

Måling av psykodynamisk fungering

Psykometriske betraktninger om instrumentet Psychodynamic Functioning Scales (PFS)

Irina Nichita



Hovedoppgave ved Psykologisk Institutt

UNIVERSITETET I OSLO

Våren 2013

Måling av psykodynamisk fungering

Psykometriske betraktninger om instrumentet Psychodynamic Functioning Scales (PFS)

© Irina Nichita

2013

Måling av psykodynamisk fungering. *Psykometriske betraktninger om instrumentet
Psychodynamic Functioning Scales (PFS).*

Irina Nichita

<http://www.duo.uio.no/>

Trykk: Reprosentralen, Universitetet i Oslo

IV

Sammendrag

Tittel: Måling av psykodynamisk fungering - Psykometriske betraktninger om instrumentet Psychodynamic Functioning Scales (PFS).

Forfatter: Irina Nichita

Hovedveileder: Anders Johan Zachrisson

Biveileder: Dag-Erik Eilertsen

Bakgrunn: Psykodynamisk fungering er et meget omfattende begrep som krever solid teoretisk og klinisk bakgrunn for å definere og anvende til hensiktsmessige kliniske og forskningsmessige formål. Psychodynamic Functioning Scales (PFS) er et av de nyere instrumenter som er utviklet av et forskerteam ledet av prof. P. Høglend for å måle forandringer i psykodynamisk fungering på pasienter som har gjennomgått korttidsdynamisk terapi. PFS består av seks skalaer, hvorav tre måler interpersonlig fungering ("Friends", "Family", "Romantic") og tre måler intrapsykisk fungering ("Tolerance affects", "Insight", "Problem solving"). De seks PFS skalaene skåres fra 1-100 poeng og er inndelt i ti kvalitative nivåer. PFS er blitt brukt som hovedutfallsmål i First Experimental Study of Transference (FEST) studien og forsøker å fange opp spesifikke forandringer som skiller seg fra symptom- eller funksjonsforandringer med et minimum antall skalaer.

Formål: Denne hovedoppgaven tar sikte på å validere PFS skalaene mot eksterne kriterier av psykopatologi slik de er målt ved hjelp av fire instrumenter som er benyttet i FEST studien: Inventory of Interpersonal Problems – Circumplex (IIP-C), Structured Clinical Interview for DSM-IV Personality Disorders II (SCID-II) Sum kriterier (SUM CRIT), Structural Analysis of Social Behaviour – Introject Surface (SASB-IS) og Symptom Checklist - 90 Revised (SCL-90-R). Per dags dato foreligger det ingen studier eller publikasjoner, verken i Norge eller internasjonalt, som har gjennomført denne typen analyse av PFS.

Metode: Hovedoppgaven benytter data som er samlet i FEST studien og har to deler. Del 1 er en evaluering av *reliabilitet* i lys av publiserte artikler som har omtalt og estimert ulike aspekter ved reliabilitet for PFS gjennom ulike estimeringsmetoder. Del 2 omhandler *validitet* og er en valideringsstudie av PFS' skalaer i forhold til de fire ovennevnte eksterne mål på psykopatologi. Den diskuterer i hvilken grad psykodynamisk fungering slik den er målt med

PFS i FEST studien kan assosieres med ulike former for psykopatologi slik de er målt fra ulike vinkler med fire eksterne instrumenter også benyttet i FEST studien. Skårer fra PFS skalaene sammenliknes med skårer fra IIP-C, SCID-II SUM CRIT, SASB-IS og SCL-90-R gjennom en analyse av korrelasjoner.

Resultater: *Del 1 - Reliabilitet:* PFS skaler oppnår meget god interrater reliabilitet når PFS brukes av tre ratere. De tre intrapsykiske skalaene viser svakest reliabilitet, antakelig pga. deres begrepskompleksitet. PFS er funnet til å være sensitiv nok til å fange opp individuelle forandringer etter terapi, noe som var en del av det opprinnelige formålet med instrumentet. *Del 2 – Validitetsstudie:* Validitetsanalysen viser en varierende grad av sammenfall mellom PFS skalaene og eksterne mål på psykopatologi, som forventet uten et tydelig mønster, da PFS ikke er designet med tanke på å måle en spesifikk type patologi. ”Romantic” skalaen har gjennomgående færrest korrelasjoner, noe som kan skyldes utydelig beskrivelse av de ti kvalitative nivåer i skalaen eller at materialet rundt seksuelle forhold ikke kommer så lett fram ved første intervju. ”Problem solving” har flest korrelasjoner, noe som kan tyde på at den fanger mer enn bare intrapsykiske aspekter affisert av patologi. PFS er tiltenkt å dekke et bredt spekter av fungeringsgrad fra svært dårlig til meget bra. Den overraskende mangel på korrelasjoner mellom PFS skalaene og den ”friske høyredelen” av SASB-IS kan tyde på at PFS ikke måler psykiske styrker og positive evner godt nok.

Konklusjon: Funnene i denne studien indikerer at sammenhengen mellom begrepet ”psykodynamisk fungering”, målt med PFS og psykopatologi, målt med fire instrumenter som er tuftet på ulike metoder, er komplekst. Resultatene viser at slik begrepet er målt med PFS, er det ingen av de fire instrumentene isolert sett som evner å belyse bedre enn de andre andelen psykopatologi som er sammenflettet med psykodynamisk fungering. Studien viser også viktigheten av å bruke ulike metoder og instrumenter i klinisk kartleggingssammenheng for å oppnå et mer meningsfullt og utfyllende bilde av pasientens psykiske lidelse og klinisk status.

Forord

En stor og varm takk retter jeg til min hovedveileder, Anders Zachrisson. Det har vært både en fornøyelse og et privilegium å oppleve den omtenkksomhet, tilgjengelighet, kunnskap og tillit du har vist meg. Min forkjærlighet til psykoanalytisk tankegods har bare vokst under denne kontakten, noe jeg er ekstra takknemlig for. Dine kunnskapsrike, beroligende og støttende kommentarer og betraktninger har vært av uvurderlig betydning for meg.

Mange takk til min kunnskapsrike biveileder, Dag-Erik Eilertsen, som har lest gjennom oppgaven og har kommet med relevante kommentarer på en veldig effektiv måte.

En stor takk går til Per Høglend. Vår kontakt gjennom få, men svært sakelige diskusjoner, har styrket min tro på at å tilhøre et solid forskerteam med faglig engasjert og nysegjerrig leder er et mål i seg selv i arbeidslivet. Takk for at jeg har hatt mulighet til å benytte meg av data fra din studie, selv om jeg er student ved Psykologisk Institutt.

Sist, men ikke minst: jeg er evig takknemlig til min kjære familie som har vist meg støtte og har hatt tro på mine evner til å klare å fullføre en runde til med seksårig universitetsutdanning.

Oslo, august, 2013.

Innholdsfortegnelse

Sammendrag	V
Forord	VII
Innholdsfortegnelse	IX
1 Innledning.....	1
1.1 Psykodynamisk fungering	2
1.2 Eksisterende instrumenter som måler psykodynamisk fungering	4
1.3 Psychodynamic Functioning Scales – et mål på psykodynamisk fungering	5
1.3.1 Beskrivelse av PFS - skalaer og innhold.....	5
1.3.2 Det dynamiske intervju	8
1.3.3 FEST studien	8
1.3.4 Selvrapporterte eller klinikerskårede mål	9
1.4 Problemstillinger for denne oppgaven.....	11
Del 1 Psykometriske egenskaper for PFS	13
2 Reliabilitet	14
2.1 Interrater reliabilitet (inter-skårer enighet)	15
2.2 Generalizability koeffisient - en avansert form for reliabilitetsestimering.....	18
Del 2 En validitetsstudie av PFS	23
3 Validitet.....	24
3.1 Begrepsvaliditet	24
3.1.1 Innholdsvaliditet.....	24
3.2 Konvergent og divergent validitet	26
4 Metode.....	29
4.1 Måleinstrumenter.....	29
4.1.1 Inventory of Interpersonal Problems – Circumplex.....	29
4.1.2 Structured Clinical Interview for DSM-IV Personality Disorders II – Sum kriterier	30
4.1.3 Structural Analysis of Social Behaviour – Introject Surface	32
4.1.4 Symptom Checklist - 90 Revised	33
4.1.5 Det dynamiske intervjuet	34
4.1.6 PFS	34
5 Resultater og diskusjon	35

5.1	Sammenheng mellom PFS og IIP - C.....	37
	Hypoteser og faktiske funn	38
5.2	Sammenheng mellom PFS og personlighetsforstyrrelser (SUM CRIT)	46
	Hypoteser og faktiske funn	46
5.3	Sammenheng mellom PFS og SASB-IS.....	50
	Hypoteser og faktiske funn	51
5.4	Sammenheng mellom PFS og SCL-90-R.....	57
	Hypoteser og faktiske funn	57
6	Begrensninger i studien	63
7	Oppsummering og konkluderende betraktninger	65
7.1	Oppsummering av Del I – reliabilitet	65
7.2	Oppsummering av Del II – validitetsstudie	66
7.3	Konkluderende betraktninger	69
	Litteraturliste	71

1 Innledning

”Vitenskap er ikke metafysikkens høyder, men erfaringens fruktbare lavland”. I. Kant.

Eysencks slakt av psykoterapi i 1952 har på mange måter markert starten på moderne psykoterapiforskning. Han mente at psykoterapi ikke produserer ønsket resultat eller at man ikke bedrer seg mer enn den naturlige spontane bedring. 23 år senere og mange hundre psykoterapiforskningsstudier rikere, felte Luborsky, Singer, og Luborsky sin ”Dodo bird” dom og konkluderte med at psykoterapier basert på teorier og teknikker fra de viktigste psykoterapiskoler fungerer og har like bra terapeutisk effekt. I spenningsfeltet mellom disse to syn på virkning av psykoterapi, som hver for seg representerer noe udifferensierte yterligheter, blir både forskere og klinere utfordret med viktige spørsmål: (1) Hva er terapiens målsetting -- hva kan kalles for effekt av behandling (forandring) i terapi? og (2) Hvordan evaluere terapiens målsetting -- hvilke målemetoder og måleinstrumenter kan benyttes for å kvantifisere forandringene?

Innenfor psykodynamisk eller psykoanalytisk behandlingstradisjon avgis det varierende svar på spørsmål (1). I klassisk psykoanalyse var det bl.a. å *bringe det ubevisste materialet i bevisstheten*, med tanke på at det vil bringe symptomletting. I psykodynamisk terapi har *innsikt* vært et sentralt mål. Strukturelle endringer, dvs. *endringer av pasientens personlighetsstruktur eller karakter* som antas å ligge til grunn for nevrotiske symptomer har vært målsetting for flere former for dynamisk terapi. I angstprovoserende korttidsterapi (Sifneos, 1979) fører fokus på *løsning av Ødipale konflikter* med seg redusert angst, bedre interpersonlige- og problemløsningsevner. Imidlertid må det understrekes at disse målsettingene hver for seg ikke behøver være tilstrekkelige for å skape endring. For eksempel, selv om økt innsikt kan virke tilfredsstillende på det intellektuelle plan, viser forskning at innsikt alene ikke er nok for å skape endring etter terapi (Holmes, 2001).

Spørsmål (2) om evaluering dreier seg om noe har verdi og innebærer en sammenlikning og en vurdering i forhold til noe. Resultatene av en terapi kan evalueres i forhold til en annen type terapi, i forhold til noen forhåndsdefinerte bedringskriterier eller kriterier for god psykisk helse eller i forhold til pasientens eget utgangspunkt. Til disse ulike formål, denne listen er ikke uttømmende, er det blitt utarbeidet flere typer design og flere instrumenter som forsøker å fange opp endring etter terapi. Psykoterapiforskningen opplever nærmest en overveldende økning i antall instrumenter og måleskalaer. Froyd, Lambert, og Froyd (1996) har

gjennomgått 348 psykoterapistudier og funnet at de har brukt 1430 utfallsmål, hvorav 840 ble brukt kun en gang. I en annen gjennomgang har Ogles, Lambert, Weight, og Payne (1990) funnet ut at i 106 studier av psykoterapi kun for agorafobi ble det brukt 98 ulike utfallsmål. Dette kan være tegn på grad av kompleksitet terapiforskningen er preget av.

Denne oppgavens hovedfokus er å analysere og diskutere ulike psykometriske aspekter ved instrumentet Psychodynamic Functioning Scales (PFS; Høglend et al., 2000) som er valgt som et av hovedmålene for utfall for psykodynamisk terapi i FEST studien (First Experiential Study of Transference). Innledningsvis skal det redegjøres for begrepet psykodynamisk fungering og dets mulige operasjonaliseringer, deretter skal det gis en beskrivelse av PFS, dets psykometriske egenskaper på bakgrunn av hvordan instrumentet har vært anvendt i FEST studien for deretter å gå over til oppgavens hoveddel, validering av PFS i lys av eksterne instrumenter som måler psykopatologi.

Grunnet det mangfoldige innhold oppgaven tar sikte på å berøre, har den prioritert et format som muliggjør en lettere forståelse av materialet og som avviker noe fra formatet til den klassiske vitenskapelige eksperimentalartikkel.

1.1 Psykodynamisk fungering

Det er umulig å finne et fenomen eller et forskningsbegrep innefor psykologien som er direkte observerbart og dermed direkte målbart, og det er knyttet mange utfordringer til det å redusere en kompleks virkelighet til en mer håndterlig størrelse.

Generelt sett representerer en persons mentale fungering en mangfoldig realitet som byr på krevende operasjonalisering. Innenfor psykoanalytisk inspirert tenkning er mer spesifikt psykodynamisk fungering en måte å betrakte individets mentale fungering på. Før jeg startet undersøkelsen av hvordan begrepet defineres (omtales) i litteraturen, stilte jeg meg selv *spørsmål hva det ville innebære å identifisere dynamiske mentale funksjoner* og hvilke elementer som ville være naturlig å assosiere psykodynamisk fungering med. Etymologisk sett stammer ordet ”dynamisk” fra den greske ”dynamikós” som betyr ”bevegelse”, ”drivende kraft”. Overført som attributt til menneskets psykologiske fungering er det ganske vanskelig å tenke seg psykologiske tilstander som er komplet statiske, som ikke har en form for bevegelse eller motiverende drivkraft bak seg, altså ”udynamiske”. På den annen side beskriver noen begreper psykisk bevegelse av raskere og mer flyktig karakter enn andre. Prosessbegreper

som for eksempel regulering av følelser, behov og drifter assosieres med større drivende kraft og grad av foranderlighet enn strukturbegreper som for eksempel forsvarsmekanismer som er langsommere forandrelige. Mine assosiasjoner til ”psykodynamisk fungering” førte med seg begreper fra begge kategorier: forsvarsmekanisme, angst, konflikt, organisering og integrering av affektive selvopplevelser og evne til å danne og opprettholde betydningsfulle mellommenneskelige relasjoner.

Alle ovennevnte begrep kan knyttes til ulike teoretiske psykoanalytisk inspirerte retninger. I egopsykologien (Freud, 1974), står *forsvarsmekanismer* som en viktig intrapsykisk prosess brukt for å beskytte individet mot smertefulle følelser, impulser og drifter. De er i følge Vaillant (Guldberg & Svendsen, 2003) dermed viktige ubevisste mekanismer for kontroll av impulser og følelser og kan være patogene så vel som adaptive. Dette henger også sammen med Menningers konflikttrekant (Menninger, 1958 i Coughlin Della Selva, 1996) hvor ”impulser/følelser”, ”angst” og ”forsvar” utgjør trekantens hjørner, og forsvar tenkes anvendt for å minske den angst som ledsager forbudte impulser/følelser. Sullivan (1953) hevder gjennom sin interpersonlige psykoanalytiske tilnærming at *interpersonlige prosesser* er i hovedsak preget og influert av *angst* og av måten mentalt innhold styres på for å håndtere angst. Som neo-Freudianer var Sullivan også med på å understreke betydningen av interpersonlige relasjoner og at deres kvalitet antas tett knyttet med psykopatologi. Winnicott adresserer *integrering av selvopplevelser* gjennom begrepet ”false self”, i forbindelse med pasienter som utvendig ser ut og handler som en person, men som ikke føler seg som en i sitt indre. (Mitchell & Black, 1995).

Selv om psykodynamisk fungering brukes i psykodynamisk litteratur er det ikke eksplisitt definert. Et søk på ”psychodynamic functioning” i PsycInfo databasen ga 35 treff og i PubMed databasen 200 treff. Ett av få treff på artikler hvor begrepet er operasjonalisert er artikkelen til Szapocznik, Rio, Murray, og Richardson (1993) som omtaler barnets psykodynamisk fungering som bestående av 8 aspekter: psykoseksuell utvikling, intellektuell fungering, ego fungering, selv-konsept, aggresjonskontroll, emosjonell tilpassning, familierelasjoner og relasjoner til jevnaldrende. Et søk i psykoanalytiske lærebøker fra ulike tidsperioder (McWilliams, 2004; Fenichel, 1945) ga eksempler på bruk av begrepet i forbindelse med kasusformulering og evalueringer, men ingen konkret definisjon. Et søk i klassiske psykoanalytisk inspirerte bøker av Sandler, Dare, og Holder (1973), Thomä og Kächele (1994), Etchegoyen (1991) ga ingen resultat, simpelthen fordi bøkene i *sin helhet*

handler om psykodynamisk fungering. Å finne en lærebokdefinisjon av et så omfattende, vanskelig og til dels elastisk begrep virker lite sannsynlig, fordi dets definisjon byr på tilsvarende vansker som å definere mennesket. I følge Malan (1995) handler psykodynamikkens vitenskap om individets ønsker, behov, undertrykte forbudte følelser, angst for å bli sanksjonert for å ha disse følelser, samt de forsvarsmekanismene som settes i gang for håndtering av konfliktfylte og smertefulle mentale tilstander. Psykodynamisk fungering omfatter således et bredt spekter av emosjonelle, kognitive og atferdsmessige reaksjonsmønstre og disposisjoner hvor begrepene ubevisste prosesser, konflikt og tilpasning av indre psykisk realitet til den ytre virkelighet er sentrale.

1.2 Eksisterende instrumenter som måler psykodynamisk fungering

Når eksperter utvikler nye måleinstrumenter, baserer de vanligvis instrumentets innhold på en kombinasjon av klinisk erfaring, litteratur som er relevant til domenet som skal måles og lignende instrumenter som er i bruk.

Freud, Nagera, og Freud (1965) og Bellak (1969) har vært blant pionerene som har utviklet instrumenter som måler dynamiske forandringer med teoretisk forankring i drive- og egopsykologi. De har fått kritikk for å være for abstrakte, vanskelig å anvende, lite reliable og for høyt korrelerte med symptom mål.

Idiografiske (individuelle) metoder utviklet av Malan, Luborsky, Horowitz har psykometriske egenskaper som ikke gjør dem anvendbare på gruppenivå (Høglend et al., 2000).

Av de senere utviklede instrumenter som måler psykodynamisk fungering på et begrepsmessig spesifikt og detaljert nivå er det tre instrumenter som peker seg ut. *Scales of Psychological Capacities* (SPC; Wallerstein, n.d.) er et empirisk validert instrument med tilfredsstillende reliabilitet som består av 36 skalaer som måler psykodynamiske egenskaper som er forskjellige fra symptomer (Huber, Brandl, & Klug, 2004). SPC skåres på en syv-poeng skala. *Karolinska Psychodynamic Profile* (KAPP; Weinryb & Rössel, 1991) er et empirisk validert instrument som har røtter i psykoanalytisk teori. Instrumentets 18 skalaer fanger opp strukturelle endringer etter psykoterapi, med fokus på de relativt stabile modaliteter av mental fungering slik de viser seg i selv-persepsjon og interpersonlige relasjoner (Weinryb, Rössel, & Asberg, 1991). KAPP skåres på en fem-poeng skala. En studie

utført av Wilczek, Barber, Gustavsson, Asberg, og Weinryb (2004) viste at kun 8 av 18 skalaer klarte å fange opp statistisk signifikante forandringer 3 år etter oppstarten av langtidspsykoanalytisk terapi. *Operationalized Psychodynamic Diagnostics* (OPD; OPD Task Force, 2001) består av fem akser, hvor tre av dem består av i alt 34 skalaer, som skåres fra 1 til 4.

1.3 Psychodynamic Functioning Scales – et mål på psykodynamisk fungering

1.3.1 Beskrivelse av PFS - skalaer og innhold

Psychodynamic Functioning Scales (Høglend et al., 2000) er et klinikerskåret instrument som er ment å måle spesifikke effekter av psykodynamisk terapi. Instrumentet er utarbeidet på bakgrunn av en allmenn patologiforståelse hentet fra psykoanalytisk teori, mest egopsykologi og objektrelasjonsteori, men i utarbeidelsen av instrumentet ble det ikke lagt vekt på en inngående forståelse av individets psykodynamiske fungering innenfor en konkret spesifikk ramme. Derfor anses instrumentet som trofast til et perspektiv preget av ”klinisk nytteverdi” framfor ”teoretisk tydelighet”. Sifneos’ foreslåtte utfallsmål (Sifneos, Apfel, Bassuk, Fishman, & Gill, 1980) for korttidsdynamisk psykoterapi har vært en av de viktigste kilder for inspirasjon i valg av skalaene og har tatt hovedsikte på å avdekke forandringer etter korttidspsykodynamisk terapi som går utover symptom- og funksjonsforandring. PFS kan sees som et produkt av 20 års forskningserfaring med korttidsdynamisk terapi og klinisk arbeid og det bygger på empirisk grunnlag fra eksisterende instrumenter omtalt i forrige avsnitt. I likhet med disse instrumenter, måler ikke PFS personlighetstrekk eller –typologi, men beskriver interne predisposisjoner, psykologiske ressurser og evner som individet kan mobilisere for å fungere adaptivt og oppnå livstilfredshet (Høglend et al., 2000). Hovedhensikten med å lage PFS er å måle de viktigste aspekter ved psykologisk fungering i løpet av de siste tre mnd. med færrest mulig antall skalaer som samtidig dekker et bredt spekter av grad av fungering fra svært dårlig til meget bra og som dermed er ”følsomme” nok til å fange reliabelt små og moderate forandringer etter korttidsdynamisk terapi.

I nåværende versjon består PFS av seks sub-skalaer.

(1) Familierelasjoner (*Quality of family relations*) måler evne til å relatere seg til familiemedlemmer. Det innebærer gjensidighet og forpliktelse, evne til å stole på andre og å

være pålitelig selv, å gi/ta, evne til å ta familiemedlemmers perspektiv og beskrive deres personlige kvaliteter, inngå i emosjonelt samspill, oppleve følelse av å være ønsket og høre til og aksept for foreldrenes begrensninger.

(2) Vennskapsrelasjoner (*Quality of friendships*) måler fungering i vennsforhold. Det innebærer gjensidighet og forpliktelse, evne til å stole på andre og å være pålitelig selv, å gi/ta, evne til å ta venners perspektiv og beskrive deres personlige kvaliteter, inngå i emosjonelt samspill, føle tilhørighet og fungere i grupper.

(3) Romantiske / seksuelle relasjoner (*Romantic/sexual relationships*) måler fungering i eksklusive relasjoner av romantisk/ seksuell karakter. Det innebærer evnen til å oppleve seksuell interesse, opphisselse og nytelse, initiativ og fleksibilitet, moden tilknytning basert på emosjonell avhengighet av en partner, evnen til å se partneren som en unik personlighet og beskrive vedkommendes personlige egenskaper, samt evne til å takle tap.

(4) Toleranse for affekter (*Tolerance of affects*) måler evne til å erfare, differensiere og uttrykke (verbalt så vel som non-verbalt) ulike affekter. Det innebærer evne til å tåle konfliktfylte følelser før de avledes av forsvarsmekanismer eller automatisert atferd. Regulering av impulser og toleranse for frustrasjon, samt alexithymia og mangel på ekthet utgjør også en del av denne skalaen.

(5) Innsikt (*Insight*) måler kognitiv og emosjonell forståelse av hoveddynamikken i egne indre konflikter. Det innebærer forståelse av sammenheng mellom mønstre i interpersonlige relasjoner og tilbakevendende atferdsmønstre i lys av tidligere erfaringer. Evne til å forstå og beskrive egen psykologisk sårbarhet og stressreaksjoner samt mestringsmekanismer inngår som en del av skalaen.

(6) Problemløsnings- og tilpasningsevne (*Problem solving and adaptive capacity*) måler evne til fleksibel håndtering av vanskelige situasjoner og naturlig selvhevdelse uten å utvikle symptomer eller uheldige unnvikende atferdsmønstre. Skalaen gjenspeiler også evnen til å integrere utforskning av nye arenaer inn i det kjente og familiære, bruke selvobservasjon og planlegging for å oppnå meningsfulle personlige mål, samt hengi seg til og nyte rekreasjonsaktiviteter.

De første tre skalaer måler interpersonlig fungering og de tre siste skalaer måler intrapsykisk fungering. Det aritmetiske gjennomsnittet av de seks skårene for PFS skalaene utgjør en total skåre for psykodynamisk fungering som heter ”Dynamiske skalaer”.

Instrumentet har flere sterke sider. Skalaene har et velkjent format som er identisk med det brukt i Global Assessment of Functioning Scale¹ (GAF; American Psychiatric Association, 1994). Det trenges derfor ikke omfattende opplæring i bruk og skåring av PFS, men det forutsettes at skårere har psykoanalytisk/psykodynamisk klinisk erfaring. Hver skala skåres fra 1-100 poeng og er inndelt i ti deskriptive funksjonsintervaller med beskrivelser av hvert funksjonsnivå, fra ekstremt desorganisert / forstyrret til fungering på høyeste nivå. Instrumentet fanger derfor opp tilstedeværelse/grad av dysfunksjon så vel som positive aspekter i pasientens psykodynamiske fungering. PFS har ikke blitt testet på et utvalg av friske personer.

