



**Próffræðilegir eiginleikar notendamiðaða
mælitækisins PSYCHLOPS (*Psychological Outcome
Profiles*)**

Helgi Héðinsson

**Lokaverkefni til Cand.psych.-gráðu
Háskóli Íslands
Heilbrigðisvísindasvið**



HÁSKÓLI ÍSLANDS

Próffræðilegir eiginleikar notendamiðaða mælitækisins PSYCHLOPS
(Psychological Outcome Profiles)

Helgi Héðinsson

Lokaverkefni til Cand.psych.-gráðu í sálfræði
Leiðbeinendur: Jakob Smári, Daníel Þór Ólason og Hafrún Kristjánsdóttir

Sálfræðideild
Heilbrigðisvísindasvið Háskóla Íslands
Júní 2010

Ritgerð þessi er lokaverkefni til Cand.psych.-gráðu í sálfræði og er óheimilt að afrita ritgerðina á nokkurn hátt nema með leyfi rétthafa.

© Helgi Héðinsson 2010

Prentun: Háskólaprent
Reykjavík, Ísland 2010

Höfundur vill koma á framfæri sérstöku þakklæti til Jakobs Smára, Daníels Þórs Ólasonar, Hafrúnar Kristjánsdóttur, Jóns Friðriks Sigurðssonar og Agnesar Agnarsdóttur fyrir gott samstarf og ábendingar við gerð verkefnisins. Höfundur þakkar starfsfólki á geðsviði Landspítala - háskólasjúkrahúss fyrir aðstoð við gagnasöfnun. Aðrir sem veittu beint eða óbeint aðstoð við gerð verkefnisins fá einnig miklar þakkir.

Útdráttur

Við mat á árangri sálfræðilegrar meðferðar er aðallega notast við mælitæki sem eru þýðismiðuð (*nomothetic*), en notendamiðuð (*client-centred*) mælitæki hafa ekki náð jafn traustri fótfestu við mat á árangri meðferðar í sálfræði og geðlæknisfræði. Markmið rannsóknarinnar var að kanna próffræðilega eiginleika íslenskrar þýðingar PSYCHLOPS (*Psychological Outcome Profiles*) sem er notendamiðuð mælitæki til notkunar við árangursmat. Mælitækið var borið saman við þýðismiðuð mælitæki, CORE-OM. Spurningalistarnir voru lagðir fyrir sjúklinga sem sóttu hugræna atferlishópmeðferð á geðsviði Landspítala eða á heilsugæslustöðvum á höfuðborgarsvæðinu. Þátttakendur voru sjúklingar sem luku meðferð og svöruðu spurningalistunum bæði fyrir og eftir meðferð. Fjöldi þátttakenda var 64, 72% konur og 28% karlar, á aldrinum 19 til 66 ára (meðalaldur 38,6 ár). Áhrifastærð (*Cohen's d*) fyrir breytingarskor á PSYCHLOPS var 1,31 en 0,60 á CORE-OM ($t = -7,101, p < 0,001$). Innri samkvæmni atriða á PSYCHLOPS var 0,85 við upphaf meðferðar og 0,83 við lok meðferðar. Breytingarskor hafði sterka fylgni við breytingarskor á CORE-OM (*Spearman's rho* = 0,64) og var einnig góð fylgni á milli samsvarandi þátta mælitækjanna. Niðurstöður gefa til kynna að íslensk þýðing PSYCHLOPS sé með viðunandi innri áreiðanleika og gott samleitni- og samtímaréttmæti. Jafnframt sýna háar áhrifastærðir fyrir breytingarskor að PSYCHLOPS er næm mæling á breytingar við lok meðferðar. Notkun PSYCHLOPS samhliða notkun þýðismiðuðra mælitækja við árangursmat sálfræðilegrar meðferðar virðist því geta aukið gæði matsins.

Efnisyfirlit

Inngangur	8
Mismunandi nálganir í mati á árangri meðferðar.....	9
Þýðismiðað mat (<i>nomothetic assessment</i>).....	9
<i>Gagnrýni á árangursmatsrannsóknir í sálfræði</i>	10
Einstaklingsmiðað mat (<i>idiographic assessment</i>).....	11
<i>Próffræðilegir eiginleikar einstaklingsmiðaðra mælitækja</i>	13
<i>Einstaklingsmiðuð mælitæki við árangursmat sálfræðilegrar meðferðar</i>	14
Samanburður á þýðismiðuðum og einstaklingsmiðuðum mælitækjum	16
<i>Takmarkanir þýðismiðaðra mælitækja</i>	16
<i>Takmarkanir einstaklingsmiðaðra mælitækja</i>	19
Samþætting mismunandi aðferða við árangursmat sálfræðilegrar meðferðar.....	20
PSYCHLOPS (<i>Psychological Outcome Profiles</i>).....	22
Próun notendamiðaða mælitækisins PSYCHLOPS	22
Uppbygging mælitækisins.....	24
Skorun mælitækisins	25
Próffræðilegir eiginleikar PSYCHLOPS	25
<i>Áreiðanleiki</i>	26
<i>Næmi fyrir breytingum</i>	26
<i>Réttmæti</i>	26
Reynsla meðferðaraðila af notkun PSYCHLOPS	27
Markmið rannsóknarinnar	28
Aðferð	29
Þátttakendur.....	29
Mælitæki	29
Rannsóknarsnið og framkvæmd.....	30
Tölfræðileg úrvinnsla gagna	30
<i>Næmi fyrir breytingum</i>	31
<i>Áreiðanleiki</i>	31
<i>Réttmæti</i>	32
Niðurstöður	35
Samanburður á þeim sem luku meðferð við þá sem luku ekki meðferð	35
Skor á PSYCHLOPS og CORE-OM	35
Næmi fyrir breytingum	36
Áreiðanleiki.....	37

Réttmæti	38
Umræða	41
Samantekt á helstu niðurstöðum	41
Takmarkanir á notkun PSYCHLOPS.....	43
Núverandi staða PSYCHLOPS og frekari rannsóknir	44
Gildi niðurstaða fyrir árangursmat sálfræðilegrar meðferðar.....	44
Notkun PSYCHLOPS á Íslandi	45
Heimildir	46

Töfluyfirlit

Tafla 1. Skor einstakra þátta og heildarskor á PSYCHLOPS og CORE-OM: Óstaðlað meðalskor og staðalfrávik ($N = 64$).	35
Tafla 2. Samanburður á áhrifastærðum breytingarskors ($N = 64$).	37
Tafla 3. Fylgni á milli staðlaðs breytingarskors PSYCHLOPS og CORE-OM ($N = 64$).	38
Tafla 4. Fylgni á milli þátta innan PSYCHLOPS og innan CORE-OM.....	39
Tafla 5. Samband á milli breytingarskors og mats þátttakenda á eigin bata við lok meðferðar	40

Myndayfirlit

Mynd 1. Samanburður á áhrifastærðum breytingarskors á PSYCHLOPS og CORE-OM.....	37
--	----

Inngangur

Á síðustu tveimur áratugum hefur eftirspurn aukist eftir stuttum, einkennamiðuðum mælitækjum sem auðveld eru í fyrirlögn, til notkunar í geðheilbrigðisþjónustu. Helstu ástæðurnar fyrir þessu eru aukin áhersla á bæði skráningu á áhrifum meðferðar og aðhald í kostnaði (Groth-Marnat, 2003). Aukinn þrýstingur er frá notendum að veitt sé hágæða þjónusta, sem felst meðal annars í því að þjónustuaðili veiti árangursríka meðferð og þróuð séu árangursrík og jafnframt skilvirk meðferðarúrræði, en jafnframt að kostnaður sé viðráðanlegur (Clement, 1996; Mirin og Namerow, 1991). Sem liður í því að hafa stjórn á ört vaxandi kostnaði við geðheilbrigðisþjónustu hafa tryggingafélög í auknum mæli aðeins veitt endurgreiðslu fyrir meðferðarúrræði sem reynst hafa árangursrík (Nelson-Gray, 1996). Einnig er ætlast til þess að meðferðaraðilar geti sýnt notendum þjónustunnar fram á, og að hve miklu leyti, tiltekið meðferðarúrræði beri árangur og að raunprófuð gögn séu því til stuðnings (Callaghan, 2001). Þetta hefur leitt til þess að rannsóknum á árangri meðferða, ásamt rannsóknum á áreiðanleika og réttmæti aðferða við árangursmat, hefur farið fjölgandi (Nelson-Gray, 1996).

Hægt er að skilgreina hágæða þjónustu þannig að veitt sé viðeigandi meðferð sem skilar góðum árangri, og er því nauðsynlegt að safna gögnum um árangur meðferða til að sýna fram á hvaða meðferðarúrræði eru þess virði að borga fyrir (Mirin og Namerow, 1991). Með því að fylgjast kerfisbundið með hvaða áhrif meðferð hefur á fólk gefast tækifæri til að auka gæði þjónustunnar. Dæmi um þætti sem árangursmat getur haft jákvæð áhrif á eru uppbygging meðferðar, greining, fjöldi meðferðartíma, greining á hvaða atriði í meðferð eru árangursríkari en önnur, viðmið fyrir lok meðferðar, gæði skráninga hjá meðferðaraðila, árangur meðferðar og fleira mætti telja (Clement, 1996). Að auki geta gögn um árangur gefið sjúklingi mikilvægar upplýsingar um batann sem hann hefur náð í meðferðinni og aukið trú hans á eigin getu við að takast á við vandamál sín (Antony og Barlow, 2002). Árangursmatsrannsóknir eru þar af leiðandi ómissandi í að greina gagnsemi sálfræðilegrar meðferðar (Clement, 1996; Mirin og Namerow, 1991). Árangur meðferðar getur verið metinn á ýmsa vegu. Þar ber helst að nefna breytingar í virkni sjúklings, mat hans á lífsgæðum og breytingar á upphaflegum einkennum og vandamálum (Panzarino, 1995). Árangursmatsrannsóknir geta þó verið erfiðar og flóknar í hönnun og framkvæmd og eru ekki allir á eitt sáttir um hvaða aðferðir henti best í árangursmati (Mirin og Namerow, 1991; Persons, 1991).

Mismunandi nálganir í mati á árangri meðferðar

Við mat á árangri sálfræðilegrar meðferðar hefur fyrst og fremst verið notast við mælitæki sem eru þýðismiðuð (*nomothetic*), en einstaklingsmiðuð (*idiographic*) mælitæki hafa aftur á móti ekki náð jafn traustri fótfestu við mat á árangri meðferðar í sálfræði og geðlæknisfræði (Nelson-Gray, 1996). Almennt er talað um að Gordon Allport hafi fyrstur gert greinamun á þessum tveimur nálgunum við mælingar innan sálfræðinnar á fjórða áratug síðustu aldar (Allport, 1937). Hugtökin sjálf eru þó upphaflega frá heimspekingnum Wilhelm Windelband en sumir telja að sálfræðingurinn Hugo Münsterberg hafi verið á undan Allport að kynna hugtökin innan sálfræðinnar (Hurlburt og Knapp, 2006). Hvernig sem því er háttað þá vísar þýðismiðað mat til þess almenna, þ.e. almennra lögmála sem eiga við hópa af fólki. Með þýðismiðaðri nálgun er gengið út frá því að einstaklingur búi yfir eiginleikum sem hann á sameiginlega með öðrum. Einstaklingsmiðað mat vísar hins vegar til þess sérstaka, þ.e. til einstaklingsins. Einstaklingsmiðuð nálgun byggist á því að einstaklingurinn sé einstakur og að hann búi yfir eiginleikum sem aðgreina hann frá öðrum einstaklingum (Cone, 1986; Haynes og O'Brien, 2000).

Þýðismiðað mat (*nomothetic assessment*)

Mikil áhersla er lögð á þýðismiðað mat innan sálfræðinnar og eru langflest sálfræðileg mælitæki af þessari tegund. Jafnframt hefur þessi tegund mats verið allsráðandi í mati á árangri sálfræðilegrar meðferðar og gegnt veigamiklu hlutverki í árangursmatsrannsóknum (Evans, Margison og Barkham, 1998; Lambert og Hawkins, 2004; Nelson-Gray, 1996).

Þeir eiginleikar sem helst einkenna þýðismiðað mat eru eftirfarandi: Í fyrsta lagi felur matið í sér aðferðir sem eru staðlaðar milli einstaklinga, þ.e. allir eru metnir með sama mælitækinu, á sama hátt, í sambærilegum aðstæðum og með eins svarblaði. Þýðismiðuð mælitæki mæla einnig sömu breytur, á sömu víddum, hjá öllum einstaklingum. Að auki felst mat sem byggt er á þýðismiðuðum mælingum að miklu leyti í að bera mælingar saman við samansafn mælinga á öðru fólki, en þannig er hægt að sjá hvort svör próftaka víkja frá fengnum viðmiðum (Haynes og O'Brien, 2000; Lambert og Hawkins, 2004). Dæmi um þýðismiðað mælitæki sem mikið er notað við árangursmat meðferða er Þunglyndiskvarði Becks (*Beck Depression Inventory-II*; BDI-II; Beck, Steer og Brown, 1996). Þetta er staðlaður spurningalisti með 21 spurningu, sem er meðal annars lagður fyrir fólk sem er í meðferð við þunglyndi. Með BDI-II er fengið heildarskor sem síðan er borið saman við fengin viðmið mælitækisins, en með þessu er verið að leggja mat á ástand fólks með tilliti

til annarra sem eru með þunglyndi og þeirra sem ekki eru þunglyndir. Jafnframt er með þessum samanburði hægt að skilgreina þunglyndisástand hvers og eins, til dæmis sem vægt, miðlungs-alvarlegt eða alvarlegt þunglyndi.

Að lokum má nefna að val á viðeigandi þýðismiðuðu mælitæki byggist á rannsóknum á fólki með sambærileg vandamál (Haynes og O'Brien, 2000). Ef sjúklingur er að glíma við einhverskonar sálræna erfiðleika er hægt að velja almennt mælitæki sem byggist á rannsóknum á fólki með almenna sálræna erfiðleika, eins og til dæmis CORE-OM (*Clinical Outcome in Routine Evaluation – Outcome Measure*; Evans o.fl., 2000). Þegar erfiðleikar eru betur skilgreindir er algengt að sértækari mælitæki séu valin, eins og til dæmis BDI-II fyrir þá sem eru með þunglyndiseinkenni.

Þýðismiðuð mælitæki eru almennt byggð þannig upp að sérfræðingar á sviðinu velja ákveðin atriði sem próftaki merkir svo við eftir því hvort atriðin eiga við hann og í hve miklum mæli. Sálfræðileg mælitæki eru því oftast en ekki hálfgerðir gátlistar fyrir samansafn einkenna (Ashworth o.fl., 2005a). Þar sem sömu spurningarnar eru lagðar fyrir alla er meðal annars hægt að skilgreina viðmið út frá vísbendingum um þýðismeðaltöl og þar af leiðandi staðla prófin. Þetta gerir það kleift að bera meðaltöl ákveðinna hópa saman, til dæmis þeirra sem eru í þunglyndismeðferð saman við þá sem ekki eru í meðferð. Með tilliti til þessara eiginleika má segja að þýðismiðuð mælitæki hafa gegnt veigamiklu hlutverki í árangursmatsrannsóknum innan sálfræðinnar (Haynes og O'Brien, 2000; Lambert og Hawkins, 2004; Meier, 2008).

Gagnrýni á árangursmatsrannsóknir í sálfræði

Persons (1991) er einn þeirra fræðimanna sem hefur gagnrýnt hvernig framkvæmd árangursmatsrannsókna innan sálfræðinnar hefur verið háttáð. Segir hún þær vera ófullnægjandi, þar sem þessar rannsóknir byggjast fyrst og fremst á þýðismiðuðum mælingum. Samkvæmt Persons þá samræmast árangursmatsrannsóknir innan sálfræðinnar ekki þeim kenningum og líkönum sem liggja meðferðum til grundvallar, þar sem í þessum rannsóknum er veitt staðlað, handbókarmiðað (*manual-based*) form af meðferð, sem stjórnast einungis af þeirri geðgreiningu sem viðkomandi fær. En í raun ætti meðferðin að vera einstaklingsmiðaðri og byggjast á greiningu á þeim vandamálum sem valda hverjum og einum mestum erfiðleikum. Sem dæmi má nefna að sálfræðileg líkön, líkt og hugrænt líkan Aaron Beck fyrir þunglyndi (Beck, 1979), eru einstaklingsmiðuð. Meðferðaraðili sem vinnur með þunglyndi út frá hugrænum kenningum Becks reynir að öðlast skilning á

erfiðleikum einstaklingsins, þar á meðal félagslegri einangrun, hugsanaskekkjum, erfiðleikum í vinnu, sambandsörðugleikum og fleira mætti telja. Hefðbundið staðlað mat eins og tíðkast í árangursmatsrannsóknum metur hins vegar aðeins brot af þessum vandamálum. Í samræmi við gagnrýni Persons þá segir Seligman (1995) að árangursmatsrannsóknir, sem einungis byggjast á þýðismiðuðum mælingum, missi af ýmsum eiginleikum sálfræðilegrar meðferðar eins og hún er yfirleitt stunduð. Semsagt að sú meðferð sem er árangursmetin endurspeglir ekki þá meðferð sem yfirleitt er stunduð (Persons, 1991; Seligman, 1995).

