

Próffræðilegir eiginleikar íslenskrar þýðingar GAD-7 í klínísku úrtaki

Daney Harðardóttir
King's College

Vaka Vésteinsdóttir, Ragnhildur Lilja Ásgeirsdóttir
Háskóli Íslands

Hafrún Kristjánsdóttir
Háskólinn í Reykjavík

Fanney Þórsdóttir
Háskóli Íslands

General Anxiety Disorder Scale (GAD-7) er sjö atriða sjálfsmatskvarði sem notaður er til að skima fyrir almennri kvíðaröskun og meta alvarleika kvíðaeinkenna. Prófanir á erlendum útgáfum GAD-7 benda flestar til þess að mælingar með kvarðanum séu áreiðanlegar en þáttabygging hans virðist ekki stöðug í ólíkum úrtökum og sértækni hefur einnig reynst ófullnægjandi. Markmið rannsóknarinnar var að kanna próffræðilega eiginleika íslenskrar þýðingar GAD-7 í klínísku úrtaki. Leitandi þáttagreining var notuð til að meta þáttabyggingu íslenskrar þýðingar GAD-7, auk þess sem innri áreiðanleiki íslenskrar þýðingar kvarðans var metinn. Gögn frá 266 sjúklingum sem leituðu til heilsugæslustöðva (188 konum og 38 körlum) og voru kvíðnir og/eða þunglyndir samkvæmt klínísku mati voru notuð. Innri áreiðanleiki var góður $\omega_t = 0,90$ en meðalskor kvarðans var lágt og einnig hlutfall þeirra sem greindust kvíðnir. Samkvæmt leitandi þáttagreiningu skýrði einn þáttur best fylgni á milli atriða GAD-7 en þátturinn skýrði 57% af heildarbreytileika fylgnifylkisins. Þáttahleðslur allra atriða voru ásættanlegar eða á bilinu 0,53 til 0,87. Fyrstu niðurstöður á próffræðilegum eiginleikum íslenskrar þýðingar GAD-7 lofa góðu um gagnsemi listans sem mælitæki fyrir kvíðaeinkenni hér á landi en sýna einnig að þörf er á lagfæringum og frekari rannsóknum.

Efnisorð: GAD-7, próffræðilegir eiginleikar, íslensk útgáfa, kvíði, klínískt úrtak.

Spitzer og félagar (2006) hönnuðu General Anxiety Disorder Scale (GAD-7), sjö atriða sjálfsmatslista sem ætlað er að skima fyrir líklegum tilfellum af almennri kvíðaröskun og meta alvarleika kvíðaeinkenna með áherslu á notkun á heilsugæslum. Atriði listans endurspeгла helstu einkenni röskunarinnar samkvæmt greiningarviðmiðum DSM-IV auk þess sem höfundar sóttu innblástur í aðra kvíðalista. Upphaflega var listinn samsettur úr 13 atriðum sem síðan var fækkað niður í 7. Hægt er að fá heildarskor á bilinu 0 til 21 stig á listanum og benda hærri skor til aukins kvíða.

Lokaútgáfa listans var prófuð í úrtaki sjúklinga sem leituðu á heilsugæslu. Áreiðanleiki listans var góður en endurprófunaráreiðanleiki var 0,83

og innri áreiðanleiki $\alpha = 0,92$. Listinn hafði háa jákvæða fylgni við Beck Anxiety Inventory ($r = 0,72$) og kvíðahluta Symptom Checklist-90 ($r = 0,74$) sem sýndi fram á viðunandi réttmæti. Réttmæti listans var auk þess stutt með tengslum heildarskora við minni virkni og heilsu samkvæmt

Daney Harðardóttir er aðstoðarmaður í rannsóknum við Cicely Saunders Institute King's College í London, Vaka Vésteinsdóttir er lektor við Sálfræðideild Háskóla Íslands, Ragnhildur Lilja Ásgeirsdóttir er doktorsnemi í megindlegri sálfræði við Sálfræðideild Háskóla Íslands, Hafrún Kristjánsdóttir er dósent og deildarforseti Íþróttfræðideildar Háskólans í Reykjavík og Fanney Þórsdóttir er dósent við Sálfræðideild Háskóla Íslands. Fyrirspurnum vegna greinarinnar skal beint til Fanneyjar Þórsdóttur, dósentis við Sálfræðideild Háskóla Ísland, Nýi Garður, Sæmundargata 12, 102 Reykjavík. Netfang: fanneyt@hi.is.

SF-20 (20-item short form survey) (SF-20) sem metur almenna heilsu á sex sviðum (Spitzer o.fl., 2006; Stewart, Hays og Ware, 1988). Auk þess fannst jákvæð fylgni við fjölda veikindadaga ($r = 0,27$), læknisheimsóknna ($r = 0,22$) og hversu mikið kvíðaeinkenni trufla daglegt líf ($r = 0,63$) út frá sjálfsmati þátttakenda. Fylgni við þunglyndismælingar með PHQ-8 (Patient health questionnaire-8), útgáfu af PHQ-9 þar sem einu atriði er sleppt, reyndist há ($r = 0,75$). Þegar listarnir voru þáttgreindir saman hlóðu öll atriði PHQ-8 á einn þátt og öll atriði GAD-7 á annan þátt og þótti það, þrátt fyrir háa fylgni, sýna fram á aðgreiningu hugsmíðanna. Framkvæmd var leitandi þáttgreining á GAD-7 og fannst einn undirliggjandi kvíðaðáttur með þáttahleðslur atriða á bilinu 0,69 til 0,81 (Spitzer o.fl., 2006). Á grundvelli þessara niðurstæðna var dregin sú ályktun að GAD-7 væri réttmætt og gagnlegt tæki til að skima fyrir almennri kvíðaröskun og meta alvarleika hennar í rannsóknum og klínísku starfi.

Próffræðilegir eiginleikar GAD-7

Fjöldi rannsókna hefur kannað próffræðilega eiginleika GAD-7 í ólíkum úrtökum og ítrekað hefur verið sýnt fram á góðan innri áreiðanleika kvarðans. Innri áreiðanleiki mælist yfirleitt á bilinu 0,8 til rúmlega 0,9 í klínískum úrtökum sem bendir til þess að innra samræmi listans sé gott (t.d. Beard og Björgvinsson, 2014; Delgadillo o.fl., 2012; Sawaya, Atoui, Hamadeh, Zeinoun og Nahas, 2016). Rannsóknir hafa einnig sýnt fram á háa fylgni GAD-7 við aðrar mælingar á kvíða í klínískum úrtökum, til að mynda við kvíðahluta HADS (Hospital anxiety and depression scale; García-Campayo o.fl., 2010) sem og við HAM-A (Hamilton anxiety rating scale; Ruiz o.fl., 2011). Auk þess hefur fundist miðlungs jákvæð fylgni við mælingar á áhyggjum með Penn State Worry Questionnaire og miðlungs til há fylgni við kvíða- og streitumælingar með DASS (Depression anxiety stress scales; Rutter og Brown, 2017). Þessar niðurstöður styðja því við réttmæti GAD-7.

Aðrar niðurstöður vekja hins vegar upp spurningar um réttmæti GAD-7 því fylgni við mælingar á annarri hugsmíð, sér í lagi þunglyndi, hefur verið á bilinu miðlungs til há. Þetta á til að mynda við

um mælingar á þunglyndi með PHQ-9 (Sawaya o.fl., 2016), þunglyndishluta HADS (García-Campayo o.fl., 2010) og þunglyndishluta DASS (Rutter og Brown, 2017). Þessar niðurstöður benda til að listanum takist ekki nógu vel að aðgreina kvíða frá þunglyndi í klínískum úrtökum. Það hefur einnig verið bent á litla sértækni listans (Beard og Björgvinsson, 2014; Donker o.fl., 2011; Kertz o.fl., 2013; Rutter og Brown, 2017; Sawaya o.fl., 2016). Léleg sértækni GAD-7 ásamt hárrí fylgni milli mælinga listans og þunglyndismælinga bendir til þess að listinn mæli mögulega tvær hugsmíðar, það er að segja kvíða og þunglyndi.