Naturlig nok, er de seks områdene i individets dynamiske fungering som PFS skalaene måler ikke ”spesifikke” områder som eksisterer helt uavhengig av hverandre. I følge Coughlin Della Selva (1996) er det alltid en interpersonlig sammenheng, i det minste i fantasien, når en følelse blir vekket. I et utviklingspsykologisk perspektiv understreker Etchegoyen (1991) at foreldre kan påvirke barnets evner til selvobservasjon og innsikt gjennom å lære barnet affektregulering. Bion (1963) understreker at evnen til å romme utålelige følelsestilstander er en nødvendig del av prosessen som transformerer disse emosjoner i α -elementer som er tett knyttet til mentalisering og erfaring av virkeligheten. Det PFS registrer er funksjoner som har en betydning for menneskets kontakt med andre mennesker, med seg selv og med virkeligheten. Beskrivende ord som familie, venner og kjærester dekker sentrale deler av menneskets interpersonlige areaner. Det som mangler er kanskje relasjoner til sjefer og medarbeidere. Innsikt, affekt regulering og evne til problemløsning er tegn på psykisk styrke og psykiske ressurser i individets forhold til virkeligheten. Og i møte med virkeligheten står andre mennesker og interaksjonen med dem helt sentralt. Det kjennes intuitivt rimelig å anta at lav innsikt og lav affekttoleranse fører med seg interpersonlige vansker. De tre intrapsyriske funksjoner har betydning for og regulerer individets forhold til familie, venner og romantiske/seksuelle partnere.

¹ GAF utgjør Akse V i DSM-IV

1.3.2 Det dynamiske intervju

Skåringen av PFS foretas på bakgrunn av et semistrukturert klinisk intervju som likner mer på en fri samtale og som starter med utsagn av typen: ” Dette er et åpent intervju og du kan fortelle alt du ønsker om deg selv” eller ”La meg starte med å spørre litt om deg selv, hvordan er din livssituasjon?” eller ”Fortell litt om hva som fikk deg til å søke terapi”. For at intervjueren skal huske alle aspekter som skal dekkes i samtalen og for å samle nok materiale for tilfredsstillende vurdering er det for hver skala utarbeidet forslag til spørsmål som tematiserer området og som dekker både kvalitative og ev. kvantitative forhold på det skalaen måler. Den psykodynamiske kliniske vurdering som intervjueren foretar må balansere mellom pasientens styrker på den ene side og psykopatologi på den annen side, noe som er felles med Malans (1995) tilnærming. Intervjuet gjennomføres av en annen kliniker enn pasientens fremtidige terapeut og det er flere klinikere som hører på intervjuet og stiller ev. tilleggsspørsmål for å sikre at alle viktige aspekter blir dekket. Intervjuets varighet varierer mellom 1 ½ og 2 timer og det tas opp på bånd. Hver kliniker som har hørt intervjuet skårer pasientens fungering på de seks PFS skalaene og GAF. På grunnlag av intervjuet formulerer de hver sin dynamiske hypotese.

1.3.3 FEST studien

Mange studier har vist at psykoterapi virker, men ikke *hvordan* og *hvorfor* (Orlinsky, Rønnestad, & Willutzki, 2004; Norcross & Lambert, 2011). FEST studien er den første RCT (Randomised Clinical Trial) som er designet for å måle langtidseffekt av overføringstolkninger i korttidodynamisk terapi. Studien er blant annet den første som viser at innsikt kan forklare mye av langtidseffekten av en spesifikk terapeutisk teknikk – overføringstolkning – slik Freud spekulerte på for mer enn 100 år siden (Hersoug, 2011). Studien ble iverksatt i 1994 av en forskningsgruppe ledet av Per Høglend og etterundersøkelsene ble avsluttet i 2005. 100 pasienter ble randomisert til ett års dynamisk psykoterapi med og uten overføringstolkninger, og ble fulgt opp i tre år etter avsluttet terapi. Studien tilfredsstillende alle kriterier for RCT, er multi-facet (multi-rater, multi-scale) og benytter dismantling design (som tillater sammenlikning av samme type behandling, med og uten én ingrediens – tolkning av overføring). Studien har ikke benyttet en kontrollgruppe, noe som kan true den interne validiteten.

Alle data som brukes i denne oppgaven stammer fra FEST studien og representerer målinger som er utført rett før starten av behandlingen (pre-treatment). I de få tilfeller hvor det ikke stemmer er dette eksplisitt nevnt. Analysene og resultatene i oppgaven er foretatt innenfor rammen av FEST studien og anses å gjelde for dets konsept, design og retningslinjer.

1.3.4 Selvrapporterte eller klinikerskårete mål

De fleste instrumenter som brukes i studier, noe som gjelder FEST studien også, er basert på selvrappotering og klinikerskårete mål. Ideelt sett skal måling av psykologiske egenskaper gjenspeile nøyaktig personens reelle psykologiske karakteristika. Dessverre finnes det en rekke kilder til feil som kan senke både reliabilitet og validitet, og som videre kan kompromittere tolkning og bruk av resultater.

Bruk av selvrapporterte mål kan introdusere potensielle feilkilder. ”Responsbias” er en av de viktigste kilder og det innebærer at pasientens systematiske responsstil på spørsmål i selvrappoteringsskjemaene fører til skårer som ikke reflekterer pasientens reelle psykologiske tilstand. De kan ha en responsstil som gjør at de konsekvent svarer i midten av skalaen og på den måten redusere skalaens sensitivitet og reliabilitet. Derimot kan pasienter med mer dramatisk responsstil besvare spørsmål med å velge de ytre punktene på skalaen. Dette gjelder både for SASB-IS, IIP-C og SCL-90-R, som alle benytter Likert skala. Lesefeil / feil tolkning av spørsmålet kan også inntreffe. Kliniske populasjoner kan ha ulike former for psykopatologi som interfererer med selvrappotering (Alden, Wiggins, & Pincus, 1990). Ulike nivåer av negativ affekt fører med seg ulik type responsbias, ifølge Vassend og Skrondal (1999). Noen pasienter unngår systematisk symptomprovoserende situasjoner og skårer dermed lavt på symptomskalaer, samtidig som de lever høyst begrensede liv pga deres psykologiske problemer. Shedler, Mayman, og Manis (1993) argumenterer for at standard selvrappotermål ikke er i stand til å skille mellom genuin mental helse og en fasade av mental helse skapt av psykologiske forsvar. De identifiserte en subgruppe pasienter som de kalte ”defensive deniers”, som så sunne ut på selvrappoterte mål, men som i realiteten viste seg å være uttrykk for en illusorisk psykologisk helse. Tilsvarende visste Varvin (2002) i sin doktorgradsavhandling om traumatiserte flyktninger at flere pasienter som tidligere hadde dissosiert bort lidelse og rapportert få symptomer, som gjennom terapi ble mindre dissosiert og mer integrert, samtidig viste en økning i antall selvrappoterte symptomer. Det er grunn til å tro at selvrappoterte helseskalaer i noen grad måler ulike ting hos ulike individer.

Pasienter kan overdrive symptomrapportering av beviste taktiske grunner (malingering) for eksempel for å være med i FEST-studien for å få tilgang til gratis psykoterapi. De kan skjule symptomer eller problematiske forhold (rusmisbruk, skambelagte seksuelle preferanser, uvanlig tvangspreget atferd) basert på sosial ønskelighet. En annen mulig feilkilde ved bruk av selvrapporteringsinstrumenter er at disse forutsetter at respondenten har en bevisst opplevelse av egen atferd og konsekvenser av denne.

Selvrapporteringsskjemaer er følsomme for stemninger respondenten for øyeblikket befinner seg i. Svarene må derfor behandles med noe varsomhet, og det kan ikke fastslås med sikkerhet i hvor stor grad de rapporterte problemene beskriver faktiske tendenser som gjør seg gjeldende i de fleste sosiale situasjoner.

Bruk av klinikerskårete mål kan også være utsatt for potensielle feilkilder. Materialet som kommer fram i intervjuet representerer ikke godt nok pasienten grunnet mangelfull intervjuing eller at pasientens narrativ er subtilt påvirket av intervjuerens holdning og ev. selvavsløringer. Sullivan (1954) hevder at både intervjuerens intense nysgjerrighet så vel som hans hånlige likegyldighet overfor pasientens problemer og historie er like uheldige for intervjuets forløp. Intervjueren kan holde seg bevisst tilbake under intervjuet for ikke å skape allianse med pasienten siden behandlingen skal foretas av en annen terapeut.

Intervjueren kan mangle god nok opplæring i gjennomføring av intervjuet, og dette er særlig viktig når begrepene som ønskes målt er vanskelig å operasjonalisere. "Halo effekten" er det fenomenet som gjør at inntrykk om spesifikke evner ved en pasient er påvirket av intervjuerens helhetsinntrykk av pasienten. Intervjueren kan ha en idiosynkratisk skåringsstil knyttet til hans personlighet som også kan introdusere bias.

Fra et psykoanalytisk ståsted har denne type intervju en konseptuel begrensning: pasienten forventes å ha grunner til å forvrengte sine erfaringer. Selv om dette ikke er et problem i seg selv, men en virkelighet som hører til det å være menneske, er det allikevel noe som vanskeliggjør vurderingen og følgelig skåringen av det kliniske materialet.

1.4 Problemstillinger for denne oppgaven

Denne oppgaven er todelt og omfatter følgende problemstillinger:

1. Del 1 retter fokuset mot *reliabilitet* og er en evaluering av reliabilitet i lys av publiserte artikler som har omtalt og estimert ulike aspekter ved reliabilitet for PFS og ulike estimeringsmetoder. Diskusjoner foretas innenfor hvert avsnitt i del 1.
2. Del 2 retter fokuset mot *validitet* og er en valideringsstudie av PFS' skalaer i forhold til eksterne mål på psykopatologi. Den diskuterer i hvilken grad psykodynamisk fungering slik den er målt med PFS i FEST studien kan assosieres med ulike former for psykopatologi slik de er målt fra ulike vinkler med fire eksterne instrumenter også benyttet i FEST studien. Skårer fra PFS skalaene vil bli sammenliknet med skårer fra IIP-C, SCID-II SUM CRIT, SASB-IS og SCL-90-R gjennom en analyse av korrelasjoner. Det foreligger ingen studier eller publikasjoner per dags dato som har gjennomført denne type validering. De forhåndsformulerte hypoteser om forventinger til resultater i forhold til hvert instrument er nevnt eksplisitt i diskusjonsdelen, i begynnelsen av hvert avsnitt om de fire instrumentene.

Del 1 Psykometriske egenskaper for PFS

Et overordnet krav til psykologisk forskningsarbeid er at det er holdbart. I oppgavens kontekst er reliabilitet og validitet generelt sett å betrakte som de viktigste aspekter ved en slik holdbarhet. Et instrument må være like konseptuelt holdbart som statistisk akseptabelt. Derfor er det svært viktig å først analysere psykometriske egenskaper ved PFS før oppgaven fortsetter videre med validering omtalt i Del 2. For at valideringen mot andre instrumenter skal kunne gjennomføres på best mulig måte, må en ha kjenneskap til psykometriske svakheter og styrker ved PFS, noe som kan gi diskusjonen rundt og tolkningen av funnene i Del 2 økt grad av nøyaktighet og mening.

For å illustrere de spesifikke betydninger reliabilitet får i den konkrete eksperimentelle studien omtalt i oppgaven, blir ulike psykometriske teorier anvendt for å danne en vitenskapelig plattform for hvordan *konstruktet* psykodynamisk fungering, *konstruktets operasjonalisering* (PFS) og *måledesign*, (FEST– RCT med dismantling design, multi-facet studie) skal forstås i en meningsfull sammenheng. Del 1 skal primært konsentrere seg om reliabilitet sett i lys av Klassisk testteori (Classical Test Theory - CTT), foreslått av Spearman i 1904, og Generalizabilityteori (GT) (Cronbach, Rajaratnam, & Gleser 1963; Brennan, 1992a).

2 Reliabilitet

Reliabilitet er en teoretisk (abstrakt) kvalitet knyttet til et instrument / målemetode og en konkret estimerbar egenskap ved *skåren* som instrumentet produserer (Furr, 2011). Innen psykologi er klassisk testteori den tradisjonelle teoretiske rammen brukt for reliabilitet og den hviler på to antakelser: (1) den observerte skåren er summen av to hypotetiske, ikke observerbare skårer, den sanne skåre og målefeil (Pedhazur & Schmelkin, 1991). (2) målefeilen anses som ”støy” som endrer seg usystematisk og tilfeldig (de systematiske målefeil er henvist til validitetsteorier). Ut ifra dette innebærer høy reliabilitet en god estimering av den sanne skåren. De fleste mål på en psykologisk egenskap er ment å fange opp variabilitet i den egenskapen, og reliabilitet er graden av variabilitet i den observerte skåren som skyldes variabilitet i den sanne skåren. Reliabilitet er definert som den del av variasjonen i en observert indikator som kan forklares med variasjon i et latent fenomen. Den forklarte variansen estimeres ved et mål på forklart variasjon, R^2 , som også heter forklart varians indeks. Reliabilitet varierer dermed mellom 0 og 1 og høyere verdier indikerer bedre psykometrisk kvalitet, altså en bedre estimering av den sanne skåren.

Innenfor ulike psykometriske teorier finnes det flere modeller som kan brukes for estimering av reliabilitet.

I CTT, er *parallel-tester* den klassiske modellen for estimering av reliabilitet og utgjør selve det teoretiske grunnlaget for CTT. Modellen estimerer reliabilitet ved Pearson's produkt-moment-koeffisient mellom to parallelle tester (for eksempel to utgaver av PFS) antatt å måle samme begrep. Testene har strenge krav knyttet til seg som ikke kan påregnes som oppfylt og anvendelsen i praksis vil innebære mye arbeid og blir derfor nokså urealistisk. En praktisk løsning på dette er blitt *split-half* og *test-retest* modellene. *Split-half* passer mest for instrumenter med testledd og det innebærer at testen deles i to ekvivalente halvdeler og det beregnes produkt-moment-koeffisienten mellom dem som er et estimat for reliabilitet for den halve testen. Denne reliabiliteten må korrigeres videre med Spearman-Brown's korreksjonsformel for å få estimering av reliabilitet for hele testen. *Test-retest* metoden adresserer spørsmålet om stabilitet og innebærer gjentatte målinger gjort med samme instrument ved ulike tidspunkt, under forutsetning av at det ikke har oppstått forandringer hos forsøkspersonene på det målte begrep mellom målingstidspunktene eller at eventuelle forandringer har gitt *like* utslag hos *alle* forsøkspersonene. PFS er blitt brukt som

måleinstrument for psykodynamisk fungering gjentatte ganger i FEST studien, både før start av behandling (T0), rett etter avsluttet behandling (T1) og ved ettårs- (T2) og treårs- (T3) oppfølging. Ingen av disse målingene kan regnes som test - retest fordi behandlingen kan ha virket ulikt på ulike pasienter, noe som vil redusere korrelasjonen mellom test og retest og vil føre til en underestimert av den sanne reliabiliteten.

I de neste to avsnitt vil det bli presentert i detalj to modeller for estimering av reliabilitet for PFS som har vært brukt i FEST studien, den første som hører til CTT og den andre som hører til en annen psykometrisk teori, GT, som adresserer en rekke svakheter og begrensinger i CTT og som representerer et avansert alternativ for estimering av reliabilitet.

2.1 Interrater reliabilitet (inter-skårer enighet)

Denne typen reliabilitet måler graden av konsensus mellom to eller flere bedømmere som skårer samme forsøksperson. Det finnes flere metoder for å undersøke dette. En metode for beregning av interrater reliabilitet som er egnet for ordinalskaler (slik PFS skalaer er) som skåres av flere enn en bedømmer (slik det er tilfellet i FEST studien) er intraklasse-korrelasjonskoeffisient (ICC- Intra-class Correlation Coefficient). Metoden analyserer hvorvidt variansen i testskåren i større grad influeres av varians over forsøkspersoner fremfor varians over bedømmere (Friborg, 2010) og kan enkelt eksemplifiseres ved formelen:

$$ICC = \frac{\textit{Patient variance}}{\textit{Patient variance} + \textit{Residual variance}} \quad (2.1)$$

En forutsetning for å oppnå en tilfredsstillende ICC verdi er at det er nok varians i det som måles ("Patient variance" i formelen over). Selv instrumenter som GAF, som ellers har en tilfredsstillende interrater reliabilitet, kan gi lav ICC dersom populasjonen som skåres er for homogen for det GAF måler (dvs. "Patient variance" er lav) og dette må tas hensyn til i tolkning og bruk av ICC. En ICC verdi på 0.70 anses å være cut-off verdien for høy reliabilitet (Fleiss, 1981; Heppner, Kivlighan, & Wampold, 1999).

ICC har fordelen av å kunne estimere både absolutt og relativ enighet mellom bedømmere, altså om bedømmere skårer alle personer likt eller bare rangordner dem likt. I den klassiske artikkelen om ICC fra 1979 diskuterer Shrout og Fleiss tre ulike modeller som kan anvendes i beregning av ICC, hver av modellene har to versjoner, avhengig av om den tar i beregning skårer fra individuelle ratere eller gjennomsnitt av skårer fra alle ratere. Gjennom konkrete

eksempler viser Shrout og Fleiss at ICC kan variere fra 0.17 til 0.91 for samme datamateriale, avhengig av hvilken modell som er valgt for beregningen av ICC.

ICC er blitt beregnet for de individuelle PFS skalaene, både ved pre- og post-treatment, i to studier:

(1) I en eldre studie (Høglend et al., 2000) som var forløper til FEST studien ble det brukt en tidligere fem-skala versjon av PFS ("Quality of family relations" og "Quality of friendships" utgjorde en skala) for skåring av psykodynamisk fungering for en gruppe på 50 pasienter. Tre klinikere skåret alle 50 pasienter før behandling og 36 av dem etter behandling. Fire andre klinikere skåret noen av disse pasienter. I beregning av reliabilitet for PFS skårer ble det valgt "two-way random effects" modellen, dvs. at bedømmerne er et tilfeldig utvalg av mulige bedømmere som skårer "n" pasienter som er tilfeldig utvalgt fra en gruppe mulige pasienter. Når kun skårer fra individuelle ratere ble brukt, viste tre av skalaene ("Tolerance of affects", "Insight" og "Problem-solving capacity") ikke tilfredsstillende ICC verdier som var under 0.70, både ved pre- og post-treatment. Den nedre grensen av 95% konfidensintervallet for disse tre skalaer var heller ikke tilfredsstillende (under 0.50). Alle skalaer viser tilfredsstillende reliabilitet dersom instrumentet brukes av tre ratere (Høglend et al., 2000).

(2) I FEST studien (Høglend, 2004) ble ICC estimert med "two-way random effects" modellen med gjennomsnitt for tre ratere (ICC(2,3)). Resultatene er veldig tilfredsstillende med unntak av "Insight" skalaen (ref. Tabell 1).

I begge studiene er ICC høyere for post- enn for pre-treatment data for samtlige PFS skalaer. Dette kan skyldes mer variabilitet i pasientgruppen som følge av behandlingen, særlig når variabiliteten som stammer fra effekt av ulik type behandling ikke skilles ut. Det kan også tenkes at pasientene var mer fortrolig med intervju situasjonen andre gang og kom med mer relevant materiale som var med på å redusere tvetydigheten i skåringsprosessen. En alternativ forklaring på dette kan være at pasienter setter pris på at noen (intervjueren) viser interesse for deres situasjon og utvikling, på tross av at terapien formelt er avsluttet, og åpner seg mer. Ratere kan også ha blitt mer enige i skåring av intervjuet andre gang som et resultat av en økt og mer samstemt forståelse for de ulike begrepene som de seks skalaer måler. Diskusjoner blant ratere om pre-treatment resultatskårene kan ha bidratt til økt grad av felles forståelse for bruken av instrumentet.

Tabell 1

The Psychodynamic Functioning Scales: Gjennomsnittsskårer for tre ratere (ICC 2,3) i FEST studien

Scale	ICC	95% CI
Quality of family relationships	0.87	0.80-0.91
Quality of friendships	0.86	0.80-0.91
Quality of romantic relationships	0.91	0.85-0.94
Tolerance of affects	0.78	0.69-0.85
Insight	0.67	0.45-0.79
Problem-solving and adaptive cap	0.81	0.73-0.86
Psychodynamic functioning	0.87	0.79-0.92
GAF	0.90	0.85-0.93

Pre-treatment (N = 100)

Scale	ICC	95% CI
Quality of family relationships	0.89	0.83-0.93
Quality of friendships	0.90	0.85-0.93
Quality of romantic relationships	0.95	0.93-0.97
Tolerance of Affects	0.86	0.79-0.90
Insight	0.85	0.77-0.90
Problem-solving and adaptive cap	0.90	0.85-0.93
Psychodynamic functioning	0.93	0.90-0.95
GAF	0.95	0.92-0.96

Post-treatment (N = 94)

Lavest reliabilitetsverdi er estimert for skalaen ”Insight”, og nest lavest for ”Tolerance of affects”, både pre- og post-treatment i begge studier. Dette kan tyde på flere ting. Det kan være vanskelig å introdusere innsikt som eget og selvstendig tema under intervjuet og dermed avdekke relevant nok materiale for skåringen. Av de seks begreper som inngår i psykodynamisk fungering er innsikt og toleranse for affekter mål på intrapsykiske dynamiske prosesser. De kan være de aller mest kompliserte å operasjonalisere og kvantifisere og dermed vanskeligst å enes om i en skåringsprosess som krever mye klinisk intuitiv vurdering. Pasientene kan også være mer ”homogene” hva angår grad av innsikt og selv-refleksjon enn på de underliggende parametre som de andre fem skalaer adresserer.

Høyest reliabilitetsverdi er estimert for skalaen ”Quality of romantic relationships”, både pre- og post-treatment i begge studier. Det kan være for tidlig å komme med en umiddelbar tolkning av dette resultatet i denne nokså begrensede konteksten, derfor henvises leseren til del 5: resultater og diskusjon hvor skalaen omtales i en bredere kontekst.

Det kan konkluderes med at PFS skaler oppnår en meget god interrater reliabilitet når PFS brukes av tre ratere. De tre intrapsykiske skalaene har svakest reliabilitet pga deres begrepskompleksitet.

2.2 Generalizability koeffisient - en avansert form for reliabilitetsestimering.

Generalizabilityteori (GT) er en omfattende psykometrisk teori som utvider perspektivet på reliabilitet fra å være kun en egenskap ved testskårer til å gjelde *bruken* av testskårer. GT kombinerer reliabilitet og validitet under et felles og nytt konseptuelt rammeverk hvor generaliserbarhet av resultater står sentralt. ”The question of ’reliability’ thus resolves into a question of accuracy of generalization, or generalizability.” (Cronbach, Gleser, Nanda, & Rajaratnam, 1972, s. 15).

Mer presist formulert, har G-teorien som mål å estimere grad av generaliserbarhet av resultater avhengig av type måledesign, kilder for målefeil, utvalg generaliseringen ønskes gjort til og slutninger som ønskes å trekkes. GT har flere fordeler i estimering av reliabilitet i forhold til CTT’s modeller og noen av disse fordelene blir presentert og forklart nedenfor i tilknytning til PFS og relatert til designet av FEST studien.

- (1) GT har mulighet til å estimere differensierte komponenter i målefeilen i motsetning til CTT som kun kan estimere én, sammensatt målefeil.

Måling og skåring av psykologiske egenskaper kan gjøres gjennom design med ulik grad av kompleksitet. En enkel-fasett målestrategi (single-facet design) foreligger når det kun er en egenskap (en fasett) som varierer i måleprosessen og antas å være den eneste kilden til feil. Innenfor GTs ramme defineres fasetten som en kilde til feilvariasjon. Et eksempel er bruk av GAF i FEST studien. Tre individuelle ratere skårer hver pasients aktuelle symptomtilstand basert på GAF. Hver rater angir en skåre per pasient basert på kun en skala (GAF) . Den eneste fasetten som varier og introduserer målefeil er ratere og forskeren kan analysere i hvilken grad variabiliteten i fasetten påvirker målekvaliteten på pasientens symptomtilstand. Siden CTT fastslår at variansen i den observerte skåren består av kun to komponenter: varians i den sanne skåren og varians i målefeilen, hvor sistnevnte ikke er differensiert, er CTT et godt nok psykometrisk rammeverk til estimering av reliabilitet i single-facet design eksemplet

brukt ovenfor, hvor målefeilen antas å stamme fra kun den ene fasetten (ratere). Reliabiliteten kan estimeres i dette tilfellet ved bruk av ICC modellen (slik det ble omtalt i forrige avsnitt).

I måling av psykologiske egenskaper ved hjelp av andre mer kompliserte typer design kan variabiliteten i den målte skåren potensielt bli påvirket av flere fasetter (feilkilder) i målingsprosessen, og hver av disse kan separat påvirke målekvaliteten. Denne situasjonen er særlig aktuell i komplekse, mange-fasett målestrategier (multi-facet design) hvor det er flere egenskaper (fasetter) som varierer i måleprosessen og påvirker skåren. Et eksempel på dette er bruk av PFS i FEST studien: for hver pasient måles psykodynamisk fungering med PFS, som består av *seks ulike skalaer* skåret av *tre individuelle ratere*. I en slik two-facet design er det kun G-teorien som kan identifisere og skille mellom to ulike kilder for feil i målingen: ratere og skalaer (Shavelson & Webb, 1991), en mulighet som ikke finnes i CTT.

(2) GT adresserer et viktig reliabilitetsspørsmål: i hvilken grad kan et begrenset antall målinger generaliseres til et tilsvarende teoretisk endeløst antall målinger? Enklere formulert, i hvilken grad kan et sett med målinger (resultater) generaliseres over tid og på tvers av situasjoner?

