

# Félagslega æskileg svörun: Þýðing og próffræðilegir eiginleikar **Balanced Inventory of Desirable Responding**

Ragnhildur Lilja Ásgeirsdóttir, Vaka Vésteinsdóttir, Fanney Þórsdóttir  
Háskóli Íslands

Balanced Inventory of Desirable Responding (BIDR) er eitt mest notaða mælitækið á félagslega æskilegri svörun. Það samanstendur af tveimur undirkvörðum, sjálfsblekkingu (SB) og ímyndarstjórnun (ÍS), sem innihalda hvor um sig 20 fullyrðingar sem svarað er á sjö punkta kvarða (1=Ekki satt, 4=Að einhverju leyti satt, 7=Mjög satt). Tilgangur þessarar rannsóknar var að þýða BIDR kvarðann og kanna próffræðilega eiginleika íslenskrar útgáfu BIDR með staðfestandi þáttagreiningu og ígrunduðum viðtölum (e. *cognitive interviews with probing*). Í fyrri fasa rannsóknarinnar (N=321) er þýðingarferlinu og próffræðilegum eiginleikum lýst. Niðurstöður bentu til þess að meðaltöl og fylgni milli undirkvarða væru sambærileg þeim sem finnast í erlendum rannsóknum og áreiðanleiki mælitækisins væri viðunandi. Niðurstöður úr staðfestandi þáttagreiningu voru einnig sambærilegar þeim sem fengist hafa í erlendum rannsóknum, þó fram hafi komið vandamál sem tengjast vissum atriðum kvarðans. Í seinni fasa rannsóknarinnar var fjallað um niðurstöður viðtala (N=20) þar sem farið var ítarlega í íslenska þýðingu kvarðans. Niðurstöður þeirrar rannsóknar bentu til þess að ekki væru til staðar alvarleg vandamál varðandi þýðingu kvarðans.

**Efnisorð:** Félagslega æskileg svörun, staðfestandi þáttagreining, próffræði, þýðing, BIDR

**F**élagslega æskileg svörun hefur löngum verið áhyggjuefni við túlkun sálfræðilegra prófa. Í starfsmannaráðningum hefur þetta til að mynda valdið hugarangri þar sem niðurstöður slíkra prófa eru hafðar til hliðsjónar við val á starfsmanni. Þrátt fyrir að ýmsar rannsóknir hafi stutt notkun persónuleikaprófa í ráðningarferlinu (Barrick og Mount, 1991; Mount, Barrick og Stewart, 1998; Salgado, 2003) hafa aðrar rannsóknir bent til þess að umsækjendur hafi tilhneigingu til að ýkja jákvæða eiginleika sína á slíkum prófum og draga úr neikvæðum eiginleikum og niðurstöður slíkra prófa gefi því ekki rétta mynd af umsækjendum (Barrick og Mount, 1996; Ones og Viswesvaran, 1998; Zerbe og Paulhus, 1987). Þessi tilhneiging hefur verið

nefnd félagslega æskilega svörun (FÆS).

Reynt hefur verið að bregðast við þessari tilhneigingu fólks með ýmsum aðferðum, svo sem með því að þvinga fólk til að velja milli tveggja eða fleiri staðhæfinga sem taldar eru álíka félagslega æskilegar (e. *forced-choice format*), með því að spyrja einhvern nátengdan viðkomandi í stað þess að spyrja viðkomandi sjálfan eða þá með því að nota aðferðir við fyrirlögn sem bjóða upp á nafnleysi, til dæmis með því að leggja spurningalistann fyrir í gegnum tölvu (sjá t.d. Paulhus, 1991 fyrir

---

Ragnhildur L. Ásgeirsdóttir og Vaka Vésteinsdóttir eru doktorsnemar við sálfræðideild Háskóla Íslands. Fanney Þórsdóttir er lektor við sálfræðideild Háskóla Íslands. Fyrirspurnum vegna greinarinnar skal beina til Ragnhildar L. Ásgeirsdóttur. Netfang: ragnhild@hi.is.

nánari umfjöllun). Með þessum aðferðum er reynt að koma í veg fyrir eða lágmarka FÆS en einnig hefur verið reynt að leiðrétta fyrir FÆS með notkun prófa sem meta hugtakið sjálf. Í gegnum tíðina virðast próf sem meta FÆS helst hafa verið notuð á þrenna vegu til að kanna áhrif slíkrar svörunar á niðurstöður sjálfsmatskvarða í rannsóknum sem fela í sér að spurningalisti sem metur félagslega æskilega hegðun er lagður fyrir á sama tíma og sjálfsmatskvarðinn (Beretvas, Meyers og Leite, 2002). Í fyrsta lagi hefur verið reiknuð fylgni milli spurningalista sem meta FÆS og annarra mælikvarða. Ef fylgnin er lág er gert ráð fyrir að FÆS hafi lítil áhrif á niðurstöður hins mælikvarðans. Í öðru lagi hafa niðurstöður prófanna tveggja verið þáttagreindar til að kanna hvort um sé að ræða tvo aðskilda þætti. Í þriðja og síðasta lagi hafa sumir rannsakendur valið að eyða svörum þeirra sem skora hátt á spurningalistum sem meta FÆS, þar sem þeir álíta að svör þeirra við sjálfsmatskvarðanum litist of mikið af FÆS.

Þau próf sem þekktust eru við mat á FÆS eru Marlowe-Crowne (MC) kvarðinn og *Balanced Inventory of Desirable Responding* (BIDR) kvarðinn (Crowne og Marlowe, 1960; Paulhus, 1984). Marlowe-Crowne kvarðinn er uppruninn frá árinu 1960 og hefur löngum verið algengasti kvarðinn á þessu sviði (Beretvas o.fl., 2002; Leite og Beretvas, 2005). Tilgangurinn með þróun MC kvarðans var að hanna mælitæki sem meta átti félagslega æskilega svörun án þess að innihalda atriði sem lýstu klínískum heilkennum (e. *pathological implications*). MC kvarðinn gerir ráð fyrir að félagslega æskileg svörun sé einvitt hugtak (Crowne og Marlowe, 1960). Fljótlega komu fram kenningar um að hugtakinu mætti betur lýsa með tveggja þátta líkani sem Block og Wiggins nefndu *Alpha* og *Gamma* og voru síðar nánar útlistaðir af Damarin og Messick (sjá Cattell og Scheier, 1961; Edwards, Diers

og Walker, 1962; Jackson og Messick, 1962; Messick, 1962; Wiggins, 1959; Wiggins, 1964 í Paulhus, 2002). *Alpha* lýsir ómeðvitaðri svarskekkju í sjálfsmatskvörðum og *Gamma* lýsir viljandi fölsun. Samkvæmt Paulhus (2002) er BIDR kvarðinn uppruninn frá tilraunum hans til að samþætta kenningar um tveggja þátta líkan FÆS og sjálfsmatskvarða Sackeim og Gur (1979) sem skiptist í tvö mælitæki, *Self-Deception Questionnaire* (SDQ) og *Other-Deception Questionnaire* (ODQ). SDQ atriðin eru spurningar sem almennt eru sannar en fólk vill síður viðurkenna, svo sem „*Hefur þú einhvern tímann hugsað um að drepa einhvern?*“. ODQ atriðin lýsa hegðun sem talin er félagslega æskileg en er í raun sjaldgæf, til dæmis „*Ég skil aldrei eftir mig rusl*“. Upprunalega voru atriði BIDR valin út frá leitandi þáttagreiningu á atriðum nokkurra kvarða sem meta FÆS, Edwards Social Desirability kvarðanum, MMPI-lygakvarðanum, Social Desirability kvarða Wiggins, MC kvarðanum, SDQ kvarðanum og ODQ kvarðanum (sjá Edwards, 1957; Meehl og Hathaway, 1946; Wiggins, 1959; Crowne og Marlowe, 1964; Sackeim og Gur, 1978 í Paulhus, 1984) sem leiddi af sér tvo meginþætti. Af þeim tíu atriðum sem hlóðu hæst á fyrri þáttinn voru fimm úr SDQ kvarðanum og fimm af þeim tíu atriðum sem hlóðu hæst á síðari þáttinn voru úr ODQ kvarðanum, sem Paulhus taldi styrkja tveggja þátta líkan hans um FÆS. Mörg þessara atriða voru síðar umorðuð, atriðum bætt við og atriði fjarlægð sem leiddi að lokum til BIDR kvarða Paulhus. Samkvæmt Paulhus (1984, 1991, 2002) er annars vegar um að ræða sjálfsblekkingu<sup>1</sup> (e. *self-deceptive enhancement*) og hins vegar ímyndarstjórnun (e. *impression management*). Bæði sjálfsblekking og ímyndarstjórnun vísa til þeirrar tilhneigingar fólks að gefa fegraða mynd af sér sem lýsir sér annað hvort með því að fólk hefur ýkta jákvæða mynd af sjálflu

<sup>1</sup> Hefð hefur skapast fyrir því í erlendum ritum að nota hugtakið *self-deception* til einföldunar fyrir þennan þátt og var því ákveðið að velja íslensku þýðinguna, sjálfsblekkingu, þrátt fyrir að það sé ekki bein þýðing á hugtakinu *self-deceptive enhancement*.

sér, það er sjálfsblekking eða með því að fólk reynir viljandi að ýkja jákvæða eiginleika sína og draga úr neikvæðum eiginleikum sínum, það er ímyndarstjórnun. Skilgreining á hugtakinu FÆS er þó enn fremur umdeild og ekki ljóst hvort eins, tveggja eða jafnvel þriggja þátta lífan lýsi hugtakinu best (Helmes og Holden, 2003).