Persons (1991) færir rök fyrir annarri nálgun við árangursmat, sem byggist á kenningarstýrðri lýsingu á þeim vanda sem hver og einn er að glíma við (*case formulation approach*). Þessi nálgun felur í sér að val á matsaðferðum, mælingum og aðferðum sem beitt er í meðferðinni er einstaklingsmiðað og kenningardrifið. Hún veltur jafnframt á eiginleikum þeirra vandamála sem sjúklingurinn er að eiga við og því líkani sem meðferðin byggist á. Meðferðaraðili notar upplýsingarnar sem fást með matinu og þróar vinnutilgátu um hvað liggur að baki erfiðleikum og einkennum sjúklingsins, og er þessi tilgáta kenningarstýrð lýsing á vandanum (*case formulation*) sem unnið er út frá í meðferðinni. Með þessari einstaklingsmiðuðu aðferð er Persons að reyna tvíanna saman sálfræðilega meðferð og mat á árangri (Persons, 1991). Gagnrýni á umfjöllun Persons um þetta efni má finna í Silverman (1991) og Wilson (1996).

Einstaklingsmiðað mat (*idiographic assessment*)

Eins og komið hefur fram er notkun þýðismiðaðra mælitækja við mat á árangri meðferðar allsráðandi innan sálfræðinnar. Nokkur einstaklingsmiðuð mælitæki hafa þó komið fram en hefur þessi tegund mælinga aldrei náð teljandi vinsældum í árangursmati sálfræðilegrar meðferðar. Það sem virðist helst hafa aftrað útbreiðslu einstaklingsmiðaðra mælitækja er meðal annars sú gagnrýni að þau séu óáreiðanleg, of löng og flókin, og að meðferðaraðili þurfi að veita aðstoð við matið (Fitzpatrick, Davey, Buxton og Jones, 1998; Nelson-Gray, 1996).

Helstu eiginleikar einstaklingsmiðaðra matsaðferða eru í mörgum tilfellum andstæður við eiginleika þýðismiðaðra aðferða. Nánar tiltekið þá felur einstaklingsmiðað mat á fólki í sér að aðferðir, aðstæður, tegund mælinga og atriði mælinga, til dæmis atriði á spurningalista, eru sérstaklega sniðin að vandamálum sjúklings (Cone, 1986; Haynes og O'Brien, 2000). Sem dæmi um einstaklingsmiðaða nálgun má nefna, að þegar sjúklingur er

fenginn til að gera atferlistilraun til að vinna á fælni þá eru áreiti og aðstæður sem unnið er með í tilrauninni eingöngu valin út frá hans eigin erfiðleikum. Með því að velja þau atriði sem skipta mestu máli fyrir erfiðleika sjúklingsins er líklegra að mælingar nái að fanga það sem skiptir mestu máli fyrir hann og þá erfiðleika sem hann er að fá meðferð við. Að auki byggist klínískt mat fyrst og fremst á viðmiðum sem varða sjúklinginn persónulega, eins og til dæmis fyrri mælingum á einkennum hans eða hegðun. Klínískt mat tekur þó oft einnig tillit til þýðisviðmiða eins og þýðismiðuð nálgun gerir ráð fyrir en ekki að jafn miklu leyti (Haynes og O'Brien, 2000).

Einn helsti styrkleiki einstaklingsmiðaðra mælinga er næmi fyrir breytingum (*sensitivity to change*). Til að mæla árangur meðferðar á áreiðanlegan hátt verður mælitæki að vera næmt fyrir breytingum (Antony og Barlow, 2002). Næmi fyrir breytingum stendur fyrir það að hve miklu leyti mælingar ná að fanga þær raunverulegu breytingar sem verða á líðan eða vandamálum fólks sem markmiðið var að breyta með meðferðinni. Ástæðan fyrir því að einstaklingsmiðaðar mælingar geta verið mjög næmar fyrir breytingum, næmari en þýðismiðaðar mælingar, er sú að mælingarnar eru sniðnar að þeim einkennum og vandamálum sem valda mestum erfiðleikum. Í meðferðinni er síðan unnið að því að hafa áhrif á einmitt þessi einkenni og vandamál. Þar af leiðandi eru minni líkur á að matið feli í sér atriði sem eiga ekki við erfiðleika einstaklingsins, sem myndu draga úr næmi fyrir raunverulegum breytingum (Ashworth o.fl., 2007a; Deyo og Inui, 1984; Haynes og O'Brien, 2000; Hunter o.fl., 2004).

Eins og sést þegar eiginleikar einstaklingsmiðaðra og þýðismiðaðra mælinga eru útlistaðir þá standa þessar tvær tegundir mælingar fyrir andstæðar nálganir við mælingar. Í flestum tilfellum eru styrkleikar þýðismiðaðra mælinga eiginleikar sem einkenna takmarkanir einstaklingsmiðaðra mælinga og svo öfugt (Ashworth o.fl., 2007a; Haynes og O'Brien, 2000). Þessar lýsingar sem gefnar hafa verið á þýðismiðuðum og einstaklingsmiðuðum mælingum eru frekar einföld mynd af eiginleikum þessara tveggja aðferða. Þó svo mælingar séu yfirleitt annað hvort þýðismiðaðar eða einstaklingsmiðaðar í grunninn þá hafa þýðismiðaðar mælingar stundum ákveðna einstaklingsmiðaða eiginleika og hafa einstaklingsmiðaðar mælingar á sama hátt stundum þýðismiðaða eiginleika. Nánar tiltekið eru þýðismiðaðar mælingar oft á tíðum einstaklingsmiðaðar á þann hátt að viðeigandi mælitæki eru valin út frá þeim vanda sem sjúklingur er að eiga við. Sem dæmi um þetta, ef valið væri að leggja BDI-II fyrir sjúkling vegna kvartana hans um þunglyndislík einkenni, eða mælitæki sem mæla á kvíðaeinkenni væri valið fyrir einhvern sem kvartar undan kvíða. Dæmi um hvernig einstaklingsmiðuð mælitæki geta verið með

Þýðismiðaða eiginleika eru mælitæki sem innihalda staðlaða kvarða fyrir svörun (Haynes og O'Brien, 2000).

Próffræðilegir eiginleikar einstaklingsmiðaðra mælitækja

Mikið er lagt upp úr próffræðilegum eiginleikum þýðismiðaðra mælitækja, en þeir eru ekki síður mikilvægir fyrir einstaklingsmiðuð mælitæki. Próffræðileg markmið í þróun einstaklingsmiðaðra mælitækja er að byggja upp traust á réttmæti mælinga og átta sig á uppruna villna sem geta haft áhrif á mælingarnar og styrkja þannig réttmæti þeirra ályktana sem dregnar eru ef mælingunum (Haynes, 2006).

Innihaldsréttmæti (*content validity*) er líklega sú tegund réttmætis sem skiptir mestu máli fyrir einstaklingsmiðuð mælitæki. Markmið einstaklingsmiðaðra mælinga er eins og áður segir að mæla ýmsa eiginleika sem hafa mikla þýðingu fyrir einstaklinginn (hegðun, líðan, vandamál, meðferðarmarkmið og fleira) og stendur innihaldsréttmæti fyrir að hve miklu leyti þeir eiginleikar sem mældir eru endurspegla þá eiginleika sem hafa mikla þýðingu fyrir einstaklinginn, þ.e. sem mælitækið á að mæla (Haynes, 2006; Haynes og O'Brien, 2000).

Einnig er mikilvægt að samleitniréttmæti (*convergent validity*) og aðgreiningarréttmæti (*discriminant validity*) sé viðunandi. Samleitniréttmæti er metið út frá þeirri fylgni sem mælingar hafa við aðrar mælingar, sem eiga að mæla sömu eiginleika, eða sömu hugsmíð. Há fylgni við slíkar mælingar er til marks um gott samleitniréttmæti. Aðgreiningarréttmæti er metið út frá fylgni sem mælingar hafa við aðrar ólíkar mælingar, sem eiga að mæla óskylda eiginleika eða hugsmíð. Lág fylgni er til marks um gott aðgreiningarréttmæti (Aiken og Groth-Marnat, 2006; Haynes, 2006). Þegar samleitniréttmæti einstaklingsmiðaðra mælitækja er skoðað er tilvalið að bera skor þeirra saman við skor á þýðismiðuðu mælitæki sem á að mæla sömu eiginleika og sem hefur trausta próffræðilega eiginleika samkvæmt rannsóknnum (Haynes, 2006).

Önnur tegund réttmætis sem horft er til í þróun einstaklingsmiðaðra mælitækja er samtímaréttmæti (*concurrent validity*), en það er ein tegund af viðmiðsbundnu réttmæti (*criterion related validity*). Samtímaréttmæti er metið út frá þeirri fylgni sem mælingar hafa við aðrar mælingar sem framkvæmdar eru á sama tíma, sem eiga að mæla sömu eiginleika eða hugsmíð. Með þessari aðferð er til dæmis hægt að sjá hvort skor fólks sem er með alvarleg geðræn vandamál sé frábrugðið skori fólks sem ekki er að glíma við geðræn vandamál og þá hvort mælitækið nái að greina þar á milli (Aiken og Groth-Marnat, 2006).

Annað dæmi um mat á samtímaréttmæti mælitækis er að skoða hvort þær breytingar sem hafa orðið á vandamálum sjúklings samkvæmt mælingum, séu í samræmi við mat sjúklingsins á þessum breytingum (Ashworth o.fl., 2005a; Paterson, 1996). Hafa þarf þó í huga að einstaklingsmiðuð mælitæki eru mjög breytileg í sniðum og því misjafnt hvaða aðferðir við réttmætisprófanir henti hverju mælitæki.

Sérstæði einstaklingsmiðaðra mælinga hefur það hins vegar í för með sér að erfitt getur reynst að draga almennar ályktanir um réttmæti mælinga af þessari tegund. En til að styrkja réttmæti einstaklingsmiðaðra mælitækja, og jafnframt auka traust á klínísku mati sem byggist á þeim mælingum, er meðal annars hægt að nota staðlaðar aðferðir við uppbyggingu mælitækisins, tryggja að öll atriði sé vel skilgreind og skýr og notast við aðrar matsaðferðir samhliða mælitækinu (Cone, 1986; Haynes, 2006; Haynes og O'Brien, 2000).

Auk réttmætis er einnig mikilvægt að einstaklingsmiðuð mælitæki hafi viðunandi áreiðanleika (*reliability*), þ.e. að mælingarnar séu nógu áreiðanlegar til að mæla það sem þær eiga að mæla (Aiken og Groth-Marnat, 2006). Lítil áreiðanleiki hefur verið mikið vandamál í mörgum mælitækjum af þessari tegund, en það ræðst fyrst og fremst af því hversu stöðluð uppbygging mælitækisins er (Cone, 1986; Haynes og O'Brien, 2000). Stöðluð uppbygging mælitækis af þessari tegund felst yfirleitt í því að einhver atriði eru metin á stöðluðum kvarða, til dæmis Likert-kvarða eða sjónrænum samanburðarkvarða (*visual analogue scale*) (Bilsbury og Richman, 2002). Ef leiðbeiningar eru á mælitækinu sjálfu um hvernig eigi að svara spurningum, sem sagt sömu leiðbeiningarnar fyrir alla, og sömu atriðin eru lögð fyrir alla, þá eru það einnig dæmi um staðlaða uppbyggingu mælitækis (Bilsbury og Richman, 2002; Haynes og O'Brien, 2000). Með því að nota staðlaðar aðferðir við uppbyggingu einstaklingsmiðaðra mælitækja er hægt að auka áreiðanleika þeirra til muna, sem stuðlar að betra réttmæti þeirra (Cone, 1986; Haynes og O'Brien, 2000).

Einstaklingsmiðuð mælitæki við árangursmat sálfræðilegrar meðferðar

Einstaklingsmiðuð nálgun í mati á árangri sálfræðilegrar meðferðar hefur ekki náð traustri fótfestu og eru tiltölulega fá einstaklingsmiðuð mælitæki í boði samanborið við mikið úrval þýðismiðaðra mælitækja (Nelson-Gray, 1996). Hér á undan hefur verið gerð grein fyrir eiginleikum einstaklingsmiðaðs mats, en mælitæki af þessari tegund eru yfirleitt nánar skilgreind sem notendamiðuð (*client-centred/patient-centred*) mælitæki. Með notendamiðuðum mælitækjum eru mælingar byggðar á erfiðleikum hvers einstaklings fyrir

sígg og sér meðferðaraðili og/eða einstaklingurinn sjálfur um að skilgreina þá erfiðleika sem á að mæla. Sum notendamiðuð mælitæki eru þannig að próftaki skilgreinir sjálfur hvað eigi að mæla (*client-generated/patient-generated*). Hann velur sjálfur hvaða vandamál eða erfiðleikar skipta mestu máli og byggist mat á árangri meðferðar á þeim breytingum sem verða á því sem hann tilgreinir (Ashworth o.fl., 2004; Donnelly og Carswell, 2002).

Eitt fyrsta notendamiðuð mælitækið fyrir árangursmat í geðheilbrigðisgeiranum er *Personal Questionnaire* (PQ) sem hannað var af Shapiro (1961). Þegar mælitækið er lagt fyrir byrjar próftaki á því að gefa meðferðaraðilanum lýsingu á erfiðleikum sínum. Því næst er skilgreint hvað felst í bata frá erfiðleikunum og að lokum er framför skilgreind, sem fellur mitt á milli núverandi erfiðleika og bata frá þeim. Spurningalistinn er sniðinn að þessum lýsingum og inniheldur eingöngu atriði sem tengjast þeim. PQ þótti hins vegar flókið og tímafrekt mælitæki í fyrirlögn og náði þar af leiðandi aldrei miklum vinsældum. Þó er talið að hugmyndafræðin á bak við mælitækið hafi haft mikil áhrif á þróun annarra notendamiðuðra mælitækja sem komu í kjölfarið (Bilsbury og Richman, 2002). Dæmi um mælitæki sem byggð hafa verið út frá PQ eru *Ordinal Personal Questionnaire* (Phillips, 1970) og *Personal Questionnaire Rapid Scoring Technique* (PQRST; Chalkley og Mulhall, 1991) en með þróun þessara mælitækja var markmiðið að einfalda umfang mælinganna.

Eitt vinsælasta notendamiðuð mælitækið sem notað er við árangursmat er *Goal Attainment Scaling* (GAS; Kiresuk og Sherman, 1968). GAS var hannað til að mæla hegðun, hugsun, tilfinningar og vandamál sem skipta mestu máli fyrir einstaklinginn og sem ætlunin er að vinna með í meðferð. Í upphafi meðferðar eru meðferðarmarkmið skilgreind og eru búnir til fimm-punkta mælikvarðar fyrir hvert og eitt þeirra. Því næst er núverandi virkni próftaka metin á hverjum mælikvarða fyrir sig. Virknin er síðan endurmetin aftur við lok meðferðar, og jafnvel í eftirfylgd þegar ákveðinn tími er liðinn frá lokum meðferðar. Í rannsókn Kiresuk og Sherman (1968) reyndist samtímaréttmæti lágt til viðunandi og við mat á áreiðanleika kom í ljós viðunandi innanflokks fylgni.

Target Complaints (TC; Battle, Imber, Hoehn-Sario, Nash og Frank, 1966) er notendamiðuð mælitæki sem þróað var til að meta breytingar í líðan við meðferð. Í TC er próftaki beðinn um að gera grein fyrir þeim kvörtunum, þ.e. þeim erfiðleikum, sem eru ástæður fyrir því að hann leitaði hjálpar. Síðan er alvarleiki hvernar kvörtunar metinn. Við lok meðferðar er aftur lagt mat á alvarleika þeirra og breytingin sem verður á alvarleikamælingunni táknar árangur meðferðarinnar. TC er stutt og einfalt í bæði skorun og fyrirlögn og talið næmara fyrir breytingum en þýðismiðuð mælitæki. Samt sem áður hefur TC ekki orðið eftirsóknarvert til notkunar í árangursmati, en ein ástæða fyrir því hefur

verið sögð hversu lítil stöðlun er í þeim aðferðum sem notaðar eru til að safna kvörtunum (Donnelly og Carswell, 2002). Fáar rannsóknir hafa verið gerðar á próffræðilegum eiginleikum TC. Í rannsókn Bond, Bloch og Yalom (1979) var samtímaréttmæti TC viðunandi en áreiðanleiki lágur, en það er líklega vegna þess hversu lítil stöðlun er í uppbyggingu mælitækisins.

Psychological Outcome Profiles (PSYCHLOPS; Ashworth o.fl., 2004) er nýlegt notendamiðað mælitæki sem hannað var til að mæla árangur sálfræðilegrar meðferðar. PSYCHLOPS svipar nokkuð til TC þar sem spurt er um lýsingu á núverandi erfiðleikum, í stað þess að spyrja spurninga sem varða væntingar sjúklings um árangur meðferðar, líkt og í GAS. Nánar tiltekið er próftaki beðinn um að gefa skriflega lýsingu á því vandamáli sem veldur honum mestum erfiðleikum og metur hann síðan hversu mikil áhrif vandamálið hefur haft á hann undanfarna viku, á sex-punkta kvarða. Við lok meðferðar er próftaka birt lýsingin sem hann hafði sjálfur gefið á vandamálinu og metur hann aftur hversu mikil áhrif það hefur haft á hann undanfarna viku. Nánari umfjöllun um PSYCHLOPS hefst á bls. 22.