I sistnevnte eksempel, kan en forestille seg at psykodynamisk fungering for pasientene i FEST studien er skåret ikke bare av de tre raterne, men av alle ratere som muligvis kan skåre, dvs. av hele universet av ratere. Dersom alle skårer aggregeres, får man pasientens ”univers”skåre (tilsvarende den ”sanne” skåren i CTT). Innenfor GTs ramme, er de tre raterne å anses som et lite utvalg fra et univers bestående av alle ratere. Det underliggende spørsmålet er: hvor godt kan PFS målinger foretatt av de tre ratere i FEST studien generaliseres til andre (alle) mulige ratere? Ved bruk av tilsvarende rasjonale, kan det tenkes at operasjonaliseringen av begrepet ”psykodynamisk fungering” ikke er begrenset kun til de aspekter som de seks PFS skalaene omtaler, slik at de seks kan anses som et lite utvalg fra et univers bestående av alle mulige aspekter som antas kan representere ”psykodynamisk fungering”. Det underliggende spørsmålet er: hvor godt kan målinger av psykodynamisk fungering foretatt med de seks aspekter i PFS skalaer i FEST studien generaliseres til alle aspekter som betegner ”psykodynamisk fungering”?

G-teorien hjelper å generalisere over begge fasetter: *ratere* og *aspekter* som betegner ”psykodynamisk fungering”. Altså man kan analysere de observerte resultater i forhold til hva forskeren ville ha fått om pasientene var skåret av alle tenkelige *ratere* på alle tenkelige

aspekter som betegner begrepet ”psykodynamisk fungering”. Imidlertid må det presiseres at generaliseringen over begge fasetter innebærer en større usikkerhet enn om forskeren bare hadde generalisert over én av fasettene.

For å estimere reliabilitet gjennom G-teorien må det gjøres noen forutsetninger om fasettene. I FEST studien er ratere en tilfeldig (”random”) fasett, dvs. de tre ratere er tilfeldig trukket og det kunne ha blitt trukket andre, færre eller flere ratere. I motsetning til det, er de seks aspektene som betegner ”psykodynamisk fungering” knyttet til skalaene i PFS festet (”fixed”), de kan hverken erstattes av andre skalaer eller reduseres /økes i antall da det er den konseptuelle operasjonaliseringen av begrepet ”psykodynamisk fungering” i PFS som har styrt valget av både antall skalaer og deres innhold som er ment å dekke begrepet ”psykodynamisk fungering”. Spørsmål om hvorvidt det å begrense ”psykodynamisk fungering” til de seks aspekter valgt i PFS kan føre til en teoretisk uforsvarlig innsnevring av begrepet hører hjemme under drøfting av begrepsvaliditet (ref. 3.1), men en detaljert diskusjon om det vil gå utover oppgavens rammer. Samtidig må det tas hensyn til den kompleksiteten som preger operasjonalisering av begrepet ”psykodynamisk fungering” og til det pragmatiske perspektiv som ligger til grunn for designet av PFS, som velger å dekke ”psykodynamisk fungering” med færrest mulig antall skalaer.

Det konkluderes derfor med at å manipulere antall PFS skalaer for å estimere generalisering langs denne fasetten ikke kjennes konseptuelt meningsfylt og vil heller ikke føre til noen empirisk nyttige tolkninger. Å betrakte ratere som tilfeldige er derimot høyst relevant empirisk da instrumentet er ment å brukes av ulike team som kommer til å innholde ulikt antall ratere. En er interessert i å vite i hvilken grad ratere som er benyttet i FEST studien og som har målt psykodynamisk fungering med PFS, har produsert skårer som er generaliserbare til alle mulige ratere. Presiseringen at FEST studien er en to - fasett mixed modell design (med én festet og én tilfeldig fasett) er viktig for videre bruk av G-teorien.

Innenfor G-teori (Shavelson & Webb, 1991) estimeres reliabilitet kvantitativt ved to koeffisienter (a) generalizability koeffisient ($\hat{\rho}^2$) og (b) error/tolerance forhold ($(E/T)^2$).

(a) reliabilitetsmål eller generalizability koeffisient (G-koeffisient - $\hat{\rho}^2$) viser hvor nøyaktig en generalisering fra personens observerte skåre til personens ”univers” skåre er. Rasjonalet bak beregning av G-koeffisienten stammer fra CTT som definerer reliabilitet som andel av sann varians i den totale variansen:

$$r_{xx} = \frac{\text{Varians i sann skåre}}{\text{Varians i sann skåre} + \text{Feilvariens}} \quad (2.2)$$

I GT beregnes G-koeffisienten på liknende måte:

$$\hat{\rho}^2 = \frac{\text{Varians i universskåren}}{\text{Varians i universskåren} + \text{Feilvariens}} \quad (2.3)$$

Universskåren i (2.3) er gjennomsnittet av alle tenkelige skårer for en person langs alle fasetter og det er tilnærmet den sanne skåre i CTT. Feilen i (2.3) er differensiert og stammer fra fasettene som er definert som "random" i det respektive design, i motsetning til feilen i (2.2) som er, som tidligere nevnt, udifferensierbar i CTT.

Beregninger av reliabilitet og dens presisjon innenfor GTs rammeverk som er presentert videre, adresserer også spørsmål om evaluering, hva en evaluerer i forhold til. I FEST studien er G-koeffisienten for PFS gjennomsnittsskåren "Dynamiske skalaer" beregnet både for *relative* skåreverdi ($\hat{\rho}_{rel}^2$), altså når pasientens skåre ses i forhold til de andres skårer i gruppen (rangordning) og for *absolutte* skåreverdi ($\hat{\rho}_{abs}^2$), altså når pasienten er sin egen målestokk, og hans skåre ses i forhold til egen tidligere skåre. Dette kalles også for differanse- eller forandringsskåre (change scores) og de brukes mye i forskning på forandring i terapi.

(b) error/tolerance forholdet ($(E/T)^2$) representerer presisjonen til estimering av G-koeffisienten (Kane, 1996). Jo lavere $(E/T)^2$ verdi, desto høyere presisjon. $(E/T)^2$ verdier på under .10 hører til høy presisjon.

Tabell 2

G-koeffisienter og estimert E/T forhold for 1, 2 og 3 ratere – pre-treatment, post-treatment og differanseskåre (change score) for gjennomsnittsskåren "Dynamiske skalaer" (N=51)

Type of estimate	Pre-treatment score			Post-treatment score			Change score		
	No. of raters			No. of raters			No. of raters		
	1	2	3	1	2	3	1	2	3
est $(E/T)_{rel}^2$.39	.19	.13	.29	.15	.10	.67	.34	.22
$\hat{\rho}_{rel}^2$.72	.84	.89	.78	.87	.91	.60	.75	.82
est $(E/T)_{abs}^2$.18	.09	.06
$\hat{\rho}_{abs}^2$.85	.92	.95

Tabell 2 (Hagtvat & Høglend, 2008) viser at for *relative* skårer er både reliabilitet ($\hat{\rho}_{rel}^2$, mellom .72 og .91) og presisjon i reliabilitetens estimering (est $(E/T)_{rel}^2$, mellom .13 og .39) høyere ved bruk av tre ratere, enn ved bruk av to eller én, hvilket er forventet. Videre merkes det at reliabiliteten og dens presisjon er høyere for både pre- og post-treatment enn for change score. Reliabiliteten til en differanseskåre avhenger av reliabiliteten til de to målingene og er lavere enn reliabiliteten til hver av de to skårene som inngår i differansen (Crocker & Algina, 1986; Rogosa & Willett, 1983).

Den beste presisjonen i estimering av G-koeffisienten, samt de høyeste G-koeffisienter ble oppnådd for *absolutte* skårer: for tre ratere er est $(E/T)_{abs}^2 = .06$ og $\hat{\rho}_{abs}^2 = .95$ som regnes som bra verdier, i motsetning til uakseptable verdier for *relative* skårer som er est $(E/T)_{rel}^2 = .22$ og $\hat{\rho}_{rel}^2 = .82$.

Målefeil for absolutte forandringskårer er større enn for relative forandringskårer. Men toleransen for feil har økt vesentlig for absolutte skårer fordi man tar med graden av endring for hele pasientgruppen.

Dette viser også ett av GT's styrker nevnt i innledningen, det å knytte reliabilitet ikke kun til skårer men til skårenes *bruksområde*. Resultatene gir en indikasjon på at PFS, når dets totalskåre "Dynamiske skalaer" brukes, er et sensitivt instrument som i FEST studien har produsert generaliserbare mål på forandringer. PFS blir dermed interessant for bruk i klinisk arbeid og forskning hvor en først og fremst vil være opptatt av en spesifikk pasients endring, ikke hvordan vedkommende endres sammenliknet med en gruppe.

Del 2 En validitetsstudie av PFS

Denne delen er en valideringsstudie av PFS' skalaer i forhold til eksterne mål på psykopatologi. Den diskuterer i hvilken grad psykodynamisk fungering, slik den er målt med PFS i FEST studien, kan assosieres med ulike aspekter for psykopatologi, slik de er målt fra ulike vinkler med fire eksterne instrumenter også benyttet i FEST studien.

I oppgavens innledning ble det nevnt at formatet for oppgaven vil avvike noe fra det klassiske formatet brukt for vitenskapelige artikler. Dette innebærer at resultat- og diskusjonsdelen er slått sammen og videre inndelt i fire avsnitt, ett per instrument. Hypotesene om samsvar mellom PFS skalaer og hvert av de fire instrumenter er nevnt i hvert avsnitt for hvert av de fire instrumentene og er etterfulgt av faktiske funn. Med denne strukturen vil hypoteser, resultater og diskusjon komme fortløpende i teksten og vil skape en mer lettfattelig fremstilling og forbindelse mellom dem.

Først skal det redegjøres for ulike typer validitet, videre skal de fire instrumenter som måler ulike aspekter av psykopatologi presenteres. Deretter i resultat- og diskusjonsdelen skal det sammenliknes skårer fra PFS skalaene med skårer fra IIP-C, SCID-II Sum kriterier, SASB-IS og SCL-90-R gjennom en analyse av korrelasjoner og hvor funn diskuteres. Oppgaven avsluttes med metodiske begrensninger og konklusjon.

3 Validitet

Reliabilitet er en nødvendig, men ikke tilstrekkelig forutsetning for validitet. Dette betyr at et mål ikke kan være valid dersom det ikke er reliabelt, men at et mål er reliabelt betyr ikke at det er valid. Selv om et instrument har bra reliabilitet, kan dets psykologiske nytte være kompromittert dersom validiteten ikke er tilfredsstillende og det gjør validering av et instrument til en svært viktig del av forskningen. Validitet, slik som reliabilitet, er ikke en egenskap som hører til måleinstrument. Den knyttes til *tolkningen av resultater* og til *slutningen* forskeren gjør fra skårer til det begrepet man tror man måler, og det finnes verken én definisjon eller ét tall som alene kan betegne validitet. Validitet er avhengig av den spesifikke konteksten den omtales i (type design, forsøkspersoner, hva ønskes målt og til hvilket formål) og det er ikke fullstendig enighet blant statistikere om typer validitet og deres definisjon. Ut fra bruk av PFS innenfor FEST studiens forskningskontekst, samt som formålstjenelig introduksjon til valideringsstudien, har jeg valgt å omtale følgende typer validitet: begrepsvaliditet, innholdsvaliditet, divergent validitet og konvergent validitet.

3.1 Begrepsvaliditet

Et begrep uten virkelig innhold er tomt og data uten begrep er blinde data ifølge Immanuel Kant.

Begrepsvaliditet (eller konstruktvaliditet) er det overordnede teoristyrte kriterium for en god operasjonalisering og den avhenger av hvordan fenomenet en ønsker å undersøke/måle står i forhold til andre fenomener (Loevinger, 1957). Den oppsto med bakgrunn i diskusjonen startet av Cronbach (1949) som mente at validitet bør være knyttet til *slutningene* som trekkes fra resultatene til det operasjonaliserte begrepet som måles, heller enn til *instrumentet*. Estimering av begrepsvaliditet involverer integrering av flere typer validitet, som oftest innholdsvaliditet, konvergent og divergent validitet (Haynes, Nelson, & Blaine, 1999).

3.1.1 Innholdsvaliditet

Innholdsvaliditet sier noe om i hvilken grad instrumentet dekker et representativt utvalg av aspekter som beskriver det hypotetiske konstrukt som skal måles (Kleven, 2002). Dette rasjonale er nært knyttet til Domain sampling theory (Ghiselli, Campbell, & Zedeck, 1981) og Generalisability theory som anser hvert begrep som et univers av indikatorer / aspekter

som man kan trekke ut fra når man lager et instrument. Spørsmålet som innholdsvaliditeten adresserer, er om man har samlet indikatorer fra et relevant domene/univers.

Innholdsvaliditet er knyttet til begrepsvaliditet, men hovedforskjellen er at den første omfatter og gjelder kun for et fenomen og knytter validitet nærmere instrumentet, mens den siste ser på hvordan fenomenet står i relasjon til andre fenomener og knytter validitet nærmere slutningene som kan trekkes fra instrumentets skårer. Den aller enkleste form for validitet, en "pseudo" validitet som kan virke beslektet med innholdsvaliditet er face-validity som sier noe om at testen "ser ut til å" måle det den påstår å måle, med andre ord om testen virker rimelig. For et instrument som PFS som er ment å brukes av flere klinikere utover forskningsteamet som har dannet det, kan det tenkes at høy "face validity" kan være en motivasjonsfaktor for klinikere som ikke har brukt instrumentet før, for å vurdere å ta instrumentet i bruk.

Undersøkelse av innholdsvaliditet kan gjøres ved hrasjonelle vurderinger. Cook og Campbell (1979) har identifisert ti hovedtrusler mot innholdsvaliditeten som kan grupperes i to kategorier: (a) konstruktet er underrepresentert i operasjonaliseringen (dvs. deler av konstruktet som ønskes målt ikke er dekket av operasjonaliseringen) og (b) operasjonaliseringen er preget av overskudd av irrelevante aspekter ved konstruktet (i verste fall er det andre konstrukter som egentlig måles).

Typen trusler omtalt under pkt.(a) er knyttet opp mot diskusjonen startet i avsnitt 2.2. om antall aspekter ved psykodynamisk fungering som er valgt for PFS og generaliserbarhet til et større univers av aspekter. I lys av GT presentert tidligere kan innholdsvaliditet formuleres slik: i hvilken grad er de aspektene ved begrepet "psykodynamisk fungering" som er benyttet i PFS, representative for alle de aspektene som en hadde hatt bruk for dersom begrepet skulle vært fullgodt representert (Kleven, 2002). GTs estimeringsapparat kunne ha blitt benyttet til å estimere innholdsvaliditet dersom det hadde vært mulig å operasjonalisere "psykodynamisk fungering" ved hjelp av et tilfeldig utvalg av alle mulige aspekter som representerer begrepet. I dette tilfellet hadde grad av generaliserbarhet vært det samme som grad av innholdsvaliditet. I virkeligheten har man ikke oversikt over alle mulige aspekter som representerer "psykodynamisk fungering". Derfor vet man heller ikke hvilket univers man generaliserer til.

En av grunnene til at PFS har fått den nåværende struktur er at de eksisterende tester som måler psykodynamisk fungering (ref. avsnitt 1.2.) er for kompliserte og omfattende og innholder for mange skalaer. Det viser seg derfor å være et vanskelig prosjekt å komme med

rasjonelle vurderinger om generaliserbarhet av de seks aspekter i PFS som forskningsteamet Høglend et al. (2002) har valgt som dekkende for begrepet ”psykodynamisk fungering”.

En annen rasjonell vurdering av innholdsvaliditet er ofte en subjektiv ekspertvurdering av hvor godt instrumentet måler konstruktet. Bøgwald og Dahlbender (2004) har testet innholdsvaliditet for både de seks PFS skalaer og GAF ved hjelp av Q-sort metoden. Alle skalaene ble brudd ned i totalt 96 ulike beskrivende utsagn og 45 selvstendige psykoterapeuter fra tre ulike land ble bedt om å rekonstruere skalaene slik de trodde de opprinnelig var. Enigheten mellom individuelle bedømmere og fasit (de opprinnelige skalaer) ble beregnet ved interrater reliabiliteten ved bruk av ”two-way random effects” modellen. Den ble funnet høyest for ”Romantic/sexual relations” ($ICC(2,1)=0.95$) og lavest for ”Tolerance of affects” og GAF ($ICC(2,1)=0.89$). Viktigheten og relevansen av å velge akkurat de seks PFS skalaene som mål på psykodynamisk fungering ble bedømt av ekspertpanelet på en skala fra 0 til 100 på ca. 80. De tre intrapsykiske skalaen skåret høyest (82.8 - 84.4) og GAF lavest (64.4). Dette kan tyde på at PFS er et instrument som måler psykiske evner knyttet til psykodynamisk fungering og som i tillegg skiller seg fra symptom- og funksjonsmål. En metodesvakheter er at det ikke ble inkludert andre, alternative mål for psykodynamisk fungering i ekspertenes vurderinger. Når det gjelder overlapp mellom de ulike skalaer, var den største overlapp mellom GAF og ”Problem solving”, noe som også er påvist i en annen studie (Høglend et al, 2000).

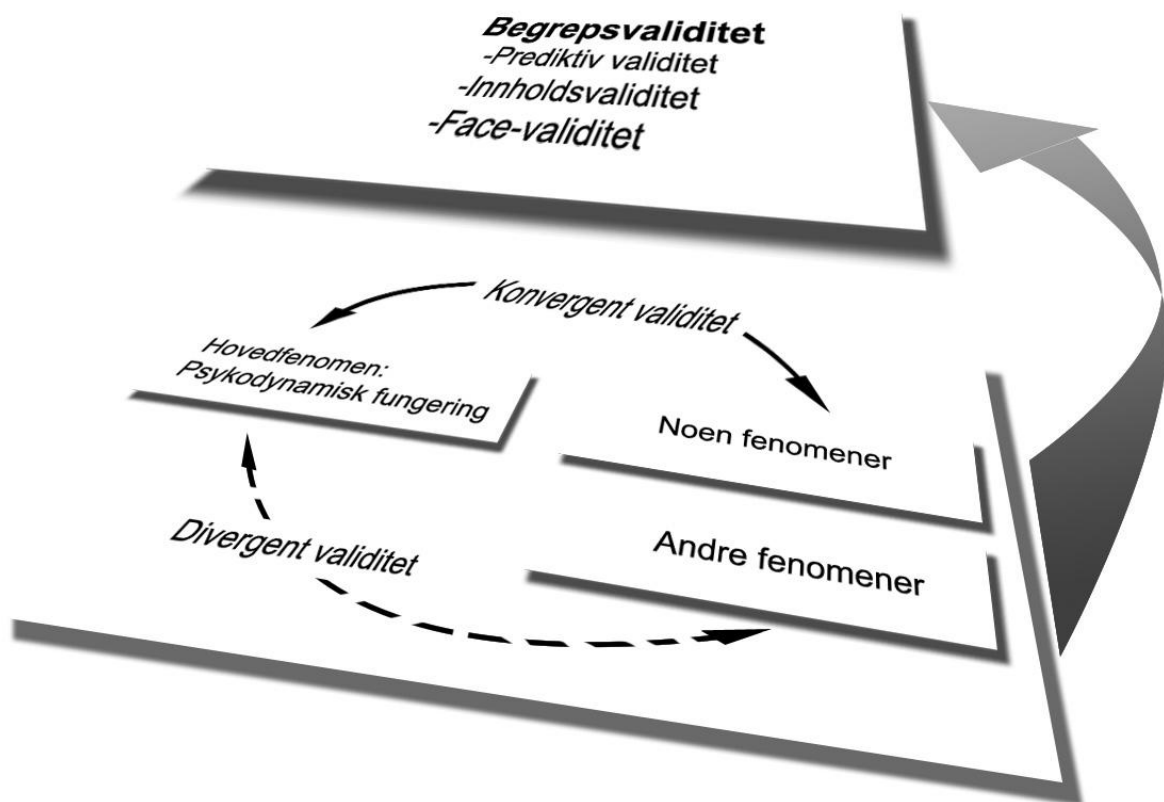
3.2 Konvergent og divergent validitet

Begrepsvaliditet kan undersøkes også ved ytre vurderinger som sier noe om hvordan og om fenomenet instrumentet måler oppfører seg som teoretisk forventet i relasjon til fenomener målt med andre instrumenter/metoder. Det kan undersøkes om PFSs testskårer korrelerer rimelig høyt med andre målinger som det teoretisk er meningsfylt grunnlag for å korrelere med, noe som kalles for konvergent validitet. Det kan også undersøkes om PFSs testskårer er uavhengige (ikke korrelerer) med målinger av andre begrep som PFS ikke er forventet å ha noen sammenheng med, noe som kalles for divergent validitet.

Denne tankegangen har sine teoretiske røtter i Campbell og Fiske (1959) sin multitrekk-multimetode-matrise som tar sikte på å skille mellom varians som stammer fra *fenomenet* som måles og varians som stammer fra *typen metode* som brukes i måling av fenomenet. For å fylle ut matrisen i sin komplette variant må forskeren identifisere minst to (eller flere)

målemetoder som kan måle det fenomenet han er interessert i og samtidig identifisere flere distinkte (uavhengige) fenomener som målemetodene kan måle. Deretter måles fenomenene med hver målemetode på samme utvalg og det beregnes korrelasjoner for hvert par målinger. Hver korrelasjonskoeffisient i matrisen tilhører en av følgende tre typer: (1) reliabilitetskoeffisienter, som er korrelasjoner mellom skårer fra samme metode som måler samme fenomen og de bør være ganske høye; (2) konvergent validitetskoeffisienter som er korrelasjoner mellom skårer fra ulike metoder som måler samme fenomen og de bør være forholdsvis høye; (3) divergent validitetskoeffisienter som enten er korrelasjoner mellom skårer fra samme metode som måler ulike fenomener, eller korrelasjoner mellom skårer fra ulike metoder som måler ulike fenomener og de bør være lave/null.

Imidlertid må det konstateres at modellen er for krevende til praktisk anvendelse og det er veldig sjelden at det foreligger data fra modellen. Derfor anses den som et teoretisk ideal som forskningen i praksis prøver å tilnærme seg.



Figur 1. Oversikt over ulike typer validitet og forholdet mellom dem.

4 Metode

For å validere psykodynamisk fungering slik den er operasjonalisert gjennom PFS seks skalaer mot eksterne mål på psykopatologi er det brukt fire instrumenter som er videre beskrevet. FEST studien som er den eksperimentelle og metodologiske rammen for valideringen er beskrevet under punkt 1.3.3.

4.1 Måleinstrumenter

Måling av utfall i psykoterapi kan hovedsakelig organiseres rundt to aspekter: 1) perspektiv - hvem foretar målingen? og 2) forandringsområde – hva anses å være forandret hos pasienten etter gjennomgått psykoterapi?

Det finnes tilstander som skaper dysfunksjon i pasientens liv, men som ikke oppfattes som plagsomme eller problematiske av pasienten selv. De kalles for ego-syntone og de oppfattes som en del av selvet, dvs. i harmoni med en selv, og pasienten vil dermed ha vansker med å skille dem fra tilstander som ikke skaper problemer, og å selvrapportere dem i et spørreskjema. I kontrast, opplever pasienten de så kalte ego-dystone tilstander som for eksempel angst og nedstemthet som plagsomme, og de utgjør ofte utgangspunktet for psykiatriske symptomer. Dette taler for nødvendigheten av å bruke ulike perspektiver i evalueringen av en pasient (Weinryb et al., 1991), da det overordnede mål er å fange det som er essensielt i pasientens psykiske tilstand. De ulike mål på psykopatologi som er brukt i FEST studien er foretatt både fra kliniker- og pasientperspektiv.

Fra de mange instrumenter som er benyttet i FEST studien har jeg valgt ut IIP-C, SCID-II, SASB-IS og SCL-90-R. De adresserer psykopatologi fra litt ulike perspektiver, de er blant de mest brukt både i klinisk forskning og i klinisk arbeid og alle har fått sine psykometriske egenskaper tilfredsstillende dokumentert.

4.1.1 Inventory of Interpersonal Problems – Circumplex

Inventory of Interpersonal Problems – Circumplex (IIP-C; Alden et al., 1990) er et selvrappoteringsinstrument utviklet for å måle subjektive opplevelser av interpersonlige problemer og er en forkortet versjon av Inventory of Interpersonal Problems (IIP; Horowitz, Rosenberg, Baer, Ureña, & Villaseñor, 1988). IIP-C har sin teoretiske forankring i Learys (1957) ”Interpersonal Circumplex”, en interpersonlig begrepsmessig modell som var inspirert

hovedsakelig av Sullivan (1953) og i noe grad av LaForge og Suczek (1955). Den ”Interpersonal Circumplex” modellen anser interpersonlige samhandlingsmønstre å være regulert av to grunnleggende former for motivasjon: behov for kontroll, makt og dominans og behov for nærhet, kjærlighet og vennskap (Horowitz, Alden, Wiggins, & Pincus, 2000). Disse utgjør modellens to ortogonale dimensjoner (akser): den horisontale er omsorgsaksen (fiendtlig/kald – vennlig/varm) og den vertikale er dominansaksen (underdanig/underkastende - kontrollerende/dominerende). Aksene definerer fire kvadranter. Videre deles de fire kvadranter i to som gir en oktant-modell med åtte poler. Disse polene definerer IIP’s åtte subskalaer som måler normal så vel som avvikende atferd. De 8 subskalaer er: Dominerende (PA), Mistenksom (BC), Kald (DE), Sosialt unnvikende (FG), Lite selvhevdende (HI), Utnyttbar (JK), Overomsorgsfull (LM) og Påtrengende (NO)². IIP-C kan sies å være problem-versjonen av ”Interpersonal Circumplex” modellen siden den fokuserer på selvrapporterte mellommenneskelige vansker, snarere enn typiske, normale væremåter. Det er 64 utsagn som inngår i IIP-C og de er formulert som selvopplevde problemer i forhold til ”en eller annen betydningsfull person i ditt liv”. De første 39 utsagnene begynner med ”Det er vanskelig for meg å...” og de resterende 25 utsagnene begynner med ”Følgende ting du gjør for mye ...” og de besvares av pasienten på en 5-punkts Likert-skala (0= Ikke i det hele tatt, 1= Litt, 2= Moderat, 3= Ganske mye, 4= Veldig). Metoden gir også en globalskåre, IIP-C Global, som er den aritmetiske summen av de 64 ledd og er et tallmessig uttrykk for den generelle grad av rapporterte mellommenneskelige problemer.