Sjötta útgáfa BIDR-kvarðans kom út árið 1988 og inniheldur 20 atriði sem meta sjálfsblekkingu (SB) og 20 atriði sem meta ímyndarstjórnun (ÍS) sem svarað er á sjö punkta einvíðri stiku (Paulhus, 1991). Atriðunum er svarað á sjö punkta einvíðri stiku sem fá orðgildin; Ekki satt (1), Að einhverju leyti satt (4) og Mjög satt (7). Paulhus (1994) hefur bæði heimilað samfellda skorun<sup>2</sup>, þar sem svör við hverri spurningu eru lögð saman og tvíkosta skorun, þar sem svargildin 6 og 7 fá gildið 1 en önnur svargildi fá gildið 0. Við spurningar sem eru neikvætt orðaðar fá svargildin 1 og 2 gildið 1 en önnur svargildi fá gildið 0. Helmingur spurninga BIDR-6 eru neikvætt orðaðar og svör við þeim atriðum því endurkóðuð við útreikning heildarskors. Þrátt fyrir að Paulhus (1991) hafi mælt með tvíkosta skorun út frá þeim rökum að ætlunin sé að bera kennsl á þá sem ýkja svör sín og niðurstöður staðfestandi þáttagreiningar í fyrri rannsóknnum (Leite og Beretvas, 2005) stutt tvíkosta skorun, hafa aðrar rannsóknir bent til þess að samfelld skorun geti verið ákjósanlegri, með tilliti til áreiðanleika, samleitniréttmætis, fylgni við persónuleikapróf og þáttabyggingar (Cervellione, Lee og Bonanno, 2009; Stöber, Dette og Musch, 2002). Uppgefnir áreiðanleiki (alfa) prófsins frá útgefanda er frá 0,67 til 0,77 fyrir SB og frá 0,77 til 0,85 fyrir ÍS (Paulhus, 1994) en allsherjargreining Li og Bagger (2007) á 110 fyrri rannsóknnum á BIDR þar sem áreiðanleikastuðull var gefinn upp gaf til kynna að meðaltalið væri 0,68 fyrir SB (sf = 0,09; spönn 0,27-0,92) og 0,74 fyrir ÍS

(sf = 0,09; spönn 0,32-0,88). Ekki var fylgni milli skorunaraðferðar og áreiðanleikastuðuls. Samkvæmt viðmiðum DeVellis (2003) um að áreiðanleikastuðull kvarða skuli vera að lágmarki 0,65 fyrir notkun í rannsóknnum og á hópum er áreiðanleiki BIDR-6 því viðunandi. Rannsóknir Vispoel og Tao (2013) benda til þess að endurtekningaráreiðanleiki SB sé 0,78 við tvíkosta skorun og 0,83 við samfellda skorun en fyrir ÍS sé endurtekningaráreiðanleikinn 0,83 við tvíkosta skorun og 0,86 við samfellda skorun. Niðurstöður fyrri rannsókna benda til þess að SB og ÍS hafi fylgni við aðrar hugsmíðar, svo sem undirklarða NEO-FFI; taugaveiklun (SB), úthverfu (SB), samviseksemi (SB og IM) og samvinnuþýði (IM) (Costa og McCrae, 1992; Pauls og Crost, 2004; Pauls og Stemmler, 2003; Stöber o.fl., 2002).

Þar sem BIDR-6 er annað af tveimur algengustu mælitækjum á sviði félagslega æskilegrar svörunar og fyrri rannsóknir benda til þess að áreiðanleiki og réttmæti sé viðunandi er talið mikilvægt að þróa vandaða íslenska útgáfu af kvarðanum. Þrátt fyrir að sjöunda útgáfa kvarðans hafi verið gefin út árið 1998 sem nefnd hefur verið *Paulhus Deception Scales* er sú útgáfa einungis fánleg gegn greiðslu og hefur BIDR-6 kvarðinn því verið mun meira notaður í rannsóknnum (Stöber o.fl., 2002). Helsti munur á sjöttu og sjöundu útgáfu kvarðans er að svarmöguleikum var fækkað úr sjö í fimm og nokkur atriði lagfærð eða fjarlægð. Rannsóknir benda þó til að þess að sjötta og sjöunda útgáfa kvarðans séu sambærilegar (Lanyon og Carle, 2007).

### Markmið rannsókna

BIDR-6 kvarðinn hefur mikið verið notaður í rannsóknnum erlendis (Li og Bagger, 2006, 2007; Stöber o.fl., 2002) en ekki er vitað til þess að unnið hafi verið að íslenskri þýðingu kvarðans. Markmið þessarar rannsóknar er því

<sup>2</sup> Til einföldunar er fjallað um skorunina sem samfellda skorun og vísað til hefðar í erlendum rannsóknnum á BIDR að fjalla um samfellda skorun (e. *continuous scoring*), þrátt fyrir að strangt til tekið sé skorunin ekki samfelld þar sem um sjö punkta stiku er að ræða.

að þýða og forprófa íslenska þýðingu BIDR-6 kvarðann. Þörf er á þýddu mælitæki sem hægt er að nota á Íslandi við mat á FÆS þar sem erlendar rannsóknir hafa bent til þess að FÆS hafi áhrif á mælingar með sjálfsmatskvörðum sem notaðir eru á Íslandi (Pauls og Crost, 2004; Pauls og Stemmler, 2003; Stöber o.fl., 2002). Í fyrri fasa rannsóknarinnar var kvarðinn þýddur og próffræðilegir eiginleikar hans kannaðir við tvíkösta og samfellda skorun. Við þýðingu mælitækja er mikilvægt að vandað sé til verks og reynt að tryggja að íslensk þýðing mælitækisins sé sambærileg upprunalegri útgáfu. Þær aðferðir sem algengast hafa verið í gegnum tíðina við þýðingar á mælitækjum hér á landi eru annars vegar þýðing og bakþýðing, sem í sinni einföldustu mynd felur í sér að tvítyngdur aðili þýðir próf, annar tvítyngdur aðili bakþýðir yfir á upprunalega tungumálið og að lokum er sú útgáfa borin saman við upprunalegu útgáfu prófsins og hins vegar að framkvæma tvær sjálfstæðar þýðingar, sem eru samræmdar og bornar undir óháðan tvítyngdan sérfræðing á sviðinu (Behling og Law, 2000; Einar Guðmundsson, 2005-2006). Báðar þessar aðferðir hafa þó verið gagnrýndar í gegnum tíðina. Bent hefur verið á að ef þýðendur eru meðvitáðir um að bakþýða eigi þýðingu þeirra, hafi þeir tilhneigingu til að velja fremur orðalag sem líklegra er að leiði til réttar bakþýðingar, fremur en að velja það orðalag sem hentar best (Hambleton, Merenda og Spielberger, 2005). Ef sérfræðingur sinnir bakþýðingu eru einnig líkur á því að þeir geti áttað sig á inntaki og hugtökum vegna þekkingar sinnar á efninu. Bakþýðingin gæti þannig orðið mjög sambærileg upprunalegri útgáfu þrátt fyrir að þýðingu sé ábótavant, ásamt því að þýðingin sjálf er ekki metin, aðeins frumútgáfan og bakþýðingin (sjá Brislin, 1970 í Behling og Law, 2000). Þegar að þýðingar eru unnar af sérfræðingum er jafnframt sú hættu að orðalag þeirra einkennist af fagmáli og er því mikilvægt að fleiri komi að þýðingarferlinu. Ein leið til að bregðast við þessu er með viðtölum (e. *cognitive interviews*) þar sem skilningur fólks á atriðum

er kannaður og borin kennsl á vandamál tengd þýðingu (Beatty og Willis, 2007; Presser og Blair, 1994). Í síðari fasa rannsóknarinnar voru tekin viðtöl þar sem skilningur viðmælenda á atriðum BIDR-6 kvarðans var kannaður, með það að leiðarljósi að bera kennsl á vandamál tengd þýðingu kvarðans. Niðurstöður þessarar rannsóknar verða notaðar til að leggja mat á gæði íslenskrar þýðingar á BIDR-6 kvarðanum.

## Aðferð

### Þátttakendur

*Fasi eitt.* Í fyrri fasa rannsóknarinnar var hentugleikaúrtak notað sem samanstóð af 321 þátttakanda. Upprunalegur fjöldi þátttakenda var 375 en 54 (14,4%) svöruðu ekki öllum atriðum og voru svör þeirra fjarlægð úr gagnasafninu. Stuðst var við hefð í erlendum rannsóknnum á BIDR-6 kvarðanum við þá ákvörðun (Li og Reb, 2009) til að auðvelda samanburð við erlendar niðurstöður. Kynjahlutfallið skiptist þannig að 83,8% þátttakenda voru kvenkyns og 16,2% voru karlkyns. Þátttakendur voru á aldrinum 19 til 71 árs ( $M=35,75$ ,  $sf=10,45$ ) og hafði 63,3% þátttakenda lokið við háskólagráðu. Þátttakendur samanstóðu af háskólanemendum (30%) og starfsmönnum nokkurra fyrirtækja og stofnana.

*Fasi tvö.* Í seinni fasa var hentugleikaúrtak notað sem samanstóð af 20 þátttakendum. Kynjaskipting var jöfn; tíu konur og tíu karlmenn. Menntun þátttakenda var breytileg og höfðu viðmælendur ýmist lokið grunnskólaprófi, verkmenntaprófi, stúdentsprófi eða háskólaprófi. Mikil dreifing var á aldri, yngsti þátttakandinn var 21 árs en sá elsti var 58 ára. Við val á þátttakendum var reynt að stuðla að því að hafa kynjahlutfall jafnt og dreifingu sem breiðasta með tilliti til aldurs og menntunar.

### Mælitæki

Uppgefin meðaltöl BIDR-6 frá útgefanda eru 7,5 (kk) og 6,8 (kvk) fyrir SB en 4,3 (kk) og

4,9 (kvk) fyrir ÍS (Paulhus, 1994). Meðaltöl í erlendum rannsóknum þar sem notast var við upprunalega útgáfu listans fyrir SB hafa verið á bilinu 3,9 til 6,2 við tvíkosta skorun en 77,2 til 87.9 fyrir samfellda skorun og kynjamunur oft ekki verið til staðar. Erlend meðaltöl fyrir ÍS hafa verið á bilinu 4,4 til 6,5 við tvíkosta skorun en 66,3 til 79,7 við samfellda skorun (Holden o.fl., 2000; Li og Reb, 2009; Stöber o.fl., 2002; Vispoel og Tao, 2013).

