

Nivå av personlighetsfungering i Den
alternative modellen for
personlighetsforstyrrelser: vurdering ved
SCID-AMPD modul 1

Marianne Arnesen og Karoline Koritzinsky



Hovedoppgave ved Psykologisk Institutt

UNIVERSITETET I OSLO

18.10.2016

© Marianne Arnesen og Karoline Koritzinsky

2016

Nivå av personlighetsfungering i Den alternative modellen for personlighetsforstyrrelser:
vurdering ved SCID-AMPD modul 1.

Marianne Arnesen og Karoline Koritzinsky

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

Trykk: Reprosentralen, Universitetet i Oslo

Sammendrag

Forfattere: Marianne Arnesen og Karoline Koritzinsky

Tittel: Nivå av personlighetsfungering i Den alternative modellen for personlighetsforstyrrelser: vurdering ved SCID-AMPD modul 1.

Veileder: Svenn Torgersen

Bakgrunn: Det har vært en debatt om personlighetsforstyrrelser skal kategoriseres kategorisk eller dimensjonalt (Frances, 1982). I dagens diagnosesystem benyttes det en kategorisk tilnærming. Forskning har vist en rekke svakheter med den kategoriske modellen, og det er foreslått at en dimensjonal modell vil beskrive personlighetspatologi på en bedre måte (Widiger & Samuel, 2005). I utviklingen av DSM-5 forsøkte man å utvikle en ny dimensjonal modell for personlighetsforstyrrelser. Resultatet ble en hybridmodell, den alternative modellen for personlighetsforstyrrelser (AMPD), som består av en dimensjonal vurdering av alvorlighetsgrad, en dimensjonal vurdering av personlighetstrekk, samt seks spesifikke personlighetsforstyrrelseskategorier. Kritikken mot modellen gjorde imidlertid at den ikke ble vedtatt, men ble inkludert i seksjon III sammen med konsepter som behøver videre forskning (APA, 2013). Studier har vist at alvorlighetsgrad er den viktigste prediktoren for nåværende og fremtidig fungering (Hopwood et al., 2011, Tyrer, 2005). Vurderingen av alvorlighetsgrad i modellen blir kalt nivå av personlighetsfungering (LPFS). Noen studier har undersøkt validitet og klinisk nytte for denne delen av modellen. Få studier har foreløpig sett på reliabiliteten. De siste årene har også flere utviklet måleinstrumenter rettet spesifikt mot å vurdere nivå av personlighetsfungering (LPFS). Denne studien undersøker interrater reliabiliteten til et av disse måleinstrumentene, et intervju kalt SCID-AMPD modul 1 (Bender et al., 2014).

Metode: Denne studien er en delstudie under DSM-5 prosjektet ledet av overlege Benjamin Hummelen ved Seksjon for personlighetspsykiatri, Ullevål. Studien undersøkte interrater reliabiliteten til måleinstrumentet SCID-AMPD modul 1 ved et test-retest design. SCID-AMPD modul 1 ble gjennomført to ganger av to ulike intervjuere på samme deltager med omtrent en ukes mellomrom. Deltagerne ble rekruttert fra fire gruppenheter i Dagbehandlingsnettverket, samt fra en allmennpoliklinikk. Enigheten i skåringer mellom

intervjuerne ble undersøkt ved korrelasjonsanalyser; Pearson produkt-moment koeffisienten og two way random intraclass korrelasjoner.

Resultat: Resultater fra denne studien viste jevnt over høy interrater reliabilitet for SCID-AMPD modul 1, med noen lave reliabilitetsestimater for deler av intervjuet.

Konklusjon: Resultater fra denne studien gir støtte for reliabiliteten til måleinstrumentet SCID-AMPD modul 1. Nivå av personlighetsfungering (LPFS) kan i hovedsak vurderes med høy grad av enighet med dette intervjuet. Imidlertid viser også resultatene at deler av intervjuet bør videreutvikles for å sikre reliable vurderinger. I forskningsøyemed bidrar denne studien med støtte for måleinstrumentet SCID-AMPD modul 1 og den dimensjonale vurderingen av alvorlighetsgrad i den alternative modellen for personlighetsforstyrrelser. Med utgangspunkt i vår studie kan det være nyttig for videre forskning å undersøke validiteten til intervjuet, samt klinisk nytteverdi.

Forord

Arbeidet med denne hovedoppgaven startet i fjor høst, da Karoline var i praksis ved Personlighetspoliklinikken. Her ble hun introdusert for forskningsprosjektet om Den alternative modellen for personlighetsforstyrrelser. Vi hadde snakket om å skrive oppgave sammen, og ble veldig glade for tilbudet om å skrive oppgave for akkurat dette prosjektet. Det var ekstra spennende at den alternative modellen for personlighetsforstyrrelser er en helt ny måte å tenke på omkring personlighetsproblematikk. Vi er utrolig takknemlige for å ha fått være med på dette prosjektet fra start til slutt; med utforming av design, koordinering av intervjuene, presentasjon av prosjektet på ulike enheter, gjennomføring av intervjuer, seminarer med DSM-5 forskningsgruppen og dataanalyser. Det har til tider vært mye arbeid, men det har absolutt vært verdt det!

Vi vil først takke Benjamin Hummelen, som med glede og iver har inkludert oss i stor grad i prosjektet. Vi vil også takke veilederen vår Svenn Torgersen for gode innspill til prosjektet og oppgaven, samt artige anekdoter på veiledning. Takk til resten av DSM-5 forskningsgruppen for å ha tatt godt imot oss. Takk til de andre intervjuerne og enhetene som har rekruttert deltagerne. En stor takk til deltagerne våre som har stilt opp ved to anledninger, og delt av sin personlighet og historie. Takk til kjæreste og ektemann for god støtte i stressende perioder.

Vi er utrolig stolte og takknemlige over å ha fått være med på noe som kan bidra til endring av diagnostisering og forståelse av personlighetspatologi. Det blir utrolig spennende å følge utviklingen i fagfeltet!

Når vi ser tilbake på denne prosessen, husker vi spesielt godt bilturene våre sammen til DPSer i Akershus. De er farget av øvelseskjøring, kartlesingsproblemer, latter og gode samtaler. Det hadde ikke vært det samme uten hverandre.

Med hilsen

Marianne Arnesen og Karoline Koritzinsky

Innholdsfortegnelse

1	Innledning.....	1
1.1	Definisjon av personlighetsforstyrrelser i DSM-V.....	1
1.2	Nåværende klassifisering av personlighetsforstyrrelser	2
1.3	Epidemiologi	3
1.4	Utviklingen til personlighetsforstyrrelser i DSM	3
1.5	Dimensjonal versus kategorisk tilnærming.....	6
1.5.1	Kategorisk tilnærming	6
1.5.2	Dimensjonal tilnærming	8
1.5.3	Hybridmodeller.....	9
1.6	Utviklingen av DSM-V.....	9
2	Den alternative modellen for personlighetsforstyrrelser.....	12
2.1	Oversikt	12
2.2	Generelle kriterier for personlighetsforstyrrelser	12
2.3	Kriterium A: Nivå av personlighetsfungering.....	12
2.4	Kriterium B: Patologiske personlighetstrekk	14
2.5	Personlighetsforstyrrelser.....	15
2.6	Personlighetsforstyrrelsesdiagnose.....	16
2.7	Rasjonale for nivå av personlighetsfungering.....	17
2.7.1	Alvorlighetsgrad	17
2.7.2	Selv og interpersonlig fungering	17
2.7.3	Utvikling av skalaen	18
2.8	ICD 11 – forslaget.....	19
3	Foreløpige publikasjoner på Den alternative modellen for personlighetsforstyrrelser	21
3.1	Klinisk nytte	21
3.2	Validitet	22
3.3	Reliabilitet	23
3.4	Måleinstrumenter for LPFS.....	24
4	Målsetningen med studiet.....	26
5	Metode.....	29
5.1	Formål og design.....	29
5.2	Deltagere	29
5.2.1	Utvalg.....	29
5.2.2	Demografi.....	30
5.3	Prosedyrer	30
5.3.1	Rekruttering.....	30
5.3.2	Datainnsamling	31
5.4	Intervjuere.....	32
5.4.1	Opplæring.....	32
5.5	Måleinstrument.....	32
5.5.1	Strukturert klinisk intervju for den alternative (seksjon III) modellen for personlighetsforstyrrelser i DSM-5 (SCID-AMPD/SCID-III)	33
5.6	Statistiske analyser.....	36
5.6.1	Preliminære analyser og beskrivende statistikk.....	36

5.6.2	Intern konsistens.....	36
5.6.3	Endring i skårer: t-test.....	37
5.6.4	Interrater reliabilitet	37
5.6.5	Tilleggsanalyser	38
6	Resultater.....	39
6.1	Oppsummering resultater	39
6.2	Preliminære dataanalyser.....	39
6.2.1	Manglende data.....	39
6.2.2	Normalitet, linearitet og homoscedasticity	39
6.2.3	Ekstremskårer	40
6.3	Beskrivende statistikk for utvalget.....	40
6.4	Intern konsistens.....	40
6.4.1	Cronbachs alpha	40
6.5	Endring i skårer mellom testene.....	42
6.5.1	T-tester	42
6.6	Enighet i skåringer mellom intervju 1 og intervju 2	43
6.6.1	Pearson r	43
6.6.2	Intraclass korrelasjoner	44
6.7	Tilleggsanalyse	46
6.7.1	Grad av erfaring	46
7	Diskusjon.....	47
7.1	Endring i skåringer mellom intervjuene	47
7.2	Intervjuets interne konsistens.....	48
7.3	Enighet mellom intervjuerne	48
7.3.1	Hva kan ha ført til god reliabilitet?	52
7.4	Styrker og begrensninger	53
7.4.1	Styrker ved studien	54
7.4.2	Begrensninger ved studien.....	55
7.5	Implikasjoner	57
7.6	Videre forskning	58
	Litteraturliste	60
	Vedlegg	72
	Vedlegg A: Nivå av personlighetsfungeringsskalaen.....	72
	Vedlegg B: Forekomst av symptomidelser i utvalget vårt.....	75
	Vedlegg C: Forekomst av personlighetsforstyrrelser i utvalget vårt.....	76
	Vedlegg D: Oversikt over rotasjon mellom intervjuerne.....	77
	Vedlegg F: Oppbygning av SCID-AMPD modul 1.....	78

1 Innledning

I 2013 publiserte American Psychiatric Association (APA) den femte utgaven av Diagnostic and Statistical Manual of Mental Disorders (APA, 2013). Denne var et produkt av mer enn 13 års arbeid. En omfattende revisjon av DSM hadde da ikke funnet sted siden 1994 (APA, 1994). I forkant av publikasjonen var det forventet store endringer i klassifiseringssystemet, spesielt med tanke på diagnostiseringen av personlighetsforstyrrelser. Kupfer et al. (2002) oppfordret til et paradigme skifte, og så for seg at personlighetsforstyrrelser var feltet hvor en kunne teste ut dimensjonale diagnoser. Da DSM-V var et faktum, var det imidlertid få endringer som var gjort fra den fjerde utgaven. Modellen for personlighetsforstyrrelser var tilnærmet lik den i DSM-IV. Dette på tross av at personlighet- og personlighetsforstyrrelsesarbeidsgruppen hadde jobbet iherdig med en ny modell. En modell som på oppløpssiden ble forkastet, og plassert i seksjon III av manualen for modeller som behøver videre forskning. Den alternative modellen i DSM-5 er en hybrid modell, bestående av både dimensjonale og kategoriske elementer. Mye av grunnen til at den ble forkastet var at det var gjort for lite forskning på den (Zachar, Krueger & Kendler, 2016). Dette vil vi bidra med. Studien vår ser på interrater reliabiliteten til måleinstrumentet SCID-AMPD Modul 1 (Bender, Skodol, First & Oldham, 2014). SCID-AMPD Modul 1 er basert på en del av den alternative modellen: en dimensjonal vurdering av personlighetsfungering kalt Nivå av personlighetsfungeringsskalaen.

Denne innledningen skal først ta kort for seg hva personlighetsforstyrrelser er, klassifisering og epidemiologi. Deretter går den gjennom utviklingen til klassifiseringen av personlighetsforstyrrelser i DSM. Så utviklingen av en ny personlighetsforstyrrelsesmodell i DSM-5 og den alternative modellen. Innledningen avslutter med foreløpige publikasjoner på Nivå av personlighetsfungeringsskalaen, og hensikten med studiet vårt.

1.1 Definisjon av personlighetsforstyrrelser i DSM-V

I DSM-5 blir en personlighetsforstyrrelse definert som "et vedvarende mønster av indre erfaring og atferd som markert avviker fra forventningene i et individs kultur..." (APA, 2013). For å oppfylle de generelle diagnostiske kriteriene må mønsteret komme til uttrykk i to eller flere av de følgende områdene: kognisjoner, affekt,

interpersonlig fungering og / eller impuls kontroll. Videre kriterier er at mønsteret er rigid og vises i en rekke sosiale og personlige situasjoner; fører til klinisk signifikant ubehag eller nedsatt sosial eller yrkesmessig fungering; er stabilt og av lang varighet; med debut i ungdom eller tidlig voksenalder. Mønsteret skal ikke være bedre forklart av en annen psykisk lidelse, substansbruk eller en annen medisinsk tilstand (APA, 2013).

1.2 Nåværende klassifisering av personlighetsforstyrrelser

Dagens diagnosesystemer, ICD-10 og DSM-5, benytter seg begge av en kategorisk klassifikasjon av personlighetsforstyrrelser. En kategorisk klassifikasjon deler personlighetsforstyrrelsene inn i klart atskilte lidelser (Tyrer, Reed, Crawford, 2015). Modellen på personlighetsforstyrrelser i hoveddelen av DSM-5 manualen, beskriver ti kategorier av personlighetsforstyrrelser. Disse er organisert i tre klustere basert på deskriptive likheter. Kluster A består av paranoid, schizoid og schizotyp personlighetsforstyrrelse. Kluster B består av antisosial, borderline, histrionisk og narsissistisk personlighetsforstyrrelse. Kluster C består av unnvikende, avhengig og tvangspregget personlighetsforstyrrelse.

Individer diagnostisert med en kluster A personlighetsforstyrrelse kan ofte oppfattes som rare eller eksentriske. Disse lidelsene er fenomenologisk og etiologisk assosiert med psykotiske lidelser (Maier, Lichtermann, Minges & Heun, 1994). Det som skiller dem fra de psykotiske lidelsene er mangelen på vedvarende hallusinasjoner eller vrangforestillinger (Hopwood & Thomas, 2012). Kluster B personlighetsforstyrrelser er gruppert sammen fordi de kan fremstå som dramatiske, emosjonelle eller oppfarende. Individer med disse lidelsene kan oppføre seg impulsivt og ha problemer med emosjonsregulering. Mens personer med en kluster C diagnose kan fremstå som redde eller engstelige. I tillegg finnes det en restkategori kalt "annen spesifisert personlighetsforstyrrelse" eller "uspesifisert personlighetsforstyrrelse." Denne skal for eksempel brukes hvis individer møter de generelle kriteriene for en personlighetssforstyrrelse, men har trekk fra flere enn en kategori (APA, 2013).

ICD-10 har identifisert åtte personlighetsforstyrrelseskategorier (WHO, 1993). Disse kategoriene er svært like kategoriene i DSM-5. Hovedforskjellene er at narsissistisk personlighetsforstyrrelse ikke er en del av klassifikasjonen, at schizotyp

personlighetsforstyrrelse blir regnet som en del av schizofreni spekteret, og at borderline personlighetsforstyrrelse blir regnet som en underkategori av emosjonelt ustabil personlighetsforstyrrelse (Tyrer, Crawford & Reed, 2015). I tillegg benytter de to klassifiseringssystemene seg av noe ulike navn på samme kategori. Tvangspreget personlighetsforstyrrelse kalles anankastisk, og antisosial kalles dysosial i ICD-10 (WHO, 1993).

For å sette en personlighetsforstyrrelsesdiagnose benytter man gjerne et strukturert eller semistrukturert intervju i kombinasjon med annen klinisk informasjon. Strukturerte og semistrukturerte intervju har vist seg å være mer reliable enn ustrukturerte intervjuer (Hopwood & Thomas, 2012). Det mest brukte intervjuet i klinisk praksis i dag er: Structured Clinical Interview for DSM-IV TR Axis II Personality Disorders, kallt SCID-II (First et al., 1997). Intervjuet er basert på DSM kriteriene for personlighetsforstyrrelsene.

1.3 Epidemiologi

Prevalens estimater indikerer at mer enn 10% av individer lider av en personlighetsforstyrrelse i løpet av livet. I den generelle populasjonen er personlighetsforstyrrelser tilnærmet like vanlige blant kvinner som blant menn (Coid et al., 2006; Torgersen et al., 2001). Estimatenes for spesifikke personlighetsforstyrrelser er mellom 1% og 2%. Personlighetsforstyrrelser er svært vanlig i kliniske settinger. Forskning tyder på at nesten halvparten av polikliniske pasienter møter kriteriene for en personlighetsforstyrrelse, og mer enn halvparten av pasienter som er innlagt møter kriteriene for en personlighetsforstyrrelse. Dette gjør personlighetsforstyrrelser til en av de vanligste psykiske lidelsene i psykisk helsevern. Avhengig, borderline, tvangspreget, unnvikende og schizotyp personlighetsforstyrrelse er overrepresentert i pasient populasjoner. Mens individer med antisosial, schizoid og paranoid personlighetsforstyrrelse er underrepresentert (Torgersen, 2009).

1.4 Utviklingen til personlighetsforstyrrelser i DSM

Siden den første utgaven av DSM i 1952 har personlighetsforstyrrelser vært inkludert i diagnosemanualen. Antall personlighetsforstyrrelser og ulike typer har imidlertid vært i stadig endring. Noen personlighetsforstyrrelsestyper har blitt lagt til, andre har blitt tatt

bort. Av typene er det kun paranoid, tvangspreget og antisosial personlighetsforstyrrelse som har vært faste DSM medlemmer. Det teoretiske grunnlaget har også forandret seg mye i løpet av de ulike versjonene (Skodol, Bender, Gunderson & Oldham, 2014).

DSM-I (APA, 1952, ref. i Oldham, 2009) definerte personlighetsforstyrrelser som mer eller mindre permanente mønstre av atferd og menneskelig interaksjon etablert i tidlig voksenalder. Personlighetsforstyrrelsene ble regnet som motsetningen til symptom-nevrosene. Tanken var at nevrosene var preget av angst og ubehag, mens personlighetsforstyrrelsene var ego-syntone og ble derfor ikke gjenkjent av individene som hadde diagnosen. Hovedsakelig ble personlighetsforstyrrelser sett på som mangelpatologi. I DSM-I ble personlighetsforstyrrelser delt inn i tre grupper: personlighetsmønstre forstyrrelser, personlighetstrekk forstyrrelser, og sosiopatiske personlighetsforstyrrelser. Personlighetsmønstre forstyrrelser ble regnet som de mest fastlåste lidelsene som selv med behandling var vanskelige å endre. Denne gruppen inkluderte inadekvat personlighet, schizoid personlighet, cyclo-tym personlighet og paranoid personlighet. Personlighetstrekk forstyrrelsene var tenkt å være mindre gjennomgripende og hemmende. Disse pasientene kunne fungere relativt godt i fravær av stress. I svært stressende situasjoner ville derimot pasienter med emosjonelt ustabile, passiv-aggressive og tvangspregete personligheter oppleve mye emosjonelt ubehag og nedsatt funksjon. Individuer med en av disse forstyrrelsene ble regnet som varierende motivert og mottagelige for behandling. Sosiopatisk personlighetsforstyrrelse inkluderte atferd som ble sett på som sosialt avvikende på den tiden. Dette var antisosial- og dysosial personlighetsforstyrrelse, seksuelle avvik og avhengighet (Oldham, 2009).

DSM-II (APA, 1968, ref. i Oldham, 2009) ble utviklet på bakgrunn av publiseringen av ICD-8 og APAs ønske om å ha liknende terminologi som det internasjonale diagnosesystemet. Det ble forsøkt å bevege seg bort fra teoribaserte personlighetsdiagnoser, og over til mer observerbare og målbare deler av personligheten. DSM-II (APA, 1968, ref. i Oldham, 2009) la vekt på at personlighetsforstyrrelser også involverer emosjonelt ubehag og nedsatt fungering, ikke bare sosialt upassende atferd. De tre DSM-I gruppene: personlighetsmønstre, personlighetstrekk og sosiopatisk personlighetsforstyrrelse, ble forkastet. Asthenic

personlighetsforstyrrelse ble lagt til, for så å bli fjernet i neste versjon av DSM (Oldham, 2009).

I DSM-III (APA, 1980, ref. i Oldham 2009) ble det gjort store forandringer både i konseptualiseringen og klassifiseringen av personlighetsforstyrrelser. Tidligere hadde DSM vært basert på psykoanalytisk teori. Nå ble det foretatt et skifte til en ikke-teoretisk deskriptiv tilnærming (Skodol, Bender, Gunderson & Oldham, 2014). I arbeidet med DSM-III ble det lagt enda større vekt på å øke reliabiliteten til diagnoser. Personlighetsforstyrrelser ble dermed plassert på en egen akse, akse II, atskilt fra symptomlidelsene. Formålet med dette var å fremheve viktigheten av personlighetsforstyrrelsene, samt å vise at personlighetsforstyrrelser kunne forekomme samtidig som andre psykiske lidelser. Det ble også definert spesifikke diagnostiske kriterier for hver personlighetsforstyrrelseskategori (Frances, 1980). Tidligere versjoner av DSM hadde bare hatt en kort, narrativ beskrivelse av personlighetsforstyrrelsene (Verheul & Widiger, 2004), så spesifikke kriterier var en klar forbedring av reliabilitet.

Det skjedde også endringer i antall og type kategorier. Et par kategorier ble fullstendig fjernet, andre omgjort til akse I lidelser. Schizoid personlighetsforstyrrelse ble vurdert som en for bred kategori og omgjort til tre kategorier: Schizoid, schizotyp og unnvikende personlighetsforstyrrelse. To nye diagnoser ble lagt til i DSM-III. Dette var borderline personlighetsforstyrrelse og narsissistisk personlighetsforstyrrelse (Oldham, 2009). De nå 11 personlighetsforstyrrelseskategoriene ble gruppert i tre klustere basert på deskriptive likheter. Kluster A besto av paranoid, schizoid og schizotyp personlighetsforstyrrelse karakterisert av rar eller eksentrisk atferd. Kluster B besto av histrionisk, narsissistisk, antisosial og borderline personlighetsforstyrrelse karakterisert av emosjonell eller uberegnelig atferd. Kluster C besto av unnvikende, avhengig, tvangspregget og passiv-aggressiv personlighetsforstyrrelse karakterisert av redd eller engstelig atferd. Det er viktig å merke seg at grupperingen skjedde på bakgrunn av klinisk intuisjon, ikke med bakgrunn i empiri (Frances, 1980). Det ble også innført en restkategori kalt personality disorder not otherwise specified (PD-NOS). Denne var for personer med åpenbar personlighetsproblematikk, men som ikke oppfylte kriteriene til en av de spesifikke kategoriene (Verheul & Widiger, 2004).

Det ble kun gjort små endringer på modellen i DSM-III-R (APA, 1987, ref. i Oldham, 2009). Kategoriene forble de samme, men det ble gjort noe endringer på de

diagnostiske kriteriene. Kriteriene ble nå gjort polytetiske for alle personlighetsforstyrrelseskategoriene, i motsetning til å være monotetiske for noen av kategoriene og polytetiske for andre. Det vil si at for å få en diagnose behøvde et individ nå kun å oppfylle en del av diagnose kriteriene, ikke alle (Oldham, 2009).

Personlighetsforstyrrelsesmodellen som ble lansert i DSM-III stimulerte til mye forskning. Funn fra en voksende empirisk litteratur ble benyttet til å utvikle DSM-IV. DSM-IV introduserte et sett med generelle diagnostiske kriterier for alle personlighetsforstyrrelsene. Kriteriene definerer personlighetsforstyrrelser som "et vedvarende mønster av indre erfaring og atferd som avviker fra kulturens forventninger." De stadfester at lidelsen må være av lang varighet, ha tidlig debut og være gjennomgripende i en rekke situasjoner. Med unntak av passiv-aggressiv personlighetsforstyrrelse som ble flyttet til appendikset, beholdt man kategoriene fra DSM-III og grupperingen i klustere (Oldham, 2009).

