

Pýðing og staðfærsla sálfræðilegra prófa

Einar Guðmundsson
Háskóli Íslands

Þýdd og óstöðluð sálfræðileg próf og kvarðar eru algeng í klínísku starfi hérlandis. Eigi að síður eru litlar eða engar upplýsingar um hvernig staðið hefur verið að þýðingu og staðfærslu þeirra og hverjir próffræðilegir eiginleika þeirra eru. Skortur á upplýsingum um þessi atriði er alvarlegur, þegar hagsmunir skjólstaðinga sálfræðinga eru hafðir í huga. Mikil þörf er á einhverjum viðmiðum, reglum eða leiðbeiningum, um hvaða lágmarksskilyrði tiltekið matstæki þarf að uppfylla til að notkun þess í klínísku starfi, í formlegu eða óformlegu samhengi, geti talist fagleg og siðleg. Megintilgangur þessarar greinar er að fjalla um slík viðmið. Í greininni eru sett fram níu viðmið til að meta hvort tiltekið sálfræðilegt próf eða kvarði geti talist þýtt og staðfært. Færð eru rök fyrir því að einungis þegar þýðing og staðfærsla á tilteknu matstæki hefur verið unnin að lágmarki í samræmi við þessi viðmið og gögn verið birt á viðeigandi vettvangi sé réttlætanlegt að tala um að þýdd og staðfærð útgáfa matstækisins sé til. Af því leiðir jafnframt að einungis er faglega og siðlega réttlætanlegt að nota óstaðlaða útgáfu prófs eða kvarða í klínísku starfi þegar þýðing og staðfærsla þess hefur verið unnin í samræmi við öll viðmiðin nfu. Þá er miðað við að það sé metið svo, á grundvelli raunvísra gagna, að það sé betra að nota óstaðlað, þýtt og staðfært matstæki, fremur en að nota það ekki. Lokatakmark allra þýðinga og staðfærslu sálfræðilegra prófa og kvarða hlýtur þó að vera stöðlun þeirra.

Notkun þýddra og óstaðlaðra sálfræðilegra prófa er algeng í klínísku starfi sálfræðinga hérlandis. Með þýddu prófi er átt við þýðingu á erlendum texta yfir á íslensku þar sem viðurkenndu vinnuferli við þýðingu matstækja hefur verið fylgt eða ekki. Með óstöðluðu prófi er eingöngu átt við matstæki sem hefur verið staðlað í heimalandi en ekki hérlandis. Hér er því ekki verið að fjalla um matstæki sem hvergi hafa verið stöðluð enda varasamt að nota slík matstæki í klínísku starfi sálfræðinga eða í öðru samhengi. Staðlað próf vísar til þess að form, innihald, matsreglur og reglur um fyrirlögn prófsins eru fyrirfram ákveðin (staðlað) og sömuleiðis að viðeigandi gögnum hefur verið safnað til að útbúa norm fyrir prófið. Þegar þýdd og óstöðluð próf eru notuð hérlandis heyrir til undantekninga að hreyft sé við framangreindum atriðum enda varhugavert að gera það án undangenginnar rannsóknar. Þegar próf eru stöðluð hérlandis er

aftur á móti algengt að breyta þurfi innihaldi prófs, reglum um mat á svörum og reglum um fyrirlögn. Þetta hefur í flestum tilvikum áhrif á notkun og túlkun prófsins. Að auki er safnað gögnum til að útbúa norm í úrtaki sem endurspeglar sérkenni þýðis.

Fá matstæki eru til fyrir fullorðna hérlandis sem uppfylla ströngustu skilyrði staðlaðra prófa. Fleiri stöðluð sálfræðileg próf eru til handa börnum og unglingum. Þýdd og óstöðluð sálfræðileg próf og kvarðar af ýmsu tagi eru þó mest notuð í fjölbreytilegu starfi með börnum, unglingum og fullorðnum.

Þó flestir sálfræðingar geri sér grein fyrir vanköntum á notkun óstaðlaðra matstækja er það líklega mat jafnmargra, með réttu eða röngu, að það sé betri kostur að nota þau en ekki í klínísku starfi. Útbreidd notkun þeirra bendir til þess. Þetta er skiljanlegt í ljósi þess hversu stór hluti af starfi sálfræðinga er greiningarvinna og mat á árangri

meðferðarstarfs. Þar gegna áreiðanleg og réttmæt matstæki lykilhlutverki. Vel samin, stöðluð, áreiðanleg og réttmæt, sálfræðileg próf eru ein af meginundirstöðum faglegra vinnubragða sálfræðinga. Gott framboð slíkra matstækja er því eitt af skilyrðum fagmennsku í klínísku starfi þeirra.

Margar spurningar vakna þegar óstöðluð sálfræðileg matstæki eru skoðuð hérlandis. Það er nánast regla að engar skriflegar upplýsingar eru til um hvernig staðið hefur verið að þýðingu og staðfærslu þeirra. Í flestum tilvikum eru engin aðgengileg svör við spurningum um: (a) ástæður þess að tiltekið matstæki var þýtt, (b) hvaða aðferð var notuð við þýðingu, (c) hvort matstækið var eingöngu þýtt eða hvort einhverjar staðfærslur voru gerðar, (d) hæfni þýðenda, eða (e) hæfni og sérþekkingu aðstandenda tiltekinnar þýðingar á efni matstækis.

Upplýsingar um framangreind atriði eru frumskilyrði þess að hægt sé að meta hvort vinnubrögð og hæfni við þýðingu og staðfærslu matstækja eru viðunandi. Þær ættu því ávallt að vera aðgengilegar ef óstaðlað matstæki er notað í klínísku starfi. En fleira þarf að koma til. Það er nánast regla að ófullnægjandi eða engar upplýsingar liggja fyrir um: (a) áreiðanleika og réttmæti þýddra matstækja, (b) hversu víðeigandi erlend norm eru til að túlka niðurstöðu þýddra matstækja, og (c) skekkjur sem fylgja túlkun þversniða í þýddum matstækjum út frá erlendum gögnum. Skortur á upplýsingum um þessi atriði er alvarlegur, sérstaklega þegar hagsmunir skjólstæðinga sálfræðinga eru hafðir í huga.

Þeir sem þýða og staðfæra tiltekið matstæki, með eða án formlegs leyfis til þess, stuðla að útbreiðslu þess og notkun. Það er því eðlilegt að gera þá kröfu að hinir sömu axli faglega og siðferðilega ábyrgð af slíkri útbreiðslu matstækja. Ein leið til þess er að veita skriflegar upplýsingar um framangreind atriði. Með því móti getur væntanlegur notandi metið hversu réttlæt看legt er að nota matstækið og hvernig eða hvort eigi að nota það í formlegu eða óformlegu mati og greiningu. Þegar engar

skriflegar eða formlegar upplýsingar liggja fyrir um þýðingu prófs ætti ekki að nota það í formlegu samhengi, og takmarkað eða ekki í óformlegu samhengi. Útbreiðsla slíkra þýðinga, án nauðsynlegra upplýsinga eða gagna til að meta sérkenni þeirra, er sjaldan og sennilega aldrei réttlæt看leg.

Erfitt er að gera sér nákvæma grein fyrir því hvaða áhrif og afleiðingar óvissa um eiginleika óstaðlaðra þýddra prófa og kvarða hefur á störf sálfræðinga og skjólstæðinga þeirra. Erlendar reglur um faglega notkun sálfræðilegra prófa gera ráð fyrir að þau séu ekki notuð óstöðluð í klínísku starfi nema að sérstök rök komi til (American Educational Research Association, American Psychological Association, og National Council on Measurement in Education, 1999). Þessar reglur undirstrika því þann almenna skilning að ekki sé hægt að reikna með því án undangenginnar rannsóknar hvort erlend norm séu víðeigandi á milli landa. Þegar rök fyrir notkun norma við túlkun prófa eru höfð í huga er ljóst hversu varasöm sú iðja er. Líklegt er að slík notkun norma sé aldrei víðeigandi.

Vitneskju um afleiðingar þess að nota erlend norm við túlkun niðurstaðna úr þýddum prófum hérlandis er helst að finna í nokkrum réttmætisathugunum á eiginleikum þýddra greindarprófa og persónuleikaprófsins MMPI. Niðurstöður þessara réttmætisathugana undirstrika almennt að margvísleg skekkja, umfram skekkju sem áreiðanleikastuðlar endurspegla, er fylgífiskur notkunar þýddra, óstaðlaðra prófa hérlandis (Einar Guðmundsson, Sigurður J. Grétarsson, Sveinborg Kristjánsdóttir og Valka Jónsdóttir, 1993 – 1994; Einar Guðmundsson, Ásdís Clausen, Berglind S. Ásgeirsdóttir og Birgir Þór Guðmundsson, 2004; Evald Sæmundsen, Margrét Arnljótsdóttir og Jónas Halldórsson, 1990; Gylfi Ásmundsson, 1990). Sömu leiðis þarf að breyta viðmiðum um greiningu og túlkun þegar próf hafa verið stöðluð (Einar Guðmundsson og Birgir Þór Guðmundsson, 2004). Erlendar rannsóknir benda í sömu átt.

Pegar bandarísk norm voru notuð til að túlka frammistöðu 1100 kanadískra barna á WISC-III kom í ljós að frammistaða þeirra var yfir meðaltali á heildartölu greindar, öllum þáttum prófsins og flestum undirprófum (Saklofske, 2003). Þessar niðurstöður eru í samræmi við niðurstöður íslenskrar rannsóknar á áhrifum þess að nota bandaríska norm við túlkun á þýddri útgáfu WISC-III hérlandis. Þar kom meðal annars fram að hérlandis mælir prófið of hátt á heildartölu greindar, öllum prófhlutum og flestum undirprófum (Einar Guðmundsson, Ásdís Clausen, Berglind S. Ásgeirsdóttir og Birgir Þór Guðmundsson, 2004). Stöðlun á WISC-III í öðrum löndum undirstrikar enn frekar nauðsyn á staðfærslu margra atriða í þessu greindarprófi, frá einu landi til annars, og skekkjum sem fylgja notkun norma á milli landa (sjá nánar í Georgas, Weiss, van de Vijver og Saklofske, 2003).