*Þýðing BIDR-6 kvarðans.* Íslensk þýðing listans var unnin af þremur þýðendum sem hafa lokið meistaraþáttu í Sálfræði eða eru við það að ljúka slíkri gráðu. Þýðendur hafa allir búið í enskumælandi löndum í tvö til tólf ár og eru tveir þeirra tvítýngdir. Við val á þýðendum var stuðst við ráðleggingar Geisinger (1994) um að ákjósanlegast sé að þýðendur hafi sérþekkingu á því sviði sem mælitækið fellur undir, séu tvítýngdir og hafi þekkingu á menningu í upprunalega heimalandi mælitækisins og þar sem nota á mælitækið. Stuðst var við ráðleggingar Alþjóðlegu prófanefndarinnar (Einar Guðmundsson, 2005-2006; Hambleton o.fl., 2005) og Behling og Law (2000) við hönnun þýðingarferlisins. Þýðingarferlið fór þannig fram að útbúnar voru þrjár sjálfstæðar þýðingar. Þýðendur hittust svo til að samræma þýðingar, samanber ráðleggingar Geisinger (1994). Farið var yfir hverja staðhæfingu og orðalag samræmt. Í þeim tilvikum þar sem verulegur munur var til staðar á merkingu atriða voru atriðin skoðuð nánar og komist að sameiginlegri niðurstöðu um þýðingu atriðis. Sem dæmi um atriði sem erfitt reyndist að þýða má nefna atriði þrjú („*I don't care to know what other people really think of me*“). Ekki náðu allar tillögur að þýðingum að endurspeglar merkingu fyrri hluta þessa atriðis („*I don't care to know*“) en á endanum náðist sátt um þýðinguna „*Ég kærir mig ekki um að vita hvaða álit aðrir hafa í raun á mér*“. Atriði sex reyndist þýðendum einnig flókið („*When my emotions are aroused, it biases my thinking*“). Ekki var talið ljóst hvort um væri að ræða eingöngu neikvæðar tilfinningar

eða bæði neikvæðar og jákvæðar tilfinningar. Á endanum varð þýðingin „*Þegar ég er í tilfinningalegu uppnámi hugsa ég ekki skýrt*“ fyrir valinu. Atriði 34 var einnig flókið í þýðingu („*I never read sexy books or magazines*“). Það sem reyndist þýðendum erfiðast varðandi þetta atriði var hvaða íslenska lýsingarorð samsvaraði enska orðinu „*sexy*“ í þessu tilviki, þá sérstaklega í styrkleika og notkun í daglegu tali. Íslenska þýðingin „*Ég les aldrei bækur eða tímarit með kynþokkafullu innihaldi*“ var að lokum valin. Þar sem talið var að hægt væri að skilja þetta atriði á mjög ólíkan hátt, það er með tilliti til þess hvað flokkist undir „*bækur eða tímarit með kynþokkafullu innihaldi*“ var ákveðið að bera þessa þýðingu sérstaklega undir 10 aðila sem fengu að sjá bæði íslenska og enska útgáfu þessa atriðis. Þeir aðilar voru beðnir um að mynda sér skoðun á því hvað þeir teldu flokkast undir „*sexy books or magazines*“ annars vegar og „*bækur eða tímarit með kynþokkafullu innihaldi*“ hins vegar. Flestir svöruðu þessum spurningum á sama hátt sem talin n var rökstuðningur fyrir því að velja þessa þýðingu. Mikil breidd var þó í því hvað þátttakendur töldu falla undir slík rit.

Þegar að fyrstu drög að þýðingu voru tilbúin var tvítýngdur Íslendingur með BA gráðu í Sálfræði, sem starfað hefur við þýðingar og búið í enskumælandi löndum mest alla ævi, fenginn til að kanna hvort áberandi merkingarmunur væri til staðar á enskri og íslenskri útgáfu listans, sem ekki reyndist vera. Að því loknu var listinn lagður fyrir þrjá yngri viðmælendur (19-24 ára) með tilliti til orðalags. Talið var mikilvægt að fá sjónarmið yngri þátttakenda á þessu stigi. Þeir viðmælendur voru beðnir um að lesa listann yfir með tilliti til þess hversu auðskilið orðalagið væri. Einu atriði var breytt í kjölfar þess þar sem tveir viðmælendur skildu ekki orðatiltækið „*ad jafna metin*“. Þýðingu á atriði 25 („*I sometimes try to get even rather than forgive and forget*“) var því breytt frá „*Ég reyni stundum að jafna metin frekar en að fyrirgefa*“ í „*Ég reyni stundum að hefna mín frekar en að fyrirgefa*“. Þýðendur höfðu áhyggjur af því að orðalag

á íslenskri þýðingu atriðisins væri sterkara og því væri fólk tregara til að samþykkja staðhæfinguna og var það kannað sérstaklega í fasa tvö. Í lokin var íslensk og ensk útgáfa listans borin undir aðila með doktorspróf í Sálfræði sem samþykkti þýðinguna eftir að þrjár smávægilegar lagfæringar voru gerðar á orðalagi íslensku útgáfunnar.

### Framkvæmd

*Fasi eitt.* Í fyrri fasa var BIDR-6 kvarðinn lagður fyrir á veraldarvefnum. Þátttakendum var sendur vefpóstur með beiðni um þátttöku í rannsókninni, ásamt því að þeir voru beðnir um að senda spurningalistann áfram á aðra mögulega þátttakendur. Vegna fyrirkomulags við fyrirlögn listans var ekki mögulegt að hafa stjórn á þeim aðstæðum sem þátttakendur voru í við svörun listans né reikna út svarhlutfall en kostur þessa fyrirkomulags var að úrtakið var nokkuð fjölbreytt sem var talið vega þyngra. Lögð var áhersla á það við þátttakendur að þeim væri ekki skylt að svara öllum spurningum og hætta mætti þátttöku hvenær sem er. Engin umbun var veitt fyrir þátttöku.

*Fasi tvö.* Í seinni fasa voru tekin viðtöl við þátttakendur þar sem kafað var ofan í skilning þeirra á hverju atriði BIDR-6 kvarðans. Viðtölin voru hálfstöðluð en spurt var sérstaklega um atriði sem reynst höfðu flókin í þýðingu (e. *cognitive interviews with probing*). Aðrar aðferðir eru mögulegar við forprófun spurningalista, svo sem hefðbundin forprófun þar sem spurningalisti er lagður fyrir takmarkaðan fjölda þátttakenda, líkt og um hefðbundna fyrirlögn væri að ræða í þeim tilgangi að bera kennsl á áberandi vandamál (e. *conventional pretests*), atferliskóðun (e. *behavior coding*) eða þá að rýnihópur sérfræðinga er fenginn til að yfirfara spurningalistann (e. *expert panels*) (Presser og Blair, 1994). Hver aðferð hefur sína kosti og galla, svo sem varðandi kostnað, tíma eða nákvæmni niðurstaðna en talið var að ígrunduð viðtöl með ítarspurningum (e. *cognitive interviews with probing*) hentuðu best í þessari rannsókn þar sem markmiðið

var að bera kennsl á vandamál við þýðingu atriðanna og fá sem ítarlegastar upplýsingar um skilning ólíkra þátttakenda á atriðunum. Flokka má ítarspurningar í fjóra flokka (Beatty og Willis, 2007), það er fyrirsjáanlegar (e. *anticipated*; fyrirfram ákveðnar af rannsakanda), ófyrirsjáanlegar (e. *spontaneous*; ákveðnar á staðnum af rannsakanda), tilfallandi (e. *emergent*; ákveðnar á staðnum undir áhrifum frá svaranda) og skilyrtar (e. *conditional*; fyrirfram ákveðnar en háðar svaranda). Fjórir rannsakendur komu að undirbúningi viðtalanna og samningu handrits sem stuðst var við í viðtölunum. Ráðleggingum Fowler (1995) var fylgt við framkvæmd viðtala. Þyrjað var á opnum spurningum um atriði og hvernig viðkomandi hefði komist að niðurstöðu varðandi val á svarmöguleika. Því næst var farið ítarlega í ákveðin atriði ef viðkomandi hafði ekki þegar minnst á þau, svo sem varðandi skilning viðkomandi á ákveðnum orðum eða orðasamböndum. Í lokin voru þátttakendur enn fremur spurðir að því hvort þeir teldu atriðið hafa verið óljóst að einhverju leyti eða hvort orða mætti atriðið betur.

Viðtölin fóru fram á heimili rannsakanda, heimili viðmælanda eða vinnustað viðmælanda. Til að koma í veg fyrir truflun fóru viðtölin fram í lokuðu eða aðskildu herbergi þegar þátttakendur höfðu nægan tíma. Þar sem listinn samanstendur af 40 staðhæfingum var ákveðið að viðtölin færu fram í tveim lotum. Í fyrri viðtalinu var farið yfir atriði sjálfsblekkingar og í seinna viðtalinu var farið yfir atriði ímyndarstjórnunar. Að meðaltali tók hvert viðtal um tvær klukkustundir í framkvæmd og því samtals um fjórar klukkustundir fyrir hvern þátttakanda. Í upphafi hvers viðtals voru viðmælendur beðnir um að lesa hefðbundin fyrirmæli sem fylgja BIDR-6 kvarðanum og útskýrt fyrir þeim að tilgangur þessarar rannsóknar væri einungis að varpa ljósi á skilning viðmælanda á atriðum listans. Því næst fengu þátttakendur tíma og næði til að svara listanum. Að lokum fór rannsakandi ásamt viðmælanda yfir hvert og eitt atriði kvarðans. Viðtölin voru hljóðrituð og síðar skrifuð upp

orðrétt. Unnið var úr þeim gögnum til að bera kennsl á vandamál sem gætu verið til staðar við þýðingu atriðanna. Þátttakendur voru beðnir um leyfi fyrir upptökunum gegn því loforði að gögnum yrði eytt að rannsókn lokinni.