PQ, GAS og TC eru meðal þeirra notendamiðuðu mælitækja sem mest hafa verið notuð í árangursmati sálfræðilegrar meðferðar, en vafalaust hefði verið hægt að telja upp fleiri. Mörg þessara mælitækja hafa hins vegar haft það sameiginlegt að búa yfir ýmsum eiginleikum sem hafa gert það að verkum að þau hafa ekki þótt eftirsóknarverð til notkunar við árangursmat. Lítil reynsla er komin á notkun PSYCHLOPS og binda höfundar mælitækisins vonir við að það búi yfir þeim eiginleikum sem þurfi til að njóta samþykkis meðferðaraðila og verði eftirsóknarvert til notkunar við árangursmat (Ashworth o.fl., 2004).

Samanburður á þýðismiðuðum og einstaklingsmiðuðum mælitækjum

Takmarkanir þýðismiðaðra mælitækja

Þýðismiðuð mælitæki eru almennt viðurkennd sem hentug mælitæki til þess að bera saman hópa og safna viðmiðum, en með notkun þeirra við árangursmat fylgja ákveðnar takmarkanir (Haynes og O'Brien, 2000; Lambert og Hawkins, 2004). Einn helsti veikleiki þessara mælitækja er að þau innihalda yfirleitt ýmis atriði sem tengjast ekki þeim vandamálum sem próftaki á í erfiðleikum með. Sem dæmi, ef lagður er fyrir spurningalisti sem inniheldur níu mismunandi atriði fyrir mat á þunglyndi, er reiknað heildarskor út frá öllum atriðunum. Við mat á árangri meðferðar fyrir einhvern sem á aðeins í erfiðleikum með þrjú af þessum atriðum, er í raun ekki verið að gera ráð fyrir breytingum af hendi meðferðarinnar á hinum sex atriðum spurningalistans. Með því að hafa þessi atriði með í

matinu, atriði sem lýsa ekki erfiðleikunum á neinn hátt, er verið að draga úr næmi mælitækisins til að mæla breytingar á þeim erfiðleikum sem próftaki er að eiga við (Cone, 1986; Haynes og O'Brien, 2000). Annað sem hafa þarf í huga varðandi þýðismiðuð mælitæki er að þau geta verið misjafnlega yfirgripsmikil. Þegar þýðismiðuð mælitæki eru yfirgripsmikil, þ.e. innihalda mörg atriði, eru auknar líkur á að þau innihaldi atriði sem tengjast beint erfiðleikum próftaka, en aftur á móti eru einnig auknar líkur á að þau innihaldi fjölda atriða sem koma erfiðleikum hans ekkert við (Ashworth o.fl., 2007a; Fitzpatrick o.fl., 1998). Þetta getur haft veruleg áhrif á næmi mælitækisins fyrir breytingum, þar sem næmi fyrir breytingum veltur aðallega á því hversu vel atriði á mælitæki eru tengd þeim vandamálum sem notandi þess á í erfiðleikum með. Eftir því sem fleiri atriði eru sem tengjast ekki erfiðleikum viðkomandi, því minna verður næmi mælitækisins fyrir breytingum (Vermeersch, Lambert og Burlingame, 2000). Annað atriði varðandi mælitæki sem innihalda mörg atriði er að það getur verið mjög tímafrekt að leggja þau fyrir. Mörg þýðismiðuð mælitæki eru styttri, þ.e. innihalda færri atriði, og eru þar af leiðandi notendavænni hvað varðar tímann sem tekur að fylla þau út. Brugðið hefur verið á það ráð að stytta mörg lengri þýðismiðuðu mælitækjanna, þ.e. draga úr fjölda atriða, til að koma til móts við notendur þeirra. Með þessu sparast tími fyrir bæði meðferðaraðilann og notandann, auk þess sem minni líkur eru á óviðkomandi atriðum. Með færri atriðum eru hins vegar auknar líkur á að mælitækið innihaldi ekki þau atriði sem eiga best við erfiðleika próftakans, en með því að fara á mis við mikilvæg atriði er líklegt að vandinn verði vanmetinn (Ashworth o.fl., 2007a; Fitzpatrick o.fl., 1998). Í versta falli getur það leitt til þess að farið sé algjörlega á mis við þá erfiðleika sem skipta mestu máli. Þetta getur haft í för með sér að árangursmat verði byggt á atriðum sem endurspegla ekki þá erfiðleika sem leitað var meðferðar við, sem dregur bæði úr réttmæti matsins og næmi fyrir breytingum af hendi meðferðarinnar (Ashworth o.fl., 2007a; Hunter o.fl., 2004).

Þeir erfiðleikar sem valda fólki vanlíðan, og sem fólk leitar sér meðferðar við, eru það fjölbreytilegir að mögulegt er að eitt þýðismiðuð mælitæki nægir ekki til að ná utan um vandann (Lambert og Hawkins, 2004). Í rannsókn Lambert, Hatch, Kingston og Edwards (1986) voru þrjú þýðismiðuð mælitæki fyrir þunglyndi borin saman og metið hvort samræmi væri á milli þeirra í mælingum á þunglyndiseinkennum. Þessi mælitæki voru *Beck Depression Inventory* (BDI; Beck, Ward, Mendelson, Mock og Erbaugh, 1961), *Zung Self-rating Scale for depression* (ZSRS; Zung, 1965) og *Hamilton Rating Scale for Depression* (HRSD; Hamilton, 1960). Framkvæmd rannsóknarinnar fólst í því að skoða eldri rannsóknir þar sem þessi þrjú mælitæki voru lögð fyrir sömu hópana og reikna áhrifastærðir

fyrir hvert þeirra. Mælitækin eru öll einkennamiðuð en samt sem áður kom í ljós að þau gáfu mismunandi niðurstöður um hversu miklar breytingar höfðu orðið á einkennum fólks eftir meðferð. Niðurstöður þeirra er hægt að túlka sem svo að fjölþættur vandi, eins og þunglyndi yfirleitt er, verði ekki almennilega metinn með einni aðferð, heldur þurfi fleiri nálganir til að ná utan um vandann (Lambert og Hawkins, 2004; Nelson-Grey, 1996). Þar af leiðandi er mikilvægt að fara varlega í túlkun niðurstaða í árangursmatsrannsóknum þegar fáar þýðismiðaðar mælingar eru notaðar við mat á árangri (Lambert og Hawkins, 2004). Þunglyndi er gott dæmi um geðröskun sem getur stuðlað að margvíslegum erfiðleikum sem koma fram á ýmsum lífssviðum. En þar sem mörg þýðismiðuð mælitæki meta ekki þessi vandamál, til dæmis virkni í samskiptum, virkni í starfi og félagslegan stuðning, er ekki tekið tillit til breytinga á þessum vandamálum þegar árangur meðferðar er metinn (Antony og Barlow, 2002).

Það fer ekki milli mála að mikill munur er á þýðismiðuðum og einstaklingsmiðuðum mælitækjum. Vöntun hefur verið á að rannsóknir séu gerðar þar sem munurinn á milli þessara tveggja nálgana er raunprófaður, þ.e. þýðismiðuð og einstaklingsmiðuð mælitæki lögð fyrir samtímis og þau borin saman. Í rannsókn Hunter o.fl. (2004) á um það bil 700 langtímanotendum geðheilbrigðisþjónustu voru tvö þýðismiðuð mælitæki, *Health of the Nation Outcome Scale* (HoNOS) og *Avon Mental Health Measure* (AVON), borin saman við einstaklingsmiðað mælitæki, *Outcome of Problems of Users of Service* (OPUS). Einstaklingsmiðaða mælingin fólst í því að skrifa niður þrjú mikilvægustu vandamál sjúklunga sem raunhæft væri að vinna með í meðferð. Bæði sjúklingar og þeirra aðal meðferðaraðili svöruðu spurningalistunum til að athuga samræmi þeirra á milli og kom í ljós að 25% sjúklunga og 27% meðferðaraðila lýstu vandamálum á einstaklingsmiðaða mælitækinu (OPUS) sem ekki komu fram á þýðismiðuðu mælitækjunum.

Í rannsókn Ashworth o.fl. (2007a) var greining gerð á þemu svara (*thematic analysis*) sem fram komu á einstaklingsmiðaða (notendamiðaða-) mælitækinu PSYCHLOPS. Þátttakendur voru 215 manns sem sóttu viðtalsmeðferð á heilsugæslu. Tilgangur rannsóknarinnar var tvíþættur. Í fyrsta lagi voru þemu í svörum þátttakenda sem voru með skor yfir frávísunarviðmiðum, þ.e. klínískt skor, á þýðismiðaða mælitækinu CORE-OM, borin saman við skor þeirra þátttakenda sem voru með skor undir frávísunarviðmiðum (frávísunarviðmið fyrir skor á CORE-OM eru notuð til að greina hvort skor sé einkennandi fyrir klínískt þýði eða almennt þýði). Í öðru lagi var athugað hvort þau þemu sem komu fram í svörum þátttakenda á einstaklingsmiðaðri mælingu (PSYCHLOPS)

myndu einnig koma fram á þýðismiðaðri mælingu (CORE-OM). Svör þátttakenda voru greind í alls 61 þemu. Í ljós kom að ekki reyndist marktækur munur á þemum þeirra sem voru með klínískt skor á CORE-OM og þeirra sem voru með skor undir klínískum viðmiðum. Það var semsagt ekkert sem benti til þess að vandamál þeirra sem voru verr staddir samkvæmt CORE-OM væru annars eðlis en vandamál þeirra sem voru með lægra skor. Einnig kom í ljós að 27 af 61 þemum (44%) komu ekki fram í svörum þátttakenda á CORE-OM. Jafnframt voru 128 þátttakendur (60%) með að minnsta kosti eitt svar sem ekki kom fram á CORE-OM. Þau þemu sem oftast komu fram í svörum á PSYCHLOPS, sem ekki komu fram á CORE-OM, voru sambandsörðugleikar, vinnutengd vandamál, að „halda áfram“ (*moving on*), ástvinamissir og einbeiting. Út frá þessum niðurstöðum, ásamt niðurstöðum Hunter o.fl. (2004), er ljóst að þýðismiðuð mælitæki geta í einhverjum tilfellum farið á mis við fjölda atriða sem tengjast beint eða óbeint þeim vandamálum sem fólk er að glíma við. Þessar niðurstöður Hunter o.fl. (2004) og Ashworth o.fl. (2007a) eru í samræmi við samantekt Froyd, Lambert og Froyd (1996) á þýðismiðuðum mælitækjum sem notuð hafa verið við árangursmat sálfræðilegrar meðferðar. Samkvæmt þeirra greiningu voru 75% mælitækjanna sem mældu tilfinningatengda þætti (*intrapersonal*), en aðeins 17% þeirra mældu félagslega þætti (*interpersonal*), eins og til dæmis samskipti, og 1% mældu þætti sem varða félagslegt öryggi (*social role performance*).

Þetta vekur frekari spurningar um hvort þýðismiðaðar mælingar séu fullnægjandi þegar kemur að árangursmati sálfræðilegrar meðferðar (Ashworth o.fl., 2007a; Ashworth o.fl., 2009; Bilsbury og Richman, 2002). Í ljósi þess hversu fjölbreytileg sálræn vandamál eru, sem fólk leitar sér hjálpar við á heilsugæslustöðvum og göngudeildum geðdeilda (Clark, Hook og Stein, 1997), má færa rök fyrir því að einstaklingsmiðaðri nálgun í árangursmati meðferðar sé æskileg. Með einstaklingsmiðaðri nálgun er áherslan ekki á vandamál eða einkenni sem fyrirfram eru skilgreind af fagmanni, heldur er áherslan á þau vandamál eða einkenni sem sjúklingurinn sjálfur telur sig vera að glíma við (Cone, 1986; Haynes og O'Brien, 2000). Öfugt við þýðismiðað mat, þá má segja að einstaklingsmiðuð nálgun í mati viðurkenni sjúklinginn sem sérfræðing í vandamálum sínum, með því að leita skilnings á vandamálunum frá hans eigin sjónarhóli (Mead og Bower, 2000).

Takmarkanir einstaklingsmiðaðra mælitækja

Fjallað hefur verið um takmarkanir þýðismiðaðra mælitækja í mati á árangri meðferðar og þá hvernig einstaklingsmiðaðar mælingar geta að sumu leyti hentað betur. En

einstaklingsmiðuð mælitæki hafa einnig sínar takmarkanir sem ekki má gera lítið úr. Til dæmis eru einstaklingsmiðaðar mælingar oft óstaðlaðar, en í þeim tilfellum er aukin hættá að villur séu gerðar í bæði uppbyggingu mælitækisins og hvernig það er lagt fyrir. Dæmi um villur af þessu tagi er þegar matsaðili velur rangar breytur fyrir próftaka, rangan svörunarkvarða, skilgreinir illa þær breytur sem eru mældar og gefur óljósar eða rangar leiðbeiningar. Annað atriði sem vert er að nefna í sambandi við takmarkanir einstaklingsmiðaðra mælinga er að erfitt er, og jafnvel ómögulegt í sumu tilfellum, að bera mælingar saman við viðmið. Einstaklingsmiðuð mælitæki eru þó misjafnlega mikið stöðluð. Eftir því sem staðlaðari aðferðir eru notaðar við mælingar því auðveldara getur reynst að bera gögn saman við viðmið og draga þar af leiðandi ályktanir út frá þeim (Clement, 1996; Haynes og O'Brien, 2000). Samkvæmt Cone (1986) er mjög mikilvægt að einstaklingsmiðuð mælitæki séu byggð upp að hluta til með stöðluðum aðferðum. Til dæmis er gott að spurningalistar sem eru einstaklingsmiðaðir innihaldi einhverja staðlaða kvarða, til dæmis Likert-kvarða eða sjónræna samanburðarkvarða. Annað dæmi um stöðlun sem hentar einstaklingsmiðuðum mælitækjum er ef allir sem svara mælitækinu fá sömu leiðbeiningarnar um hvernig eigi að svara og að sömu atriðin séu lögð fyrir alla (Bilisbury og Richman, 2002; Cone, 1986). Þetta er gott dæmi um hvernig einstaklingsmiðaðar mælingar geta stundum haft ákveðna þýðismiðaða eiginleika. Með því að byggja mælitæki upp með þessum aðferðum, til dæmis með stöðluðum kvörðum, er meðal annars auðveldara að bera saman mælingar fyrir og eftir meðferð á hlutlægari hátt og greina þá breytingu sem verður þar á (Cone, 1986; Haynes og O'Brien, 2000). GAS og PSYCHLOPS eru dæmi um mælitæki sem eru að hluta til byggð upp með stöðluðum aðferðum.

Samþætting mismunandi aðferða við árangursmat sálfræðilegrar meðferðar

Það er ljóst að þýðismiðaðar og einstaklingsmiðaðar mælingar á árangri meðferðar hafa bæði sína styrkleika og takmarkanir. Því er greinilegt að það skiptir miklu máli að meta á faglegan hátt í hverju tilviki fyrir sig, hvaða aðferð henti best hverju sinni, en ekki einskorða sig við að nota aðeins aðra þeirra. Þýðismiðuð mælitæki henta til dæmis vel þegar mat á að vera víðtækt og þá sérstaklega á upphafsstigum greiningar þegar ennþá er verið að átta sig á erfiðleikum sjúklings. Með þýðismiðuðum mælitækjum, eins og til dæmis hegðunartengdum gátlistum og stöðluðum viðtölum, er hægt að greina vandamál sem valda erfiðleikum, sem síðan eru metin nánar með einstaklingsmiðuðum aðferðum. Þannig að einstaklingsmiðaðar mælingar á árangri meðferðar geta falist í mati á breytum

sem taldar eru skipta mestu máli út frá niðurstöðum úr þýðismiðuðu mati (Haynes og O'Brien, 2000). Algengt er að ferlið sé á þennan hátt, þ.e. að einstaklingsmiðaðar mælingar komi í kjölfar þýðismiðaðra mælinga. Einnig er þó hægt að nota þýðismiðuð mælitæki á einstaklingsmiðan hátt. Það er meðal annars hægt með því leggja áherslu á ákveðin atriði sem mælitæki inniheldur. Sem dæmi, ef BDI-II er notað við að meta árangur meðferðar er hægt að skoða ákveðin atriði á BDI-II sem markmiðið er að hafa áhrif á, frekar en að einblína á heildarskor (Haynes og O'Brien, 2000). Mat á árangri myndi þá byggjast á þeim breytingum sem verða á þessum tilteknu atriðum, en ekki á heildarskori.