Niðurstöður þessara rannsókna undirstrika með ótvíræðum hætti hversu varasamt er að nota erlend norm við túlkun niðurstaðna úr þýddum greindarprófum (WISC-III). Það sama á mjög líklega við um önnur matstæki. Þannig kom til dæmis fram í rannsókn Gylfa Ásmundssonar (1990) á eiginleikum íslenskrar þýðingar MMPI-prófsins að allir kvarðar prófsins mældu of hátt í íslensku úrtaki. Erlendis hafa margir gagnrýnt klíníska notkun MMPI-2 persónuleikaprófsins utan Bandaríkjanna vegna margvíslegs vanda við þýðingu og staðfærslu prófsins á milli menningarsvæða (sjá t.d. Dana, 1993, 1996; Nicholas, Padilla og Gomez-Maqueo, 2000). Alvarlegust er þó notkun þessa persónuleikaprófs, atferlislista og hliðstæðra matstækja án nokkurra réttmætisathugana á hugsanlegum skekkjum og ályktunarvillum samfara notkun erlendra norma og greiningarviðmiða við túlkun þeirra.

Pegar þýdd og óstöðluð matstæki eru notuð í klínísku starfi skiptir mestu að fyrir liggi raunvís gögn um hvaða skekkjur það eru sem fylgja notkun þeirra. Slík gögn eru jafnframt grundvöllur ákvarðanatöku um hvort réttlætanlegt er að nota óstaðlað matstæki

í klínísku starfi. Með réttmætisathugunum þýddra matstækja er hægt að draga úr margvíslegum ályktunarvillum við notkun þeirra. Það er því brýn þörf á því að fjölga rannsóknum af þessu tagi á meðan beðið er staðlaðrar útgáfu þeirra. En réttmætisathuganir á þýddum, óstöðluðum matstækjum geta þó aldrei komið í staðinn fyrir stöðluð matstæki eða bætt upp vankanta þeirra. Niðurstöður úr réttmætisathugunum ættu aftur á móti að hafa áhrif á umgengni og túlkun þessara matstækja í klínísku starfi á meðan beðið er stöðlunar þeirra. Og í sumum tilvikum að leiða til þess að óstöðluð matstæki séu ekki notuð í klínísku starfi. Lágmarksskilyrði er að sýnt sé fram á réttmæti þess að nota óstöðluð matstæki í klínísku starfi.

Hérlandis eru hvorki í gildi leiðbeinandi reglur um þýðingu og staðfærslu sálfræðilegra prófa né heldur hvaða lágmarksskilyrði þau þurfa að uppfylla til að réttlæta óstaðlaða notkun þeirra í klínísku starfi. Mikil þörf er á einhverjum viðmiðum, reglum eða leiðbeiningum, um hvaða lágmarksskilyrði tiltekið matstæki þarf að uppfylla til að notkun þess í klínísku starfi, í formlegu eða óformlegu samhengi, geti talist fagleg og siðleg. Megintilgangur þessarar greinar er að fjalla um slík viðmið.

Hér á eftir er gefið almennt yfirlit vinnuferlis við þýðingu og staðfærslu sálfræðilegra prófa og kvarða. Athyglinni er beint að því hvaða gögn og upplýsingar þurfa að liggja fyrir áður en þýtt og staðfært matstæki er notað í klínísku starfi. Lýst er þeim lágmarksskilyrðum sem þarf að uppfylla til að erlent matstæki geti talist þýtt og staðfært. Sú umfjöllun er á engan hátt tæmandi en er fyrsta skref í átt að einhvers konar faglegu samkomulagi um hvað felst í því að þýða og staðfæra próf og kvarða á milli landa.

Þýðing og staðfærsla sálfræðilegra próf og kvarða

Talsvert hefur verið ritað um þýðingu og staðfærslu matstækja milli menningarsvæða

(Bracken og Barona, 1991; Brislin, 1980, 1986; Geisinger, 1994; Hambleton, 1993, 1994; Tanzer og Sim, 1999; van de Vijver, 2000, 2003; van de Vijver og Hambleton, 1996; van de Vijver og Leung, 1997). Einkennandi fyrir þessa umfjöllun er þó skortur á raunvísunum gögnum um kosti og galla tiltekinna aðferða við þýðingu og staðfærslu matstækja.¹ Meginreglan er sú að aðferðir við þýðingu og staðfærslu prófa byggjast á uppsafnaðri reynslu af notkun þeirra fremur en formlegum rannsóknum á ágæti þessara aðferða. Jafnframt er byggt á þekktum tölfræðiaðferðum við athugun á eiginleikum og jafngildi prófa (*test equivalence*).

Reglur Alþjóðlegu prófaneftdarinnar

Fram til skamms tíma voru fáar leiðbeinandi reglur um þýðingu og staðfærslu prófa. Á þessu hefur orðið sú breyting á síðari árum að Alþjóðlega prófaneftdin (*International Test Commission*) hefur þróað sérstakar reglur í þessu skyni (Hambleton, 1994; van de Vijver og Hambleton, 1996; Tanzer og Sim, 1999).² Þær eru það nákvæmasta sem til er á þessu sviði og miðast við samningu jafngilda matstækja á tveimur eða fleiri tungumálum. Enn er ekki mikil reynsla af notkun þeirra (sjá þó Hambleton, Yu og Slater, 1999; Jeanrie og Bertrand, 1999). Eigi að síður er eðlilegt að taka mið af þessum reglum þegar matstæki eru þýdd og staðfærð hérlandis.

Reglurnar eru samtals 22 og er skipt í fjóra flokka eftir innihaldi þeirra: (a) samhengi, tvær reglur, (b) samning og þróun matstækja, tíu reglur, (c) fyrirlögn, sex reglur, og (d) tæknilegar upplýsingar, fjórar reglur.

Samhengi. Fyrstu tvær reglurnar miða að því að draga úr skekkjum við þýðingu

og staðfærslu matstækja sem tengjast alhæfingargildi hugsmíða og sérkennum prófa eins og kunnugleika próftaka í nýjum heimkynnum matstækisins af aðferð við matið. Skekkjur vegna hugsmíða koma til þegar skilgreining þeirra er skekkt af sérkennum menningar í heimalandi matstækisins. Þegar beita á þessum tveimur reglum við þýðingu og staðfærslu sálfræðilegra prófa og kvarða, sem ekki á að nota til samanburðar á milli landa, felst það fyrst og fremst í því að meta hvort að tiltekið matstæki geti falið í sér skekkjur af þessu tagi. Líta verður á skilgreiningu hugsmíða í matstækinu og meta sérkenni þeirra, aðferð við mat á þeim og efni einstakra atriða sem notuð eru til að meta þær. Niðurstaða þess mats hefur síðan áhrif á það hvort ákveðið er að þýða og staðfæra tiltekið próf eða ekki. Reglurnar tvær miðast við að matstæki séu samin í samanburðarrannsóknum og hægt sé að hafa áhrif á þessi atriði. Sú er ekki alltaf raunin ef útgefið matstæki í einu landi er til dæmis eingöngu þýtt. Þá hefur lang líklegast ekki verið hugað að þessu sérstaklega og í sumum tilvikum eru takmarkanir á því frá hendi erlendra útgefenda hversu mikið má hreyfa við innihaldi prófs. Aftur á móti eru möguleikar til að bregðast við skekkjum af framangreindu tagi með staðfærslu atriða, samningu nýrra atriða, breytingu á kvörðum og aðferð við að fá svör frá próftökum. Þegar allir þessir kostir blasa við í einhverju verkefni er sennilega ávallt skynsamlegra að íhuga hvort ekki sé betra að þýða annað matstæki sem metur sömu eða hliðstæðar hugsmíðar.

Samning og þróun matstækja. Reglurnar tíu um samningu og þróun matstækja lúta að samningu, þýðingu og staðfærslu margvíslegra atriða í prófum. Í raun endurspegla reglurnar hefðbundið samningsferli sálfræðilegra prófa en þó með sérstakri áherslu á þýðingu og

¹ Undantekningu á þessu er hægt að finna í Brislin (1970).

² Að baki reglum ITC eru ýmis alþjóðleg og evrópsk samtök sálfræðinga. Þar á meðal: European Association of Psychological Assessment, International Association of the Evaluation of Educational Achievement, International Association of Applied Psychology, International Association of Cross-Cultural Psychology, International Language Testing Association, og International Union of Psychological Science.

staðfærslu þeirra á milli landa. Þessar reglur nýtast allar með einum eða öðrum hætti í þýðingu og staðfærslu sálfræðilegra prófa og kvarða, sem eingöngu á að nota í einu landi, en ekki til samanburðar á milli landa. Raunar er efni flestra regnanna klassískt dæmi um atriði sem huga þarf sérstaklega að við staðfærslu prófa. Þar á meðal mætti til dæmis nefna mikilvægi þess að: (a) þeir sem þýða og staðfæra próf séu tvítyngdir og hafi þekkingu á menningu heimalands matstækis og nýrra heimkynna þess, (b) fyrirmæli og leiðbeiningar við fyrirlögn matstækisins séu viðeigandi, (c) spurningagerðir og kvarðar séu ekki framandi fyrir þá sem eiga að taka tiltekið próf eða kvarða, (d) tiltekin áreiti eða myndefni í prófi séu viðeigandi í nýjum heimkynnum matstækis, (e) beinar þýðingar leiði ekki til breytinga á merkingu atriða í nýjum heimkynnum, og (f) mikilvægi aðferðar við þýðingu matstækisins. Jafnframt eru reglur um hvaða aðferðum á að beita til að skoða jafngildi matstækja á milli landa. Þegar matstæki á eingöngu að nota í einu landi, en ekki til samanburðar á milli landa, er ekki nauðsynlegt í öllum tilvikum að staðfesta jafngildi matstækja með jafn rækilegum hætti og reglurnar mæla fyrir um. Mestu skiptir að staðfesta próffræðieiginleika matstækisins í nýjum heimkynnum þess.