I FEST studien er det benyttet en versjon av IIP-C som er oversatt til norsk av Stiles og Høglend (1994). Monsen, Hagtvatn, Havik, og Eilertsen (2006) fant at den norske versjonen av IIP-C har akseptabel reliabilitet, en sirkumpleks struktur lik den originale versjonen og begrepsvaliditet ved at dens resultater samsvarer med resultater fra et semistrukturert diagnostisk intervju for personlighetsforstyrrelser (SCID-II; First, Gibbon, Spitzer, Williams, & Benjamin, 1997).

4.1.2 Structured Clinical Interview for DSM-IV Personality Disorders II – Sum kriterier

Structured Clinical Interview for DSM-IV Personality Disorders II (SCID-II; First et al., 1997) er et semistrukturert klinisk intervju som kartlegger pasienters personlighetsforstyrrede

² Bokstavene i parentes er etablerte betegnelse på aksene som går igjen i det meste av litteraturen.

(PF) trekk i henhold til akse II i Diagnostisk og Statistisk Manual for Mentale Forstyrrelser (DSM-IV-TR; American Psychiatric Association, 2000). Basert på dette intervjuet beregnes det et totalt antall kriterier på akse II ("SCID-II Sum kriterier", videre kalt SUM CRIT) som summen av alle diagnostiske kriterier som en pasient oppfyller for hver av de ti spesifikke PF omtalt i DSM-IV. SUM CRIT er en kontinuerlig variabel i tråd med den teoretiske oppfatning fremmet av Widiger (1992) som hevder at kumulative skårer gir et bedre bilde på personlighetspatologi enn kategorielle skårer (selvstendige PF diagnoser). Ved å angi kun et tall, adresserer den også flerdimensjonaliteten i personlighetspatologi og er bedre i stand til å fange opp det unike ved individet. SUM CRIT er også påvist å ha et lineært forhold til livskvalitet og grad av dysfunksjonalitet i interpersonlige relasjoner, uavhengig av spesifikke PF diagnoser (Cramer, Torgersen, & Kringlen, 2006) og den er også direkte relatert til grad av rigiditet i pasientens væremåte (J. T. Monsen, personlig kommunikasjon, 15. mars 2012). Rigiditet i denne konteksten må ikke forstås kun som den komponenten som er knyttet til tvangspreget personlighetsforstyrrelsesdiagnosen (rigiditet og stahet). Rigiditeten som SUM CRIT implisitt sikter mot må ses på i sin allmennhet, som en grunnleggende mangel på evner til å justere seg etter situasjon og kontekst, med andre ord et begrenset responsrepertoar.

SUM CRIT vil brukes videre i denne oppgaven som en kvantitativ indikator på mengden av generell personlighetspatologi.

I FEST studien er SCID-II blitt administrert til pasienter av én kliniker og det er ikke blitt beregnet intraklasse-korrelasjonskoeffisient (ICC) på skårene. Gjennom flere studier er det blitt påvist at den kvantitative indikatoren SUM CRIT har bedre reliabilitet enn sin tilsvarende kategorielle motpart - selvstendige PF diagnoser (Zanarini et al., 2000) og den adresserer tap av relevant klinisk informasjon som er tilfelle for pasienter som på tross av tydelige kliniske tegn og symptomer, ikke når den kliniske terskelen som kreves i DSM diagnosesystemet for å få en PF diagnose. Bruk av kvantitative indikatorer er imidlertid ikke uproblematisk da PF kriterier ikke opprinnelig var laget med tanke på høy intern konsistens, som fører til diagnostisk sammenfall mellom PF diagnoser ettersom de deler noen felles trekk og atferd (Pfohl, Coryell, Zimmerman, & Stangl, 1986).

4.1.3 Structural Analysis of Social Behaviour – Introject Surface

Structural Analysis of Social Behaviour (SASB; Benjamin, 1974) er en kompleks modell som representerer interpersonlig og intrapsykisk atferd langs to dimensjoner (interdependence og affiliation) og på tre plan med ulik interpersonlig fokus (1-individets atferd rettet mot andre; 2- individets atferd rettet mot andre som respons til andres atferd; 3- individets atferd rettet mot seg selv). Modellen bygger på den kliniske tradisjonen fra interpersonlig psykologi og objektrelasjons-teori og en beskrivelse av hele SASB modellen vil overskride oppgavens rammer, derfor blir fokuset videre kun på det tredje plan som er brukt i FEST studien.

Structural Analysis of Social Behaviour – Introject Surface (SASB-IS; Benjamin, 1974, 1988, 1994) er et selvrapporeringsmål som beskriver atferd som individet retter mot seg selv som et resultat av introjisering (internalisering) av andres atferd rettet mot individet tidligere i livet. Det er en sirkumpleks modell (Benjamin, 1996b) organisert rundt to ortogonale dimensjoner (akser). Den vertikale akse representerer gjensidig avhengighet (interdependence) som spenner mellom ytterpolene autonomi (autonomy) og kontroll (control). Den horisontale akse representerer tilknytning (affiliation), som spenner mellom ytterpolene kjærlighet (love) og fiendtlighet (attack). Aksene utgjør fire kvadranter som dekker instrumentets 8 skalaer (klustere) som er: Self-emancipate, Self-affirm, Self-love, Self-protect, Self-control, Self-blame, Self-attack og Self-neglect.

I FEST studien er det brukt en norsk oversettelse (Monsen, von der Lippe, Havik, Halvorsen, & Eilertsen, 2007) av den lange versjonen av SASB-IS selvrapporeringsskjema, som ber pasienten å beskrive seg selv slik hun/han vanligvis er gjennom å krysse av på 36 utsagn på en 11-punkts Likert-skala (0= Passer ikke i det hele tatt, 10= Passer perfekt). I Benjamins (1988) originale versjon, var instruksjonen å beskrive seg selv på sitt beste / dårligste, men denne ble forandret i den norske oversettelsen for å bl.a. kunne fange opp mer stabile aspekter ved introjektet.

Monsen et al. (2007) fant at den norske versjonen av SASB-IS tilfredsstillende de strukturelle krav til den "ideelle" sirkumplekse struktur på data fra ett klinisk utvalg, men ikke fra ett normal utvalg. Det siste er imidlertid ikke ansett å ha praktisk betydning i den kliniske anvendelsen av instrumentet. Reliabiliteten er funnet tilfredsstillende med unntak av Self-emancipate og Self-protect (ibid.), noe som blir nærmere sett på under resultat- og

diskusjonsdelen. Begrepsvaliditeten er funnet tilfredsstillende ved sammenlikning med eksterne mål på psykopatologi (ibid.).

4.1.4 Symptom Checklist - 90 Revised

Symptom Checklist - 90 Revised (SCL-90-R; Derogatis, 1994) er et selvrapporteringsinstrument som kartlegger et her-og-nå nivå av psykiske symptomer. Instrumentet har sin bakgrunn i Hopkins Symptom Checklist (HSCL; Derogatis, Lipman, Rickels, Uhlenhuth, & Covi, 1974) og er konstruert for å kunne fange opp psykologisk symptomstatus over et vidt spektrum av individer; fra de ikke-symptomatiske ("normale") individer til individer med ulike psykiatriske forstyrrelser.

Instrumentet består av 90 utsagn av typen "I løpet av siste 7 dager, hvor mye har du vært plaget av ...". Hvert av disse 90 utsagnene skåres på en 5-punkts Likert skala (0=Ikke i det hele tatt, 1=Litt, 2=Måtelig, 3=Ganske mye og 4=Veldig mye). Testen er flerdimensjonal og måler psykiske symptomer i form av ni hovdesymptomdimensjoner og tre globale indekser. De ni symptomområdene er: Kroppslige plager, Konsentrasjonsplager, Mellommenneskelig overfølsomhet, Depresjon, Angst, Raseri, Fobisk angst, Mistenksomhet og Fremmedgjøring). De globale indeksene ble utviklet for å vise en samlet lidelses status, hvert fra et noe ulikt perspektiv (Derogatis, 1994). De er "Global Severity Index" (GSI), "Positive Symptom Distress Index (PSDI) og "Positive Symptom Total" (PST). Her vil kun GSI bli omtalt da dette er det mest brukte globale mål og den indeks Derogatis (1994) selv anbefaler. Hill og Lambert (2004) omtaler GSI som den beste enkeltindikatoren på aktuell lidelsestrykk (distress). GSI er gjennomsnittet av de 90 leddene og beskriver pasienten med hensyn til generell grad av symptombelastning. Omfattende undersøkelser av referansegrupper har etablert klinisk signifikante cut-off skårer for GSI, der verdier over 0.97 (med konfidensintervaller: 0.76 – 1.19) betraktes som indikatorer på alvorlig psykopatologi (Tingey, Lambert, Burlingame, & Hansen, 1996). Instrumentet har en relativ høy grad av intern konsistens og høy test-retest reliabilitet (Derogatis, 1994; Derogatis & Unger, 2010) og i følge Derogatis og Unger (2010) er det publisert over 2000 ulike studier hvor dette instrumentet er brukt.

I FEST- studien er det brukt en norsk oversettelse av SCL-90-R (Nielsen & Vassend, 1994). I følge Vassend, Lian, og Andersen (1992) er det funnet tilfredsstillende høy reliabilitet målt ved indre-konsistens metoden, noe som er i samsvar med funn fra andre studier som viser

moderat til høy test-retest reliabilitet. Uliker aspekter av validitet (begrepsvaliditet, innholdsvaliditet) er funnet tilfredsstillende gjennom en rekke andre studier (Vassend et al., 1992).

4.1.5 Det dynamiske intervjuet

Det dynamiske intervju er utførlig beskrevet under punkt 1.3.2.

4.1.6 PFS

PFS er utførlig beskrevet under punkt 1.3.1.

5 Resultater og diskusjon

For å undersøke i hvilken grad det finnes samvariasjonen mellom skårer fra de seks PFS-skalaene, samt den globale PFS skåren ”Dynamiske skalaer” og skårer fra de andre instrumenter, ble Pearson’s produkt-moment-koeffisient beregnet. Data er presentert per instrument i tabeller med korrelasjonskoeffisienter. For de av instrumenter som har en sirkumpleks modell blir data fremstilt også ved hjelp av en sirkel for lettere å kunne observere eventuelle mønstre og / eller profiler som kan knytte PFS til psykopatologi.

Presentasjon av data og diskusjonen er slått sammen og gjøres per instrument i separate avsnitt. Resultater diskuteres i lys av de opprinnelige hypoteser som ble formulert i forbindelse med hvert instrument. Hypotesene nevnes i begynnelsen av hvert avsnitt. Alle hypoteser som legges til grunn i denne studien har til felles forventninger om å finne forholdsvis høye, positive så vel som negative korrelasjoner mellom de ulike instrumenters skalaer og PFS skalaer med beslektet tematisk innhold (konvergent validitet) og lave (insignifikante) eller ingen korrelasjoner mellom de ulike instrumenters skalaer og PFS skalaer som har ulikt tematisk innhold (divergent validitet).

I prosessen som ledsager en hypotesedannelse, er det rimelig å involvere bl.a. både tidligere empiriske funn og teoretiske fundamentet som kan støtte opp under hypotesen. Ettersom dette er den første studien som analyserer PFS i forhold til de fire instrumenter, finnes det ikke noen tilsvarende empiriske studier eller data å henvise til når det gjelder hypotesedannelse, eller å sammenlikne med når det gjelder tolkning av resultater. Resultatene i form av korrelasjonsdata mellom hvert instrument og PFS skal derfor analyseres i lys av den *teoretiske forståelse for psykopatologi* som legges til grunn for hvert instrument.

Utgangspunkt for omtale av en korrelasjonsstørrelse er Cohens (1988) standard som anser r mellom 0,10 og 0,30 som en svak korrelasjon, r mellom 0,30 og 0,50 som en moderat korrelasjon, og r over 0,50 som en sterk korrelasjon. I diskusjonsdelen omtales kun funn som er statistisk signifikante, altså korrelasjoner med $p < 0,05$, 2-tailed.

For tolkning av resultater er det nyttig å kjenne til korrelasjonsverdier mellom de seks PFS skalaskårer. Tabell 3 viser at korrelasjonene mellom de tre intrapsykiske skalaene (4, 5 og 6) ligger mellom 0,68 - 0,82 for data fra pre-treatment og mellom 0,72 - 0,80 for data fra post-treatment. De tre intrapsykiske skalaer viser sterke og stabile korrelasjoner innbyrdes, noe

som kan tyde på at de er mer homogene, at de måler samme forhold og representerer én komponent ved konstruert psykodynamisk fungering. Derimot er korrelasjonene mellom de tre interpersonlige skalaene (1, 2 og 3) svakere og ligger mellom 0,19 - 0,56 for data fra pre-treatment og mellom 0,20 - 0,62 for data fra post-treatment. De tre interpersonlige skalaer representerer derimot mer distinkte komponenter av konstruert.

Ut i fra korrelasjoner mellom de seks PFS skalaene beregnet både ved pre- og post-treatment, ser konstruert ”Psykodynamisk fungering” ikke ut til å være et homogent konstrukt (Hagtvet & Høglend, 2008).

Tabell 3

Pearsons korrelasjonsverdier mellom PFS skalaskårer – pre-treatment og post-treatment

PFS scales	1	2	3	4	5	6
1. Family	-	0,62	0,27	0,58	0,52	0,53
2. Friends	0,56	-	0,20	0,41	0,42	0,54
3. Romantic	0,19	0,34	-	0,31	0,37	0,47
4. Tolerance affects	0,43	0,58	0,44	-	0,72	0,80
5. Insight	0,25	0,48	0,44	0,74	-	0,76
6. Problem solving	0,40	0,56	0,62	0,82	0,68	-

*N = 51

Tall i **fet skrift: Korrelasjoner post-treatment

***Tall i normal skrift: Korrelasjoner pre-treatment

5.1 Sammenheng mellom PFS og IIP - C

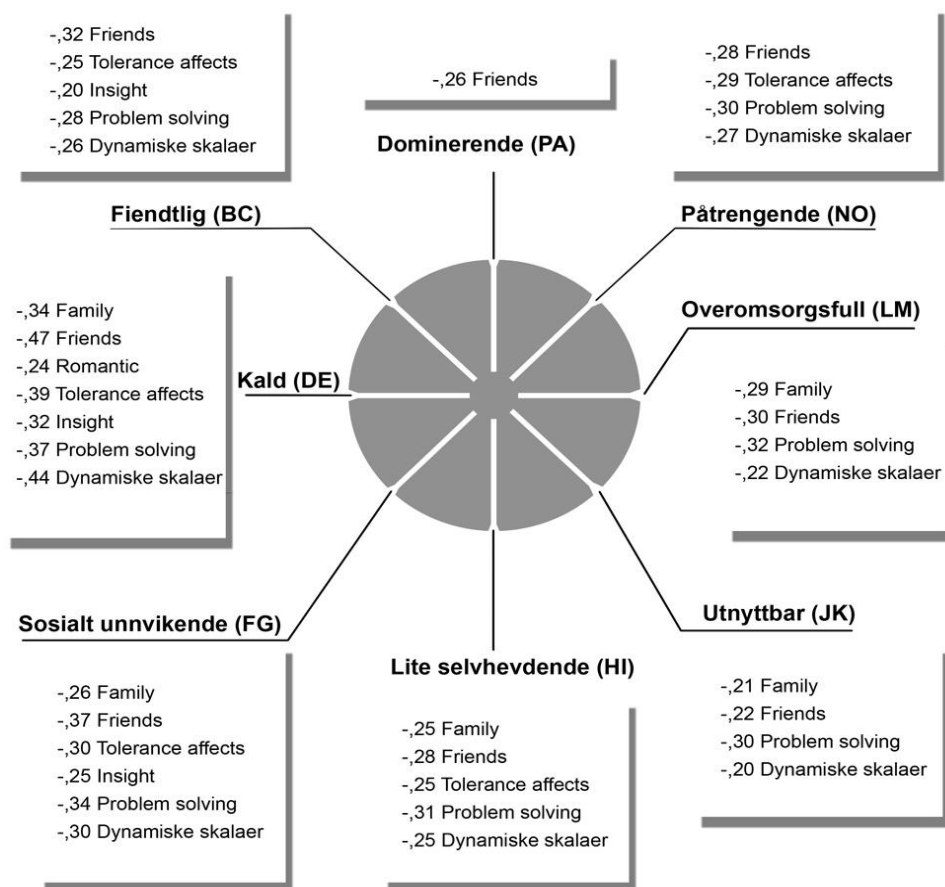
Tabell 4

Pearsons korrelasjonsverdier mellom IIP-C råskårer & IIP-C Global råskårer og PFS skalaskårer i FEST studien

PFS scales	Family	Friends	Romantic	Tolerance affects	Insight	Problem solving	Dynamiske skalaer
IIP-C Scales							
IIP-C Global	-0,34	-0,44	-0,06	-0,35	-0,2	-0,43	-0,37
PA Dominerende	-0,16	-0,26	-0,02	-0,15	-0,08	-0,17	-0,18
BC Fiendtlig	-0,20	-0,32	-0,04	-0,25	-0,20	-0,28	-0,26
DE Kald	-0,34	-0,47	-0,24	-0,39	-0,32	-0,37	-0,44
FG Sosialt unnvikende	-0,26	-0,37	0,00	-0,30	-0,25	-0,34	-0,30
HI Lite selvhevdende	-0,25	-0,28	-0,05	-0,25	-0,09	-0,31	-0,25
JK Utnyttbar	-0,21	-0,22	0,00	-0,18	-0,05	-0,30	-0,20
LM Overomsorgsfull	-0,29	-0,30	0,02	-0,19	-0,02	-0,32	-0,22
NO Påtrengende	-0,16	-0,28	-0,07	-0,29	-0,15	-0,30	-0,27

*N mellom 89 og 100

Tall i **fet skrift: signifikante korrelasjoner på 0,05 nivå (2-tailed)



Figur 2. Oversikt over korrelasjoner mellom Inventory of Interpersonal Problems – Circumplex (IIP-C) råskårer og PFS skalaskårer. Kun signifikante korrelasjoner på 0.05 nivå (2-tailed).

Hypoteser og faktiske funn

Hypotese 1. Som klinisk instrument har IIP-C sitt utspring i den interpersonlige tilnærmingen i forståelsen av psykiske problemer. Den interpersonlige tilnærmingen bygger på en rekke viktige antakelser. En av disse stammer fra Sullivans (1953) versjon av interpersonlig teori som sier at en sentral kilde til opphav og utforming av psykopatologi er av interpersonlig art. Sullivan hevder også at spesifikke former for interpersonlig atferd har en tendens til å fremkalle spesifikke reaksjoner fra andre, noe som over tid vil resultere i stabile interaksjonsmønstre i interpersonlig atferd. Horowitz (2004) bringer inn antakelsen at all mellommenneskelig atferd er motivert. Når viktige interpersonlige mål eller motiver blir kronisk frustrerte, vil det føre til mellommenneskelige problemer og skape negativ affekt (Horowitz, 2004; Horowitz et al., 2006). Frustrasjon som er assosiert med utsagn av typen "Det er vanskelig for meg å få venner" eller "Jeg befinner meg for mye alene" tyder på subjektivt opplevde mellommenneskelige problemer. I følge Horowitz et al. (1988) kan IIP-C gi en annen type informasjon enn hva rene symptom mål kan, da instrumentet er bl.a. sensitivt for endring utover det SCL-90 er. IIP-C kan dermed hjelpe med å skille beskrivelser på symptomnivå fra beskrivelser av interpersonlige problemer. PFS har også som mål å fange opp aspekter som skiller seg fra symptom- eller funksjonsforandringer og dets første tre skalaer måler interpersonlig fungering. På et mer abstrakt nivå melder følgende spørsmål seg: måler PFS og IIP-C liknende eller ulike aspekter ved menneskelig interpersonlig fungering og vansker? I lys av ovennevnte teori blir dette spørsmålet konkretisert i følgende hypotese:

- 1) For skårer på alle 8 IIP-C subskalaer og IIP-C Global er det forventet å finne flere og sterkere negative korrelasjoner med PFS' interpersonelle skalaskårer enn med PFS' intrapsykiske skalaskårer.

Faktiske funn. En visuell gjennomgang av Tabell 4 og Figur 2 viser at det er ingen umiddelbart lett gjenkjennelig mønster som kan støtte opp under den opprinnelige hypotesen 1.

"Friends" variabelen er den eneste PFS interpersonlige variabel som korrelerer signifikant med alle IIP-C variabler, inkludert gjennomsnittsskåren IIP- C Global, og blir derfor beskrevet mest inngående. Korrelasjonene kan tyde på flere ting. På den ene siden er venns relasjoner noe som går igjennom i kliniske intervjuer og undersøkelser av voksne pasienter i FEST studien, og det er den relasjon som pasientene i utvalget har flest av. Venns relasjoner er altså den mest prototypiske mellommenneskelige relasjon. På den

annen side skal IIP-C utsagn besvares i forhold til ”en eller annen betydningsfull person i ditt liv”, altså i forhold til andre mer generelt, ikke i forhold til noen spesifikke undergrupper som for eksempel familie eller kjæreste/partner.

”Friends” variabelen samvarierer sterkest med IIP-C variabelen Kald (DE, $r = -0,47$, $p < 0,05$, 2-tailed), som også utgjør den høyeste negative korrelasjon i hele Tabell 4. Høye skårer på Kald (DE) dimensjonen indikerer en person som har vansker med å uttrykke varme følelser og å føle kjærlighet overfor en annen person, en som beskriver seg selv som en ”ensom ulv” (Horowitz et al., 2000). Personen har imidlertid begynt å innse de vansker og problemer som skapes av egen begrenset evne til å tilgi, være sjenerøs, forplikte seg overfor en annen over lang tid, vise varme, medfølelse og omsorg (ibid.). Et karakteristisk utsagn for denne profilen er: ”Jeg holder andre personer for mye på avstand” (Horowitz, Rosenberg, & Bartholomew, 1993). Det er rimelig å anta at høyere skårer på Kald (DE) variabelen henger sammen med dårligere fungering i vennskapsforhold, noe som gjenspeiler seg i den forholdsvis høye negative korrelasjonen med PFS variabelen ”Friends”, som måler bl.a. evnen til å gi/ta, føle tilhørighet, inngå i emosjonell samspill og stole på andre.

”Friends” variabelen samvarierer nest sterkest med IIP-C Global ($r = -0,44$, $p < 0,05$, 2-tailed). Denne indikatoren er oftest benyttet i kliniske vurderinger (Pedersen, 2002) og har vist seg å være relatert til alvorlighetsgrad av symptomer, selvbedrag og negativ affekt (Hill, Zrull, & McIntire, 1998), noe som kan være med på å påvirke negativt evnen til vennskapelige relasjoner.

”Friends” variabelen har tredje sterkest negative korrelasjon på $-0,37$ ($p < 0,05$, 2-tailed) med variabelen IIP-C variabelen Sosialt unnvikende (FG). Høye skårer på denne IIP-C skala indikerer en person som føler seg engstelig og forlegen i andres nærvær, er hemmet i sosiale situasjoner, har vansker for å ta kontakt og være sammen med andre, samt er hypersensitiv for negativ tilbakemelding (Horowitz et al., 2000), noe som forventet gjenspeiler seg i den moderate negative korrelasjon med ”Friends” variabelen. Et karakteristisk utsagn for denne skalaen er: ”Det er vanskelig for meg å omgås med andre mennesker” (Horowitz et al., 1993).

”Romantic” er, i sterk motsetning til ”Friends”, den PFS variabelen som viser svakest korrelasjonsmønster da den kun korrelerer med én IIP-C variabel: Kald (DE, $r = -0,24$, $p < 0,05$, 2-tailed). Dette funnet kan tilsynelatende virke som nokså uventet, da ”Romantic” er en interpersonlig skala, men som allerede påpekt, tar IIP-C sikte på å måle relasjonelle

vansker i forhold til 'andre', altså ikke nødvendigvis romantiske partnere. Her er det betimelig å påpeke at Kald (DE) variabelen er den eneste IIP-C variabel som korrelerer med alle PFS variablene, inkludert PFS totalskåren "Dynamiske skalaer". En interpersonlig stil preget av høye skårer på Kald (DE) skalaen kan tenkes å ha en gjennomgripende innvirkning på personens fungering på mange områder i livet, noe som belyses av de åtte negative korrelasjonsverdier som ligger i spennet mellom $r = -0,24$ og $r = -0,47$.

"Family" variabelen har de høyeste tre negative korrelasjonsverdier med IIP-C variablene Kald (DE, $r = -0,34$, $p < 0,05$, 2-tailed), IIP-C Global ($r = -0,34$, $p < 0,05$, 2-tailed) og Overomsorgsfull (LM, $r = -0,29$, $p < 0,05$, 2-tailed).

Av de intrapsykiske PFS skalaer viser "Problem solving" flest og sterkest negative korrelasjonsverdier med IIP-C variablene, som ligger i spennet mellom $r = -0,28$ og $r = -0,43$. En tolkning av dette kan være at fleksibiliteten i håndteringen av vanskelige situasjoner, som "Problem solving" bl.a. måler, er direkte relatert til fleksibiliteten i interpersonlige relasjoner. "Problem solving" variabelen korrelerer sterkest med IIP-C Global variabelen ($r = -0,43$, $p < 0,05$, 2-tailed) som også utgjør den nest sterkeste korrelasjon for IIP-C Global (etter "Friends"). Dette kan tyde på at "Problem solving" er knyttet til et bredere fungeringsområde som også innebærer en sterkere interpersonlig komponent.

Totalskåren "Dynamiske skalaer" korrelerer negativt med alle IIP-C variabler med unntak av Dominerende (PA), med korrelasjonsverdier som ligger i spennet mellom $r = -0,20$ og $r = -0,44$, noe som gir mening da skåren er det aritmetiske gjennomsnittet av de seks skårene for PFS skalaene.