1.5 Dimensjonal versus kategorisk tilnærming

1.5.1 Kategorisk tilnærming

Den kategoriske modellen i DSM-III ble kritisert nesten umiddelbart etter publisering i 1980 (Frances, 1980). Frances mente at den dårlige reliabiliteten til personlighetsforstyrrelsesdiagnosene, skyldtes at personlighetsforstyrrelser ikke er separate enheter, men kontinuerlige konstrukter. Kritikken var en del av en større debatt om diagnoser er bedre representert av dimensjoner, enn kategorier. Dette er en debatt som har omfattet de fleste diagnoser i DSM, men som har vært spesielt tydelig for personlighetsforstyrrelsesdiagnosene (Widiger & Samuel, 2005). Tilhengere av den kategoriske modellen for personlighetsforstyrrelser har påpekt flere fordeler. Et kategorisk system er til stor hjelp når klinikere skal bestemme om individer behøver behandling eller ikke; de som har en diagnose behøver behandling, de som ikke har en diagnose behøver ikke en behandling. En annen mye sitert fordel er at kategoriske systemer er lett å bruke når man skal kommunisere diagnoser. Det er mye informasjon som er innbakt i en enkel diagnostisk merkelapp (Trull & Bennett, 2005). Det tredje er at det kategoriske systemet er noe klinikere er godt kjent med. De fleste nåværende og tidligere diagnoser har vært kategoriske (Widiger, 1991).

Forskning har imidlertid bidratt med å kartlegge en rekke svakheter ved den kategoriske modellen: Komorbiditet av ulike personlighetsforstyrrelser har blitt observert i en rekke studier (Coid, et al., 2009; Grant, Stinson, Dawson, Chou & Ruan, 2005; Zimmerman & Rotschild, 2005; Oldham, Skodol, Kellman, Hyler, & Rosnick, 1992). Uttrykket komorbiditet viser til forekomsten av flere diagnoser på samme tid (Widiger & Samuel, 2005). Det er vanligere for pasienter å bli diagnostisert med flere personlighetsforstyrrelsesdiagnoser, enn å bli diagnostisert med bare en (Grant et al., 2005; Zimmerman & Rotschild, 2005). Dette skaper forvirring rundt hvilken behandling som skal tilbys, og gjør inkludering i forskningsprosjekter vanskelig. Forskere kan ikke vite om pasienter med mer enn en diagnose virkelig er representative for den diagnostiske kategorien (Morey, Benson, Busch & Skodol, 2015).

For det andre, er det stor grad av variasjon (heterogenitet) blant pasienter som får den samme diagnosen (Trull & Durett, 2005). Dette skyldes at dagens personlighetsforstyrrelsesdiagnoser er polytetiske. Det vil si at en pasient bare trenger å ha noen, og ikke alle symptomene som en diagnose spesifiserer. For å bli diagnostisert med borderline personlighetsforstyrrelse må man for eksempel oppfylle fem av ni diagnosekriterier. Dette gjør at det er hele 256 måter å oppfylle kriteriene for en slik diagnose på. To individer med samme diagnose kan dermed ha få diagnostiske trekk til felles (Johansen, Karterud, Pedersen, Gude, & Falkum, 2004). I slike tilfeller vil behandlingsindikasjonen være svært forskjellig, og den kliniske nytteverdien av diagnosen vil være liten (Morey et al., 2015).

For det tredje er de diagnostiske tersklene for hver personlighetsforstyrrelsesdiagnose basert på ekspert uttalelser, ikke empiri. Det gjør at de er svært vilkårlige (Widiger & Trull, 2007). Man kan spørre seg hvor forskjellige en pasient som oppfyller 5 av 8 kriterier for avhengig personlighetsforstyrrelse er (diagnostisk terskelverdi), fra en pasient som oppfyller 4 av 8 (subterskel) (Skodol, Bender, Gunderson & Oldham, 2014).

Videre kan det virke som DSM-III modellen ikke dekker personlighetspatologi på en god måte, på tross av at det er kriterier for ti ulike personlighetsforstyrrelser. En metaanalyse gjort av Verheul og Widiger (2004) viste at PD-NOS er en av de vanligste personlighetsforstyrrelsesdiagnosene i forskningssettinger og den mest brukte personlighetsforstyrrelsesdiagnosen i klinisk praksis.

Tilslutt, har ulike måleinstrument vist dårlig konvergerende validitet (Clark, Livesley, Morey, 1997), og det har også vært svak temporal reliabilitet ved diagnostisering av personlighetsforstyrrelser gjort på ulike tidspunkt. Dette til og med når tidsintervallet er så lite som en uke (Zimmerman, 1994).

1.5.2 Dimensjonal tilnærming

Svakhetene ved den kategoriske modellen foreslår at en dimensjonal modell kan være en bedre måte å fremstille personlighetspatologi på (Widiger & Samuel, 2005).

Det er funnet lite empirisk støtte for at personlighetsforstyrrelser er klinisk separate enheter, selv i store kliniske utvalg som har brukt statistiske metoder for å avdekke kategorier (Eaton, Krueger, South, Simms & Clark, 2011). Det er i motsetning, mye data som tyder på at menneskelig personlighet varierer langs et kontinuum. Disse resultatene kommer fra studier gjort på personlighetstrekk i normalbefolkningen. Personlighetstrekk har vist seg å ha en hierarkisk dimensjonal struktur, funnet i både selv og observatør data, og i en rekke kulturer (McCrae & Costa, 1997; McCrae & Terracino, 2005). Dimensjonale modeller bidrar også til å gjøre fenomener som komorbiditet og heterogenitet mer forståelig (Skodol, 2012). Heterogenitet kan forklares av at mennesker kan variere på en rekke personlighetsdimensjoner. Komorbiditet foreslår tilstedeværelsen av noe felles, en delt patologi, som for eksempel en felles dimensjon (Kendler, Prescott, Myers & Neale, 2003).

Ulike dimensjonale tilnærminger til personlighetsforstyrrelser har blitt foreslått som alternativer til DSM kategorier (Widiger & Simonsen, 2005). En tilnærming er å gjøre om de eksisterende diagnostiske kategoriene til dimensjoner. En fordel med en slik tilnærming er at den vil beholde de eksisterende kategoriene, og på denne måten gjøre overgangen til et dimensjonalt system lettere. En annen tilnærming er å reorganisere de eksisterende diagnostiske kriteriene til gyldige dimensjoner. En tredje tilnærming er å identifisere spektrum som går på tvers av både symptomlidelser og personlighetsforstyrrelser (Widiger & Simonsen, 2005), som internalisering og eksternalisering (Kruger et al., 2001). En fjerde tilnærming foreslår personlighetsforstyrrelser på et kontinuum med generell personlighetsfungering. Det vil si, de er ekstremvarianter av normale personlighetstrekk. Dette bidrar til å viske ut grensen mellom normal og abnormal personlighetsfungering. Det er en rekke trekk modeller som har blitt relatert til DSM-IV kategoriene (Widiger & Simonsen, 2005). Den

mest kjente er fem faktor modellen som består av domeneene: nevrotisisme, ekstroversjon, åpenhet til erfaring, medmenneskelighet og samvittighetsfullhet (Costa & McCrae, 1990).

1.5.3 Hybridmodeller

En løsning på debatten om psykiske lidelser bør klassifiseres som kategorier eller dimensjoner, er hybrid modeller. Hybrid modeller er en kombinasjon av dimensjonale og kategoriske elementer. Noen personlighetsforstyrrelseseksperter (Blashfield, 1993; Benjamin, 1993) foreslo dette allerede før utgivelsen av DSM-IV. Blashfield (1993) foreslo at oppdelingen av klassifikatoriske modeller i to typer er overforenklet, og mente at det finnes flere varianter å velge mellom. I 2007 utførte Bernstein et al. (2007) også en spørreundersøkelse av personlighetsforstyrrelseseksperter. Han fant at 87% av ekspertene mente at personlighetspatologi er et dimensjonalt fenomen, 74% mente at den kategoriske modellen i DSM-IV burde bli endret, og 70% mente at det beste alternativet ville være en kombinasjon av kategorier og dimensjoner. Dette var noe som lå til grunn da utarbeidingen av DSM-5 startet.

1.6 Utviklingen av DSM-V

Forberedelsene til DSM-V startet allerede i 1999 med en DSM-V planleggingskonferanse sponset av APA og NIMH. Et resultat av denne konferansen var seks "white papers" som skulle stimulere til forskning og diskusjon i feltet forut for revisjonsprosessen. Artiklene ble publisert i boken "A Research Agenda for DSM-V." I introduksjonen stiller Kupfer et al. (2002) seg kritiske til den kategoriske tilnærmingen til psykiske lidelser. Forfatterne antyder at et paradigme skifte må til for å forbedre klassifiseringssystemet. I samme bok identifiseres syv hovedtema for utviklingen av DSM-V. Et av temaene var å vurdere om psykiske lidelser heller burde bli representert av dimensjoner, enn av kategorier. Rounsaville et al. (2002) utdyper dette:

" There is a clear need for dimensional models to be developed and their utility compared with that of existing typologies in one or more limited fields, such as personality. If a dimensional system performs well and is acceptable to clinicians, it might be appropriate to explore dimensional approaches in other domains (e.g., psychotic or mood disorders.)" (s. 13)

Personlighetsforstyrrelser ble altså identifisert som en mulig test arena for utviklingen av dimensjonale diagnosemodeller.

I 2007 ble personlighet og personlighetsforstyrrelse arbeidsgruppen opprettet. Andrew E. Skodol ble utnevnt som leder for gruppen, og fikk i oppgave å velge de ti andre medlemmene i samarbeid med DSM-V Task Force. Det ble valgt representanter fra både psykologi og psykiatri, fra ulike teoretiske bakgrunner og med klinisk- eller forskningsekspertise (Skodol, Bender, Oldham, 2009). Ledelsen i DSM ba medlemmene om å bevege seg bort fra DSM-IV og mot en mer dimensjonal tilnærming. Gruppen satte seg som mål å lage en empirisk basert modell som skulle eliminere eller redusere en del av svakhetene med den kategoriske modellen (Zachar, Krueger & Kendler, 2016). Prosessen ble dessverre langt fra enkel. Den ble preget av både indre uenigheter og mye ytre kritikk (Torgersen, 2014).

Arbeidet begynte med en litteraturgjennomgang av hver DSM-IV personlighetsforstyrrelse. Medlemmene ble så enige om hvilke typer som hadde mest empirisk støtte. Uenigheter begynte imidlertid raskt å komme til syne. Et av de første spørsmålene var om en dimensjonal tilnærming skulle baseres på allerede eksisterende modeller, eller om det skulle lages en helt ny modell. Dette ble det ikke enighet om.

Høsten 2009 introduserte Skodol sitt første forslag. I forslaget ble personlighetsforstyrrelser definert som mangler ved selv og interpersonlig fungering, tilstedeværelse av patologiske personlighetstrekk og fem DSM-IV personlighetsforstyrrelser. Personlighetsforstyrrelsene skulle bli vurdert ved hjelp av en prototypisk tilnærming. Dette innebar å matche en pasient opp mot en narrativ beskrivelse. Forslaget fikk en blandet mottakelse i arbeidsgruppen. Noen medlemmer følte at deres input og ekspertise ikke hadde blitt hørt, og at modellen var for kompleks til å kunne brukes i klinisk praksis. Andre mislikte de narrative prototypene. Da modellen ble publisert på nettsiden til DSM-5 fikk den også mye kritikk utenfra. (Zachar, Krueger, Kendler, 2016). Kritikere hevdet at forslaget var for radikalt, for komplekst, manglet empirisk grunnlag (Gunderson, 2013) og at trekk delen av modellen ikke ville være nyttig i klinisk praksis (Shedler et al., 2010).

Kritikken gjorde DSM Task Force svært bekymret for forslaget. De etablerte en komitee som skulle bistå arbeidsgruppen med rådgivning (Zachar, Kreuger & Kendler, 2016). I 2011 publiserte arbeidsgruppen en hybrid modell med dimensjonale elementer for å imøtekomme ny forskning, og kategorier for å bevare kontinuitet med modellen i

DSM-IV (Morey, Skodol, Oldham, 2014). De droppet de omstridte prototypene til fordel for diagnostiske kriterier. Istedenfor skulle hver av de gjenværende personlighetsforstyrrelse nå bli definert av spesifikke svekkelser i selv og interpersonlig fungering, samt små grupper av dimensjonale personlighetstrekk. Dette forslaget fikk sin støtte fra DSM Task Force (Zachar, Kreuger & Kendler, 2016).

To gruppemedlemmer var imidlertid svært misfornøyd med hybridmodellen. Så misfornøyd at de valgte å forlate gruppen i protest. Dette var John Livesley og Roel Verheul (Zachar, Kreuger & Kendler, 2016). Begge kritiserte modellen for å være for kompleks, mangle helhet og mangle kontinuitet med DSM-IV (Verheul, 2012; Livesley, 2012). Undertittelen på Verheuls kritikk "a heroic and innovative but nevertheless fundamentally flawed attempt to improve DSM-IV" (Verheul, 2012) sier mye om hvor sterkt han mislikte modellen.

Da hybrid modellen ble publisert på DSMs nettsider var de offisielle DSM feltforsøkene allerede i gang. Personlighet og personlighetsforstyrrelsens arbeidsgruppen ville undersøke modellens validitet, men dette var det ikke tid eller mulighet for i feltforsøkene. To vitenskapelige godkjenningsskomiteer ble engasjert for å vurdere forslaget. Den første kom frem til at det var gjort for lite forskning på modellen. Den andre komiteen satte spørsmålstegn ved klinisk nytteverdi. Begge rådet APA til å ikke vedta den nye modellen. APA inngikk et kompromiss. De valgte å beholde DSM-IV modellen i hoveddelen av DSM-5, og plassere hybrid modellen i seksjon III av manualen. Hybrid modellen fikk navnet den alternative modellen for personlighetsforstyrrelser, for å antyde bruk av denne modellen som et alternativ (Zachar, Krueger, Kendler, 2016).

2 Den alternative modellen for personlighetsforstyrrelser

2.1 Oversikt

Den alternative modellen for personlighetsforstyrrelser består av fire deler. Dette er a, en dimensjonal vurdering av svekkelser i personlighetsfungering; b, en dimensjonal vurdering av patologiske personlighetstrekk; c, seks spesifikke personlighetsforstyrrelser, samt en trekk spesifikk personlighetsforstyrrelse og d, generelle kriterier for personlighetsforstyrrelser (Skodol et al., 2011).

2.2 Generelle kriterier for personlighetsforstyrrelser

De generelle kriteriene i den alternative DSM-5 modellen slår fast at for å sette en personlighetsforstyrrelsesdiagnose må det foreligge en moderat eller mer alvorlig svekkelse i personlighetsfungering (kriterium A) og det må forekomme patologiske personlighetstrekk (kriterium B). Videre uttrykker kriteriene at svekkelsen i personlighetsfungering og individets personlighetstrekk må komme til uttrykk i en rekke kontekster og situasjoner (kriterium C); være relativt stabile over tid og kunne bli sporet tilbake til ungdomstid eller tidlig voksenalder (kriterium D); ikke være bedre forklart av en annen psykisk lidelse (kriterium E); ikke skyldes substansbruk eller en annen medisinsk tilstand (kriterium F); og ikke regnes som normalt for et individs utviklingstrinn eller sosiokulturelle miljø (kriterium G) (APA, 2013).

2.3 Kriterium A: Nivå av personlighetsfungering

Nivå av personlighetsfungering er et mål for å vurdere tilstedeværelsen og alvorlighetsgraden av personlighetspatologi (APA, 2013). DSM-5 arbeidsgruppen kom frem til at selve kjernen i personlighetspatologi er forstyrrelser i måten å tenke om seg selv og andre på (Skodol et al., 2011). Modellen deler derfor personlighetsfungering inn i selvfungering og interpersonligfungering. Selvfungering består av elementene identitet og målrettethet, mens interpersonligfungering består av elementene empati og nærhet. Identitet blir definert som "å oppleve seg selv som unik, med klare grenser mellom selv og andre; å ha stabil selvaktelse med presis selv vurdering; å ha kapasitet

for og evne til å regulere et bredt spekter av emosjoner” (APA, 2013, s. 762).

Måltrettethet blir definert som ”å forfølge meningsfylte kortsiktige og livslange mål; å anvende passende og prososiale indre standarder for atferd; å ha evne til selvrefleksjon.” (APA, 2013, s. 762) Empati blir definert som ”forståelse og verdsettelse av andres erfaringer og motivasjoner; toleranse for andres perspektiver; å være klar over effekten egne handlinger har på andre.” (APA, 2013, s. 762) Tilslutt, blir nærhet definert som ”å ha dype og varige forhold med andre; ønske om og evne til nærhet; å strebe etter gjensidighet i forhold” (APA, 2013, s. 762).

Grad av svekkelse måles på et kontinuum bestående av fem nivåer. Nivå 0 er ingen eller lite svekkelse, nivå 1 er noe svekkelse, nivå 2 er moderat svekkelse, nivå 3 er alvorlig svekkelse og nivå 4 er meget alvorlig / ekstrem svekkelse. For hvert element (identitet, måltrettethet, empati, nærhet) er det skrevet en beskrivelse for hvert nivå av personlighetsfungering. En tabell med 20 slike beskrivelser er inkludert i DSM-5 manualen. Beskrivelsene består av psykologiske trekk som er tenkt å være typiske for kombinasjonene av element og nivå. Beskrivelsene for identitet er for eksempel:

Nivå 0: ”Har vedvarende bevissthet om et unikt selv; opprettholder passende roller og adekvate grenser. Har stabil og selvregulert positiv selvaktelse, med presis selvvurdering. Evner å erfare, tolerere og regulere et bredt spekter av emosjoner”.

Nivå 1: ”Har relativt intakt selvfølelse, med noe nedsatt bevissthet rundt grenser når man erfarer sterke emosjoner og mental uro. Selvaktelse til tider nedsatt, i overkant kritisk eller noe forvrengt selvvurdering. Sterke emosjoner kan være plagsomme, forbundet med en innskrenking i omfanget av emosjonelle opplevelser”.

Nivå 2: ”Er i stor grad avhengig av andre for å definere egen identitet, med tilsvarende svekket grenseoppgang til andre. Har sårbar selvaktelse drevet av overdreven engstelse rundt andres evaluering, med ønske om å bli godtatt. Har følelse av ufullstendighet eller mindreverdighet, med resulterende forstørret eller minsket selvaktelse. Emosjonell regulering er avhengig av positive vurderinger utenfra. Trusler mot selvaktelsen kan føre til sterke følelser som aggresjon eller skam”.

Nivå 3: ”Har svak opplevelse av autonomi og egenvilje; opplever mangel på identitet, eller tomhet. Grenser til andre er dårlige eller rigide: kan være overidentifisert med andre, legge for stor vekt på uavhengighet fra andre, eller pendle mellom disse. Skjør selvaktelse som lett påvirkes av hendelser, og selvbildet mangler sammenheng. Selvvurdering er unyansert: selvhat, selvforherligelse, eller en ulogisk, urealistisk

kombinasjon. Emosjoner kan skifte raskt, eller en kronisk, urokkelig følelse av fortvilelse”.

Nivå 4: ”Opplevelse av et unikt selv og følelse av autonomi og egenvilje er nærmest fraværende, eller er organisert rundt en oppfattet ytre forfølgelse. Forvirrende eller manglende grenser overfor andre. Har svakt eller fordreid selvbilde som lett trues av interaksjoner med andre; vesentlig forstyrrelser og forvirring rundt selvvurdering. Emosjoner er ikke i overenstemmelse med kontekst eller indre opplevelse. Hat og aggresjon kan være dominerende affekter, selv om de kan forkastes og attribueres til andre” (APA, 2013, s.775-778, tabell 2).

Tabellen med de 20 beskrivelsene er Nivå av personlighetsfungeringsskalaen / level of personality functioning scale (LPFS)(se vedlegg A). I forskningslitteraturen benyttes gjerne forkortelsen LPFS ved bruk av Nivå av personlighetsfungeringsskalaen (Few et al., 2013; Zimmerman et al., 2015). Nivå av personlighetsfungeringsskalaen (LPFS) finner grad av svekkelse i personlightefungering.

Ved bruk av personlighetsfungeringsskalaen velger klinikerer det nivået som passer best med individets grad av svekkelse. For å kunne sette en personlighetsforstyrrelsesdiagnose må det foreligge en moderat svekkelse i personlighetsfungering (nivå 2) (APA, 2013). Denne terskelen er valgt basert på en studie gjort av Morey, Bender & Skodol (2013). Studien konkluderte med at ved en moderat eller mer alvorlig svekkelse er det 84,6 % sensitivitet og 72,7% spesifisitet for å identifisere pasienter som oppfyller kravene for en DSM-IV personlighetsforstyrrelse. Modellen forklarer at graden av svekkelse predikerer om et individ har mer enn en personlighetsforstyrrelse eller det som blir regnet som en av de mer alvorlige personlighetsforstyrrelsene (APA, 2013).

2.4 Kriterium B: Patologiske personlighetstrekk

Et personlighetstrekk er en tendens til å føle, tenke, oppfatte og oppføre seg på en bestemt måte over tid og i ulike situasjoner. Personlighetstrekk er dimensjonale fenomen, og alle individer har mer eller mindre av et trekk. Noen personlighetstrekk er svært spesifikke og beskriver et smalt spekter av atferd, mens andre er brede og beskriver en rekke atferder. Brede trekk dimensjoner blir kalt domener, mens

spesifikke trekk dimensjoner blir kalt fasetter. Trekk domener består av en rekke trekk fasetter som ofte opptrer sammen (APA, 2013).

Kriterium B består av 25 patologiske personlighetstrekk som er organisert inn i fem brede domener. De fem domenene er negativ affekt (versus emosjonell stabilitet), distansering (versus ekstroversjon), antagonisme (versus medmenneskelighet), manglende hemning (versus samvittighetsfullhet) og psykotisisme (versus klarhet). Manualen beskriver de fem brede domenene som "maladaptive variants of the five domains of the extensively validated and replicated personality model known as the "Big Five", or Five Factor Model of personality (FFM)..." (APA, 2013, s. 773) og trekkene i parentes som de sunne, adaptive, resiliente personlighetstrekkene som er motpolene til de patologiske personlighetstrekkene (APA, 2013). Big Five består av domenene ekstroversjon, medmenneskelighet, samvittighetsfullhet, emosjonell stabilitet og åpenhet til erfaring (Norman, 1963; McCrae & Costa, 1985). Kriterium B ser derfor ut til å bestå av varianter av fire av disse domenene, ikke fem, og i tillegg psykotisisme som ikke er en del av Big Five.

Klinisk bruk av trekkmodellen består i å vurdere alle de fem brede personlighetsdomenene. Dette kan man gjøre ved hjelp av måleinstrumentet Personality Inventory for DSM-5 (PID-5). PID-5 består både av en selv-rapport form og en informant-rapport form. Sistnevnte kan bli fylt ut av en som kjenner pasienten godt, for eksempel en ektefelle. En grundig undersøkelse vil være å innhente både en selv-rapport og en informant-rapport på alle de 25 fasettene. Hvis ikke dette er mulig, kan man vurdere kun de fem domenene. Dette er godt nok hvis man bare behøver en generell beskrivelse av en pasients personlighet. Hvis det er personlighetsproblemer som er fokus for behandlingen, anbefales derimot den grundigere undersøkelsen (APA, 2013). Skodol, Morey, Bender & Oldham (2015) skriver at et trekk som vurderes som moderat eller svært beskrivende for en person regnes som tilstede.