### Tölfræðileg úrvinnsla

Við tölfræðilega úrvinnslu í fasa eitt var notast við tölfræðiforritin SPSS 20 og LISREL 9.10 (Jöreskog og Sörbom, 2012). SPSS 20 var notað við útreikning á lýsandi tölfræði og atriðagreiningu en staðfestandi þáttagreining var gerð með LISREL í þeim tilgangi að kanna þáttabyggingu íslenskrar útgáfu BIDR. Þar sem breytur voru á raðkvarða og skekktar (Shapiro-Wilk próf voru marktæk) var aðferð minnstu veginna ferninga, þar sem leiðrétt er fyrir smæð úrtaks (e. *diagonally weighted least squares*), notuð til að greina raðkvarðafylgnifylki (e. *polychoric correlation*; *tetrachoric* fylgnifylki fyrir tvíkosta skorun) og aðfellusamdreifnifylki (e. *asymptotic covariance matrix*) (Flora og Curran, 2004). Með staðfestandi þáttagreiningu er kannað hve vel gögnin falla að þeim mællíkönum sem sett eru fram. Þannig er hægt að kanna gæði beggja þátta BIDR-6 kvarðans. Til að skoða heildarmátgæði líkana er mikilvægt að styðjast við fleiri en einn mátgæðastuðul og verður stuðst við niðurstöður Satorra-Bentler kí-kvaðratprófs ( $SB\chi^2$ ), RMSEA (e. *Root Mean Square Error of Approximation*), CFI (e. *Comparative Fit Index*) og NNFI (e. *Non-normed Fit Index*), einnig þekkt sem TLI (e. *Tucker-Lewis Index*), í þessari rannsókn samkvæmt ábendingum frá Hu og Bentler (1998), West, Taylor og Wu (2012) og Yu (2002).

Satorra-Bentler kí-kvaðratprófið er notað í stað hefðbundins kí-kvaðratprófs þegar að aðferð minnstu veginna ferninga (e. *diagonally weighted least squares*) er notuð og leiðréttir sá stuðull fyrir skekkju í dreifingu (Yu, 2002). Ef niðurstöður kí-kvaðratprófs eru háar og stuðullinn marktækur ( $p > 0,05$ ) er núlltilgátunni hafnað sem bendir til þess að mátgæði líkansins séu ófullnægjandi en ef niðurstöðurnar eru

lágur og stuðullinn ómarktækur ( $p < 0,05$ ) eru rök fyrir því að ekki sé hægt að hafna núlltilgátunni og mátgæðin séu góð (Albright og Park, 2009; West o.fl., 2012). Helsti ókostur við kí-kvaðrat stuðullinn er sá að stuðullinn er háður úrtaksstærð og er því líklegur til þess að vera marktækur ef úrtak er miðlungsstórt, óháð mátgæðunum (Hu og Bentler, 1999; West o.fl., 2012). Í minni úrtökum,  $N \leq 250$ , getur verið varhugavert að líta eingöngu á niðurstöður SB kí-kvaðratprófsins og er því einnig mælt með notkun RMSEA, CFI og NNFI í þeim tilvikum (Yu, 2002). Í þessari rannsókn er úrtakið aðeins lítillega yfir því viðmiði og verður því einnig litið til niðurstaðna RMSEA, CFI og NNFI við mat á gæðum líkana. RMSEA stuðullinn hefur þann kost að vera nokkuð ónæmur fyrir úrtaksstærð og segir til um það hversu illa líkanið fellur að gögnunum (Bentler, 1990). CFI stuðullinn lýsir því hversu mikið betri mátgæði mællíkansins eru samanborið við mællíkan sem gerir ekki ráð fyrir sambandi milli breytanna. NNFI stuðullinn ber saman mátgæði mællíkans við mátgæði líkans sem gerir ekki ráð fyrir sambandi milli mælibreyta og leiðréttir fyrir stærð líkans. Gildi NNFI falla oftast á bilinu núll til einn en stuðullinn getur gefið af sér gildi sem falla fyrir utan þá spönn. Því hærri sem stuðullinn er því betri eru mátgæði mællíkansins. Almenn er gert ráð fyrir að viðmiðunargildin  $\geq 0,95$  fyrir NNFI og CFI og  $\leq 0,08$  fyrir RMSEA bendi til þess að líkan falli vel að gögnunum (Bentler, 1990; Browne og Cudeck, 1993; Hu og Bentler, 1998, 1999; Steiger, 1990; West o.fl., 2012; Yu, 2002; Yuan, 2005). En í ljósi þess að varað hefur verið við of strangri túlkun þessara viðmiðunargilda (Markland, 2007; West o.fl., 2012) og jafnvel mælt með því að færa viðmiðunargildin fyrir NNFI og CFI niður fyrir 0,90 þegar staðfestandi þáttagreining er notuð til að meta margvið sálfræðileg mælitæki (Marsh, Hau og Wen, 2004), verður að túlka þau á varfærinn hátt.

## Niðurstöður

### Fasi eitt

*Lýsandi tölfraði.* Líkt og sjá má í 1. töflu voru meðaltöl sjálfsblekkingar (SB) í íslenski útgáfu BIDR-6 kvarðans sambærileg meðaltölum úr erlendum rannsóknum en meðaltöl ímyndarstjórnunar (ÍS) voru örlítið hærri en í erlendum rannsóknum (Holden o.fl., 2000; Li og Reb, 2009; Stöber o.fl., 2002).

Ekki var til staðar kynjamunur á meðaltölum SB og ÍS, hvorki með tvíkosta skorun ( $t(313)=0,706, p=0,481$ ) fyrir SB og ( $t(313)=-0,133, p=.894$ ) fyrir ÍS né samfelldri skorun ( $t(313)=0,674, p=0,501$ ) fyrir SB og ( $t(313)=-0,215, p=0,830$ ) fyrir ÍS. Dreifingin er fremur jákvætt skekkt við tvíkosta skorun

bæði fyrir SB og ÍS en víkur aðeins lítillega frá normaldreifingu við samfellda skorun. Neikvætt ris dreifingar undirþátta við tvíkosta og samfellda skorun merkir að dreifingin er örlítið lægri og halar dreifingarinnar lengri en í normaldreifingu.

Fylgni milli undirþátta BIDR kvarðans, SB og ÍS, má sjá í 2. töflu. Fylgnin er sambærileg fylgni milli undirþátta í erlendum rannsóknum (Stöber o.fl., 2002) og er hæst milli ÍS við tvíkosta og samfellda skorun og lægst milli SB við tvíkosta skorun og ÍS við samfellda skorun. Fylgnistuðlar fyrir sömu hugsmíð með ólíkum skorunaraðferðum eru í báðum tilfellum háir, 0,88 fyrir ÍS við tvíkosta og samfellda skorun og 0,78 fyrir SB við tvíkosta og samfellda skorun.

**1. tafla.** Lýsandi tölfraði fyrir undirþætti BIDR-6 kvarðans við tvíkosta og samfellda skorun (N=321)

		Meðaltal	Staðalfrávik	Skekka	Ris dreifingar
Sjálfsblekking	Tvíkosta skorun	5,08	3,23	0,70	0,04
	Samfelld skorun	83,59	12,08	0,05	-0,22
Ímyndarstjórnun	Tvíkosta skorun	6,83	3,76	0,43	-0,37
	Samfelld skorun	81,56	16,03	0,03	-0,02

**2. tafla.** Fylgni milli undirþátta BIDR-6 við tvíkosta og samfellda skorun

		Sjálfsblekking tvíkosta	Sjálfsblekking samfelld	Ímyndarstjórnun tvíkosta
Sjálfsblekking	Tvíkosta skorun			
	Samfelld skorun	0,78*		
Ímyndarstjórnun	Tvíkosta skorun	0,31*	0,32*	
	Samfelld skorun	0,22*	0,36*	0,88*

\* $p < 0,001$

**3. tafla.** Áreiðanleiki undirþátta BIDR-6 kvarðans við tvíkosta og samfellda skorun

		N	Alfa
Sjálfsblekking	Tvíkosta skorun	321	0,70
	Samfelld skorun	321	0,71
Ímyndarstjórnun	Tvíkosta skorun	321	0,76
	Samfelld skorun	321	0,77



*Próffræðilegir eiginleikar BIDR-6.* Í 3. töflu má sjá áreiðanleika (alfa) undirþátta BIDR-6 kvarðans. Líkt og sjá má er hann viðunandi (DeVellis, 2003) fyrir SB og ÍS bæði við tvíkosta og samfellda skorun. Áreiðanleikastuðlar eru hærri fyrir ÍS en SB og örlítið hærri við samfellda skorun beggja kvarða. Áreiðanleikastuðlar eru sambærilegir áreiðanleikastuðlum úr erlendum rannsóknum og nálægt meðaltali allsherjargreiningar Li og Bagger (2007), þar sem meðaltalið var 0,68 fyrir SB og 0,74 fyrir ÍS.

Réttmæti íslenskrar þýðingar BIDR-6 kvarðans var metið með staðfestandi þáttagreiningu. Borin voru saman tvö líkön þar sem dreifni þátta var fest í einum, annars vegar líkan fyrir SB og hins vegar líkan fyrir ÍS. Algengt er að BIDR-6 kvarðinn sé ekki lagður fyrir í heild sinni og aðeins annar hvor kvarðinn notaður og eru það helstu rökin fyrir að tvö líkön voru metin. Niðurstöður staðfestandi þáttagreiningar fyrir SB og ÍS, annars vegar við tvíkosta skorun og hins vegar við samfellda skorun, má sjá í 4. töflu.

Niðurstöður staðfestandi þáttagreiningar benda til þess að tvíkosta skorun komi betur út en samfelld skorun, bæði fyrir SB og ÍS, sem er í samræmi við niðurstöður Leite og Beretvas (2005). RMSEA, CFI og NNFI stuðlar benda til þess að mátgæði mæilíkans fyrir ÍS kvarðann með tvíkosta skorun séu ásætlanleg. Mátgæði annarra mæilíkana eru þó fremur slök miðað við hefðbundin viðmiðunargildi en þetta er í samræmi við niðurstöður erlendra rannsókna (Cervellione o.fl., 2009; Lanyon og Carle, 2007; Leite og Beretvas, 2005).