Fyrirlögn. Reglurnar sex um fyrirlögn matstækis lúta að sérkennum í aðferð við fyrirlögn og nákvæmum fyrirlagnarreglum í handbók með matstækinu. Hefðbundin leið til að skoða þetta er forprófun með því að leggja matstækið fyrir lítinn hóp próftaka. Sú aðferð gagnast vel til að lagfæra óskýr fyrirmæli eða annað í fyrirlögn matstækisins og þeim aðstæðum sem fyrirlögn fer fram í. Allar þessar reglur nýtast vel þegar eingöngu á að nota tiltekið matstæki í einu landi.

Tæknilegar upplýsingar. Reglurnar fjórar um tæknilegar upplýsingar matstækis lúta að skráningu allra upplýsingasemerunauðsynlegrar til að túlka niðurstöðu tiltekins matstækis. Þar á meðal eru til dæmis upplýsingar um aðferð við þýðingu, umfang og eðli staðfærslu efnis í matstækinu. Birting tæknilegra upplýsinga

(áreiðanleiki, réttmæti og túlkun mælitalna) í handbók um matstæki sem ekki eru stöðluð gegnir lykilhlutverki við túlkun þeirra og skilgreinir jafnframt hver umgengni við þessi matstæki á að vera eða hvort þau eigi að nota í tilteknum aðstæðum. Þessi atriði í reglunum fjórum eiga því vel við þegar eingöngu á að nota þýtt og staðfært próf í einu landi en ekki í samanburði á milli landa.

Leiðbeinandi reglur Alþjóðlegu prófanefndarinnar um þýðingu og staðfærslu prófa bera þess merki að þær eru hugsaðar almennt í tengslum við samanburðarrannsóknir á milli þjóða og menningarsvæða. Reglurnar beinast því að margvíslegum skilyrðum sem þarf að uppfylla til að auka líkur á því að matstækin séu jafngild á milli landa og réttlæti samanburð á milli þeirra. Dæmi um þetta eru fjölþjóðlegar samanburðarrannsóknir á kunnáttu barna og ungmenna í tilteknum námsgreinum. Þegar markmiðið er samanburður á kunnáttu milli landa er lykilatriði að þau atriði sem notuð eru í prófunum séu jafngild eða nálgest það. Þar skiptir máli að orðalag, hugtök, setningaskipan og svo framvegis séu nálægt því að vera hliðstæð. Einnig skiptir máli að jafngildi atriða í prófunum sé athugað á milli landa með tilliti til þess hvort sömu líkur eru á því hvort einstaklingar, með sömu hæfni eða færni í mismunandi löndum, svari tilteknum atriðum rétt (*differential item functioning*). Með öðrum orðum að líkur á réttu svari séu jafnar í hliðstæðum getuhópum í mismunandi löndum burt séð frá ólíkum menningarlegum sérkennum þeirra eða kynferði. Reglur Alþjóðlegu prófanefndarinnar miðast við þessar aðstæður og svo samanburðarrannsóknir í sálfræði á milli menningarsvæða (*cross-cultural psychology*) eða jafngildi mismunandi útgáfa sama matstækis í sama landi. Mikið af rannsóknum í Bandaríkjunum á þessu sviði miðast einmitt við athugun á eiginleikum sálfræðilegra prófa og kvarða í alls kyns minnihlutahópum þar (sjá t.d. í Neisser o.fl., 1996; Reynolds og Kaiser, 1990).

Aftur á móti er ekki fjallað sérstaklega um þýðingu og staðfærslu matstækja milli

landa í þessum reglum þegar eingöngu á að nota matstækið í einu landi en ekki í beinum samanburði á milli þeirra. En þetta er einmitt markmiðið í flestum tilvikum þegar sálfræðileg próf og kvarðar eru þýddir og staðfærðir hérlendis. Það er því full ástæða til að huga að einstökum verkþáttum í slíkum verkefnum og útlista nánar í hverju þeir felast.

Þegar sálfræðileg próf og kvarðar eru þýdd og staðfærð og eingöngu á að nota þau í einu landi en ekki til samanburðar á milli landa þarf að lágmarki að huga að öllum eftirtöldum þáttum: (1) val á matstæki til þýðingar og staðfærslu, (2) hæfni þýðenda, (3) fagleg hæfni og sérþekking á efni matstækis, (4) aðferð við þýðingu, (5) aðferð við staðfærslu, (6) sérstöðu þýddra og staðfærðra matstækja, (7) færni og sérþekkingu á aðferðum til að athuga skekkjur í þýddum og staðfærðum matstækjum, (8) færni og þekkingu til að framkvæma forprófanir, og (9) færni og þekkingu til að gera réttmætisathuganir.

Hér á eftir eru færð rök fyrir því að einungis þegar þýðing og staðfærsla á tilteknu matstæki hefur verið unnin að lágmarki í samræmi við framangreinda níu verkþætti og gögn um þá hafa verið birt á viðeigandi vettvangi er réttlæt看legt að tala um að þýdd og staðfærð útgáfa matstækisins sé til. Af því leiðir jafnframt að einungis er faglega réttlæt看legt að nota óstaðlaða útgáfu prófs í klínísku starfi þegar þýtt og staðfært matstæki hefur verið unnið í samræmi við alla framangreinda verkþætti og gögn hafa verið birt um þá. Hér er þá miðað við að það sé metið svo, á grundvelli fyrirbyggjandi gagna, að það sé betra að nota óstaðlað, þýtt og staðfært matstæki, fremur en að nota það ekki. Raunar er eðlilegast að það sé verkefni nefndar á vegum fagfélags að meta hvaða óstöðluð, þýdd og staðfærð, matstæki uppfylli lágmarksskilyrði til notkunar í klínísku starfi. Lokatakmark allra þýðinga og staðfærslna prófa og kvarða hlýtur þó að vera stöðlun þeirra.

1. Val á matstæki

Fyrsta skrefið við þýðingu og staðfærslu (og

stöðlun) sálfræðilegs prófs eða kvarða er athugun á próffræðilegum eiginleikum þess í heimalandi. Hægt er að ganga út frá því sem vísu að eiginleikar matstækis í heimalandi muni við bestu aðstæður vera svipaðir í nýjum heimkynnum þess. Líklegra er að þeir verði eitthvað lakari (Einar Guðmundsson og Hólmfríður Ólafsdóttir, 2003). Þegar upp kemur vandi í tengslum við alhæfingargildi hugsmíða eða annars sem rekja má til menningarþátta er líklegast að það leiði til þess að þýtt matstæki hafi lakari próffræðilega eiginleika í nýjum heimkynnum.

Brýnast er að gaumgæfa eftirtalda eiginleika matstækis í heimalandi áður en ákvörðun er tekin um hvort á að þýða það eða ekki: (a) áreiðanleika undirprófa og heildartalna, (b) hugsmíða-, viðmiðs-, og aðgreiniréttmæti, (c) gólf- og rjáfuráhrif í normum, (d) breidd aldursbilla í normum, (e) aldur norma, (f) gæði stöðlunarúrtaks, (g) sérkenni þýðis, og (h) eiginleika mælitalna. Öll þessi atriði hafa veruleg áhrif á notagildi matstækis í heimalandi en einnig í nýjum heimkynnum.

Almennt má ganga út frá því að klínískt notagildi prófs fari eftir próffræðieiginleikum þess. Mikilvægi réttmætisupplýsinga og samspil áreiðanleika og réttmætis eru vel þekkt. Aftur á móti eru ef til vill færri sem átta sig á að því lægri sem áreiðanleiki undirprófa matstækis er þeim mun líklegra er að undirprófin hafi ekki sérstakt skýringargildi umfram þann þátt sem þau tilheyra. Það dregur auðvitað verulega úr notagildi matstækisins og hefur áhrif á þversniðsgreiningu undirprófa. Áherslan á eiginleika norma í upptalningunni hér að framan er augljóslega mikilvæg þegar erlend norm eru notuð til að túlka niðurstöðu úr þýddum matstækjum. Reikna má með að öll sérkenni norma yfirfærast í ný heimkynni þegar þau eru notuð til að túlka niðurstöðu þýddra og óstaðlaðra matstækja. Að auki bætast við skekkjur samfara notkun erlendra norma í framandi heimkynnum t.d. að mælitölur undirprófa og heildartölur eru kerfisbundið of háar eða lágar. Áður en ákvörðun er tekin um að þýða tiltekið matstæki þarf að liggja fyrir

fullvissa um að framangreind atriði séu í lagi og sömuleiðis að ekki sé völ á betri hliðstæðum matstækjum.