Þáttabygging GAD-7 hefur verið skoðuð með leitandi hætti í klínískum úrtökum og stundum hefur fengist stöðug eins þáttar lausn (García-Campayo o.fl., 2010; Sawaya o.fl., 2016; Sousa o.fl., 2015) en rannsókn á enskri útgáfu listans í klínísku úrtaki benti til að tveir þættir henti betur en einn (Beard og Björgvinsson, 2014). Atriði 1, 2, 3 og 7: „Verið spennt/-ur á taugum, kvíðin/-n eða hengd/-ur upp á þráð“, „ekki tekist að bægja frá þér áhyggjum eða hafa stjórn á þeim“, „haft of miklar áhyggjur af ýmsum hlutum“ og „verið hrædd/-ur eins og eitthvað hræðilegt gæti gerst“, hlóðu á fyrri þáttinn sem snýr að hugrænni og tilfinningalegri upplifun á kvíða. Atriði 4, 5 og 6: „Átt erfitt með að slaka á“, „verið svo eirðarlaus að þú áttir erfitt með að sitja kyrr“ og „orðið gröm/gramur eða pirruð/pirraður af minnsta tilefni“ hlóðu á seinni þáttinn sem tengist líkamlegri upplifun og pírringi. Niðurstöður staðfestandi þáttgreiningar í klínískum úrtökum benda einnig til að eins þátta lausn passi illa (Kertz o.fl., 2013; Rutter og Brown, 2017) Þessar niðurstöður benda til að þörf sé á frekari athugunum á réttmæti GAD-7 og að ekki megi gera ráð fyrir að eins þáttar þáttabygging haldi í nýjum úrtökum.

Samantekt og markmið

Fyrri rannsóknir hafa sýnt fram á góðan innri áreiðanleika mælinga með GAD-7. Auk þess hefur verið sýnt fram á viðunandi réttmæti þegar fylgni við mælitæki sem mæla sömu hugsmíð er skoðuð en fylgni við mælingar á þunglyndi hefur einnig verið há og sundurgreiningarhæfni listans því ábótavant. Þar að auki hafa rannsakendur (t.d.

Beard og Björgvinsson, 2014) lýst yfir áhyggjum um virkni listans sem skimunartæki vegna háshlutfalls falskra jákvæðra og ófullnægjandi sér-tækni. Þáttaformgerð GAD-7 hefur einnig reynst nokkuð óstöðug. Í upphaflegri rannsókn Spitzer og félaga (2006) myndaði listinn einn kvíðabátt og hefur sú niðurstaða stundum verið endurtekin en í öðrum tilfellum hefur einn báttur passað illa. Ef til vill benda þessar niðurstöður til ólíkrar virkni listans milli ólíkra hópa eða til mögulegrar skekkju í mælitækinu. Niðurstöður þáttagreininga á GAD-7 sýna skýrt að rannsaka þarf þáttaformgerð listans nánar. Auk þess sýna niðurstöður fram á mikilvægi þess að kanna þáttaformgerð mælitækja í nýjum þýðingum, en vitneskja um þáttaformgerð þýddra sjálfsmatslista er grundvallarskilýrði fyrir notkun þeirra og auk þess nauðsynleg til að notkunin sé gagnleg (Einar Guðmundsson, 2011).

Markmið eftirfarandi rannsóknar er að kanna innri áreiðanleika og þáttabyggingu íslenskrar þýðingar GAD-7. Slíkt hefur ekki verið gert áður þrátt fyrir að notast sé við íslenska þýðingu listans til að mynda í klínísku starfi (Inga Hrefna Jónsdóttir o.fl., 2013) og rannsóknarstarfi (t.d. Erla Svansdóttir o.fl., 2018). Sökum óstöðugleika þáttabyggingar GAD-7 í öðrum útgáfum og þar sem íslenska útgáfan hefur ekki verið skoðuð áður verður framkvæmd leitandi þáttagreining. Notast verður við tölfræðiaðgerðir sem henta raðkvarða til að draga meg upp sem skýrasta mynd af undirliggjandi þáttaformgerð listans. Þannig verður stigið fyrsta skrefið í athugunum á prófkræðilegum eiginleikum íslenskrar útgáfu GAD-7 og munu niðurstöður nýtast þeim sem hyggja á að rannsaka mælitækið nánar sem og þeim sem vilja nýta sér það í skimun, klínísku starfi eða rannsóknum.

Aðferð

Þátttakendur

Þátttakendur voru hluti af stærra rannsóknarverkefni á vegum Landsspítalans og Háskóla Íslands. Þeim var vísað á rannsóknina ef heilbrigðisstarfsfólk á heilsugæslum taldi þá kvíðna og/eða þunglynda í kjölfar klíníks mats. Núverandi

vímuefnaneysla eða fíknivandi, geðvefræn einkenni eða merki um geðrofseinkenni útilokuðu þátttöku. Í upphafi rannsóknar fóru þátttakendur í greiningarviðtal hjá klínískum sálfræðingum þar sem þeir svöruðu fjölda sjálfsmatslista þar á meðal GAD-7 sem fjallað er um hér. Í heildina svöruðu 226 þátttakendur GAD-7 þar af 188 konur og 38 karlar. Þátttakendur voru fæddir á árunum 1933 til 1993 og voru þeir að minnsta kosti 18 ára við þátttöku.

Mælitæki

Notast var við íslenska þýðingu GAD-7. Sjálfsmatslistinn samanstendur af sjö atriðum sem ætlað er að skima fyrir almennum kvíða. Atriðum listans er stillt upp sem staðhæfingum sem lýsa kvíðatengdum einkennum og er þátttakandi beðinn um að meta hversu oft þessi einkenni hafa plagað hann síðastliðnar tvær vikur. Atriðin eru eftirfarandi: 1) „verið spennt/-ur á taugum, kvíðin/-n eða hengd/-ur upp á þráð“, 2) „ekki tekist að bægia frá þér áhyggjum eða hafa stjórn á þeim“, 3) „haft of miklar áhyggjur af ýmsum hlutum“, 4) „átt erfitt með að slaka á“, 5) „verið svo eirðarlaus að þú áttir erfitt með að sitja kyrr“, 6) „orðið gröm/gramur eða pirruð/pirraður af minnsta tilefni“ og 7) „verið hrædd/-ur eins og eitthvað hræðilegt gæti gerst“. Atriðum er svarað á fjögur punkta Likert-kvarða með svarmöguleikana „aldrei“, „nokkra daga“, „oftar en helming daganna“ og „næstum daglega“.

Öll atriði listans eru skorud jákvætt og hægt er að fá á bilinu 0 til 3 stig fyrir hvert atriði. Heildarskor geta því verið á bilinu 0 til 21 stig, þar sem hærri stigafjöldi bendir til aukins kvíða. Samkvæmt Spitzer og félögum (2006), höfundum listans, má miða við að 5 stig eða fleiri bendi til milds kvíða, 10 stig til meðalmikils og 15 stig eða fleiri til alvarlegra kvíðaeinkenna. Þau leggja einnig til að viðmið fyrir skimun sé heildarskor að minnsta kosti 10 stig. Plummer og félagar (2016) leggja aftur á móti til að viðmiðunarskor sé 8 stig.

Framkvæmd

Heimilislæknar á nokkrum heilsugæslustöðvum vísuðu þátttakendum 18 ára eða eldri í rannsóknina ef þeir töldu þá vera með kvíða og/eða þunglyndi. Öllum sem vísað var í rannsóknina var

gert að gangast undir ítarlegt mat hjá klínískum sálfræðingi og var þeim í kjölfar þess boðið að taka þátt í meðferðarhluta rannsóknarinnar eða vísað annað. Í þessu klíníska mati var meðal annars lagt fyrir MINI greiningarviðtalið sem og GAD-7 og fjöldi annarra sálfræðilegra mælitækja. Gögnum var safnað á árunum 2006 til 2011. Rannsóknin var framkvæmd með leyfi Vísindasiðanefndar (VSNb2005090003/03-15) og Persónuverndar (S2602/2005).

Tölfræðileg úrvinnsla

Öll tölfræðileg úrvinnsla var framkvæmd í R. Athugun á brottfalli leiddi í ljós sex brottfallsgildi sem mynduðu um 0,3% af heildarfjölda mæligilda og dreifðust á atriði 1, 2, 4 og 5. Þar sem brottfallsgildi voru minniháttar og að því er virtist tilfallandi var tilreiknað (*imputed*) eitt gildi fyrir hvert þeirra byggt á gildum annarra atriða GAD-7. Forritapakinn *mice* (van Buuren og Groothuis-Oudshoorn, 2011) var notaður til að tilreikna brottfallsgildi.

Reiknuð var lýsandi tölfræði fyrir bæði heildarskor þátttakenda á GAD-7 og fyrir einstaka atriði listans með forritapakkanum *psych* (Revelle, 2017). Einnig var skoðuð tíðni bakgrunnsbreyta á borð við kyn, aldur og menntun. Til að meta dreifingu atriða og heildarskora var notast við Jarque-Bera prófið (Jarque og Bera, 1987) sem metur hvort að dreifing breytu víki marktækt frá normaldreifingu byggt á skekkju og risi hennar. Núlltilgáta prófsins er sú að dreifing sé normal og að gildi skekkju sé núll og gildi riss þrír (Yap og Sim, 2011). Í kjölfarið voru reiknuð próf sem meta gildi riss og skekkju í sitthvori lagi. Normalpróf skekkjustuðulsins hefur þá núlltilgátu að skekkja sé núll og dreifing normal en aðaltilgátuna að dreifing víki frá normaldreifingu vegna skekkju. Normalpróf risstuðulsins hefur núlltilgátuna að ris jafngildi þremur og dreifing sé normal en aðaltilgátuna að dreifing víki frá normaldreifingu sökum riss (D'Agostino, Balanger og D'Agostino Jr., 1990). Miðað var við að $\alpha = 0,05$ fyrir marktækt þessara prófa. Notast var við forritapakkan *normtest* í R (Gavrilov og Pusev, 2014) til að framkvæma ofangreind próf. Auk þess var dreifing einstaka atriða og heildarskora skoðuð myndrænt.