Í 5. töflu má sjá atriðin 40 í íslenski þýðingu og staðlaðar þáttahleðslur hvers atriðis við tvíkosta og samfellda skorun. Meðaltöl og staðalfrávik atriða eru einnig birt. Þáttahleðslur fyrir SB við tvíkosta skorun eru á bilinu 0,18 til 0,67. Allar þáttahleðslur voru tölfraðilega marktækar en hleðslur þriggja atriða (3, 7 og 13) voru lægri en 0,3, sem túlka má sem lágur (Floyd og Widaman, 1995). Við samfellda skorun eru þáttahleðslur frá 0,00 til 0,61. Tvær þáttahleðslur voru ekki marktækar, atriði 3 og 7 og hleðslur fimm atriða voru lægri en 0,3. Þetta voru atriði 1, 4, 8, 13 og 14. Í öllum tilfellum nema þremur eru þáttahleðslur hærri fyrir tvíkosta skorun en samfellda skorun. Þáttahleðslur fyrir ÍS voru frá 0,31 til 0,71 við tvíkosta skorun. Við samfellda skorun voru þáttahleðslur frá 0,26 til 0,61. Allar þáttahleðslur voru marktækar en eitt atriði var undir 0,3 við samfellda skorun, atriði 34.

### Fasi tvö

Niðurstöður viðtala bentu til þess að mögulega mætti bæta þýðingu atriða 13, 16, 19, 23, 25 og 39 og hugsanlega er orðalag atriða 13, 23, 25 og 39 of sterkt. Hafa ber þó í huga að mörg atriðanna eru viljandi mjög sterkt orðuð í upprunalegri útgáfu kvarðans þar sem ætlunin er að bera kennsl á ýkta svörun en mikilvægt er að bera kennsl á atriði sem mögulega geta verið orðuð á enn sterkari hátt í íslenski þýðingu.

Orðið „meta“ í atriði 16 „Ég kann sjaldnast að meta gagnrýni“ vafðist fyrir einhverjum þátttakendum. Viðmælendur lögðu tvenns konar skilning á orðið, annars vegar skildu þeir hugtakið á þann veg að átt væri við hvort

**4. tafla.** Niðurstöður staðfestandi þáttagreiningar við tvíkosta og samfellda skorun

		SB $\chi^2$	df	RMSEA	Öryggisbil RMSEA (90%)	CFI	NNFI
SB	Tvíkosta skorun	517,66*	170	0,133	0,125-0,140	0,90	0,88
	Samfelld skorun	397,14*	170	0,147	0,139-0,154	0,84	0,83
ÍS	Tvíkosta skorun	393,20*	170	0,094	0,087-0,102	0,95	0,94
	Samfelld skorun	397,74*	170	0,115	0,107-0,122	0,90	0,88

\* $p < 0,01$

**5. tafla.** Meðaltöl, staðalfrávik og þáttahleðslur atriða í staðfestandi þáttgreiningu við tvíkosta skorun (T) og samfellda skorun (S)

	Þáttahleðslur		Meðaltal (staðalfrávik)		
	T	S	T	S	
<i>Sjálfsblekking</i>					
1.	Það álit sem ég mynda mér á fólki við fyrstu kynni reynist yfirleitt rétt.	0,30	0,26	0,19(0,39)	4,44(1,18)
2.	Það væri erfitt fyrir mig að losa mig við einhvern af mínum ósiðum. *	0,45	0,43	0,23(0,42)	4,27(1,46)
3.	Ég kæri mig ekki um að vita hvaða álit aðrir hafa í raun á mér.	0,18	0,03 <sup>a</sup>	0,09(0,29)	3,27(1,45)
4.	Ég hef ekki alltaf verið heiðarleg/ur gagnvart sjálfri/sjálfum mér. *	0,34	0,21	0,17(0,37)	3,73(1,67)
5.	Ég veit alltaf af hverju mér líkar við eitthvað.	0,66	0,46	0,27(0,45)	4,45(1,50)
6.	Þegar ég er í tilfinningalegu uppnámi hugsá ég ekki skýrt. *	0,67	0,57	0,12(0,33)	3,53(1,57)
7.	Þegar ég hef gert upp hug minn geta aðrir sjaldan fengið mig til að skipta um skoðun.	0,23	-0,01 <sup>a</sup>	0,12(0,32)	3,56(1,51)
8.	Ég er ekki öruggur bílstjóri þegar ég keyri yfir hámarks- hraða. *	0,35	0,20	0,21(0,41)	3,76(1,81)
9.	Ég hef fullkomna stjórn á eigin örlögum.	0,32	0,39	0,15(0,36)	3,69(1,59)
10.	Það er erfitt fyrir mig að bæjga frá mér óþægilegum hugsunum. *	0,65	0,60	0,26(0,44)	4,11(1,69)
11.	Ég sé aldrei eftir ákvörðunum mínum.	0,45	0,38	0,06(0,24)	2,90(1,40)
12.	Ég missi stundum af tækifærum því að ég er ekki nógu fljót/ur að gera upp hug minn. *	0,44	0,37	0,22(0,42)	4,06(1,62)
13.	Ég kys vegna þess að atkvæði mitt getur skipt sköpum.	0,26	0,16	0,58(0,49)	5,47(1,52)
14.	Foreldrar mínir voru ekki alltaf sanngjarnir þegar þeir refsuðu mér. *	0,34	0,28	0,39(0,49)	4,69(1,81)
15.	Ég er fullkomlega skynsöm manneskja.	0,61	0,45	0,30(0,46)	4,53(1,52)
16.	Ég kann sjaldnast að meta gagnrýni. *	0,51	0,34	0,37(0,48)	4,92(1,32)
17.	Ég hef mikla trú á dómgreind minni.	0,66	0,46	0,49(0,50)	5,31(1,13)
18.	Ég hef stundum efast um hæfni mína sem elskhugi. *	0,43	0,46	0,29(0,46)	4,17(1,77)
19.	Það er allt í lagi mín vegna þó einhverju fólk kunni að líka illa við mig.	0,40	0,31	0,34(0,48)	4,66(1,63)
20.	Ég veit ekki alltaf af hverju ég geri það sem ég geri. *	0,59	0,61	0,23(0,42)	4,07(1,57)
<i>Ímyndarstjórnun</i>					
21.	Ég lýg stundum ef þörf krefur. *	0,71	0,61	0,31(0,46)	4,36(1,63)
22.	Ég hylmi aldrei yfir mistök mín.	0,64	0,43	0,26(0,44)	4,09(1,66)
23.	Það hefur komið fyrir að ég hef notfært mér einhvern. *	0,65	0,60	0,33(0,47)	4,58(1,55)
24.	Ég blóta aldrei.	0,37	0,35	0,08(0,28)	2,23(1,67)
25.	Ég reyni stundum að hefna mín frekar en að fyrirgefa. *	0,55	0,54	0,63(0,48)	5,56(1,45)

**5. tafla.** Meðaltöl, staðalfrávik og þáttahleðslur...framhald

	Þáttahleðslur		Meðaltal (staðalfrávik)	
	T	S	T	S
26. Ég fylgi alltaf lögum þó ólíklegt sé að það komist upp um mig.	0,41	0,41	0,50(0,50)	4,95(1,67)
27. Ég hef baktalað vin/vinkonu. *	0,69	0,51	0,22(0,42)	3,75(1,75)
28. Þegar ég heyrri fólk tala saman í trúnaði forðast ég að hlusta á samræðurnar.	0,53	0,37	0,36(0,48)	4,60(1,70)
29. Ég hef fengið of mikinn pening til baka í verslun án þess að segja afgreiðslufólkinu frá því. *	0,36	0,32	0,46(0,50)	4,65(2,18)
30. Ég tel alltaf fram allan tollskyldan varning.	0,54	0,42	0,20(0,40)	3,15(2,09)
31. Þegar ég var ung/ur stal ég stundum. *	0,38	0,40	0,50(0,50)	4,79(2,19)
32. Ég hef aldrei skilið eftir mig rusl á götum.	0,38	0,34	0,25(0,43)	3,27(2,15)
33. Ég ek stundum yfir hámarkshraða. *	0,46	0,43	0,11(0,31)	2,74(1,72)
34. Ég les aldrei bækur eða tímarit með kynþokkafullu innihaldi.	0,31	0,26	0,14(0,35)	2,83(1,88)
35. Ég hef gert hluti sem ég segi ekki öðru fólki frá. *	0,32	0,31	0,15(0,35)	2,73(1,89)
36. Ég tek aldrei hluti sem ég á ekki.	0,48	0,33	0,59(0,49)	5,16(1,96)
37. Ég hef tekið mér veikindaleyfi frá vinnu eða skóla þrátt fyrir að vera ekki veik/ur. *	0,46	0,42	0,46(0,50)	4,39(2,31)
38. Ég hef aldrei skemmt bókasafnsbók eða vöru í verslun án þess að láta vita af því.	0,62	0,43	0,72(0,45)	5,60(1,93)
39. Sumar venjur mínar eru mjög slæmar. *	0,56	0,48	0,26(0,44)	3,93(1,76)
40. Ég slúðra ekki um mál annarra.	0,59	0,43	0,30(0,46)	4,19(1,70)

\* Svör við þessum atriðum eru endurköðuð við tölfræðilega úrvinnslu, <sup>a</sup> Þáttahleðsla ekki marktæk (>0,05), T = tvíkosta skorun, S = samfelld skorun, gildi í sviga eru staðalfrávik.

