

Samkvæmni í mati á Orðskilningi og Líkingum í WASI^{IS}

Arna Rún Oddsdóttir og Einar Guðmundsson

Háskóli Íslands

Nákvæmni greindartalna byggist meðal annars á hlutlægum matsreglum. Markmið rannsóknarinnar er að athuga samkvæmni matsmanna á undirprófunum Orðskilningur og Líkingar í WASI^{IS}. Úrlausnir 70 einstaklinga voru valdar af handahófi úr stöðlunarúrtaki ($N = 700$) prófsins hérlendis. Þrjár sálfræðingar og þrjár sálfræðinemar mátu 70 svör í Orðskilningi og þrjár sálfræðingar og fjórir sálfræðinemar mátu 70 svör í Líkingum. Sálfræðinemar gefa að meðaltali fleiri stig heldur en sálfræðingar í Orðskilningi en í Líkingum er ekki munur á hópunum. Þegar mælitölur eru bornar saman er ekki munur á mati sálfræðinga og sálfræðinema í Líkingum en í Orðskilningi eru mælitölur hærri hjá sálfræðinimum en sálfræðingum. Áreiðanleikastuðull mats (e. *intra-class correlation coefficient*) í stigagjöf er 0,98 fyrir Orðskilning og 0,99 fyrir Líkingar. Samkvæmni mælitalna er 0,99 fyrir Orðskilning og 0,98 fyrir Líkingar. Niðurstöðurnar sýna að hægt er að meta svör eftir matsreglum í WASI^{IS} með áreiðanlegum hætti.

Efnisorð: WASI^{IS}, samkvæmni matsmanna, orðskilningur, líkingar, greindarpróf

Í stöðluðum greindarprófum er reynt að stuðla að samkvæmni í fyrirlögn og mati á úrlausnum með því að fylgja ítarlegum fyrirlagnar- og matsreglum. Hefðbundin leið til að kanna hlutlægni matsreglna er að athuga samkvæmni matsmanna (e. *inter-rater reliability*) sem meta sömu svör eða úrlausnir út frá sömu matsreglum (Shultz, Whitney og Zickar, 2014).

Margar rannsóknir hafa verið gerðar á algengi og tegundum villna þegar úrlausnir eru metnar í greindarprófum Wechslers fyrir börn (t.d. Alfonso, Johnson og Patinella, 1998; Belk, LoBello, Ray og Zachar, 2002; Loe, Kadlubek og Marks, 2007; Slate og Chick, 1989; Slate, Jones, Coulter og Covert, 1992) og fullorðna (t.d. Schwartz, 1966; Slate og Jones, 1989; Slate, Jones, Murray og Coulter, 1993). Niðurstöður þessara rannsókna sýna meðal annars að í munnlegum undirprófum á greindarprófum Wechslers er

ósamkvæmni í mati mest á undirprófunum Orðskilningur, Líkingar og Skilningur. Villur í mati á úrlausnum í greindarprófum Wechslers fyrir börn og fullorðna eru af ýmsu tagi. Algengt er til dæmis að svör séu ekki skráð á skráningarblöð sem gerir mat ónákvæmt eða jafnvel ómögulegt þegar fyrirlögn er lokið. Einnig má nefna ófullnægjandi fyrirlagnir, til dæmis þegar ekki er spurt nánar um óljós svör. Þá kemur fyrir að vitlaust sé gefið fyrir svör og villur gerðar í samlagningu stiga fyrir undirpróf eða þá að stigum er breytt vitlaust í mælitölur (Belk o.fl., 2002; Loe, Kadlubek og Marks, 2007).

Á Íslandi hafa nokkrar rannsóknir verið gerðar á samkvæmni matsmanna á nokkrum

Arna Rún Oddsdóttir cand. psych. er sálfræðingur og dr. Einar Guðmundsson er prófessor í sálfræði við sálfræðideild Háskóla Íslands. Fyrirspurnum um greinina skal beina til Örnú Rúnar Oddsdóttur (Tölvupóstur: arnarun@hotmail.com) eða Einars Guðmundssonar, Oddi, Sturlugata, 101 Reykjavík (Tölvupóstur: eing@hi.is).