Til þess að meta hvort gögnin hæfi til þáttagreiningar var fylgnifylki skoðað. Miðað var við að fylgnistuðlar skyldu vera 0,3 eða hærri til að fylgnifylkið væri þáttagreinanlegt (Tabachnick og Fidell, 2013). Auk þess var framkvæmt Bartlett próf en það prófar núlltilgátuna að fylgnifylkið sé einingarfylki (*identity matrix*) og þar með hvort að fylgnistuðlar víki marktækt frá núlli og er prófið sérstaklega gagnlegt þegar úrtök eru lítil (Tabachnick og Fidell, 2013). Miðað var við að Bartlett próf skyldi vera marktækt miðað við $\alpha = 0,05$ en marktækt þess gefur til kynna að gögn séu þáttagreinanleg (Williams, Onsman og Brown, 2010). Kaiser-Meyer-Olkin (KMO) stuðullinn var einnig skoðaður en hann metur sameiginlega dreifingu milli mældra breyta. Miðað var við að stuðullinn skyldi vera hærri en 0,5 til að þáttagreining væri viðeigandi, en því nær einum sem gildið er þeim mun betur henta gögnin til þáttagreiningar (Beavers o.fl., 2013). KMO stuðullinn og gildi Bartlett prófsins voru reiknuð með forritapakkanum *psych* (Revelle, 2017).

Við framkvæmd leitandi þáttagreiningar var einnig stuðst við forritapakkan *psych* (Revelle, 2017). Þáttagreining var framkvæmd á raðkvarðafylgnifylki (*polychoric correlation matrix*) sem er viðeigandi þegar mældar breytur eru fjölkosta atriði mæld á raðkvarða og undirliggjandi breytur eru samfelldar (Timmerman og Lorenzo-Seva, 2011). Ef notast er við Pearson fylgnifylki við þáttagreiningu slíkra gagna á sér stað vanmat á fylgnistuðlum auk þess sem þáttahleðslur verða skekktar (Baglin, 2014). Því ætti að fást skýrari þáttalaun þegar stuðst er við raðkvarðafylgni. Hermanir hafa einnig sýnt fram á að þegar mældar breytur eru á raðkvarða nálgast þáttagreining raunverulega þáttabyggingu betur ef notuð er raðkvarðafylgni í stað Pearson fylgni (Holgado-Tello, Chacón-Moscoso, Barbero-García og Vila-Abad, 2010). Þar sem atriðin sem unnið er með hér eru mæld á Likert-kvarða er við hæfi að þáttgreina raðkvarðafylgnifylki.

Við val á fjölda þátta er hægt að beita mörgum ólíkum aðferðum. Samkvæmt Henson og Roberts (2006) er algengasta aðferðin regla Kaiser (1960), hún segir að draga skuli alla þá þætti sem hafa eigin gildi hærri en einn. Þættir með eigin gildi jafnt og einn skýra jafn mikla dreifingu og stök

mæld breyta og telja bæði Kaiser (1960) og Guttman (1954) lítið gagn í að draga þætti sem skýra minni dreifingu en mældar breytur (Raïche, Walls, Magis, Riopel og Blais, 2013; Zwick og Velicer, 1986). Þrátt fyrir vinsældir og einfaldleika reglunnar hefur hún verið gagnrýnd til að mynda á þeirri forsendu að um sé að ræða tilbúna ákvörðunarreglu með skýrt og strangt viðmið sem ekki hefur mikið upplýsingagildi. Til dæmis væri þáttur með eigingildi 1,01 dreginn en ekki þáttur með eigingildi 0,99 þrátt fyrir að varla liggi neinn raunverulegur munur á milli skýringarmáttis þáttanna og að sá munur byggji líklega á úrtaksskekkju (Ruscio og Roche, 2012). Auk þess hefur fjöldi rannsókna sýnt að regla Kaiser eigi það til að ofmeta fjölda þátta (Fabrigar, Wegener, MacCallum og Strahan, 1999; Henson og Roberts, 2006; Zwick og Velicer, 1986).

Önnur algeng aðferð til að ákvarða fjölda þátta er skriðupróf Cattell (1966). Þá eru eigingildi teiknuð upp og síðan tengd saman svo úr verði ferill sem kallast skriðuplott. Síðan er fundinn sá punktur á skriðuplottinu þar sem sléttist úr ferlinum og fjöldi þátta fyrir þann punkt dreginn, en þessi punktur er talinn aðskilja mikilvæga þætti frá þeim ómerkilegu (Ledesma og Valero-Mora, 2007). Skriðuprófið felur því í sér hlutlæga ákvörðun um fjölda þátta en hefur þó reynst vel, sérstaklega þegar skýra undirliggjandi þætti er að finna í gögnunum (Fabrigar o.fl., 1999; Henson og Roberts, 2006). Þetta er þó ekki án vandkvæða en oft reynist erfitt að lesa úr skriðuplottinu og leiðir það til ólíkra ákvarðana meðal ólíkra rannsakenda (Zwick og Velicer, 1986). Til þess að bregðast við þessu og reyna að auðvelda og samræma túlkun á skriðuplotti hafa verið hannaðar ákvörðunarreglur sem ekki byggja á sjónrænni túlkun um hvernig best sé að staðsetja sléttuna. Til að mynda besta hnit (*optimal coordinate, OC*) sem felur í sér að fyrir hvert eigingildi sé reiknað væntigildi með aðhvarfsgreiningu. Síðan eru dregnir þeir þættir sem hafa raunverulegt eigingildi sem er hærra en eða jafnt og væntigildið (Raïche o.fl., 2013). Önnur slík ákvörðunarregla byggir á hröðunarþætti (*acceleration factor, AF*) sem leggur áherslu á staðinn á grafinu þar sem halli ferilsins breytist skyndilega og er fjöldi þátta ákvarðaður út frá

eigingildi á undan því sem sýnir mesta hröðun (Raïche o.fl., 2013).

Seinasta aðferðin til að ákvarða fjölda þátta sem fjallað verður um hér er samhliðagreining Horn (1965). Samhliðagreining byggir á því að búin eru til tilviljunarúrtök af sömu stærð og með jafn margar breytur og raunverulegt úrtak rannsakanda. Þessi úrtök eru tekin úr þýði þar sem engin undirliggjandi þáttaformgerð er til staðar og fylgnifylkið því einingarfylki. Reiknuð eru eigingildi fyrir þessi tilbúna gagnasöfn og síðan fundin meðaltöl þeirra fyrir hvern þátt. Meðaltölin eru síðan borin saman við eigingildi raunverulegu gagnanna og dregin sá fjöldi þátta þar sem raunveruleg eigingildi eru hærri en meðaleigingildi úr tilviljunarkenndum gögnum. Oft er þetta sett fram myndrænt (Ledesma og Valero-Mora, 2007). Helsti gallinn við samhliðagreiningu er að líkt og á við um reglu Kaiser er um handahófskennt viðmið að ræða. Samhliðagreining hefur þó þann kost fram yfir reglu Kaiser að tekið er tillit til úrtaksskekkju með því að bera saman raunveruleg gögn við tilviljunargögn dregin úr þýði með enga undirliggjandi þáttaformgerð (Ruscio og Roche, 2012).

Zwick og Velicer (1986) könnuðu ólíkar aðferðir til að ákvarða fjölda þátta. Regla Kaiser sýndi stöðugt ofmat á þáttum og sýndi skriðuprófið töluvert betri niðurstöðu þrátt fyrir að matsmenn hafi ekki alltaf verið sammála í túlkun sinni á því. Þá sýndi skriðuprófið, líkt og viðmið Kaiser, frekar tilhneigingu til ofmats en vanmats á þáttum. Samkvæmt niðurstöðu þeirra sýndi samhliðagreining mestan stöðugleika og fékkst oftast rétt niðurstaða þegar stuðst var við hana. Ruscio og Roche (2012) skoðuðu einnig gagnsemi ólíkra aðferða til að velja fjölda þátta í hermunarrannsókn þar sem skoðuð voru 10.000 gagnasöfn. Niðurstöður þeirra sýndu að með samhliðagreiningu fékkst rétt þáttalaun í um 76% tilfella, með besta hnit fékkst rétt niðurstaða í 74% tilvika og í 46% tilfella fékkst rétt niðurstaða með hröðunarþætti. Samkvæmt reglu Kaiser fannst rétt þáttabygging í aðeins tæplega 9% tilfella. Mæla þessar niðurstöður með notkun samhliðagreiningar og besta hnit en gegn reglu Kaiser. Það ber þó að hafa í huga að engin ofangreindra viðmiða um val á fjölda þátta nær fram rétttri þáttaformgerð í 100% tilvika og að þessar

aðferðir geta sýnt ólíkar niðurstöður þegar þær eru notaðar á sama gagnasafn (Henson og Roberts, 2006). Með þetta í huga var stuðst við öll þau viðmið sem fjallað var um hér að ofan við val á fjölda þátta. Þetta er regla Kaiser, skriðupróf, besta hnit, hröðunarpáttur og samhliðagreining. Notast var við forritapakka *psych* (Revelle, 2017) og *nFactors* (Raíche, 2011) í R til að beita þessum aðferðum og til að setja niðurstöðu fram myndrænt var notast við forritapakka *ggplot2* (Wickham, 2016).