2.5 Personlighetsforstyrrelser

Utfra den alternative modellen for personlighetsforstyrrelser kan man diagnostisere seks spesifikke personlighetsforstyrrelser: antisosial, unnvikende, borderline, narcissistisk, tvangspreget og schizotyp personlighetsforstyrrelse. Hver personlighetsforstyrrelse er definert av typiske svekkelser i personlighetsfungering, og

karakteristiske patologiske personlighetstrekk (APA,2013, s. 763). DSM-5 arbeidsgruppen valgte å beholde akkurat disse seks kategoriene fordi de har vist seg å ha mest empirisk støtte. Kategoriene er like DSM-IV diagnosene med samme navn, men de er ikke identiske (Skodol et al, 2011).

Tilnærmingen inkluderer også en trekk-spesifikk personlighetsforstyrrelse (PD-TS). Denne diagnosen kan bli benyttet hvis man tenker at en personlighetsforstyrrelse er tilstede, men kriteriene for en spesifikk forstyrrelse ikke er oppfylt. De nåværende personlighetsforstyrrelsesdiagnosene som ikke er med i den alternative modellen, kan også bli representert av PD-TS. Dette er paranoid, schizoid, histrionisk og avhengig personlighetsforstyrrelse samt restkategorien PD-NOS (APA, 2013; Skodol et al., 2011).

2.6 Personlighetsforstyrrelsesdiagnose

Den alternative modellen for personlighetsforstyrrelser ble designet for å være en fleksibel og stegvis måte å vurdere personlighetsforstyrrelser på (Skodol et al., 2015). Først vurderer man om det foreligger en moderat eller mer alvorlig svekkelse i minst to av personlighetsfungerings-elementene: identitet, målrettethet, empati og nærhet (APA, 2013, s. 771). Under de generelle kriteriene kommer det ikke tydelig frem at svekkelsen må være i to av elementene. Dette blir spesifisert senere under de diagnostiske kriteriene for hver personlighetsforstyrrelsesdiagnose og under skårings algoritmene (APA, 2013, s. 771). I følge Skodol et al. (2015) vil vurdering av personlighetsfungering gi en kliniker en god pekepinn på om en person har en personlighetsforstyrrelse eller ikke. Deretter vurderer man patologiske personlighetstrekk. For hver personlighetsforstyrrelsesdiagnose er det satt diagnostiske terskler for antall trekk som må være tilstede (APA, 2013).

Det tredje steget er å se om individet har et mønster av svekkelse i personlighetsfungering og patologiske personlighetstrekk som matcher en av de seks personlighetsforstyrrelsesdiagnosene. Hvis et individ har patologiske personlighetstrekk som ikke er nødvendig for å sette diagnosen, men som kan ha klinisk relevans, kan disse bli notert som spesifikke trekk (APA, 2013, s. 771).

Individer som har et personlighetsfungering og trekk mønster som er ulikt de seks spesifikke personlighetsforstyrrelsene, kan bli diagnostisert med trekk-spesifikk

personlighetsforstyrrelse (PD-TS). Da spesifiserer man grad av personlighetsfungering og de patologiske personlighetstrekkene som er tilstede (APA, 2013).

Det siste steget er å vurdere de generelle kriteriene. Hvis disse betingelsene er tilstede, kan man i følge Skodol et al. (2015) sette en personlighetsforstyrrelsesdiagnose.

2.7 Rasjonale for nivå av personlighetsfungering

2.7.1 Alvorlighetsgrad

Noe av det DSM-V arbeidsgruppen la til grunn da de utviklet en skala som måler alvorlighetsgraden til personlighetspatologi, var Hopwood et al. (2011) sitt studie av pasienter med personlighetsforstyrrelser (Skodol et al, 2011). Hopwood et al. (2011) fant ut at alvorlighetsgrad er den viktigste enkeltfaktoren som predikerer samtidig og fremtidig fungering. Forfatterne konkluderte med at personlighetsforstyrrelser er best beskrevet av et alvorlighetsgrads kontinuum og stilmessige symptom uttrykk. En rekke eksperter mener at alvorlighetsgrad må være en del av enhver dimensjonal modell som skal vurdere personlighetspatologi (Tyrer, 2005) og at det burde være en sentral komponent i enhver fremtidig diagnostisk manual (Bornstein 1998, Bornstein & Huprich 2011, Crawford, Koldobsky, Mulder & Tyrer 2011, Leising og Zimmermann 2011, Widiger og Trull 2007). Tyrer (2005) mente det var svært viktig å vite hvilke personer som har de mest alvorlige personlighetsforstyrrelsene, for å kunne ta vare på deres behov. Resultater fra forskning tyder også på at alvorlighetsgraden til personlighetsforstyrrelser har mer betydning for behandlingsplanlegging, forløp og prognose, enn hvilken type personlighetsforstyrrelse en pasienter har (Karterud et al., 2003, Hopwood et al., 2011, Tyrer og Johnson, 1996, Yang, Coid & Tyrer, 2010).

Det nåværende diagnosesystemet er blitt kritisert for å ikke ha med et mål på alvorlighetsgraden til personlighetsforstyrrelser (Bernstein, 1998). De generelle kriteriene for personlighetsforstyrrelser er blitt beskrevet som for vage (Livesley, 1998) og GAF er ikke spesifikk nok for personlighetspatologi (Bender, Morey, Skodol, 2011). Arbeidsgruppen valgte dermed å utarbeide et nytt mål (Skodol et al., 2011).

2.7.2 Selv og interpersonlig fungering

Arbeidsgruppen bestemte at en personlighetsfungeringsskala skulle bestå av psykologiske prosesser som er felles for alle personlighetsforstyrrelser. De utførte en

litteraturgjennomgang og fant ut at dårlig tilpassete mentale representasjoner av selv og andre er det som ligger til grunn for personlighetspatologi (Bender, Morey & Skodol, 2011). En tilnærming de la vekt på var Kernbergs (1970/1989) klassifikasjon av personlighetspatologi. Kernberg klassifiserte personlighetsforstyrrelser langs et kontinuum basert på blant annet kvaliteten av et individs mentale representasjoner av selv og andre. Han mente at tvangspregete karakterforstyrrelser er høyt fungerende, narcissistiske karakterforstyrrelser er middels fungerende mens antisosiale, paranoide, schizoide og borderline karakterforstyrrelser er dårlig fungerende.

Arbeidsgruppen viste også til omfattende empirisk litteratur som viser at personlighetsforstyrrelser er assosiert med forstyrret tenkning om selv og andre (Bender, Morey & Skodol, 2011). Salvatore, Nicolo og Dimaggio (2005) fant for eksempel at pasienter med paranoid personlighetsforstyrrelse vanligvis ser på seg selv som svake og inadekvate, mens de ser på andre som fiendtlige og svikefulle. Jovev og Jackson (2004) **demonstrerte** at individer med unnvikende personlighetsforstyrrelse tenker på seg selv som ødelagte, og forventer å bli forlatt på grunn av sine mangler. Likeledes har studier som har brukt Ekmans ansikter som stimuli, vist at pasienter med borderline personlighetsforstyrrelse lettere tilegner nøytrale ansikter negative emosjoner (Donegan et al., 2003; Wagner & Linehan, 1999). Med basis i dette begynte arbeidsgruppen å utarbeide et mål på alvorlighetsgrad av personlighetspatologi.

2.7.3 Utvikling av skalaen

Første steget i utformingsprosessen av et dimensjonalt mål var en litteraturgjennomgang av aktuelle måleinstrumenter. Bender, Morey og Skodol (2011) plukket ut måleinstrumenter basert på en rekke kriterier. Hvert instrument skulle a, bestå av dimensjoner, ikke kategorier; b, ha et selv-andre fokus; c, ha blitt brukt i studier med kliniske utvalg; d, inneholde konsepter som er nyttige for en rekke klinikere; e, kunne brukes til å skåre klinisk intervju materiale og f, ha publiserte psykometriske data. Måleinstrumentene som møtte disse kriteriene var "The Quality of Object Relations Scale", "Personality Organization Diagnostic Form, versions I and II", "Object Relations Inventory", "Social Cognition and Object Relations Scale", og "Reflective Functioning Scale" (Bender, Morey, Skodol, 2011). Gjennomgangen viste at et selv-andre perspektiv er nyttig både i å forstå mentale prosesser, planlegge behandling og måle endring underveis i behandling (Bender, Morey & Skodol, 2011).

De fleste måleinstrumentene som ble vurdert var designet for forskning, og krevde mye opplæring før anvendelse. Det var dermed ikke praktisk å innføre et av disse til klinisk bruk i DSM-5. Flere av konseptene i måleinstrumentene hadde imidlertid god reliabilitet, validitet og viste seg nyttige for å beskrive grad av personlighetspatologi. Disse fungerte som grunnlag for å lage en ny skala (Bender, Morey & Skodol, 2011; Skodol et al. 2011).

Morey et al. (2011) utførte sekundære dataanalyser for å validere skalaen og bidra med empirisk støtte. Resultatene støttet en global alvorlighetskala som var assosiert med sannsynligheten for å få en DSM-IV diagnose og sannsynligheten for å få flere DSM-IV diagnoser. Forfatterne konkluderte med at selv- og interpersonlig fungering er viktige komponenter på en slik dimensjon.

Etter datanalysen forenklet man skalaen ved å beholde kun de komponentene som hadde svært god reliabilitet. Tilslutt gikk arbeidsgruppen gjennom skalaen. De endret på språket der de anså det som for komplisert eller der de anså at det tilhørte en spesifikk teoretisk sjargong. De slo sammen konsepter som virket overlappende og fjernet konsepter som var overflødige. Resultatet ble nivå av personlighetsfungeringsskalaen (Bender, Morey & Skodol, 2011).

2.8 ICD 11 – forslaget

Det har også blitt foreslått en dimensjonal klassifikasjon av personlighetsforstyrrelser for ICD-11. Ved første øyekast virker ICD-11 forslaget svært ulikt AMPD, men de er konseptuelt like (Tyrer, Reed & Crawford, 2015). ICD-11 forslaget legger i enda større grad vekt på hvor alvorlig personlighetsforstyrrelsen er, enn i AMPD. Det skilles mellom mild personlighetsforstyrrelse, moderat personlighetsforstyrrelse og alvorlig personlighetsforstyrrelse. Disse er karakterisert av økende grad av interpersonlige problemer og problemer i yrkesliv. Ved moderat og alvorlig personlighetsforstyrrelse forekommer det også skade av en selv og andre. Det har blitt foreslått en subterskel grad av lidelsen, kalt personlighetsvansker. Personlighetsvansker referer til en forstyrrelse som kun er tilstede periodevis, i spesifikke miljøer eller omstendigheter (for eksempel under stress). ICD-forslaget inkluderer, slik som AMPD, også personlighetstrekk for å beskrive individer. Fire av disse er nesten de samme som i AMPD. Dette er negativ affekt, antisosiale trekk, manglende hemming, og atskillelse.

Istedenfor trekkdomenet psykotisisme, har ICD-11 et tvangspreget domene (Tyrer, Reed & Crawford, 2015).

Det som i størst grad skiller ICD-11 forslaget fra AMPD, er at man har fjernet alle de type spesifikke kategoriene. ICD-11 vurderer kun om det er tilstedeværelse av personlighetsforstyrrelse, og grad av dette. ICD-11 forslaget vurderer heller ikke selvpatologi, kun interpersonlig patologi. Dette er noe ICD-11 forslaget har blitt kritisert for (Tyrer, Reed, Crawford, 2015).

3 Foreløpige publikasjoner på Den alternative modellen for personlighetsforstyrrelser

Den alternative modellen for personlighetsforstyrrelser (AMPD) er plassert i seksjon III i DSM-5, med formålet at det skal appellere til mer forskning på modellen (APA, 2013). En av årsakene til at den ikke ble vedtatt som hovedmodell for diagnostisering av personlighetsforstyrrelser var begrenset empiri omkring de ulike elementene i modellen, blant annet LPFS. Det har blitt etterlyst forskning på modellens validitet, reliabilitet og kliniske nytte (Clarkin & Huprich, 2011). Modellen har vært kritisert for at den er for kompleks, mangler helhet, stiller for høye krav til raterne og er vanskelig å lære (Leising & Zimmermann, 2011, Morey, Skodol & Oldham, 2014, Pilkonis, Hallquist, Morse & Stepp, 2011, Tyrer, 2012, Zimmermann et al., 2014). I følge Morey, Benson, Busch og Skodol (2015) har imidlertid de tidlige studiene på modellen vært lovende. Videre vil vi oppsummere noe av den foreløpige forskningen på modellen med hovedfokus på LPFS.

3.1 Klinisk nytte

Studier som har sett på den kliniske nytten til AMPD har gjennomgående argumentert for at den er nyttig for klinisk forståelse, vurdering av prognose og behandlingsplanlegging (Bach, Markon, Simonsen & Krueger, 2015, Morey, Bender & Skodol, 2013, Simonsen & Simonsen, 2014) Bach et al. (2015) fokuserte spesielt på at modellen gjør at man tydeligere ser hvordan personen lider utover at den har en personlighetsforstyrrelsesdiagnose. I tillegg poengterte de at modellen kan anvendes på tvers av ulike perspektiver og terapeutiske tilnærminger.

AMPD har også blitt vurdert som like eller mer klinisk nyttig enn DSM-IV kategoriene (Morey, Bender & Skodol, 2013, Morey, Skodol & Oldham, 2014). Morey, Skodol og Oldham (2014) fant at dette spesielt var knyttet til forståelse av problematikken, kommunikasjon med pasienter, beskrivelse av pasientene og behandlingsplanlegging. Imidlertid ble DSM-IV vurdert som mer nyttig med tanke på kommunikasjon med andre profesjonelle og at den er lettere å bruke. Forfatterne argumenterte for at det kan skyldes at den nåværende kategoriske tilnærmingen er så

kjent for klinikere. I deres studie trakk de også frem at trekkkomponenten i AMPD muligens oppleves som mer nyttig enn LPFS (Morey, Skodol & Oldham, 2014). Det er også mulig at noe av den kliniske nytten som er funnet så langt er knyttet til at klinikeren har god kjentskap til pasienten fra før (Morey, Bender & Skodol, 2013).

3.2 Validitet

Aspekter ved validiteten til AMPD, og mer spesifikt validiteten til LPFS er blitt undersøkt i flere studier. Validiteten er avgjørende for om LPFS faktisk måler det den skal. En viktig faktor er dermed om LPFS er relatert til faktisk svekkelse hos pasienter med personlighetsforstyrrelser. Dette kan gjøres ved å undersøke om LPFS vurderingen er assosiert med andre mål på personlighetsfungering. Resultater fra studier gir støtte for at dette er tilfelle. DSM-IV personlighetsforstyrrelsesdiagnoser er assosiert med høyere skårer på LPFS (Few et al., 2013, Zimmermann et al.,2014). Det er også en sammenheng mellom antall DSM-IV personlighetsforstyrrelser og LPFS skårer. De med flere komorbide diagnoser kommer altså ut med dårligere nivå av personlighetsfungering.

Few et al. (2013) fant at DSM-IV diagnosene borderline, schizotyp, antisosial og paranoid personlighetsforstyrrelse var sterkest assosiert med svekkelse på LPFS, mens tvangspreget, histrionisk og narsissistisk viste svakest assosiasjoner. Dette er i tråd med annen forskning som har vist at de med schizotyp, paranoid, unnvikende, schizoid og borderline personlighetsforstyrrelser har størst svekkelse i livskvalitet, mens tvangspreget, histrionisk og narsissistisk er mindre assosiert med svekkelse (Cramer, Torgersen & Kringlen, 2006, Cramer, Torgersen & Kringlen, 2007). Nivåene for personlighetsfungering har også vist seg å kunne predikere alvorlighetsgraden ved antisosial og borderline personlighetsforstyrrelse (Amini, Pourshahbaz, Mohammadkhani, Reza, Ardakani & Lotfi, 2015). Oppsummert tyder dermed disse funnene på at nivå av personlighetsfungering (LPFS) fanger opp svekkelse ved personlighetspatologi.

Et annet viktig aspekt ved LPFS er om kriteriet bidrar med inkrementell validitet for vurdering av personlighetspatologi relativt til de andre kriteriene i AMPD, spesielt trekkkomponenten. Noen resultater tyder på at det er begrenset hvor mye tilleggsinformasjon man får fra LPFS sammenlignet med å kun vurdere

trekkkomponenten (Baastiansen et al., 2016, Few et al., 2013), og det er dermed stilt spørsmål ved hvorvidt en separat vurdering av personlighetsvekkelse og patologiske trekk er nødvendig (Few et al., 2013). En mulig forklaring kan være at personlighetsvekkelse kanskje ikke er så atskilt fra patologiske trekk, noe forskning tyder på (Few et al., 2013, Hill, Fudge, Harrington, Pickles & Rutter, 2000). Et annet studie fant imidlertid inkrementell validitet for LPFS konstrukter kontrollert for patologiske trekk ved bruk av andre vurderingsmetoder (Hentschel & Pukrop, 2015). Dette er dermed et område det trengs mer forskning på.

Det er funnet støtte for å inndeling av nivå av personlighetsfungering i en to-faktorstruktur slik som selv og interpersonlig. Imidlertid er det flere faktorer under disse med kryssladninger på tvers av de to.

3.3 Reliabilitet

Reliabiliteten er avgjørende for validiteten. Man kan ikke ha god validitet uten reliabilitet, selv om man kan finne reliabilitet uten god validitet. Dette er dermed et viktig fokus for forskning på LPFS dersom dette kriteriet, og dermed AMPD, skal kunne tas i bruk i klinisk praksis. Reliabiliteten til LPFS har hittil i hovedsak blitt undersøkt ved å se på interrater reliabilitet (Few et al., 2013, Hutsebaut et al., 2016b, Thylstrup et al., 2016, Zimmermann et al., 2014,), det vil si om to klinikere vil skåre den samme pasienten likt. Few et al. (2013) undersøkte dette ved å sammenligne ulike raters LPFS skåringer etter at de hadde sett videoopptak av SCID-II intervjuer. 109 voksne dagpasienter ble intervjuet. De analyserte grad av enighet mellom raterne ved intraclass korrelasjoner (ICC). Intraclass korrelasjonene for domeneene var: ICC = .49 for identitet, ICC = .47 for målrettethet, ICC = .49 for identitet, og ICC = .47 for nærhet. De argumenterte for at mangel på kjentskap med modellen, og komplekse beskrivelser på LPFS kan ha gjort at de ikke fant høyere grad av enighet. Samtidig argumenterte de for at resultatene ga moderat støtte for interraterreliabiliteten til LPFS.

Zimmermann et al. (2014) ønsket å undersøke om uerfarne ratere kunne vurdere pasienters nivå av personlighetsfungering på en presis og reliabel måte. De lot 22 studenter se på 10 videoopptak av erfarne klinikere som intervjuet pasienter med et operasjonalisert psykodynamisk intervju (OPD Task Force, 2008). Studentene ratet pasientene på alle subdomeneene på LPFS ut i fra informasjonen som kom frem i

intervjuet. Subdomeneskåringene ble aggregert til de fire domeneskåringene, og videre til en LPFS totalskåre. Interrater reliabiliteten ble deretter analysert ved intraclass korrelasjoner. Resultatene viste akseptable intraclass korrelasjoner på LPFS total for individuelle ratere med en intraclass korrelasjon = .51 for individuelle ratere. Når interrater reliabiliteten ble aggregert for alle de 22 raterne, fant de en svært høy intraclass korrelasjon = .96. På domenenivå varierte imidlertid reliabiliteten for individuelle ratere fra intraclass korrelasjon = .25 for empati til ICC = .63 for nærhet. Når de så på gjennomsnittet av ratere økte estimatene også her betraktelig og varierte mellom ICC = .88 for empati til ICC = .97 for nærhet. De konkluderte med at funnene deres gir støtte for at det ikke trengs omfattende klinisk ekspertise eller trening for å anvende LPFS.

3.4 Måleinstrumenter for LPFS

De foreløpige validitets – og reliabilitetsfunnene har vist lovende resultater for modellen, men også pekt på områder som kan bedre. Det er for eksempel pekt på forbedringer som kan øke den kliniske nytteverdien til AMPD, blant annet: validitetsskalaer knyttet til pasienters responsstil, mulighet for å inkludere informasjon fra flere informanter, og et instrument for å gjøre LPFS vurderinger på en reliabel, meningsfull og klinisk nyttig måte (Back, Markon, Simonsen & Krueger, 2015). Fravær av instrumenter for å vurdere LPFS har også vært kritikk mot foreløpig forskning (Bender et al., 2014).

De siste årene har imidlertid dette vært fokus for mer forskning, og flere slike instrumenter har blitt utviklet: The Semi-Structured Interview for Personality Functioning *DSM – 5* (STiP-5.1) (Hutsebaut, Kamphuis, Feenstra, Weekers & De Saeger, 2016), the Clinical Assessment of the Levels of Personality Functioning Scale (CALF) (Thylstrup et al., 2016), The LPFS Brief Form (Hutsebaut, Feenstra & Kamphuis, 2016a) og SCID-AMPD modul 1 (Bender et al., 2014). Alle disse instrumentene er utviklet spesifikt for å kartlegge nivå av personlighetsfungering på LPFS. STiP-5.1, CALF og SCID-AMPD modul 1 er kliniske intervjuer, mens LPFS Brief Form er et selvrappport spørreskjema. Validiteten og reliabiliteten til STiP-5.1 og CALF har blitt undersøkt i to studier. Et studie har også sett litt på LPFS Brief Form, og funnet noe foreløpig støtte for validitet og reliabilitet. Ingen studier er foreløpig undersøkt SCID-AMPD modul 1.

Hutsebaut et al. (2016) fant lovende resultater for både validiteten og reliabiliteten til STiP-5.1 som et instrument for å kartlegge nivå av personlighetsfungering på LPFS. De så i likhet med tidligere studier mer spesifikt på interrater reliabiliteten ved rating av videointervjuer. Intervjuerne fikk i forkant opplæring om LPFS og STiP-5.1. Enigheten ble undersøkt ved intraclass korrelasjoner. Resultatene viste gode interraterreliabilitets estimater fra ICC = .58 til ICC = .80 for det klinisk utvalget på 40 deltagere. Når en ikke klinisk gruppe ble inkludert i analysene, økte alle estimatene.

CALF intervjuet er også undersøkt i en studie av Thylstrup et al. (2016). De benyttet også et videodesign, men gjorde to ulike analyser: en sammenligning av skåren til intervjueren og rateren som så videointervjuet ved Pearson r , samt en sammenligning av to ratere som så på samme videointervju ved intraclass korrelasjoner. Klinikerne fikk ingen trening utenom å få tildelt LPFS og intervjuet. Pearson r korrelasjonen mellom ratingen til intervjuerne og ratingen til de som så det samme intervjuet på video totalskåren på LPFS var $r = .59$. For domenene varierte den mellom $r = .16$ for identitet og $r = .66$ for målrettethet. Intraclass korrelasjonen mellom to videorater varierte mellom ICC = .31 for identitet til ICC = .60 for målrettethet for den kliniske gruppen. Thylstrup et al. (2016b) argumenterte for at korrelasjonene var sterke nok til å støtte videre forskning på LPFS, men ga uttrykk for noe forbehold om bruken av CALF som et diagnostisk instrument, spesielt for identitetsdomenet.

Det er foreløpig ingen instrumenter som er offisielt godkjent, og det vil sannsynligvis kreve mer forskning på validiteten og reliabiliteten til både måleinstrumenter og LPFS før det tas i bruk i praksis. Imidlertid er dette et forskningsfelt der det skjer mye, og resultatet av forskningen vil mest sannsynlig være svært aktuelt når det gjelder implementering av Den alternative modellen for personlighetsforstyrrelser, og dermed kartlegging av fungering på LPFS i klinisk praksis.

4 Målsetningen med studiet

Vi ønsker å undersøke reliabiliteten til intervjuet SCID-AMPD modul 1 (Bender, Skodol, First & Oldham, 2014). På denne måten vil vi også kunne bidra med mer kunnskap om nivå av personlighetsfungering i Den alternative modellen for personlighetsforstyrrelser i DSM-5 (APA, 2013). SCID-AMPD modul 1 er som nevnt utviklet som et måleinstrument for å kartlegge personlighetsfungering ved LPFS. Det å kunne gjøre reliable vurderinger av LPFS er viktig, da foreløpige studier tyder på at LPFS er klinisk nyttig for å identifisere tilstedeværelse og grad av personlighetspatologi, prognose, behandlingsplanlegging, terapeutisk allianse, og å studere behandlingsforløp og utfall (APA, 2013, Bender, Morey & Skodol, 2011). Å kartlegge personlighetsfungering kan bidra til at behandlere i større grad kan tilpasse terapeutiske intervensjoner etter den individuelle pasientens behov (Bach, Markon, Simonsen & Krueger, 2015) og dermed skape en bedre terapeutisk prosess (Alvarez, 2010, Horowitz, 2013). Det er også vist at dimensjonale målinger kan fange opp personlighetspatologi bedre enn det nåværende kategoriske systemet (McCrae & Costa, 1997).