þeir leggðu mat á gagnrýni og tækju gagnrýni til umhugsunar og hins vegar hvort þeir væru ánægðir með að fá gagnrýni. Mögulegt er að þetta megi rekja til vandamáls við þýðingu en jafnframt gæti verið að sama vandamál sé til staðar í upprunalegri útgáfu kvarðans. Orðið „*appreciate*“ er þar notað en túlka má það orð á samskonar hátt. Flestir töldu fullyrðinguna einungis eiga við um gagnrýni á þá sjálfa en einn viðmælandi taldi að einnig væri átt við gagnrýni í garð annarra. Þrátt fyrir að þátttakendur hafi ekki talið sig lenda í vandræðum með atriði 19 „*Það er í lagi mín vegna þó einhverju fólki kunnir að líka illa við mig*“ lögðu ekki allir sama skilning á

hugtakið „*einhverju fólki (e. some people)*“. Hluti viðmælanda taldi að eingöngu væri átt við ókunnuga en aðrir töldu að þeir sem væru nákomnir gætu einnig fallið undir þennan hóp. Nokkrir viðmælandur töldu orðið „*sumir*“ eiga betur við í þessu tilviki þar sem þeir töldu að með því hugtaki væri bæði átt við ókunnuga og þá sem stæðu þeim nær. Aðrir viðmælandur töldu ekki skipta máli hvort orðið væri notað og töldu orðin hafa sömu merkingu. Talið er að þetta vandamál megi þó frekar rekja til upprunalegrar útgáfu atriðisins fremur en þýðingu. Orðið „*notfært (e. taken advantage)*“ í atriði 23 „*Það hefur komið fyrir að ég hef notfært mér aðra*“ var túlkað á ólíkan hátt

meðal þátttakenda. Sumir viðmælendur töldu orðið vísa til þess að nýta sér aðstoð annarra en aðrir töldu að orðið hefði neikvæðari merkingu og hefði slæmar afleiðingar fyrir hinn aðilann. Þeir þátttakendur sem töldu að orðið hefði neikvæða merkingu töldu atriðið of sterkt orðað. Atriði 25 „*Ég reyni stundum að hefna mín frekar en að fyrirgefa*“ var túlkað á ólíkan hátt meðal þátttakenda. Í fyrsta lagi töldu þátttakendur orðið „*hefna*“ eiga við um misalvarlega atburði. Sumir töldu orðið eiga við um mjög alvarlega hegðun á meðan að aðrir tengdu orðið við léttvægari hegðun, svo sem að hefna sín í fótbolta. Í öðru lagi var breytilegt hvernig þátttakendur túlkuðu orðið „*fyrirgefa*“ en sumir töldu sig vera búna að fyrirgefa þegar þeir hættu að hugsa um atburðinn en aðrir töldu fyrirgefningu fela í sér að ræða við viðkomandi aðila. Þátttakendur voru því ekki alveg sammála um skilgreiningu hefndar og fyrirgefningar. Við þýðingu þessa atriðis var ákveðið að þýða atriðið „*I sometimes try to get even rather than forgive and forget*“ ekki beint þar sem það þótti ekki nægilega auðskiljanlegt á íslensku máli. Þýðendur voru meðvitaðir um að íslensk þýðing atriðisins væri ef til vill of sterkt orðuð og er því mikilvægt að líta til niðurstöðu tölfræðilegar greiningar við ákvörðun á því hvort breyta þurfi þýðingu þessa atriðis. Sumir þátttakendur töldu atriði 39 „*Sumar venjur mínar eru mjög slæmar*“ mögulega vera of sterkt orðað og töldu að fjarlægja þyrfti orðið „*mjög*“ þar sem það gæfi til kynna venjur sem hefðu mikil áhrif á líf og heilsu viðkomandi. Aðrir töldu sig ekki lenda í vandræðum með þessa fullyrðingu og nefndu mörg dæmi um slíkar venjur, svo sem reykingar, hreyfingarleysi, áfengisvandamál, óstundvísi, kæki og fleira. Hafa ber þó í huga að í upprunalegri útgáfu kvarðans er fjallað um „*pretty awful habits*“ þannig að ekki er talið að íslensk þýðing sé of sterkt orðuð. Atriði 13 „*Ég kys vegna þess að atkvæði mitt getur skipt sköpum*“ var túlkað af flestum viðmælendum á þann veg að atkvæði skipti sköpum ef það ræður úrslitum kosninga en nokkrir töldu atkvæðið ekki þurfa að ráða

úrslitum. Sumir viðmælendur töldu að atriðið væri það sterkt orðað að erfitt væri að svara staðhæfingunni á þann veg að hún ætti að öllu leyti við um viðkomandi. Upprunalega útgáfan notar enska orðalagið „*make a difference*“ og er því mögulegt að íslenska þýðingin sé of sterkt orðuð. Töldu þessir viðmælendur að orða þyrfti þessa staðhæfingu á vægari hátt til að fá meiri dreifingu á svörum. Atriði 13 var eitt þeirra atriða sem var erfitt í þýðingu.

Álitaefni höfðu einnig komið upp við þýðingu annarra atriða og voru þau skoðuð sérstaklega í viðtölunum. Þetta voru atriði 3, 6 og 24. Niðurstöður viðtala veittu frekari upplýsingar um skilning þátttakenda á þeim atriðum. Atriði 3, „*Ég kæri mig ekki um að vita hvaða álit aðrir hafa í raun á mér*“, reyndist nokkuð flókið í þýðingu. Nokkrir þátttakendur töldu orðið „*vita*“ hafa áhrif á skilning þeirra á atriðinu, þar sem þeir vildu síður vita álit annarra þrátt fyrir að hafa áhyggjur af álit þeirra. Enska orðið „*know*“ er þó notað í upprunalegri útgáfu listans þannig að ekki er talið að það vandamál tengist þýðingu kvarðans. Alvarlegasta vandamálið við þetta atriði virðist þó stafa af tvöfaldri neitun sem einnig er til staðar í upprunalegri útgáfu kvarðans. Af tuttugu þátttakendum, svöruðu sjö þátttakendur atriðinu öfugt. Talið er að þetta vandamál sé til staðar í fleiri atriðum í kvarðanum og er því þörf á að skoða nánar vandamál sem kunna að stafa af innihaldi atriðanna fremur en þýðingu atriðanna í framhaldi af þessari rannsókn. Atriði 6, „*Þegar ég er í tilfinningalegu uppnámi hugsa ég ekki skýrt*“, reyndist einnig fremur flókið í þýðingu en niðurstöður viðtala bentu til þess að skilningur þátttakenda hafði verið svipaður og túlkun þýðenda á upprunalegri útgáfu atriðins, það er að „*tilfinningalegt uppnám*“ gæti átt við jákvætt og neikvætt tilfinningalegt ástand. Ekki var því talin ástæða til að breyta þýðingu þessa atriðis. Atriði 34, „*Ég les aldrei bækur eða tímarit með kynþokkafullu innihaldi*“, reyndist fremur erfitt í þýðingu. Þátttakendur túlkuðu þessa spurningu á ólíkan hátt og þá sérstaklega hvers kyns ritefni félli undir þessa

skilgreiningu. Þátttakendur töldu þó samskonar bækur og tímarit falla undir íslenska þýðingu atriðisins og upprunalega útgáfu atriðisins þannig að ekki er talið að vankantar séu á þýðingu atriðisins. Vinsældir bókar á þessu sviði er þó talin hafa áhrif á svör þátttakenda við þessu atriði þar sem flestir nefndu ákveðna bók sem dæmi um slíkt ritefni. Enn fremur er mögulegt að vinsældir þeirrar bókar hafi haft áhrif á lága fylgni atriðisins við önnur atriði þar sem þátttakendur voru opinskáir varðandi lestur bókarinnar og virtust ekki upplifa þá hegðun sem félagslega óæskilega.

## Umræða

BIDR-6 kvarðinn á að mæla félagslega æskilega svörun (FÆS) og er yfirleitt notaður til að kanna hvort mælingar með öðrum sjálfsmatskvörðum séu bjagaðar vegna slíkrar svörunar. Ástæða er til að ætla að bjögun vegna FÆS sé vandamál þegar íslenskir sjálfsmatskvarðar eru notaðir ekki síður en erlendir. Því er þörf á tæki eins og BIDR-6 kvarðanum hér á landi, sem mælir FÆS á áreiðanlegan og réttmætan hátt. Markmið rannsóknarinnar var að þýða BIDR-6 úr ensku yfir á íslensku og kanna síðan próffræðilega eiginleika íslenskrar útgáfu kvarðans. Helstu niðurstöður benda til þess að þýðing mælitækisins hafi tekist nokkuð vel. Meðaltöl voru á svipuðu bili og erlendis en ímyndarstjórnun mældist þó aðeins hærri hér á landi en erlendis (Holden o.fl., 2000; Li og Reb, 2009; Stöber o.fl., 2002). Nokkrar mögulegar ástæður eru taldar geta útskýrt mun á meðaltölum ÍS sem kanna þyrfti nánar. Í fyrsta lagi gæti þýðingu atriðanna sem mynda þáttinn verið ábótavant. Í öðru lagi gæti verið til staðar menningarmunur á skilningi atriðanna, tíðni hegðunar og hversu tilbúið fólk er til að viðurkenna hegðunina en rannsóknir benda til þess að menningarmunur geti verið í svörum við BIDR (Lalwani, Shavitt og Johnson, 2006; Lalwani, Shrum og Chiu, 2009). Í þriðja lagi gæti framkvæmd rannsóknar hafa haft áhrif og þeir þátttakendur sem notaðir voru í úrvinnslu einfaldlega frekar haft tilhneigingu

til að reyna að ýkja jákvæða eiginleika í eigin fari en erlendir samanburðarhópar. Aðeins voru þeir þátttakendur notaðir í úrvinnslu gagna sem svöruðu öllum atriðum á listanum. Jákvæð fylgni hefur verið á milli samviskusemi á NEO-FFI persónuleikaprófinu og ÍS í fyrri rannsóknum (Stöber o.fl., 2002) og neikvæð fylgni milli samviskusemi og þess að sleppa þátttöku eða einstökum atriðum (Rogelberg, Conway, Sederburg, Spitzmüller, Aziz og Knight, 2003). Mögulegt er að þeir þátttakendur séu hærri á ÍS en þeir sem ekki svöruðu öllum atriðum og gæti það mögulega skýrt herra meðaltal á íslenska kvarðanum. Í þeim erlendu úrtökum sem notuð hafa verið til samanburðar hefur kvarðinn oftast verið lagður fyrir í kennslustund þar sem svarhlutfall hefur verið mun herra (t.d. Pauls og Crost, 2004 og Pauls og Stemmler, 2003). Innri áreiðanleiki beggja undirkvarða var hærri en 0,7 og bendir það til þess að áreiðanleiki íslenskrar útgáfu BIDR-6 sé ásættanlegur (DeVellis, 2003).