Notuð var meginásapáttagreining (*principal axis factoring*) til að draga þáttalaun. Um er að ræða tegund af leitandi þáttagreiningu sem er mikið notuð og hentar vel þegar gögn víkja frá normaldreifingu og til þess að koma auga á veikburða þætti (de Winter og Dodou, 2012; Fabrigar o.fl., 1999). Notast var við leitandi þáttagreiningu fram yfir meginhlutagreiningu (*principal components analysis*) þar sem slíkt var viðeigandi fyrir markmið rannsóknarinnar. Markmið þáttagreiningar er að skýra fylgni milli mældra breyta með undirliggjandi hugsmíðum sem ekki er hægt að mæla með beinum hætti og gert er ráð fyrir að hafi áhrif á breyturnar (Fabrigar og Wegener, 2012). Þetta er gert með því að beina sjónum að sameiginlegri dreifingu mældra breyta og útiloka sértæka dreifingu þeirra þar sem slík dreifing orsakast af öðru en undirliggjandi þáttum. Í meginhlutagreiningu er öll dreifing skoðuð, það er bæði sameiginleg og sértæk dreifing (Costello og Osborne, 2005; Tabachnick og Fidell, 2013). Markmið slíkrar greiningar er að einfalda safn mældra breyta með því að sameina skor á þeim línulega í færri breytum sem kallaðar eru meginhlutar. Í meginhlutagreiningu er því ekki gert ráð fyrir undirliggjandi hugsmíðum sem hafa áhrif á breytur heldur er heildardreifingu breyta gerð skil í einfaldaðri mynd (Costello og Osborne, 2005; Fabrigar og Wegener, 2012).

Þrátt fyrir að leitandi þáttagreining og meginhlutagreining séu sitt hvor aðferðin með sitt hvort markmiðið eru þær oft notaðar líkt og um sömu aðferð væri að ræða og oft talað um meginhlutagreiningu sem afbrigði af þáttagreiningu (Fabrigar og Wegener, 2012). Þetta er varasöm venja þar sem meginhlutagreining er ekki viðeigandi ef markmiðið er að finna undirliggjandi þætti og

rannsóknir sýna að val á aðferð hefur áhrif á niðurstöðuna (t.d. Fabrigar o.fl., 1999; Widaman, 1993). Hér er viðeigandi að nota þáttalíkan þar sem gert er ráð fyrir undirliggjandi og ómælanlegri hugsmíð, það er kvíða, sem hefur áhrif á mældar breytur, það er svör þátttakenda á GAD-7.

Til að meta niðurstöðu þáttagreiningar var hér stuðst við eftirfarandi viðmið. Þáttahleðslur hærri en 0,71 voru metnar framúrskarandi, hærri en 0,63 mjög góðar, 0,55 góðar, 0,45 sæmilegar og 0,32 lélegar, en lægri þáttahleðslur voru ekki teknar til greina (Tabachnick og Fidell, 2013). Til að meta þáttaskýringar (*communalities*) var stuðst við viðmið Fabrigar og Wegener (2012). Samkvæmt þeim eru þáttaskýringar jafnt og eða hærri en 0,70 góðar, þáttaskýringar á milli 0,40 til 0,70 sæmilegar og allt undir 0,40 lélegt. Eftir að þáttalaun var dreginn voru leifar athugaðar sem reiknaðar eru með því að draga raunverulega fylgni atriða frá spáfylgni þáttalausnar. Miðað var við að færri en helmingur leifa skyldi vera hærri en 0,05 til að þáttalaun væri talin ásættanleg (Yong og Pearce, 2013).

Auk þáttagreiningar var innri áreiðanleiki mælinga metinn. Innri áreiðanleiki GAD-7 hefur yfirleitt verið metinn með Cronbach's alfa. Sá stuðull hvílir á þeirri forsendu að hallatölur sem lýsa sambandi atriða listans við undirliggjandi þátt séu þær sömu alls staðar (*essentially tau-equivalent*). Það telst hæpið að þessi forsenda standist í sálfræðilegum mælitækjum og því er gert ráð fyrir að alfa vanmeti réttan áreiðanleika mælinga (McNeish, 2018; Teo og Fan, 2013). Til þess að forðast vanmat var innri áreiðanleiki mælinga með GAD-7 hér metinn með omega total (ω_t). Omega stuðullinn gefur betra mat á áreiðanleika mælinganna þar sem hann hvílir ekki á ofangreindri forsendu, auk þess gerir stuðullinn ráð fyrir að atriði séu mæld á samfelldum kvarða og því hægt að reikna hann út frá raðkvarðafylgnifylki (McNeish, 2018). Lágmarksviðmið fyrir ásættanlegan áreiðanleika mældan með omega stuðli er 0,5 en þó er mælt með að stuðullinn sé 0,75 eða hærri (Reise, 2012; Reise, Bonifay og Haviland, 2013). Alfa stuðullinn var einnig reiknaður til þess að auðvelda samanburð við fyrri rannsóknir. Omega og alfa voru reiknaðir með forritapakkanum *psych* (Revelle, 2017).

Niðurstöður

Lýsandi tölfraði

Upplýsingar um úrtakið er að finna í töflu 1. Eins og sjá má í töflunni er meirihluti úrtaksins konur (83,2%) og flestir þátttakendur (77,4%) fæddir á árunum 1961 til 1990. Auk þess hafa flestir þátttakendur lokið starfsnámi (33,6%) og meirihluti úrtaks er í launaðri vinnu (57,1%).

Lýsandi tölfraði fyrir atriði GAD-7 í úrtakinu má sjá í töflu 2. Meðaltöl atriða voru öll á bilinu 1,1 og 1,6 að tveimur atriðum undanskildum, það eru atriði 5 og 7 og höfðu bæði þessi atriði meðaltöl lægri en einn. Lægst var meðaltal atriðis 7 ($M = 0,76$) og hæst meðaltal hafði atriði 4 ($M = 1,58$).

Öll atriði tóku gildi á bilinu 0 til 3 og nýttu því mögulega dreifingu. Við myndræna skoðun dreifingar sást að algengasta gildi allra atriða var ýmist 0 eða 1 stig. Dreifing atriða var könnuð með Jarque-Bera prófi og samkvæmt því vék dreifing atriðis 1 sem og atriða 4 til 7 marktækt frá normaldreifingu. Myndræn skoðun á dreifingu atriðis 4 leiddi í ljós að fleiri svörðu því með seinasta svarmöguleikanum „næstum daglega“ (sem gefur 3 stig) en á við um önnur atriði. Bendir þetta til þess að atriði 4 sé það atriði listans sem auðveldast er fyrir þátttakendur að samþykka. Ris þessa atriðis vék marktækt frá normaldreifingu og sama á við um atriði 1, 2 og 3. Ef gildi riss eru skoðuð í töflu 2 sést að dreifingin er alls staðar flatari en í normaldreifingu.

Atriði 5 til 7 sem og atriði 1 vöktu athygli við myndræna athugun þar sem þar er að sjá jákvæða skekkju og var skekkja þessara atriða marktæk

Tafla 1 Einkenni þátttakenda ($n = 226$).

Breytur	Fjöldi	Hlutfall (%)
Kyn		
Karlar	38	16,8
Konur	188	83,2
Fæðingarár		
1933-1940	2	0,9
1941-1950	11	4,9
1951-1960	29	12,8
1961-1970	44	19,5
1971-1980	55	24,3
1981-1990	76	33,6
1991-1993	8	3,5
Ótilgreint	1	0,4
Námi lokið		
Skuldunám	60	26,5
Starfsnám	76	33,6
Sérskólanám	14	6,2
Háskólanám	67	29,6
Annað nám	8	3,5
Ótilgreint	1	0,4
Launuð vinna		
Já	129	57,1
Nei	95	42,0
Ótilgreint	2	0,9

Tafla 2 Meðaltal (M), Staðalfrávik (SF), lægsta og hæsta gildi, ásamt skekkju og risi fyrir atriði GAD-7.