I lys av ovennevnte betraktninger og uten å ta hensyn til at IIP-C er en sirkumpleks modell (noe som denne oppgaven ikke tar sikte på å verifisere) kan det sies at den opprinnelige hypotesen 1 delvis er bekreftet. "Friends" variabelen er den som samvarierer med alle IIP-C variabler og denne sammenhengen gir mening ut i fra den teoretiske forståelsen av IIP-C, ettersom instrumentet er intendert å måle subjektivt opplevde interpersonlige vansker i forhold til andre mer generelt, ikke i forhold til noen spesifikke undergrupper (familie, kjæreste/partner). På bakgrunn av dette er det ikke overraskende at "Family" og "Romantic" variablene viser en svakere samvariasjon med IIP-C variabler. Korrelasjonsdataene i Tabell 4 har et nokså sammensatt mønster – for eksempel har "Problem solving" variabelen flere og sterkere korrelasjoner med IIP-C variabler enn det "Family" og "Romantic" variablene har.

Dermed gir data ikke entydig støtte til antakelsen at IIP-C variabler korrelerer sterkere med PFS' interpersonelle skalaskårer enn med PFS' intrapsykiske skalaskårer. Slikt sett kan bruken av IIP-C og PFS interpersonlige skalaer sies å kunne belyse ulike aspekter ved interpersonlig fungering, noe som også kan kaste nytt lys over hva andre PFS skalaer måler.

Hypotese 2. Funn fra en rekke studier (Horowitz et al., 1993; Horowitz et al., 1988) viser at pasienter i korttidspyskodyamisk terapi som strever med interpersonlige problemer knyttet til intimitet (pasienter med en IIP-C profil som heller mot polene Kald (DE), Fiendtlig (BC), Dominerende (PA)), opplever mindre forandring i sine interpersonlige problemer enn de som strever med interpersonlige problemer knyttet til selvhevdelse og selvsikkerhet (pasienter med en IIP-C profil som heller mot polene Lite selvhevdende (HI), Utnyttbar (JK) og Overomsorgsfull (LM)). Dette kan tyde på at den første gruppen plages av mer langvarige vansker, som interpersonlig tilnærming knytter til både psykopatologi og psykoterapeutisk behandling (Leary, 1957). I lys av dette og basert på det faktum at PFS også er tiltenkt som mål på den delen av menneskelig psykodynamisk fungering som delvis sammenfaller med psykopatologi, blir følgende hypotese dannet:

- 2) For subskalaene Kald (DE), Fiendtlig (BC), Dominerende (PA) er det forventet å finne flere og sterkere negative korrelasjoner med alle PFS'skaler, samt med PFS's gjennomsnittsskåre "Dynamiske skalaer" enn det er forventet å finne for subskalaene Lite selvhevdende (HI), Utnyttbar (JK) og Overomsorgsfull (LM).

Faktiske funn. En visuell gjennomgang av Tabell 4 og Figur 2 viser at det ikke er noe enkelt identifiserbart og tydelig mønster som kan støtte opp under den opprinnelige hypotesen 2.

Skalaene HI, JK og LM sett sammen under ett (Figur 2) og som tilhørende omsorgsfull/underkastende kvadranten viser en mer "homogent" distribuert antall korrelasjoner med PFS skalaer (fem for HI, fire for JK og fire for LM) enn skalaene DE, BC, PA som utgjør den fiendtlig/dominerende kvadranten (syv for DE, fem for BC og én for PA). En mulig forklaring kan være at HI, JK og LM skalaene viser høyest korrelasjon med hverandre sammenliknet med korrelasjoner mellom skalaer som utgjør de andre tre IIP-C kvadranter, noe som kan tyde på større empirisk og logisk overlapp mellom HI, JK og LM (Monsen et al., 2006). På en annen side viser disse tre skalaene ulike aspekter av avhengig interpersonlig stil, hvor Lite selvhevdende (HI) er mest forbundet med gjennomgripende vansker i livet og psykisk besvær, som for eksempel depresjon (Pincus & Gurtman, 1995)

En kunne også forvente sterkere motsetninger i korrelasjonsmønsteret mellom HI, JK og LM kvadranten og PFS skalaer på den ene siden og DE, BC og PA kvadranten og PFS skalaer på den andre siden. Dette ser ikke ut til å være tilfellet av flere mulige grunner. IIP-C skårene som er brukt i beregning av korrelasjonene er rå, altså ikke ipserte (ipsering utdypes mot slutten av avsnittet). Det kan være noe spesifikt ved utvalget sånn som underrapportering av negative interpersonlige egenskaper, eller at disse er blitt ego-syntone og oppleves dermed ikke forstyrrende ”nok” for pasientene for å rapporteres. Det kan også hende at subskalaenes interkorrelasjoner og faktorstruktur ikke oppfyller kravene til sirkumplekse modeller for dette utvalget.

Som tidligere nevnt, skiller variabelen Kald (DE) seg som den eneste IIP-C variabel som korrelerer med alle PFS variablene, inkludert PFS totalskåren ”Dynamiske skalaer”. Sett per PFS skala, viser Kald (DE) de sterkeste korrelasjoner av alle IIP-C variabler, med unntak av ”Problems solving” hvor korrelasjonen med Kald (DE) variabelen er nest sterkest, forbigått av korrelasjonen med IIP-C Global. ”Friends” variabelen korrelerer sterkere med både Fiendtlig (BC, $r = -0,32$, $p < 0,05$, 2-tailed) og Dominerende (PA, $r = -0,26$, $p < 0,05$, 2-tailed) variablene enn den gjør med Lite selvhevdende (HI, $r = -0,25$, $p < 0,05$, 2-tailed) og Utnyttbar (JK, $r = -0,21$, $p < 0,05$, 2-tailed). Dette er med på å bekrefte hypotese 2.

I motsetning til Kald (DE), korrelerer Dominerende (PA) variabelen med kun én PFS variabel, ”Friends”. Høye skårer på Dominerende (PA) indikerer en person som har problemer med å slippe kontrollen, og kan beskrive seg selv som manipulative og aggressive i deres forsøk på å kontrollere andre (Horowitz et al., 2000). Det kan oppsummeres med utsagnet: ”Jeg forsøker for mye å forandre andre” (Horowitz et al., 1993). Imidlertid må det understrekkes at hovedtyngden i denne skalaen ligger på kontroll og ikke på aggressivitet. Dette kan danne en ”snillere” profil som fungerer mer hensiktsmessig i interpersonlige relasjoner ettersom dominans fremkaller underkastelse, altså en komplementær respons fra motparten, i motsetning til en fiendtlig profil som fremkaller fiendtlighet fra motparten (Tracey, 2005). Dominans kan medvirke til å skape litt mindre giftige interpersonlige relasjoner enn det fiendtlighet gjør og assosieres dermed med lavere lidelsestrykk (distress). Dette kan være med på å forklare fravær av forventede signifikante negative korrelasjoner med de andre PFS interpersonlige variabler. En alternativ forklaring kan også finnes i karakteristika ved pasientutvalget i studien som hadde svært få problemer knyttet til dominans. Monsen et al. (2006) har funnet at skalaene Påtrengende (NO) og Dominerende

(PA) har lavest reliabilitet, mens Kald (DE), Sosialt unnvikende (FG), Lite selvhevdende (HI) og Utnyttbar (JK) har høyest reliabilitet, noe som samsvarer med funn fra den opprinnelige versjonen (Alden et al., 1990).

Høye skårer på Fiendtlig (BC) skalaen beskriver en person som er opptatt av hevngjerrighet, som uttrykker mye sinne og mistenksomhet og som har svekket evne til å tilgi og bry seg om andres behov og lykke (Horowitz et al., 2000). Et karakteristisk utsagn for denne skalaen er: ”Jeg slåss (engelsk ’fight’) for mye med andre mennesker” (Horowitz et al., 1993). Fiendtlig (BC) variabelen korrelerer med fem av de syv PFS variabler i spennet mellom $r = -0,20$ og $r = -0,32$.

Høye skårer på Lite selvhendende (HI) skalaen beskriver en person med lav selvfølelse og svekket evne til å gjøre sine behov kjent for andre, som er utilpass i autoritetsroller og som tviler på seg selv (Horowitz et al., 2000). Et karakteristisk utsagn for denne skalaen er: ”Det er vanskelig for meg å være selvsikker (’assertive’) overfor en annen person ” (Horowitz et al., 1993). Lite selvhendende (HI) variabelen korrelerer med fem av de syv PFS variabler i spennet mellom $r = -0,25$ og $r = -0,31$.

Høye skårer på Utnyttbar (JK) skalaen beskriver en person med vansker for å uttrykke sinne av frykt for å fornærme andre, har et overdrevent ønske om å tilfredsstille andre og har vansker med å si ”nei” for å beholde vennlige relasjoner, noe andre kan dra fordel av (Horowitz et al., 2000). Et karakteristisk utsagn for denne skalaen er: ”Jeg lar andre mennesker utnytte meg i for stor grad” (Horowitz et al., 1993). Utnyttbar (JK) variabelen korrelerer med fire av de syv PFS variabler i spennet mellom $r = -0,20$ og $r = -0,30$.

Høye skårer på Overomsorgsfull (LM) skalaen beskriver en person som prøver å tilfredsstille andre i for stor grad, er for omsorgsfull, generøs, empatisk, tillitsfull og har vansker med å sette grenser (Horowitz et al., 2000). Et karakteristisk utsagn for denne skalaen er: ”Jeg setter i for stor grad andres behov foran mine egne” (Horowitz et al., 1993). Overomsorgsfull (LM) variabelen korrelerer med fire av de syv PFS variabler i spennet mellom $r = -0,22$ og $r = -0,32$.

Det viser seg å være nokså problematisk å sammenlikne korrelasjonsverdier for hver PFS variabel på tvers av de seks utvalgte IIP-C variabler for å undersøke om det finnes støtte for hypotese 2 da det ikke er samme PFS variabler som korrelerer med hver av de seks IIP-C variabler.

I et forsøk på å forenkle denne analysen skal fokuset rettes hovedsakelig mot PFS's gjennomsnittsskåre "Dynamiske skalaer" som oppsummerer alle seks PFS skalaer. Data viser at korrelasjonen mellom gjennomsnittsskåren "Dynamiske skalaer" og Fiendtlig (BC, $r = -0,26$, $p < 0,05$, 2-tailed) og Kald (DE, $r = -0,44$, $p < 0,05$, 2-tailed) på den ene siden er litt sterkere enn korrelasjonen mellom gjennomsnittsskåren "Dynamiske skalaer" og Lite selvhverdende (HI, $r = -0,25$, $p < 0,05$, 2-tailed), Utnyttbar (JK, $r = -0,20$, $p < 0,05$, 2-tailed) og Overomsorgsfull (LM, $r = -0,22$, $p < 0,05$, 2-tailed) på den andre siden. Forskjellene er allikevel ikke store nok for å kunne danne en tydelig konklusjon. Funnene fra studiene til Horowitz et al. (1993) og Horowitz et al. (1988) som den opprinnelige hypotesen 2 baserer seg på og som tilsier at pasienter som strever med interpersonlige problemer knyttet til intimitet opplever mindre forandring i sine interpersonlige problemer enn de som strever med interpersonlige problemer knyttet til selvhverdelse og selvsikkerhet - noe som videre kan knyttes til mengde psykopatologisk lidelse - kan ha alternative forklaringer. Selvhverdelse involverer en mer åpenbar, lettere observerbar atferd mens intimitet innebærer private, mindre åpenbare opplevelser som nærhet, tillit, forpliktelse, noe som gjør det sistnevnte vanskeligere å avdekke og mer komplekst å håndtere i en behandlingssituasjon hvor pasienter med intimitetsvansker for eksempel kan ha en redusert evne til å inngå en allianse med terapeuten.

Ved tolkning av resultatene må det også taes hensyn til at selvrappotering kan skje asymmetrisk på de ulike IIP-C områder. En sammenlikning av selvrappoterte IIP-C skårer med skårer fra jevnaldrende observatører utført av Hill et al. (1998) har vist at forsøkspersonene var mer tilbøyelig til å rapportere interpersonlige problemer i Overomsorgsfull (LM) og Utnyttbar (JK) kvadrantene mens observatørene la mer fokus på emosjonell avstand, aggressiv, kontrollerende og hevnjerrig atferd, altså interpersonlige problemer som hører til Kald (DE), Fiendtlig (BC) og Dominerende (PA) kvadrantene. Dette kan tyde på at individet er mindre klar over innvirkningen de sistnevnte typer atferd har på andre, og fenomener som self-serving bias og attribusjonsstil kan ytterligere komplisere bildet. Som William James skrev i 1890: vi viser oss ikke for våre barn på samme måte som for venner i festlige lag, for våre kunder som overfor arbeidere vi ansetter, for våre overordnede og arbeidsgivere som overfor våre nære venner.

Som avsluttende betraktninger, felles for både hypotese 1 og 2, må det nevnes at data ikke støtter IIP-C som en sirkumpleks modell, da positivt korrelerte PFS skalaer (ref. Tabell 3) korrelerer ganske likt med motsatte poler i IIP-C. En av hovedgrunnene til det kan være at

IIP-C skårene brukt i denne oppgaven ikke er ipserte. Ipseringen eliminerer en g-faktor eller en såkalt ”problem/plagefaktor” som ved faktoranalyse har vist seg å lade på alle skalaer (Alden et al., 1990; Tracey, Rounds, & Gurtman, 1996) og som kan vanskeliggjøre tolkningen av korrelasjonskoeffisientene. Da en analyse av IIP-C’s sirkumplekse egenskaper ville gå utover oppgavens rammer og omfang, ble IIP-C’s egenskaper som en sirkumpleks modell sett bort fra allerede ved hypotesedannelse. De IIP-C råskårene, altså ikke ipserte skårene i Tabell 4 er dermed influert av en generell problemfaktor og korrelasjonene tyder på at IIP-C måler i hovedsak *mengde* av interpersonlige problemer. I lys av dette gir de negative korrelasjonene mellom PFS skårer og IIP-C skårer mening, ettersom større mengde interpersonlige problemer er relatert til svakere psykodynamisk fungering. Det er betimelig å nevne her at g-faktoren også kan reflektere en spesifikk responsstil eller en strategisk selv-representasjon. Det kan derfor tenkes at den tapper en annen essensiell dimensjon ved pasienten, i tillegg til mengde interpersonlige problemer.

Følgende sitat av Leary (1957) er vanskelig å ikke si seg enig i da den oppsummerer kompleksiteten i menneskelig (og implisitt mellommenneskelig) atferd: ”Human behavior is not a unified single process. It is rather a shifting, conflicted multifaceted complex of motives, overt and covert” (s. 40).

Dette støtter opp under det faktum at komplekse konstrukter fordrer bruk av flere uavhengige metoder og måleinstrumenter for å kunne belyse ulike aspekter av fenomenet som ønskes analysert.

5.2 Sammenheng mellom PFS og personlighetsforstyrrelser (SUM CRIT)

Tabell 5

Pearsons korrelasjonsverdier mellom SUM CRIT skårer og PFS skalaskårer i FEST studien

PFS scales	Family	Friends	Romantic	Tolerance affects	Insight	Problem solving	Dynamiske skalaer
SUM CRIT	-0,30	-0,45	-0,31	-0,51	-0,45	-0,45	-0,50

*N mellom 79 og 87

Tall i **fet skrift: signifikante korrelasjoner på 0,05 nivå (2-tailed)

Hypoteser og faktiske funn

I ICD-10 (World Health Organization, 1992) omfatter personlighetsforstyrrelser (PF) dypt innarbeidede og vedvarende atferdsmønstre som kommer til uttrykk gjennom rigide reaksjoner på et bredt spekter av personlige og sosiale situasjoner. De representerer enten ekstreme eller betydelige avvik fra måten et gjennomsnittsindivid i en gitt kultur opplever, tenker, føler og særlig forholder seg til andre individer på. PF kommer ofte til uttrykk i interpersonlige relasjoner. Selv om både DSM-IV-R og ICD-10 er ment å være ateoretiske diagnostiske klassifikasjonssystemer, er kriteriene for de fleste PF formulert som vedvarende problemer i forhold til andre mennesker, noe som grenser mot en teoretisk interpersonlig tilnærming. PF assosieres med grad av dysfunksjonalitet i interpersonlige relasjoner og er funnet å ha sterkere statistisk prediktiv verdi for livskvalitet enn demografiske variabler, somatisk helse og Akse-I diagnoser (Cramer et al., 2006). Flere studier finner at tilstedeværelse av en co-morbid PF har en negativ innvirkning på både tilbakefall og varighet av flere Akse-I diagnoser (Skodol et al., 2005; Newton-Howes, Tyrer, & Johnson, 2006).

I FEST studien er 46% av pasientene diagnostisert med én personlighetsforstyrrelse og 19% har mer enn én PF (Høgland et al., 2006). SUM CRIT gjenspeiler antall overordnede kriterier på at pasienten har reaksjonsmåter som er utilfredsstillende og som fører til subjektiv lidelse, funksjonssvikt eller problemer med og for andre. Cut-off skåren ble satt til 9.

Hypotese 1. Det er forventet å finne forholdsvis negative korrelasjoner mellom SUM CRIT og alle seks PFS skalaer, samt gjennomsnittsskåren ”Dynamiske skalaer”.

Faktiske funn. Korrelasjonsverdier i Tabell 5 gir full støtte for hypotesen da alle PFS variabler, inkludert gjennomsnittsvariabelen, korrelerer med SUM CRIT i spennet mellom $r = -0,30$ og $r = -0,51$.

I følge Pincus og Wiggins (1990) regnes PF og interpersonlige problemer for å være relaterte til hverandre. De påpeker en voksende konsensus i litteraturen om personlighetsforstyrrelser som tyder på at dysfunksjonell mellommenneskelig adferd enten er definerende trekk ved eller en stor del av mange PF. Tanken støttes også av en rekke andre teoretikere og klinikere (Horowitz, 2004; Ward, 2004) som beskriver individer diagnostisert med ulike PF som har de vanligste vanskeligheter knyttet til mellommenneskelige relasjoner. Det gir derfor mening at SUM CRIT korrelerer negativt med de tre interpersonlige PFS variabler, "Family", "Friends" and "Romantic". Funnet er ikke overraskende og støttes av studien til Johansson et al. (2010) som viser at 60% av FEST studiens' pasienter med lave skårer på en skala som kartlegger historikk og kvalitet på interpersonlige relasjoner har hatt én eller flere personlighetsforstyrrelser. Dette er videre i tråd med studien til Skodol et al. (2005) som har funnet at pasienter med PF opplever betydelig større vansker i fungering, særlig på den sosiale arenaen, noe som har vist seg å være mer vedvarende enn PF symptomer og PF trekk.

PF er vanligvis forbundet med betydelige problemer i forhold til yrkesmessige funksjoner og ferdigheter. Dette kan være med på å støtte opp under den forholdsvis sterke korrelasjonen mellom SUM CRIT variabelen og "Problem solving" ($r = -0,45$), da en pasient med omfattende og gjennomtrengende mistilpassede (maladaptive) atferdsmønstre kan ha problemer med å håndtere nye og vanskelige situasjoner, samt planlegge og oppnå realistiske personlige mål. Interpersonlige vansker har også negative innvirkninger på pasientens engasjement i fritidsaktiviteter, da mange av dem innebærer mellommenneskelig samhandling.

Frustrasjon av måloppnåelse i mellommenneskelig samhandling - som den interpersonlige tilnærmingen legger til grunn i forståelse av PF - fører med seg en del negativ affekt (Horowitz, 2004). Affekt står også sentralt knyttet til mål og behov i nyere psykoanalyse og psykodynamisk tenkning, og den anses å innramme alle motiv, behov og relasjoner til ytre personer (objekter) eller indre objektrepresentasjoner (Kernberg, 1984; Gullestad & Killingmo, 2005). Angående de regulatoriske mekanismer for affektiv selvregulering, vil en del forskere og teoretikere hevde at det skjer gjennom relasjonen med andre. Dersom relasjonen med andre er preget av maladaptive (mistilpassede) atferdsmønstre, blir det

opprinnelige motivet ytterligere frustrert, noe som kan være med på å forsterke og tydeliggjøre individets ubalanse på det affektive plan. I følge Malmquist (1991), kan personlighets”defekter” føre til visse atferdssymptomer som kan være direkte uttrykk for affektive forstyrrelser. SUM CRIT’s sterkeste korrelasjon er med PFS variabelen ”Tolerance affects”, ($r = -0,51$, $p < 0,05$, 2-tailed) noe som kan forstås på bakgrunn av det ovennevnte og at affektregulering, differensiering og integrering av affekter, samt toleranse for frustrasjon er en del av kjernen i PF diagnoser.

SUM CRIT har en korrelasjon på $r = -0,45$ med ”Insight” variabelen. Innsikt er et kompleks begrep som kan defineres på ulike måter. I klassisk freudiansk tenkning betyr innsikt å gjøre det ubevisste bevisst for en selv, gjennom en prosess som anses ikke å være overførbar (individet kan oppnå innsikt kun i seg selv). Innsikt er også en link til ny mening, som forandrer individet’s syn på seg selv og på virkeligheten, gjennom tilgang til ny kunnskap som tidligere ikke var tilgjengelig (Rapaport i Etchegoyen, 1991, s. 657). Oppnåelse av ny innsikt mobiliserer nytt atferdsrepertoar med responser som er mer adaptive (Etchegoyen, 1991). Autentisk innsikt er også blitt sett på som en kombinasjon av intellektuell og emosjonell forståelse av de dypere kilder til egen psykisk opplevelse, en krevende prosess som i sin genuine form er ledsaget av forandringer i individets intrapsykiske balanse (Kernberg, 1984). I forbindelse med behandling av PF, betrakter Benjamin (1996) innsikt som en læringsprosess som avdekker og setter sammen pasientens kognitive mønstre på nye måter i forhold til persepsjoner, responser og internaliseringer. Relaterte og mer moderne begreper er ”psychological mindedness” (aspektene som angår det å være seg selv bevisst), ”mentalisering /reflexive functioning”. Det er en nærliggende tanke å knytte lavere innsikt med alvorlighetsgrad av PF, men bildet kompliseres av de nokså ulike syn på innsikt som følgelig byr på heterogent operasjonalisering. Ulik operasjonalisering av begrepet gjør det problematisk å sammenlikne resultater fra ulike studier på tvers av PF på en meningsfull måte. Selv om videre nyansering av den negative sammenhengen mellom SUM CRIT og ”Insight” vanskeliggjøres av manglende empirisk sammenlikningsgrunnlag, kan en allikevel si seg enig med Westen’s (1997) utsagn at manglende innsikt er ”diagnostisk”, dvs. karakteristisk, for mange PF. Korrelasjonen må ses på som en indikasjon på at de to variabler har til felles og måler en del psykopatologi som er knyttet til lavere innsikt ved tilstedeværelse av PF.

SUM CRIT viser nest sterkest korrelasjon med ”Dynamiske skalaer” ($r = -0,50$), noe som tyder på at en del av psykopatologi knyttet til PF som SUM CRIT reflekterer, også kan måles gjennom gjennomsnittet av PFS’ skalaer.

Oppsummert viser resultatene forholdsvis sterke korrelasjoner mellom SUM CRIT og alle PFS variabler og dette kan ha flere alternative tolkninger. Det kan tyde på at personlighetspatologi er en gjennomgripende psykisk tilstand som kan ha innvirkninger på og komme til uttrykk i mange viktige livsarenaer som PFS skalaer måler. En annen mulig tolkning av korrelasjonene kan ligge i det metodiske valget for evaluering av PF. SUM CRIT er en kontinuerlig variabel beregnet kumulativt. På denne måten kan den gi et tydeligere og sterkere bilde på personlighetspatologi hos pasienten ettersom SUM CRIT fanger opp tilfeller der pasienter, til tross for tydelige kliniske symptomer, ikke når den kliniske terskelen som kreves i DSM diagnosesystemet for å få en PF diagnose. Det kan også tenkes at noen av pasientene får et kunstig forhøyet SUM CRIT tall dersom de oppfyller de kriterier som er felles for flere diagnoser. For eksempel i DSM-IV, står sinne oppført som et kriterium for både Borderline PF, Narsissistisk PF og Antisosial PF (Benjamin, 1996). Dersom noen av de oppfylte kriteriene er overlappende og summeres på tvers av PF-er, kan dette føre til en kunstig høy spredning i data som ytterligere kan resultere i sterke, men spuriøse korrelasjoner. Et annet aspekt som er betimelig å nevne er Perry’s (1992) forskningsfunn som viser til større uoverensstemmelse når bruk av selvrapporterte mål sammenliknes med kliniske (semi)strukturete intervjuer ved diagnostisering av PF enn ved sammenlikning av kliniske intervjuer innbyrdes. Dessuten peker studien mer spesifikt på manglende pålitelighet ved kliniske vurderinger i diagnostiseringen av PF og mer generelt på lav begrepsvaliditet i DSM-IV’s kriterier for PF, noe som også støttes av funnene til Westen (1997) og argumentasjonen til Clark (1999).