Niðurstöður staðfestandi þáttgreiningar styðja réttmæti ÍS kvarðans með tvíkösta skorun en ákveðin vandamál virðast vera til staðar í tengslum við mátgæði ÍS með samfelldri skorun og SB með tvíkösta og samfelldri skorun sem skoða þarf nánar. Þetta vandamál hefur einnig komið fram í erlendum rannsóknum sem notað hafa staðfestandi þáttgreiningu þannig að þetta virðist vera vandamál tengt upprunalega listanum fremur en þýðingu kvarðans (Cervellione o.fl., 2009; Lanyon og Carle, 2007; Leite og Beretvas, 2005). Mikilvægt væri þó að skoða próffræðilega eiginleika kvarðans nánar með öðru og stærra úrtaki. Enn fremur mætti skoða nánar ákveðin atriði sem höfðu lágur þáttahleðslur. Niðurstöður viðtala bentu til þess að vandamál við þýðingu gætu verið til staðar varðandi atriði 13, 16, 23, 25 og 39. Talið var að orðalag atriða 13, 23, 25 og 39 gæti verið of sterkt í íslenskri þýðingu. Af þessum atriðum er það helst atriði 13 sem kom jafnframt ekki nægilega vel út úr tölfraðilegri greiningu. Þegar svör fólks við því atriði voru skoðuð nánar bentu niðurstöður þó til þess að yfir helmingur svarenda valdi tvo sterkustu

svarmöguleikana og voru því mjög sammála fullyrðingunni. Meðaltal þessa atriðis var með því hæsta í rannsókninni af öllum atriðum og studdi því tölfraðileg greining ekki að atriðið væri of sterkt orðað. Atriði 23, 25 og 39 komu einnig ágætlega út úr tölfraðilegri greiningu og bentu meðaltöl atriða ekki til þess að atriðin væru of sterkt orðuð. Mögulegt er að hægt sé að bæta þýðingu atriða 16, sem er líklega of flókið í íslenskri þýðingu en tölfraðileg greining benti þó til þess að atriðið væri ásættanlegt.

Niðurstöður staðfestandi þáttgreiningar bentu einnig til þess að tvíkosta skorun komi betur út en samfelld skorun sem er í samræmi við niðurstöður Leite og Beretvas (2005). Samkvæmt Paulhus (1984) er ætlunin með BIDR að bera kennsl á ýkta svörum og benda niðurstöður þessarar rannsóknar til þess að tvíkosta skorun geti verið ákjósanlegri. Niðurstöður viðtala benda jafnframt til þess að fjöldi svarmöguleika sé ef til vill of mikill og fólki finnist erfitt að velja á milli sjö valmöguleika þegar kemur að því hversu sönn fullyrðing er. Innri áreiðanleikinn var þó örlítið hærri fyrir samfellda skorun í samræmi við niðurstöður Stöber og félaga (2002). Mikilvægt væri því að framkvæma frekari rannsóknir til að kanna mun á tvíkosta og samfelldri skorun þar sem niðurstöður annarra rannsókna (Stöber o.fl., 2002) benda til þess að samfelld skorun sé ákjósanlegri með tilliti til áreiðanleika, samleitniréttmætis og fylgni við persónuleikapróf.

Fyrstu niðurstöður benda því til að íslensk útgáfa BIDR-6 kvarðans sé sambærileg upprunalegri útgáfu kvarðans og að áreiðanleikinn sé ásættanlegur. Réttmæti ÍS kvarðans með tvíkosta skorun er viðunandi en nauðsynlegt er að kanna nánar réttmæti íslensku útgáfunnar. Rétt er að benda á að þessi rannsókn er einungis fyrsta skrefið í mati á próffraðilegum eiginleikum íslenskrar þýðingar BIDR-6 kvarðans. Því er mikilvægt að gera frekari rannsóknir á áreiðanleika og

réttmæti BIDR-6 kvarðans áður en hægt er að mæla með almennri notkun hans. Það er til að mynda mikilvægt að kanna próffraðilega eiginleika hans í breiðara og stærra úrtaki. Einnig er nauðsynlegt að kanna réttmæti kvarðans með fjölbreyttari aðferðum eins og fylgni við aðra kvarða sem mæla kenningarlega tengd hugtök eða gagnsemi hans við að greina á milli þeirra sem svara á félagslega æskilegan hátt á öðrum mælingum. Eins er talið mikilvægt að framkvæma ítarlegri tölfraðigreiningar, svo sem með notkun svarferlálkanna (e. *item response theory*), þar sem kanna mætti nánar hvernig svarmöguleikarnir sjö nýtast. Einnig væri gagnlegt að fá frekari upplýsingar um skilning fjölbreyttari hóps á hugtökum í upprunalegri útgáfu kvarðans þrátt fyrir að tvítyngdir þýðendur hafi komið að þýðingu þessa mælitækis. Gagnlegt hefði verið að hafa aðgang að slíkum gögnum svo hægt væri að leggja betur mat á það hvort vandamál stafi af þýðingu eða innihaldi mælitækja. Mögulegt væri sem dæmi að leggja upprunalegu útgáfu mælitækisins fyrir nokkra viðmælendur sem annaðhvort hafa það tungumál sem móðurmál eða hafa dvalið lengi í landi þar sem tungumálið er talað. Að lokum gæti verið áhugavert að kanna hvort útbúa mætti styttri útgáfu af mælitækinu með því að fjarlægja atriði sem virðast eiga lítið sameiginlegt með öðrum atriðum. Dæmi eru um að styttrar útgáfur af kvarðanum hafi verið notaðar erlendis af einstaka rannsakendum en þær útgáfur hafa verið próaðar út frá einföldum aðferðum (Leite og Beretvas, 2005; Pauls og Stemmler, 2003; Stöber o.fl., 2002). Mælitækið er fremur langt og samanstandur af 40 atriðum og virðist því vera þörf á styttri útgáfu kvarðans. Niðurstöður viðtala benda enn fremur til þess að þrátt fyrir að ekki sé talið að vandamál sé til staðar varðandi þýðingu atriða séu ákveðin vandamál með upprunalegri atriði kvarðans, svo sem varðandi tvöfalda neitun. Mikilvægt er talið að kafa frekar ofan í þau atriði með tilliti til innihalds frekar en þýðingar.

## **Socially desirable responding: The psychometric properties of the Icelandic version of the Balanced Inventory of Desirable Responding**

The Balanced Inventory of Desirable Responding (BIDR) is one of the most commonly used measures of socially desirable responding. It consists of two scales, Self-Deceptive Enhancement (SDE) and Impression Management (IM), containing 20 statements each, answered on a 7-point scale (1=Not true, 4=Somewhat true, 7=Very true). The purpose of this paper was to translate the BIDR to Icelandic and examine the psychometric properties of the Icelandic translation using Confirmatory Factor Analysis (CFA) and

cognitive interviews with probing. The first phase of the study (N=321) focuses on the translation process and the psychometric properties of the BIDR. The results indicated that the means, intercorrelation between the two scales, and CFA results are comparable to those found in previous studies and the reliability is acceptable, although the results indicated certain items were problematic. In the second phase of the study (N=20) cognitive interviews with probing were used in order to identify potential problems with regards to the translation. The results suggest the Icelandic translation is adequate.