En god reliabilitet er helt nødvendig for brukbarheten til SCID-AMPD modul 1 (Cohen & Swerdlik, 2010, s. 151). Det er viktig at intervjuet kan brukes på konsistente og forutsigbare måter, slik at skårene ikke endrer seg med mindre det har vært en faktisk endring i det man måler (DeVellis, 2012, s.31). Dette vil være viktig for både forskning og klinisk praksis (Chmielewski, Clark, Bagby & Watson, 2015). Mer spesifikt vil vi undersøke intervjuets interrater reliabilitet, det vil si hvor stor grad av enighet eller konsistens det er i vurderingene til to intervjuere. Det er viktig at en pasient vil vurderes tilnærmet likt av to ulike klinikere, da dette kan påvirke diagnostisering, samt type og omfang av behandling. Andre studier som har sett på interrater reliabiliteten til LPFS på bakgrunn av andre vurderinger har vist lovende funn, men også noe varierende reliabilitet (Few et al., 2013, Hutsebaut et al., 2016b, Thylstrup et al., 2016, Zimmermann et al., 2014). Vårt studie er det første som undersøker interrater reliabiliteten til vurdering av LPFS ved SCID-AMPD-1. Hypotesen vår er at vi vil finne gode reliabilitetsestimater, med moderate til høye interrater reliabilitet på LPFS total og domenenivå, og moderate estimater på subdomenenivå.

Vi har valgt å studere interrater reliabiliteten ved et test-retest design. Dette innebærer at to klinikere gjennomfører to uavhengige intervjuer på samme pasient, og skårene fra de to intervjuene sammenlignes. Intervjuerne er dermed blinde for hverandres spørsmål og konklusjoner (Chmielewski et al., 2015). Tidligere studier har i motsetning til vårt i stor grad brukt et audio/videodesign når de har undersøkt interrater reliabilitet (Chmielewski et al., 2015, Few et al., 2013). Ved denne metoden gjennomfører en kliniker et intervju og rater pasienten, mens en eller flere andre "blinde" klinikere deretter gir et uavhengige skårer basert på opptakene av intervjuet. Alle studiene som hittil har sett på LPFS interrater reliabilitet har benyttet en type videodesign (Few et al., 2013, Hutsebaut et al., 2016b, Thylstrup et al., 2016, Zimmermann et al., 2014).

Imidlertid er denne metoden kritisert for å muligens gi kunstig høye reliabilitetsestimater. Klinikere vil for eksempel ha tilgang til akkurat den samme informasjonen, og den er begrenset til det intervjueren velger å spørre om. Enighet kan dermed oppstå uten at videorateren har mulighet til å innhente tilstrekkelig informasjon for en egen uavhengig vurdering. Metoden tar heller ikke hensyn til at to klinikere kan oppnå svært ulike responser ved to separate intervjuer (Chmielewski et al., 2015, Kraemer, Kupfer, Clarke, Narrow & Regier, 2012, Zimmermann, 1994).

DSM-5 prosjektet benyttet i starten et slikt videodesign. Ved de preliminare analysene fant de i tråd med argumentasjonen nevnt ovenfor, svært gode interrater reliabilitetsestimater. Dette ble satt i sammenheng med videometoden i kombinasjon med at SCID-AMPD modul 1 er såpass strukturert at det er lett å se til hvor intervjueren på videoen skårer pasienten. DSM-5 arbeidsgruppen bestemte seg i samråd med eksperter på området at designet burde endres til et test-retest design (Benjamin Hummelen, personlig kommunikasjon, november 2016). Det er foreslått at det å undersøke interraterreliabiliteten ved dette vil være et strengere og renere design (Few et al., 2013), og gi en mer realistisk vurdering av reliabiliteten (Chmielewski et al., 2015). Imidlertid vil man mest sannsynlig kunne forvente lavere reliabilitetsestimater ved test-retest enn et videodesign (Chmielewski et al., 2015, Thylstrup et al., 2016), noe studier har bekreftet (Chmielewski et al., 2015).

Ingen studier har foreløpig undersøkt LPFS interraterreliabiliteten ved test-retest metoden, og kun to tidligere studier har undersøkt interraterreliabiliteten til et intervju for å kartlegge nivå av personlighetsfungering ved LPFS. Dette er da det første

studiet som undersøger SCID-AMPD modul 1 intervjuet, og det første studiet som undersøger LPFS interraterreliabilitet ved et test-retest design.

5 Metode

5.1 Formål og design

Dette studiet er en del av et større forskningsprosjekt, DSM-5 prosjektet, som undersøker reliabiliteten og validiteten til den alternative modellen for personlighetsforstyrrelser i DSM-5, seksjon III. Forskningsprosjektet er ledet av overlege og forsker Benjamin Hummelen ved seksjon for personlighetspsykiatri, Ullevål.

Formålet med dette delstudiet er å undersøke interrater reliabilitet (IRR) til måleinstrumentet basert på Den alternative modellen for personlighetsforstyrrelser DSM-5, SCID-AMPD modul 1 (Bender, Skodol, First & Oldham, 2014). For å gjøre dette benyttet vi oss av et test-retest design. SCID-AMPD modul I ble utført to ganger av to ulike intervjuere på samme deltager, med et intervall på omtrent en uke (maks to uker). Vi forventet at IRR ville være god (mellom .75 og .85), men ikke utmerket.

Målet var å rekruttere mellom 30 og 50 deltagere. Vi ønsket å rekruttere både pasienter med personlighetsforstyrrelser og med symptomidelser. Dette for å få deltagere på alle nivå av personlighetsfungering og dermed god spredning i materialet.

5.2 Deltagere

5.2.1 Utvalg

Utvalget som benyttes i denne hovedoppgaven består av 33 deltagere med fullverdige nivå av personlighetsfungering skårer fra to intervjuetidspunkt. Deltagerne ble hovedsaklig rekruttert fra gruppeenheter i Dagbehandlingsnettverk Østlandet. Dagbehandlingsnettverk er et klinisk forskingssamarbeid mellom behandlingsenheter som driver behandlingsprogram for pasienter med personlighetsforstyrrelser. Dagbehandlingsnettverk har som formål å opprettholde forskningsaktiviteter på personlighetsforstyrrelser og bidra med faglig oppdatering slik at behandlingsenhetene kan tilby gode behandlingsprogram. Behandlingsenhetene har felles praktiske rutiner, standardiserte datasett for innsamling av klinisk informasjon, samt et egenutviklet datasystem hvor pasientdata blir lagret (DN2000). Per dags dato er 21 behandlingsenheter i Sør-Norge med i Dagbehandlingsnettverk (<http://www.dagbehandlingsnettverk.no>). Vi hentet deltagere fra gruppeenhetene ved

Øvre Romerike DPS (n = 4), Nedre Romerike DPS (n = 3), Lovisenberg DPS (n = 5) og Seksjon for personlighetspsykiatri (n = 19). Allmennpoliklinikken ved Nydalen DPS ble også inkludert (n=2), med tanke på å rekruttere pasienter med symptomidelser. Allmennpoliklinikken ved Nydalen DPS er ikke en del av Dagbehandlingsnettverk.

5.2.2 Demografi

Utvalget på 33 deltagere bestod av 24 kvinner og 9 menn. Alder på deltagere var mellom 20 og 55 år, gjennomsnittlig alder var 29 år (SD=8,54). Antall symptomidelsesdiagnoser per deltager var mellom 0 og 4, gjennomsnittet var 1,52 (SD=1,21). De vanligste symptomidelsesdiagnosene var tilbakevendende depressiv lidelse (10 tilfeller), sosial fobi (6 tilfeller), dystymi (5 tilfeller), generalisert angstlidelse (5 tilfeller), panikklidelse (4 tilfeller) og PTSD (4 tilfeller) (Se tabell i vedlegg B). Antall personlighetsforstyrrelsesdiagnoser per deltager var mellom 0 og 3, gjennomsnittet var 0,91 (SD=0,51). De vanligste personlighetsforstyrrelsesdiagnosene var borderline personlighetsforstyrrelse (13 tilfeller), blandet personlighetsforstyrrelse (8 tilfeller), og unnavikende (6 tilfeller) personlighetsforstyrrelse (Se tabell i vedlegg C). Fem deltagere var ikke diagnostisert med en personlighetsforstyrrelse.

5.3 Prosedyrer

5.3.1 Rekruttering

Datainnsamlingen ble utført i tidsperioden januar 2016 til juli 2016, etter godkjenning av REK (prosjektnummer 2014/1696/REK sør-øst A) i november 2015. Terapeuter ved gruppeenhetene ved Øvre og Nedre Romerike DPS (Ahus Universitetssykehus), gruppeenheten ved Lovisenberg DPS, gruppeenheten ved seksjon for personlighetspsykiatri (Oslo universitetssykehus) og allmennpoliklinikken ved Nydalen DPS, rekrutterte pasienter som deltagere i studien. Inklusjonskriteriene innebar at pasientene var blitt utredet med SCID-II de siste fem ukene og GAF de siste tre ukene, at pasientdata var lagret eller kunne bli lagret i Dagbehandlingsnettverket og at deltager var villig til å delta på to intervjuer. Det var å foretrekke at deltagerne var pasienter under utredning, men også pasienter som var under behandling ble inkludert. Eksklusjonskriteriene var psykoselidelse, senvirkninger etter hjerneskade, vedvarende utviklingsforstyrrelse (for eksempel: autismespektrum lidelse), psykisk

utviklingshemming, pågående substans avhengighet og manglende forståelse av norsk språk.

Terapeutene gav pasientene noe muntlig informasjon om prosjektet, samt en liten skriftlig brosjyre. Pasientene gav tillatelse til å bli kontaktet av en prosjektmedarbeider for å få mer informasjon om studien. Heretter kunne vedkommende bestemme seg for å delta. Hvis vedkommende ønsket å delta avtalte prosjektmedarbeider og pasient to intervjuetidspunkter. Pasienten fikk samtykkeerklæring og et selvutfyllingsskjema utdelt fra sin henvisende instans. Noen ganger ble disse delt ut ved første intervju.

5.3.2 Datainnsamling

Hver deltager ble intervjuet to ganger av to ulike intervjuere. Hvem som intervjuet samme deltager roterte mellom de seks intervjuerne. Hvem som intervjuet først og sist alternerte også (se tabell i vedlegg D). Forut for intervjuet fikk rater kun vite navnet til deltager og behandlingssted. Rater fikk ingen annen informasjon om deltager. Dette for å unngå forventningsbias (Rosenthal, 1976). De to raterne som intervjuet samme deltager ble oppfordret til å ikke prate om deltager mellom intervjuene. I mars 2016 mottok forskningsprosjektet frie forskningsmidler fra avdeling for forskning og utvikling ved klinikk psykisk helse og avhengighet, Oslo Universitetssykehus. Vi hadde dermed mulighet til å tilby deltager et gavekort på 500 kroner etter begge intervjuene var gjennomført

Etter første intervju gav rater tilbakemelding til henvisende terapeut. Tilbakemeldingen kunne legges inn i journal, og benyttes i utredning og behandling. Terapeut ble bedt om å ikke gi deltager tilbakemeldingen før etter begge intervjuene var utført. Etter terapeut hadde fått tilbakemeldingen ble han/hun bedt om å fylle ut et skjema om den klinisk nytteverdien av tilbakemeldingen.

Alle innsamlede data ble av-identifisert ved at hver deltager ble tildelt en identifikasjonskode som benyttes gjennom hele prosjektet. Forskningsdata er registrert og lagret i Dagbehandlingsnettverkets database (DN2000). Ferdig utfylte samtykkeerklæringer, selvutfyllingsskjema, SCID-AMPD-I og terapeutskjema oppbevares i låste arkivskap på Seksjon for Personlighetspsykiatri.

5.4 Intervjuere

Syv intervjuere deltok på prosjektet. Dette var: Tore Buer Christensen, overlege ved Sørlandet sykehus; Kjetil Bremer og Kristoffer Walther, psykologspesialister ved Seksjon for Personlighetspsykiatri; Kjerstin Holde, 5. års medisinstudent ved UiO; Tuva Langjord, Karoline Koritzinsky og Marianne Arnesen, 6. års psykologstudenter ved UiO. Siden intervjuerne besto av både klinikere og studenter, var det stor spredning i både alder (R=25-53 år, M=36 år) og antall års klinisk erfaring (R=0-20 år, M= ca. 5 år). Studentene hadde noe erfaring med bruk av utredningsverktøy (MINI Nevropsykiatrisk Intervju, SCID-II) fra praksis. To av psykologstudentene hadde hatt et halvt års hovedpraksis på Personlighetspoliklinikken.

5.4.1 Opplæring

Klinikkerne fikk opplæring i bruk av SCID-AMPD Modul 1 av Donna Bender som har vært med å utvikle måleinstrumentet. Kurset gikk over to dager og bestod av både teori og praktiske øvelser. Forelesningene var en innføring i nivå av personlighetsfungeringsskalaen og SCID-AMPD Modul 1. De praktiske øvelsene var skåring og diskusjon av ni pasientkasuistikker hentet fra DSM-IV-TR Casebook og original eksempler skrevet av Donna Bender, skåring og diskusjon av intervjuer tatt opp på video, og skåring og diskusjon av intervjuer gjort på stedet.

Studentene fikk opplæring i bruk av SCID-AMPD-I av Kjetil Bremer og Benjamin Hummelen. Kurset ble gjennomført over fire ettermiddager i november og desember 2015. Dette fulgte samme mal som kurset klinikerne var på, men med noen tilpasninger. I tillegg til å skåre Donna Benders pasientkasuistikker skåret studentene kasus vignetter skrevet av Tore Buer Christensen, Tor Erik Nysæter (psykologspesialist), Kristoffer Walter og Kjetil Bremer. Studentene fikk observere norske videointervjuer foretatt av Kjetil Bremer og Kristoffer Walter. Det var dessverre ikke mulighet for å intervju pasienter på kurset, men kursholderne regnet med at skåring av ekstra pasientkasuistikker veide opp for dette.

5.5 Måleinstrument

5.5.1 Strukturert klinisk intervju for den alternative (seksjon III) modellen for personlighetsforstyrrelser i DSM-5 (SCID-AMPD/SCID-III)

Strukturert klinisk intervju for den alternative (seksjon III) modellen for personlighetsforstyrrelser i DSM-5 (SCID-AMPD) Modul 1: Skala for nivå av personlighetsfungering (Bender et al., 2014), er et strukturert klinisk intervju basert på kriterium A: Nivå av personlighetsfungering i Den alternative (seksjon III) modellen for personlighetsforstyrrelser i DSM-5 (Bender, Skodol, First & Oldham, 2014). Dette vil videre i oppgaven betegnes som SCID-AMPD modul 1. Intervjuet er utviklet av Donna Bender, Andrew E. Skodol, Michael First og John M. Oldham. Norsk oversettelse er foretatt av Espen Folmo og Tuva Langjord i samarbeid med Forskningsgruppen for personlighetspsykiatri, Oslo Universitetssykehus. Intervjuet foreligger per dags dato som et utkast, det er ikke publisert enda. Det finnes dermed ikke noen brukermanual til intervjuet. Det er heller ikke publisert noe om hvordan intervjuet er utviklet (Bender & Skodol, personlig kommunikasjon, september 2016). I en artikkel skrevet av Andrew Skodol i 2014 (Skodol, 2014), nevnes det at han er medforfatter på SCID-AMPD.

Intervjuet vurderer de fire elementene, kalt domener i SCID-AMPD Modul 1, som utgjør kriterium A: Nivå av personlighetsfungering i den alternative modellen for personlighetsforstyrrelser. Vurderingen gjøres på en fem-punkts skala hvor 0 er lite eller ingen svekkelse, 1 er mild svekkelse, 2 er moderat svekkelse, 3 er alvorlig svekkelse og 4 er ekstrem svekkelse. Intervjuet skal brukes til å vurdere pågående problematikk (Bender et al., 2014).

Oppbygning (se vedlegg E)

For å standardisere mengde og type forkunnskap man har før intervjuet, har Donna Bender laget en innledende del med spørsmål om demografi, utdanning, arbeidsbakgrunn, og spørsmål om pågående og tidligere episoder med psykopatologi. Det er ikke nødvendig å benytte seg av denne delen, hvis informasjonen er hentet inn ved en tidligere anledning. Selvet intervjuet starter med noen generelle spørsmål som skal kartlegge den intervjuedes syn på seg selv, grunnleggende tilnærming til livet og kvaliteten på vedkommendes mellommenneskelige relasjoner (Bender et al, 2014).

Hoveddelen av intervjuet består av fire deler: identitetsdomenet, domenet for målrettethet, empatidomenet og nærhetsdomenet. Under hvert av domenene er det tre subdomener. Dette blir 12 subdomener til sammen (se figur i vedlegg F). Hvert subdomene starter med spesifikke screenings spørsmål, fulgt av nivåbestemmende spørsmål. Det er mellom ett og seks spørsmål per nivå. For hvert funksjonsnivå er det en liten beskrivelse av typiske psykologiske trekk ved dette nivået. Dette er de samme beskrivelsene som er i Nivå av personlightefungeringsskalaen (LPFS) i DSM-5 manualen (APA, 2013). Beskrivelsene skal være til hjelp for intervjuer ved skåring av intervjuobjektet.

Del 1 vurderer identitet. Ifølge nivå av personlighetsfungering i DSM-5 Manualen er identitet å oppleve seg selv som unik, ha klare grenser mellom selv og andre, ha stabil selvtillit og nøyaktig selv vurdering, samt evne til å regulere en rekke emosjoner (APA, 2013, s. 762). Subdomenene under identitetsdomenet i SCID-AMPD Modul 1, er selvopplevelse, selv vaktelse, og emosjonell bredde og regulering. Eksempler på nivåbestemmende spørsmål under disse subdomenene er: "Er du som regel klar på hvem du er og hva ditt perspektiv er?" (Selvopplevelse, nivå 1), "Har du det dårlig med deg selv dersom du ikke får anerkjennelse fra andre?" (Selv vaktelse, nivå 2), "Er du vanligvis svært humørsyk og har mange følelsmessige opp- og nedturer?" (Emosjonell bredde og regulering, nivå 3) (Bender et al, 2014).

Del 2 vurderer målrettethet. Målrettethet i SCID-AMPD Modul 1, består av subdomenene evne til å forfølge meningsfylte mål, konstruktive, prososiale indre normer for atferd og selvreflektiv fungering. Eksempler på nivåbestemmende spørsmål under disse subdomenene er: "Blir du vanligvis veldig forvirret og overveldet når du tenker på å skulle prøve å oppnå ting i livet?" (Evne til å forfølge meningsfylte mål, nivå 3), "Føler du nesten alltid at du ikke aner hvordan du bør opptre eller hva som forventes av deg?" (Konstruktive, prososiale indre normer for atferd, nivå 4), "Er det ofte vanskelig for deg å finne ut av ditt eget sinn, og hva du tenker om ting?" (Selvreflektiv fungering, nivå 2) (Bender et al, 2014).

Del 3 vurderer empati. Empati i SCID-AMPD Modul 1, består av subdomenene å forstå og sette pris på andres erfaringer og motivasjoner, toleranse for annerledes perspektiver, og forståelse av egen atferds effekt på andre. Eksempler på nivåbestemmende spørsmål under disse subdomenene er: "Går du til ytterligheter for å forstå andre, slik at du kan fortsette å bli likt av dem?" (Å forstå og sette pris på andres

erfaringer og motivasjoner, nivå 2), "Er du generelt i stand til å forstå og sette pris på andre menneskers synspunkter når de ikke er enig med deg?" (Toleranse for annerledes perspektiver, nivå 0), "Havner du ofte i ubehagelige eller forvirrende situasjoner med andre mennesker?" (Forståelse av egen atferds effekt på andre, nivå 3) (Bender et al, 2014).

Del 4 vurderer nærhet. Nærhetsdomenet i SCID-AMPD Modul 1, består av subdomenene dybde og varighet av relasjoner, ønske om og evne til nærhet og gjensidig respekt reflektert i atferd. Eksempler på nivåbestemmende spørsmål under disse subdomenene er: "Føler du det slik at du må sette opp en fasade, eller at du bare spiller en "rolle" i dine forhold til andre eller i sosiale sammenhenger?" (Dybde og varighet av relasjoner, nivå 2), "Holder du deg mest for deg selv, fordi det alltid ender med at andre sårer eller utnytter deg?" (Ønske om og evne til nærhet, nivå 4), "Gjør du hovedsakelig ting for eller med andre bare dersom det gagnar deg på noen måte?" (Gjensidig respekt reflektert i atferd, nivå 2) (Bender et al, 2014).

Bruk

Intervjuer starter med å stille generelle spørsmålene om hvordan intervjuobjektet er som person, for å få en grunnleggende oppfatning av den intervjuedes syn på seg selv og andre. Basert på denne informasjonen og klinisk skjønn setter så intervjuer en innledende global skåre fra 0 til 4.

På hvert subdomene starter intervjuer med å stille de spesifikke screeningsspørsmålene. Ut fra svaret på screeningsspørsmålene gjør intervjuer en vurdering av hvilket nivå han/hun tror personen ligger på. Intervjuer begynner å stille nivåbestemmende spørsmål på dette funksjonsnivået. Hvis intervjuobjektet bekrefter spørsmålene fortsetter intervjuer å stille spørsmål på økende nivå av svekkelse, helt til den intervjuede åpenbart ikke oppfyller kravene for svekkelsesnivået. Hvis den intervjuede avkrefter spørsmålene på dette funksjonsnivået, stiller intervjuer spørsmål på minkende nivå av svekkelse. Intervjuer stiller så mange av de nivåbestemmende spørsmålene som trengs for å avgjøre fungering. For hvert spørsmål skal intervjuer be om eksempler og utdypninger slik at han/hun får nok informasjon til å kunne gjøre en vurdering. Negative svar bør ikke sees på som tilstrekkelig bevis for at nivået ikke stemmer. Personens atferd under intervjuet bør også tas med i betraktning når man stiller tilleggsspørsmål. I dette prosjektet har vi skåret svarene på de nivåbestemmende spørsmålene slik som på SCID-II: 1 er nei, 2 er litt, 3 er ja og 9 er ikke nok informasjon.

Intervjuer kommer frem til en skåre fra 0 til 4 på hvert subdomene etter hvilket nivå han/hun vurderer at intervjuobjektet befinner seg på. For å finne skåren på domenenivå, regner intervjuer gjennomsnittet av de tre subdomene skårene. Total skåren for intervjuet (LPFS total) finner intervjuer ved å regne gjennomsnittet av skårene på de fire domene. Intervjuer sitter altså igjen med fire skårer: LPFS Total, Identitet, Målrettethet, Empati og Nærhet. Mot slutten av intervjuet noterer intervjuer ned tid brukt på intervjuet og avslutter med å spørre pasienten om hvordan vedkommende opplevde intervjuet.

5.6 Statistiske analyser

Skårene på SCID-AMPD Modul 1 intervjuene ble plottet inn i en excel fil. Statistiske analyser ble så utført i SPSS (IBM SPSS Statistics versjon 22).

5.6.1 Preliminære analyser og beskrivende statistikk

Først undersøkte vi om datamaterialet vårt var normalfordelt eller ikke. Vi undersøkte dette for LPFS total for intervju 1 (n=33), intervju 2 (n=33) og gjennomsnittet av intervju 1+2 (n=33). Dette gjorde vi ved hjelp av deskriptive mål for skjevhet og kurtose, samt laget histogram. Skjevhet gir en indikasjon på symmetrien til populasjonen, kurtose gir informasjon om brattheten (Pallant, 2010). Ved hjelp av histogrammet undersøkte vi for ekstrem skårer. Deretter fant vi gjennomsnitt, standardavvik, og standardfeil for LPFS total på intervju 1 (n=33), intervju 2 og gjennomsnittet av intervju 1+2 (n=33).