Einstaklingsmiðaðar mælingar geta komið til móts við þær takmarkanir sem fylgja þýðismiðuðum mælingum og í ljósi þess getur besta matsaðferðin í sumum tilfellum verið falin í samhliða notkun beggja aðferða (Nelson-Gray, 1996). Þegar fólk leitar meðferðar er það venjulega vegna ákveðinna vandamála sem það vill fá hjálp við. Meðferðaraðili og sjúklingur vinna saman í meðferðinni að því að létta á erfiðleikunum og koma sér saman um meðferðarmarkmið með tilliti til vandamálanna. Þar af leiðandi er æskilegt að mat á árangri meðferðarinnar sé sniðið sérstaklega að þessum tilteknu vandamálum (Fiske o.fl., 1970). Með þeim sveigjanleika sem einstaklingsmiðaðar mælingar búa yfir myndi mælitækið vera sniðið að þeim vandamálum sem sjúklingurinn vill fá hjálp við og þar af leiðandi væri hægt að mæla breytingar á þeim erfiðleikum sem hann vill raunverulega breyta (Ashworth o.fl., 2007a; Haynes og O'Brien, 2000). Auk þess að stuðla að réttmætum mælingum á áhrifum meðferðar þá er það mikilvægur þáttur í sjálfri meðferðinni að sjúklingur skilgreini vanda sinn sem eigin orðum. Með því að ná fram, viðurkenna og innleiða það í meðferðina hvernig sjúklingur skilgreinir sinn eigin vanda öðlast meðferðaraðilinn betri skilning á erfiðleikum sjúklingsins. Þessar upplýsingar til viðbótar við sýn meðferðaraðilans getur aukið líkurnar á að betri árangur náist í meðferðinni, auk þess sem það stuðlar að réttmætari mælingum á áhrifum meðferðarinnar ef tekið er mið af því hvað sjúklingur telur skipta mestu máli (Macran, Ross, Hardy og Shapiro, 1999). En eins og gert hefur verið grein fyrir fylgja þýðismiðuðum mælingum ýmsir kostir sem eru eftirsóknarverðir, sem einstaklingsmiðaðar mælingar búa ekki yfir. En með því að nota einstaklingsmiðað mælitæki samhliða er hægt að leggja mat á þau vandamál sem sjúklingurinn telur skipta mestu máli fyrir sig, án þess að glata þeim kostum sem fylgja notkun þýðismiðaðra mælinga (Nelson-Gray, 1996).

Það er algjört lykilatriði að matsaðilar velji viðeigandi mælitæki þegar meta á árangur meðferðar. Algengar verklagsreglur í vali á mælitækjum gefa til kynna að mælitæki til notkunar í árangursmati meðferðar eigi að vera áreiðanleg, auðveld í fyrirlögn og túlkun

og síðast en ekki síst næm fyrir breytingum (Antony og Barlow, 2002). Eins og komið hefur fram þá hafa einstaklingsmiðuð mælitæki ýmsa eiginleika fram yfir þýðismiðuð mælitæki, sem gera þau að sama skapi afar hentug til að meta breytingar hjá sjúklingi (Ashworth o.fl., 2005a; Ashworth o.fl., 2009; Nelson-Gray, 1996). Hins vegar hafa mælitæki af þessari tegund löngum verið gagnrýnd fyrir að vera of óáreiðanleg, flókin og tímafrek í fyrirlögn og hafa þar af leiðandi ekki notið vinsælda við árangursmat (Fitzpatrick o.fl., 1998; Nelson-Gray, 1996). Sem tilraun til að bæta úr þessu var einstaklingsmiðaða mælitækið PSYHLOPS þróað (Ashworth o.fl., 2004). Markmiðið með þróun PSYCHLOPS var að bjóða upp á mælitæki til notkunar í árangursmati sálfræðilegrar meðferðar sem er bæði stutt og einfalt í fyrirlögn, næmt fyrir breytingum og jafnframt með góða próffræðilega eiginleika (Ashworth o.fl., 2004; Ashworth o.fl., 2005a).

PSYCHLOPS (*Psychological Outcome Profiles*)

PSYCHLOPS er notendamiðað mælitæki sem ætlað er til notkunar við mat á árangri sálfræðilegrar meðferðar (Ashworth o.fl., 2004). Notendamiðuð mælitæki hafa lítið verið notuð við árangursmat meðferðar í heilsugæslu (Ashworth, 2007a; Nelson-Gray, 1996) en með tilkomu PSYCHLOPS vilja höfundar mælitækisins að breyting verði þar á. Markmið þeirra er að PSYCHLOPS muni bætast við hóp þeirra mælitækja sem reglulega eru notuð við mat á árangri meðferðar og auka gæði árangursmats með notkun mælitækisins samhliða þýðismiðuðum mælitækjum (Ashworth o.fl., 2007a). Þar sem PSYCHLOPS er notendamiðað, hefur það ekki sömu greiningargetu og þýðismiðuð mælitæki, en kemur hins vegar til móts við ákveðnar takmarkanir þýðismiðaðra mælitækja með því að mæla þá erfiðleika sem fólki finnst skipta mestu máli fyrir sig (Ashworth o.fl., 2005a; Ashworth, 2007c; Ashworth o.fl., 2009).

Stefnt er að víðtækri dreifingu á PSYCHLOPS og hefur mælitækið nú þegar verið þýtt yfir á frönsku, úrdú, farsí, arabísku og kúrdísku til notkunar við árangursmat í vinnu með flóttamönnum. Að auki hefur mælitækið verið þýtt yfir á hollensku til notkunar í rannsóknarstarfi (Ashworth, 2007b; Ashworth, 2007c). Nú síðast var mælitækið þýtt yfir á íslensku til notkunar í þessari rannsókn.

Þróun notendamiðaða mælitækisins PSYCHLOPS

Vinna að þróun PSYCHLOPS hófst árið 1999 og var markmiðið með þróun þess að búa til mælitæki sem væri einstaklingsmiðaðra en þau mælitæki sem hafa tíðkast í árangursmati

sálfræðilegrar meðferðar. PSYCHLOPS var þróað af hóp fagmanna úr heilsugæslu frá *Department of General and Primary Care, King's College London*. Vinnan hófst á því að safna saman mælitækjum sem notuð hafa verið við árangursmat meðferðar á sviði geðheilbrigðis. Markmiðið var að auka þátt notenda þjónustunnar í matinu, einfaldlega vegna þess að skortur er á þannig mælitækjum og var niðurstaða faghópsins að þróa mælitæki sem væri notendamiðað. Jafnframt var ákveðið að hafa mælitækið notendamiðað á þann hátt að meðferðaraðili sér ekki um að skilgreina vanda sjúklingsins, heldur eru atriði valin af sjúklingnum sjálfum, þ.e. hann ákveður sjálfur hvað er mikilvægt, og byggjast mælingar á árangri á því sem hann tilgreinir (*client-generated*) (Ashworth o.fl., 2004; Ashworth, 2007c).

Einstaklingsmiðuð mælitæki hafa ekki náð vinsældum í mati á sálrænum vanda. Það sem virðist helst hafa aftrað útbreiðslu þeirra er meðal annars sú gagnrýni að þau séu of löng og flókin og að meðferðaraðili þurfi að veita aðstoð við matið (Fitzpatrick o.fl., 1998; Nelson-Gray, 1996). Með þetta í huga var markmiðið við þróun PSYCHLOPS að búa til mælitæki sem væri stutt, auðvelt í notkun og jafnframt næmt á breytingar í líðan fólks við meðferð, en einmitt þessir eiginleikar einkenna mælitæki sem henta best við árangursmat (Antony og Barlow, 2002). Þessir eiginleikar voru taldir forsenda þess að mælitækið yrði notendavænt og líklegra til að njóta samþykkis meðferðaraðila til notkunar innan geðheilbrigðisgeirans. Ákveðið var að nota mælitækið *Measure Your Medical Outcome Profile* (MYMOP; Paterson, 1996) sem fyrirmynd við uppbyggingu PSYCHLOPS. MYMOP er mælitæki sem er notað við mat á árangri læknisfræðilegrar meðferðar. MYMOP er lagt fyrir í upphafi og við lok meðferðar og metur breytingar sem verða á líkamlegri heilsu og einkennum við meðferð. Ástæðurnar fyrir því að MYMOP var valið eru að mælitækið er stutt, auðvelt í fyrirlögn, hefur komið vel út í réttmætisprófunum og er næmt fyrir breytingum (Paterson, 1996; Paterson o.fl., 2000). Í MYMOP er próftaki einfaldlega beðinn um skriflega lýsingu á því vandamáli eða vandamálum sem valda mestum erfiðleikum og hvaða áhrif það hafi á virkni hans. Að auki metur hann alvarleika þessara einkenna og áhrif á virkni á sjö-punkta kvarða. Að lokum er almenn vellíðan metin á sjö-punkta kvarða. Við lok meðferðar eru próftaka sýndar þessar lýsingar á vandamálum sínum og virkni og hann beðinn um að meta alvarleika þeirra aftur, ásamt því að spurningin um almenna vellíðan er endurtekin (Paterson, 1996).

Þar sem MYMOP var hannaður til að meta líkamlega heilsu var þörf á að gera nokkrar breytingar, aðallega í orðalagi, til að aðlaga nýja mælitækið að mati á sálrænum erfiðleikum. Ráðleggingar um orðalag og uppsetningu PSYCHLOPS, til að gera mælitækið

notendavænna, voru fengnar frá góðgerðasamtökunum *Depression Alliance* (DA; www.depressionalliance.org) og *Primary Care Mental Health Education* (PRIMHE; www.primhe.org.uk). Að auki voru ráðleggingar fengnar frá *Plain English Campaign* um hvernig hægt væri að gera mælitækið meira aðlaðandi fyrir notendur með skýru og einföldu orðalagi og framsetningu. Eftir breytingar fékk PSYCHLOPS svokallað „*Crystal Mark*“ frá *Plain English Campaign* sem viðurkenningu fyrir skýrt og einfalt orðalag (Ashworth o.fl., 2004).

Uppbygging mælitækisins

PSYCHLOPS er stuttur einnar síðu sjálfsmatskvarði sem samanstendur af fjórum spurningum sem meta þrjá þætti. Vandamálaþáttur, sem metinn er með tveimur spurningum, virkniþáttur sem metinn er með einni spurningu og vellíðunarþáttur sem einnig er metinn með einni spurningu. Á spurningalistanum sem lagður er fyrir í upphafi meðferðar er fyrst spurt hvaða vandamál valdi mestum erfiðleikum og gefur próftaki lýsingu á því skriflega. Því næst leggur hann mat á hversu mikil áhrif vandamálið hefur haft á hann undanfarna viku, á sex-punkta kvarða, frá engum áhrifum til mjög mikilla áhrifa. Loks er lagt mat á hversu lengi vandamálið hefur verið til staðar, á fimm-punkta kvarða, frá „fyrir minna en mánuði síðan“ til „fyrir meira en fimm árum síðan“. Í næstu spurningu er spurt á sama hátt um annað vandamál sem veldur próftakanum erfiðleikum. Virkni er síðan metin með því að biðja próftaka um að greina frá einhverju sem hann hefur átt í erfiðleikum með að gera vegna vandamálsins (eða vandamálanna). Því næst leggur hann mat á hversu erfitt það hefur verið undanfarna viku, á sex-punkta kvarða, frá alls ekki erfitt til mjög erfitt. Að lokum er vellíðan metin með því að spyrja hvernig líðan hefur verið undanfarna viku, á sex-punkta kvarða, frá mjög vel til mjög illa (Ashworth o.fl., 2005a).

Við lok meðferðar er próftaki beðinn um að meta hversu mikil áhrif vandamálið, sem hann gerði grein fyrir við upphaf meðferðar, hefur haft á líðan hans undanfarna viku. Svo ekki fari á milli mála hvert vandamálið er sem gerð var grein fyrir í upphafi meðferðar þá afritar meðferðaraðili orðrétt þá lýsingu sem próftaki gaf á vandamáli sínu í upphafi meðferðar og setur í þar til gerðan reit á „við lok meðferðar“ eintaki PSYCHLOPS. Áhrif vandamálsins eru síðan metið líkt og í upphafi meðferðar á sex-punkta kvarða, frá engin áhrif til mjög mikil áhrif. Seinna vandamálið og virkniþátturinn er síðan metið á sama hátt. Því næst er spurningin um vellíðan undanfarna viku endurtekin. Að lokum er próftaki spurður hvernig hann myndi lýsa líðan sinni við lok meðferðar á sex-punkta kvarða, frá

miklu betri til miklu verri. Þessi spurning er notuð sem réttmætisathugun, þar sem hægt er að skoða hvort svar próftaka við henni sé í samræmi við þá breytingu sem verður á heildarskori hans, út frá hinum spurningunum. Sem dæmi þá ætti skor þeirra sem segjast vera „miklu betri“ við lok meðferðar að vera búið að taka meiri jákvæðum breytingum frá því meðferð hófst, en hjá þeim sem segjast vera „aðeins betri“ (Ashworth o.fl., 2005a; Ashworth, 2007c).

Skorun mælitækisins

Þegar skor á PSYCHLOPS er reiknað er tekið saman skor fyrir alla þrjá þætti mælitækisins og reiknað út heildarskor. Alls eru fjórar spurningar sem mæla þættina þrjá og eru öll svör skorud á sex-punkta kvarða, frá 0 til 5. Til að fá skor fyrir vandamálapátt eru spurningarnar tvær sem þátturinn inniheldur lagðar saman og er deilt með tveimur til að fá meðalskor fyrir þáttinn. Hins vegar ef tilgreint er aðeins eitt vandamál (annarri spurningunni sleppt) er sú tala látin standa. Hinir tveir þættirnir, virkni og vellíðan, innhalda aðeins eina spurningu hvor og gefa svör við spurningunum þar af leiðandi skor þáttanna án frekar útreikninga. Heildarskor er síðan fengið með því að leggja saman meðalskor þáttanna og deila með fjölda þeirra (Ashworth o.fl., 2007c).

Mat á árangri meðferðar með PSYCHLOP byggist á þeirri breytingu sem verður á heildarskori milli mælinga. Til að fá breytingarskor er heildarskor fyrir meðferð einfaldlega dregið frá heildarskori eftir meðferð og stendur mismunurinn fyrir þann árangur sem náðst hefur í meðferðinni. Ef breytingarskor er í mínus er það tákn um bata (Ashworth o.fl., 2007c).

Próffræðilegir eiginleikar PSYCHLOPS

Í rannsókn Ashworth o.fl. (2005a) voru próffræðilegir eiginleikar PSYCHLOPS fyrst kannaðir, en birtar hafa verið tvær rannsóknir sem meta próffræðilega eiginleika mælitækisins. Skor þeirra 110 þátttakenda sem svöruðu PSYCHLOPS bæði fyrir og eftir meðferð í rannsókn Ashworth o.fl. (2005a) voru borin saman við skor þátttakenda á þýðismiðaða mælitækinu CORE-OM, sem er viðurkennt mælitæki við árangursmat sálfræðilegrar meðferðar. Próffræðilegir eiginleikar PSYCHLOPS voru einnig metnir í rannsókn Ashworth, Evans og Clement (2009), þar sem mælitækið var borið saman við þýðismiðaða mælitækið *Hospital Anxiety and Depression Scale* (HADS; Zigmond og Snaith, 1983). Í báðum þessum rannsóknum er um að ræða klínískt úrtak þátttakenda.

Samkvæmt þessum tveimur rannsóknum eru próffræðilegir eiginleikar PSYCHLOPS eftirfarandi:

Áreiðanleiki

Við mat á áreiðanleika PSYCHLOPS var reiknaður alfastuðull (*Cronbach's alpha*), en alfastuðull gefur vísbendingar um innri samkvæmni atriða á mælitækinu. Í rannsókn Ashworth o.fl. (2005a) var alfastuðull fyrir atriðin fjögur á PSYCHLOPS fyrir meðferð 0,79 og eftir meðferð 0,87, samanborið við alfastuðlana 0,94 fyrir meðferð og 0,95 eftir meðferð á CORE-OM. Sambærilegar niðurstöður komu fram í Ashworth o.fl. (2009) en þar voru alfastuðlarnir fyrir meðferð 0,75 og eftir meðferð 0,83, samanborið við alfastuðlana 0,67 fyrir meðferð og 0,79 eftir meðferð á HADS. Þegar alfastuðlar þeirra sem luku ekki meðferð (þ.e. svöruðu einungis PSYCHLOPS fyrir meðferð) voru bornir saman við þá sem luku meðferð reyndist ekki marktækur munur þar á (Ashworth o.fl., 2009).

Almennt er talið að alfastuðlar yfir 0,70 séu til marks um viðunandi innri samkvæmni, en þegar alfastuðlar eru komnir yfir 0,90 bendir það til einsleitni atriða á mælitæki (Nunnally og Bernstein, 1994). Þessar niðurstöður Ashworth o.fl. (2005a) og Ashworth o.fl. (2009) gefa til kynna að innri samkvæmni atriða á PSYCHLOPS sé góð.

Næmi fyrir breytingum

Til að leggja mat á næmi PSYCHLOPS fyrir breytingum í líðan fólks hafa áhrifastærðir (*Cohen's d*) fyrir breytingarskor á PSYCHLOPS verið bornar saman við þýðismiðuð mælitæki. Með breytingarskori er átt við mismuninn á skori á PSYCHLOPS fyrir meðferð og eftir meðferð. Sýnt hefur verið fram á að áhrifastærðir breytingarskora á PSYCHLOPS eru mun stærri en áhrifastærðir breytingarskora á bæði CORE-OM (PSYCHLOPS, $d = -1,53$; CORE-OM, $d = -1,06$) (Ashworth o.fl., 2005a) og HADS (PSYCHLOPS, $d = -1,61$; HADS, $d = -1,15$) (Ashworth o.fl., 2009). Breytingar sem mældar eru með PSYCHLOPS eru semsagt allt að 50% meiri en þær breytingar sem mælingar á CORE-OM og HADS gefa til kynna og má álykta út frá þessum niðurstöðum að PSYCHLOPS sé mjög næmt fyrir breytingum í líðan fólks við meðferð (Ashworth o.fl., 2005a; Ashworth o.fl., 2009).