Raunvís gögn um notagildi matstækis í greiningu og meðferð skipta máli við val á matstæki til þýðingar. Aftur á móti ætti útbreiðsla matstækis eða vinsældir þess ekki að ráða úrslitum um hvort það er þýtt. Þetta er augljóst ef veikleikar eru til staðar í einhverjum próffræðieiginleikum þess. Eigi að síður virðast vinsældir eða útbreiðsla matstækja ráða miklu um að þau eru þýdd fremur en hliðstæð matstæki með betri eiginleika. Um það vitna þýdd og óstöðluð matstæki sem eru í notkun hérlandis.³ Raunar er líklegast að athyglin hafi beinst að matstækinu einmitt vegna útbreiðslu þess eða vinsældum. Í öðrum tilvikum hefur viðkomandi kynnst matstækinu við vinnu eða framhaldsnám erlendis og þykir eftirsóknarvert að nota það hérlandis einnig. Í sumum tilvikum standast þessi matstæki ítarlegt mat á próffræðieiginleikum í samanburði við hliðstæð matstæki. En oft gera þau það ekki. Það er því ástæða til að minna á mikilvægi skipulegrar leitar að sálfræðilegum prófum og kvörðum í handbókum og gagnaböndum. Með því móti einu fæst bitastætt yfirlit matstækja á tilteknu sviði. Þegar haft er í huga hversu mikil vinna biður þeirra sem þýða og staðfæra matstæki skilar slíkt yfirlit miklu og hefur áhrif á hvaða matstæki er valið á endanum og hvers vegna það er valið.

2. Hæfni þýðenda

Þegar matstæki eru þýdd er æskilegast að þýðendur uppfylli þrjú meginskilyrði: (a) tvítyngi, (b) sérþekkingu á því efni sem á að þýða, og (c) þekkingu á menningarsvæði heimalands matstækis og nýrra heimkynna þess (Geisinger, 1994). Með öðrum orðum er æskilegast að öll skilyrðin rúmist í einum þýðenda. Ef þýða ætti bandarískt greindarpróf yfir á íslensku væri æskilegast að fá til þess tvítyngdan sálfræðing með sérhæfingu á sviði greindar- og hæfnisprófa sem hefði

búið bæði í Bandaríkjunum og á Íslandi. Næstbesti kostur væri íslenskur sálfræðingur með góða enskukunnáttu sem hefði búið í Bandaríkjunum án sérhæfingar á sviði hæfnis- og greindarprófa. Og þannig mætti halda áfram lengra og lengra frá því sem æskilegast er. Almenn má gera ráð fyrir því að hættur á alvarlegum þýðingarvillum aukist eftir því sem lengra er farið frá því sem æskilegt er. Gott dæmi um þetta er að finna í Jeanrie og Bertrand (1999).

Í raunverulegum aðstæðum getur verið erfitt að sameina alla framangreinda þætti í einum og sama þýðandanum. Þegar sú er raunin verður að laga þýðingarferli og staðfærslu að þeim veruleika. Það felst meðal annars í því að fjölga þeim einstaklingum sem koma að verkinu. Endanleg útgáfa þýðingar gæti t.d. verið unnin af tveimur hópum, hópi þýðenda og hópi sérfræðinga á tilteknu sviði.

Mikilvægt er að meta með formlegum eða óformlegum hætti hæfni þýðenda áður en þeir eru fengnir til verksins. Slíkar upplýsingar eru einnig gagnlegar þegar gerð er almenn eða fræðileg grein fyrir aðferð við þýðingu og hverjir koma að henni. Formleg próf í tungumálum (t.d. háskólagráða í tungumáli), löggilding til skjalapýðinga, starfsreynsla við þýðingar, útgefnar þýðingar og annað þess háttar kemur að gagni. Eins koma að gagni upplýsingar um tvítyngt uppeldi, lengd dvalar í erlendu landi, námskeið eða menntun í menningu tiltekins lands, formleg námskeið í viðkomandi tungumáli og svo framvegis.

3. Fagleg hæfni og sérþekking

Þegar sálfræðilegt matstæki er þýtt og staðfært á milli landa er sérþekking á innihaldi og sérsviði þess grundvallaratriði. Einnig þarf að vera til staðar færni í því að túlka og umgangast margvísleg sérkenni matstækisins í nýjum heimkynnum. Endanlegt markmið þýðingar og staðfærslu prófs ætti að vera stöðlun þess. Þegar að því kemur þarf að byggja á sérþekkingu og reynslu við þá vinnu.

³ Þessi staðhæfing er byggð á óbirtum gögnum sem nemendur í framhaldsnámi við sálfræðiskor Háskóla Íslands söfnuðu um sálfræðileg próf og kvarða í notkun á vinnustöðum sálfræðinga hérlandis.

Með því móti einu er hægt að búast við gagnlegri leiðsögn í umgengni við matstækið til væntanlegra notenda. Í mörgum tilvikum er ekki hægt að búast við því að einn einstaklingur búi yfir færni á öllum þessum sviðum. Samvinna sérfræðinga á framangreindum sviðum, í einhverri mynd, ætti því að vera regla fremur en undantekning. Hér gildir sennilega sú gullvæga regla fagmennskunnar að þekkja takmörk eigin sérþekkingar og vitneskja um hvar sérþekking annarra tekur við.

Í flestum þýðingum matstækja þarf að staðfæra atriði, fella einhver brott og semja ný. Sérþekking á tilteknum hugsmíðum ræður miklu um það hvernig til tekst í þessum aðstæðum. Það sama á við þegar kemur að því að gera fyrirvara eða leiðbeinandi reglur við túlkun á matstækinu í nýjum heimkynnum. Þar þarf sérþekkingu í tölfræði og gagnavinnslu.

Þeir sem þýða og staðfæra tiltekið matstæki og hafa til þess sérstakt leyfi erlendra útgefenda, ætla venjulega að nota það sjálfir og heimila öðrum aðgang að því. Í þessum aðstæðum skiptir sérhæfð leiðsögn aðstandenda þýðinga miklu máli til að útgefin þýðing og staðfærsla geri meira gagn en ógagn.

Þegar framangreind skilyrði um faglega hæfni og sérþekkingu á efni prófs eru ekki virt getur það haft slæm áhrif á framboð þýðdra, staðfærðra og staðlaðra matstækja hérlandis. Það getur til að mynda leitt til þess að eftirsóknarverð matstæki eru ekki þýdd og staðfærð við bestu skilyrði vegna þess að einhverjir hafa að eigin frumkvæði tekið verkið að sér án þess að ákjósanleg þekking sé til staðar til að vinna það. Þegar formleg leyfi til þýðinga sálfræðilegra prófa eða kvarða liggja fyrir frá útgefendum byggjast slík leyfi yfirleitt á því að prófin séu notuð í rannsóknarskyni en ekki í klínísku starfi. Gerður er skýr greinarmunur á rannsókn á eiginleikum matstækis annars vegar og klínískri notkun þess. Í samræmi við bandarískar reglur um þetta efni (American Educational Research Association o.fl., 1999) er gert ráð fyrir að próf séu stöðluð í viðkomandi landi áður en þau eru notuð klínískt. Langur vegur er frá því að þeir

sem þýða matstæki af einhverjum ástæðum séu í öllum tilvikum færir um, faglega og fjárhagslega, að ljúka stöðlun á því. Sennilega á það við í fæstum tilvikum. Þess vegna verður að skapa þá faglegu hefð hérlandis að þegar sálfræðileg próf og kvarðar eru ekki stöðluð sé ekki óeðlilegt að fleiri en einn geti þýtt þau og staðfært sjálfstætt og athugað eiginleika þeirra.

4. Aðferðir við þýðingu

Aðferðir við þýðingu og staðfærslu prófa eru einkum tvær: (a) þýðing og bakþýðing, og (b) tvær sjálfstæðar þýðingar og athugun á nákvæmni þýðinga hjá þriðja aðila. Fyrri aðferðin hefur mest verið notuð í samanburðarrannsóknnum í sálfræði á milli menningarsvæða en einnig mikið við þýðingu og staðfærslu sálfræðilegra prófa og kvarða milli landa. Sú síðari hefur mikið verið notuð í umfangsmiklum samanburðarrannsóknnum á kunnáttu barna og ungmenna í tilteknum námsgreinum. Þessi aðferð var einnig notuð hérlandis við þýðingu og staðfærslu greindarprófs Wechslera handa leikskóla- og grunnskólabörnum (WPPSI-R¹⁵) og reyndist þar vel (Einar Guðmundsson og Hólmfríður Ólafsdóttir, 2003). Fleiri aðferðir eru þekktar en hafa lítið verið notaðar við þýðingu og staðfærslu sálfræðilegra prófa og kvarða (sjá nánar í Brislin, 1980; van de Vijver og Leung, 1997).

Þýðing og bakþýðing matstækja felst í því að tveir tvítyngdir þýðendur þýða tiltekið matstæki óháð hvor öðrum. Annar þýðir mælitækið yfir á tungumál nýrra heimkynna en hinn þýðir aftur yfir á upprunalegt tungumál (Brislin, 1970; Brislin, Lonner og Thorndike, 1973; Werner og Campbell, 1970). Þýðendur og bakþýðendur geta þó verið fleiri en tveir. Sömu leiðis getur ferli þýðinga og bakþýðinga verið endurtekið í nokkur skipti. Mælikvarði á ágæti tiltekinnar þýðingar er síðan hversu lík bakþýðingin er upprunalegri þýðingu. Þegar talið er nauðsynlegt að staðfæra einhver atriði matstækisins eru notuð til þess sérstök aðferð (*decentering*) (Brislin, Lonner, og

Thorndike, 1973; Werner og Campbell, 1970).