5.6.2 Intern konsistens

For å undersøke reliabiliteten til SCID-AMPD modul 1, begynte vi med å se på intern konsistens. Intern konsistens referer til homogeniteten til enheter på en skala. Høy intern konsistens indikerer at enhetene måler den samme underliggende variabelen (DeVellis, 2012). Vi undersøkte intern konsistens av SCID-AMPD modul 1 ved bruk av Cronbachs alpaha (α) koeffisienten. Ideelt sett skal Cronbachs alpaha (α) koeffisienten til en skala være over .7 (DeVellis, 2012). Vi vurderte intern konsistens for hele skalaen, for komponentene selvfungering og interpersonligfungering, for domene **og**

subdomenene. Dette gjorde vi for skårene fra intervju 1 (n=33) for seg, skårene fra intervju 2 (n=33) for seg og gjennomsnitt av skårene fra intervju 1 og 2 (n=33).

5.6.3 Endring i skårer: t-test

For å undersøke om det er en statistisk signifikant forskjell i skårer fra intervju 1 til intervju 2 utførte vi en paired sample t-test. En paired sample t-test brukes når man har et enkelt utvalg av deltagere og innhenter data fra dem på to ulike tidspunkt (Howitt & Cramer, 2011, s. 125). Slik er det i vårt studie. Vi foretok en t-test for skårene på LPFS total, komponent, domene og subdomene nivå. Signifikans nivå ble satt til .05.

5.6.4 Interrater reliabilitet

Hovedanalysene våre bestod av to typer korrelasjonsanalyser for å finne ut av om det var en sammenheng mellom skårene på intervju 1 og intervju 2. Vi valgte å benytte oss av Pearson product-moment korrelasjon. Pearson product-moment korrelasjon blir ofte brukt når det er snakk om et test-retest design (DeVellis, 2012). Vi kalkulerte Pearson product-moment korrelasjon for sammenhengen mellom intervju 1 og intervju 2 for den totale LPFS skåren, for selvfungering, for interpersonlig fungering, for de fire domene og for de tolv subdomene. Statistisk signifikans ble satt til 0.01 nivået. Vi laget også et scatterplot for å se på forholdet mellom intervju 1 og intervju 2, og se etter «uteliggere». Scatterplottet ble laget for forholdet mellom skåren LPFS total på intervju 1 (n=33) og LPFS total på intervju 2 (n=33).

Vi vurderte også interrater reliabilitet ved bruk av en annen analyse: intraclass korrelasjoner (ICC). Intraclass korrelasjoner er et nyttig estimat på interrater reliabilitet, hvis man har flere enn to intervjuere eller hvis intervjuerne varierer etter intervjuobjekt (Landers, 2015). Vi utførte en two-way random ICC analyse. Vi valgte two-way random fordi denne analysen regner med at både intervjuer og intervjuobjekt har en effekt på skårene (Landers, 2015). Vi var interessert i både reliabiliteten til en enkelt intervjuer, og reliabiliteten til gjennomsnittet av de syv intervjuerne. Derfor utførte vi både en two-way random single measures, og en two-way random average measures. Disse analysene utførte vi for sammenhengen mellom intervju 1 (n=33) og intervju 2 (n=33) for skåren LPFS total, for selvfungering, interpersonlig fungering, domene og subdomene.

5.6.5 Tilleggsanalyser

Det ble utført en tilleggsanalyse, hvor vi så på forholdet mellom intervjuernes grad av erfaring og absolutt forskjell mellom skåringene på intervju 1 og intervju 2. Først beregnet vi absoluttverdi for forskjellen i skårer mellom intervju 1 og intervju 2 på alle intervjuoparene. Deretter kodet vi intervjuoparene etter om det var to eksperter som hadde utført intervjuene (2) (n=4), en ekspert og en uerfaren som hadde utført intervjuene (1)(n=12), og to uerfarne som hadde utført intervjuene (0)(n=16). Pearson product-moment korrelasjonskoeffisient ble beregnet for dette forholdet. Det manglet informasjon om hvem som hadde intervjuet deltager for intervjuene på en deltager (n=32).

6 Resultater

6.1 Oppsummering resultater

Interrater reliabilitetsanalysene for SCID-AMPD-1 viser høye reliabilitetsestimater på domene, komponent og LPFS totalnivå. For subdomenene varierer interrater reliabiliteten mellom lave og høye estimater.

6.2 Preliminære dataanalyser

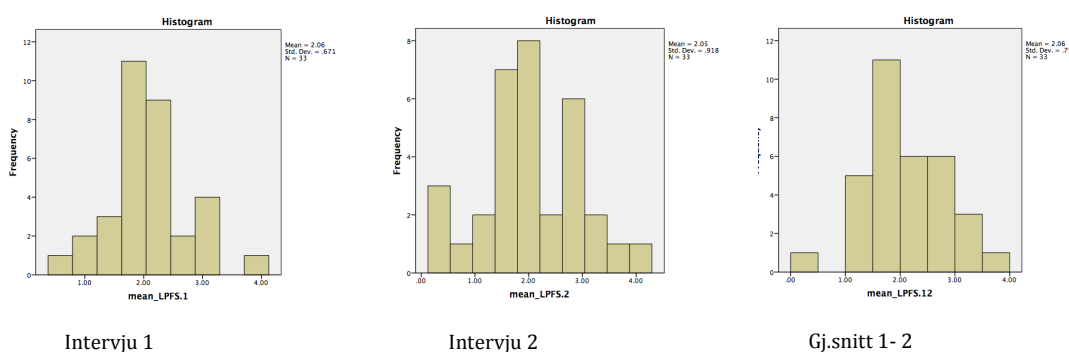
6.2.1 Manglende data

Ingen deltakere droppet ut fra første til andre intervju. Det var heller ingen manglende data i materialet når vi gjennomførte analysene.

6.2.2 Normalitet, linearitet og homoscedasticity

Preliminære analyser ble utført for å sikre at det ikke var noen brudd på antagelser om normalitet, linearitet og homoscedasticity. Datamaterialet ble undersøkt for hvorvidt variablene var normalfordelte eller ikke ved histogrammer og deskriptive mål for skjevhet og kurtose. Figur 1 viser fordelingen av totalskåreer på LPFS for de første og andre intervjuene, samt for skåringen av begge intervjuene slått sammen (Gj.snitt 1-2). Skåringene for intervjuene slått sammen er beregnet ut i fra gjennomsnittet av subdomeneneskåringene på de to intervjuene. Skårene for LPFS total viste en tilnærmet normalfordeling, med en tydeligere normalfordelingskurve for gjennomsnittet av de første og andre intervjuene.

Figur 1: Fordeling av LPFS totalskårene ved de første og de andre intervjuene, samt gjennomsnittet av skåringene på begge intervjuene



6.2.3 Ekstremskårer

Preliminære analyser viste fire deltagere som tydelig skilte seg ut i diskrepansen mellom skårene for intervju 1 og 2 og ble definert som ekstremskårer. Intervjuene og plottingen av data ble gjennomgått for å sikre at det ikke var registrert noen feil. En av deltakerne ble ekskludert fra datamaterialet, da deltageren oppfylte en eksklusjonsdiagnose. De tre resterende ble beholdt i materialet, da vi ikke fant noen årsaker som kunne forklare diskrepansen annet enn variabler knyttet til responsstil og spørsmål som var stilt i intervjuet. Når vi leste gjennom disse intervjuprotokollene, fant vi spesielt to tendenser som kan ha påvirket skåringene: ulike svar og ordlegging fra pasientene, og at det var stilt spørsmål på ulike nivåer.

6.3 Beskrivende statistikk for utvalget

Beskrivende statistikk for gjennomsnittlig totalskåre på LPFS for de første og de andre intervjuene, samt gjennomsnittsskåre for begge intervjuene slått sammen vises i Tabell 1. 33 deltagere ble inkludert i alle analysene våre. Gjennomsnittlig totalskåre for LPFS er over terskelen for å sette en personlighetsforstyrrelsesdiagnose (APA, 2013). Gjennomsnittlig skåre for intervju 1, intervju 2 og gj.snitt intervju 1-2 er tilnærmet lik, med noe mer variasjonsbredde for intervju 2.

Tabell 1: Beskrivende statistikk for utvalget

	N	R	M	SD	SE
Intervju 1	33	3.33	2.06	.67	.12
Intervju 2	33	3.67	2.05	.92	.16
Gj.snitt. 1-2*	33	3.50	2.06	.75	.13

Note: *Gj.snitt 1-2: beregnet ut i fra gjennomsnittet av subdomeneskåringene på intervju 1 og 2.

6.4 Intern konsistens

6.4.1 Cronbachs alpha

Intervjuets indre konsistens som et mål på reliabilitet ble undersøkt ved Cronbachs alpha (α) (Cronbach, 1951). Alphaverdier ble beregnet for de første og andre intervjuene, samt for begge intervjuene slått sammen beregnet ut i fra gjennomsnittet av skårene på subdomenene. Alphakoeffisientene er vist i Tabell 2.

Tabell 2: SCID-AMPD-1 indre konsistens beregnet ved Cronbachs alpha basert på vårt utvalg

	Ant. subdomener	Intervju 1 N = 33	Intervju 2 N = 33	Gj.snitt 1-2 N = 33
Identitet	3	.77	.82	.86
Målnettethet	3	.74	.82	.84
Empati	3	.85	.91	.91
Nærhet	3	.77	.82	.87
Selvet	6	.87	.91	.92
Interpersonlig	6	.88	.92	.93
LPFS total	12	.89	.95	.95

Note. Cronbachs alpha beregnet ved reliabilitetsanalyse for domener (identitet, målnettethet, empati og nærhet), komponenter (selvet og interpersonlig) og LPFS totalt.

*Gj.snitt 1-2: beregnet ut i fra gjennomsnittet av subdomeneskåringene på intervju 1 og 2

Alphakoeffisientene varierer fra .74 (målnettethet, intervju 1) til .95 (LPFS total, intervju 2 og gj.snitt intervju 1-2). Dette blir vurdert som respektable til veldig gode alphaverdier (DeVellis, 2012). Ingen alphaverdier avviker betydelig i noen retning. Dataene viser konsekvent høyere alphaverdier for de andre intervjuene sammenlignet med de første. Imidlertid er forskjellen liten.

Alphaverdiene øker med antall ledd som ble inkludert i analysene. De laveste alphaverdiene finner vi dermed på domenenivå, og den høyest på LPFS total, som er beregnet ut i fra henholdsvis 3 og 12 subdomener.

Ved nærmere gjennomgang av analysene, var det noen subdomener som viste seg å svekke den interne konsistensen. Alphaverdien for intervju 1 LPFS total ville økt dersom man fjernet evnen til å forfølge meningsfylte mål. Intervju 1 identitet ville økt uten selvpålevelse, og intervju 1 nærhet ville økt uten gjensidig respekt reflektert i atferd. For gjennomsnittet av intervju 1 og 2, ville alphaen for målnettethet økt uten evne til å forfølge meningsfylte mål, noe som dermed også ville økt alphaverdien for selvet og LPFS total. Oppsummert visert dataene våre god intern konsistens på både domene, komponent og LPFS totalnivå. Subdomenene evne til å forfølge meningsfylte

mål, selvopplevelse og gjensidig respekt reflektert i atferd påvirker intervjuets interne konsistens negativt.

6.5 Endring i skårer mellom testene

6.5.1 T-tester

En paired sample t-test (tohalet) ble gjennomført for å evaluere sammenhengen mellom skåringene på de første og de andre intervjuene. T-test for domenene, komponentene og LPFS total er vist i tabell 3. T-test for subdomenene er vist i tabell 4. En høy t-verdi gir lave p-verdier. Det var ingen t-tester som viste en statistisk signifikant endring på gjennomsnittsskåringen fra intervju 1 til intervju 2. P verdiene varierte mellom $p = .19$ (emosjonell bredde og regulering) og $p = 1.00$ (selvreflektiv fungering). Der hvor t-testen nærmer seg p-verdien på .005, er subdomenene: ønske om og evne til nærhet ($p = .16$), emosjonell bredde og regulering ($p = .19$) og forståelse av egen atferd effekt på andre ($p = .12$). Subdomenenet selvrefleksjon viste en signifikansverdi på $p = 1.00$, som vil si at det var null gjennomsnittlig endring i skårer mellom intervju 1 og intervju 2. Standardavviket for de to intervjuene er ulikt, noe som vil si at skåringene kommer fra to ulike datasett og ikke skyldes en feilregistrering.

T-testen viser heller ingen systematisk forskjell i endring i skårene fra det første til det andre intervjuene. En foreløpig konklusjon er dermed at retest fenomenet (Arrindell, 2000) ikke ser ut til å ha påvirket dataene.

Tabell 3: Gjennomsnittlig endring i skårer fra intervju 1 til intervju 2 for domener, komponenter og LPFS total beregnet ved two-tailed t-test

	Antall subdomener	Intervju 1			Intervju 2			<i>t</i>	<i>p</i>
		M	SD	R	M	SD	R		
Identitet	3	2.56	.66	3.00	2.47	.89	3.33	.65	.52
Målrettethet	3	2.20	.79	3.33	2.19	.98	3.67	.10	.92
Empati	3	1.43	.97	4.00	1.52	1.12	4.00	-.63	.53
Nærhet	3	2.06	.88	4.00	2.00	1.03	4.00	.42	.68
Selvet	6	2.38	.69	3.17	2.33	.90	3.50	.45	.66
Interpersonlig	6	1.75	.86	4.00	1.76	1.02	4.00	-.12	.91
LPFS total	12	2.06	.67	3.33	2.05	.92	3.67	.15	.88

Tabell 4: Gjennomsnittlig endring i skårer fra intervju 1 til intervju 2 for subdomenene beregnet ved two tailed t-test

	Intervju 1			Intervju 2			t	p
	M	SD	R	M	SD	R		
Selv - opplevelse	2.27	.98	3.00	2.36	1.17	4.00	-.53	.60
Selvaktels e	2.73	.72	3.00	2.64	1.02	3.00	.50	.62
Emosjons - regulering ¹	2.67	.69	3.00	2.42	.902	3.00	1.61	.19
Meningsfylte mål ²	2.58	1.00	4.00	2.48	1.20	4.00	.62	.54
Indre normer ³	1.79	.99	4.00	1.85	1.15	4.00	-.36	.72
Selvrefleksjon	2.24	.90	4.00	2.24	1.06	3.00	.00	1.00
Forståelse andre ⁴	1.79	1.11	4.00	1.97	1.16	4.00	-.95	.35
Toleranse ⁵	1.36	1.17	4.00	1.18	1.16	4.00	1.10	.28
Forståelse egen atferd ⁶	1.15	1.03	4.00	1.42	1.35	4.00	-1.60	.12
Relasjoner ⁷	2.30	1.07	4.00	2.21	1.29	4.00	.49	.63
Evne nærhet ⁸	2.48	.97	4.00	2.24	1.15	4.00	1.44	.16
Respekt ⁹	1.39	1.14	4.00	1.55	1.17	4.00	-.61	.55

Note. Fullstendig beskrivelse av variablene: ¹ Emosjonell bredde og regulering ² Evne til å forfølge meningsfylte mål ³ Konstruktive prososiale indre normer for atferd ⁴ Å forstå og sette pris på andres erfaringer og motivasjoner ⁵ Toleranse for annerledes perspektiver ⁶ Forståelsen av egen atferds effekt på andre ⁷ Dybde og varighet av relasjoner ⁸ Ønske om og evne til nærhet ⁹ Gjensidig respekt reflektert i atferd

6.6 Enighet i skåringer mellom intervju 1 og intervju 2

6.6.1 Pearson r

Forholdet mellom det første og det andre SCID-AMPD modul 1 intervjuet ble først undersøkt ved bruk av Pearson produkt-moment korrelasjonskoeffisient (Pearson *r*). Pearson *r* for domenene, komponentene og LPFS total er vist i Tabell 5. Pearson *r* for subdomenene er vist i Tabell 6. Sammenhengen mellom skåringene for intervju 1 og

intervju 2 for deltagerne er fremstilt i Figur 2. Pearson r korrelasjonen varierte mellom $r = .23$ ($p > .005$) og $r = .81$ ($p < .001$). Det var en sterk, positiv korrelasjon mellom de første og de andre intervjuene på LPFS totalskåringen, $r = (.78)$, $n = 33$, $p < .001$.

Alle subdomenene med unntak av tre var positivt og signifikante korrelert med hverandre på et 1 % nivå. Emosjonell bredde var signifikant på et 5 % nivå. Subdomenene som ikke var signifikant korrelerte var: selvaktelse, $r = (.33)$, og gjensidig respekt reflektert i atferd, $r = (.23)$. Korrelasjonene for LPFS domenenivå var $r = .61$ ($p < .001$) for identitet, til $r = .81$ ($p < .001$) for målrettethet. På komponentnivå fant vi omtrent tilsvarende estimater som for domeneene. I følge Cohen (1988, pp. 79 – 81) tilsvarer dette høye ($r = .5 - 1.0$) korrelasjonskoeffisienter på domene, komponent og LPFS total nivå. På subdomenenivå varierer estimatene fra lave ($r = .10 - .29$) til høye ($r = .5 - 1.0$) korrelasjonskoeffisienter.

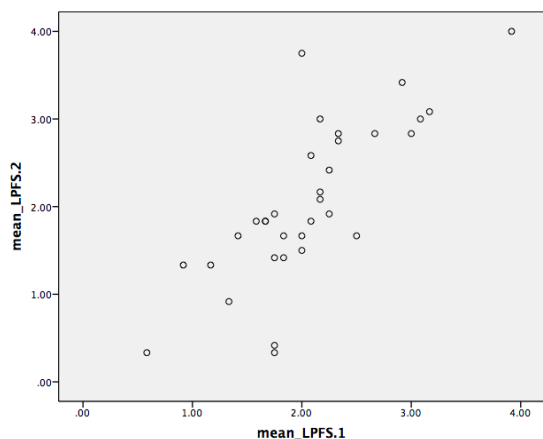
6.6.2 Intraclass korrelasjoner

For å videre undersøke interraterreliabiliten gjennomførte vi two-way random consistency intraclass korrelasjonsanalyser (ICC). Tabell 7 viser ICC for domeneene, komponentene og LPFS total. Tabell 8 viser ICC for subdomenene. Reliabiliteten til individuelle ratinger for LPFS total var ICC [2, 1] = .74, 95 % CI [.54, .86]. På domenenivå varierte reliabiliteten mellom ICC [2, 1] = .58, 95 % CI [.31, .78] for identitet, og ICC [2, 1] = .79, 95 % CI [.62, .89] for målrettethet.

Alle subdomenene med unntak av tre var signifikante og positivt assosiert med hverandre. Intraclass korrelasjonene for subdomenene varierte fra ICC [2, 1] = .53, 95 % CI [.23, .74] (å forstå og sette pris på andres erfaringer og motivasjoner) til ICC [2, 1] = .71, 95 % CI [.49, .85] (evne til å forfølge meningsfylte mål). De tre subdomenene som demonstrerte lave og ikke signifikante reliabilitetsestimater var: selvaktelse: ICC [2, 1] = .31, 95 % CI [-.03, .59], emosjonell bredde og regulering: ICC [2, 1] = .42, 95 % CI [.09, .17] og gjensidig respekt reflektert i atferd ICC [2, 1] = .23, 95 % CI [-.19, .53].

Alle korrelasjonene øker når vi ser på gjennomsnittlig reliabilitet for intervjuerne (ICC average). Målrettethet viser da spesielt høy ICC [2, 7] = .88, 95 % CI [.76, .94]. Intraclass korrelasjonen for LPFS total øker også for ICC average: ICC [2, 7] = .85, 95 % CI [.70, .93].

Figur 2: Sammenhengen mellom skåringer på intervju 1 og intervju 2 for alle deltagerne



Tabell 5: Pearson r og ICC two-way random consistency for domener, komponenter og LPFS total

	Antall subdomener	Pearson r N = 33	ICC single N = 33	ICC average N = 33
Identitet	3	.61**	.58**	.74**
Målrettethet	3	.81**	.79**	.88**
Empati	3	.70**	.69**	.82**
Nærhet	3	.63**	.62**	.77**
Selvet	6	.76**	.74**	.85**
Interpersonlig	6	.69**	.68**	.81**
LPFS total	12	.78**	.74**	.85**

Note. ** $p < .001$ (2-tailed)

Tabell 6: Pearson r og ICC two-way random consistency for subdomenene

	Pearson r N = 33	ICC single N = 33	ICC average N = 33
Selv - opplevelse	.60**	.59**	.74**
Selvaktelse	.33	.31	.47
Emosjons - regulering ¹	.43*	.42	.59
Meningsfylte mål ²	.72**	.71**	.83**
Indre normer ³	.60**	.60**	.75**
Selvreflektiv fungering	.72**	.71**	.83**
Forståelse andre ⁴	.53**	.53**	.69**
Toleranse ⁵	.67**	.67**	.80**

Forståelsen egen atferd ⁶	.69 **	.67 **	.80 **
Relasjoner ⁷	.60 **	.59 **	.75 **
Evne nærhet ⁸	.59 **	.58 **	.74 **
Respekt ⁹	.23	.23	.37

Note. ** $p < .001$ (2-tailed) * $p < .005$ (2-tailed)

Fullstendig beskrivelse av variablene: ¹ Emosjonell bredde og regulering ² Evne til å forfølge meningsfylte mål ³ Konstruktive prososiale indre normer for atferd ⁴ Å forstå og sette pris på andres erfaringer og motivasjoner ⁵ Toleranse for annerledes perspektiver ⁶ Forståelsen av egen atferds effekt på andre ⁷ Dybde og varighet av relasjoner ⁸ Ønske om og evne til nærhet ⁹ Gjensidig respekt reflektert i atferd

6.7 Tilleggsanalyse

6.7.1 Grad av erfaring

Sammenhengen mellom grad av erfaring hos intervjuerne og absolutt forskjell mellom skåringene fra intervju 1 og intervju 2 ble undersøkt ved Pearson r . Vi fant en negativ, ikke signifikant korrelasjon mellom grad av erfaring og forskjell i skåringene på de to intervjuene, $r = (-.31)$, $n = 33$, $p = .09$. Større grad av erfaring førte dermed til mindre diskrepans i skåringene mellom de to intervjuene.

7 Diskusjon

Målsetningen med studiet var å undersøke interrater reliabiliteten til SCID-AMPD modul 1 som et instrument for å kartlegge nivå av personlighetsfungering beskrevet i Den alternative modellen for personlighetsforstyrrelser (AMPD) i DSM-5 (APA, 2013). Vi valgte å benytte et test-retest design for å gjøre en grundig vurdering og ta hensyn til kritikk mot det mye brukte videodesignet (Chmielewski et al., 2015). Skalaen som benyttes for å beskrive Nivå av personlighetsfungering i DSM-5 manualen kalles Levels of Personality Functioning Scale (LPFS) (APA, 2013). SCID-AMPD modul 1 intervjuet bygger på manualen og beskrivelsene gitt i LPFS, men skiller seg også ut som et eget instrument. Når vi snakker om LPFS i diskusjonen, viser vi da til de nivåbestemte beskrivelsene som beskrevet i DSM-5 manualen. Når vi nevner SCID-AMPD modul 1, viser dette til selve intervjuet som er et instrument for å vurdere LPFS.

Hypotesen vår var at SCID-AMPD modul 1 ville vise moderate til høye reliabilitetsestimater for LPFS total, komponent og domenenivå. På subdomenenivå forventet vi moderate reliabilitetsestimater. Vi vil videre diskutere funnene våre med hovedfokus på reliabilitet, samt adressere begrensninger ved studiet, implikasjoner av funnene og muligheter for videre forskning.