undirprófum í greindarprófum Wechslers. Anna Sigríður Jökulsdóttir, Einar Guðmundsson, Gyða Haraldsdóttir, Rúnar Helgi Andrason og Ævar Árnason (2004) gerðu rannsókn á samkvæmni matsmanna á undirprófunum Orðskilningur og Líkingar í WPPSI-R^{IS} (*Wechsler Preschool and Primary Scale of Intelligence – Revised*) (Einar Guðmundsson og Hólmfríður Ólafsdóttir, 2003; Wechsler, 1989). Rannsóknin leiddi í ljós viðunandi samkvæmni milli matsmanna eða 0,88 fyrir Orðskilning og 0,96 fyrir Líkingar í staðlaðri útgáfu prófsins hérlendis. Í hliðstæðri rannsókn á samkvæmni í mati á tveim verklegum undirprófum í WPPSI-R^{IS}, Myndflötum og Völundarhúsum, var samkvæmni matsmanna heldur lægri (0,82 fyrir Myndfletti og 0,87 fyrir Völundarhús; Einar Guðmundsson, Áslaug Kristinsdóttir, Haukur Pálmason, Sigurður Levy og Þuríður Pétursdóttir, 2004). Eins hefur samkvæmni matsmanna fyrir Orðskilning og Líkingar verið athuguð í óstaðlaðri íslenskri þýðingu á WAIS-III (*Wechsler Adult Intelligence scale – Third Edition*) en þar var samkvæmnin óviðunandi í Orðskilningi (0,58) en meira samræmi í mati á Líkingum (0,88) (Helgi Sigurður Karlsson og Karl Fannar Gunnarsson, 2008). Með hliðsjón af því hversu miklar breytingar þarf að gera á fyrirlagnar- og matsreglum við þýðingu, staðfærslu og stöðlun greindarprófa koma þessar niðurstöður ekki á óvart. Ósamkvæmni í mati mismunandi matsaðila á óstöðluðum greindarprófum veldur skekkju í greindartölum og er ein af mörgum ástæðum þess að greindarpróf á ætíð að staðfæra og staðla áður en þau eru notuð í nýjum heimkynnum (Einar Guðmundsson, 2006, 2011).

WASI^{IS} (*Wechsler Abbreviated scale of Intelligence*) er stutt greindarpróf sem nýlega hefur verið staðlað hérlendis og gefið út til notkunar (Einar Guðmundsson, 2015). WASI^{IS} er fyrsta staðlaða greindarprófið fyrir fullorðna sem hefur verið gefið út á Íslandi. Prófið samanstendur af fjórum undirprófum (Rökþrautir, Orðskilningur, Litaletir og Líkingar), tveim greindarþáttum (munnlegum og verklegum) og heildartölu greindar.

Hægt er að reikna heildartölu greindar út frá tveim (Orðskilningur og Rökþrautir) eða fjórum undirprófum. Þó er ætlast til þess að tvö undirpróf séu aðeins notuð í undantekningartilvikum til að áætla heildartölu greindar. Próffræðilegir eiginleikar íslenskrar stöðlunar prófsins eru ágætir og sambærilegir við hliðstæða eiginleika í bandarískri útgáfu prófsins (Wechsler, 1999). Miðgildi áreiðanleikastuðla á öllum aldursbilum fyrir greindarþættina tvo og heildartölu greindar eru á bilinu 0,90 til 0,94 og þáttabygging prófsins (tveir greindarþættir) er í samræmi við erlenda útgáfu prófsins (Einar Guðmundsson, 2015, 2016). Í bandarískri útgáfu prófsins er fylgni milli hliðstæðra greindartalna í WASI og WAIS-III á bilinu 0,84 (munnleg greindartala) til 0,92 (heildartala greindar-fjögur undirpróf).

Samkvæmni í mati á undirprófunum Orðskilningur og Líkingar var athuguð í bandarískri útgáfu WASI. Niðurstöður sýndu að samkvæmni matsmanna í undirprófinu Orðskilningur var 98% og í Líkingum 99% (Wechsler, 1999). Þessar niðurstöður benda til þess að þó álitamál séu fleiri við mat á þessum undirprófum í WASI en í verklegum undirprófum (Rökþrautir og Litaletir), þá er hægt að meta svör með áreiðanlegum hætti eftir bandarískum matsreglum. Í handbók með bandarískri útgáfu WAIS-III (Wechsler, 1997) eru birtar niðurstöður úr rannsókn á samkvæmni milli matsmanna. Í rannsókninni voru þau undirpróf skoðuð þar sem líkur eru mestar á villum í mati á úrlausnum, þar á meðal í undirprófinu Líkingar. Þrjú einstaklingar mátu 60 úrlausnir í undirprófinu og var samkvæmni milli matsmanna 0,93.