Réttmæti

Við mat á samleitniréttmæti PSYCHLOPS var í rannsókn Ashworth o.fl. (2005a) könnuð fylgni breytingarskora á öllum þremur þáttum PSYCHLOPS við þrjá af fjórum þáttum

CORE-OM. Bæði mælitækin hafa þrjá samsvarandi þætti, þ.e. vandamálþátt, virkniþátt og vellíðunarþátt, en að auki hefur CORE-OM áhættuþátt. Fylgnistuðlar voru reiknaður með *Spearman's rho*. Niðurstaðan var að nokkuð sterk fylgni var á milli samsvarandi þátta á mælitækjunum báðum, þ.e. vandamál = 0,48, virkni = 0,40 og vellíðan = 0,52. Jafnframt var góð fylgni á milli heildar breytingarskors PSYCHLOPS og heildar breytingarskors CORE-OM (að undanskildum áhættuþætti), eða 0,61. Hins vegar kom í ljós að fylgnin var mjög sambærileg á milli aðskildra þátta mælitækjanna, sem getur bent til þess að aðgreining þáttanna á PSYCHLOPS er veik. Höfundar benda hins vegar á, að samkvæmt Evans o.fl. (2002) og Lyne, Barrett, Evans og Barkham (2006) hafa þættirnir á CORE-OM háa fylgni sín á milli, þ.e. að aðgreining þátta á CORE-OM sé veik. Há fylgni milli þátta innan CORE-OM getur að einhverju leyti skýrt hvers vegna þættir á PSYCHLOPS höfðu álíka mikla fylgni við sambærilega og ósambærilega þætti á CORE-OM.

Í rannsókn Ashworth o.fl. (2009) kom í ljós sterk fylgni breytingarskors á PSYCHLOPS við heildar breytingarskor á HADS. Fylgnistuðlar voru reiknaður með *Pearson's r*. Vandamálþáttur PSYCHLOPS hafði fylgnina 0,58 við HADS, virkniþáttur 0,51, vellíðunarþáttur 0,69 og þegar allir þrír þættirnir voru teknir saman var fylgnin 0,65 við heildar breytingarskor HADS.

Við mat á samtímaréttmæti PSYCHLOPS var í rannsókn Ashworth o.fl. (2005a) bornar saman áhrifastærðir breytingarskors þeirra þátttakenda sem mátu líðan sína við lok meðferðar „aðeins betri“ og þeirra sem mátu líðan sína „miklu betri“. Öðrum svarmöguleikum var sleppt vegna þess hversu fáir þátttakendur svöruðu þeim. Hærra breytingarskor, þ.e. meiri árangur, hjá þeim sem segjast vera „miklu betri“ við lok meðferðar, en hjá þeim sem segjast vera „aðeins betri“, gefur til kynna gott samtímaréttmæti. Áhrifastærðir fyrir svör á PSYCHLOPS voru marktækt stærri fyrir þá sem svöruðu „miklu betri“: -2,16 samanborið við 1,20 hjá þeim sem svöruðu „aðeins betri“ ($t = -3,78; p = 0,001$). Samtímaréttmæti var reiknað á sama hátt í rannsókn Ashworth o.fl. (2009) og voru niðurstöðurnar þar líkt og hjá Ashworth o.fl. (2005a), að áhrifastærðir fyrir svör voru marktækt stærri fyrir þá sem svöruðu „miklu betri“: -2,18 samanborið við -0,94 fyrir þá sem svöruðu „aðeins betri“ ($t = -8,14; p = 0,001$).

Reynsla meðferðaraðila af notkun PSYCHLOPS

Sem hluti af réttmætisprófunum á PSYCHLOPS fengu Ashworth o.fl. (2005b) fjóra meðferðaraðila til að segja frá reynslu sinni af notkun mælitækisins. Að auki voru þeir

beðnir um að bera mælitækið saman við CORE-OM. Meðferðaraðilum þótti PSYCHLOPS alveg jafn viðeigandi og þýðismiðuð mælitæki við að meta árangur og jafnframt að það væri vel boðlegt fyrir fólk í meðferð þar sem enginn hafi neitað að svara. Þó töldu þeir að meiri kröfur væru gerðar til fólks með því að láta það gera skriflega grein fyrir erfiðleikum sínum, í stað þess að einfaldlega merkja við atriði líkt og í CORE-OM. Fram kom að einn kostur PSYCHLOPS var að gefa meðferðaraðilanum innsýn í hvernig fólk skilgreinir sjálft erfiðleika sína, sem síðan væri hægt að nýta í meðferðinni. Mat meðferðaraðilanna gaf til kynna að eiginlegar upplýsingar sem fengnar væru með PSYCHLOPS væru góð viðbót við meginlegar upplýsingar úr CORE-OM og þar af leiðandi væri hentugt að nota þessi mælitæki saman. Þetta mat meðferðaraðila á reynslu þeirra af notkun PSYCHLOPS gefur til kynna gott sýndarréttmæti (*face validity*) PSYCHLOPS og að mælitækið hafi ýmsa eftirsóknarverða eiginleika sem geti ýtt undir það að meðferðaraðilar velji það til notkunar í árangursmati sálfræðilegrar meðferðar.

Markmið rannsóknarinnar

Eins og fjallað hefur verið um er nær eingöngu notast við þýðismiðuð mælitæki við mat á árangri sálfræðilegrar meðferðar. Hins vegar hafa einstaklings-/notendamiðuð mælitæki ákveðna eiginleika sem þykja eftirsóknarverðir og geta aukið gæði árangursmats, til viðbótar við hefðbundin þýðismiðuð mælitæki. Markmið rannsóknarinnar er að meta próffræðilega eiginleika íslenskrar þýðingar notendamiðaða mælitækisins PSYCHLOPS. Næmi fyrir breytingum, áreiðanleiki og samleitni- og samtímaréttmæti PSYCHLOPS verður metið og mun skor verða borið saman við skor á þýðismiðaða mælitækinu CORE-OM. Próffræðilegir eiginleikar PSYCHLOPS hafa áður verið metnir með samanburði við CORE-OM (Ashworth o.fl., 2005a), líkt og gert verður í þessari rannsókn. Í rannsókninni verður leitast eftir að endurtaka niðurstöður Ashworth o.fl. (2005a), en mikilvægt er að það sé gert af aðilum sem ekki eru tengdir upprunalegri þróun PSYCHLOPS og þannig auka vægi niðurstaðanna (Ashworth, 2007d).

Aðferð

Þátttakendur

Þátttakendur í rannsókninni voru sjúklingar sem sóttu hugræna atferlishópmeðferð á geðsviði Landspítala eða á heilsugæslustöðvum á höfuðborgarsvæðinu. Þátttakendur sóttu eitthvert af eftirfarandi hópmeðferðarúrræðum: Níu vikna HAM við lágu sjálfsmati, sex vikna HAM við þunglyndi og kvíða, sex vikna HAM við þunglyndi og kvíða fyrir 25 ára og yngri eða fimm vikna ósérhæfða HAM í heilsugæslu við þunglyndi og kvíða. Fjöldi meðferðartíma var á bilinu fimm til níu, mismunandi eftir hvaða hópmeðferð þátttakendur sóttu, en ekki voru til tölur yfir hversu vel þátttakendur sóttu meðferðartíma. Við upphaf rannsóknarinnar var stefnt að því að ná sem flestum þátttakendum, eða allt að 100, sem myndu ljúka meðferð og svara spurningalistum bæði fyrir og eftir meðferð.

Alls voru 135 sem svöruðu PSYCHLOPS og CORE-OM við upphaf meðferðar, þar af 98 konur (72,6%) og 37 karlar (27,4%). Fjöldi sjúklinga sem luku meðferð og svöruðu báðum spurningalistum við bæði upphaf og lok meðferðar voru 64 talsins (47% af þeim sem hófu meðferð), þar af 45 konur (70,3%) og 19 karlar (29,7%). Þessir þátttakendur voru á aldrinum 19 til 66 ára og var meðalaldur þeirra 38,6 ár ($sf = 13,0$). Niðurstöður rannsóknarinnar eru byggðar á þessum 64 þátttakendum.

Mælitæki

Mælitækin sem lögð voru fyrir voru í heildina níu talsins, en það var misjafnt eftir tegund hópmeðferðar hvaða mælitæki voru lögð fyrir hverju sinni. Rannsóknin byggist einungis á gögnum tveggja þessara mælitækja, PSYCHLOPS og CORE-OM, og voru þau lögð fyrir í öllum tilfellum. Önnur mælitæki voru BDI-II, *Beck Anxiety Inventory* (BAI), *Depression, Anxiety and Stress Scale* (DASS), *Robson Self-Esteem Questionnaire* (RSEQ), Mat á eigin líðan og *General Health Questionnaire* (GHQ), en ekki verður fjallað nánar um þau.

PSYCHLOPS (Ashworth o.fl., 2004) er einstaklingsmiðaður sjálfsmatskvarði sem samanstendur af fjórum atriðum og er notaður við mat á árangri sálfræðilegrar meðferðar. Árangur er metinn út frá heildarskori sem byggist á þremur þáttum, þ.e. vandamálþætti (2 atriði), virkniþætti (1 atriði) og vellíðunarþætti (1 atriði). Nánari lýsingu á PSYCHLOPS er að finna á blaðsíðum 22-28. Íslensk þýðing á PSYCHLOPS var í höndum Hafrúnar Kristjánsdóttur, Brynjars Halldórssonar, Jón Friðriks Sigurðssonar, Jakobs Smára og Helga Héðinssonar. Jakob Smári sá um bakþýðingu.

CORE-OM (Evans o.fl., 2000) er þýðismiðaður sjálfsmatskvarði sem samanstendur af 34 atriðum og er notaður við mat á árangri sálfræðilegrar meðferðar. Árangur er metinn á fjórum þáttum. Vandamálþáttur sem hefur undirþættina þunglyndi, kvíði, líkamleg vandamál og áföll (alls 12 atriði), virkniþáttur sem hefur undirþættina almenn virkni, nán samskipti og félagsleg samskipti (alls 12 atriði), áhættuþáttur sem metur áhættuhegðun gagnvart öðrum og manni sjálfum (6 atriði) og að lokum vellíðunarþáttur sem metur almenna vellíðan (4 atriði). Atriðin á *CORE-OM* eru sett fram sem fullyrðingar um líðan síðastliðna viku og eru þau metin á fimm-punkta kvarða þar sem notandi metur hversu oft honum leið á viðkomandi hátt síðastliðna viku. Svarmöguleikar eru alls ekki, stöku sinnum, stundum, oft og næstum alltaf eða alltaf. Sýnt hefur verið fram á að *CORE-OM* hafi góðan innri áreiðanleika og viðunandi áreiðanleika endurtekinnar prófunar (*test-retest*), gott samleitni- og aðgreiningarréttmæti og að mælitækið hafi gott næmi fyrir breytingum (Evans o.fl., 2002; Lyne o.fl., 2006). Einnig hafa próffræðilegir eiginleikar íslenskrar þýðingar *CORE-OM* verið skoðaðir, bæði meðal háskólanema og sjúklinga (Eva Dögg Gylfadóttir, 2007). Samkvæmt rannsókn Evu Daggar Gylfadóttur (2007) eru próffræðilegir eiginleikar íslensku þýðingarinnar sambærilegir próffræðilegum eiginleikum upprunalegu útgáfunnar. Íslensk þýðing á *CORE-OM* var í höndum þýðingarteymis á geðsviði Landspítala.

Rannsóknarsnið og framkvæmd

Framkvæmd var langsniðsrannsókn, þar sem þátttakendur voru beðnir um að svara spurningalistum við upphaf meðferðar og aftur við lok hennar. Þátttakendum var kynntur tilgangur fyrirlagnarinnar og voru allir látnir vita að þátttaka í rannsókninni væri valbundin. Leyfi fyrir rannsókninni var fengið frá Vísindasiðanefnd. Gagnasöfnun fór fram á tímabilinu október 2009 til apríl 2010.

Tölfræðileg úrvinnsla gagna

Við mat á próffræðilegum eiginleikum *PSYCHLOPS* var mælitækið borið saman við *CORE-OM*, sem er viðurkennt mælitæki við árangursmat sálfræðilegrar meðferðar (Evans o.fl., 2000). Unnið var úr skori á *CORE-OM* eins og lýst er í handbók mælitækisins (*CORE System Group*, 1998). Eins og komið hefur fram hefur *PSYCHLOPS* þrjá þætti (vandamál, virkni, vellíðan; alls fjögur atriði) og hefur *CORE-OM* fjóra (vandamál, virkni, vellíðan, áhætta; alls 34 atriði). Gögn voru í flestum tilfellum greind út frá þeim þremur þáttum sem

sameiginlegir eru milli mælitækjanna, sem þýðir að áhættuþættinum á CORE-OM (alls sex spurningar af 34) var sleppt.

Mismunandi matskvarðar eru á mælitækjunum og var skor beggja mælitækja þar af leiðandi staðlað með umbreytingu í z -skor. Til að breyta hráskori í z -skor var meðalskor fyrir meðferð dregið frá skori hvers og eins þátttakenda og var síðan deilt í þá tölu með staðalfrávikum skors fyrir meðferð. Með þessari stöðlun er upprunalegu dreifingunni breytt í dreifingu þar sem meðal z -skor fyrir meðferð verður 0, með staðalfrávikum 1,0. Með z -skori er skor þátttakenda túlkað út frá fjölda staðalfrávikum sem skor er frá meðalskori dreifingarinnar (Howell, 2007). Úrvinnsla gagna var unnin í tölfræðiforritinu SPSS 16,0 fyrir Mac.

Næmi fyrir breytingum

Við mat á næmi mælitækjanna fyrir breytingum voru reiknaðar áhrifastærðir Cohens (*Cohen's d*) fyrir breytingarskor. Áhrifastærð er í raun mælikvarði á áhrif frumbreytu á fylgibreytu og eru áhrifastærðir mjög mikið notaðar til að meta áhrif inngripa á heilsutengda þætti í klínískum rannsóknum. Með áhrifastærð er að auki lagt mat á hvort áhrif skipta máli, þ.e. hvort þau séu klínískt mikilvæg. Með áhrifastærðum eru breytingar metnar á staðlaðan hátt, sem gerir það kleift að bera saman niðurstöður sem byggjast á mismunandi mælingum (Kazis, Anderson og Meenan, 1989). Áhrifastærð var reiknuð með því að draga meðalskor fyrir meðferð frá meðalskori eftir meðferð (sem gefur breytingarskor) og deila með staðalfrávikum fyrir meðferð (Cohen, 1988). Sem dæmi þá gefur áhrifastærðin $d = -1,0$ til kynna að meðalskor hefur lækkað um sem nemur einu staðalfrávikum. Samkvæmt viðmiðum Cohens (1988) gefur áhrifastærð upp á 0,2 til kynna lítil áhrif, 0,5 eru miðlungs áhrif og 0,8 og yfir gefur til kynna mikil áhrif. Notað var parað t -próf (*paired sample t-test*) við að bera áhrifastærðir breytingarskors PSYCHLOPS saman við CORE-OM. Þegar skoðaðar voru dreifingar breytanna komu í ljós nokkur fráviksgildi (*outliers*), en þau geta ógnað forsendum t -prófa. En þegar úrtaksstærð er yfir 40 er talið að t -próf sé nógu traust (*robust*) gagnvart nokkrum fráviksgildum eða skekkju (Moore og McCabe, 2006).

Áreiðanleiki

Við mat á innri samkvæmni beggja mælitækja voru reiknaðir alfastuðlar (*Cronbach's alpha*), bæði fyrir og eftir meðferð. Fyrir skor á PSYCHLOPS voru alfastuðlar reiknaðir út

frá þeim fjórum atriðum sem heildarskor byggist á. Þar sem aðeins ein spurning liggur að baki bæði virkni- og vellíðunarþætti og tvær spurningar að baki vandamálþætti, er ekki við hæfi að leggja mat á innri áreiðanleika hvers þáttar fyrir sig. Fyrir skor á CORE-OM voru alfastuðlar reiknaðir fyrir heildarskor, bæði með og án atriða fyrir áhættuþátt. Að auki voru reiknuð öryggisbil fyrir alfastuðla beggja mælitækja, með aðferð sem lýst er í Feldt, Woodruff og Salih (1987). Þetta gerir samanburð á alfastuðlum réttmætari þar sem mikill munur er á fjölda atriða mælitækjanna, en alfastuðlar byggjast meðal annars á fjölda atriða mælitækis (Fan og Thompson, 2001; Feldt o.fl., 1987). Almennt er talið að alfastuðlar yfir 0,70 séu til marks um viðunandi innri samkvæmni, en þegar alfastuðlar eru komnir yfir 0,90 bendir það til of mikillar einsleitni atriða á mælitækinu (Nunnally og Bernstein, 1994).

Réttmæti

Við réttmætisprófanir á PSYCHLOPS var mat lagt á bæði samleitniréttmæti og samtímaréttmæti. Samleitniréttmæti var metið með því að skoða fylgni á milli áhrifastærða breytingarskors samsvarandi þátta og heildarskors á mælitækjunum tveimur, en eins og áður segir eru bæði mælitækin með vandamálþátt, virkniþátt og vellíðunarþátt. Ef góð fylgni er á milli breytingarskors mælitækjanna, og samsvarandi þátta innan þeirra, er það til marks um samleitniréttmæti mælitækisins (Aiken og Groth-Marnat, 2006; Haynes, 2006). Þegar breyturnar voru skoðaðar kom í ljós að nokkur fráviksgildi voru í dreifingum breytanna og var þess vegna ákveðið að reiknað fylgni með *Spearman's rho* frekar en *Pearson's r*. Þegar frávillingar eru til staðar getur notkun *Pearson's r* vanmetið verulega styrkleika sambands milli breyta, en *Spearman's rho* er hins vegar traust gagnvart þessu.