7.1 Endring i skåringer mellom intervjuene

Vi fant ingen systematisk forskjell i endring av gjennomsnittsskåringene fra de første til de andre intervjuene. Man kunne imidlertid forventet det motsatte, da man ofte finner en retest effekt i test-retest studier (Arrindel, 2000, Jorm, Duncan-Jones & Scott, 1989, Ormel, Koeter & van den Brink, 1989). En retest effekt vil si at det er en tendens til at patologiskåren er lavere ved tidspunkt nummer to enn ved tidspunkt nummer en. I behandlingsstudier kan dette for eksempel føre til at man tenker det er intervensjonen som skaper bedring hos pasientene, mens det i virkeligheten kan spores tilbake til retest effekten som ville vært tilstede uavhengig av intervensjonen. Det er foreløpig noe usikkert hva som fører til disse systematiske endringene (Arrindel, 2000, Ormet et al., 1989), men det har vært diskutert om de blant annet kan skyldes en legitimeringsprosess der pasientene overdriver symptomene første gangen eller sosial ønskelighet, (Ormel et al., 1989). Det at vi ikke fant en retest effekt i vår studie,

indikerer høy kvalitet over intervjuet og antyder at intervjuet mest sannsynlig vil ha god reliabilitet.

7.2 Intervjuets interne konsistens

Resultatene våre gir også støtte for intervjuets interne konsistens, da vi fant Cronbachs alphaverdier som vurderes som akseptable til veldig gode. Disse funnene er spesielt bra siden Cronbachs alpha øker med antall ledd og vi finner gode alphaverdier også på domenenivå som består av kun tre subdomener. Imidlertid bør ikke alphaverdiene være for høye, da dette kan tyde på at de ulike elementene i modellen måler overlappende konstrukter og muligens er overflødige. Likevel ser vi i vår studie at alphaverdiene øker proporsjonalt med antall subdomener som er inkludert i analysene, slik at verdiene er lavest på domenenivå og høyest på LPFS total. Dette gjør det mindre problematisk at alphaverdiene ligger på over .90 på flere av komponentene (selv og interpersonlig), og LPFS total.

I noen studier ser man kun på Cronbachs alpha for å vurdere reliabiliteten til et instrument. Vi gikk imidlertid mer i dybden og undersøkte intervjuets interraterreliabilitet ved Pearson produkt moment korrelasjonskoeffisienten, og intraclass korrelasjoner for å kunne si mer omkring reliabiliteten. Videre i diskusjonsdelen vil vi i hovedsak rapportere ICC korrelasjoner, da dette er en anerkjent og mye brukt metode for å beregne interraterreliabilitet. Vi vil også fokusere på analysene for LPFS total, domene og subdomenenivå, da komponentene som vi har undersøkt ikke er med som en egen skåre i intervjuet. Når vi viser til ICC, viser vi til ICC single estimer, med mindre vi spesifiserer annet.

7.3 Enighet mellom intervjuerne

Resultatene viser at den totale LPFS skåren kan bli vurdert ved høy grad av enighet, både ved Pearson $r = .78$ og intraclass korrelasjon (ICC) = $.74$. Dette anses som en veldig god korrelasjon ut i fra retningslinjene brukt for test-retest interrater reliabilitet i DSM-5 feltforsøkene (Regier et al., 2013). Ut i fra disse tolkningene av reliabilitetsestimater, regner estimer under som $.20$ uakseptable, mellom $.20 - .39$ er tvilsomme, mellom $.40 - .59$ er gode, mellom $.60 - .79$ er veldig gode og mellom $.80-1$ er utmerket. Vi har valgt å vurdere våres resultater ut i fra dette. Andre har argumentert for en strengere

vurdering, og at test-retest reliabilitet under .60 vil være bekymringsverdig (Spitzer, Williams & Endicott, 2012). Imidlertid er våre interrater reliabilitetsestimater for LPFS totalskåren godt over hva begge disse retningslinjene aksepterer som en akseptabel reliabilitet.

Vår metode med test-retest design er krevende for å få like skåringer, og det forventes en lavere reliabilitet enn ved andre typer design (Chmielewski et al., 2015, kilde). Når man dermed i tillegg ser estimatene våre i sammenheng med at vi valgte et test-retest design tyder det på at resultatene er gode.

Man kan ikke forvente likhet mellom vårt og tidligere studier, da alle de foreløpige studiene på interrater reliabiliteten til LPFS har brukt videodesign, som skiller seg fra vårt test-retest design. Direkte sammenligning av interrater reliabilitets estimater fra andre studier på LPFS er dermed ikke mulig. Tidligere studier som har undersøkt interraterreliabiliteten vil likevel nevnes med tanke på hvilke domener som jevnt over relativt til andre domener har vist seg å være enklere eller vanskeligere å vurdere likt. Dette kan gi oss nyttig informasjon om hvilke konstrukter ved LPFS som fungerer godt. Andre studier som har benyttet ved skåring av videointervjuer for å undersøke enigheten mellom ratere har funnet interraterreliabilitet for LPFS totalskåringer mellom ICC = .51 (Zimmermann et al., 2014), og ICC = .71 (Hutsebaut et al., 2016b) i kliniske utvalg.

Når vi analyserte enigheten på domenenivå, fant vi at målrettethet viste best enighet med ICC = .81, etterfulgt av empati med ICC = .70, nærhet med ICC = .63 og identitet med ICC = .58. Komponentene viste estimater på ICC = .68 (interpersonlig), og ICC = .74 (selvet). Reliabiliteten til målrettethet kan dermed beskrives som utmerket (Regier et al., 2013). En mulig forklaring på den høye korrelasjonen for målrettethet kan være at konseptet målrettethet er mer knyttet til konkrete aspekter ved fungering enn de andre domenene. I SCID-AMPD-1 stilles det for eksempel i subdomenet evne til å forfølge meningsfylte mål, mange spørsmål rettet mot mål. Mest sannsynlig vil det da også være oppfølgingsspørsmål knyttet til konkret hva dette er. En kan tenke seg at det da da er lettere å både svare og skåre slike svar sammenlignet med konsepter som er mer abstrakte.

Thylstrup et al. (2016) fant også at domenet målrettethet hadde klart høyest enighet relativt til de andre domenene i studien deres. Resultater fra andre studier (Few et al., 2013, Hutsebaut et al., 2016b, Zimmermann et al., 2014) tyder også på at dette

domenet jevnt over synes å være stabilt relativt til de andre domenene. Dette kan indikere at konseptet målrettethet er enklere å vurdere likt enn de andre konseptene på domenenivå i AMPD, der reliabilitetsfunnene for disse domenene spriker mer på tvers av studier. Dette er i tråd med Thylstrup et al. (2016) sin argumentasjon om hvorfor de fant at målrettethet målt ved CALF ble vurdert mer likt enn andre konsepter.

Empati viste veldig god reliabilitet med ICC = .69. Nærhet har noe lavere enighet enn empati med fremdeles med veldig god ICC = .62. Her er det spesielt et subdomene: gjensidig respekt reflektert i atferd, som svekker domenet med en så lav reliabilitet som ICC = .23. Imidlertid viser de andre subdomenene akseptable reliabilitetsestimater, noe som gjør at domenet totalt har en akseptabel grad av enighet. Identitet er det domenet i vårt studie der det er lavest enighet mellom intervjuerne. Når vi ser på subdomenene på identitetsdomenet, finner vi to av de med lavest intraclass korrelasjoner på ICC = .33 for selvaktelse og ICC = .43 for emosjonell bredde og regulering. Emosjonell bredde og regulering var signifikant på et 5 % nivå ved Pearson r, men ikke ved ICC. Det gir dermed mening at identitetsdomenet har lave korrelasjoner. Til tross for disse svakhetene med subdomenene, fyller imidlertid domenet kriteriet for en god reliabilitet (Regier et al., 2013)

Det er en sammenheng mellom hvor vi finner de laveste og høyeste subdomenene og hvilke domener som har lavest og høyest reliabilitet. Det gir dermed mening at de to laveste domenene er de som har de tre lavest korrelerte subdomenene. Dermed ser man også hvilke som bør videreutvikles. De tre subdomenene som ikke var signifikante og hadde lav enighet mellom intervjuerne (selvaktelse, emosjonell bredde og regulering og gjensidig respekt reflektert i atferd) er problematiske dersom de skal brukes for å vurdere enkeltindivider i klinisk praksis. Det vil for eksempel kunne føre til at de kan få svært ulike vurderinger av alvorlighetsgrad avhengig av hvilken kliniker de møter.

Gjensidig respekt reflektert i atferd var også et av subdomenene som svekket intervjuets interne konsistens, noe som kan tyde på at det er et spesielt problematisk subdomene som bør revideres for å sikre en bedre reliabilitet. Test-retest designet vårt gjør også at svakhetene ved spørsmålene og subdomenene sannsynligvis blir tydeligere enn om vi hadde benyttet et design der ratere vurderte akkurat det samme intervjuet. Dette er samtidig viktig, da det vil tydeliggjøre hvilke deler av intervjuet som kan være sårbart slik det skal brukes i klinisk praksis.

At disse tre subdomenene er dårligere enn resten kan skyldes flere ting. For det første kan man kanskje si at gjensidig respekt reflektert i atferd er at av de subdomenene som synes å være preget av at vurderingen i stor grad er basert på hva som er "riktig" måte å oppføre. Her kan det muligens fremkomme en bias der hver enkelt intervjuers individuelle oppfatning av som er "riktig" eller "bra nok" atferd kan prege hvordan de skårer pasientene.

For det andre kan det være en tendens til at noen av spørsmålene på disse subdomenene gjør at pasientene får behov for å fremstå mer sosialt ønskelig enn det som er tilfelle, spesielt ved gjensidig respekt reflektert i atferd og selvaktelse. På subdomenet selvaktelse får deltagerne blant annet spørsmålet: "Tror du at du er bedre, smartere eller mer attraktiv enn de fleste andre?". De færreste vil nok svare umiddelbart ja på dette spørsmålet. Imidlertid blir intervjuerne oppfordret til å vurdere helheten og atferd gjennom intervjuet. Noen intervjuere kan dermed skåre personen ut i fra informasjonen som kommer frem, mens andre vektlegger atferd underveis i intervjuet, noe som kan skape forskjeller i vurderingene. Det er dermed et spørsmål om slike deler av personlighetsfungering egner seg for å vurderes med direkte spørsmål.

En tredje forklaring kan være at subdomenene selvaktelse og emosjonell bredde og regulering måler fenomener som muligens er vanskelige for pasientene å rapportere tydelig selv. Sammenlignet med for eksempel evne til å forfølge meningsfylte mål under målrettethet, er dette mer abstrakte å forholde seg til.

En fjerde forklaring kan være at det på nivå 2 og 3 ved emosjonell bredde og regulering, kun er to spørsmål. Dette gjør at intervjueren får begrenset mengde informasjon på disse nivåene, og mest sannsynlig trenger å stille en del oppfølgingsspørsmål for å kunne skåre pasienten. På denne måten blir vurderingen her mindre strukturert enn for nivåer og subdomener der det er opptil seks spørsmål. Dette gjør at informasjonen som fremkommer enten er mer ulik som følge av forskjellige oppfølgingsspørsmål, eller at man har mindre informasjon å basere vurderingen sin på.

En siste mulig årsak vi vil løfte frem er at det for flere av domenene, spesielt emosjonsregulering, kan spekuleres i om dagsformen kan påvirke hvordan pasienten rapporterer sin egen fungering.

Oppsummert viser resultatene våre god til utmerket enighet på domene, komponent og LPFS totalnivå, noe som gir støtte for bruk av intervjuet i klinisk praksis. På subdomenenivå var det imidlertid store forskjeller i intertaterreliabiliteten, hvor noen

subdomener viser utmerkede reliabilitetsestimater, mens andre har tvilsom reliabilitet (Regier et al., 2013). Foreløpig kan vi da konkludere med at SCID-AMPD-1 kan brukes reliabelt på domene – og LPFS total nivå, mens det på subdomenenivå trengs revidering av flere subdomener før de kan brukes reliabelt.

7.3.1 Hva kan ha ført til god reliabilitet?

Til tross for at vi brukte et strengt test-retest design, fant vi jevnt over god enighet mellom de første og andre intervjuene. Det kan være flere årsaker til dette. For det første benyttet vi et intervju spesifikt rettet mot å vurdere nivå av personlighetsfungering på LPFS, der spørsmålene ved intervjuet er tilpasset dette formålet. Dette vil sannsynligvis øke reliabiliteten sammenlignet med å vurdere LPFS ut i fra informasjon som ikke er tilpasset beskrivelsene på skalaen. Dersom en bruker andre kilder for informasjon, må rateren fortolke hva pasientens respons vil tilsvare på LPFS, og det kan dermed være større rom for ulike vurderinger. Det at SCID-AMPD-1 er såpass strukturert kan også ha vært med å øke reliabiliteten.

For det andre fikk intervjuerne trening i hvordan de skulle skåre intervjuene. Dette kan også ha økt reliabiliteten. Det var også en konsensus i intervjuergruppen om hvordan man skulle skåre intervjuerne, da alle hadde fått omtrent den samme treningen. En retningslinje var at man for eksempel skulle skåre pasienten på nivå 2 dersom den svarte bekreftende på nivå 1 og nivå 3 svar, men ingen nivå 2 svar. Dette var etablert etter muntlig avtale med Donna Bender (Benjamin Hummelen, personlig kommunikasjon, november 2015). Dersom trening bidrar til å øke reliabiliteten, vil i så fall dette være viktig informasjon hvis intervjuet skal implementeres i praksis. Few et al. (2013) påpekte at klinikernes relative mangel på kjentskap med modellen i deres studie kan ha vært noe av årsaken til at de ikke fant høyere reliabilitet. Imidlertid er det viktig å poengtere at vi ikke hadde en gruppe med intervjuere uten trening å sammenligne med, så vi kan ikke konkludere med at det var treningen som skapte god reliabilitet. Ingen tidligere studier har heller gjort det, og det blir problematisk å sammenligne grad av trening på tvers av studier på grunn av ulike metoder og instrumenter for å vurdere LPFS. Det vil likevel være et viktig tema videre.

Dette kan imidlertid settes i sammenheng med grad av erfaring, som man jo også kan anse som en form for trening, selv om den er generell. Som vi nevnte i metode – og resultatdelen, så vi også på om det var noen sammenheng mellom hvor ulikt raterne

skåret samme pasient, og grad av erfaring. Det var ingen signifikant sammenheng, men signifikansverdi var ikke svært høy ($p = .088$), og dette i kombinasjon med Pearson r på $r = -.306$ vurderer vi som en rimelig høy korrelasjon når vi studerer noe slikt.

Resultatene tyder dermed på at grad av erfaring gjør at to ulike ratere sannsynligvis vil vurdere samme pasient mer likt enn det uerfarne vil gjøre. Dette er noe man også kunne forvente med et slikt klinisk instrument. Imidlertid er dette i vårt studie kun en tilleggsanalyse, da vi ikke har nok ratere i hver ratergruppe til å kunne trekke noen videre konklusjoner utover studiet vårt. Likevel er det et interessant funn, og det vil være spennende å utforske dette videre med tanke på hvem som eventuelt skal ta i bruk SCID-AMPD modul 1 og behov for trening.

Funnet kan settes i sammenheng med kritikken som har vært rettet mot LPFS ved at konstruktene er for komplekse til å bli reliabelt vurdert når raterne mangler klinisk ekspertise (Zimmermann et al., 2014). Det funnene våre tyder på er at reliabiliteten blir høyere ved mer klinisk erfaring. Imidlertid er det fremdeles god reliabilitet til tross for at studien inkluderer uerfarne studenter som intervjuere. Dette er i tråd med funnene til Zimmermann et al. (2014), som også fant at det ikke trengs flere års klinisk erfaring eller at modellen er for komplisert for å ta i bruk. Imidlertid må disse to studiene sees uavhengig av hverandre på grunn av metode.

Mest sannsynlig skyldes reliabiliteten at intervjuer fungerer godt. Til tross for et strengt design, fant vi høye estimer. Dette kan tyde på at intervjuet i seg selv er et reliabelt måleinstrument for å kartlegge nivå av personlighetsfungering på LPFS. Flere elementer ved intervjuet kan gjøre at det fungerer godt som for eksempel: at det er strukturert, at det på flere av nivåene har mange spørsmål som gjør at intervjuerne kan samle mye informasjon, og at spørsmålene er gode til å fange opp funksjonsnivået ved personlighetspatologi som beskrevet ved LPFS. Vi kan dermed si at resultatene våre gir foreløpig støtte for bruken av intervjuet i klinisk sammenheng, spesielt på LPFS total, komponent og domenenivå. Imidlertid er det visse subdomenenivåer som vi nevner tidligere bør gjennomgås. Det at intervjuet er så reliabelt med hensyn til disse konstruktene i Den alternative modellen for personlighetsforstyrrelser, gir også foreløpig støtte for brukbarheten til selve konstruktene.

7.4 Styrker og begrensninger

7.4.1 Styrker ved studien

Dette er ett av få studier som har undersøkt reliabiliteten til kriterium A: nivå av personlighetsfungering på den alternative modellen for personlighetsforstyrrelser i DSM-5. Det er det eneste studiet, som vi vet om, som har undersøkt reliabiliteten til måleinstrumentet SCID-AMPD Modul 1. SCID-AMPD modul 1 er et måleinstrument, skreddersydd til kriterium A. Dette gjør studiet vårt nyskapende og originalt, og resultatene bidrar til dagens forskningsfelt.

Vi har benyttet oss av en anerkjent måte å se på reliabilitet på: et test-retest design (DeVellis, 2012). Test-retest er et passende mål for egenskaper som er stabile over tid, slik som personlighet (Cohen, Swerdlik, Sturman, 2013, s.150). Studiet vårt er i tråd med anbefalingene for utforming av test-retest design fra DSM-5 feltforsøkene (Kraemer et al., 2012).

Studiet vårt har flere styrker. Utvalget vårt bestod av 33 deltagere som vi gjorde doble intervjuer på. Dette er et relativt stort utvalg, med tanke på studiets design. Utvalget er hentet fra en klinisk populasjon og består av deltagere med variert patologi, både i type og alvorlighetsgrad. Deltagere diagnostisert med borderline, unnvikende, avhengig, narsissistisk, paranoid og blandet personlighetsforstyrrelse er representert i utvalget. Det er også deltagere som kun er diagnostisert med symptomlidelser. Angående alvorlighetsgrad, viste de deskriptive analysene at vi har deltagere på alle nivå av personlighetsfungering. De overnevnte styrkene gjør at vi kan generalisere resultatene våre til en klinisk populasjon.

En annen styrke ved utvalget vårt, er at vi har valgt å ikke inkludere mennesker fra en normalpopulasjon. For det første, skal ikke måleinstrumentet brukes i en normalpopulasjon. For det andre, vil en intervjuer raskt merke om han/hun vurderer en kontroll og dermed skåre deltageren friskere. Dette kan føre til kunstig høye korrelasjoner.

For å unngå en form for forsker bias kalt forventningseffekt (Rosenthal, 1976) fikk intervjuer kun vite deltagers navn på forhånd. Det vil si at intervjuer var blind med tanke på alder, bakgrunnsinformasjon, og diagnoser. Intervjuerne i studien vår var klinikere med ulik faglig bakgrunn og grad av erfaring. Dette er svært likt slik det er i klinisk praksis, og øker studiens økologiske validitet. Hvilke to intervjuere som intervjuet samme deltagere rullerte etter en oppsatt tabell. Hvem som intervjuet først

og sist byttet vi også systematisk på. På denne måten har vi unngått at reliabiliteten kan forklares av et/flere intervjuer med høy grad av samsvar i skårer.

7.4.2 Begrensinger ved studien

Begrensninger med design

I et test-retest design kan tidsintervallet mellom de to test tidspunktene være en kilde til feilvarians (Cohen, Swerdlik, Sturman, 2013, s. 150). I vårt studie hadde vi et relativt kort tidsintervall (fra 2-14 dager), noe som kan ha ført til en kunstig forhøyet inter-rater reliabilitet. Faktorer som erfaring, øvelse, hukommelse, trøtthet og motivasjon kan påvirke et reliabilitetsestimat (Cohen, Swerdlik, Sturman, 2013, s. 151). Siden tidsintervallet var såpass kort, kan deltagerne ha husket spørsmålene og svarene de gav fra det første intervjuet til det andre. På den andre siden, har empiriske studier vist at personlighetspatologi ikke er et så stabilt fenomen som tidligere antatt (Grilo et al., 2004; Zanari et al., 2006). Dermed kunne et lengre tidsintervall ført til en faktisk endring i personlighetspatologi og ført til et kunstig lavt reliabilitetsestimat.

Begrensninger ved intervjuerne

Intervjuerne var ikke et tilfeldig utvalg av psykologer og psykiatere. De besto av en psykiater og to psykologspesialister som kan omtales som eksperter på personlighetsforstyrrelser. Psykiater Tore Buer Christensen har forsket mye på antisosial personlighetsforstyrrelse, mens psykologspesialistene arbeider på personlighetspoliklinikken på Ullevål. Dette kan ha bidratt til et høyere reliabilitetsestimat enn hvis det hadde vært et tilfeldig utvalg klinikere fra en allmennpoliklinikk. På den andre siden kan inkludering av studenter som intervjuere ha bidratt til å redusere reliabilitetsestimatet. Studenter mangler klinisk erfaring, og erfaring med måleinstrumenter. Det at vi rullerte på hvem som intervjuet sammen, kan ha kontrollert for effekten av både eksperter og studenter.

Begrensninger ved utvalg

Det er flere begrensninger med tanke på utvalg og hvem vi kan generalisere resultatene til. For det første bestod utvalget av pasienter som hadde blitt spurt av deres terapeuter om å delta, og som så hadde sagt seg villige til dette. Forskning har vist at individer som deltar frivillig i studier er forskjellige fra individer som ikke velger å delta (Rosenthal &

Rosnow, 1975, ref. i Bordens & Abbott, 2011). I tillegg ble deltagere tilbudt et gavekort på 500,- for å delta i begge intervjuene. Dette kan ha påvirket deltagerens motivasjon for å delta, samt hvem som deltok.

For det andre går det an å stille spørsmålsteget ved representativiteten til utvalget vårt. Hele 85% av alle deltagerne var diagnostisert med en personlighetsforstyrrelsesdiagnose. Dette er mye mer enn hva prevalensen vanligvis er i kliniske settinger. Dette kan nok skyldes at de fleste deltagerne ble hentet fra Dagbehandlingsnettverk. Samtidig er det personlighetsfungering vi ønsker å undersøke og studiet må da inneholde en god del deltagere med personlighetspatologi. Noen personlighetsforstyrrelser som antisosial, tvangspreget, schizoid, og schizotyp var desverre ikke representert i utvalget. I tillegg var det en skjev fordeling med tanke på kjønn, 73% var kvinner. Dette er ikke representativt for fordelingen av personlighetsforstyrrelser i befolkningen, hvor denne har vist seg å være tilnærmet lik for kvinner og menn (Coid et al., 2006; Torgersen et al., 2001). I kliniske settinger er derimot kvinner overrepresentert (Tyrrer, Reed, Crawford, 2015).

Begrensninger ved måleinstrumentet

SCID-AMPD modul 1 er et helt nytt, upublisert instrument. Vi har merket oss noen potensielle svakheter ved instrumentet. Noen svakheter er knyttet til oppbygningen av instrumentet. Vi har blant annet stusset på om instrumentet er fullstendig dimensjonalt, eller om prototypisk tenkning preger nivåene. Dette gjenspeiles i beskrivelsene på hvert nivå. For eksempel passer nivå 1 beskrivelser tvangspreget personlighetsproblematikk, nivå 2 passer narsissistisk problematikk, nivå 3 ligner mer borderline vansker, og nivå 4 schizotyp personlighetsforstyrrelse. Et eksempel på dette er under subdomenet selvaktelse. Her er beskrivelsen på nivå 1: "Til tider minsket selvaktelse, med overdrevent kritisk eller noe fordreid vurdering." På nivå 2: "Har sårbar selvaktelse med overdrevent fokus på ytre evaluering, med ønske om å bli anerkjent. Har følelse av ufullstendighet eller mindreverdighet, med tilsvarende oppblåst eller minsket selv vurdering." Nivå 3: "Sårbar selvaktelse som lett påvirkes av hendelser. Selvbildet mangler sammenheng. Selvvurdering er unyansert, med selvforakt, grandiositet, eller en ulogisk, urealistisk kombinasjon." Nivå 4: "Har svakt eller forvrengt selvbilde som lett trues ved samhandling med andre; vesentlige forvrengninger og forvirring rundt selvvurdering." (Bender, et al., 2014). Beskrivelsene er identiske med beskrivelsene i

nivå av personlighetsfungeringsskalaen i DSM-5 (APA, 2013). Dette er dermed ikke bare en svakhet ved måleinstrumentet, men også mot den underliggende modellen.