Atriði	M	SF	Lægsta	Hæsta	Skekkja	Ris
1	1,37	0,95	0	3	0,41	2,24
2	1,41	0,90	0	3	0,19	2,30
3	1,54	0,89	0	3	0,07	2,23
4	1,58	0,99	0	3	0,03	1,93
5	0,85	0,91	0	3	0,86	2,90
6	1,11	0,88	0	3	0,57	2,71
7	0,76	0,89	0	3	0,87	2,73

Tafla 3 Meðaltal (M), staðalfrávik (SF), hæsta og lægsta gildi, ásamt skekkju og risi eftir kyni fyrir heildarskor á GAD-7.

	M	SF	Lægsta	Hæsta	Skekkja	Ris
Konur	8,96	4,79	0	21	0,31	2,54
Karlar	6,95	4,48	0	17	0,59	2,45
Heild	8,62	4,79	0	21	0,36	2,51

samkvæmt normalprófi skekkjustuðuls. Sérstaklega vöktu atriði 5 og 7 athygli en þar merktu flestir við vægasta valmöguleikann sem bendir til gólfhrifa í þessum atriðum. Þrátt fyrir marktækt þessara prófa eru frávik í skekkju eða risi hvergi veruleg ef litið er til ris- og skekkjustuðla, því vakna ekki áhyggjur varðandi frekari úrvinnslu gagnanna að svo stöddu.

Heildarskor þátttakenda á GAD-7 voru reiknuð og má sjá lýsandi tölfræði þeirra eftir kyni í töflu 3. Heildarskor dreifast frá 0 upp í 21 stig og er heildardreifing listans því nýtt en flestir þátttakendur (38%) voru með heildarskor á bilinu 4 til 7 stig. Dreifing heildarskora er lítillega jákvætt skekkkt með meðaltal 8,62 stig (SF = 4,79). Skekkjan er marktæk samkvæmt Jarque-Bera prófi og samkvæmt prófi á skekkjustuðli. Ef að dreifingin er skoðuð fyrir kynin í sitthvoru lagi vikur hún þó ekki marktækt frá normaldreifingu en sjá má talsverðan mun á dreifingu heildarskora eftir kyni. Dreifing heildarskora karla sýnir meiri skekkju en kvenna samkvæmt skekkjustuðli og auk þess er möguleg dreifing heildarskora ekki fullnýtt þar sem dreifingin nær aðeins upp í 17 stig meðal karla. Þar að auki mælast karlar að meðaltali með minni kvíða en konur og er munurinn marktækur $t(55,52) = 2,49, p = 0,016$. Gefur þessi niðurstaða til kynna mögulega kynjaskekkju í mælitækinu eða raunverulegan mun á tíðni kvíðaeinkenna karla og kvenna.

Samkvæmt viðmiðum Spitzer og félaga (2006) mælast 87 einstaklingar (38%) með mild kvíða-einkenni, 63 (28%) með meðalmikinn kvíða og 28 (12%) með alvarleg einkenni kvíða. Ef miðað er við 10 stig eða fleiri skimast um 40% þátttakenda í áhættuhóp en um 51% ef miðað er við 8 stig.

Innri Áreiðanleiki

Innri áreiðanleiki mælinga með GAD-7 var góður. Áreiðanleiki mældist $\omega = 0,90$ með omega og $\alpha = 0,87$ með alfa. Munurinn á stuðlunum tveimur

er lítill en þó bendir munurinn til vanmats með alfastuðli. Áreiðanleiki var hér í góðu samræmi við aðrar mælingar með GAD-7 í ólíkum þýðingum (t.d. Beard og Björgvinsson, 2014; Donker o.fl., 2011; Hinz o.fl., 2017).

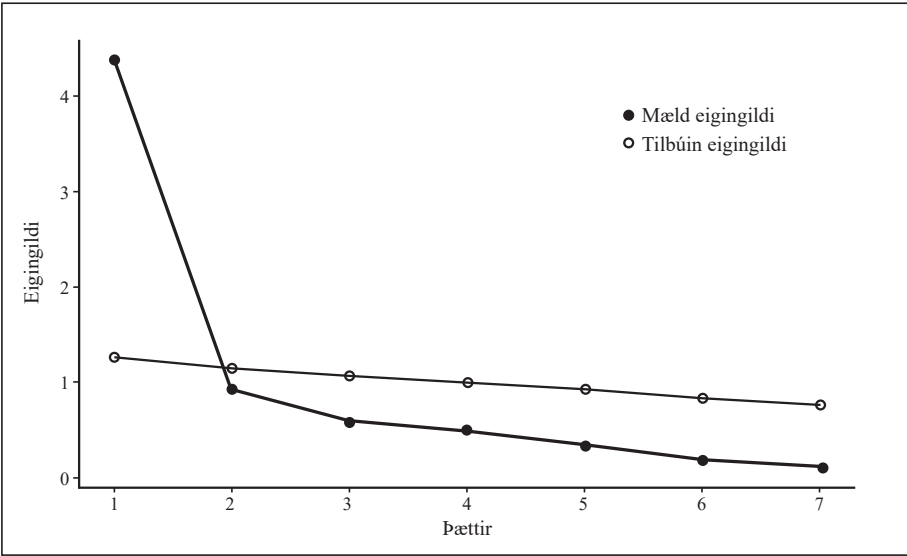
Leitandi þáttgreining

Þáttgreinanleiki fylgnifylkis

Raðkvarðafylgni milli atriða var í öllum tilfellum hærri en 0,3 sem bendir til að fylgnifylkið, sem sjá má í töflu 4, sé þáttgreinanlegt. Fylgnin var lægst á milli atriða 2 og 5 (0,34) og hæst á milli atriða 2 og 3 (0,86). Þetta er töluvert há fylgni og jafnframt eini fylgnistuðullinn sem nær yfir 0,8 en flestir fylgnistuðlar eru á bilinu 0,4 til 0,8. Almenn sýnir atriði 6 lægsta fylgni við önnur atriði en fylgnistuðlar þess mældust á bilinu 0,37 til 0,50. Bartlett próf var marktækt, $p < 0,001$, sem gefur til kynna að fylgnifylkið sé frábrugðið einingarfylki og þar af leiðandi að fylgni milli atriða viki marktækt frá núlli. KMO gildi var 0,81 sem bendir til þess að sameiginleg dreifing breyta sé viðunandi til þáttgreiningar. Allt bendir þetta til að gögnin hæfi þáttgreiningu.

Fjöldi þátta

Fyrstu tvö eigingildi gagnanna voru 4,39 og 0,92 og því fannst aðeins eitt eigingildi sem var herra en einn. Þegar mynd 1 er skoðuð sést að það sléttist úr ferli skriðuplottsins við annað eigingildið og því bendir skriðuprófið til eins þáttar. Sömuleiðis gefa báðar ákvörðunarreglurnar, það er besta hnit og hröðunarþáttur, til kynna að eins þáttar lausn falli best að gögnunum. Eins og sjá má á mynd 1 er eins þáttar lausn viðeigandi samkvæmt samhliðagreiningu þar sem það er eingöngu fyrsta eigingildið sem er herra en eigingildi tilbúinna gagna. Öll viðmið benda því til sömu niðurstöðu, það er að gögnunum sé best lýst með einum undirliggjandi kvíðapætti.



Mynd 1 Skriðupróf og samhliðagreining sem bendir til þess að eins þáttar lausn sé við hæfi.

Tafla 4 Raðkvarðafylgnifylki fyrir atriði GAD-7.

	1	2	3	4	5	6	7
1	1,00						
2	0,73	1,00					
3	0,68	0,86	1,00				
4	0,78	0,70	0,73	1,00			
5	0,45	0,34	0,40	0,63	1,00		
6	0,38	0,37	0,46	0,46	0,50	1,00	
7	0,64	0,58	0,62	0,52	0,43	0,40	1,00

Tafla 5 Þáttahleðslur og þáttaskýringar fyrir eins þáttar lausn GAD-7.

Atriði	Þáttahleðsla	h2
3. Miklar áhyggjur	0,87	0,76
4. Erfitt að slaka á	0,87	0,76
1. Spenna og kvíði	0,84	0,71
2. Erfitt að stjórna áhyggjum	0,83	0,69
7. Hræðsla	0,70	0,49
5. Eirðarleysi	0,57	0,33
6. Gremja og pirringur	0,53	0,28

Þáttalausn

Notast var við meginásalþáttagreiningu til þess að draga þáttalausn. Dreginn var einn þáttur sem öll atriði hlaða á og skýrir hann 57% af heildardreifingu atriða. Athugun á leifum leiddi í ljós að minna en helmingur þeirra eða 43% var hærri en 0,05 sem bendir til að þáttalausnin sé viðunandi. Þáttahleðslur ásamt þáttaskýringum má sjá í töflu 5.