Samtímaréttmæti var metið með þremur mismunandi aðferðum. Í fyrsta lagi var samtímaréttmæti metið út frá því hvort samræmi væri á milli þess hvernig þátttakendur lýstu líðan sinni við lok meðferðar á PSYCHLOPS (spurning númer 6 á PSYCHLOPS) og þess hversu mikil breyting hafði orðið á skori þeirra við lok meðferðar. Með samræmi er átt við að skor þeirra sem segjast vera „miklu betri“ við lok meðferðar ætti að vera búið að taka meiri jákvæðum breytingum en hjá þeim sem segjast vera „mun betri“ og að sama skapi ætti skor þeirra sem segjast vera „mun betri“ að vera búið að taka meiri jákvæðum breytingum en hjá þeim sem segja vera „aðeins betri“ við lok meðferðar (Ashworth o.fl., 2005a; Ashworth, 2007c). Valið var að nota *t*-próf óháðra úrtaka (*independent sample t-test*) við mat á þessum tengslum, þar sem óháða breytan var lýsing þátttakenda á líðan sinni við lok meðferðar og háða breytan áhrifastærðir breytingarskors. Var *t*-prófið valið þrátt

fyrir nokkur fráviksgildi í dreifingum breytanna, en eins og komið hefur fram er *t*-próf nógu traust þrátt fyrir nokkur fráviksgildi þegar úrtaksstærð er yfir 40 (Moore og McCabe, 2006). Við mat á breytingarskori var notast við áhrifastærðir, en það var gert til að fá staðlaðan samanburð á milli mælitækjanna.

Samtímaréttmæti var einnig metið með því að athuga hversu mikil tengsl væru á milli heildarskors PSYCHLOPS og CORE-OM við upphaf meðferðar, við lok meðferðar og breytingarskors. Reiknuð var fylgni (*Spearman's rho*) við mat á tengslunum. Þetta er sambærileg aðferð og notuð var við mat á samleitniréttmæti.

Við mat á samtímaréttmæti voru tengsl á milli skors á PSYCHLOPS og CORE-OM könnuð með einni aðferð til viðbótar. Á CORE-OM er hægt að skilgreina klínískt skor, klínískt marktæka breytingu (*clinically significant change*) og áreiðanlega breytingu (*reliable change*) á skori (Evans o.fl., 1998). Klínískt marktæk breyting er þegar skor við upphaf meðferðar, sem einkennandi er fyrir klínískt þýði, breytist við lok meðferðar yfir í skor sem einkennandi er fyrir almennt þýði (Jacobson og Truax, 1991). Til að reikna klínískt marktæka breytingu þarf frávísunarviðmið (*cut-off*) fyrir klínískt skor og setja Jacobsons og Truax (1991) fram þrjú mismunandi viðmið til að reikna það út, viðmið A, B og C. Með viðmiði A er metið hvort sjúklingur sé við lok meðferðar kominn meira en tvö staðalfrávik frá meðalskori sem einkennandi er fyrir klínískt þýði og með viðmiði B er metið hvort sjúklingur sé innan við tvö staðalfrávik frá meðalskori sem einkennandi er fyrir almennt þýði. Viðmið C tekur mið af bæði viðmiði A og B og fellur mitt á milli þeirra. Til að reikna viðmið A þarf því einungis viðmið fyrir klínískt þýði og fyrir viðmið B þarf einungis viðmið fyrir almennt þýði, en viðmið C krefst hvoru tveggja. Í þessari rannsókn voru bæði viðmið fyrir klínískt og almennt þýði fyrir hendi, þar sem úrtakið í rannsókninni var notað fyrir klínískt þýði og var tekið mið af skori í almennu þýði úr rannsókn Evu Daggar Gylfadóttur (2007). Samkvæmt Jacobson og Traux (1991) er æskilegt að nota viðmið B eða C þegar viðmið fyrir skor í almennu þýði er fyrir hendi og ef dreifing skors fyrir klínískt og almennt þýði skarast er æskilegt að nota viðmið C. Skörun var á dreifingu skors fyrir klínískt og almennt þýði og var þar af leiðandi ákveðið að nota viðmið C til að reikna klínískt marktæka breytingu.

Áreiðanleg breyting er mælikvarði á hvort sú breyting sem hefur orðið á skori við lok meðferðar sé nógu mikil að ólíklegt er að breytingin sé tilkomin vegna óstöðugleika í mælingum (Jacobson og Truax, 1991). Sem mat á réttmæti PSYCHLOPS var skor á CORE-OM greint út frá þessum þáttum og borið saman við skor á PSYCHLOPS. Samtímaréttmæti var metið með því að athuga hvort þeir þátttakendur sem væru með

klínískt skor, þ.e. yfir frávísunarviðmiði, á CORE-OM væru með hærra skor á PSYCHLOPS en þeir sem væru með almennt skor á CORE-OM. Einnig var athugað hvort þátttakendur sem sýndu áreiðanlega og klínískt marktæka breytingu í skori á CORE-OM höfðu hærra breytingarskor en aðrir á PSYCHLOPS.

Niðurstöður

Samanburður á þeim sem luku meðferð við þá sem luku ekki meðferð

Af þeim 135 sjúklingum sem svöruðu spurningalistum við upphaf meðferðar var um helmingur þeirra sem lauk ekki meðferð. Til að skoða hvort munur væri á þeim sjúklingum sem luku meðferð og þeim sem ekki luku meðferð var upphafsskor þeirra borið saman, ásamt því að kanna hvort munur væri á aldri og kynjahlutfalli hópanna. Munur var metinn með því að reikna *t*-próf óháðra úrtaka. Þegar upphafsskor sjúklinga sem luku meðferð var borið saman við upphafsskor þeirra sem luku ekki meðferð reyndist hvorki munur á skori á PSYCHLOPS ($p = 0,14$) né CORE-OM ($p = 0,43$). Þetta bendir til þess að erfiðleikar þeirra sjúklinga sem luku ekki meðferð hafi verið sambærilegir erfiðleikum þeirra sem luku meðferð. Auk þess var ekki munur á aldri ($p = 0,50$) og kynjahlutfalli ($p = 0,51$) þeirra.

Tafla 1. Skor einstakra þátta og heildarskor á PSYCHLOPS og CORE-OM: Óstaðlað meðalskor og staðalfrávik ($N = 64$).

<i>Þættir</i>	PSYCHLOPS			CORE-OM		
	Fyrir meðferð M (<i>sf</i>)	Eftir meðferð M (<i>sf</i>)	Breyting M (<i>sf</i>)	Fyrir meðferð M (<i>sf</i>)	Eftir meðferð M (<i>sf</i>)	Breyting M (<i>sf</i>)
Vandamál	7,06 (2,14)	4,63 (2,46)	-2,43 (2,49)	1,98 (0,82)	1,54 (0,85)	-0,44 (0,64)
Virgni	3,33 (1,14)	2,08 (1,40)	-1,25 (1,39)	1,59 (0,84)	1,13 (0,77)	-0,46 (0,61)
Vellíðan	3,05 (1,09)	1,97 (1,26)	-1,08 (1,01)	2,00 (0,96)	1,41 (0,88)	-0,59 (0,75)
Áhætta	-	-	-	0,35 (0,53)	0,26 (0,56)	-0,09 (0,41)
Heildarskor (með áhættuþ.)	-	-	-	1,56 (0,71)	1,15 (0,71)	-0,41 (0,52)
Heildarskor (án áhættuþ.)	3,36 (0,94)	2,17 (1,08)	-1,19 (0,95)	1,82 (0,79)	1,34 (0,78)	-0,48 (0,60)

Skor á PSYCHLOPS og CORE-OM

Heildarskor þátttakenda á PSYCHLOPS fór úr 3,36 að meðaltali við upphaf meðferðar niður í 2,17 að meðaltali við lok meðferðar og varð lækkun á meðalskori á öllum þáttum PSYCHLOPS (sjá í töflu 1). Heildarskor þátttakenda á CORE-OM (án áhættuþátta) fór úr 1,82 að meðaltali við upphaf meðferðar niður í 1,34 að meðaltali við lok meðferðar og varð

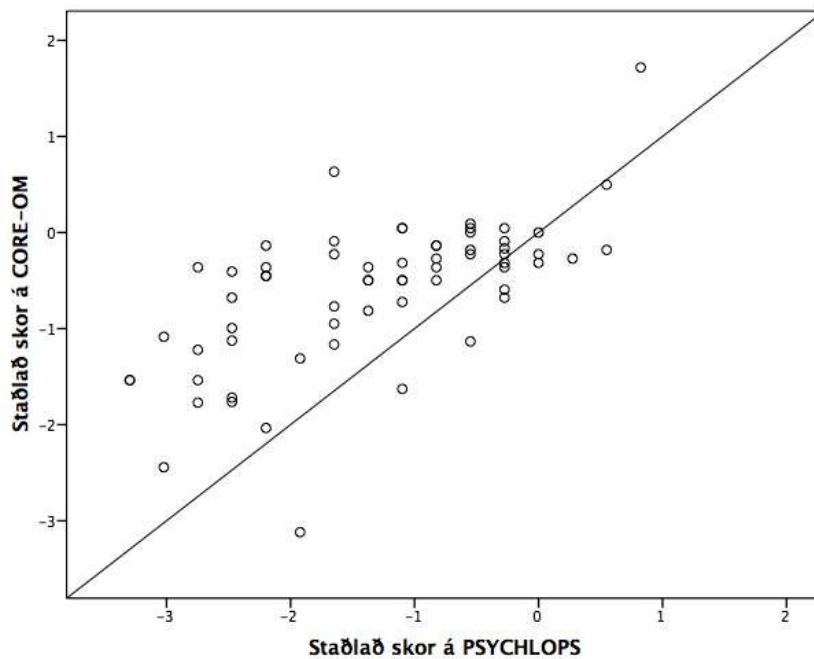
einnig lækkun á meðalskori allra þátta (sjá í töflu 1). Reiknað var út frávísunarviðmið fyrir klínískt skor á CORE-OM með aðferð Jacobsons og Truax (1991). Við upphaf meðferðar voru 76,6% þátttakenda með klínískt skor en 53,1% þátttakenda voru með klínískt skor við lok meðferðar. Viðmið fyrir áreiðanlega breytingu á skori var 0,49, en viðmiðið var reiknað út frá staðalfráviki skors fyrir meðferð (0,79) og áreiðanleikastuðli (0,95) (Evans o.fl., 1998; Jacobson og Truax, 1991). Ef skor hækkar eða lækkar um 0,49 eða meira er um að ræða áreiðanlega breytingu. Samkvæmt þessu viðmiði varð áreiðanleg breyting til batnaðar hjá 23 þátttakendum (35,9%), en tveimur þátttakendum (3,1%) hrakaði. Breytingar í skori hjá hinum 39 þátttakendum (60,9%) voru ekki nógu miklar til að teljast áreiðanlegar samkvæmt viðmiðinu.

Næmi fyrir breytingum

Við mat á næmi fyrir breytingum PSYCHLOPS voru áhrifastærðir breytingarskors bornar saman við CORE-OM og var reiknað parað t -próf til að bera saman einstaka þætti á milli mælitækjanna og heildarskor þeirra (sjá í töflu 2). Helstu niðurstöður voru þær að áhrifastærð heildar breytingarskors á PSYCHLOPS ($d = -1,31$) var mun stærri en áhrifastærð fyrir breytingarskor á CORE-OM ($d = -0,60$), ($t(63) = -7,101$, $p < 0,001$). Á Mynd 1 má sjá samanburð á dreifingu áhrifastærða fyrir skor á PSYCHLOPS og CORE-OM. Skor sem er við núll gefur til kynna að ekki var breyting á skori við lok meðferðar, en skor sem er í mínus gefur til kynna bata. Eins og sést á myndinni tók skor á PSYCHLOPS almennt meiri breytingum en skor á CORE-OM. Neðst í vinstra horni myndarinnar sést greinilega að hjá þeim þátttakendum sem náðu árangri í meðferð samkvæmt mælingum, mælist árangurinn meiri með PSYCHLOPS en með CORE-OM. Einnig voru áhrifastærðir fyrir þætti á PSYCHLOPS í öllum tilfellum marktækt stærri en áhrifastærðir fyrir samsvarandi þætti á CORE-OM, en niðurstöður marktæktarprófa má sjá nánar í töflu 2. Þegar nokkrir samanburðir eru reiknaðir á þennan hátt aukast líkurnar á höfnunarmistökum (*type I error*), þ.e. að rétttri núlltilgátu sé hafnað. Til að minnka líkurnar á að fram kæmi marktækur munur fyrir mistök í einhverjum á þeim samanburðum sem gerðir voru var reiknuð Bonferroni leiðrétting fyrir marga samanburði. Formúlan fyrir leiðréttinguna var alfa ($\alpha = 0,05$) deilt með fjölda samanburða ($k = 4$). Með Bonferroni leiðréttingu voru alfamörkin (α) lækkuð úr 0,05 í 0,0125. Þrátt fyrir leiðrétt alfamörk reyndist marktækur munur í öllum samanburðum.

Tafla 2. Samanburður á áhrifastærðum breytingarskors ($N = 64$).

Þættir	PSYCHLOPS	CORE-OM	Munur milli PSYCHLOPS og CORE-OM	
	Áhrifastærðir	Áhrifastærðir	Mismunur meðaltala (95% öryggisbil)	Niðurstöður marktektarprófa
Vandamál	M (sf) -1,20 (1,23)	M (sf) -0,55 (0,79)	-0,65 (-0,91, -0,39)	$t(63) = -5,070, p < 0,001$
Virkni	-1,17 (1,30)	-0,54 (0,71)	-0,63 (-0,95, -0,31)	$t(63) = -3,919, p = 0,001$
Vellíðan	-0,97 (0,91)	-0,62 (0,79)	-0,35 (-0,57, -0,13)	$t(63) = -3,233, p = 0,002$
Áhætta	-	-0,17 (0,76)	-	-
Heildarskor (án áhættup.)	-1,31 (1,04)	-0,60 (0,76)	-0,71 (-0,91, -0,51)	$t(63) = -7,101, p < 0,001$



Mynd 1. Samanburður á áhrifastærðum breytingarskors á PSYCHLOPS og CORE-OM.

Áreiðanleiki

Reiknaðir voru alfastuðlar (*Cronbach's alpha*) fyrir bæði mælitækin til að leggja mat á innri samkvæmni atriða. Fyrir PSYCHLOPS reyndist alfastuðullinn 0,85 (95% öryggisbil: 0,78-0,90) fyrir meðferð og 0,83 (95% öryggisbil: 0,75-0,89) eftir meðferð. Sem frekara mat á innri samkvæmni PSYCHLOPS var skoðuð fylgni (*Spearman's rho*) á milli spurninganna tveggja sem vandamálaþátturinn byggist á. Fylgni á milli spurninganna var

0,74 ($p < 0,001$) fyrir meðferð og 0,68 ($p < 0,001$) eftir meðferð. Þessar niðurstöður eru til marks um að innri samkvæmni atriða á PSYCHLOPS er viðunandi. Fyrir CORE-OM (alls 34 atriði) reyndist alfastuðullinn 0,95 (95% öryggisbil: 0,93-0,97) fyrir meðferð og 0,96 (95% öryggisbil: 0,94-0,97) eftir meðferð. Þegar áhættuþátturinn var ekki hafður með (alls 28 atriði) voru niðurstöðurnar þær sömu, 0,95 (95% öryggisbil: 0,94-0,97) fyrir meðferð og 0,96 (95% öryggisbil: 0,94-0,97) eftir meðferð.

Réttmæti

Við mat á samleitniréttmæti PSYCHLOPS var reiknuð fylgni á milli samsvarandi þátta mælitækjanna og heildar breytingarskors (sjá í töflu 3). Í ljós kom að há fylgni var á milli heildar breytingarskors mælitækjanna ($\rho = 0,64$). Miðlungs- eða há fylgni var milli samsvarandi þátta, en aftur á móti var fylgni milli ósamsvarandi þátta álíka mikil.

PSYCHLOPS leggur ekki mat á áhættu og hafði breytingarskor á PSYCHLOPS mjög lága fylgni við áhættuþáttinn á CORE-OM.

Í ljósi þess hversu há fylgni var á milli ósamsvarandi þátta á mælitækjunum var ákveðið að skoða fylgni á milli þátta á hvoru mælitæki fyrir sig. Ef há fylgni er á milli þátta innan mælitækis verður aðgreining þáttanna veikari. Niðurstöður má sjá í töflu 4, en eins og sést er nokkur fylgni á milli þátta innan beggja mælitækja. Fylgni á milli þátta innan CORE-OM er þó töluvert hærri en á milli þátta innan PSYCHLOPS.

Tafla 3. Fylgni á milli staðlaðs breytingarskors PSYCHLOPS og CORE-OM ($N = 64$).