Med tanke på kvaliteten til spørsmålene i måleinstrumentet, inneholder noen av disse flere spørsmål i ett. Et eksempel på dette er under subdomenet konstruktive, prososiale indre normer for atferd. Spørsmålet på nivå 1 er som følger: "Føler du at du har faste verdier og at du vet hva du skal gjøre for å komme overens med andre og ha et tilfredstillende liv?" Dette gjør det vanskelig for intervjuobjektet å svare, og for intervjuere å skåre. Spørsmål i et måleinstrument bør være entydige (DeVellis, 2012).

Utover dette mangler SCID AMPD modul 1 en validitetsskala, samt avviker fra DSM-5 modellen ved at det ikke finner skårer for selv og interpersonligfungering (APA, 2013).

Andre svakheter er knyttet til bruk av instrumentet. Etter screeningsspørsmålene er det opp til intervjuer å vurdere hvilket nivå av fungering han/hun tror intervjuobjektet befinner seg på. En ulik innledende vurdering kan føre til at en intervjuer begynner på nivå 2, mens en annen begynner på nivå 3. Intervjuer nummer 1 kan få avkrefteende svar på nivå 2, bevege seg ned til nivå 1 og 0, og få bekrefteende svar her. Intervjuer nummer 2 kan få bekrefteende svar på nivå 3, og dermed skåre personen på dette nivået. Den innledende vurderingen, og påfølgende valg av spørsmål kan føre til at samme person blir skåret svært ulikt av to intervjuere. Intervjuet er dermed sårbart for subjektive vurderinger, samt at ikke alle spørsmål stilles. Å stille alle de nivåbestemmende spørsmålene ville imidlertid vært svært tidkrevende. Det blir en avveining om hvorvidt en skal ha et tidsbesparende intervju med noe lavere reliabilitet, eller et grundigere og mer tidsomfattende intervju med muligens høyere reliabilitet. Måleinstrumentet er i utgangspunktet tidkrevende. I vårt studie brukte intervjuerne mellom 40 og 130 minutter å gjennomføre intervjuet (M=80 min). Dette er imidlertid tilnærmet like lang tid som klinikere bruker på dagens semistrukturerte intervjuer (Tyrer, Geoffrey, Reed, 2015).

7.5 Implikasjoner

Dette studiet tyder på at måleinstrumentet SCID-AMPD modul 1 er et reliabelt måleinstrument. For SCID-AMPD modul 1 betyr dette at skåren på måleinstrumentet ikke skal endre seg, så lenge nivå av personlighetsfungering ikke endrer seg. Dette er

svært viktig i klinisk praksis da det kan føre til større enighet mellom ulike klinikere ved diagnostisering og bedre oppfølging av behandlingseffekt.

SCID-AMPD modul 1 bygger på kriterium A: nivå av personlighetsfungering, i den alternative modellen for personlighetsforstyrrelser. Det at måleinstrumentet har vist seg å være reliabelt, tyder på at det underliggende konstruktet også er reliabelt. Reliabilitet er en nødvendig, men ikke tilstrekkelig faktor for validitet. I den grad at reliabilitet kan si noe om validitet, tyder resultatene våre på at måleinstrumentet kan avgjøre tilstedeværelse og alvorlighetsgrad av personlighetspatologi. Et studie av Hopwood et al. (2011) har vist at alvorlighetsgrad er den viktigste prediktoren for nåværende og fremtidig fungering. Det har vært en svakhet at tidligere utgaver av DSM ikke har inneholdt en alvorlighetskala for personlighetsforstyrrelser (Bernstein, 1998). Vårt studie støtter opp om at SCID-AMPD modul 1 kan skille mellom god og dårlig fungering. Dette kan igjen føre til korrekt identifisering av mennesker som behøver behandling. Hvis SCID-AMPD modul 1 fortsetter å vise god reliabilitet i flere studier samt tilfredsstillende validitet, kan dette være et nytt måleinstrument for bruk i klinisk praksis.

Studien vår støtter opp om brukervennligheten til dimensjonale diagnoser. Flere forskere har hevdet at dimensjonale modeller for personlighetsforstyrrelser er for komplekse (Verheul, 2012; Livesley, 2012) for klinisk praksis. Resultatene fra vår studie motstrider dette. Både uerfarne og erfarne klinikere lærte seg måleinstrumentet på et todagers kurs og fikk reliable målinger. Det at vårt studie har vist at en dimensjonal modell er mer brukervennlig enn antatt, kan i ytterste konsekvens bidra til at den alternative modellen for personlighetsforstyrrelser blir vedtatt som ny modell for diagnostisering av personlighetsforstyrrelser. Innføringen av den alternative modellen vil føre til en ny forståelse av personlighetspatologi, en omlegging av diagnosesystemet og klinisk praksis.

7.6 Videre forskning

Det er behov for mer forskning på den alternative modellen for personlighetsforstyrrelser i DSM-5 (APA, 2013). Dette kan avgjøre modellens fremtid. Noe av grunnen til at den ikke ble vedtatt i hoveddelen av manualen var nettopp at det var gjort for lite forskning på den (Zachar et al., 2016). Vårt studie har bidratt med forskning på et måleinstrument, SCID-AMPD Modul 1 (Bender et al., 2014), basert på Kriterium A: nivå av personlighetsfungering.

Det blir viktig å replisere våre funn i fremtidige studier, for å støtte opp om at SCID-AMPD modul 1 er et reliabelt måleinstrument. Reliabilitetsstudier kan gjerne foretas med større utvalg av deltagere hentet fra en mer variert klinisk populasjon, og med flere intervjuere hentet fra et tilfeldig utvalg av klinikere. Videre blir det viktig å undersøke måleinstrumentets validitet opp mot tidligere validerte mål på personlighetsforstyrrelser som for eksempel SCID-II (First et al., 1997). Hvilken klinisk nytteverdi et mål på alvorlighetsgrad av personlighetspatologi har, er også viktig å etablere. Nivå av personlighetsfungering er imidlertid bare en del av den alternative modellen for personlighetsforstyrrelser. Kriterium B: patologiske personlighetstrekk, og de seks personlighetsforstyrrelsene (APA, 2013), behøver også å bli undersøkt med tanke på validitet, reliabilitet og klinisk nytteverdi.

Litteraturliste

- Alvarez, A. (2010). Levels of analytic work and levels of pathology: The work of calibration. *International Journal of Psychoanalysis*, 91(4), 859 – 878.
- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Arlington, VA: American Psychiatric Publishing.
- Amini, M., Pourshahbaz, A., Mohammadkhani, P., Reza, M., Ardakani, K., & Lotfi, M. (2015). The DSM-5 Levels of Personality Functioning and Severity of Iranian Patients With Antisocial and Borderline Personality Disorders. *Iranian Red Crescent Medical Journal*, 17(8).
- Arrindell, W. A. (2000). Changes in waiting-list patients over time: Data on some commonly-used measures. Beware! *Behaviour Research and Therapy*, 39(10), 1227-1247.
- Baastiansen, L., Hopwood, C. J., Van den Broeck, J., Rossi, G., Schotte, C. & De Fruyt, F. (2016). The Twofold Diagnosis of Personality Disorder: How Do Personality Dysfunction and Pathological Traits Increment Each Other at Successive Levels of the Trait Hierarchy? *Personality Disorders Theory*, 7(3), 280-292.
- Bach, B., Markon, K., Simonsen, E. & Krueger, R. F. (2015). Clinical Utility of the DSM-5 Alternative Model of Personality Disorders: Six Cases from Practice. *Journal of Psychiatric Practice*, 21(1), 3-25.
- Bender, D. S., Morey, L. C., & Skodol, A. E. (2011). Toward a Model for Assessing Level of Personality Functioning in DSM-5, Part I: A review of Theory and Methods. *Journal of Personality Assessment*, 93(4), 332-346.
- Bender, D. S., Skodol, A. E., First, M. B. og Oldham, J. M. (2014). *Modul 1: Strukturert klinisk intervju for den alternative modellen for personlighetsforstyrrelser i DSM-5 (SCID-AMPD)*. New York: New York State Psychiatric Institute.
- Benjamin, L. S. (1993). Dimensional, Categorical, or Hybrid Analyses of Personality: A Response to Widiger's Proposal. *Psychological Inquiry*, 4(2), 95-98.
- Berghuis, H., Kamphuis, J. H. & Verheul, R. (2014). Specific Personality Traits and General Personality Dysfunction as Predictors of the Presence and Severity of Personality Disorders in a Clinical Sample. *Journal of Personality Assessment*, 96(4), 410-416.

- Berghuis, H., Kamphuis, J. H., Boedijn, G., & Verheul, R. (2009). Psychometric properties and validity of the Dutch Inventory of Personality Organization (IPO-NL). *Bulletin of the Menninger Clinic*, 73(1), 44-60.
- Berghuis, H., Kamphuis, J. H., Verheul, R., Larstone, R., & Livesley, J. (2013). The General Assessment of Personality Disorder (GAPD) as an instrument for Assessing the Core Features of Personality Disorders. *Clinical Psychology and Psychotherapy*, 20(6), 544-557.
- Bernstein, R. F. (1998). Reconceptualizing Personality Disorder Diagnosis in the DSM-V: The Discriminant Validity Challenge. *Clinical Psychology: Science and Practice*, 5(3), 333-343.
- Bernstein, D. P., Iscan, C., Maser, J., & the Boards of Directors of the Association for Research in Personality Disorders and the International Society for the Study of Personality Disorders (2007). Opinions of personality disorder experts regarding the DSM-IV personality disorders classification system. *Journal of Personality Disorders*, 21(5), 536-51.
- Blashfield, R. K. (1993). Variants of Categorical and Dimensional Models. *Psychological Inquiry*, 4(2), 95-98.
- Bordens, K.S., & Abbott, B.B. (2011). *Research Design and Methods*. New York: McGraw Hill.
- Bornstein, R. F. & Huprich, S. K. (2011). Beyond dysfunction and threshold-based classification: A multidimensional model of personality disorder diagnosis. *Journal of Personality Disorders*, 25(3), 331-337.
- Bornstein, R. F. (2011). Toward a Multidimensional Model of Personality Disorder Diagnosis: Implications for DSM-5. *Journal of Personality Assessment*, 93(4), 362-369.
- Calabrese, W. R., & Simms, L. J. (2014). Prediction of Daily Ratings of Psychosocial Functioning: Can Ratings of Personality Disorder Traits and Functioning Be Distinguished? *Personality Disorders: Theory, Research and Treatment*, 5(3), 314-322.
- Chmielewski, M., Clark, L. A., Bagby, R. M. & Watson, D. (2015). Method Matters: Understanding Diagnostic Reliability in DSM-IV and DSM-5. *Journal of Abnormal Psychology* 124(3), 764-769.

- Clark, L. A., Livesley, W. J., & Morey, L. (1997). Special feature: Personality Disorder Assessment: The Challenge of Construct Validity. *Journal of Personality Disorders*, 11(3), 205-231.
- Clarkin, J. F. & Huprich, S. K. (2011). Do DSM-5 Personality Disorder Proposals Meet Criteria for Clinical Utility? *Journal of Personality Disorders*, 25(2), 195-205.
- Costa, P. T. & Widiger, T. A. (Eds). (1994). *Personality Disorders and the Five-Factor Model of Personality*. Washington DC: APA Books.
- Costa, P. T., & McCrae, R. R. (1990). Personality Disorders and the Five-Factor Model of Personality. *Journal of Personality Disorders*, 4(4), 362-371.
- Cramer, V., Torgersen, S., & Kringlen, E. (2007). Socio-Demographic Conditions, Subjective Somatic Health, Axis 1 Disorders and Personality Disorders in the Common Population: The Relationship to Quality of Life. *Journal of Personality Disorders*, 21(5), 552-567.
- Cramer, V., Torgersen, S, & Kringlen, E. (2006). Personality Disorders and Quality of Life. A Population Study. *Comprehensive Psychiatry*, 47(3), 178-184.
- Crawford, M. J., Koldobsky, N., Mulder, R., & Tyrer, P. (2011). Classifying Personality Disorder According to Severity. *Journal of Personality Disorders*, 25(3), 321-330.
- Cronbach, L. J. (1951). Coefficient Alpha and the Internal Structure of Tests. *Psychometrika*, 16(3), 297-334.
- Cohen, J. W. (1988). *Statistical power analysis for the behavioral sciences (2. utg)*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Cohen, J. R., & Swerdlik, M. E. (2010). *Psychological Testing and Assessment. An introduction to Tests & Measurements (7. utg)*. New York, NY: The McGraw-Hill Companies.
- Cohen, R.J., Swerdlik, M.E., & Sturman, E.D. (2013). *Psychological Testing and Assessment. An introduction to Tests & Measurements (7. utg)*. New York, NY: The McGraw-Hill Companies.
- Coid, J., Moran, P., Bebbington, P., Brugha, T., Jenkins, R., Farrell, M., Singleton, N., & Ullrich S. (2009) The Co-Morbidity of Personality Disorder and Clinical Syndromes in Prisoners. *Criminal Behavior and Mental Health*, 19(5), 321-33.
- Coid, J., Yang, M., Tyrer, P., Roberts, A., & Ullrich, S. (2006). Prevalence and correlates of personality disorder in Great Britain. *British Journal of Psychiatry*, 188(5), 423-431.

- Cuthbert, B. N. (2005). Dimensional models of psychopathology: Research agenda and clinical utility. *Journal of Abnormal Psychology, 114*(4), 565-569.
- <http://www.dagbehandlingsnettverk.no> Hentet 13.10.16
- DeVellis, R. F. (2012). *Scale development: theory and applications (3. utg)*. Thousand Oaks, California: Sage Publications.
- Doering, S., Burgmer, M., Heuft, G., Menke, D., Bäumer, B., Lübking, M., Feldman, M., Hörz, S., Schneider, G. (2013). Reliability and validity of the German version of the Structured Interview of Personality Organization (STIPO). *BMC Psychiatry, 13*(210).
- Donegan, N. H., Sanislow, C. A., Blumberg, H. P., Fulbright, R. K., Lacadie, C., Skudlarski, P., Gore, J.C., Olson, I.R., McGlashan, T.H., & Wexler, B. E. (2003). Amygdala hyperreactivity in borderline personality disorder: Implications for emotional dysregulation. *Biological Psychiatry, 54*(11), 1284–1293.
- Eaton, N. R., Krueger, R. F., South, S. C., Simms, L. J., & Clark, L. A. (2011). Contrasting Prototypes and Dimensions in the Classification of Personality Pathology: Evidence that Dimensions, but not Prototypes, are Robust. *Psychological Medicine, 41*(6), 1151-63.
- Few, L. R., Miller, J. D., Rothbaum, A. O., Meller, S., Maples, J., Terry, D. P., Collins, B. & MacKillop, J. (2013). Examination of the Section III DSM-5 diagnostic System for Personality Disorders in an Outpatient Clinical Sample. *Journal of Abnormal Psychology, 122*(4), 1057-1069.
- First, M.B., Gibbon, R.L., Spitzer, R.L., Williams, J.B.W., & Benjamin, L.S. (1997). *Structured Clinical Interview for DSM-IV Axis II personality disorders (SCID-II)*. Washington, DC: American Psychiatric Press.
- Frances, A. (1980). The DSM-III Personality Disorders Section: A Commentary. *American Journal of Psychiatry, 137*(9), 1050-1054.
- Grant, B. F., Stinson, F. S., Dawson, D. A., Chou, S. P., Ruan, W. J. (2004). Co-Occurrence of DSM-IV Personality Disorders in the United States: Results from the National Epidemiologic Survey on Alcohol and Related Conditions. *Comprehensive Psychiatry, 46*(1), 948-958.
- Grilo, C.M., Sanislow, C.A., Gunderson, J.G., Stout, R., Pagano, M.E., Yen, S., Morey, L.C., Zanarini, M.C., & McGlashan, T.H. (2004). Two-year stability and change of

schizotypal, borderline, avoidant, and obsessive-compulsive personality disorders. *Journal of Consulting and Clinical Psychology*, 72 (5), 767-775.

Gunderson, J. G. (2013). Seeking clarity for future revisions of the personality disorders in DSM-5. *Personality Disorders: Theory, Research, and Treatment*, 4, 368–376.

Hentschel, A. G. & Pukrop, R. (2015). The Essential Features of Personality Disorder in DSM-5. *The Journal of Nervous and Mental Disease*, 202(5), 412-418.

Heumann, K. A. & Morey, L. C. (1990). Reliability of Categorical and Dimensional Judgements of Personality Disorder. *American Journal of Psychiatry*, 147(4), 498-500.

Hill, J., Fudge, H., Harrington, R., Pickles, A., & Rutter, M. (2000). Complementary Approaches to the Assessment of Personality Disorder: The Personality Assessment Schedule and Adult Personality Functioning Assessment compared. *The British Journal of Psychiatry*, 176(5), 434-438.

Hopwood, C.J. & Thomas, K.M. (2012). Personality Disorders. I M. Hersen & D.C. Beidel (Eds.) *Adult Psychopathology and Diagnosis*. Hoboken, New Jersey: John Wiley and Sons, Inc.

Hopwood, C.J., Malone, J. C., Ansell, E. B., Sanislow, C. A., Grilo, C. M., McGlashan, T. H., Pinto, A., Markowitz, J. C., Shea, T., Skodol, A. E., Gunderson, J. G., Zanarini, M. C. & Morey, L. C. (2011). Personality Assessment in DSM-V: Empirical support for rating severity, style, and traits. *Journal of Personality Disorders*, 25(3), 305-320.

Horowitz, M. (2013). Disturbed Personality Functioning and Psychotherapy Technique. *Psychotherapy*, 50(3), 438-442.

Howitt, D., & Cramer, D. (2011). *Introduction to Statistics in Psychology*, 5th Edition. Essex, England: Pearson Education.

Hutsebaut, J., Feenstra, D. J., & Kamphuis, J. H. (2016a). Development and Preliminary Psychometric Evaluation of a Brief Self-Report Questionnaire for the Assessment of the DSM-5 Level of Personality Functioning Scale: The LPFS Brief Form (LPFS-BF). *Personality Disorders: Theory, Research, and Treatment*, 7(2), 192-197.

Hutsebaut, J., Kamphuis, J. H., Feenstra, D. J., Weekers, L. C., & De Saeger, H. (2016b). Assessing DSM-5-Oriented Level of Personality Functioning: Development and Psychometric Evaluation of the Semi-Structured Interview for Personality

- Functioning DSM - 5 (STiP-5.1). *Personality Disorders: Theory, Research, and Treatment*, June 2016.
- Johansen, M., Karterud, S., Pedersen, G., Gude, T. & Falkum, E. (2004). An Investigation of the Prototype Validity of the Borderline DSM-IV Construct. *Acta Psychiatrica Scandinavica*, 109(4), 289-98.
- Jorm, A. F., Duncan-Jones, P., & Scott, R. (1989). An Analysis of the Re-Test Artefact in Longitudinal Studies of Psychiatric Symptoms and Personality. *Psychological Medicine*, 19(2), 487 - 493.
- Jovev, M., & Jackson, H. J. (2004). Early Maladaptive Schemas in Personality Disordered Individuals. *Journal of Personality Disorders*, 18(5), 467-478.
- Karterud, S., et al. (2003). Day Treatment of Patients with Personality Disorders: Experiences from a Norwegian Treatment Research Network. *Journal of Personality Disorders*, 17(3), 243-262.
- Kendler, K. S., Prescott, C. A., Myers, J., & Neale, M. C. (2003). The Structure of Genetic and Environmental Risk Factors for Common Psychiatric and Substance use Disorders in Men and Women. *Archives of General Psychiatry*, 60(9), 929-937.
- Kernberg, O.F. (1989). A psychoanalytic classification of character pathology. In R. Lax (Ed), *Essential papers on character neurosis and treatment* (s. 191-210) New York, NY: New York University Press.
- Kraemer, H. C., Kupfer, D. J., Clarke, D. E., Narrow, W. E., & Regier, D. A. (2012). DSM-5: How Reliable Is Reliable Enough? *The American Journal of Psychiatry*, 169(1), 13-15.
- Krueger, R. F., McGue, M., Iacono, W. G. (2001). The Higher-Order Structure of Common DSM Mental Disorders: Internalization, Externalization, and their Connections to Personality. *Personality and Individual Differences*, 30(7), 1245-59.
- Krueger, R. F., Hopwood, C. J., Wright, A. G. C. & Markon, K. E. (2014). Challenges and Strategies in Helping the DSM Become More Dimensional and Empirically Based. *Current Psychiatry Report*, 16(515).
- Kupfer, D. J., First, M. B., & Regier, D. A. (2002). *A Research Agenda for DSM-V*. Washington, DC: American Psychiatric Association.
- Landers, R. (2015). Computing Intraclass Correlations (ICC) as Estimates of Interrater Reliability in SPSS. *The Winnower*, 2.

- Leising, D. & Zimmermann, J. (2011). An Integrative Conceptual Framework for Assessing Personality and Personality Pathology. *Review of General Psychology*, 15(4), 317-330.
- Livesley, J. (2012). Tradition versus empiricism in the current DSM-5 proposal for revising the classification of personality disorders. *Criminal Behaviour and Mental Health*, 22(2), 81-90.
- Livesley, W. J. (1998). Suggestions for a Framework for an Empirically Based Classification of Personality Disorder. *Canadian Journal of Psychiatry*, 43(2), 137-147.
- Maier, W., Lichtermann, D., Minges, J., & Heun, R. (1994). Personality disorders among the relatives of schizophrenia patients. *Schizophrenia Bulletin*, 20(3), 481-493.
- McCrae, R.R., & Costa, P.T. (1985). Updating Norman's adequate taxonomy: Intelligence and personality dimensions in natural language and in questionnaires. *Journal of Personality and Social Psychology*, 49(3), 710-721.
- McCrae, R. R., & Costa, P. T. (1997). Personality Trait Structure as a Human Universal. *The American Psychologist*, 52(5), 509-16.
- McCrae, R. R., & Terracciano, A. (2005). Personality Profiles of Cultures Project. *Journal of Personality and Social Psychology*, 89(3), 407-25.
- Morey, L. C., Bender, D. S., & Skodol, A. E. (2013). Validating the Proposed Diagnostic and Statistical Manual of Mental Disorders, 5th Edition, Severity Indicator for Personality Disorder. *The Journal of Nervous and Mental Disease*, 201(9), 729-735.
- Morey, L. C., Benson, K. T., Busch, A. J., & Skodol, A. E. (2015). Personality Disorders in DSM-5: Emerging Research on the Alternative Model. *Current Psychiatry Reports*, 17(24).
- Morey, L. C., Berghuis, H., Bender, D. S., Verheul, R., Krueger, R. F., & Skodol, A. E. (2011). Toward a Model for Assessing Level of Personality Functioning in DSM-5, Part II: Empirical Articulation of a Core Dimension of Personality Pathology. *Journal of Personality Assessment*, 93(4), 347-353.
- Morey, L. C., Hopwood, C. J., Gunderson, J. G., Skodol, A. E., Shea, M. T., Yen, S et al. (2007). Comparison of Alternative Models for Personality Disorders. *Psychological Medicine*, 37(7), 983-994.