5.3 Sammenheng mellom PFS og SASB-IS

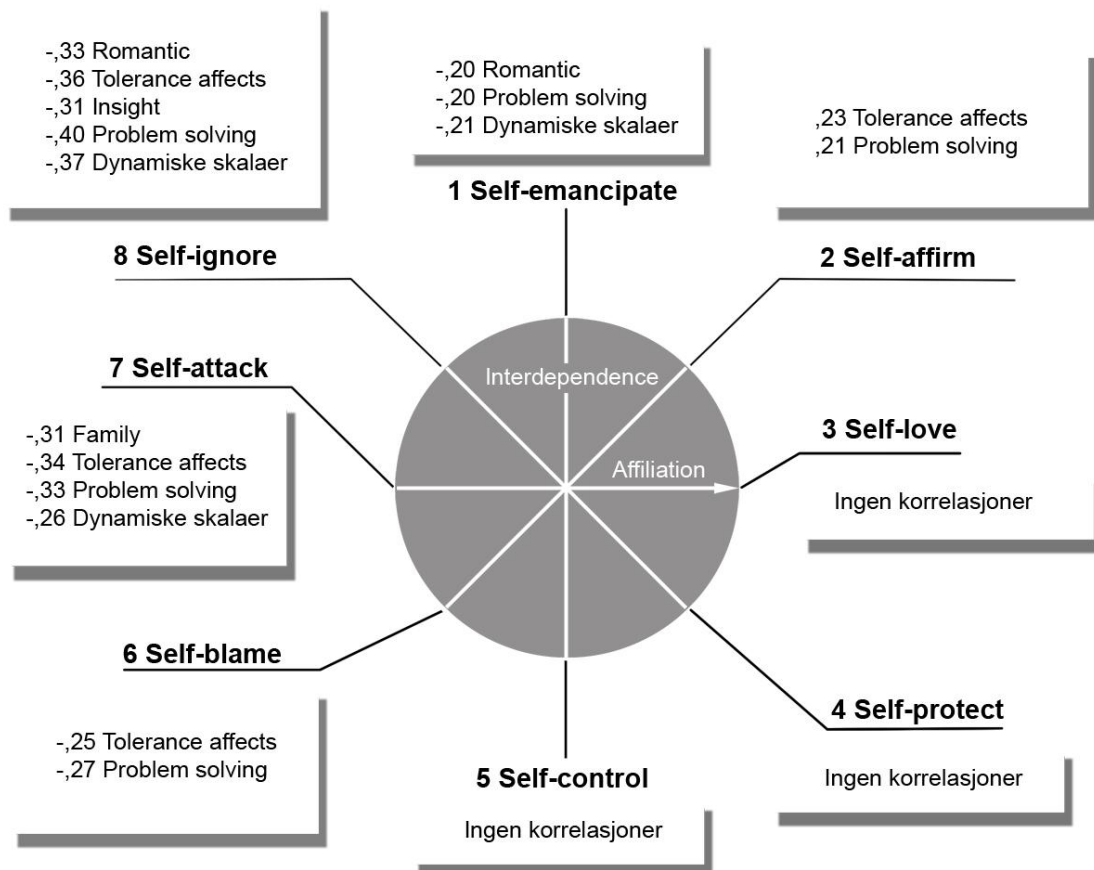
Tabell 6

Pearsons korrelasjonsverdier mellom SASB-IS kluster råskårer og PFS skalaskårer i FEST studien

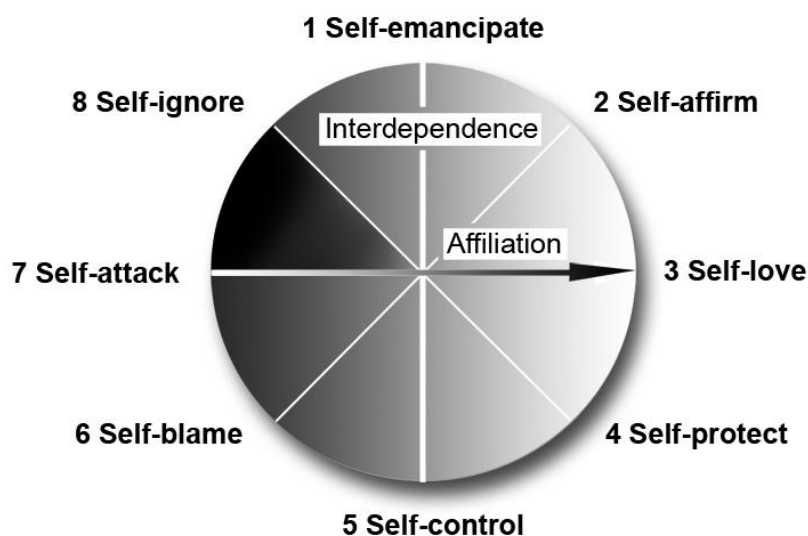
PFS scales	Family	Friends	Romantic	Tolerance affects	Insight	Problem solving	Dynamiske skalaer
SASB Introject scales							
Self emancipate	-0,16	-0,09	-0,20	-0,16	-0,19	-0,20	-0,21
Self affirm	0,14	0,18	-0,07	0,23	0,09	0,21	0,19
Self love	0,17	0,14	0,02	0,13	0,04	0,13	0,12
Self protect	0,06	0,00	0,04	0,04	0,02	0,09	0,05
Self control	0,04	-0,05	0,05	-0,10	-0,03	-0,09	-0,03
Self blame	-0,18	-0,11	-0,03	-0,25	-0,03	-0,27	-0,18
Self attack	-0,31	-0,11	-0,09	-0,34	-0,16	-0,33	-0,26
Self ignore	-0,18	-0,14	-0,33	-0,36	-0,31	-0,40	-0,37

*N mellom 89 og 99

Tall i **fet skrift: signifikante korrelasjoner på 0,05 nivå (2-tailed)



Figur 3. Oversikt over korrelasjoner mellom The Structural Analysis of Social Behavior – Introject Surface (SASB-IS) klusterskårer og PFS subskalaskårer. Kun signifikante korrelasjoner på 0.05 nivå (2-tailed).



Figur 4. The Structural Analysis of Social Behavior – Introject Surface: Distribusjon av psykopatologi langs horisontalaksen Affiliation. Venstre halvsirkel (Self-ignore, Self-attack, Self-blame) assosieres med psykopatologi, høyre halvsirkel (Self-affirm, Self-love, Self-protect) assosieres med psykisk sunnhet.

Hypoteser og faktiske funn

På bakgrunn av funn fra mange empiriske studier (Benjamin, 1994; Henry, 1996; Ruiz, Pincus, & Bedics, 1999) som viser et konsistent mønster som indikerer en sterk relasjon mellom Affiliation skalaen i SASB-IS og grad av psykopatologi og basert på det faktum at SASB-IS er et intrapsykisk mål, dannes følgende to hypoteser:

Hypotese 1. For *Self-ignore*, *Self-attack* og *Self-blame* klustere er det forventet å finne: (a) Negative korrelasjoner med PFS's gjennomsnittsskåre "Dynamiske skalaer"; (b) Flere og sterkere negative korrelasjoner med PFS' intrapsyriske skalaer enn med PFS' interpersonelle skalaer.

Faktiske funn. En rask visuell gjennomgang av Figur 3 viser at de fleste korrelasjoner (alle negative) befinner seg i den venstre halvsirkelen, hvor *Self-ignore*, *Self-Attack* og *Self-blame* regnes som de "negative", patologibetonete klusterne i SASB-IS. Disse korrelerer med flest skalaer i PFS, primært med "Tolerance affects" og "Problem solving".

Høye skårer på *Self-ignore* resulterer fra en blanding av uvennlig, destruktiv kritiskhet rettet mot seg selv, forsømmelse av og likegyldighet i forhold til eget utviklingspotensiale og

ignorering av egne behov (Benjamin, 1996). Dette kan tenkes å henge sammen med svake evner til å planlegge og oppnå meningsfylte mål, manglende retningsans i livet og en utpreget laissez-faire holdning, noe som vil resultere i en lav skåre på PFS' "Problem solving" skalaen (som gjenspeiler seg i en forholdsvis sterk korrelasjon på $-0,40$, $p < 0,05$, 2-tailed). I Figur 3, er *Self-ignore* klusteret plassert ganske langt unna *Self-control* klusteret og ganske nær *Self-attack* klusteret, en posisjon som beskriver en dumdristig person med et ikke godt nok definert selv og som utøver ukontrollert impulsivitet rettet mot seg selv, noe som kan henge sammen med dårlig affekt- og impulsregulering, samt et ganske unyansert affektrepertoar. Dette gjenspeiler seg i en ganske sterk korrelasjon med "Tolerance affects" på $-0,36$ ($p < 0,05$, 2-tailed). Den negative korrelasjonen på $-0,33$ ($p < 0,05$, 2-tailed) med "Romantic" gir mening sett i lys av at personen er styrt av ukontrollert impulsivitet rettet mot seg selv og ignorerer egne behov. Det siste kan henge sammen med lav selvvinnsikt, gjenspeilet i den negative korrelasjonen på $-0,31$ ($p < 0,05$, 2-tailed) med "Insight".

Self-attack er ytterpolen på hoveddimensjonen Affiliation, og i følge mange empiriske studier viser den sterkest assosiasjon med psykopatologi av de 8 SASB-IS klusterne (Monsen et al., 2007). En person med høye skårer på *Self-attack* avviser, ødelegger og angriper seg selv på en hatefull måte. Dette tyder på intrapsykiske tilstander preget av vansker for å romme og tåle konfliktfylte følelser og mangel på differensiering av affekt, noe som vil gi en ganske lav skåre på "Tolerance affects" skalaen (noe som gjenspeiler seg i den negative korrelasjonen med "Tolerance affects" på $-0,34$, $p < 0,05$, 2-tailed). En slik indre, nokså turbulent affekttilstand, kan tenkes å påvirke negativt evnen til problemløsning, noe som gjenspeiler seg i den negative korrelasjonen på $-0,33$ ($p < 0,05$, 2-tailed) med "Problem solving" skalaen.

Høy skåre på *Self-attack* kombinert med lav skåre på *Self-control* regnes som den mest patologiske introjekt tilstanden (Henry, 1996), noe som gjør *Selv-attack / Self-ignore* "hjørnet" til det meste patologiske i hele sirkumplekset. Her passer det å merke seg at *Selv-attack* og *Self-ignore* er de eneste klusterne PFS skåren "Dynamiske skalaer" korrelerer med, med korrelasjoner på $-0,26$ ($p < 0,05$, 2-tailed) , respektivt $-0,37$ ($p < 0,05$, 2-tailed), noe som kan være med på å styrke antakelsen at PFS sin gjennomsnittsskåre er en generell indikator på psykopatologi.

Høye skårer på *Self-blame* klusteret indikerer en person som er ganske selvdestruktiv, straffer, undertrykker og bebreider seg selv og viser mye tvil i forhold til egen person. Dette gjøres med en viss grad av selvkontroll (ref. Figur 3, hvor klusteret *Self-blame* er nærmere klusteret

Self-control), noe som tegner et litt mildere patologibilde enn klusteret *Self-ignore*, hvor innslag av impulsivitet er mer fremtredende. Her passer det å merke seg at korrelasjonsverdiene for PFS intrapsykiske skalaer ”Tolerance affects” og ”Problem solving” er lavere for *Self-blame* enn for *Self-ignore*, noe som kan styrke antakelsen at de to PFS intrapsykiske skalaer er et nyansert mål på psykopatologi. Samtidig må det huskes at det er blitt beregnet relativt lav reliabilitet på *Self-ignore* klusteret (Lorr & Strack, 1999).

Hypotese 2. For *Self-affirm*, *Self-love* og *Self-protect* klustere er det forventet å finne: (a) Positive korrelasjoner med PFS’s gjennomsnittsskåre ”Dynamiske skalaer”; (b) Flere og sterkere positive korrelasjoner med PFS’ intrapsykiske skalaer enn med PFS’ interpersonelle skalaer.

Faktiske funn. En rask visuell gjennomgang av Figur 3 viser et veldig lavt antall korrelasjoner mellom PFS skalaer og *Self-affirm*, *Self-love* og *Self-protect*, som regnes som de ”positive”, ”friske” (minst knyttet mot patologi) klusterne i SASB-IS.

Høye skårer på *Self-affirm* indikerer en person som er spontan med en balansert selvaksept. Dette kan tenkes å samsvare med gode mekanismer for affekt- og selvregulering som gjenspeiler seg i den positive korrelasjon på 0,23 ($p < 0,05$, 2-tailed) med ”Tolerance affects”. Personen er i tillegg åpen, nysgjerrig og ganske selvsikker, noe som kan antas å gi gode evner til å være selvhevdende uten utvikling av symptomer eller unnvikende atferd og fordrer (oppmuntrer til?) utforskning av nye arenaer. Slike egenskaper gir høye skårer på ”Problem solving” skalaen, noe som gjenspeiler seg i en moderat positiv korrelasjon på 0,21 ($p < 0,05$, 2-tailed).

Høye skårer på *Self-love* indikerer en person som har gode evner til å berolige seg selv, more seg, beskytte og pleie eget selv. Det kan virke rimelig å anta at en slik person er i besittelse av brukbare evner til følelsesmessig selvregulering og har et nyansert følelsesmessig repertoar, noe som kan gjenspeile seg i høyere skårer på ”Tolerance affects” og følgelig resultere i en positiv signifikant korrelasjon med *Self-love*. Resultatene viser at det ikke er funnet noen signifikante korrelasjoner med noen av PFS skalaene, noe som avkrefter den opprinnelige hypotesen formulert for dette SASB-IS klusteret.

Høye skårer på *Self-protect* indikerer en person som analyserer og gransker seg selv, bestreber seg på å bli fullendt og dyktig, ser velvillig på seg selv, håndterer og kontrollerer sitt eget selv på en beskyttende og positiv måte og som er åpen og nysgjerrig. Det kan virke rimelig å anta

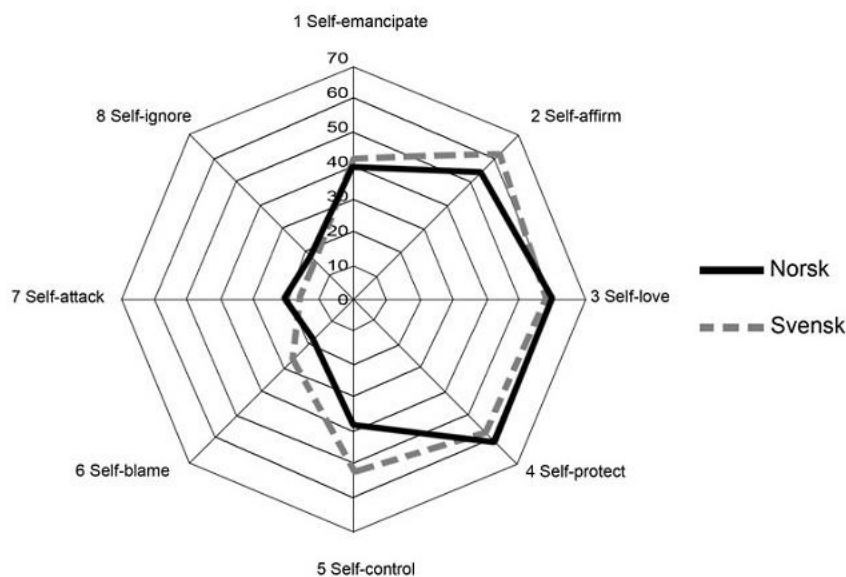
at en slik person er i besittelse av brukbare evner til følelsesmessig selvregulering og god selvinnsikt, med aksept for egne tanker og følelsesliv, som samtidig har god evne til planlegging og oppnåelse av personlige mål, dvs. skårer høyt på de tre intrapsykeiske PFS skalaer ("Tolerance affects", "Insight", "Problem solving"). Resultatene viser at det ikke er funnet noen signifikante korrelasjoner med noen av PFS skalaene, noe som avkrefter den opprinnelige hypotesen formulert for dette SASB-IS klusteret.

At de positive korrelasjoner mellom PFS og *Self-love / Self-protect* uteblir kan skyldes flere faktorer. Studiene til Lorr og Strack (1999) og Monsen et al. (2007) har vist at *Self-protect* har en lav reliabilitet og at den har en uklar faktorstruktur. Det kan tenkes at selve pasientutvalget brukt i FEST studien er skjevt i den forstand at det ikke finnes nok variasjon (spredning) i skårene for de sunne aspekter som SASB-IS måler, at variasjon i "sunnhet" er for fraværende i pasientutvalget for at den kan la seg fange opp. Over 58% av pasientene i utvalget er diagnostisert med ulike former for affektive lidelser – depresjon, dystymi, tilpasningsforstyrrelser, og 27% med ulike former for angstlidelse - panikklidelse, generalisert angstlidelse, sosiale fobier og agorafobi (Høglend et al., 2006). Fenomenet er kjent som "Restriction of range" (Alexander, 1988), hvor et for snevert utvalg av forsøkspersoner i forhold til en variabel eller et fenomen fører til reduserte eller fraværende korrelasjoner. Det kan også tenkes at PFS skalaene er bedre på å fange opp psykopatologi enn psykisk sunnhet, noe som i så fall vil være en svakhet ved PFS skalaene, da de har som mål å dekke et bredt spekter av fungeringsgrad fra svært dårlig til meget bra.

Benjamin (2003) hevder at profiler som befinner seg mot ytterpolene *Self-emancipate* og *Self-control* på aksene Gjensidig avhengighet representerer tegn på psykopatologi. De opprinnelige hypoteser listet i FEST studien tar ikke stilling til det fordi Benjamins antakelse ikke ser ut til å ha blitt empirisk validert i andre studier. I FEST data har *Self-emancipate* signifikante negative korrelasjoner med både PFS' "Dynamiske skalaer" (-0,21, $p < 0,05$, 2-tailed), "Romantic" (-0,20, $p < 0,05$, 2-tailed) og "Problem solving" (-0,20, $p < 0,05$, 2-tailed). Høye skårer på *Self-emancipate* betegner en altfor selvstendig person som er lite knyttet til andre mennesker, og den negative korrelasjonen med "Romantic" er derfor ikke uventet. Det er imidlertid flere studier som har påvist at klusteret *Self-emancipate* viser en lav reliabilitet (Lorr & Strack, 1999; Monsen et al., 2007) og dette må tas hensyn til i analysen av resultater. I tillegg viser polen *Self-emancipate* en to-delt faktor struktur (M. S. Halvorsen, personlig

kommunikasjon, 28. februar 2012) som er en metodesvakhet i SASB-IS sirkumpleks og som ytterligere kompliserer tolkningen av resultatene.

Generelt sett er det lettere å tolke skårer langs den horisontale aksene Affiliation som går fra *Self-attack* til *Self-love*, som et klarere mål på psykopatologi, hvor lave skårer på *Self-attack* og høye skårer på *Self-love* indikerer mental helse. Den vertikale skalaen Interdependence derimot er vanskeligere å tolke, og det er ikke lett å anta om det er bra å være lav eller høy på *Self-Emancipate* og *Self-control*. Skårer fra normalutvalget i den norske multisenterstudien til Havik et al. (1995) og i den svenske studien til Armelius utført i 2006 (M. S. Halvorsen, personlig kommunikasjon, 28. februar 2012) plottet i Figur 5, indikerer en plassering i sirkumplekset på de åtte klusterne som kan være assosiert med psykisk sunnhet: på den horisontale aksene er det skårer mot høyre, mens på den vertikale aksene er det skårer i midtområdet. Som nevnt tidligere, er *Self-attack* / *Self-ignore* sirkumpleksets mest patologiske ”hjørne”.



Figur 5³. The Structural Analysis of Social Behavior – Introject Surface skårer fra normalutvalg i en norsk og en svensk studie.

³ Mottatt på e-post fra M. S. Halvorsen, 28. februar 2011 (brukt i K. Armelius' presentasjon på M. S. Halvorsens' disputas, 15. desember 2006).

Oppsummert viser resultatene at "Tolerance affects" og "Problem solving" er de to PFS intrapsykiske dimensjoner som er mest fremtredende i korrelasjonene med SASB-IS klusterne. Dette gir mening ettersom SASB-IS teoretisk er ment som et mål på intrapsykiske aspekter og funnene tyder på en del konvergent validitet mellom disse instrumentene. Fravær av korrelasjoner mellom PFS skalaer (inkludert gjennomsnittsskåren "Dynamiske skalaer") og den "sunne" høyre halvsirkelen av SASB-IS sirkumplekset er det mest interessante og overraskende funnet. En forklaring kan være at PFS ikke måler psykiske styrker og positive evner godt nok, at det kan ligge noe uklart i formuleringene av beskrivelsene for de høyeste funksjonsintervallene (som gir retningslinjer for de "sunne", positive aspekter i pasientens psykodynamiske fungering) for hver PFS skala og som fører til at terapeutene ikke tildeler høye skårer selv om intervjumaterialet skulle tilsi det. Samtidig stammer data fra et pasientutvalg og det er ikke urimelig å tenke seg at pasientenes PFS skårer ligger klumpet sammen i lavere områder av psykodynamisk fungering, altså at det ikke er nok "friskere" pasienter til å danne en stor nok spredning i data som kan gjenspeile seg i større korrelasjonsverdier. Fra en teoretisk synsvinkel har PFS en større bredde enn SASB-IS, som har åtte subskalaer som alle går på den internaliserte atferden og som assosierer psykisk helse med vennlig selvaksept, og emosjonelle problemer med selvforsømmelse og selvkritiskhet (Henry, 1996).

PFS er ikke blitt validert på et referanseutvalg som kunne gi en indikasjon på spredning og gjennomsnittsskårer i en frisk populasjon. Dette skal i utgangspunktet ikke være nødvendig da alle PFS skalaer definerer hva som er normalt. De overraskende funnene vedrørende mangel på korrelasjoner med SASB-IS "friske" klustere, *Self-affirm*, *Self-love* og *Self-protect*, kan tilsi et behov for et valideringsprosjekt av PFS på et referanseutvalg.

5.4 Sammenheng mellom PFS og SCL-90-R

Tabell 7

Pearsons korrelasjonsverdier mellom SCL-90-R utvalgte skårer og PFS skalaskårer i FEST studien

PFS scales	Family	Friends	Romantic	Tolerance affects	Insight	Problem solving	Dynamiske skalaer
SCL-90-R							
Global Symptom Indeks (GSI)	-0,32	-0,24	-0,07	-0,47	-0,35	-0,48	-0,37
Depresjon	-0,28	-0,23	-0,13	-0,41	-0,28	-0,48	-0,35
Angst	-0,18	-0,13	-0,02	-0,42	-0,30	-0,35	-0,27
Mellommenneskelig overfølsomhet	-0,33	-0,24	-0,03	-0,43	-0,29	-0,44	-0,33

*N mellom 88 og 98

Tall i **fet skrift: signifikante korrelasjoner på 0,05 nivå (2-tailed)

En valideringsanalyse av PFS i forhold til alle de ni hovedsymptomdimensjoner og tre indekser som SCL-90-R består av ville bli for omfattende. Dermed ble det valgt Depresjon, Angst, Mellommenneskelig overfølsomhet og GSI. Rasjonalet bak utvalg av SCL-90-R variabler henger sammen med hypotesene som blir presentert nedenfor. Reliabiliteten for skalaene Depresjon, Angst, Mellommenneskelig overfølsomhet og GSI er 0,89, 0,87, 0,87 og 0,97 (Vassend et al., 1992) og anses dermed høyt nok for at skalaene kan analyseres selvstendig, løsrevet fra selvrapporterings skjemaets kontekst (Derogatis, 1994).

Hypoteser og faktiske funn

Hypotese 1. Mellommenneskelig overfølsomhet er en av de tre SCL-90-R symptomskalaer som har en markant interpersonlig karakter (Karterud et al., 1995). Leddene som inngår i skalaen fokuserer på følelser av personlig utilstrekkelighet og mindreverd, spesielt i sammenligninger med andre, følelser av uro og merket ubehag under mellommenneskelig samhandling. Ut i fra dette, blir følgende hypotese formulert:

- 1) Mellommenneskelig overfølsomhet variabelen er forventet å ha flere og sterkere negative korrelasjoner med PFS' interpersonelle variabler enn med PFS' intrapsykiske variabler.

Faktiske funn. Personer med høy score på mellommenneskelig overfølsomhet rapporterer forhøyet selvbevissthet og negative forventninger i kommunikasjon og samhandling med andre. Eksempler på ledd som inngår i skala: ”Føler seg blyg eller engstelig i forhold til det

motsatte kjønn”, ”Føler deg lett såret”, ”Føler at andre ikke forstår deg eller bryr seg om deg”, ”Føler deg underlegen eller mindreverdige”. En visuell gjennomgang av Tabell 7 viser at korrelasjonsverdiene ikke gir støtte til hypotesen. Mellommenneskelig overfølsomhet samvarierer sterkest med ”Problem solving”(r = - 0,44, p<0,05, 2-tailed), deretter med ”Tolerance affects” (r = - 0,43, p<0,05, 2-tailed), ”Dynamiske skalaer” og ”Family” (r = - 0,33, p<0,05, 2-tailed), etterfulgt av ”Insight” (r = - 0,29, p<0,05, 2-tailed) og ”Friends” (r = - 0,24, p<0,05, 2-tailed). Disse resultatene peker heller i motsatt retning i forhold til hypotesen, altså de viser at mellommenneskelig overfølsomhet variabelen har flere og sterkere negative korrelasjoner med PFS’ intrapsykiske variabler enn med PFS’ interpersonelle variabler. Dette funnet er imidlertid ikke lett å tolke. En mulig forklaring kan være en utydelig, ustabil faktorstruktur i SCL-90-R som er funnet i flere faktoranalytiske studier oppsummert av Cyr, McKenna-Foley, og Peacock (1985) og senere av Vassend og Skrandal (1999). Et alternativ er å problematisere, ut i fra de ledd som inngår i skalaen, i hvilken grad mellommenneskelig overfølsomhet adresserer intrapsykiske prosesser. I følge Karterud et al. (1995) relateres høye skårer til økt sårbarhet av selvet. I fravær av ytterligere teoretisk og empirisk grunnlag risikerer denne tolkningen å grense mot spekulative antakelser.

Hypotese 2. GSI indeksen er valgt fordi, slik det også er omtalt under beskrivelsen av instrumentet, den utgjør den beste enkeltindikatoren av aktuelt lidelsestrykk. Den kombinerer informasjon om antall symptomer og lidelses-intensitet (Derogatis, 1994).

- 2) GSI er forventet å korrelere sterkere med alle PFS variabler enn de øvrige SCL-90-R variabler (Mellommenneskelig overfølsomhet, Depresjon og Angst).