PSYCHLOPS	CORE-OM				
	Vandamál	Virgni	Vellíðan	Áhætta	Heildarskor (án áhættuþ.)
Vandamál	0,59**	0,56**	0,42**	0,29*	0,61**
Virgni	0,48**	0,37**	0,20	0,28*	0,41**
Vellíðan	0,49**	0,53**	0,44**	0,14	0,55**
Heildarskor	0,64**	0,60**	0,41**	0,34**	0,64**

** Spearman's rho fylgnistuðlar marktækir, miðað við $p < 0,01$.

* Spearman's rho fylgnistuðlar marktækir, miðað við $p < 0,05$.

Við mat á samtímaréttmæti PSYCHLOPS voru könnuð tengsl á milli skors þátttakenda á PSYCHLOPS og CORE-OM með því að reikna fylgni á milli heildarskors fyrir meðferð, eftir meðferð og breytingarskors. Fylgni á milli mælitækjanna var í öllum tilvikum sterk, eða 0,66 (*rho*) fyrir meðferð, 0,76 (*rho*) eftir meðferð og var fylgnin milli breytingarskors 0,64 (*rho*). Við upphaf meðferðar, voru þátttakendur sem voru með klínískt skor á CORE-OM, þ.e. skor yfir frávísunarviðmiði, með meðalskor 3,60 (hráskor) á PSYCHLOPS. Þeir sem voru með almennt skor á CORE-OM voru hins vegar með mun lægra skor á PSYCHLOPS við upphaf meðferðar, eða 2,58 ($t(62) = -4,110, p < 0,001$). Við lok meðferðar, voru þátttakendur sem voru með klínískt skor á CORE-OM með meðalskorið 2,79 á PSYCHLOPS, en þeir sem voru með almennt skor á CORE-OM voru með meðalskorið 1,46 á PSYCHLOPS ($t(62) = -6,229, p < 0,001$). Þátttakendur sem sýndu áreiðanlega breytingu úr klínísku skori í almennt skor á CORE-OM við lok meðferðar, þ.e. klínískt marktæka breytingu, voru með mun hærra breytingarskor á PSYCHLOPS (-2,19), samanborið við aðra þátttakendur (-0,96), $t(62) = -4,661, p < 0,001$).

Tafla 4. Fylgni á milli þátta innan PSYCHLOPS og innan CORE-OM

Þættir	PSYCHLOPS			CORE-OM			Áhætta
	Vandamál	Virkni	Vellíðan	Vandamál	Virkni	Vellíðan	
Vandamál	-	0,52**	0,30**	-	0,69**	0,60**	0,32*
Virkni	0,52**	-	0,35**	0,69**	-	0,55**	0,27*
Vellíðan	0,30**	0,35**	-	0,60**	0,55**	-	0,07

** Spearman's rho fylgnistuðlar marktækir, miðað við $p < 0,01$.

* Spearman's rho fylgnistuðlar marktækir, miðað við $p < 0,05$.

Við mat á samtímaréttmæti PSYCHLOPS voru einnig skoðuð svör þátttakenda við því hvernig þeir lýstu líðan sinni við lok meðferðar (spurning númer 6 á PSYCHLOPS). Svör við spurningunni dreifðust lítið á milli svarmöguleika (sjá töflu 4), nokkrir þátttakendur sögðust vera „miklu betri“ við lok meðferðar, en flestir sögðust annað hvort vera „mun betri“ eða „aðeins betri“. Fáir sögðu að líðan væri óbreytt eða að henni hafi farið aftur við lok meðferðar. Gert var *t*-próf óháðra úrtaka til að athuga hvort samræmi væri á milli breytingarskors þátttakenda og líðanar þeirra við lok meðferðar. Út frá skori á

PSYCHLOPS, voru áhrifastærðir þeirra þátttakenda sem sögðu að líðan þeirra væri „miklu betri“ við lok meðferðar marktækt stærri en þeirra sem sögðu að líðan væri „mun betri“ ($t(32) = -2,215, p = 0,034$), eða „aðeins betri“ ($t(33) = -2,564, p = 0,015$), við lok meðferðar. Hins vegar var ekki munur á áhrifastærðum þeirra sem sögðu að líðan væri „mun betri“ eða „aðeins betri“ ($p > 0,05$). Sambærilegar niðurstöður komu í ljós þegar skoðaðar voru áhrifastærðir þátttakenda á CORE-OM, þar sem áhrifastærðir þeirra sem sögðust vera „miklu betri“ voru marktækt stærri en þeirra sem sögðu við lok meðferðar að líðan þeirra væri „mun betri“ ($t(32) = -2,318, p = 0,027$), eða „aðeins betri“ ($t(33) = -3,223, p = 0,003$).

Tafla 5. Samband á milli breytingarskors og mats þátttakenda á eigin bata við lok meðferðar

Líðan við lok meðferðar ($N = 64$)	n	PSYCHLOPS	CORE-OM
		Áhrifastærðir (<i>Cohen's d</i>)	Áhrifastærðir (<i>Cohen's d</i>)
		M (<i>sf</i>)	M (<i>sf</i>)
Miklu betri	12	-2,04 (1,10)	-1,19 (0,71)
Mun betri	22	-1,22 (0,98)	-0,46 (0,95)
Aðeins betri	23	-1,21 (0,80)	-0,52 (0,51)
Óbreytt líðan	5	-0,99 (1,51)	-0,33 (0,23)
Aðeins verri	2	0,14 (0,19)	-0,25 (0,03)
Miklu verri	0	-	-

Umræða

Megintilgangur þessarar rannsóknar var að meta próffræðilega eiginleika íslenskrar þýðingar PSYCHLOPS. PSYCHLOPS er notendamiðað mælitæki sem ætlað er til notkunar í mati á árangri sálfræðilegrar meðferðar og var það borið saman við þýðismiðað mælitæki sem einnig er notað í árangursmati, CORE-OM. Sambærileg rannsókn hefur áður verið gerð af Ashworth o.fl. (2005a) þar sem þessi sömu mælitæki voru borin saman, en höfundar PSYCHLOPS telja mikilvægt að niðurstöður þeirra séu endurteknar af aðilum sem ótengdir eru upprunalegri þróun mælitækisins (Ashworth, 2007d).

Samantekt á helstu niðurstöðum

PSYCHLOPS reyndist vera mjög næmt fyrir breytingum við lok meðferðar. Reiknaðar voru áhrifastærðir fyrir breytingarskor beggja mælitækja til að fá staðlaðan samanburð milli mælitækjanna. Breytingarskor á PSYCHLOPS reyndist töluvert herra en breytingarskor á CORE-OM og var munurinn marktækur. Niðurstöðurnar eru í samræmi við niðurstöður Ashworth o.fl. (2005a) þar sem áhrifastærðir fyrir PSYCHLOPS voru marktækt stærri en áhrifastærðir fyrir CORE-OM. Áhrifastærð breytingarskors á PSYCHLOPS var 1,31 sem er aðeins minna en áhrifastærðir í fyrri rannsóknum, 1,53 í rannsókn Ashworth o.fl. (2005a) og 1,61 í rannsókn Ashworth o.fl. (2009). Aftur á móti kom á óvart að áhrifastærð breytingarskors á CORE-OM var ekki nema 0,60, sem er talsvert lægra en í rannsókn Ashworth o.fl. (2005a) þar sem áhrifastærðin var 1,06. Meðferðin skilaði samt sem áður töluverðum árangri, en þegar áhrifastærðir eru 0,8 eða stærri eru áhrif meðferðar talin vera mikil (Kazis o.fl., 1989). Nánari umræða um þennan mun á áhrifastærðum má sjá á næstu síðu.

Niðurstöður í mati á áreiðanleika gaf til kynna að PSYCHLOPS hefur viðunandi innri samkvæmni. Alfastuðlar (*Cronbach's alpha*) voru 0,85 fyrir meðferð og 0,83 eftir meðferð sem er sambærilegt niðurstöðum Ashworth o.fl. (2005a) (0,79; 0,87) og Ashworth o.fl. (2009) (0,75; 0,83). Mat á innri samkvæmni CORE-OM kom einnig vel út og voru alfastuðlarnir nánast þeir sömu og í rannsókn Ashworth o.fl. (2005a) og rannsókn Evans o.fl. (2002). Lágur áreiðanleiki er vandamál sem hefur fylgt einstaklingsmiðuðum mælitækjum (Haynes og O'Brien, 2000), en það virðist hins vegar ekki eiga við PSYCHLOPS.

Stuðningur við samleitniréttmæti PSYCHLOPS kom fram með sterkri fylgni breytingarskors við breytingarskor á CORE-OM. Mest var fylgnin á milli heildar

breytingarskors mælitækjanna, en einnig var góð fylgni á milli samsvarandi þátta. Ekki tókst að sýna fram á aðgreiningarréttmæti þar sem fylgni milli ósamsvarandi þátta var sambærileg þeirri fylgni sem var á milli samsvarandi þátta. Þó hafði heildarskor og skor einstakra þátta á PSYCHLOPS litla fylgni við áhættuþátt á CORE-OM. Það gefur vísbendingu um aðgreiningarréttmæti þar sem PSYCHLOPS er ekki ætlað að meta áhættu sérstaklega. Þegar könnuð var fylgni á milli þátta innan beggja mælitækja kom í ljós að sterk fylgni var á milli þátta innan CORE-OM. Nokkuð há fylgni var á milli þátta innan PSYCHLOPS en þó töluvert lægri en innan CORE-OM. Þessar niðurstöður benda til að þættir á CORE-OM, að undanskildum áhættuþætti, eru of tengdir til að hægt sé að meta aðgreiningarréttmæti út frá þeim eins og hér var reynt að gera. Lág fylgni áhættuþáttar við aðra þætti innan CORE-OM getur til kynna góða aðgreiningu þáttarins gagnvart öðrum þáttum. Þessar niðurstöður eru í samræmi við niðurstöður Evans o.fl. (2002) og Lyne o.fl. (2006) um veika aðgreiningu á milli vandamála-, virkni- og vellíðunarþáttar á CORE-OM. Sambærilegar niðurstöður komu einnig fram í rannsókn Evu Daggar Gylfadóttur (2007) á íslenskri þýðingu CORE-OM. Þáttagreining í rannsókn Lyne o.fl. (2006) bendir til að best sé að reikna skor á CORE-OM á tveimur aðskildum þáttum. Annars vegar að reikna skor fyrir áhættuþátt og hins vegar reikna skor fyrir sálræna erfiðleika sem einn þátt, sem er samanlagt skor hinna þriggja þáttanna.

Há fylgni skors á PSYCHLOPS við skor á CORE-OM var til marks um samtímaréttmæti mælitækisins. Jafnframt voru þátttakendur sem voru með klínískt skor á CORE-OM, þ.e. skor yfir frávísunarviðmiði, með marktækt hærra skor á PSYCHLOPS og þeir sem sýndu áreiðanlega og klínískt marktæka breytingu í skori á CORE-OM við lok meðferðar sýndu einnig marktækt meiri breytingu í skori á PSYCHLOPS. Stuðningur við samtímaréttmæti PSYCHLOPS kom einnig fram með marktækt hærra breytingarskori hjá þeim sem sögðu við lok meðferðar að líðan þeirra væri miklu betri, en hjá þeim sem sögðu líðan vera mun eða aðeins betri. Sami stuðningur við samtímaréttmæti mælitækisins kom fram í niðurstöðum Ashworth o.fl. (2005a) og Ashworth o.fl. (2009).

Próffræðilegir eiginleikar íslenskrar þýðingar PSYCHLOPS koma almennt vel út og eru í góðu samræmi við niðurstöður fyrri rannsókna á mælitækinu (Ashworth o.fl., 2005a; Ashworth o.fl., 2009). Helsti munurinn á þessari rannsókn og fyrri rannsóknum var að áhrifastærðir fyrir bæði mælitækin voru nokkuð minni í þessari rannsókn, en minni áhrifastærðir gefa til kynna að meðferðin var að skila minni árangri. Líklegasta skýringin á þessum mun í áhrifastærðum er mismunandi meðferðarform. Í þessari rannsókn fengu þátttakendur hugræna atferlismeðferð í hóp, en í rannsóknum Ashworth o.fl., 2005 og

Ashworth o.fl. (2009) var veitt einstaklingsmeðferð í heilsugæslu. Gera má ráð fyrir að þátttakendur í rannsóknum þeirra hafi þar af leiðandi fengið meðferð sem sniðin var að vandamálum hvers og eins og stuðlað þannig að betri árangri.

Með niðurstöðum þessarar rannsóknar er kominn aukinn stuðningur við að PSYCHLOPS hafi ýmsa jákvæða eiginleika til notkunar við árangursmat sálfræðilegrar meðferðar. Mælitækið er mjög næmt fyrir breytingum og hefur viðunandi próffræðilega eiginleika (Ashworth o.fl., 2005a; Ashworth o.fl., 2009). Meðferðaraðilar hafa sagt mælitækið hafi ýmsa eftirsóknarverða kosti sem gefa til kynna gott sýndarréttmæti. Meðal annars að PSYCHLOPS sé alveg jafn viðeigandi og þýðismiðuð mælitæki við að meta árangur og að það gefi upplýsingar um hvernig fólk skilgreinir sjálft erfiðleika sína sem síðan er hægt að nýta í meðferðinni. Jafnframt að eigindlegar upplýsingar sem fengnar eru með PSYCHLOPS séu góð viðbót við megindlegar upplýsingar sem fást með þýðismiðuðum mælitækjum eins og CORE-OM og þar af leiðandi sé hentugt að nota þessi mælitæki saman (Ashworth o.fl., 2005b).

Takmarkanir á notkun PSYCHLOPS

Þrátt fyrir ýmsa jákvæða eiginleika PSYCHLOPS, þá eru ákveðnar takmarkanir sem fylgja notkun mælitækisins sem mikilvægt er að gera grein fyrir. Eins og fjallað var um í inngangi þá henta einstaklingsmiðuð mælitæki, eins og PSYCHLOPS, ekki jafn vel og þýðismiðuð mælitæki til rannsókna þar sem einhver samiburður á fólki fer fram (Haynes og O'Brien, 2000). Einstaklingsmiðuð mælitæki eru yfirleitt þannig uppbyggð að þau henta best við að meta breytingar hjá einstaklingum, en ekki mun á milli þeirra. Til dæmis eru engin viðmið fyrir svör á PSYCHLOPS, sem gerir það að verkum að erfitt er að túlka hvaða þýðingu upphafsskor hefur fyrir próftakann. PSYCHLOPS metur einungis einstaklingsbundna erfiðleika, sem eru einstakir fyrir hvern og einn, en ekki erfiðleika sem mældir eru með stöðluðum spurningum líkt og þýðismiðuð mælitæki gera. Því er erfitt að bera saman skor tveggja eða fleiri próftaka. Þar af leiðandi er ekki hægt að nota PSYCHLOPS til að úrskurða um hvort skor sé yfir klínískum viðmiðum, né til að þróa viðmiðsskor til að greina á milli klínískra og heilbrigðra hópa, og ekki er hægt að leggja mat á alvarleika þeirra erfiðleika sem sjúklingurinn er að eiga við. Aðeins er hægt að nota þýðismiðuð mælitæki til þess, líkt og CORE-OM. Þessar takmarkanir eiga samt ekki að hafa áhrif á notagildi PSYCHLOPS. En eins og fram hefur komið þá vilja höfundar mælitækisins ekki að það

komi í stað annarra mælinga í árangursmati, heldur að það verði notað sem viðbót við aðrar hefðbundnari mælingar (Ashworth o.fl., 2007a; Ashworth, 2007c).

Núverandi staða PSYCHLOPS og frekari rannsóknir

Nú hafa próffræðilegir eiginleikar PSYCHLOPS verið skoðaðir út frá samanburði við tvö þýðismiðuð mælitæki, CORE-OM og HADS. Einnig hefur sýndarréttmæti verið metið út frá reynslu meðferðaraðila af notkun mælitækisins. Góður stuðningur er við próffræðilega eiginleika PSYCHLOPS en þörf er á frekari rannsóknum. Mikilvægt er að niðurstöður þessara rannsókna séu endurteknar af óháðum aðilum, en með þessari rannsókn hefur það nú að hluta til verið gert. Búið er á safna gögnum fyrir mat á áreiðanleika endurtekinna prófunar (*test-retest reliability*) út frá úrtaki háskólanemenda, en eftir er að gefa niðurstöðurnar út (Ashworth, 2007d). Hugmyndir eru um að skoða hugsníðaréttmæti mælitækisins nánar með því að meta hvort samræmi sé á milli mats meðferðaraðila og mats sjúklingsins um batann sem sjúklingurinn hefur náð að lokinni meðferð (Ashworth o.fl., 2005a).