Samkvæmni cand. psych. nema í mati á undirprófunum Orðskilningur og Líkingar í WASI^{IS} hefur verið athuguð í litlum úrtökum hérlendis. Þrjú cand. psych. nemar og einn sálfræðingur mátu svör 21 einstaklings í Orðskilningi (Hafdís Rósa Sæmundsdóttir, 2013) en fjórir cand. psych. nemar mátu svör sömu einstaklinga í Líkingum (Emanúel Geir Guðmundsson, 2013). Niðurstöðurnar benda til þess að samkvæmni í íslenskri útgáfu á

undirprófunum Orðskilningur og Líkingar sé viðunandi eða 0,97 fyrir Orðskilning og 0,98 fyrir Líkingar. Þó er nauðsynlegt að staðfesta þessar niðurstöður í stærra og fjölbreyttara úrtaki þar sem fleiri úrlausnir eru metnar og reyndir sálfræðingar meta svör.

Í þessari grein eru birtar niðurstöður rannsóknar á samkvæmni sex matsmanna við mat á 70 úrlausnum á undirprófinu Orðskilningur og sjö matsmanna á undirprófinu Líkingar í WASI^{IS}. Við þýðingu, staðfærslu og stöðlun WASI^{IS} var lögð mikil áhersla á hlutlægni matsreglna til að stuðla að hlutlægni niðurstöðu prófsins (Einar Guðmundsson, 2015, 2016). Þetta var meðal annars gert með því að hafa mörg dæmi um rétt og röng svör fyrir hvert atriði í Orðskilningi og Líkingum. Í rannsókninni er athugað hvort matsmenn með mismikla reynslu af fyrirlögn og mati á greindarprófum meti svör sömu einstaklinga á undirprófunum Orðskilningur og Líkingar svipað og um leið hvort niðurstaða greindarprófunar með WASI^{IS} sé óháð því hver leggur prófið fyrir og metur svör í munnlegum undirprófum. Í úrtaki matsmanna voru bæði starfandi sálfræðingar með reynslu af fyrirlögn og mati á svörum í greindarprófum Wechslers og cand. psych. namar í sálfræði með litla reynslu á þessu sviði. Samanburður á milli þessara hópa gefur góða mynd af hlutlægni matsreglna á undirprófunum Orðskilningur og Líkingar í WASI^{IS}.

Aðferð

Þátttakendur

Dregið var handahófsúrtak 70 úrlausna fyrir Orðskilning og Líkingar úr stöðlunarúrtaki ($N = 700$) WASI^{IS} fyrir fólk á aldrinum 17 til 64 ára. Úrlausnirnar dreifðust jafnt á sjö aldursbil prófsins. Tíu úrlausnir voru dregnar af tilviljun á hverju aldursbili (17-19 ára, 20-24 ára, 25-29 ára, 30-34 ára, 35-44 ára, 45-54 ára og 55-64 ára).

Matsmenn fyrir Orðskilning. Hver úrlausn var metin af sex matsmönnum. Þrír þeirra voru starfandi sálfræðingar og þrír sálfræðinemar

á öðru ári í cand. psych. námi við Háskóla Íslands. Allir matsmenn voru konur. Matsmenn tóku ekki þátt í stöðlun prófsins og voru því að meta úrlausnir í WASI^{IS} í fyrsta skipti.

Matsmenn fyrir Líkingar. Hver úrlausn var metin af sjö matsmönnum. Þar af voru þrír starfandi sálfræðingar (aðrir en þeir sem mátu svör í Orðskilningi) og fjórir sálfræðinemar (aðrir en þeir sem mátu svör í Orðskilningi) á öðru ári í cand. psych. námi við Háskóla Íslands. Allir matsmenn voru konur. Í heild tóku 13 matsmenn þátt í að meta úrlausnir. Enginn matsmaður mat svör bæði í Orðskilningi og Líkingum. Matsmenn tóku ekki þátt í stöðlun prófsins og voru því að meta úrlausnir í WASI^{IS} í fyrsta skipti.

Mælitæki

Stöðlun WASI^{IS}. Lagskipt kvótaúrtak fullorðinna á aldrinum 17 til 64 ára á landinu öllu var notað við stöðlun prófsins. Stöðlunarúrtak WASI^{IS} ($N = 700$) samsvarar vel þýði fullorðinna á Íslandi eftir búsetu, kyni, aldri og menntun (Einar Guðmundsson, 2015, 2016). Í þessari rannsókn voru notuð bráðabirgðaviðmið til að breyta stigum í mælitölur þar sem endanleg aldursviðmið með staðlaðri útgáfu prófsins voru ekki tilbúin þegar rannsóknin fór fram. Summu stiga fyrir undirprófin tvö var breytt í T-tölur. Meðaltal normaldreifðra T-talna er 50 og staðalfrávik 10. Mælitölurnar geta verið á bilinu 20 til 80. Meðaltal normaldreifðra greindarþátta og heildartölur greindar er 100 og staðalfrávik 15.

Öll svör við atriðum í Orðskilningi og Líkingum voru skráð í Excel. Munnleg undirpróf (Orðskilningur og Líkingar) í WASI^{IS} voru notuð í þessari rannsókn.