Faktiske funn. I undersøkelsen av denne hypotesen er man ikke opptatt av spørsmålet hvor godt SCL-90-R er egnet til å skille mellom de ulike psykiatriske diagnoser som er assosiert med det generelle lidelsestrykket, altså hvor godt det fungerer som et screeningsverktøy for psykiske diagnoser. Her er man, tvert imot, interessert i å bruke den kvantitative indikasjonen GSI gir på generelt psykologisk besvær og relatere den til PFS’s skalaer. Fra psykometrisk teori er det kjent at reliabiliteten øker når testlengden øker (under forutsetning av at de delene som blir lagt til måler det samme som den opprinnelige delen). GSI, som er gjennomsnittet av 90 testledd, kombinerer antall symptomer med intensiteten i det generelle lidelsestrykket og har dermed høyere reliabilitet enn de enkelte subskalaene når det gjelder indikasjon på generell lidelsestrykk (Vassend et al., 1992), noe som kan være med på å styrke funnene.

Dataene i Tabell 7 gir full støtte til hypotesen gjennom korrelasjonsverdier som ligger i spennet mellom $r = -0,24$ og $r = -0,48$, med unntak av "Romantic". Høyest korrelasjon registreres med "Problem solving" ($r = -0,48$, $p < 0,05$, 2-tailed), et resultat som tilsynelatende virker mindre intuitivt da beskrivelsen av "Problem solving" skalaen ikke gir de sterkeste umiddelbare assosiasjoner med psykopatologi av alle PFS skalaene. Samtidig adresserer "Problem solving" evner og ressurser som inngår i mange viktige livsområder, som for eksempel håndtering av vanskelige situasjoner, utforskning og integrering av det ukjente i det familiære, nytelse av rekreasjonsaktiviteter og fleksibel selvhevdelse. Disse områder kan være påvirket av psykiske problemer og det er antakelig dette som belyses gjennom denne korrelasjonen.

"Tolerance affects" viser også en ganske sterk negativ korrelasjonsverdi med GSI ($r = -0,47$, $p < 0,05$, 2-tailed). Vassend (1989) har funnet assosiasjoner mellom negativ affekt – subjektive følelser av nervøsitet, spenning, bekymring og sårbarhet - og GSI. GSI betraktes som en skala som dekker stort sett alle aspekter av negativ affekt (ibid.). At korrelasjonen 'kun' er på $-0,47$ betyr at PFS skalaen måler andre, mer psykodynamisk relaterte, aspekter ved affekter som for eksempel konfliktfylte følelser, mangel på ekthet og regulering av impulser.

"Dynamiske skalaer" har en korrelasjon $r = -0,37$ ($p < 0,05$, 2-tailed) med GSI, noe som tyder på at dynamisk fungering, sett fra de ulike aspekter som inngår i gjennomsnittsskåren "Dynamiske skalaer", er assosiert med subjektivt opplevd lidelsestrykk.

Hypotese 3. Depresjon og Angst symptomskalaene ble valgt fordi utvalget benyttet i FEST studien besto av 100 pasienter hvorav 58% var diagnostisert med Akse-I affektive lidelser og 27% med Akse-I ulike former for angst lidelse (Høglend et al., 2006). Funn fra en stor studie foretatt av Schmitz, Kruse, Heckrath, Alberti, og Tress (1999) viser at SCL-90-R har særlig god presisjon når det gjelder symptomområdene depresjon og angst og at instrumentet på en god måte identifiserer psykiske lidelser sett i forhold til en tysk versjon av SCID. En annen studie (Margo, Dewan, Fisher, & Greenberg, 1992) viser at SCL-90-R identifiserer klinisk depresjon like godt som Beck Depression Inventory (BDI; Beck & Beamesderfer, 1974). I lys av dette melder følgende spørsmål seg: hvor mye overlapp er det mellom psykisk lidelse (angst og depresjon) målt gjennom de to SCL-90-R variablene og PFS' variabler?

- 3) For Depresjon og Angst er det ikke formulert noen konkrete forventninger da PFS ikke er designet med tanke på å måle en spesifikk type patologi. Derfor foretas det en eksplorerende tilnærming til resultatene.

Faktiske funn. I undersøkelsen av denne hypotesen må det først problematiseres hvor godt SCL-90-R er egnet til å skille mellom de ulike psykiatriske diagnoser (i dette tilfellet angst og depresjon) som er assosiert med det generelle lidelsestrykket. Dette er i hovedsak et spørsmål om begrepsvaliditet. Selv om studiene til Schmitz et al. (1999) og Margo et al. (1992) viser at SCL-90-R kan fungere som et screeningsverktøy for psykiske diagnoser, viser andre studier (Pedersen & Karterud, 2004) at assosiasjonene mellom SCL-90-R's symptomskalaer og tilsvarende diagnoser ikke er sterke nok for en tilfredsstillende klinisk diskriminering.

En rask gjennomgang av data i Tabell 7 viser at Depresjon samvarierer i større grad med flere av PFS' skalaer enn med Angst. Depresjon har moderate korrelasjoner med alle PFS skalaer med unntak av "Romantic", mens Angst har ingen korrelasjoner med de tre interpersonlige PFS skalaene. Gjennomsnittsvariabelen "Dynamiske skalaer" korrelerer $r = -0,35$ ($p < 0,05$, 2-tailed) med Depresjon og $r = -0,27$ ($p < 0,05$, 2-tailed) med Angst, noe som antyder en moderat samvariasjon med stemnings- og angstlidelser. Imidlertid er det grunn til å utøve varsomhet i tolkningen av disse resultatene. For det første har SCL-90-R omdiskuterte egenskaper som screening verktøy (slik det er nevnt tidligere i avsnittet). For det andre er instrumentets faktorstruktur også omdiskutert, både når det gjelder stabilitet og antall. Dette belyses i Cyr et al.'s (1985) oppsummering av faktoranalytiske studier som konkluderer med at det er en høyst ustabil faktorstruktur i SCL-90-R, både på tvers av kjønn og på tvers av populasjoner og at instrumentet i hovedsak fanger opp én faktor: generell psykisk symptombelastning. Det sistnevnte samsvarer også med resultater fra studien til Vassend et al. (1992) som finner at den norske oversettelsen av SCL-90-R i hovedsak måler en generell negativ affekt-faktor. SCL-90-R skårene brukt i korrelasjonsberegningen er rå, altså ikke ipserte. Det ble ikke gjennomført noe strukturert intervju i forbindelse med diagnostiseringen av Akse-I lidelser. Fastsettelse av Akse-I diagnose ble foretatt basert på alt tilgjengelig klinisk materiale som ble diskutert mellom terapeutene til konsensus var oppnådd i overensstemmelse med diagnostiske kriterier i DSM-III-R manualen.

Oppsummert viser resultatene at "Problem solving" og "Tolerance affects" har sterkest samvariasjon med SCL-90-R variablene, høyest av alle med GSI, noe som gir mening da GSI er et samlet uttrykk på et generelt, subjektivt opplevd psykisk ubehag. Det er imidlertid

vanskelig å tolke det faktum at "Dynamiske skalaer" ikke korrelerer høyest med GSI av alle PFS skalaer. "Romantic" korrelerer ikke med noen av SCL-90-R variabler, noe som ellers er et gjennomgående trekk ved denne skalaen.

6 Begrensninger i studien

I tillegg til de metodologiske begrensninger knyttet til måleinstrumenter presentert under 1.3.4, må resultatene i denne studien ses i forhold til en rekke begrensninger som hører til hovedoppgaven. Tilgangen til datamaterialet innsamlet i FEST studien har vært begrenset til kun korrelasjonsverdier som representerer målinger som er utført rett før starten av behandlingen (pre-treatment). PFS er et instrument som er laget med tanke på å måle spesifikke effekter av psykodynamisk terapi. En fullverdig analyse av PFS' evne til å fange opp slike forandringer ville innebære tilgang til data også fra post-treatment og follow-up.

Begrenset tilgang til datamaterialet har forhindret å teste i hvilken grad data innsamlet med IIP-C og SASB-IS for FEST studiens utvalg tilfredsstillt kravene til sirkumpleksmodeller. For disse to instrumentene har det ikke vært mulig å sammenlikne ipserte med ikke-ipserte skårer ettersom ipserte skårer ikke har vært tilgjengelig.

Det kan ligge begrensninger i tolkningene av resultatene når det gjelder overføringsverdien til andre kliniske utvalg. Valideringsstudien av PFS bygger på data samlet fra kun én studie (FEST). Dette kan virke som en begrensning når det gjelder generaliserbarheten til andre kliniske utvalg ettersom validering av et instrument er en kontinuerlig prosess som knyttes til *tolkningen av resultater* og til *slutningen* forskeren gjør fra skårer til det begrepet man tror man måler. I følge Paap et al. (2011) kan en skala ha ulike psykometriske egenskaper når den brukes på forskjellige populasjoner. Validitet er dermed avhengig av den spesifikke konteksten den omtales i (type design, forsøkspersoner, hva ønskes målt og til hvilket formål) og må derfor utsettes for gjentakende analyse og forskning, noe som indikerer en implisitt begrensning i generaliserbarhet.

En annen begrensning er at oppgaven ikke er styrt av klare teoretiske rammer eller psykologiske modeller, noe som har sitt opphav i konseptualiseringen av instrumentet PFS som har prioritert et perspektiv preget av "klinisk nytteverdi" framfor "teoretisk tydelighet". Manglende analyser fra tilsvarende studier som kunne ha gitt mulighet til empiriske sammenlikninger og utvidede teoretiske drøftinger bidrar til at denne studien må betraktes som eksplorerende.

Utvalget i studien er på 100 pasienter. Det er rimelig å spørre seg om hvordan resultatene hadde sett ut dersom utvalget hadde vært større. Et annet spørsmål er om utvalget er

representativt nok for norske pasienter med psykiske lidelser for en valideringsstudie som fokuserer på patologi. Deltakerne i studien har meldt seg frivillig og det i seg selv innebærer en seleksjonsprosess som kan indusere en skjevhet i utvalget.

Signifikansnivået for resultatene er satt til 5%, og i praksis betyr det at det alltid er 5% sjanse til stede for at signifikante korrelasjoner er tilfeldige (Wampold, 2001).

Psykopatologi begrepet er ikke blitt klart definert i oppgaven og det henger sammen med det faktum at PFS ikke er designet med tanke på å måle en spesifikk type patologi.

7 Oppsummering og konkluderende betraktninger

På et overordnet nivå forsøker denne studien å behandle et av psykologiens grunnproblem. Det handler om det faktum at begreper som innebærer komplekse prosesser slik ”psykodynamisk fungering” er, ikke uten videre kan reduseres til noen enkle størrelser som kan måles ved et enkeltstående instrument. Siden fenomenene ikke kan måles direkte, må en arbeide med fortolkninger av observasjoner. Psykometriske teorier prøver å danne en vitenskapelig plattform for hvordan konstrukt, tester og måling av konstruktets operasjonalisering skal forstås i en betydningsbærende sammenheng. I en slik kontekst kommer reliabilitet og validitet inn som nødvendige vitenskapelige konsepter som må undersøkes.

Resultatene fra denne studien er todelt og en oppsummering per del er presentert nedenfor.

7.1 Oppsummering av Del I – reliabilitet

I første del er det blitt gjennomgått reliabilitetsegenskaper til PFS instrumentet i lys av publiserte artikler. Det er omtalt to typer reliabilitetsestimater innenfor rammene av to ulike psykometriske teorier.

(1) Innenfor Klassisk testteori (CTT) ble reliabilitet belyst ved hjelp av intraklasse-korrelasjonskoeffisient (ICC), som estimerer interrater reliabilitet. ICC er et egnet mål for ordinalskalaer som skåres av flere bedømmere. Lavest reliabilitetsverdi ble estimert, både pre- og post-treatment, for skalaene ”Insight” og ”Tolerance of affects”. Tilsvarende resultater ble funnet i en eldre studie som benyttet PFS. Høyest reliabilitetsverdi ble estimert for ”Quality of romantic relationships”, både pre- og post-treatment, i begge studier.

Det kan konkluderes med at PFS skaler oppnår meget god interrater reliabilitet når PFS brukes av tre ratere, og at de tre intrapsykiske skalaene viser svakest reliabilitet, antakelig pga deres begrepskompleksitet.

(2) Innenfor Generalizabilityteori (GT) følger FEST studien en to-fasett mixed modell, med fasetten ’PFS skalaer’ festet (”fixed”) til seks skalaer og fasetten ’ratere’ tilfeldig. Reliabilitet

ble estimert ved hjelp av generalizability koeffisienten (G-koeffisienten) $\hat{\rho}^2$, samt error/tolerance forholdet ($(E/T)^2$), som representerer presisjon i estimering av G-koeffisienten.

I FEST studien er G-koeffisienten for PFS gjennomsnittsskåren ”Dynamiske skalaer” beregnet både for *relative* skåreverdi ($\hat{\rho}_{rel}^2$), altså når pasientens skåre ses i forhold til de andres skårer i gruppen (rangordning) og for *absolutte* skåreverdi ($\hat{\rho}_{abs}^2$), altså når pasienten er sin egen målestokk, og hans skåre ses i forhold til egen tidligere skåre. Dette kalles også for differanseskåre (change scores) og de brukes mye i forskning på individuell forandring i terapi. Den høyeste G- koeffisienten og den beste presisjonen i dens estimering er oppnådd for absolutte skårer. Dette tyder på at PFS er sensitiv nok til å fange opp generaliserbare forandringer etter terapi, noe som var den opprinnelige tanken ved å lage instrumentet. Større antall ratere gir både større reliabilitet og høyere presisjon i reliabilitetens estimering.

7.2 Oppsummering av Del II – validitetsstudie

Denne delen av oppgaven har prøvd å kaste lys over spørsmålet i hvilken grad ”psykodynamisk fungering”, slik den er operasjonalisert gjennom PFS’ seks skalaer og gjennomsnittsskåren, kan assosieres med psykopatologi ved å undersøke sammenhengen mellom PFS og fire instrumenter som måler ulike aspekter av psykopatologi. Å undersøke sammenhengen mellom to fenomener representerer en innledende forståelse av fenomenene og kan være viktig for videreutvikling av teoretisk og metodisk grunnlag for ytterligere undersøkelse og måling av fenomenet (Kazdin, 1999), noe som kan anses å være en overordnet implikasjon av denne studien. En svært høy samvariasjon mellom ulike mål kan bety at disse bare er alternative mål for samme fenomen. Ved korrelasjoner mellom PFS og andre mål på psykopatologi som er over +/- 0,80, vil de fleste forskere og klinikere si at det er bortkastet å lage og bruke PFS. En kan like gjerne bruke det andre instrumentet i stedet.

Resultatene som har kommet fram av validitetsanalysen viser en varierende grad av sammenfall mellom PFS skalaene og eksterne mål på psykopatologi, noe som på et overordnet nivå verken er overraskende eller regnes som et spesielt funn. Selv om målene fanger opp ulike fenomener, vil en imidlertid forvente en viss samvariasjon mellom målene. Dette fordi fenomenene instrumentene måler er gjensidig avhengig av hverandre og fanger

opp en del ”generell patologi”. PFS’ skalaene adresserer individets evner og ressurser på essensielle områder som mellommenneskelige relasjoner, affektregulering, innsikt, fungering i dagliglivet gjennom problemløsning og fleksibel tilpasning til nye situasjoner. Dette er områder som antas å bli affisert ved tilstedeværelse av psykopatologi. Det vil derfor forventes en moderat samvariasjon mellom skårer på de ulike instrumentene.

Det er vanlig at det oppnås lavere korrelasjoner når det benyttes instrumenter som måler fra ulike perspektiv, for eksempel selvrapporterte mål og semistrukturerte intervju (Perry, 1992) og selvrapporterte mål og observatørmål (Hill et al., 1998). I tillegg viser funn fra mange studier at terapeuten og pasienten ser ulikt på mange områder knyttet til psykopatologi og terapi (Weiss, Rabinowitz, & Spiro, 1996; Hersoug, 2010; Hersoug, Høglend, Monsen, & Havik, 2001; Hersoug, Høglend, Havik, von der Lippe, & Monsen, 2009), noe som bidrar til en implisitt svekkelse av sammenhengen mellom målene.

Følgende er en oppsummering av de mest interessante funn ut ifra data i Tabell 4, 5, 6 og 7.

PFS ”Romantic” skalaen har gjennomgående, på tvers av data fra alle fire tabeller, færrest korrelasjoner av alle PFS skaler. Den fremstår i denne konteksten som en skala som fanger opp forhold som ikke finnes i de andre instrumentene og det kan reises spørsmål om hva denne skalaen egentlig måler. Resultatet er enda mer overraskende ettersom skalaen har fått høyest konsensus på innhold (Bøgwald & Dahlbender, 2004), noe som tyder på høy innholdsvaliditet. Det kan tenkes at beskrivelsen av de ti kvalitative nivåer i skalaen ikke er tydelig nok differensiert eller at materialet rundt seksuelle forhold ikke kommer så lett fram ved første intervju.

I motsatt ende ligger PFS ”Problem solving” skalaen som har flest korrelasjoner av alle PFS skalaer, flere enn den aritmetiske gjennomsnittsskåren ”Dynamiske skalaer”. Dette kan reise spørsmål om hva denne skalaen egentlig måler – fanger den noe av kjernen ved menneskelig fungering som er gjennomgående affisert av psykopatologi? Selv om den er definert som en intrapsykisk skala, kan det tenkes at den indirekte fanger opp andre aspekter i tillegg (som for eksempel interpersonlige aspekter) som gjør at den korrelerer mer gjennomgående med mål på psykopatologi. Det er noe som må undersøkes nærmere i andre studier.

Det er en overraskende mangel på korrelasjoner mellom PFS skalaene og den ”friske høyredelen” av SASB-IS (Tabell 6, Figur 3). Dette melder seg som et signifikant spørsmål siden PFS er tiltenkt å dekke et bredt spekter av grad av fungering fra svært dårlig til meget

bra. Resultatet kan tyde på at PFS ikke måler psykiske styrker og positive evner godt nok, at det kan ligge noe uklart i formuleringene av beskrivelsene for de høyeste funksjonsintervaller som PFS skåringen baserer seg på. I følge Benjamin (1996) bør en god teori forklare forskjellen mellom normalitet og patologi og som nevnt tidligere mangler PFS et entydig solid teoretisk grunnlag. ”Tolerance affects” og ”Problem solving” er de to PFS intrapsykiske dimensjoner som er mest fremtredende i korrelasjonene med SASB-IS klusterne og det gir mening siden SASB-IS er et intrapsykisk mål. Samtidig korrelerer ”Tolerance affects” høyest med ”Problem solving” av alle PFS skalaer (Tabell 3), noe som kan tyde på at de fanger opp fenomener som er altfor relaterte og at ikke begge skalaer er nødvendige.

I Tabell 7 viser PFS skalaene flere og til dels høyere korrelasjoner med SCL-90-R depresjonskala enn med angstskala (forskjellen er ikke signifikanstestet). Disse funn må ses i lys av at PFS ikke er designet med tanke på å måle en spesifikk type patologi. Denne tanken støttes også opp av at de fleste, også de høyeste korrelasjoner er med GSI, som er en enkeltindikator på aktuelt lidelsestrykk.

På tvers av data fra de fire tabeller, viser PFS ”Dynamiske skalaer” høyest korrelasjon med SUM CRIT. En mulig tolkning er at SUM CRIT, som er indikasjon på PF, belyser en tilstand som ikke kjennetegnes av en brå start, men bærer med seg en viss historikk av psykisk ubehag, noe som gjør SUM CRIT til et ikke ”her og nå” mål. Denne karakteristikken er felles med PFS, som måler evner, tendenser, historikk – for eksempel, kan man skåre høyt på ”Romantic” skalaen selv om en ikke har noen partner på nåværende tidspunkt, men har vist tidligere å ha evnen til å danne en god romantisk relasjon (P. Høglend, personlig kommunikasjon, 03. mai 2012). I motsetning til det, er IIP-C, SCL-90-R og SASB-IS status deskriptive mål, som viser hvordan pasienten har det ”nå”. Det kan også tenkes at tilstedeværelse av PF som er assosiert med økte interpersonlige problemer er gjennomgripende på andre arenaer av mer intrapsykisk karakter. På et mer abstrakt plan, reiser dette spørsmål om hvor dypt skillet mellom interpersonlige og intrapsykiske evner slik de er definert i og målt med PFS går. Dette grenser mot en diskusjon av operasjonaliseringen av konstruktet ”psykodynamisk fungering”, noe som er utenfor denne studiens rammer.

7.3 Konkluderende betraktninger

Funnene i denne studien indikerer et ganske kompleks sammenheng mellom PFS og psykopatologi. PFS synes til å fange opp en viss mengde generell psykopatologi gjennom korrelasjonene med SCL-90. Dette belyses særlig gjennom korrelasjonene med GSI, som omtales som den beste enkeltindikatoren av aktuelt lidelsestrykk.

Videre viser PFS et slektskapsforhold med patologi-relaterte dimensjoner i de andre tre instrumenter som for eksempel Self-attack i SASB-IS, Kald i IIP-C og SUM CRIT relatert til PF. På tross av dette, er det fortsatt uklart hvilke mer konkrete typer patologi PFS måler.

Vanskene med å besvare dette spørsmålet er knyttet blant annet til det faktum at instrumentene i seg selv ikke har et klart definert psykopatologi-område som de er mål på og fokuserer på forskjellige aspekter av menneskelig fungering. For eksempel er SASB-IS ikke relatert til noen spesifikke symptomområder og fokuserer på den faktiske atferden, uten å knytte den mot noen prototypiske interpersonlige problemer (Benjamin, 1974, 1994, 1996b). IIP-C derimot fokuserer på motivasjonen som ligger bak og utgjør utgangspunktet for ulike typer atferd (Horowitz, 2004), med fokus på interpersonlige problemer.

Dette viser at ulike instrumenter gir ulikt bidrag i prosessen av kartlegging av psykopatologi. Ulike målemetoder har ulike feilkilder og omfanget av målefeil kan variere fra instrument til instrument (Haynes et al., 1999). I klinisk kartleggingssammenheng er det interessant å få data som kan gi et mer meningsfullt og utfyllende bilde om pasientens psykiske lidelse og klinisk status, noe som ikke kan oppnås ved bruk av kun et instrument. Med teoretiske røtter i Campbell og Fiske (1959) sin multitrekk-multimetode-matrise sikter man på å skille mellom varians som stammer fra *fenomenet* som måles og varians som stammer fra *typen metode* som brukes i måling av fenomenet. Et biprodukt av analysen av konvergent og divergent validitet er viktigheten av å ha flere ulike perspektiver og metoder som måler til dels relaterte, men også til dels ulike begreper.

Denne oppgaven finner at det spesielle ved psykodynamisk fungering slik den er målt med PFS er at ingen av de fire instrumentene isolert sett evner å belyse bedre enn de andre andelen psykopatologi som er sammenflettet med psykodynamisk fungering.

Dette kan tilsi at det er relevant å bruke flere og ulike metoder og instrumenter i klinisk arbeid. Funnene indikerer også at det kan virke hensiktsmessig å vurdere psykodynamisk

fungering som et bredt konstrukt; et konstrukt som både er relatert til psykopatologi og til ikke-patologisk fungering.

Den komplekse sammenhengen mellom PFS og psykopatologi vil gi implikasjoner for benyttelsen av PFS i både forskningsmessige og kliniske settinger. Studien kan også bidra til en bredere forståelse av begrepet ”psykodynamisk fungering” slik den er operasjonalisert gjennom PFS. Studien belyser relasjonen mellom PFS og ulike mål på psykopatologi. Den gir en utvidet beskrivelse av PFS’ skalaer og hvordan de kan brukes sammen med andre instrumenter ved kartlegging av forandring etter terapi. Studien representerer således et bidrag til å belyse en forbindelse mellom forskningsområdene psykometri og kartlegging av psykopatologi og effekter av terapi.

”Selv om analysen avsluttes, innebærer ikke det at det ikke er mer å analysere”. S. Freud

Litteraturliste

- Alden, L. E., Wiggins, J. S., & Pincus, A. L. (1990). Construction of circumplex scales for the Inventory of Interpersonal Problems. *Journal of Personality Assessment*, *55*, 521–536. doi: 10.1080/00223891.1990.9674088
- Alexander, R. A. (1988). Group homogeneity, range restriction, and range enhancement effects on correlations. *Personnel Psychology*, *41*, 773–777. doi: 10.1111/j.1744-6570.1988.tb00653.x
- American Psychiatric Association (1994). *Diagnostic and Statistical Manual of Mental Disorders: DSM-IV* (4. utg.). Washington, DC: American Psychiatric Association.
- American Psychiatric Association (2000). *Diagnostic and Statistical Manual of Mental Disorders: DSM-IV-TR* (4. utg., revidert tekst). Washington, DC: American Psychiatric Association.
- Beck, A. T., & Beamesderfer, A. (1974). Assessment of depression: The Depression Inventory. I P. Pichot & R. Oliver-Martin (Red.), *Psychological measurements in psychopharmacology: Modern problems of pharmacopsychiatry* (s. 151-169). Basel: S. Karger.
- Bellak, L. (1969). Research on ego function patterns: A progress report. Rating scales for scoring ego functions from clinical interview material. I L. Bellak & L. Loeb (Red.), *The schizophrenic syndrome* (s. 11-65, 789-833). New York: Grune & Stratton.
- Benjamin, L. S. (1974). Structural Analysis of Social Behavior. *Psychological Review*, *81*, 392-425. doi: 10.1037/h0037024
- Benjamin, L. S. (1988). *Intrex users' manual*. Madison, WI: Intrex Institute.
- Benjamin, L. S. (1994). SASB: A bridge between personality theory and clinical psychology. *Psychological Inquiry*, *5*, 273-316. doi: 10.1207/s15327965pli0504_1
- Benjamin, L. S. (1996). *Interpersonal diagnosis and treatment of personality disorders*. New York: Guilford Press.