Gildi niðurstaða fyrir árangursmat sálfræðilegrar meðferðar

Aukin áhersla er á að meðferðaraðilar fylgist með og skrái niður áhrif þeirrar meðferðar sem þeir veita (Groth-Marnat, 2003). Mikilvægt er að notaðar séu aðferðir við mælingar sem bæði eru réttmætar og áreiðanlegar, en velja þarf viðeigandi mælitæki hverju sinni. Algeng viðmið fyrir val á mælitækjum, auk viðunandi próffræðilegra eiginleika, eru að þau séu auðveld í fyrirlögn og túlkun, og næm fyrir breytingum í líðan (Antony og Barlow, 2002). Einstaklingsmiðuð mælitæki hafa ýmsa eiginleika fram yfir þýðismiðuð mælitæki, sem geta að sama skapi gert þau afar hentug til að mæla áhrif meðferðar á fólk (Ashworth o.fl., 2005a; Ashworth o.fl., 2009; Nelson-Gray, 1996). Aftur á móti hafa þau oft verið gagnrýnd fyrir að vera óáreiðanleg, flókin og tímafrek í fyrirlögn, og þar af leiðandi lítið verið notuð við árangursmat (Fitzpatrick o.fl., 1998; Nelson-Gray, 1996). Þessi gagnrýni á hins vegar ekki við PSYCHLOPS, þar sem það tekur mjög stuttan tíma að leggja það fyrir, það er auðvelt í túlkun og hefur viðunandi próffræðilega eiginleika (Ashworth o.fl., 2005a; Ashworth o.fl., 2005b, Ashworth o.fl., 2009).

Sýnt hefur verið að PSYCHLOPS nær að draga fram ýmsar upplýsingar um erfiðleika fólks sem hefðbundið þýðisbundin mælitæki mæla ekki (Ashworth o.fl., 2007a). Þegar spurt er á PSYCHLOPS hvaða vandamál valdi mestum erfiðleikum er mjög algengt

að þau séu tengd félagslegum þáttum, eins og til dæmis erfiðleikar í samskiptum (Ashworth o.fl., 2007a; Robinson, Ashworth, Shepherd og Evans, 2006), en þýðismiðuð mælitæki mæla yfirleitt ekki beint þá þætti (Froyd o.fl., 1996). Með þýðismiðuðum mælitækjum er þar af leiðandi líkur á að farið sé á mis við atriði sem skipta máli (Ashworth o.fl. 2007a; Hunter o.fl., 2004). PSYCHLOPS getur því reynst hentugt mælitæki við að ná fram mikilvægum atriðum sem annars yrðu sniðgengin og stuðlað þannig að því að mat á árangri verði betur sniðið að þeim erfiðleikum sem sjúklingur telur skipta mestu máli fyrir sig. Auk þess geta upplýsingar um hvernig sjúklingur skilgreinir sinn eigin vanda skipt máli. Samkvæmt Macran o.fl. (1999) er nauðsynlegt að ná fram, viðurkenna og nota það hvernig sjúklingur skilgreinir sinn eigin vanda, en öðruvísi sé erfitt að stuðlað að réttmætum mælingum á áhrifum meðferðar.

Þróunin innan sálfræðinnar hefur verið á þá leið að þýðismiðuð mælitæki eru mjög mikið notuð ein og sér við mat á árangri, óháð því hver geðgreining eða meðferðarmarkmið sjúklinga er. Mikil þörf er á að gæði verkferla séu aukin með meiri áherslu á einstaklingsmiðaðar mælingar til að nota samhliða þýðismiðuðum mælingum og þannig bæta upp fyrir þær takmarkanir sem þýðismiðaðar mælingar hafa. Þannig verða meiri líkur á að þau vandamál sem mestu máli skipta fyrir sjúklinginn séu mæld og þá hvaða áhrif meðferð hefur á þessi tilteknu vandamál (Nelson-Gray, 1996). Notendamiðaða mælitækið PSYHLOPS er auðvelt í fyrirlögn, hefur viðunandi próffræðilega eiginleika og er mjög næmt fyrir breytingum. Út frá þessum eiginleikum mælitækisins og umfjölluninni hér að ofan má álykta að PSYCHLOPS gæti orðið góð viðbót við þær aðferðir sem nú þegar eru notaðar við árangursmat sálfræðilegrar meðferðar.

Notkun PSYCHLOPS á Íslandi

Nú þegar PSYCHLOPS hefur verið þýtt yfir á íslensku er mælitækið tilbúið til frekari prófunar á Íslandi. Réttmætisprófanir á mælitækinu, með samanburði við CORE-OM sem nú þegar er í notkun hér á landi, koma vel út og eru í samræmi við niðurstöður erlendra rannsókna (Ashworth o.fl., 2005a; Ashworth o.fl., 2009). Skilyrði fyrir því að mæling geti verið réttmæt er viðunandi áreiðanleiki og þær kröfur uppfyllir íslensk þýðing PSYCHLOPS. PSYCHLOPS er ef til vill tilbúið til notkunar hér á landi en þó er þörf á frekari rannsóknum á íslenskri þýðingu mælitækisins. Meðal annars þarf að gera rannsókn þar sem mælitækið er lagt fyrir í einstaklingsmeðferð, en mælitækið var upphaflega hannað til notkunar í einstaklingsmeðferð og nýtist vafalaust best til notkunar þar.

Heimildir

- Aiken, L. R. og Groth-Marnat, G. (2006). *Psychological testing and assessment* (12. útg.). Boston: Allyn & Bacon.
- Allport, G. W. (1937). *Personality: A psychological interpretation*. New York: Holt.
- Antony, M. M. og Barlow, D. (2002). *Handbook of assessment, treatment planning, and outcome evaluation: Empirically supported strategies for psychological disorders*. New York: Guilford Press.
- Ashworth, M., Shepherd, M., Christey, J., Matthews, V., Wright, K., Parmentier, H., Robinson, S. og Godfrey, E. (2004). A client-generated psychometric instrument: The development of 'PSYCHLOPS'. *Counselling and Psychotherapy Research*, 4, 27-33.
- Ashworth, M., Robinson, S., Godfrey, E., Shephard, M., Evans, C., Seed, P., Parmentier, H. og Tylee, A. (2005a). Measuring mental health outcomes in primary care: The psychometric properties of a new patient-generated outcome measure, 'PSYCHLOPS' ('psychological outcome profiles'). *Primary Care Mental Health*, 3, 261-270.
- Ashworth, M., Robinson, S., Godfrey, E., Parmentier, H., Shephard, M., Christey, J., Wright, K. og Matthews, V. (2005b). The experiences of therapists using a new client-centred psychometric instrument, PSYCHLOPS (Psychological Outcome Profiles). *Counselling and Psychotherapy Research*, 5, 37-42.
- Ashworth, M., Robinson, S., Evans, C., Shephard, M., Conolly, A. og Rowlands, G. (2007a). What does an idiographic measure (PSYCHLOPS) tell us about the spectrum of psychological issues and scores on a nomothetic measure (CORE-OM)? *Primary Care and Community Psychiatry*, 12, 7-16.
- Ashworth, M. (2007b). PSYCHLOPS – a psychometric outcome measure that is finding a niche. *Counselling and Psychotherapy Research*, 7, 201-202.

- Ashworth, M. (2007c). *PSYCHLOPS ('Psychological Outcome Profiles')*. Sótt 20. mars 2010, af <http://www.psychlops.info/>.
- Ashworth, M. (2007d). PSYCHLOPS: A patient-centred outcome measure. *Progress in Neurology and Psychiatry, 1*, 37-41.
- Ashworth, M., Evans, C. og Clement, S. (2009). Measuring psychological outcomes after cognitive behaviour therapy in primary care: A comparison between a new patient-generated measure “PSYCHLOPS” (Psychological Outcome Profiles) and “HADS” (Hospital Anxiety and Depression Scale). *Journal of Mental Health, 18*, 169-177.
- Battle, C., Imber, S., Hoehn-Sario, A., Nash, E. og Frank, J. (1966). Target complaints as criteria of improvement. *American Journal of Psychotherapy, 20*, 184-192.
- Beck, A. T. (1979). *Cognitive therapy and the emotional disorders*. New York: Penguin Books.
- Beck, A. T., Steer, R. A. og Brown, G. K. (1996). *Beck Depression Inventory- Second Edition manual*. San Antonio: Psychological Corporation.
- Beck, A. T., Ward, C. H., Mendelson, M., Mock, J. og Erbaugh, J. (1961). An inventory for measuring depression. *Archives of General Psychiatry, 4*, 561-571.
- Bilsbury, C. D. og Richman, A. (2002). A staging approach to measuring patient-centred subjective outcomes. *Acta Psychiatrica Scandinavica, Supplementum, 104*, 5-40.
- Bond, G., Bloch, S. og Yalom, I. (1979). The evaluation of a “Target Problem” approach to outcome measurement. *Psychotherapy: Theory and Research Practice, 16*, 48-54.
- Callaghan, G. M. (2001). Demonstrating clinical effectiveness for individual practitioners and clinics. *Professional Psychology: Research and Practice, 32*, 289-297.

- Chalkley, A. J. og Mulhall, D. J. (1991). The PQRSTUV: The Personal Questionnaire Rapid Scaling Technique--'Ultimate Version'. *British Journal of Clinical Psychology*, 30, 181-183.
- Clark, A., Hook, J. og Stein, K. (1997). Counsellors in primary care in Southampton: A questionnaire survey of their qualifications, working arrangement, and casemix. *British Journal of General Practice*, 47, 613-617.
- Clement, P. W. (1996). Evaluation in private practice. *Clinical Psychology: Science and Practice*, 3, 146-159.
- Cohen, J. (1988). *Statistical Power Analysis for the Behavioural Science* (2. útg.). New York: Academic Press.
- Cone, J. D. (1986). Idiographic, nomothetic and related perspectives in behavioral assessment. Í R. O. Nelson og S. C. Hayes (Ritstj.), *Conceptual foundations of behavioral assessment*. New York: Guilford Press.
- CORE System Group (1998). *CORE System (Information Management) Handbook*. Leeds: CORE System Group.
- Deyo, R. A. og Inui, T. S. (1984). Toward clinical applications of health status measures: Sensitivity of scales to clinically important changes. *Health Services Research*, 19, 275-289.
- Donnelly, C. og Carswell, A. (2002). Individualized outcome measures: A review of the literature. *The Canadian Journal of Occupational Therapy*, 69, 84-94.
- Eva Dögg Gylfadóttir (2007). *CORE-OM árangursmælitækið: Próffræðilegir eiginleikar og tengsl við persónuleikaraskanir*. Óbirt Cand.psych ritgerð: Háskóli Íslands, Félagsvísindadeild.

- Evans, C., Connell, J., Barkham, M., Margison, F., McGrath, G., Mellor-Clark, J. og Audin, K. (2002). Towards a standardised brief outcome measure: Psychometric properties and utility of the CORE-OM. *British Journal of Psychiatry*, 180, 51-60.
- Evans, C., Mellor-Clark, J., Margison, F., Barkham, M., Audin, K., Connell, J. o.fl. (2000). CORE: Clinical Outcomes in Routine Evaluation. *Journal of Mental Health*, 9, 247-255.
- Evans, C., Margison, F. og Barkham, M. (1998). The contribution of reliable and clinically significant change methods to evidence-based mental health. *Evidence Based Mental Health*, 1, 70-72.
- Fan, X. og Thompson, B. (2001). Confidence intervals about score reliability coefficients, please: An EPM guidelines editorial. *Educational and Psychological Measurement*, 61, 517-531.
- Feldt, L. S., Woodruff, D. J. og Salih, F. A. (1987). Statistical inference for coefficient alpha. *Applied Psychological Measurement*, 11, 93-103.
- Fiske, D. W., Hunt, H. F., Luborsky, L., Orne, M. T., Parloff, M. B., Reiser, M. F. o.fl. (1970). *Planning of research on effectiveness of psychotherapy*, 25, 727-737.
- Fitzpatrick, R., Davey, C., Buxton, M. og Jones, D. R. (1998). Evaluating client-based outcome measure for use in clinical trials. *Health Technology Assessment*, 2, 1-74.
- Froyd, J. E., Lambert, M. J. og Froyd, J. D. (1996). A review of psychotherapy outcome measurement. *Journal of Mental Health*, 5, 11-15.
- Groth-Marnat, G. (2003). *Handbook of psychological assessment* (4. útg.). New York: John Wiley and Sons.
- Hamilton, M. (1960). A rating scale for depression. *Journal of Neurology, Neurosurgery and Psychiatry*, 23, 56-62.

- Haynes, S. N. (2006). Psychometric considerations. Í M. Hersen (Ritstj.), *Clinician's handbook of adult behavioral assessment*. London: Elsevier Academic Press.
- Haynes, S. N. og O'Brien, W. H. (2000). *Principles and practice of behavioral assessment*. New York: Kluwer Academic/Plenum Publishers.
- Howell, D. C. (2007). *Statistical methods for psychology* (6. útg.). Belmont: Thomson Wadsworth.
- Hunter, R., McLean, J., Peck, D., Pullen, I., Greenfield, A., McArthur, W. o.fl. (2004). The Scottish 700 outcomes study: A comparative evaluation of the Health of the Nation Outcome Scale (HoNOS), the Avon Mental Health Measure (AVON), and an Idiographic Scale (OPUS) in adult mental health. *Journal of Mental Health, 13*, 93-105.
- Hurlburt, R. T. og Knapp, T. J. (2006). Münsterberg in 1898, not Allport in 1937, introduced the terms 'idiographic' and 'nomothetic' to American psychology. *Theory and Psychology, 16*, 287-293.
- Jacobson, N. S. og Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology, 59*, 12-19.
- Kazis, L., Anderson, J. og Meenan, R. (1989). Effect sizes for interpreting changes in health status. *Medical Care, 27*, 178-189.
- Kiresuk, T. J. og Sherman, R. E. (1968). Goal attainment scaling: A general method for evaluating comprehensive community mental health programs. *Community Mental Health Journal, 4*, 443-453.
- Lambert, M. J., Hatch, D. R., Kingston, M. D. og Edwards, B. C. (1986). Zung, Beck, and Hamilton rating scales as measures of treatment outcome: A meta-analytic comparison. *Journal of Consulting and Clinical Psychology, 54*, 54-59.

- Lambert, M. J. og Hawkins, E. J. (2004). Use of psychological tests for assessing treatment outcome. Í M. E. Maruish (Ritstj.), *The use of psychological testing for treatment planning and outcome assessment*. New Jersey: Lawrence Erlbaum Associates.
- Lyne, K. J., Barrett, P., Evans, C. og Barkham, M. (2006). Dimensions of variation on the CORE-OM. *British Journal of Clinical Psychology*, 45, 185-203.
- Macran, S., Ross H., Hardy, G. E. og Shapiro, D. A. (1999). The importance of considering clients' perspectives in psychotherapy research. *Journal of Mental Health*, 8, 325-337.
- Mead, N. og Bower, P. (2000). Patient-centredness: A conceptual framework and review of the empirical literature. *Social Science and Medicine*, 51, 1087-1110.
- Mirin S. M. og Namerow, M. J. (1991). Why study treatment outcome? *Hospital and Community Psychiatry*, 42, 1007-1013.
- Meier, S. T. (2008). *Measuring change in counseling and psychotherapy*. New York: Guilford Press.
- Moore, D. S. og McCabe, G. P. (2006). *Introduction to the practice of statistics* (5. útg.). New York: W. H. Freeman and Company.
- Nelson-Gray, R. O. (1996). Treatment outcome measures: Nomothetic or idiographic? *Clinical Psychology: Science and Practice*, 3, 164-167.
- Nunnally, J. og Bernstein, I. (1994). *Psychometric theory* (3. útg.). New York: McGraw-Hill.
- Panzarino, P. (1995). Behavioral healthcare report card indicators. *Behavioral Healthcare Tomorrow*, 4, 49-54.

- Paterson, C. (1996). Measuring outcome in primary care: A patient-generated measure, MYMOP, compared to the SF-36 health survey. *British Medical Journal*, *312*, 1016-1020.
- Paterson, C., Langan, C. E., McKaig, G. A., Anderson, P. M., Maclaine, G. D. H., Rose, L. B. o.fl. (2000). Assessing patient outcomes in acute exacerbations of chronic bronchitis: The measure your medical outcome profile (MYMOP), medical outcomes study 6-item general health survey (MOS-6A) and EuroQol (EQ-5D). *Quality of Life Research*, *9*, 521-527.
- Persons, J. B. (1991). Psychotherapy outcome studies do not accurately represent current models of psychotherapy: A proposed remedy. *American Psychologist*, *46*, 99-106.
- Phillips, J. P. N. (1970). A new type of personal questionnaire. *British Journal of Social and Clinical Psychology*, *9*, 241-256.
- Robinson, S., Ashworth, M., Shepherd, M. og Evans, C. (2006). In their own words: A narrative-based classification of clients' problems on an idiographic outcome measure for talking therapy in primary care. *Primary Care Mental Health*, *4*, 165-173.
- Shapiro, M. B. (1961). A method of measuring psychological changes specific to the individual psychiatric patient. *British Journal of Medical Psychology*, *34*, 151-155.
- Seligman, M. E. P. (1995). The effectiveness of psychotherapy: The consumer reports study. *American Psychologist*, *50*, 965-974.
- Silverman, W. K. (1991). Person's description of psychotherapy outcome studies does not accurately represent psychotherapy outcome studies. *American Psychologist*, *46*, 1351-1352.
- Vermeersch, D. A., Lambert, M. J. og Burlingame, G. M. (2000). Outcome Questionnaire: Item sensitivity to change. *Journal of Personality Assessment*, *74*, 242-261.
- Wilson, G. T. (1996). Manual-based treatments: The clinical application of research findings. *Behaviour Research and Therapy*, *34*, 295-314.

Zigmond, A. S. og Snaith, R. P. (1983). The Hospital Anxiety and Depression Scale. *Acta Psychiatrica Scandinavica*, 67, 361-370.

Zung, W. W. K. (1965). A self-rating depression scale. *Archives of General Psychiatry*, 12, 63-70.