Í töflu 5 sést að atriði 1 til 4 sýna framúrskarandi þáttahleðslu samkvæmt viðmiðum Tabachnick og Fidell (2013). Auk þess sýna þessi atriði góða þáttaskýringu samkvæmt viðmiðum Fabrigar og Wegener (2012) að atriði 2 undanskildu sem þó er á mörkum þess að sýna góða þáttaskýringu. Atriði 7 „verið hrædd/-ur eins og eitthvað hræðilegt gæti gerst“ sýnir lægri þáttahleðslu sem þó flokkast sem mjög góð en þáttaskýring atriðisins er sæmileg. Atriði 5 og 6 sýna þáttahleðslur sem flokkast sem góð annars vegar og sæmileg hins vegar samkvæmt viðmiðum. Aftur á móti má sjá að þetta eru áberandi verri þáttahleðslur en önnur atriði sýna og þáttaskýringar þessara atriða eru lélegar. Þetta eru atriðin „verið svo eirðarlaus að þú áttir erfitt með að sitja kyrr“ og „orðið gröm/gramur eða pirruð/pirraður af minnsta tilefni“, en það seinna sýnir örlítið lægri þáttahleðslu og -skýringu. Þáttalausn nær því ekki að gera dreifingu þessara atriða góð skil og þau tengjast öðrum atriðum listans illa. Niðurstöður benda því til þess að einhver önnur hugsmíð umfram kvíða hafi áhrif á svörun við seinustu þremur atriðum listans og/eða að þau sýni einhverskonar kerfisbundna mæli- eða svarskekkju. Atriði 5 og 6 sérstaklega geta að svo stöddu ekki talist gagnleg mæling á magn kvíða og vekja niðurstöður því upp spurningar um notagildi íslenskrar þýðingar GAD-7.

Umræða

Markmið rannsóknarinnar var að kanna í fyrsta skipti próffræðilega eiginleika íslenskrar þýðingar GAD-7 og þá sér í lagi innri áreiðanleika og þáttaformgerð listans. Listinn var skoðaður í klínísku úrtaki og voru niðurstöður að nokkru leyti í samræmi við fyrri rannsóknir. Mælingar sýndu að venju góðan innri áreiðanleika bæði mælt með omega og alfa, því er innra samræmi

íslenskrar þýðingar stutt. Auk þess fékkst skýr niðurstaða um eins þáttar formgerð listans sem styður við réttmæti íslensku þýðingarinnar.

Meðaltöl heildarskora í fyrri rannsóknum sem skoða GAD-7 í klínískum úrtökum eru yfirleitt hærri en 10 stig sem er gjarnan notað sem viðmiðunarskor í skimun (Beard og Björgvinsson, 2014; García-Campayo o.fl., 2010; Kertz o.fl., 2013; Sawaya, 2016; Sousa o.fl., 2015). Þetta er ekki í samræmi við þá niðurstöðu sem fékkst hér en meðaltal heildarskora var 8,62 og jafnframt var í öllum atriðum mest merkt við þá svarmöguleika sem gefa til kynna engin eða mild kvíðaeinkenni. Þar sem um er að ræða klínískt úrtak hefði verið eðlilegra að sjá dreifðari svörun og hærri meðaltal. Þar að auki var hlutfall þátttakenda sem skimast í áhættuhóp lægra en mætti vænta miðað við tegund úrtaks. Þegar miðað var við upphaflegt tíu stiga viðmiðunarskor Spitzer og félaga (2006) skimuðust aðeins um 40% þátttakenda og þegar miðað var við átta stig líkt og Plummer og félagar (2016) lögðu til skimaðist rétt rúmlega helmingur þátttakenda með kvíða. Þessar niðurstöður setja spurningarmerki við gagnsemi íslenskrar þýðingar GAD-7 og mæla gegn notkun listans í skimun. Þó ber að hafa í huga að val í úrtak gæti orsakað niðurstöðu af þessu tagi hafi viðmið heilbrigðisstarfsmanna sem vísuðu þátttakendum á rannsóknina verið of lág. Það er einnig mikilvægt að hafa í huga að þátttakendur leituðu til heilsugæslunnar og var þaðan vísað í meðferð fyrir vægan til miðlungsmikinn vanda. Þeir sem eiga í miklum vanda eru þar af leiðandi ekki líklegir til að vera hluti af úrtakinu. Þetta getur skýrt lágt meðaltal í þessari rannsókn.

Leitandi þáttagreining á íslenskri þýðingu GAD-7 leiddi skýrt í ljós einn undirliggjandi kvíðapátt sem hefur áhrif á svörun þátttakenda. Slík þáttalausn er bæði í samræmi við þáttaformgerð í upphaflegri rannsókn á listanum (Spitzer o.fl., 2006) sem og niðurstöður sem fengist hafi í öðrum þýðingum listans (Donker o.fl., 2011; García-Campayo o.fl., 2010; Hinz o.fl., 2017; Sawaya o.fl., 2016; Sousa o.fl., 2015). Hér var notast við raðkvarðafylgnifylki í þáttagreiningu sem kemur í veg fyrir vanmat á fylgnistuðlum sem vill verða þegar notast er við Pearson fylgni með gögnum á raðkvarða. Því ætti hér að koma fram

skýr þáttalausn þar sem þáttahleðslur atriða eru hærri en í þeim tilfellum þar sem Pearson fylgni er notuð, líkt og fékkst hjá Sawaya og félögum (2016) en þar var raðkvarðafylgnifylki einnig þáttgreint. Ef þáttalausnin er aftur á móti skoðuð sést að þetta á aðeins við um sum atriðin, en skýr niðurstaða sést fyrir fyrstu fjögur atriði listans, þau sýna mjög háar þáttahleðslur og þáttaskýringu og eru því vel studd í íslenskri þýðingu listans. Sömu sögu er ekki að segja um atriði 5, 6 og 7. Atriði 5 og 6 sérstaklega sem og atriði 7 sýna áberandi verri þáttahleðslur og þáttaskýringar en hin atriðin og því ljóst að hugsmíðinni kvíða tekst ekki að gera svörun við þeim jafn góð skil og öðrum atriðum listans. Mögulega er hér eitthvað annað en kvíði að hafa áhrif á svörunina. Þegar þessi atriði eru skoðuð nánar sést að þau hafa lægri meðaltöl en önnur atriði og sýna þau jafnframt mesta skekkju og jafnvel gólfhrif. Bendir þetta til þess að þessi atriði séu erfiðustu atriði listans – einkennin eru óalgengari og þurfi þátttakendur því að vera kvíðnari til að sammælask þeim en öðrum atriðum. Eðlilegt er að atriði listans séu miserfið þar sem slíkt aðgreinir þá sem sýna mildan kvíða frá þeim sem sýna alvarlegri kvíðaeinkenni en þó ætti í úrtaki sem þessu að sjást aukin nýting á þyngri svarmöguleikum allra atriða, líka þeirra sem eru erfiðari. Bendir þetta því allt til að skoða þurfi þessi atriði nánar.

Athyglisvert er að atriði 5 „verið svo eirðarlaus að þú áttir erfitt með að sitja kyrr“ sýni jafn laka niðurstöðu og raun ber vitni þegar atriði 4 „átt erfitt með að slaka á“ sem sýnir framúrskarandi niðurstöðu virðist snúa að sambærilegu ef ekki sama einkenni. Bæði atriðin beinast að erfiðleikum með slökun og ættu þátttakendur því röklega að svara þeim á sambærilegan hátt og atriðin þar af leiðandi að sýna áþekka niðurstöðu í þáttagreiningu. Það er þó ekki raunin og ef til vill er röðun atriðanna hér að hafa áhrif. Fyrst eru þátttakendur beðnir um að meta almennt hvort þeir hafi átt erfitt með að slaka á og virðast þeir ekki eiga í erfiðleikum með það. Strax í kjölfarið eru þeir spurðir hvort þeir eigi erfitt með að sitja kyrrir sökum eirðarleysis, en því reynist erfiðara að svara. Seinna atriðið er sértækara en snýr að öðru leyti að sama einkenni. Því er atriði 4 ef til vill að hafa áhrif á túlkun atriðis 5, en

þátttakendur gætu túlkað seinna atriðið á enn sértækari máta þegar það birtist í kjölfar atriðis 4 en á almennari máta ef það væri frammar á listanum. Þar að auki gæti birtingarmynd eirðarleysis verið ólík eftir þátttakendum og atriðið því gripið fleiri þátttakendur ef það væri orðað á almennari hátt.