- Benjamin, L. S. (1996b). Introduction to the special section on Structural Analysis of Social Behavior. *Journal of Consulting and Clinical Psychology, 64*, 1203-1212. doi: 10.1037/0022-006X.64.6.1203
- Benjamin, L. S. (2003). *Interpersonal reconstructive therapy: Promoting change in nonresponders*. New York: Guilford Press.
- Bion, W. R. (1963). *Elements of psychoanalysis*. London: Karnac.
- Brennan, R. L. (1992a). *Elements of generalizability theory*. Iowa City, Iowa: ACT.
- Bøggwald, K. P., & Dahlbender, R. (2004). Procedures for testing some aspects of the content validity of the Psychodynamic Functioning Scales and the Global Assessment of Functioning Scale. *Psychotherapy Research, 14*, 453-468. doi:10.1093/ptr/kph038
- Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin, 56*, 81-105. doi: 10.1037/h0046016
- Clark, L. A. (1999). Dimensional approaches to personality disorder assessment and diagnosis. In C. R. Cloninger (Red.), *Personality and psychopathology* (s. 219-244). Washington, DC: American Psychiatric Press.
- Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2. utg.). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- Cook, T. D., & Campbell, D. T. (1979). *Quasi-experimentation: Design and analysis issues for field settings*. Boston: Houghton Mifflin.
- Coughlin Della Selva, P. (1996). *Intensive short-term dynamic psychotherapy: Theory and technique synopsis*. New York: Wiley & Sons.
- Cramer, V., Torgersen, S., & Kringlen, E. (2006). Personality disorders and quality of life. A population study. *Comprehensive Psychiatry, 47*, 178– 184. doi: 10.1016/j.comppsy.2005.06.002
- Crocker, L., & Algina, J. (1986). *Introduction to classical and modern test theory*. Fort Worth: Holt, Rinehart and Winston.
- Cronbach, L. J. (1949). *Essentials of psychological testing*. New York: Harper.

- Cronbach, L. J., Gleser, G. C., Nanda, H., & Rajaratnam, N. (1972). *The dependability of behavioral measurements. Theory of generalizability for scores and profiles*. New York: Wiley & Sons, Inc..
- Cronbach, L. J., Rajaratnam, N., & Gleser, G. C. (1963). Theory of generalizability: A liberalization of reliability theory. *British Journal of Statistical Psychology*, *16*, 137-163. doi: 10.1111/j.2044-8317.1963.tb00206.x
- Cyr, J. J., McKenna-Foley, J. M., & Peacock, E. (1985). Factor structure of the SCL-90-R: Is there one? *Journal of Personality Assessment*, *49*, 571-578. doi: 10.1207/s15327752jpa4906_2
- Derogatis, L. R. (1994). *SCL-90-R: Administration, scoring and procedures manual* (3. utg.). Minneapolis, MN: NCS Pearson.
- Derogatis, L. R., Lipman, R. S., Rickels, K., Uhlenhuth, E. H., & Covi, L. (1974). The Hopkins Symptom Check List (HSCL): A self-report symptom inventory. *Behavioral Science*, *19*, 1–15. doi: 10.1002/bs.3830190102
- Derogatis, L. R., & Unger, R. (2010). Symptom Checklist-90-Revised. I I. B. Weiner & W. E. Craighead (Red.), *The Corsini Encyclopedia of Psychology* (4. utg., vol. 4, s. 1743-1744). New York: John Wiley & Sons, Inc.
- Etchegoyen, R. H. (1991). *The fundamentals of psychoanalytic technique*. London: Karnac Books.
- Eysenck, H. J. (1952). The Effects of Psychotherapy: An Evaluation. *Journal of Consulting Psychology*, *16*, 319–324. doi: 10.1037/h0063633
- Fleiss, J. L. (1981). *Statistical methods for rates and proportions* (2. utg.). New York: Wiley & Sons, Inc.
- Fenichel, O. (1945). *The psychoanalytic theory of neurosis*. New York: W.W. Norton.
- First, M. B., Gibbon, M., Spitzer, R. L., Williams, J. B. W., & Benjamin, L. S. (1997). *User's guide for the Structured Clinical Interview for DSM-IV Axis II Personality Disorders, (SCID-II)*. Washington, DC: American Psychiatric Press.
- Freud, A. (1974). *Jeg'et og forsvarsmekanismene*. Oslo: Gyldendal Norsk Forlag.

- Freud, A., Nagera, H., & Freud, W. E. (1965). Metapsychological assessment of the adult personality: The adult profile. *Psychoanalytic Study of the Child*, 20, 9-41. Hentet fra <http://www.pep-web.org/document.php?id=psc.020.0009a>
- Friborg, O. (2010). Klassisk testteori og utvikling av spørreinstrumenter. I M. Martinussen (Red.), *Kvantitativ forskningsmetodologi i samfunns- og helsefag*. (s. 15-55). Bergen: Fagbokforlaget.
- Froyd, J. E., Lambert, M. J., & Froyd, J. D. (1996). A review of practices of psychotherapy outcome measurement. *Journal of Mental Health*, 5, 11-15. doi: 10.1080/09638239650037144
- Furr, R. M. (2011). *Scale construction and psychometrics for social and personality psychology*. Los Angeles: Sage.
- Ghiselli, E. E., Campbell, J. P., & Zedeck, S. (1981). *Measurement theory for the behavioral sciences*. San Francisco: W.H. Freeman.
- Guldberg, C. A., & Svendsen, Ø. (2003). Psykologiske forsvarsmekanismer. I J.O. Johannessen & B. G-R. Thorsen (Red.), *Den gode psykoterapeut* (s. 87-93). Stavanger: Hertervig Forlag.
- Gullestad, S. E., & Killingmo, B. (2005). *Underteksten:Psykoanalytisk terapi i praksis*. Oslo: Universitetsforlaget.
- Hagtvet, K., & Høglend, P. (2008). Assessing precision of change scores in psychodynamic psychotherapy: A generalizability theory approach. *Measurement and Evaluation in Counseling and Development*, 41, 162-178. Hentet fra: <http://search.proquest.com/docview/195605927/fulltextPDF?accountid=14699>
- Havik, O. E., Monsen, J. T., Høglend, P., von der Lippe, A., Lyngstad, G., Stiles, T., & Svartberg, M. (1995). *Norwegian Multi-site Study of Process and Outcomes in Psychotherapy (NMSPOP)*. [Research protocol. For information: Odd E. Havik, Department of Clinical Psychology, University of Bergen, N-5015 Bergen].

- Haynes, S. N., Nelson, K., & Blaine, D. D. (1999). Psychometric issues in assessment research. I P. C. Kendall, J. N. Butcher, & G. N. Holmbeck (Red.), *Handbook of research methods in clinical psychology* (s. 125-154). New York: John Wiley & Sons, Inc.
- Henry, W. P. (1996). Structural Analysis of Social Behavior as a common metric for programmatic psychopathology and psychotherapy research. *Journal of Consulting and Clinical Psychology, 68*, 1263-1275. doi: 10.1037/0022-006X.64.6.1263
- Heppner, P. P., Kivlighan, D. M. Jr., & Wampold, B. E. (1999). *Research design in counseling* (2. utg.). Belmont, CA: Brooks/Cole Wadsworth.
- Hersoug, A. G. (2010). A closer look at good and poor outcome in psychodynamic psychotherapy: A case comparison study. *Clinical Case Studies, 9*, 135-153. doi: 10.1177/1534650110364415
- Hersoug, A. G. (2011). Tolkning av overføringer gir innsikt og langtidseffekt. *Tidsskrift for Norsk Psykologforening, 48*, 88-89. Hentet fra: http://www.psykologtidsskriftet.no/index.php?seks_id=133396&a=4
- Hersoug, A. G., Høglend, P., Havik, O., von der Lippe, A., & Monsen, J. T. (2009). Therapist characteristics influencing the quality of alliance in long-term psychotherapy. *Clinical Psychology and Psychotherapy, 16*, 100–110. doi: 10.1002/cpp.605
- Hersoug, A. G., Høglend, P., Monsen, J. T., & Havik, O. E. (2001). Quality of working alliance in psychotherapy: Therapist variables and patient/therapist similarity as predictors. *The Journal of Psychotherapy Practice and Research, 10*, 205-216. Hentet fra <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3330657/pdf/205.pdf>
- Hill, C. E., & Lambert, M. J. (2004). Methodological issues in studying psychotherapy processes and outcomes. I M. J. Lambert (Red.), *Bergin and Garfield's Handbook of psychotherapy and behavior change*. (5. utg., s. 84-135). New York: John Wiley & Sons, Inc.
- Hill, R. W., Zrull, M. C., & McIntire, K. (1998). Differences between self- and peer ratings of interpersonal problems. *Assessment, 5*, 67-83. doi: 10.1177/107319119800500109

- Holmes, J. (2001). *The search for the secure base: Attachment theory and psychotherapy*. Hove: Brunner-Routledge.
- Horowitz, L. M. (2004). *Interpersonal foundations of psychopathology*. Washington, DC: American Psychological Association.
- Horowitz, L. M., Alden, L. E., Wiggins, J. S., & Pincus, A. L. (2000). *Inventory of Interpersonal Problems: Manual*. San Antonio, TX: The Psychological Corporation.
- Horowitz, L. M., Rosenberg, S. E., Baer, B. A., Ureña, G., & Villaseñor, V. S. (1988). Inventory of Interpersonal Problems: Psychometric properties and clinical applications. *Journal of Consulting and Clinical Psychology, 56*, 885-892. doi: 10.1037/0022-006X.56.6.885
- Horowitz, L. M., Rosenberg, S. E., & Bartholomew, K. (1993). Interpersonal problems, attachment styles, and outcome in brief dynamic psychotherapy. *Journal of Consulting and Clinical Psychology 6*, 549–560. doi: 10.1037/0022-006X.61.4.549
- Horowitz, L. M., Wilson K. R., Turan B., Zolotsev, P., Constantino, M. J., & Henderson L. (2006). How interpersonal motives clarify the meaning of interpersonal behavior: A revised circumplex model. *Personality and Social Psychology Review, 20*, 67-86. doi: 10.1207/s15327957pspr1001_4
- Huber, D., Brandl, T., & Klug, G. (2004). The scales of psychological capacities: Measuring beyond symptoms. *Psychotherapy Research, 14*, 89-106. doi: 10.1093/ptr/kph006
- Høglend, P. (2004). Analysis of transference in psychodynamic psychotherapy: A review of empirical research. *Canadian Journal of Psychoanalysis, 12*, 279-300.
- Høglend, P., Amlo, S., Marble, A., Bøgwald, K.-P., Sørbye, Ø., Sjaastad, M. C., & Heyerdahl, O. (2006). Analysis of the patient-therapist relationship in dynamic psychotherapy: An experimental study of transference interpretations. *The American Journal of Psychiatry, 163*, 1739-1746. Hentet fra: <http://ajp.psychiatryonline.org/data/Journals/AJP/3778/06aj1739.pdf>

- Høglend, P., Bøgwald, K. P., Amlo, S., Heyerdahl, O., Sørbye, Ø., Marble, A., ... Bentsen, H. (2000). Assessment of change in dynamic psychotherapy. *Journal of Psychotherapy Practice and Research*, 9, 190-199. Hentet fra: <http://vuir.vu.edu.au/19368/31/00jpr190.pdf>
- James, W. (1890). *Principles of psychology*. London: Macmillan.
- Johansson, P., Høglend, P., Ulberg, R., Amlo, S., Marble, A., Bøgwald, K.-P., ... Heyerdahl, O. (2010). The mediating role of insight for long-term improvements in psychodynamic therapy. *Journal of Consulting and Clinical Psychology*, 78, 438-448. doi:10.1037/a0019245
- Kane, M. T. (1996). The precision of measurements. *Applied measurement in education*, 9, 355-379. doi: 10.1207/s15324818ame0904_4
- Kazdin, A. E. (1999). Overview of research design issues in clinical psychology. I P. C. Kendall, J. N. Butcher, & G. N. Holmbeck (Red.), *Handbook of research methods in clinical psychology* (s. 3-31). New York: John Wiley & Sons, Inc.
- Karterud, S., Friis, S., Irion, T., Mehlum, L., Vaglum, P., & Vaglum, S. (1995). A SCL-90-R derived index of the severity of personality disorders. *Journal of Personality Disorders*, 9, 112-123. doi: 10.1521/pedi.1995.9.2.112
- Kernberg, O. F. (1984). *Severe personality disorders: Psychotherapeutic strategies*. New Haven, CT: Yale University Press.
- Kleven, T. A. (2002). Begrepsoperasjonalisering. I T. Lund (Red.), *Innføring i forskningsmetodologi* (s. 141-183). Oslo: Unipub.
- LaForge, R., & Suczek, R. F. (1995). The interpersonal dimension of personality: III. An interpersonal checklist. *Journal of Personality*, 24, 94-112. doi: 10.1111/j.1467-6494.1955.tb01177.x
- Leary, T. (1957). *Interpersonal diagnosis of personality: A functional theory and methodology for personality evaluation*. New York: The Ronald Press Company.
- Loevinger, J. (1957). Objective tests as instruments of psychological theory. *Psychological Reports*, 3, 635-694. doi: 10.2466/pr0.1957.3.3.635

- Lorr, M., & Strack, S. (1999). A Study of Benjamin's eight-facet Structural Analysis of Social Behavior (SASB) model. *Journal of Clinical Psychology, 55*, 207-215. Hentet fra: <http://web.comhem.se/u68426711/14/Lorr99.pdf>
- Luborsky, L., Singer B., & Luborsky, L. (1975). Comparative studies of psychotherapies: Is it true that "Everyone Has Won and All Must Have Prizes"?. *Archives of General Psychiatry, 32*, 995-1008. doi: 10.1001/archpsyc.1975.01760260059004
- Malan, D. H. (1995). *Individual psychotherapy and the science of psychodynamics*. (2. utg.) London: Butterworth-Heinemann.
- Malmquist, C. P. (1991). Depression and extreme violence. I D. Cicchetti & W. M. Grove (Red.), *Thinking clearly about psychology: Essays in honor of Paul E. Meehl*. Vol. 2 *Personality and psychopathology* (s. 378-409). Minneapolis: University of Minnesota Press.
- Margo, G. M., Dewan, M. J., Fisher, S., & Greenberg, R. P. (1992). Comparison of three depression rating scales. *Perceptual and Motor Skills, 75*, 144-146. doi: 10.2466/pms.1992.75.1.144
- McWilliams, N. (2004). *Psychoanalytic psychotherapy: A practitioner's guide*. New York: Guilford press.
- Mitchell, S. A., & Black, M. J. (1995). *Freud and beyond: A history of modern psychoanalytic thought*. New York: Basic Books.
- Monsen, J. T., Hagtvet, K. A., Havik, O. E., & Eilertsen, D. E. (2006). Circumplex structure and personality disorder correlates of the interpersonal problems model (IIP-C): Construct validity and clinical implications. *Psychological Assessment, 18*, 165-173. doi: 10.1037/1040-3590.18.2.165
- Monsen, J. T., von der Lippe, A. L., Havik, O. E., Halvorsen, M. S., & Eilertsen, D. E. (2007). Validation of the SASB Introject Surface in a Norwegian clinical and nonclinical sample. *Journal of Personality Assessment, 88*, 235-245. doi: 10.1080/00223890701268108

- Newton-Howes, G., Tyrer, P., & Johnson, T. (2006). Personality disorder and the outcome of depression: Meta-analysis of published studies. *The British Journal of Psychiatry*, *188*, 13-20. doi: 10.1192/bjp.188.1.13
- Nielsen, G., & Vassend, O. (1994). SCL-90-R, Norsk oversettelse.
- Norcross, J. C., & Lambert, M. J. (2011). Evidence-based therapy relationships. I J. C. Norcross (Red.), *Psychotherapy relationships that work: Evidence-based responsiveness*. (2. utg., s. 3-21). New York: Oxford University Press.
- Ogles, B. M., Lambert, M. J., Weight, D. G., & Payne, I. R. (1990). Agoraphobia outcome measurements in the 1980's: A review and meta-analysis. *Psychological Assessment*, *2*, 317-325. doi: 10.1037/1040-3590.2.3.317
- OPD Task Force (2001). *Operationalized Psychodynamic Diagnostics: Foundations and manual*. Bern: Hogrefe & Huber Publishers.
- Orlinsky, D. E, Rønnestad, H. M., & Willutzki U. (2004). Fifty years of psychotherapy process-outcome research: Continuity and change. I M. J. Lambert (Red.), *Bergin and Garfield's Handbook of psychotherapy and behavior change*. (5. utg., s. 84-135). New York: John Wiley & Sons, Inc.
- Paap, M., Meijer, R. R., Van Bebber, J., Pedersen, G., Karterud, S., Hellem, F. M., & Haraldsen, I. R. (2011). A study of the dimensionality and measurement precision of the SCL-90-R using item response theory. *International Journal of Methods in Psychiatric Research*, *20*, e39-e55. doi: 10.1002/mpr.347
- Pedhazur, E. J. & Schmelkin, L. P. (1991). *Measurement, design and analysis. An integrated approach*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Pedersen, G. A. (2002). Norsk revidert versjon av Inventory of Interpersonal Problems – Circumplex (IIP-C). *Tidsskrift for Norsk Psykologforening*, *39*, 25-34.
- Pedersen, G., & Karterud, S. (2004). Is SCL-90R helpful for the clinician in assessing DSM-IV symptom disorders? *Acta Psychiatrica Scandinavica*, *110*, 215–224. doi: 10.1111/j.1600-0447.2004.00321.x

- Perry, J. C. (1992). Problems and considerations in the valid assessment of personality disorders. *American Journal of Psychiatry*, *149*, 1645-53. Hentet fra: <http://search.proquest.com/docview/220467357?accountid=14699>
- Pfohl, B., Coryell, W., Zimmerman, M., & Stangl, D. (1986). DSM-III personality disorders: Diagnostic overlap and internal consistency of individual DSM-III criteria. *Comprehensive Psychiatry*, *27*, 21-34. doi: 10.1016/0010-440X(86)90066-0
- Pincus, A. L., & Gurtman, M. B. (1995). The three faces of interpersonal dependency: Structural analyses of self-report dependency measures. *Journal of Personality and Social Psychology*, *69*, 744-758. doi: 10.1037/0022-3514.69.4.744
- Pincus, A. L., & Wiggins, J. S. (1990). Interpersonal problems and conceptions of personality disorders. *Journal of Personality Disorders*, *4*, 342-352. doi: 10.1521/pedi.1990.4.4.342
- Rogosa, D. R., & Willett, J. B. (1983). Demonstrating the reliability of the difference score in the measurement of change. *Journal of Educational Measurement*, *20*, 335-343. doi: 10.1111/j.1745-3984.1983.tb00211.x
- Ruiz, M. A., Pincus, A. L., & Bedics, J. B. (1999). Using the Structural Analysis of Social Behavior (SASB) to differentiate young adults with borderline personality disorder features. *Journal of Personality Disorders*, *13*, 187-198. doi: 10.1521/pedi.1999.13.2.187
- Sandler, J., Dare, C., & Holder, A. (1973). *The patient and the analyst: The basis of the psychoanalytic process*. London: Allen & Unwin.
- Schmitz, N., Kruse, J., Heckrath, C., Alberti, L., & Tress, W. (1999). Diagnosing mental disorders in primary care: the General Health Questionnaire (GHQ) and the Symptom Check List (SCL-90-R) as screening instruments. *Social Psychiatry Psychiatric Epidemiology*, *34*, 360-366. doi: 10.1007/s001270050156
- Shavelson, R. J., & Webb, N. M. (1991). *Generalizability theory: A primer*. Newbury Park, CA: Sage.

- Shedler, J., Mayman, M., & Manis, M. (1993). The illusion of mental health. *American Psychologist*, 48, 1117-1131. doi: 10.1037/0003-066X.48.11.1117
- Shrout, P. E., & Fleiss, J. L. (1979). Intraclass Correlations: Uses in assessing rater reliability. *Psychological Bulletin*, 86, 420-428. doi: 10.1037/0033-2909.86.2.420
- Sifneos, P. E. (1979). *Short-term dynamic psychotherapy: Evaluation and technique*. New York, NY: Plenum Medical Book Company.
- Sifneos, P. E., Apfel, R. J., Bassuk, E., Fishman, G., & Gill, A. (1980). Ongoing outcome research on short-term dynamic psychotherapy. *Psychotherapy and Psychosomatics*, 33, 233-241. doi: 10.1159/000287436
- Skodol, A. E., Gunderson, J. G., Shea, M. T., McGlashan, T. H., Morey, L. C., Sanislow, C., ... Stout, R. L. (2005). The Collaborative Longitudinal Personality Disorders Study (CLPS): Overview and implications. *Journal of Personality Disorders*, 19, 487-504. doi: 10.1521/pedi.2005.19.5.487
- Spearman, C. (1904). The proof and measurement of association between two things. *American Journal of Psychology*, 15, 72-101. doi: 10.2307/1412159
- Stiles, T. C., & Høglend, P. (1994). IIP-64 – versjon C. Norsk oversettelse.
- Sullivan, H. S. (1953). *The interpersonal theory of psychiatry*. New York: Norton.
- Sullivan, H. S. (1954). *The psychiatric interview*. New York: Norton.
- Szapocznik, J., Rio, A. T., Murray, E., & Richardson, R. (1993). Assessing change in child psychodynamic functioning in treatment outcome studies: The Psychodynamic Child Ratings. *Revista Interamericana de Psicología*, 27, 147-162.
- Thomä, H., & Kächele, H. (1994). *Psychoanalytic practice*. Vol. 1. *Principles*. Northvale NJ: Jason Aronson.
- Thomä, H., & Kächele, H. (1994). *Psychoanalytic practice*. Vol. 2. *Clinical Studies*. Northvale NJ: Jason Aronson.

- Tingey, R. C., Lambert, M. J., Burlingame, G. M., & Hansen, N. B. (1996). Assessing clinical significance: Proposed extensions to method. *Psychotherapy Research, 6*, 109-123. doi: 10.1080/10503309612331331638
- Tracey, T. J. G. (2005). Interpersonal rigidity and complementarity. *Journal of Research in Personality, 39*, 592-614. doi: 10.1016/j.jrp.2004.12.001
- Tracey, T. J. G., Rounds, J., & Gurtman, M. B. (1996). Examination of the general factor with the interpersonal circumplex structure: Application to the Inventory of Interpersonal Problems. *Multivariate Behavioral Research, 31*, 441-466. doi: 10.1207/s15327906mbr3104_3
- Varvin, S. (2002). *Mental survival strategies after extreme traumatisation*. Oslo: Unipub Faculty of Medicine, University of Oslo.
- Vassend, O. (1989). Dimensions of negative affectivity, self-reported somatic symptoms, and health-related behaviors. *Social Science and Medicine, 28*, 29-36. doi: 10.1016/0277-9536(89)90303-1
- Vassend, O., Lian, L., & Andersen, H.T. (1992). Norske versjoner av NEO Personality Inventory, Symptom Checklist 90 Revised og Giessen Subjective Complaints List. Del I. *Tidsskrift for Norsk Psykologforening, 29*, 1150-1160.
- Vassend, O., & Skrondal, A. (1999). The problem of structural indeterminacy in multidimensional symptom report instruments. The case of SCL-90-R. *Behaviour Research and Therapy, 37*, 685-701. doi: 10.1016/S0005-7967(98)00182-X
- Wallerstein, R. S. (n.d). *Scales of Psychological Capacities*. Unpublished manuscript. The psychotherapy research project II. San Francisco. Department of Psychiatry, University of California.
- Wampold, B. E. (2001). *The great psychotherapy debate: Models, methods, and findings*. Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.
- Ward, R. K. (2004). Assessment and management of personality disorders. *American Family Physician, 70*, 1505–1512. Hentet fra <http://www.aafp.org/afp/2004/1015/p1505.pdf>

- Weinryb, R. M., & Rössel, R. J. (1991). Karolinska Psychodynamic Profile (KAPP). *Acta Psychiatrica Scandinavica*, 83(Suppl. 363), 1–23. doi: 10.1111/j.1600-0447.1991.tb10573.x
- Weinryb R. M., Rössel R. J., & Asberg M. (1991). The Karolinska Psychodynamic Profile. I. Validity and dimensionality. *Acta Psychiatrica Scandinavica*, 83, 64-72. doi: 10.1111/j.1600-0447.1991.tb05513.x
- Weiss, I., Rabinowitz, J., & Spiro, S. (1996). Agreement between therapists and clients in evaluating therapy and its outcomes: literature review. *Administration and Policy in Mental Health and Mental Health Services Research*, 23, 493-511. doi: 10.1007/BF02108686
- Westen, D. (1997). Divergences between clinical and research methods for assessing personality disorders: implications for research and the evolution of axis II. *American Journal of Psychiatry*, 154, 895-903.
Hentet fra <http://ajp.psychiatryonline.org/article.aspx?articleID=172395>
- Widiger, T. A. (1992). Categorical versus dimensional classification: Implications from and for research. *Journal of Personality Disorders*, 6, 287-300. doi: 10.1521/pedi.1992.6.4.287
- Wilczek, A., Barber, J. P., Gustavsson, J. P., Asberg, M., & Weinryb, R. M. (2004). Change After Long-Term Psychoanalytic Psychotherapy. *Journal of the American Psychoanalytic Association*, 52, 1163-1184. Hentet fra <https://www.apsa.org/portals/1/docs/japa/524/wilczek-p.1163-1184.pdf>
- World Health Organization (1992). *The ICD-10 classification of mental and behavioural disorders, clinical description and diagnostic guidelines*. Geneva: World Health Organization. Norsk utgave: (1999). *ICD-10. Psykiske lidelser og atferdsforstyrrelser. Kliniske beskrivelser og diagnostiske retningslinjer*. Oslo: Gyldendal Akademisk.
- Zanarini, M. C., Skodol, A. E., Bender, D., Dolan, R., Sanislow, C., Schaefer, E., ... Gunderson, J. G. (2000). The Collaborative Longitudinal Personality Disorders Study: Reliability of Axis I and II diagnoses. *Journal of Personality Disorders*, 14, 291-299. doi: 10.1521/pedi.2000.14.4.291