**Keywords:** Socially desirable responding, confirmatory factor analysis, psychometrics, translation, BIDR

## Heimildir

- Albright, J. J. og Park, H. M. (2009). Confirmatory factor analysis using Amos, LISREL, Mplus, and SAS/STAT CALIS. *Technical working paper*. Bloomington, IL: Indiana University.
- Barrick, M. R. og Mount, M. K. (1991). The Big Five personality dimensions and job performance: A meta-analysis. *Personnel Psychology*, 44, 1-26. <http://dx.doi.org/10.1111/j.1744-6570.1991.tb00688.x>
- Barrick, M. R. og Mount, M. K. (1996). Effects of impression management and self-deception on the predictive validity of personality constructs. *Journal of Applied Psychology*, 81, 261-272. <http://dx.doi.org/10.1037/0021-9010.81.3.261>
- Beatty, P. C. og Willis, G. B. (2007). Research synthesis: The practice of cognitive interviewing. *Public Opinion Quarterly*, 71(2), 287-311. <http://dx.doi.org/10.1093/poq/nfm006>
- Behling, O. og Law, K. S. (2000). *Translating questionnaires and other research instruments: Problems and solutions*. London: SAGE Publications. <http://dx.doi.org/10.4135/9781412986373>
- Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, 107(2), 238-246. <http://dx.doi.org/10.1037/0033-2909.107.2.238>
- Beretvas, S. N., Meyers, J. L. og Leite, W. L. (2002). *A reliability generalization study of the Marlowe-Crowne Social Desirability Scale*. *Educational and Psychological Measurement*, 62(4), 570-589. <http://dx.doi.org/10.1177/0013164402062004003>
- Cervellione, K. L., Lee, Y. S. og Bonanno, G. A. (2009). Rasch modeling of the self-deception scale of the balanced inventory of desirable responding. *Educational and Psychological Measurement*, 69(3), 438-458. <http://dx.doi.org/10.1177/0013164408322020>
- Costa, P. T. og MacCrae, R. R. (1992). *Revised NEO Personality Inventory (NEO PI-R) and NEO Five-Factor Inventory (NEO FFI) professional Manual*. Odessa, FL: Psychological Assessment Resources.
- Crowne, D. P. og Marlowe, D. (1960). A new scale of social desirability independent of psychopathology. *Journal of Consulting Psychology*, 24(4), 349-354. <http://dx.doi.org/10.1037/h0047358>
- DeVellis, R. F. (2003). *Scale development: Theory and applications*. Thousand Oaks: Sage.
- Einar Guðmundsson (2005-2006). Þýðing og staðfærsla sálfræðilegra prófa. *Sálfræðiritið – Tímarit Sálfræðingafélags Íslands*, 10 – 11, 23-40. <http://dx.doi.org/2336/11992>
- Flora, D. B. og Curran, P. J. (2004). An empirical evaluation of alternative methods of estimation for confirmatory factor analysis with ordinal data. *Psychological Methods*, 9(4), 466-491. <http://dx.doi.org/10.1037/1082-989x.9.4.466.supp>
- Floyd, F. J. og Widaman, K. F. (1995). Factor analysis in the development and refinement of clinical assessment instruments. *Psychological Assessments*, 7(3), 286-299. <http://dx.doi.org/10.1037/1040-3590.7.3.286>
- Fowler, F. J. (1995). *Improving survey questions: Design and evaluation*. Thousand Oaks: Sage.
- Geisinger, K. F. (1994). Cross-cultural normative assessment: Translation and adaptation issues influencing the normative interpretation of assessment instruments. *Psychological Assessment*, 6(4), 304-312. <http://dx.doi.org/10.1037//1040-3590.6.4.304>
- Hambleton, R. K., Merenda, P. og Spielberger, C. (2005). *Adapting educational and psychological tests for cross-cultural assessment*. Hillsdale, NJ: Lawrence S. Erlbaum.
- Helmes, E. og Holden, R. R. (2003). The construct of social desirability: one or two dimensions? *Personality and Individual Differences*, 34, 1015-1023. [http://dx.doi.org/10.1016/S0191-8869\(02\)00086-7](http://dx.doi.org/10.1016/S0191-8869(02)00086-7)
- Holden, R. R., Starzyk, K. B., McLeod, L. D. og Edwards, M. J. (2000). Comparisons among the Holden Psychological Screening Inventory (HPSI), the Brief Symptom Inventory (BSI), and the Balanced Inventory of Desirable Responding (BIDR). *Assessment*, 7(2), 163-175. <http://dx.doi.org/10.1177/107319110000700208>
- Hu, L. T. og Bentler, P. M. (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psychological Methods*, 3(4), 424-453. <http://dx.doi.org/10.1037/1082-989X.3.4.424>
- Hu, L. T. og Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, 6(1), 1-55. <http://dx.doi.org/10.1080/1070519909540118>
- Jöreskog, K. G. og Sorbom, D. (2012). LISREL 9.1 [hugbúnaður]. *Lincolnwood, IL: Scientific Software International*.
- Lalwani, A. K., Shavitt, S. og Johnson, T. (2006). What is the relation between cultural orientation and socially desirable responding? *Journal of Personality and Social Psychology*, 90(1), 165-178. <http://dx.doi.org/10.1037/0022-3514.90.1.165>



- Lalwani, A. K., Shrum, L. J. og Chiu, C. Y. (2009). Motivated response styles: The role of cultural values, regulatory focus, and self-consciousness in socially desirable responding. *Journal of Personality and Social Psychology*, 96(4), 870-882. <http://dx.doi.org/10.1037/a0014622>
- Lanyon, R. I. og Carle, A. C. (2007). Internal and external validity of scores on the Balanced Inventory of Desirable Responding and the Paulhus Deception Scales. *Educational and Psychological Measurement*, 67(5), 859-876. <http://dx.doi.org/10.1177/0013164406299104>
- Leite, W. L. og Beretvas, S. N. (2005). Validation of scores on the Marlowe-Crowne social desirability scale and the Balanced Inventory of Desirable Responding. *Educational and Psychological Measurement*, 65(1), 140-154. <http://dx.doi.org/10.1177/0013164404267285>
- Li, A. og Bagger, J. (2006). Using the BIDR to distinguish the effects of impression management and self-deception on the criterion validity of personality measures: A meta-analysis. *International Journal of Selection and Assessment*, 14(2), 131-141. <http://dx.doi.org/10.1111/j.1468-2389.2006.00339.x>
- Li, A. og Bagger, J. (2007). The Balanced Inventory of Desirable Responding (BIDR): A Reliability Generalization Study. *Educational and Psychological Measurement*, 67(3), 525-544. <http://dx.doi.org/10.1177/0013164406292087>
- Li, A. og Reb, J. (2009). A cross-nations, cross-cultures, and cross-conditions analysis on the equivalence of the Balanced Inventory of Desirable Responding. *Journal of Cross-Cultural Psychology*, 40(2), 214-233. <http://dx.doi.org/10.1177/0022022108328819>
- Marsh, H. W., Hau, K. T. og Wen, Z. (2004). In search of golden rules: Comment on hypothesis-testing approaches to setting cutoff values for fit indexes and dangers in overgeneralizing Hu and Bentler's (1999) findings. *Structural Equation Modeling*, 11(3), 320-341. [http://dx.doi.org/10.1207/s15328007sem1103\\_2](http://dx.doi.org/10.1207/s15328007sem1103_2)
- Markland, D. (2007). The golden rule is that there are no golden rules: A commentary on Paul Barrett's recommendations for reporting model fit in structural equation modeling. *Personality and Individual Differences*, 42(5), 851-858. <http://dx.doi.org/10.1016/j.paid.2006.09.023>
- Mount, M. K., Barrick, M. R. og Stewart, G. L. (1998). Personality predictors of performance in jobs involving interaction with others. *Human Performance*, 11(3), 145-166. [http://dx.doi.org/10.1207/s15327043hup1102&3\\_3](http://dx.doi.org/10.1207/s15327043hup1102&3_3)
- Ones, D. S. og Viswesvaran, C. (1998). The effects of social desirability and faking on personality and integrity assessment for personnel selection. *Human Performance*, 11(2-3), 245-269. [http://dx.doi.org/10.1207/s15327043hup1102&3\\_7](http://dx.doi.org/10.1207/s15327043hup1102&3_7)
- Paulhus, D. L. (1984). Two-component models of socially desirable responding. *Journal of Personality and Social Psychology*, 46(3), 598-609. <http://dx.doi.org/10.1037//0022-3514.46.3.598>
- Paulhus, D. L. (1991). Measurement and control of response bias. Í J. P. Robinson, P. R. Shaver og L. S. Wrightsman (Ritstj.), *Measures of Personality and Social Psychological Attitudes* (bls. 17-59). San Diego: Academic Press. <http://dx.doi.org/10.1016/b978-0-12-590241-0.50006-x>
- Paulhus, D. L. (1994). *Balanced inventory of desirable responding: Reference manual for BIDR version 6*. Óbirt handrit, University of British Columbia, Vancouver, Canada.
- Paulhus, D. L. (2002). Socially desirable responding: The evolution of a construct. Í H. I. Braun, D. N. Jackson og D. E. Wiley (Ritstj.), *The role of constructs in psychological and educational measurement* (bls. 49-69). Mahwah NJ: Lawrence Erlbaum Associates, Inc.
- Pauls, C. A. og Crost, N. W. (2004). Effects of faking on self-deception and impression management scales. *Personality and Individual Differences*, 37(6), 1137-1151. <http://dx.doi.org/10.1016/j.paid.2003.11.018>
- Pauls, C. A. og Stemmler, G. (2003). Substance and bias in social desirability responding. *Personality and Individual Differences*, 35(2), 263-275. [http://dx.doi.org/10.1016/S0191-8869\(02\)00187-3](http://dx.doi.org/10.1016/S0191-8869(02)00187-3)
- Presser, S. og Blair, J. (1994). Survey pretesting: Do different methods produce different results. *Sociological Methodology*, 24(1), 73-104. <http://dx.doi.org/10.2307/270979>
- Rogelberg, S. G., Conway, J. M., Sederburg, M. E., Spitzmüller, C., Aziz, S. og Knight, W. E. (2003). Profiling active and passive nonrespondents to an organizational survey. *Journal of Applied Psychology*, 88(6), 1104-1114. <http://dx.doi.org/10.5465/ambpp.2003.13792519>
- Sackeim, H. A. og Gur, R. C. (1979). Self-deception, other-deception, and self-reported psychopathology. *Journal of Consulting and Clinical Psychology*, 47(1), 213-215. <http://dx.doi.org/10.1037/0022-006X.47.1.213>
- Salgado, J. F. (2003). Predicting job performance using FFM and non-FFM personality measures. *Journal of Occupational and Organizational Psychology*, 76(3), 323-346. <http://dx.doi.org/10.1348/096317903769647201>

- Steiger, J. H. (1990). Structural model evaluation and modification: An interval estimation approach. *Multivariate behavioral research*, 25(2), 173-180. [http://dx.doi.org/10.1207/s15327906mbr2502\\_4](http://dx.doi.org/10.1207/s15327906mbr2502_4)
- Stöber, J., Dette, D. E. og Musch, J. (2002). Comparing continuous and dichotomous scoring of the Balanced Inventory of Desirable Responding. *Journal of Personality Assessment*, 78(2), 370-389. [http://dx.doi.org/10.1207/S15327752JPA7802\\_10](http://dx.doi.org/10.1207/S15327752JPA7802_10)
- Vispoel, W. P. og Tao, S. (2013). A generalizability analysis of score consistency for the Balanced Inventory of Desirable Responding. *Psychological Assessment*, 25(1), 94-104. <http://dx.doi.org/10.1037/a0029061>
- West, S., Taylor, A. og Wu, W. (2012). Model fit and model selection in structural equation modeling. Í R. H. Hoyle (Ritstj.), *Handbook of Structural Equation Modeling*. New York: Guilford.
- Yu, C. Y. (2002). *Evaluating cutoff criteria of model fit indices for latent variable models with binary and continuous outcomes*. Óbirt Doktorsritgerð, University of California, Los Angeles.
- Yuan, K. H. (2005). Fit indices versus test statistics. *Multivariate Behavioral Research*, 40(1), 115-148. [http://dx.doi.org/10.1207/s15327906mbr4001\\_5](http://dx.doi.org/10.1207/s15327906mbr4001_5)
- Zerbe, W. J. og Paulhus, D. L. (1987). Socially desirable responding in organizational behavior: A reconception. *Academy of Management Review*, 12(2), 250-264. <http://dx.doi.org/10.2307/258533>