Ýmsir hafa gagnrýnt þýðingu og bakþýðingu matstækja út frá reynslurökum. Hambleton (1993) hefur til dæmis bent á að þegar þýðendur eru meðvitaðir um að bakþýða eigi upprunalega þýðingu hafi þeir tilhneigingu til að nota orðalag í þýðingum sem auðveldara er að bakþýða. Van de Vijver og Leung (1997) taka undir þetta en gagnrýna jafnframt áherslu þessarar aðferðar á málfræðilega og setningafræðilega réttar þýðingar. Það dregur úr áherslum á samhengi, merkingu og skilning á texta að þeirra mati. Geisinger (1994) nefnir þann ókost bakþýðinga að þær útiloki mikla staðfærslu atriða en slíkt sé oftast nauðsynlegt í meira eða minna mæli.

Mikilvægter að undirstrika að framangreindir fræðimenn vísa ekki til raunvísra gagna í gagnrýni sinni sem veikir hana nokkuð. Eigi að síður er mikilvægt að útiloka að framangreind gagnrýnisatriði séu til staðar þegar bakþýðingar eru notaðar. Það er því varasamt að nota þessa aðferð vélrænt og sömuleiðis að nota nákvæmni bakþýðingar einvörðungu til að staðfesta ágæti þýðingar og jafngildi matstækis á tveimur tungumálum.

Hin meginaðferðin við þýðingu og staðfærslu matstækja felst í því að tvær eða fleiri sjálfstæðar þýðingar eru gerðar á sama matstæki og þær bornar saman. Ein útgáfa er gerð eftir að ósamræmi í þýðingum hefur verið jafnað af þýðendum og þeim sem hafa sérþekkingu á sviði matstækisins. Óháður aðili (tvítýngdur með sérþekkingu á viðkomandi efni) fer síðan yfir endanlega þýðingu matstækisins, ber hana saman við upprunalega útgáfu og gerir athugasemdir. Endanleg útgáfa matstækisins er síðan gerð eftir breytingar í samræmi við athugasemdir.

Þegar tvær sjálfstæðar þýðingar eru gerðar af matstæki og ein útgáfa síðan búin til á grundvelli þeirra eru staðfærslur gerðar um leið. Hópur sérfræðinga á tilteknu sviði, helst með færni í tilteknu tungumáli og þekkingu á tilteknu menningarsvæði, fer yfir frumútgáfu matstækisins og ber hana saman við upphaflega útgáfu þess. Geisinger

(1994) leggur til að hver einstaklingur í slíkum hópi geri skriflegar athugasemdir í einrúmi og hópurinn hittist og fari yfir athugasemdir hvors annars og nái síðan samkomulagi um endanlegar breytingar á frumútgáfu. Reynsla er af þessari aðferð hérlandis í fjölþjóðlegum samanburðarrannsóknum og við þýðingu greindarprófs, og óhætt að mæla með henni. Aðferðin hefur þó þann augljósa galla að mælikvarði á ágæti þýðingar er fremur veikur (samanburður eins aðila á þýddri útgáfu og útgáfu í heimalandi matstækis).

En hvora aðferðina hér að framan á að nota við þýðingu sálfræðilegra prófa og kvarða? Þessu er ekki hægt að svara með afdráttarlausum hætti. Til þess skortir rannsóknir á þessum aðferðum, kostum þeirra og göllum, og í hvaða aðstæðum þær henta best. Meginatriðið er að sú aðferð sem valin er til þýðingar matstækis tryggi að þýdd útgáfa sé hliðstæð eða jafngild upprunalegri útgáfu með tilliti til merkingar tungumálsins og sálfræðilegra hugsniðna (van de Vijver, 2000).

Reynsla hérlandis af notkun tveggja hliðstæðra þýðinga í fjölþjóðlegum samanburðarrannsóknum og greindarprófi er góð. Því er ekki ólíklegt að þessi aðferð henti almennt vel þegar á að þýða kunnáttu-, hæfnis-, greindar- og þroskapróf. Hún gæti einnig hentað í öðrum tilvikum eins og við þýðingu á sálfræðilegum kvörðum af ýmsu tagi, persónuleikaprófum, áhugasviðskönnunum og svo framvegis. Góð leið til að ákvarða vinnuferli við þýðingu og staðfærslu prófa er að kynna sér hliðstæð verkefni erlendis frá (sjá t.d. Butcher, 1996; Georgas, Weiss, van de Vijver og Saklofske, 2003).

5. Aðferðir við staðfærslu

Í flestum tilvikum, ef ekki öllum, er líklegt að staðfæra þurfi sálfræðileg próf og kvarða í kjölfar þýðinga. Reynsla hérlandis og í öðrum löndum bendir til þess (sjá t.d. Cheung, 1989; Einar Guðmundsson og Hólmfríður Ólafsdóttir, 2003; Liu o.fl., 1994; Lucio, Reyes-Lagunes, og Scott, 1994; Spielberger, Gorsuch, og Lushene, 1970). Í kunnáttu-, hæfnis-, greindar-

og þroskaprófum er hægt að þýða atriði fyrir atriði og huga sérstaklega að staðfærslum í hverju atriði fyrir sig. Í öðrum tilvikum, t.d. í persónuleikaprófum eða atferlislistum, er farsælla að skoða öll atriði, sem meta tiltekna hugsmíð, í samhengi.

Í sumum tilvikum er óhjákvæmilegt að semja ný atriði í stað þess að reyna að staðfæra þau. Þegar eingöngu á að nota matstæki í einu landi en ekki í samanburðarrannsókn er þetta oft betri kostur en að þýða og staðfæra þau atriði sem einhver vandi er tengdur. Í sumum tilvikum getur umtalsverð samning nýrra atriða verið nauðsynleg. Dæmi um þetta er stöðlun á WISC-III í Japan. Þar var eingöngu tveimur atriðum í undirprófinu Orðskilningi haldið úr amerískri útgáfu prófsins. Önnur atriði voru samin sérstaklega fyrir japanska útgáfu prófsins. Á sama hátt var eingöngu 50% amerískra atriða haldið í undirprófinu Þekking, 40% í Ófullgerðum myndum og 25% atriða í Skilningi (Ueno og Nakatani, 2003). Það er því augljóst að nýta þarf sérþekkingu á innihaldi og hugsmíðum matstækis þegar semja þarf atriði í stað þess að staðfæra þau. Farsælast er að byggja hér einnig á sérþekkingu í samningu matstækja.

Þegar matstæki eru staðfærð er meginatriði að áreiðanleiki undirprófa og heildartalna, mismunandi tegundir réttmætis (hugsmíðaréttmæti, viðmiðsréttmæti, aðgreiniréttmæti), gólf og rjáfur í normum, séu viðunandi. Það er ekki markmið í sjálfu sér að öll atriði prófs séu til dæmis af svipaðri eða sömu þyngd á milli landa. Miklu raunsærra er að gera ráð fyrir mismunandi þyngd atriða (sjá t.d. Einar Guðmundsson og Hólmfríður Ólafsdóttir, 2003; Georgas, Weiss, van de Vijver og Saklofske, 2003). Aðalatriðið er að prófið haldi próffræðilegum eiginleikum við þýðingu og staðfærslu. Breyting á þyngdarröð atriða getur verið áhugaverð og vakið spurningar um hvers vegna munur kemur fram en er ekki sérstakt áhyggjuefni þegar stöðlun fer fram á matstækinu í kjölfarið. Rökin að baki þessu er sá grunnur sem sálfræðileg próf og kvarðar hvíla á. Tiltekið atriðasafn í prófi

er úrtak allra hugsanlegra atriða í prófum en ekki endanlegt. Viðfangsefni þess sem þýðir og staðfærir próf er því að ákveða hvenær er betra að skipta út atriði í stað þess að leggja mikla vinnu í að halda því. Val á nýju atriði verður síðan að endurspegla tiltekið efnissvið. Hugsmíðaréttmæting prófsins sker síðan úr hvernig til hefur tekist.

6. Sérstaða þýddra og staðfærðra matstækja

Flest það sem á við um þýðingu og staðfærslu matstækja í fjölþjóðlegum samanburðarrannsóknum, á sálfræðilegum þáttum eða kunnáttu, á einnig við um þýðingu og staðfærslu matstækja sem eingöngu á að nota í einu landi. Aftur á móti er ekki ávallt nauðsynlegt að gera jafn strangar kröfur um að matstæki séu jafngild (*equivalent*) á milli landa þegar eingöngu á að nota þýtt (og staðlað) matstæki í einu landi eins og gera þarf þegar samanburðarrannsóknir eru gerðar. Í flestum tilvikum eru gögn um það hvort tvö matstæki séu jafngild á milli landa áhugaverð og auka notagildi þeirra. En slíkar upplýsingar eru í mörgum tilvikum ekki nauðsynlegar þegar eingöngu á að nota þýtt og staðlað matstæki í einu landi.

Þegar sálfræðileg próf og kvarðar eru þýdd á milli landa er eftirsóknarvert að þau hafi sömu próffræðieiginleika í nýjum heimkynnum og í heimalandi. Af þeim ástæðum er æskilegt að athuga þetta sérstaklega með samanburðarrannsókn á milli landa. Tryggja þarf að matstækið sé jafngilt á milli landa og laga þarf þýðingu og staðfærslu að þessum veruleika. Bera þarf til dæmis saman eiginleika atriða í hliðstæðum hópum í tveimur löndum og athuga hvort þau hafa hliðstæða eða sömu eiginleika.