Atriðið sem sýndi lakasta niðurstöðu var atriði 6 „orðið gröm/gramur eða pirruð/pirraður af minnsta tilefni“. Atriðið felur í raun í sér tvær spurningar sem flækir svörun, en annars vegar er spurt um gremju og hins vegar pírring. Slíkt getur leitt til vandkvæða í svörun til að mynda ef þátttakandi hefur upplifað pírring en ekki gremju eða á erfitt með að skilja þessi hugtök eða muninn á þeim. Ef til vill gæti atriðinu farnast betur og svörun einfaldast ef því væri skipt upp í tvö ný atriði sem spyrja annars vegar um gremju og hins vegar um pírring. Einnig væri athyglisvert að kanna skilning þátttakenda á þessu atriði, sér í lagi á hugtakinu gremja. Sé hugtakið þátttakendum torskilið gæti einfaldlega verið betra að spyrja eingöngu um pírring.

Atriði 7 „verið hrædd/-ur eins og eitthvað hræðilegt gæti gerst“ spyr um hræðslu sem vissulega ætti hugtakalega að tengjast kvíða og gerir það í mörgum tilfellum. Það er þó spurning hversu vel hræðsla á við í almennri kvíðaröskun þar sem aðaleinkenni röskunarinnar eru umfangsmiklar áhyggjur. Kenningar eru á lofti um að áhyggjur þeirra sem glíma við almennan kvíða séu í raun tegund af forðun eða bjargráðum sem koma í veg fyrir yfirgnæfandi hræðslu (Behar, DiMarco, Hekler, Mohlman og Staples, 2009). Auk þess hefur verið sýnt fram á að fólk með almenna kvíðaröskun hefur gjarnan jákvæðar hugmyndir um áhyggjur sínar, telur þær jafnvel vera dyggð eða jákvætt persónuleikaeinkenni sem hjálpar þeim að vera undirbúin fyrir allar aðstæður (Dugas, Gagnon, Ladouceur og Freeston, 1998). Í ljósi þessa er mögulegt að fólk með almennan kvíða upplifi áhyggjur sínar ekki sem hræðslu og eigi því erfitt með að svara þessari spurningu þó svo að það upplifi talsverðan kvíða. Þetta atriði á því ef til vill betur við um annars konar kvíða en almennan kvíða og því vafaatriði hvort það eigið yfirhöfuð að tilheyra GAD-7 sem ætlað er að skima fyrir almennri kvíðaröskun.

Varðandi mögulegar úrbætur á listanum almennt gæti reynst gagnlegt að líta til þeirra

atriða sem koma hvað best út, eru þetta atriði 3 og 4. Þessi atriði sýna hvað jafnasta dreifingu á svarmöguleikunum fjórum, hafa hæst meðaltöl og eru því líklegri til að grípa þá sem þjást af kvíða, jafnframt sýndu þau hæstar þáttahleðslur og þáttaskýringar. Það sem einkennir þessi atriði og aðskilur þau bæði frá seinustu þremur atriðum listans en einnig atriðum 1 og 2 sem þó koma vel út er einna helst að þau eru einföld. Þessi atriði eru stutt og hnitmiðuð og spyrja á einfaldan máta út í aðeins eitt einkenni röskunarinnar. Bendir þetta því til þess að gagnlegt gæti verið að einfalda önnur atriði listans, en einfaldleiki atriða gerir þau auðveldari í túlkun og kemur í veg fyrir misræmi í skilningi þátttakenda á þeim.

Niðurstöður rannsóknarinnar benda almennt til mögulegra vandamála í íslenskri þýðingu GAD-7 og að frekari rannsókna á listanum sé þörf. Í framhaldinu væri gagnlegt að beina sjónum betur að eiginleikum einstakra atriða og hvernig þeir sem svara listanum túlka þau. Beina þarf sjónum sérstaklega að seinustu þremur atriðum listans og mögulegum úrbótum á þeim. Einnig væri athyglisvert að skoða kynjamun í nýju úrtaki en niðurstöður þessarar rannsóknar benda til að virkni listans geti verið ólík eftir kynjum. Til að mynda er dreifing heildarskora karla takmörkuð og marktækur munur á meðaltölum kynjanna. Sambærilegar niðurstöður um kynjamun hafa fengist erlendis (Beard og Björgvinsson, 2014; Hinz o.fl., 2017). Kanna þarf nánar hvort að þessar niðurstöður endurspeglar raunverulegan kynjamun eða skekkju í mælitækinu. Ekki er þó hægt með góðu móti að skoða kynjamun í þessu úrtaki þar sem karlar eru aðeins 38 þátttakendur og missir því slík athugun marks og verður óáreiðanleg. Þetta slaka kynjahlutfall er einn helsti annmarki rannsóknarinnar ásamt stærð úrtaksins sem er talsvert lítið. Tegund úrtaks er þó einn helsti styrkleiki rannsóknarinnar og skýrir það að hluta smæð þess en erfitt getur reynst að nálgast stór klínísk úrtök. Annar styrkleiki felst í tölfræðiaðferðum sem beitt var við úrvinnslu á gögnum, en sérstaklega var tekið tillit til þess að gögnin voru á raðkvarða við tölfræðilega úrvinnslu.

Niðurstöður rannsóknarinnar lofa að mörgu leyti góðu fyrir gagnsemi íslenskrar þýðingar GAD-7 en vekja einnig upp ýmsar spurningar um próf-
fræðilega eiginleika hans og þar með spurningar

sem tengjast réttlætingu á notkun hans. Einna helst bendir rannsóknin til þess að skimunarhæfni listans sé ábótavant og að skoða þurfi hana frekar. Auk þess var komið auga á þrjú atriði sem hegða sér frábrugðið öðrum atriðum listans og lagfæra þarf til að notkun listans geti talist réttmæt.

Psychometric properties of the Icelandic version of GAD-7 in a clinical sample

GAD-7 is a seven-item self-report anxiety questionnaire designed to screen for general anxiety disorder and assess the seriousness of anxiety symptoms. The GAD-7 has been translated into a number of languages and studies on different versions of the scale indicate adequate internal reliability. However, studies have also shown an unstable factor structure and insufficient specificity. The purpose of this research was to evaluate the internal reliability and factor structure of the Icelandic version of GAD-7 in a clinical sample. Data from 266 patients in primary care (188 females and 38 males) with diagnoses of anxiety and/or depression disorders. The findings showed adequate internal consistency, $\alpha = 0,90$ but a low sum-score mean and percentage of positive screening results considering the nature of the sample. The findings also showed a clear one-factor structure, the one factor explained 57% of the variability of the correlation matrix. The factor loadings of all seven items were acceptable and ranged from 0,53 to 0,87 providing supporting evidence of the validity of the scale. Nevertheless, the factor analysis results also showed that the last three items did not function as well as the first four and possible improvements of the Icelandic translations are discussed. In conclusion the findings are promising but show that the Icelandic version of the GAD-7 needs further research and improvements.

Keywords: *GAD-7, psychometric properties, Icelandic version, anxiety, clinical sample.*

Daney Harðardóttir is a research assistant at Cicely Saunders Institute King's College in London, Vaka Vésteinsdóttir is an assistant professor at the Faculty of psychology, University of Iceland, Ragnhildur Lilja Ásgeirsdóttir is a PhD student in quantitative psychology at the Faculty of psychology, University of Iceland, Hafrún Kristjánsdóttir is an associate professor and dean of the Department of sport science, University of Reykjavík and Fanney Þórsdóttir is an associate professor at the Faculty of psychology, University of Iceland. Corresponding author is Fanney Þórsdóttir associate professor at the Faculty of psychology, University of Iceland, Nýi Garður, Sæmundargata 12, 102 Reykjavík. Email: fanneyt@hi.is

