testet
- Data som insamlats med ett normativt svarsformat på verkas även av några andra snedvridande faktorer (jasägartendens och tendens att klustra svaren), men effekterna av dessa faktorer kan också elimineras till största delen av den använda metodiken
- Lovande resultat har erhållits vid en undersökning av onyanserat avvisande av påståenden med negativt innehåll; i framtiden kan det eventuellt bli möjligt att korrigera för skönmålning utan att mäta skönmålning med hjälp av särskilda skalor

Referenser

- Brown, A., & Bartram, D. (2009). *Development and psychometric properties of the OPQ32r. Supplement to the OPQ 32 technical manual: SHL.*
- Cohen, J. (1988). *Statistical power analysis for behavioral sciences (2nd ed.)*. Hillsdale, NJ: Erlbaum.
- Crowne, D. P., & Marlowe, D. (1960). A new scale of social desirability independent of psychopathology. *Journal of Consulting and Clinical Psychology, 24*, 349-354.
- 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.
- Kleinbaum, D. G., Kupper, L. L., Nizam, A., & Muller, K. E. (2007). *Applied Regression Analysis and Multivariable Methods*. Pacific Grove, CA: Duxbury.
- Martin, B. A., Bowen, C. C., & Hunt, S. T. (2002). How effective are people at faking on personality questionnaires? *Personality and 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.
- Montgomery, D. C. (2009). *Design and Analysis of Experiments, 7th Edition*. New York: Wiley.
- Rogosa, D. R., & Willett, J. B. (1983). Demonstrating the reliability of the difference score in the measurement of change. *Journal of Educational Measurement, 20*(4), 335-343.
- Sjöberg, L. (2005). Självskattningstest: Den som ljugar fräckast vinner? (Self report tests: will the liar win?). *Personality and Individual Differences, 36*(6), 22-25.
- Sjöberg, L. (2008). *Bortom Big Five: Konstruktion och validering av ett personlighetstest* (SSE/EFI Working Paper Series in Business Administration No. 2008:7). Stockholm: Stockholm School of Economics.
- 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.
- Vrij, A. (2000). *Detecting lies and deceit*. New York: Wiley.