- Morey, L. C., A. E. Skodol, & Oldham, J. M. (2014). Clinician Judgments of Clinical Utility: A Comparison of DSM-IV-TR Personality Disorders and the Alternative Model for DSM-5 Personality Disorders. *Journal of Abnormal Psychology, 123*(2), 398-405.
- Norman, W.T. (1963). Toward an adequate taxonomy of personality attributes: Replicated factor structure in peer nomination personality ratings. *Journal of Abnormal and Social Psychology, 66*, 574-583.
- Oldham, J. M., Skodol, A. E., Kellman, H. D., Hyler, S. E., & Rosnick, L. (1992). Diagnosis of DSM-III-R Personality Disorders by Two Structured Interviews: Patterns of Comorbidity. *American Journal of Psychiatry, 149*(2), 213-20.
- Oldham, J. M. (2009). Personality Disorders. Recent History and the DSM System. In J.M. Oldham, A.E. Skodol, & D.S. Bender (Eds.), *Essentials of Personality Disorders* (s. 3-13). Washington DC: American Psychiatric Publishing.
- OPD Task Force. (Eds.) (2008). *Operationalized Psychodynamic Diagnosis OPD-2: Manual of diagnosis and treatment planning*. Cambridge, MA: Hogrefe & Huber Publishers.
- Ormel, J., Koeter, M. W. J., & van den Brink, W. (1989). Measuring change with the General Health Questionnaire (GHQ). The problem of retest effects. *Social Psychiatry and Psychiatric Epidemiology, 24*, 227-232.
- Pallant, J. (2010). *SPSS Survival Manual, 4th Edition*. Berkshire, England: McGraw Hill Education.
- Parker, G. (1997). Special feature: The Etiology of Personality Disorders: A Review and Consideration of Research Models. *Journal of Personality Disorders, 11*(4), 345-369.
- Parker, G., Hadzi-Pavlovic, D., Both, L., Kumar, S., Wilhelm, L., & Olley, A. (2004). Measuring Disordered Personality Functioning: To Love and Work Reprised. *Acta Psychiatrica Scandinavica, 110*(3), 230-239.
- Pikonis, P. A., Hallquist, M. N., Morse, J. Q., & Stepp, S. D. (2011). Striking the (Im)proper Balance Between Scientific Advantages and Clinical Utility: Commentary on the DSM-5 Proposal for Personality Disorders. *Personality Disorders: Theory, Research, and Treatment, 2*(1), 68-82.
- Regier, D. A., Narrow, W. E., Clarke, D. E., Kraemer, H. C., Kuramoto, S. J., Kuhl, E. A., & Kupfer, D. J. (2013). DSM-5 Field Trials in the United States and Canada, Part II:

- Test-Retest reliability of Selected Categorical Diagnoses. *The American Journal of Psychiatry*, 170(1), 59-70.
- Rosenthal, R. (1976). *Experimenter effects in behavioral research*. New York: Irvington.
- Rounsaville, B. J., Alarcon, R. D., Andrews, G., Jackson, J. S., Kendell, R. E., Kendler, K. S. (2002). Basic Nomenclature Issues for DSM-V. In D.J. Kupfer, M.B. First, & D.E. Regier (Eds.), *A Research Agenda for DSM-V* (s. 1-29). Washington, DC: American Psychiatric Association.
- Salvatore, G., Nicolo, G., & Dimaggio, G. (2005). Impoverished dialogical relationship patterns in paranoid personality disorder. *American Journal of Psychotherapy*, 59(3), 247-265.
- Shedler, J., Beck, A., Fonagy, P., Gabbard, G. O., Gunderson, J., Kernberg, O., Michels, R., & Westen, D. (2010). Personality disorders in DSM-5. *American Journal of Psychiatry* 167(9), 1026-1028.
- Simonsen, S. & Simonsen, E. (2014). Contemporary Directions in Theories and Psychotherapeutic Strategies in Treatment of Personality Disorders: Relation to Level of Personality Functioning. *Journal of Contemporary Psychotherapy*, 44(2), 141- 148.
- Skodol, A.E., Bender, D.S., Gunderson, J.G., & Oldham, J.M. (2014). Personality disorders. In R.E. Hales, S.C. Yudofsky, & L.W. Roberts (Eds.) *The American Psychiatric Publishing Textbook of Psychiatry*, 6th ed. (851-894). Washington DC: American Psychiatric Publishing.
- Skodol, A.E., Bender, D.S., & Oldham, J.M. (2009). Future Directions: Toward DSM-V. I J.M. Oldham, A.E. Skodol, & D.S. Bender (Ed.), *Essentials of Personality Disorders* (s. 381-392). Washington DC: American Psychiatric Publishing.
- Skodol, A.E., Bender, D.S., Morey, L.C., Alarcon, R.D., Siever, L.J., Clark, L.A., Krueger, R.F., Verheul, R., Bell, C.C., Oldham, J.M. (2011). Proposed Changes in Personality Disorder Assessment and Diagnosis for DSM-5 part 1: Description and Rationale. *Personality Disorders: Theory, Research, and Treatment*, 2(1), 4-22.
- Skodol, A.E. (2012). Personality Disorders in DSM-5. *Annual Review of Clinical Psychology*. 8, 317-344.

- Skodol, A. E., Morey, L. C., Bender, D. S., & Oldham, J. M. (2015). The Alternative DSM-5 Model for Personality Disorders: A Clinical Application. *The American Journal of Psychiatry*, 172(7), 606-613.
- Spitzer, R. L., Williams, J. B. W., & Endicott, J. (2012). Standards for DSM-5 Reliability. *The American Journal of Psychiatry*, 169(5), 537.
- Thylstrup, B., Simonsen, S., Nemery, C., Simonsen, E., Noll, J. F., Myatt, M. W., & Hesse, M. (2016). Assessment of Personality-Related Levels of Functioning: A Pilot Study of Clinical Assessment of the DSM-5 Level of Personality Functioning Based on a Semistructured Interview. *BMC Psychiatry*, 16(298).
- Torgersen, S. (2014). Kan mysterier rommes i et moderne diagnosesystem? DSM-5: Bakgrunn, karakteristika og kontroverser. *Tidsskrift for norsk psykologforening*, 51(9), 708-714.
- Torgersen, S. (2009). Prevalence, Sociodemographics, and Functional Impairment. In J.M. Oldham, A.E., Skodol & D.S. Bender (Eds.). *Essentials of personality Disorders*. Washington DC: American Psychiatric Publishing.
- Torgersen, S., Kringlen E., Cramer, V. (2001). The prevalence of personality disorders in a community sample. *Arch Gen Psychiatry*, 58, 590-596.
- Trull, T. J., & Durrett, C.A. (2005). Categorical and Dimensional Disorders of Personality Disorder. *Annual Review of Clinical Psychology*, 1, 355-80.
- Tyrer, P. (2005). The Problem of Severity in the Classification of Personality Disorder. *Journal of Personality Disorders*, 19(3), 309-314.
- Tyrer, P. (2012). Diagnostic and Statistical Manual of Mental Disorders: A Classification of Personality Disorders That Has Had Its Day. *Clinical Psychology & Psychotherapy*, 19(5), 372-374.
- Tyrer, P., Geoffrey, M.R., & Crawford, M.J. (2015). Classification, assessment, prevalence, and effect of personality disorder. *Lancet*, 385, 717-726.
- Tyrer, P. & Johnson, T. (1996). Establishing the severity of personality disorder. *American Journal of Psychiatry*, 153(12), 1539-1597.
- Verheul, R. (2012). Personality Disorder Proposal for DSM-5: A Heroic and Innovative but Nevertheless Fundamentally Flawed Attempt to Improve DSM-IV. *Clinical Psychology and Psychotherapy*, 19(5), 369-371.

- Verheul, R. & Widiger, T. A. (2004). A Meta-Analysis of the Prevalence and Usage of the Personality Disorder not Otherwise Specified (PDNOS) Diagnosis. *Journal of Personality Disorders, 18*(4), 309-319.
- Wagner, A. W., & Linehan, M. M. (1999). Facial Expression Recognition Ability Among Women with Borderline Personality Disorder: Implications for Emotion Regulation? *Journal of Personality Disorders, 13*(4), 329–344. **H**
- Westen, D. & Shedler, J. (1999). Revising and Assessing Axis II: Developing a Clinically and Empirically Valid Assessment Method. *The American Journal of Psychiatry, 156*(2), 258-272.
- WHO (1993). *The ICD-10 Classification of Mental and Behavioural Disorders. Diagnostic criteria for research*. Geneva: World Health Organization.
- Widiger, T. A. (1991). Personality Disorder Dimensional Models proposed for DSM-IV. *Journal of Personality Disorders, 5*(4), 386-398.
- Widiger, T. A. & Samuel, D. B. (2005). Diagnostic Categories or Dimensions? A Question for the Diagnostic and Statistical Manual of Mental Disorders – Fifth Edition. *Journal of Abnormal Psychology, 114*(4), 494-504.
- Widiger, T. A. & E. Simonsen (2005). Alternative dimensional models of personality disorder: Finding a common ground. *Journal of Personality Disorders, 19*(2), 110-130.
- Widiger, T. A, Simonsen, E., Krueger, R., Livesley, J., Verheul, R. (2005). Personality Disorder Research Agenda for the DSM-V. *Journal of Personality Disorders, 19*(3), 315-338.
- Widiger, T. A. & Trull, T. J. (2007). Plate Tectonics in the Classification of Personality Disorder: Shifting to a Dimensional Model. *American Psychologist, 62*(2), 71-83.
- Yang, M., Coid, J., & Tyrer, P. (2010). Personality Pathology Recorded by Severity: National Survey. *The British Journal of Psychiatry, 197*(3), 193-199.
- Zanarini, M.C., Frankenburg, F.R., Hennen, J., Reich, D.B., & Silk, K.R. (2006). Prediction of the 10-years course of borderline personality disorder. *American Journal of Psychiatry, 163*(5), 827-832.
- Zachar, P., Krueger, R. F., & Kendler, K. S. (2016). Personality Disorder in DSM-5: An Oral History. *Psychological Medicine, 46*(1), 1-10.
- Zimmermann, M. (1994). Diagnosing Personality Disorders. A Review of Issues and Research Methods. *Archives of General Psychiatry, 51*(3), 225-245

- Zimmermann, J. Z., Benecke, C., Bender, D., Skodol, A. E., Schauenberg, H., Cierpka, M., Leising, D. (2014). Assessing DSM-5 Level of Personality Functioning From Videotaped Clinical Interviews: A Pilot Study with Untrained and Clinically Inexperienced Students. *Journal of Personality Assessment*, 96(4), 397-409.
- Zimmermann, J., Böhnke, J. R., Eschstruth, R., Mathews, A., Wenzel, K., & Leising, D. (2015). The Latent Structure of Personality Functioning: Investigating Criterion A From the Alternative *Model for Personality Disorders in DSM-5*. *Journal of Abnormal Psychology*, 124(3), 532-548.
- Zimmerman, M. & Rotschild, L. (2005). The Prevalence of DSM-IV personality disorders in psychiatric outpatients. *American Journal of Psychiatry*, 162(10), 1911-1918.

Vedlegg

Vedlegg A: Nivå av personlighetsfungeringsskalaen (APA, 2013) (oversatt til norsk av av Espen Folmo og Tuva Langjord).

	SELVET		INTERPERSONLIG	
Svekkelses nivå	Identitet	Målnrettethet	Empati	Nærhet
0—Lite eller ingen svekkelse	<p>Har vedvarende bevissthet om et unikt selv; opprettholder passende roller og adekvate grenser.</p> <p>Har stabil og selvregulert positiv selvaktelse, med presis selvvurdering.</p> <p>Evner å erfare, tolerere og regulere et bredt spekter av emosjoner.</p>	<p>Setter seg og forsøker å oppnå rimelige mål, basert på en realistisk vurdering av personlige evner.</p> <p>Anvender passende standarder for atferd, og realiserer disse på mange nivåer.</p> <p>Kan reflektere over og gi konstruktiv mening til indre erfaring.</p>	<p>Har god evne til å forstå andres opplevelser og motivasjoner i de fleste situasjoner.</p> <p>Forstår og setter pris på andres perspektiver, selv om en er uenig.</p> <p>Er klar over effekten egne handlinger har på andre.</p>	<p>Opprettholder flere tilfredsstillende og varige forhold i personlig liv og samfunnsliv.</p> <p>Ønsker seg og deltar i en rekke kjærlige, nære og gjensidige forhold.</p> <p>Streber etter samarbeid og felles nytte, og responderer fleksibelt på en rekke av andres ideer, emosjoner og atferd.</p>
1—Noe svekkelse	<p>Har relativt intakt selvfølelse, med noe nedsatt bevissthet rundt grenser når man erfarer sterke emosjoner og mental uro.</p> <p>Selvaktelse til tider nedsatt, i overkant kritisk eller noe forvrengt selvvurdering.</p> <p>Sterke emosjoner kan være plagsomme, forbundet med en innskrenkning i omfanget av emosjonelle opplevelser.</p>	<p>Er overdrevent innrettet mot mål, noe målhemmet, eller har konflikter rundt mål.</p> <p>Kan ha urealistiske eller sosialt upassende personlige normer, noe som begrenser noen aspekter av måloppnåelse.</p> <p>Er i stand til å reflektere over indre erfaringer, men kan legge for stor vekt på en enkelt type selvkunnskap (f.eks. intellektuell, emosjonell).</p>	<p>Noe begrenset evne til å sette pris på og forstå andres erfaringer; kan ha tendens til å synes at andre har urimelige forventninger, eller ønsker kontroll.</p> <p>Selv om man har evne til å vurdere og forstå forskjellige perspektiver, motsetter man seg å gjøre det.</p> <p>Er mangelfullt oppmerksom på effekten egen atferd har på andre.</p>	<p>Har evne til å danne varige forhold i personlig liv og samfunnsliv, med noen begrensninger når det kommer til grad av dybde og tilfredsstillelse.</p> <p>Har evne til å danne og ønsker å danne nære og gjensidige forhold, men kan være hemmet når det kommer til å uttrykke seg om viktige ting, og til tider anstrengt og forlegen dersom det kommer til intense emosjoner eller konflikter.</p> <p>Samarbeid kan hindres av urealistiske normer og krav; noe begrenset i evne til å respektere eller respondere på andres ideer, emosjoner og atferd.</p>

	SELVET		INTERPERSONLIG	
Svekkelses nivå	Identitet	Målnrettethet	Empati	Nærhet
2—Moderat svekkelse	<p>Er i stor grad avhengig av andre for å definere egen identitet, med tilsvarende svekket grenseoppgang til andre.</p> <p>Har sårbar selvaktelse drevet av overdreven engstelse rundt andres evaluering, med ønske om å bli godtatt. Har følelse av ufullstendighet eller mindreverdighet, med resulterende forstørret eller minsket selvaktelse.</p> <p>Emosjonell regulering er avhengig av positive vurderinger utenfra. Trusler mot selvaktelsen kan føre til sterke følelser som aggresjon eller skam.</p>	<p>Mål er oftere en måte å få anerkjennelse fra andre på enn at de er selvdannede, noe som kan vitne om mangel på sammenheng og/eller stabilitet.</p> <p>Personlige normer og krav kan være urimelig høye (f.eks. et behov for å være spesiell eller behage andre) eller lave (f.eks. ikke i overensstemmelse med rådende sosiale verdier og normer). Realisering vanskeligjøres av en følelse av mangel på autenticitet.</p> <p>Har nedsatt evne til å reflektere over indre erfaringer og opplevelser.</p>	<p>Er overdrevent oppmerksom på hvordan andre opplever ting, men bare når det oppleves som relevant for en selv.</p> <p>Er i meget stor grad selvrefererende: vesentlig nedsatt evne til å sette pris på og forstå andres opplevelser og å vurdere alternative perspektiver.</p> <p>Er generelt ikke oppmerksom på, eller bryr seg ikke om hva slags effekt egen atferd har på andre, eller urealistisk vurdering av egen effekt.</p>	<p>Er i stand til å gå inn i, og ønsker seg å gå inn i forhold i personlig liv og samfunnsliv, men forholdene kan i stor grad være overflatiske.</p> <p>Nære forhold er hovedsakelig basert på at behov for selvregulering og selvaktelse skal møtes, med urealistiske forventninger om å skulle bli forstått fullt ut av andre.</p> <p>Ser for det meste ikke forhold som gjensidige, og samarbeider hovedsakelig for egen vinnings skyld.</p>
3—Alvorlig svekkelse	<p>Har svak opplevelse av autonomi og egenvilje ("agency"); opplever mangel på identitet, eller tomhet. Grenser til andre er dårlige eller rigide: kan være overidentifisert med andre, legge for stor vekt på uavhengighet fra andre, eller pendle mellom disse.</p> <p>Skjør selvaktelse som lett påvirkes av hendelser, og selvbildet mangler sammenheng.</p> <p>Selvvurdering er unyansert: selvhat, selvforherligelse, eller en ulogisk, urealistisk kombinasjon.</p> <p>Emosjoner kan skifte raskt, eller en kronisk, urokkelig følelse av fortvilelse.</p>	<p>Har vansker med å sette seg og/eller oppnå personlige mål.</p> <p>Indre normer for atferd er uklare eller motstridende. Livet oppleves som meningsløst eller farlig.</p> <p>Har vesentlig svekket evne til å reflektere over og forstå egne mentale prosesser.</p>	<p>Evne til å vurdere og forstå tanker, følelser og atferden til andre mennesker er vesentlig begrenset; kan skille ut spesifikke aspekter ved andres opplevelser, særlig sårbarhet og lidelse.</p> <p>Er generelt ikke i stand til å vurdere alternative perspektiver; svært truet av meningsforskjeller eller alternative synspunkter.</p> <p>Er forvirret over eller ikke oppmerksom på egne handlingers innvirkning på andre; ofte forvirret over andre menneskers tanker og handlinger, feilattribuerer ofte egne destruktive tilbøyeligheter til andre.</p>	<p>Ønsker i noen grad å gå inn i forhold i samfunnsliv og i det personlige liv, men evnen til å ha positive og varige forbindelser er vesentlig svekket.</p> <p>Forhold er basert på en sterk overbevisning om at man fullt ut trenger nære kjærester, og/eller forventninger om å bli forlatt eller mishandlet. Følelser rundt å involvere seg nært med andre veksler mellom frykt/avvisning og desperat behov for relasjoner.</p> <p>Liten grad av gjensidighet: Andre ses først og fremst i lys av hvordan de påvirker selvet (negativt eller positivt); forsøk på samarbeid blir ofte forstyrret og/eller avbrutt fordi det oppfattes som at man blir avvist og oversett av andre.</p>
	SELVET		INTERPERSONLIG	
Svekkelse snivå	Identitet	Målnrettethet	Empati	Nærhet
	<p>Opplevelse av et unikt selv og følelse av autonomi og egenvilje ("agency") er nærmest</p>	<p>Har dårlig evne til å skille tanker fra handlinger, slik at evnen til å sette seg</p>	<p>Er tydelig ute av stand til å vurdere og forstå andres opplevelser og motivasjoner.</p>	<p>Ønsket om å knytte seg til andre er begrenset fordi man er dypt uinteressert, eller man forventer å</p>

<p>4—Ekstrem svekkelse</p>	<p>fraværende, eller er organisert rundt en oppfattet ytre forfølgelse. Forvirrede eller manglende grenser overfor andre.</p> <p>Har svakt eller fordreid selvbilde som lett trues av interaksjoner med andre; vesentlige forstyrrelser og forvirring rundt selv vurdering.</p> <p>Emosjoner er ikke i overensstemmelse med kontekst eller indre opplevelse. Hat og aggresjon kan være dominerende affekter, selv om de kan forkastes og attribueres til andre.</p>	<p>mål er alvorlig svekket, med urealistiske eller usammenhengende mål.</p> <p>Indre normer for atferd er omtrent fraværende. Sann realisering er nærmest utenkelig.</p> <p>Er fullstendig ute av stand til å reflektere konstruktivt over egne erfaringer. Kan være blind for personlig motivasjon, og/eller den erfares som utenfor en selv.</p>	<p>Har tilnærmet ingen bevissthet rundt andres perspektiver (oppmerksomheten er overdrevent på vakt, fokusert på behovstilfredsstillelse og å unngå å komme til skade).</p> <p>Sosiale interaksjoner er forvirrende og desorienterende.</p>	<p>komme til skade. Relasjoner til andre er avsondrede, desorganiserte eller gjennomgående negative.</p> <p>Forhold ses nærmest utelukkende i lys av deres evne til å bidra med støtte og hjelp, eller å påføre smerte og lidelse.</p> <p>Sosial/interpersonlig atferd er ikke gjensidig, men søker snarere å dekke grunnleggende behov eller å flykte fra smerte.</p>
-----------------------------------	---	--	---	--

Vedlegg B: forekomst av symptomidelser i utvalget vårt

Tabell B1. Forekomst av symptomidelser i utvalget vårt

Symptomlidelse	Antall deltagere med symptomlidelsen	Andel med symptomlidelsen
Skadelig bruk av cannabinoider	1	3,0 %
Skadelig bruk av sedativer og hypnotiske stoffer	1	3,0 %
Skadelig bruk av alkohol	1	3,0 %
Alkoholavhengighet	1	3,0 %
Bipolar enkel manisk episode	2	6,1 %
Enkeltstående depressiv episode	2	6,1 %
Tilbakevendende depressiv lidelse	10	30,3 %
Uspesifisert depressiv episode	1	3,0 %
Dystymi	5	15,2 %
Agorafobi uten panikk	1	3,0 %
Sosial fobi	6	18,2 %
Panikk lidelse	4	12,1 %
Generalisert angstlidelse	5	15,2 %
OCD	1	3,0 %
PTSD	4	12,1 %
Nevrasteni	1	3,0 %
Bulemia nevrosa	1	3,0 %
Annen spiseforstyrrelse	1	3,0 %
Uspesifisert spiseforstyrrelse	1	3,0 %
Hyperkinetisk lidelse	1	3,0 %
Ingen symptomidelser	8	24,2 %
Totalt	58	175,8%

*Dette er antall registrerte symptomidelser. Her brukes ICD-10 betegnelsen, da disse brukes i klinisk praksis i Norge. Flere av deltagerne har komorbide symptomidelser, som gjør at totalt antall symptomidelser summeres til 58, mens antall deltagere er 33.

Vedlegg C: Forekomsten av personlighetsforstyrrelser i utvalget vårt

Tabell C1.

Oversikt over personlighetsforstyrrelsedagnoser

Diagnose	Antall deltagere med diagnosen	Andel med diagnosen
Paranoid	1	3,0%
Schizoid	0	0,0%
Schizotyp	0	0,0%
Antisosal	0	0,0%
Borderline	13	39,4%%
Histrionisk	0	0,0%
Narcissistisk	1	3,0%
Unnvikende	6	18,2%
Avhengig	1	3,0%
Tvangspreget	0	0,0%
Blandet	8	24,2%
Ingen PF	5	15,2%
Totalt	35*	106,1%

*Dette er antall registrerte diagnoser. En av deltagerne har tre komorbide diagnoser, som gjør at totalt antall diagnoser summeres til 35, mens antall deltagere er 33.

Vedlegg D: Oversikt over rotasjon mellom intervjuerne

Tabell C1. Oversikt over rotasjon mellom intervjuerne

1.gangs intervju	2.gangs intervju	Antall doble intervjuer
Karoline	Marianne	3
Marianne	Karoline	4
Karoline	Kjetil	1
Kjetil	Karoline	2
Karoline	Kristoffer	0
Kristoffer	Karoline	1
Karoline	Kjerstin	1
Kjerstin	Karoline	0
Karoline	Tuva	1
Tuva	Karoline	1
Karoline	Tore	1
Tore	Karoline	1
Marianne	Kjetil	0
Kjetil	Marianne	0
Marianne	Kristoffer	0
Kristoffer	Marianne	1
Marianne	Tuva	1
Tuva	Marianne	2
Marianne	Kjerstin	1
Kjerstin	Marianne	1
Marianne	Tore	0
Tore	Marianne	1
Kjetil	Kristoffer	0
Kristoffer	Kjetil	1
Kjetil	Tuva	0
Tuva	Kjetil	0
Kjetil	Kjerstin	1
Kjerstin	Kjetil	0
Kjetil	Tore	0
Tore	Kjetil	1
Kristoffer	Tuva	1
Tuva	Kristoffer	0
Kristoffer	Kjerstin	0
Kjerstin	Kristoffer	0
Kristoffer	Tore	0
Tore	Kristoffer	3
Tuva	Kjerstin	1
Kjerstin	Tuva	0
Tuva	Tore	0
Tore	Tuva	1
Kjerstin	Tore	0
Tore	Kjerstin	1
Totalt		33

Vedlegg F: Oversikt over oppbygningen til måleinstrumentet SCID-AMPD modul 1

Figur F1

Oversikt over SCID-AMPD modul 1