Heimildir

- Baglin, J. (2014). Improving your exploratory factor analysis for ordinal data: A demonstration using FACTOR. *Practical Assessment, Research & Evaluation*, 19(5).
- Beard, C. og Björgvinsson, T. (2014). Beyond generalized anxiety disorder: Psychometric properties of the GAD-7 in a heterogeneous psychiatric sample. *Journal of Anxiety Disorders*, 28(6), 547-552.
- Beavers, A. S., Lounsbury, J. W., Richards, J. K., Huck, S. W., Skolits, G. J. og Esquivel, S. L. (2013). Practical considerations for using exploratory factor analysis in educational research. *Practical Assessment, Research & Evaluation*, 18(6).
- Behar, E., DiMarco, I. D., Hekler, E. B., Mohlman, J. og Staples, A. M. (2009). Current theoretical models of generalized anxiety disorder (GAD): Conceptual review and treatment implications. *Journal of Anxiety Disorders*, 23(8), 1011-1023.
- Cattell, R. B. (1966). The screetest for the number of factors. *Multivariate Behavioral Research*, 1(2), 245-276.
- Costello, A. B og Osborne, J. W. (2005). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment, Research & Evaluation*, 10(7).
- D'Agostino, R. B., Belanger, A. og D'Agostino Jr., R. B. (1990). A suggestion for using powerful and informative tests of normality. *The American Statistician*, 44(4), 316-321.
- Delgadillo, J., Payne, S., Gilbody, S., Godfrey, C., Gore, S., Jessop, D. og Dale, V. (2012). Brief case finding tools for anxiety disorders: Validation of GAD-7 and GAD-2 in addictions treatment. *Drug and Alcohol Dependence*, 125(1-2), 37-42.
- de Winter, J. C. F. og Dodou, D. (2012). Factor recovery by principal axis factoring and maximum likelihood factor analysis as a function of factor pattern and sample size. *Journal of Applied Statistics*, 39(4), 695-710.
- Donker, T., van Straten, A., Marks, I. og Cuijpers, P. (2011). Quick and easy self-rating of generalized anxiety disorder: Validity of the Dutch web-based GAD-7, GAD-2 and GAD-SI. *Psychiatry Research*, 188, 58-64.
- Dugas, M. J., Gagnon, F., Ladouceur, R. og Freeston, M. H. (1998). Generalized anxiety disorder: A preliminary test of a conceptual model. *Behaviour Research and Therapy*, 36, 215-226.
- Einar Guðmundsson. (2011). Erlend norm og túlkun sálfræðilegra prófa. *Sálfræðiritið – Tímarit Sálfræðingafélags Íslands*, 16, 94-97.
- Erla Svansdóttir, Sesselja Hreggviðsdóttir, Björg Sigurðardóttir, Elísabet Benedíktz, Karl Andersen og Hróbjartur Darri Karlsson. (2018). Ótilgreindir brjóstverkir og tengsl við viðvarandi verkjaupplifun og vanlíðan. *Læknablaðið*, 104, 71-77.
- Fabrigar, L. R. og Wegener, D. T. (2012). *Exploratory Factor Analysis*. New York: Oxford university press.
- Fabrigar, L. R., Wegener, D. T., MacCallum, R. C. og Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods*, 4(3), 272-299.
- García-Campayo, J., Zamarano, E., Ruiz, M. A., Pardo, A., Pérez-Páramo, M., López-Gómez, V., ... Rejas, J. (2010). Cultural adaption into Spanish of the generalized anxiety disorder-7 (GAD-7) scale as a screening tool. *Health and Quality of Life Outcomes*, 8(1), 8-18.
- Gavrilov, I. og Pusev, R. (2014). normtest: Tests for normality. Forrit.
- Guttman, L. (1954). Some necessary conditions for common-factor analysis. *Psychometrika*, 19(2), 149-161.
- Henson, R. K. og Roberts, J. K. (2006). Use of exploratory factor analysis in published research: Common errors and some comment on improved practice. *Educational and Psychological Measurement*, 66(3), 393-416.
- Hinz, A., Klein, A. M., Brähler, E., Glaesmer, H., Luck, T., Reidel-Heller, S. G., ... Hilbert, A. (2017). Psychometric evaluation of the generalized anxiety disorder screener GAD-7, based on a large German general population sample. *Journal of Affective Disorders*, 210, 338-344.
- Holgado-Tello, F. P., Chacón-Moscoso, S., Barbero-García, I. og Vila-Abad, E. (2010). Polychoric versus Pearson correlations in exploratory and confirmatory factor analysis of ordinal variables. *Quality & Quantity*, 44(1), 153-166.
- Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika*, 30(2), 179-185.
- Inga Hrefna Jónsdóttir, Ingunn Hansdóttir, Ragnar Pétur Ólafsson, Sigurrós Jóhannsdóttir, Hákon Sigursteinson og Guðfinna Halla Þorvaldsdóttir. (2013). Mælitæki og sálfræðileg próf í notkun hjá sálfræðingum á Íslandi: Rafræn könnun prófnefndar Sálfræðingafélags Íslands. *Sálfræðiritið – Tímarit Sálfræðingafélags Íslands*, 18, 93-102.
- Jarque, C. M og Bera, A. K. (1987). A test for normality of observations and regression residuals. *International Statistical Review*, 55(2), 163-172.

- Kaiser, H. H. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement*, 20(1), 141-151.
- Kertz, S., Bigda-Peyton, J. og Björgvinsson, T. (2013). Validity of the generalized anxiety disorder-7 scale in an acute psychiatric sample. *Clinical Psychology & Psychotherapy*, 20, 456-464.
- Ledesma, R. D. og Valero-Mora, P. (2007). Determining the number of factors to retain in EFA: an easy-to-use computer program for carrying out parallel analysis. *Practical Assessment, Research & Evaluation*, 12(2).
- McNeish, D. (2018). Thanks coefficient alpha, we'll take it from here. *Psychological Methods*, 23(3), 412-433.
- Plummer, F., Manea, L., Trepel, D. og McMillan, D. (2016). Screening for anxiety disorders with the GAD-7 and GAD-2: a systematic review and diagnostic meta-analysis. *General Hospital Psychiatry*, 39, 24-31.
- Reise, S. P. (2012). The rediscovery of bifactor measurement models. *Multivariate Behavioral Research*, 47(5), 667-696.
- Reise, S. P., Bonifay, W. E. og Haviland, M. G. (2013). Scoring and modeling psychological measures in the presence of multidimensionality. *Journal of Personality Assessment*, 95(2), 129-140.
- Raíche, G. (2011). nFactors: Parallel analysis and non graphical solutions to the Cattell's scree test. Forrit.
- Raíche, G., Walls, T. A., Magis, D., Riopel, M., og Blais, J. G. (2013). Non-graphical solutions for Cattell's scree test. *Methodology: European Journal of Research Methods for the Behavioral and Social Sciences*, 9(1), 23-29.
- Revelle, W. R. (2017). psych: Procedures for personality and psychological research. Forrit.
- Ruiz, M. A., Zamorano, E., García-Campayo, J., Pardo, A., Freire, O. og Rejas, J. (2011). Validity of the GAD-7 scale as an outcome measure of disability in patients with generalized anxiety disorders in primary care. *Journal of Affective Disorders*, 128(3), 277-286.
- Ruscio, J. og Roche, B. (2012). Determining the number of factors to retain in an exploratory factor analysis using comparison data of known factorial structure. *Psychological Assessment*, 24(2), 282-292.
- Rutter, L. A. og Brown, T. A. (2017). Psychometric properties of the generalized anxiety disorder scale-7 (GAD-7) in outpatients with anxiety and mood disorders. *Journal of Psychopathology and Behavioral Assessment*, 39(1), 140-146.
- Sawaya, H., Atoui, M., Hamadeh, A., Zeinoun, P. og Nahas, Z. (2016). Adaptation and initial validation of the patient health questionnaire-9 (PHQ-9) and the generalized anxiety disorder-7 questionnaire (GAD-7) in an Arabic speaking Lebanese psychiatric outpatient sample. *Psychiatry Research*, 239, 245-252.
- Sousa, T. V., Viveiros, V., Chai, M. V., Vicente, F. L., Jesus, G., Carnot, M. J., ... Ferreira, P. L. (2015). Reliability and validity of the Portuguese version of the generalized anxiety disorder (GAD-7) scale. *Health and Quality of Life Outcomes*, 13(1).
- Spitzer, R. L., Kroenke, K., Williams, J. B. og Löwe, B. (2006). A brief measure for assessing generalized anxiety disorder: The GAD-7. *Arch Intern Med*, 166(10), 1092-1097.
- Stewart, A. L., Hays, R. D. og Ware, J. E. (1988). The MOS-short form general health survey. Reliability and validity in a patient population. *Medical Care*, 26(7), 724-735.
- Tabachnick, B. G. og Fidell, L. S. (2013). *Using Multivariate Statistics* (6. útgáfa). Upper Saddle River: Pearson.
- Teo, T. og Fan, X. (2013). Coefficient alpha and beyond: Issues and alternatives for educational research. *Asia-Pacific Edu Res*, 22(2), 209-213.
- Timmerman, M. E. og Lorenzo-Seva, U. (2011). Dimensionality assessment of ordered polytomous items with parallel analysis. *Psychological Methods*, 16(2), 209-220.
- van Buuren, S. og Groothuis-Oudshoorn, K. (2011). mice: Multivariate imputation by chained equations in R. *Journal of Statistical Software*, 45(3).
- Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis* (2. útgáfa). New York: Springer-verlag.
- Widaman, K. F. (1993). Common factor analysis versus principal component analysis: Differential bias in representing model parameters? *Multivariate Behavioral Research*, 28(3), 263-311.
- Williams, B., Onsmann, A. og Brown, T. (2010). Exploratory factor analysis: A five-step guide for novices. *Journal of Emergency Primary Health Care*, 8(3).
- Yap, B. W. og Sim, C. H. (2011). Comparisons of various types of normality tests. *Journal of Statistical Computation and Simulation*, 81(12), 2141-2155.
- Yong, A. G. og Pearce, S. (2013). A beginner's guide to factor analysis: focusing on exploratory factor analysis. *Tutorials in Quantitative Methods for Psychology*, 9(2), 79-94.
- Zwick, W. R. og Velicer, W. F. (1986). Comparison of five rules for determining the number of components to retain. *Psychological Bulletin*, 99(3), 432-442.