Aftur á móti á oft við að þetta er ekki meginmarkmiðið heldur hitt að matstækið hafi þekkta próffræðieiginleika í nýjum heimkynnum þannig að réttlætanlegt sé að nota það í klínísku starfi. Í slíkum tilvikum er ekki nauðsynlegt að beita flóknum tölfræðiaðferðum til að tryggja að matstækið

sé jafngilt erlendri útgáfu þess. Auðvitað er þess vænst að það hafi sömu eða svipaða próffræðilega eiginleika og í heimalandi. Og væntanlega er upphaflegur áhugi á því að þýða matstækið sprottinn vegna einhverra eiginleika þess eða notagildis. En það þarf alls ekki að rýra notagildi tiltekins matstækis að það hafi ekki nákvæmlega sömu eiginleika á milli landa. Meginatriðið er að áreiðanleiki, réttmæti og aðrir próffræðieiginleikar matstækisins séu þekktir í því landi þar sem á að nota það. Þegar það hefur verið gert þarf jafnframt að staðfesta að margvísleg túlkun á niðurstöðu matstækisins í heimalandi sé viðeigandi í nýjum heimkynnum. Ef próffræðieiginleikar matstækisins eru svipaðir á báðum stöðum er líklegt að það eigi við þó jafnframt megi búa við einhverjum breytingum. Þess vegna verður að athuga með raunvísunum gögnum hvort greiningarviðmið og aðrar reglur í túlkun tiltekins matstækis eru eins eða breyttar. Líklegast er að einhverjar breytingar þurfi að gera. Þetta undirstrikar að þýðing og staðfærsla sálfræðilegra prófa og kvarða hlítur ávallt að vera áfangi á leið til stöðlunar.

Í fáum tilvikum, ef nokkrum, er raunhæft að búa við því að hægt sé að nota norm og annað efni í tengslum við túlkun á niðurstöðu matstækis eftir þýðingu og staðfærslu þess. Þessu hefur þó stundum verið haldið fram, einkum í tengslum við einstök matstæki (sjá t.d. Butcher og Garcia, 1978). Geisinger (1994) hefur bent á hið augljósa í þessu efni að sýna verði fram á með raunvísunum hætti að hægt sé að nota bandarísk norm óbreytt þegar bandarísk sálfræðileg próf hafa verið þýdd og staðfærð. Réttmætisathuganir héraðs á bandarískum prófum styrkja þessa skoðun. Niðurstöður þessara rannsókna benda eindregið til þess að skekkja fylgi notkun bandarískra norma þegar bandarísk próf hafa verið þýdd og staðfærð héraðs (Einar Guðmundsson, Sigurður J. Grétarsson, Sveinborg Kristjánsdóttir og Valka Jónsdóttir, 1993 – 1994; Einar Guðmundsson, Ásdís Clausen, Berglind S. Ásgeirsdóttir og Birgir Þór Guðmundsson, 2004; Einar Guðmundsson og Birgir Þór Guðmundsson,

2004; Evald Sæmundsen, Margrét Arnljótsdóttir og Jónas Halldórsson, 1990).

Af því sem sagt hefur verið hér að framan má draga þá ályktun að þýðing og staðfærsla sálfræðilegra prófa og kvarða sem eingöngu á að nota í einu landi lúti sömu lögmálum og þýðing og staðfærsla matstækja í samanburðarrannsóknunum. Aftur á móti gegnir öðru máli þegar metið er hvort matstækin eru jafngild á milli landa. Þegar um samanburðarrannsóknir er að ræða er þetta grundvallaratriði en ekki nauðsynlega þegar nota á matstækið eingöngu í einu landi. Þar skiptir mestu að staðfesta ýmsa próffræðieiginleika matstækisins og staðla það.

7. Skekkjur

Þegar matstæki er þýtt og staðfært á milli landa geta margvíslegar skekkjur haft áhrif á eiginleika og notagildi þess í nýjum heimkynnum. Þess vegna er mikilvægt að gera sér grein fyrir því við upphaf vinnuferlisins hvaða þekktir skekkjuvaldar geta haft áhrif á eiginleika matstækja í þýddri og staðfærðri útgáfu. Premur meginskekkjum hefur verið lýst í þessu sambandi. Skekkjur sem tengjast hugsmíðum, atriðum í prófi og aðferð við fyrirlögn matstækis (van de Vijver og Hambleton, 1996; van de Vijver og Leung, 1997; van de Vijver og Poortinga, 1997). Í van de Vijver (2000), Allen og Walsh (2000) og van de Vijver og Tanzer (1997) er að finna margar gagnlegar aðferðir og ábendingar til að fást við skekkjur af þessu tagi þegar matstæki eru þýdd og staðfærð á milli landa.

Skekkjur vegna hugsmíða koma til þegar hugsmíðar eru ekki jafngildar á milli menningarsvæða. Skilgreining þeirra er skekkt af sérkennum menningar í heimalandi matstækisins. Við þýðingu og staðfærslu sálfræðilegra prófa og kvarða innan Vesturlanda er sennilega minni hætt á þessu en þegar meiri munur er á menningu landa. Aftur á móti er ekki hægt að útloka slíkt fyrr en það hefur verið athugað sérstaklega. Þegar hugsmíðar eru mældar með tiltölulega fáum atriðum er

til dæmis ávallt hætta á að skekkja komi til vegna vals á atriðum í heimalandi (van de Vijver og Leung, 1997). Minni hætta er á þessu þegar atriðin eru mörg og ástæða til að ætla að þau séu gott úrtak allra hugsanlegra atriða af tilteknum sviðum sem gætu metið hugsmíðina. Þegar ástæða er til að ætla að hugsmíðaskekkja komi fram í matstæki er það tilefni til efasemda um réttmæti þess að þýða það og staðfæra.

Skekkjur vegna þýddra atriða matstækis vísa almennt til þess að tiltekin atriði hafi aðra eiginleika í nýjum heimkynnum en í heimalandi. Þegar slíkar skekkjur eru til staðar geta þær haft áhrif á hugsmíðaréttmæti (van de Vijver og Leung, 1997). Staðfesting á slíkri skekkju fæst þegar eiginleikar atriða í hliðstæðum getuhópum í tveimur löndum eru ekki svipaðir (van de Vijver og Hambleton, 1996). Tölfræðiaðferðir til að meta slíka skekkju eru vel þekktar (Holland og Thayer, 1988; Holland og Wainer, 1993).

Skekkjur vegna atriða matstækis geta komið til vegna slakra eða ónákvæmra þýðinga á atriðum, flókins orðalags eða setningaskipanar, atriðin eru óviðeigandi en ekki staðfærð og svo framvegis. Með hliðsjón af þessu er augljós hætta á því að skekkjur af þessu tagi komi til við þýðingu og staðfærslu matstækja. Brislin (1986) hefur lagt til að nokkrum viðmiðum sé fylgt þegar atriði, sem á að þýða, eru samín. Þau lúta að lengd setninga, notkun mismunandi orðflokka í atriðum, setningabyggingu og fleira. Þó svo að þessi viðmið hafi verið hugsuð fyrir þá sem semja atriði, sem ætlunin er að þýða í kjölfarið, geta þau einnig gagnast þegar matstæki eru þýdd á milli landa.

Aðferð við fyrirlögn matstækis eða önnur sérkenni (t.d. í matsreglum) geta valdið skekkjum í niðurstöðutölum á þýddri útgáfu þess. Almennt þarf að gaumgæfa vel öll sérkenni matstækisins sem huganlega gætu valdið skekkju í nýjum heimkynnum. Dæmi um þetta eru kunnugleiki próftaka af leiðum til að svara (t.d. spurningagerðir og tegundir kvarða), kunnugleiki áreita af einhverju tagi (t.d. myndir eða annað efni sem er notað við fyrirlögn) og orðalag leiðbeininga við

fyrirlögn. Fleiri dæmi eru gefin í van de Vijver og Leung (1997). Van de Vijver og Hambleton (1996) benda á að hægt sé að meta skekkjur af þessu tagi með því að nota útfærslu af vel þekktri aðferð Campbell og Fiske (1959) um samleitna og sundurgreinandi fylgni hugsmíða sem metnar eru með mismunandi aðferðum (*multitrait-multimethod*). Í stað þess að nota margar hugsmíðar er ein hugsmíð notuð en nokkrar aðferðir til að meta hana (*monotrait-multimethod*). Þegar ástæða er til að ætla að skekkjur af þessu tagi geti haft áhrif á niðurstöðutölur einhverra undirprófa eða heildartölu ætti að athuga það sérstaklega. Slík gagnasöfnun myndi þá falla undir forprófun sem er næsti áfangi á leið til þýddrar og staðfærðar útgáfu matstækis.

8. Forprófanir

Líta verður á þýðingu og staðfærslu prófs sem hluta af raunvísu ferli. Því lýkur með fyrirlögnum prófsins. Samkvæmt þessu er ekki nægjanlegt að tilteknu vinnuferli sé fylgt við þýðingu atriða, reglna við fyrirlögn og matsreglna á svörum. Þegar þýdd og staðfærð útgáfa prófs liggur fyrir verður að athuga hvernig þýðing prófsins reynist í raunverulegum aðstæðum. Í engu tilviki ætti að nota þýðingu prófs í klínísku starfi þegar hún hefur ekki verið forprófuð og réttmætisathugun verið gerð á henni.

Að lágmarki þarf að gera tvær forprófanir á þýðingu matstækis. Sú fyrri til að fá upplýsingar um einstök atriði, orðalag og merkingu, matsreglur og reglur um fyrirlögn og annað sem skiptir máli við notkun matstækisins. Jafnframt er tilgangurinn að breyta framantöldum atriðum í þýddri útgáfu ef nauðsyn krefur. Fyrsta úrtak í forprófun þarf ekki að vera stórt en mikilvægast er að einstaklingarnir í því séu dæmigerðir fyrir próftaka á einhvern hátt.

Þegar breytingar hafa verið gerðar á þýðingu á grundvelli fyrstu forprófunar þarf að forprófa matstækið aftur í nýju og stærra úrtaki. Best er að þetta úrtak fari nálægt því að endurspeglar tiltekið þýði vel. Við aðra forprófun hæfnis- og

greindarprófa, svo dæmi sé tekið, er atriðum þyngdarraðað (með því að leggja öll atriði prófsins fyrir), og athugað hvort reglur um hvenær eigi að byrja og hætta fyrirLögn eru viðeigandi. Í flestum, ef ekki öllum tilvikum, er nauðsynlegt að forprófa þýdda og staðfærða útgáfu matstækis tvisvar eða oftár. Þegar það hefur verið gert er hægt að tala um að til sé þýdd og staðfærð útgáfa matstækis. Þó er einn mikilvægur verkþáttur eftir sem felst í réttmætisathugun matstækisins. Slík athugun er skilyrði þess að hægt sé að réttlæta óstaðlaða notkun þess í klínísku starfi.

9. Réttmætisathuganir

Í hefðbundnu samningferli sálfræðilegra prófa væru réttmætisathuganir gerðar eftir að stöðlun prófsins lýkur. Í ljósi þess að mun fleiri þýdd próf eru í notkun hérlandis en líklegt er að verði stöðluð á endanum er nauðsynlegt að líta svo á að þýdd og staðfærð útgáfa prófs sé ekki notuð í klínísku starfi án undangenginnar réttmætisathugunar á því. Í þessu felst að prófið er lagt fyrir hæfilega stórt úrtak sem endurspeglar vel sérkenni í þýði. Æskilegast er að stærð úrtaksins geri til dæmis þáttgreiningu undirprófa mögulega til að athuga hugsmíðaréttmæti.

Það fer eftir eðli matstækis hvaða þætti er nauðsynlegt að athuga að lágmarki. Í öllum tilvikum ætti þó að meta áreiðanleika, réttmæti og skekkjur í erlendum normum. Hægt er að meta hugsmíðaréttmæti með þáttgreiningu gagna eða fyrirögnar á tveimur eða fleiri matstækjum í einu. Ef ætlunin er að nota óstaðlaða útgáfu prófs eða kvarða er nauðsynlegt að meta skekkjur í tengslum við greiningarviðmið og önnur atriði er lúta að túlkun og ákvarðanatöku á grundvelli niðurstaðna matstækisins.

Stöðlun sálfræðilegra prófa og kvarða

Rækileg umfjöllun um stöðlun sálfræðilegra prófa fyrir börn og fullorðna er utan meginviðs þessarar greinar. Aftur á móti er viðeigandi að fara nokkrum orðum um þetta efni til að leggja

áherslu á að þýðing og staðfærsla erlendra matstækja á aldrei að vera takmark í sjálfu sér heldur áfangi á leið til stöðlunar þeirra.

Þegar sálfræðileg próf eða kvarðar handa börnum eða fullorðnum hafa verið þýdd og staðfærð í samræmi við reglurnar hér að framan er næsta skref að staðla þau. Þegar endanlegt markmið vinnu við þýðingu og staðfærslu prófs eða kvarða er stöðlun þess er hugað að stöðlun innihalds prófsins, reglum um fyrirögn og fyrirögn, á öllum stigum vinnunnar. Oft þarf þó að gera sérstakar athuganir á þessum atriðum og öðrum áður en ákveðið er hvernig haga á söfnun gagna sem norm eru byggð á. Mörg atriði bætast við framangreinda upptalningu þegar unnið er að stöðlun atferlislista sem nota á við skimun og greiningu geðraskana hjá börnum og fullorðnum. Þegar þessari vinnu er lokið er ákveðið hvernig safna á úrtaki einstaklinga sem endurspeglar sérkenni þýðis sem túlkun matstækisins miðast við.

Nokkur reynsla er komin á það hérlandis hvernig best er að haga úrtöku þegar norm eru útbúin fyrir matstæki handa börnum og unglingum. Best reynsla er af notkun lagskipts handahófsúrtaks. Þessi aðferð felst í því að landinu er skipt upp í sjö landshluta og viðeigandi hlutfall einstaklinga í hverjum landshluta valið af handhófi í úrtakið. Þessi aðferð var til dæmis notuð við stöðlun á greindarprófi David Wechsler handa börnum á leikskóla- og grunnskólaaldri (Einar Guðmundsson og Hólmfríður Ólafsdóttir, 2003), Íslenska þroskalistanum (Einar Guðmundsson og Sigurður J. Grétarsson, 1997) og Smábarnalistanum (Einar Guðmundsson og Sigurður J. Grétarsson, 2005). Aðferðin tryggir rétta dreifingu í úrtaki á þáttum sem geta haft áhrif á frammistöðu eða útkomu barna og unglinga á sálfræðilegum matstækjum. Í handbókum framangreindra matstækja kemur fram viðunandi eða góð samsetning stöðlunarúrtaka með tilliti til lykilbreyta á borð við kyn, aldur, menntun og störf. Með hliðsjón af góðri reynslu af framkvæmd við söfnun gagna og eiginleikum stöðlunarúrtaka af framangreindri gerð má hiklaust mæla með

þeim við stöðlun sálfræðilegra prófa og kvarða hérlandis.

Lítil reynsla er af mismunandi aðferðum við stöðlun matstækja fyrir fullorðna hérlandis. Þó er ljóst að lagskipt handahófsúrtök henta ekki í flestum tilvikum. Það stafar fyrst og fremst af því hversu brottfall er hátt í slíkum úrtökum í hópi fullorðinna. Sennilega er mun raunhæfara að byggja stöðlunarúrtök fullorðinna á lagskiptum kvótaúrtökum. Aðferðin byggist á því að skipta landinu upp í til dæmis sjö landshluta og safna síðan tilteknu hlutfalli einstaklinga með fyrirfram skilgreind sérkenni (t.d. með tilliti til menntunar, starfa, aldurs, þjóðernis o.s.frv.) í hverjum landshluta. Þegar þessari aðferð er beitt er raunhæft að nota stóra vinnustaði við söfnun gagna. Hjá því verður þó ekki komist að safna gögnum í öllum landshlutum.

Niðurlag

Hér að framan hefur verið fjallað um nokkur skilyrði þess að sálfræðilegt matstæki geti talist þýtt og staðfært. Með hliðsjón af hagsmunum skjólstaðinga sálfræðinga er mikilvægt að faglegt samkomulag náist um að öll sálfræðileg matstæki uppfylli tiltekin lágmarksskilyrði áður en þau eru notuð í klínísku starfi. Lagt er til að formlega sé staðið að útgáfu þýddra og staðfærðra matstækja ef þau eru notuð í klínísku starfi. Þar komi að lágmarki fram upplýsingar um: (i) eiginleika matstækisins í heimalandi og ástæður fyrir vali á matstækinu, (ii) hæfni þýðenda, (iii) sérfræðiþekkingu aðstandenda þýðingar matstækisins á efni þess, (iv) aðferð við þýðingu og staðfærslu, (v) samanburður á erlendra útgáfu matstækisins og nýrri útgáfu þess eftir þýðingu og staðfærslu, (vi) umfjöllun um sérstakan vanda við þýðingu og staðfærslu tiltekins matstækis og hugsanlega skekkjuvalda, (vii) forpröfun, hvernig hún var gerð og til hvers hún leiddi, og (viii) réttmætisathugun, hvernig hún var gerð og til hvers hún leiddi. Lagt er til að þessi viðmið gildi almennt um þýðingu og staðfærslu sálfræðilegra matstækja, óháð inntaki þeirra eða aldri próftaka. Þegar aðstandendur þýðingar og

staðfærslu sálfræðilegs matstækis, í skilningi viðmiðanna hér að framan, ætla að nota það í klínísku starfi, eingöngu sjálfir eða aðrir einnig, komi að auki fram rökstuddar tillögur að því hver umgengni við matstækið eigi að vera í klínísku starfi.

Notkun óstaðlaðra sálfræðilegra matstækja er útbreidd í klínísku starfi hérlandis. Fyrir þessu eru margar ástæður en algengast er að skorti á stöðluðum matstækjum sé borið við. Þó framboð staðlaðrar matstækja sé lítið hérlandis og því þurfi að breyta með markvissum hætti leysir það sálfræðinga ekki undan þeirri faglegu og siðferðislegu ábyrgð sem fylgir notkun óstaðlaðra matstækja í klínísku starfi. Eina raunhæfa leiðin til að fást við ófullnægjandi framboð á stöðluðum sálfræðilegum matstækjum er að umgangast óstöðluð matstæki af fagmennsku til dæmis í samræmi við þau lágmarksskilyrði sem hér hafa verið sett fram fyrir notkun óstaðlaðra matstækja. Eina varanlega lausnin á þessu ástandi er þó auðvitað fjölbreytt framboð á þýddum, staðfærðum og stöðluðum sálfræðilegum matstækjum.

Abstract

Translated but unstandardized psychological instruments are widely used by clinical psychologists in Iceland. For most of these instruments meagre information is available on the method of translation, extent of adaptations and the psychometric properties of the translated instruments. Lack of information on these issues has unforeseen consequences for decision making in a clinical context. It is argued that there is a need for some criteria or guiding rules outlining the minimum conditions to justify proper and ethical use of unstandardized psychological instruments in a clinical context. The main purpose of this paper is to outline some of the minimum criteria that could justify the use of these instruments in a clinical context.

Nine criteria are suggested as a minimum requirement that unstandardized psychological instrument have to meet before they can

be considered a legitimate translation and adaption of a foreign instrument. Further, it is argued that only when translated and adapted psychological instruments meet these nine criteria can it be justifiable to use these instruments in a clinical context. This only applies when it is considered, based on empirical evidence, a better option to use these unstandardized instruments rather than not. Finally, it is emphasized that the main aim of all translations and adaptations of psychological instruments should be a proper standardization in the population where the instruments are to be used.

Heimildaskrá

- Allen, J. og Walsh, J.A. (2000). Í R. H. Dana (Ritstjóri), *Personality and clinical psychology series* (bls. 63-85). Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.
- American Educational Research Association, American Psychological Association, and National Council on Measurement in Education (1999). *Standards for educational and psychological testing*. Washington, D.C.: American Educational Research Association.
- Bracken, B. A. og Barona, A. (1991). State of the art procedures for translating, validating, and using psychoeducational tests in cross-cultural assessment. *School Psychology International*, 12, 119-132.
- Brislin, R.W. (1970). Back translation for cross-cultural research. *Journal of Cross-cultural Psychology*, 1, 185-216.
- Brislin, R. W. (1980). Translation and content analysis of oral and written material. Í H.C. Triandis og J.W. Berry (Ritstjórar), *Handbook of cross-cultural research* (1. bindi, bls. 389-444). Boston: Allyn & Bacon.
- Brislin, R.W. (1986). The wording and translation of research instruments. Í W.J. Lonner og J.W. Berry (Ritstjórar), *Field methods in cross-cultural research* (bls. 137-164). Newbury Park, CA: Sage.
- Brislin, R.W., Lonner, W. og Thorndike, R. (1973). *Cross-cultural research methods*. New York: Wiley.
- Butcher, J.N. (1996). Translation and adaptation of the MMPI-2 for international use. Í J.N. Butcher (Ritstjóri), *International adaptations of the MMPI-2* (bls. 3-46). Minneapolis: University of Minnesota Press.
- Butcher, J. N. og Garcia, R. (1978). Cross-national application of psychological tests. *Personnel and Guidance*, 56, 472-475.
- Campbell, D.T. og Fiske, D.W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, 56, 81-105.
- Cheung, F.M. (1989). A review on the clinical applications of the Chinese MMPI. *Psychological Assessment*, 3, 230-237.
- Dana, R. H. (1993). *Multicultural assessment perspectives for professional psychology*. Needham Heights: Allyn & Bacon.
- Dana, R. H. (1996). Culturally competent assessment practice in the United States. *Journal of Personality Assessment*, 66, 472-487.

- Einar Guðmundsson og Birgir Þór Guðmundsson (2004). Algengi mismunar á munnlegri og verklegri greindartölu WPPSI-R^{IS}. *Sálfræðiritið – Tímarit Sálfræðingafélags Íslands*, 9, 101-109.
- Einar Guðmundsson og Hólmfríður Ólafsdóttir (2003). WPPSI-R^{IS}. *Greindarpróf David Wechsler handa börnum á leikskóla- og grunnskólaaldri. Endurskoðuð útgáfa*. Reykjavík: Námsmatsstofnun.
- Einar Guðmundsson og Sigurður J. Grétarsson (1997). *Íslenski þroskalistinn. Handbók*. Reykjavík: Rannsóknastofnun uppeldis- og menntamála.
- Einar Guðmundsson og Sigurður J. Grétarsson (2005). *Smábarnalistinn. Handbók*. Reykjavík: Námsmatsstofnun.
- Einar Guðmundsson, Ásdís Clausen, Berglind S. Ásgeirsdóttir og Birgir Þór Guðmundsson (2004). *Notagildi erlendra staðla við túlkun niðurstaðna úr WISC-III*. Veggspjald á málþingi Sálfræðingafélags Íslands, Háskóla Íslands og Geðsviðs Landspítala-Háskólasjúkrahús. Haldið í Odda, HÍ, 4. júní 2004.
- Einar Guðmundsson, Sigurður J. Grétarsson, Sveinborg Kristjánsdóttir og Valka Jónsdóttir (1993–1994). Greindarpróf Wechslera fyrir forskólabörn: Réttmætisathugun. *Sálfræðiritið: Tímarit Sálfræðingafélags Íslands*, 4–5, 29–32.
- Evald Sæmundsen, Margrét Arnljótsdóttir og Jónas Halldórsson (1990). Skyn- og hreyfiproski 90 íslenskra ungbarna mældur með þroskaprófi Bayleys. *Sálfræðiritið – Tímarit Sálfræðingafélags Íslands*, 1, 7-18.
- 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.
- Georgas, J., Weiss, L.G., van de Vijver, F.J.R. og Saklofske, D.H. (2003). (Ritstjórar). *Culture and children's intelligence. Cross-cultural analysis of the WISC-III*. Amsterdam: Academic Press.
- Gylfi Ásmundsson (1990). Íslensk viðmið fyrir MMPI-prófið. *Sálfræðiritið – Tímarit Sálfræðingafélags Íslands*, 1, 75-83.
- Hambleton, R.K. (1993). Translating achievement tests for use in cross-national studies. *European Journal of Psychological Assessment*, 9, 57-68.
- Hambleton, R.K. (1994). Guidelines for adapting educational and psychological tests: A progress report. *European Journal of Psychological Assessment (Bulletin of the International Test Commission)*, 10, 229-244.
- Hambleton, R.K., Yu, J., og Slater, S.C. (1999). Fieldtest of the ITC guidelines for adapting educational and psychological tests. *European Journal of Psychological Assessment (Bulletin of the International Test Commission)*, 15 (3), 270-276.

- Holland, P.W. og Thayer, D.T. (1988). Differential item performance and the Mantel-Haenszel procedure. Í H. Wainer og H.I. Braun (Ritstjórar), *Test validity* (bls. 129-145). Hillsdale, NJ: Erlbaum.
- Holland, P.W. og Wainer, H. (Ritstjórar). (1993). *Differential item functioning*. Hillsdale, NJ: Erlbaum.
- Jeanrie, C. og Bertrand, R. (1999). Translating tests with the International Test Commission's guidelines: Keeping validity in mind. *European Journal of Psychological Assessment (Bulletin of the International Test Commission)*, (3), 277-283.
- Liu, H.C., Chou, P., Lin, K.N., Wang, S.J., Fuh, J.L., Lin, H.C., Liu, C.Y., Wu, G.S., Larson, E.B., White, L.R., Graves, A.B., og Teng, E.L. (1994). Assessing cognitive abilities and dementia in a predominantly illiterate population of older individuals in Kinmen. *Psychological Medicine*, 24, 763-770.
- Lucio, E., Reyes-Lagunes, I., og Scott, R.L. (1994). MMPI-2 for Mexico: Translation and adaptation. *Journal of Personality Assessment*, 63, 105-116.
- Neisser, U., Boodoo, G., Bouchard, T.J., Jr., Boykin, A.W., Brody, N., Ceci, S.J., Halpern, D.F., Loehlin, J.C., Perloff, R., Sternber, R.J., og Urbina, S. (1996). Intelligence: Knowns and unknowns. *American Psychologist*, 51, 77-101.
- Nicholas, D.S., Padilla, J. og Gomez-Maqueo, E.L. (2000). Issues in the cross-cultural adaptation and use of the MMPI-2. Í R. H. Dana (Ritstjóri), *Personality and clinical psychology series* (bls. 247-266). Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.
- Reynolds, C. R. og Kaiser, M. (1990). Test bias in psychological assessment. Í T.B. Gutkin og C.R. Reynolds (Ritstjórar), *The handbook of school psychology* (2. útg., bls. 487-525). New York: Wiley.
- Saklofske, D. H. (2003). Canada. Í Georgas, J., Weiss, L.G., van de Vijver, F.J.R. og Saklofske, D.H. (Ritstjórar), *Culture and children's intelligence. Cross-cultural analysis of the WISC-III* (bls. 61-75). Amsterdam: Academic Press.
- Spielberger, C.D., Gorsuch, R.L., og Lushene, R.E. (1970). *State-Trait anxiety inventory*. Palo Alto, CA: Consulting Psychologists Press.
- Tanzer, N.K. og Sim, C.Q.E. (1999). Adapting instruments for use in multiple languages and cultures: A review of the ITC guidelines for test adaptations. *European Journal of Psychological Assessment (Bulletin of the International Test Commission)*, 15 (3), 258-269.
- Ueno, K. og Nakatani, I. (2003). Japan. Í Georgas, J., Weiss, L.G., van de Vijver, F.J.R. og Saklofske, D.H. (Ritstjórar), *Culture and children's intelligence. Cross-cultural analysis of the WISC-III* (bls. 215-225). Amsterdam: Academic Press.

- Van de Vijver, F. (2000). The nature of bias. Í R. H. Dana (Ritstjóri), *Personality and clinical psychology series* (bls. 87-112). Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.
- Van de Vijver, F.J.R. (2003). Principles of adaptation of intelligence tests to other cultures. Í Georgas, J., Weiss, L.G., van de Vijver, F.J.R. og Saklofske, D.H. (Ritstjórar), *Culture and children's intelligence. Cross-cultural analysis of the WISC-III* (bls. 255-276). Amsterdam: Academic Press.
- Van de Vijver, F. og Tanzer, N.K. (1997). Bias and equivalence in cross-cultural assessment: An overview. *European Review of Applied Psychology*, 47, 263-280.
- Van de Vijver, F. og Hambleton, R. K. (1996). Translating tests: Some practical guidelines. *European Psychologist*, 1 (2), 89-99.
- Van de Vijver, F. og Leung, K. (1997). *Methods and analysis for cross-cultural research*. London: Sage.
- Van de Vijver, F. og Poortinga, Y.H. (1997). Towards an intergrated analysis of bias in cross-cultural assessment. *European Journal of Psychological Assessment*, 13, 21-29.
- Werner, O. og Campbell, D.T. (1970). Translating, working through interpreters, and the problem of decentering. Í R. Narol and R. Cohen (Ritstjórar), *A handbook of cultural anthropology* (bls. 398-419). New York: American Museum of Natural History.