Orðskilningur. Í undirprófinu Orðskilningur eru 36 orð lesin upp og próftaki getur auk þess lesið orðin á spjaldi um leið og atriðin eru lögð fyrir. Próftakar eru beðnir að útskýra munnlega hvað hvert orð þýðir. Svör próftaka eru færð á skráningarblað og stig eru gefin fyrir hvert svar eftir matsreglum prófsins. Tvö stig eru gefin fyrir rétt og fullnægjandi svar, eitt stig er gefið fyrir svar sem er ófullkomið en þó

ekki rangt og ekkert stig er gefið fyrir rangt eða óviðeigandi svar. Allir próftakar byrja á fyrsta atriði undirprófsins. Þegar próftaki hefur svarað rangt fimm sinnum í röð er fyrirlögn undirprófsins hætt. Ekki eru sérstök tímamörk í þessu undirprófi.

Líkingar. Í undirprófinu Líkingar eru 22 atriði. Tvö orð eru lesin upp og próftaki útskýrir hvernig þau eru lík eða eins. Svör próftaka eru skráð á skráningarblað og stig eru gefin fyrir hvert svar eftir matsreglum prófsins. Tvö stig eru gefin fyrir rétt og fullnægjandi svar, eitt stig er gefið fyrir svar sem er ófullkomið en þó ekki rangt og ekkert stig er gefið fyrir rangt eða óviðeigandi svar. Í Líkingum er byrjað á því að leggja fyrir sjötta atriðið. Ef ekkert eða eitt stig er gefið fyrir atriði sex og sjö eru atriði eitt til fimm lögð fyrir í öfugri röð þar til tvö stig hafa verið gefin tvisvar í röð. Ef svarað er rangt fjórum sinnum í röð er fyrirlögn undirprófsins hætt. Ekki eru sérstök tímamörk í þessu undirprófi.

Framkvæmd

Sálfræðingar og cand. psych. nemar í sálfræði voru beðnir um að vera matsmenn. Matmenn sem samþykktu þátttöku fengu leiðbeiningar um framkvæmd matsins og matsreglur fyrir Orðskilning (sex matsmenn) og Líkingar (sjö matsmenn). Þeim var bent á að hafa til hliðsjónar bæði almennar reglur um stigagjöf og dæmi um svör við matið, bæði í Orðskilningi og Líkingum. Í matsreglunum kemur fram að tvö stig eru gefin ef svar er jafngott eða betra en dæmi um tveggja stiga svar í matsreglunum. Ekkert stig er gefið ef svar er jafn slakt eða slakara en dæmi um núll stigs svar í matsreglunum. Á sama hátt er gefið eitt stig ef svar er jafngott og dæmi um eins stigs svar í matsreglunum. Loks kemur fram í matsreglunum að leggja eigi áherslu á inntak svara en ekki glæsileik í málfari eða orðskrá. Ekki á að refsja fyrir slakan framburð eða slaka málnotkun í stigagjöf á munnlegum undirprófum í WASI^{IS}.

Úrlausnir 70 einstaklinga úr stöðlun prófsins fyrir Orðskilning og Líkingar voru fluttar

í Excel. Úrlausnum var ruglað þannig að niðurstöður sem tilheyrðu vissu aldursbili voru ekki hlið við hlið. Þetta skjal fengu matsmenn sent til sín í tölvupósti og í það skráðu þeir þann fjölda stiga sem þeir gáfu fyrir hvert svar. Matsmenn sáu hvorki mat hvers annars né hvernig svör voru metin í stöðluninni. Engin tímamörk voru sett til að ljúka matinu en matsmenn voru beðnir um að skila matinu eins fljótt og hægt væri.

Matsmenn voru beðnir um að kynna sér matsreglur undirprófsins vel sem þeir fengu úthlutað áður en þeir byrjuðu matið. Matsmenn fengu að vita að hægt væri að gefa tvö stig fyrir rétt svar, eitt stig fyrir svar sem sýnir takmarkaðan skilning og ekkert stig fyrir rangt svar. Matsmenn voru beðnir um að ljúka stigagjöf fyrir hvern þátttakanda áður en þeir byrjuðu á þeim næsta. Þetta var gert til að líkja sem mest eftir raunverulegum matsaðstæðum.

Að síðustu voru matsmenn beðnir um að senda niðurstöður sínar á umsjónarmann verkefnis í tölvupósti. Þeir voru einnig beðnir að eyða Excel skjalinu sem þeir fengu sent í tölvupósti úr pósthólfinu sínu sem og í tölvunni sinni. Engin auðkenni fylgdu gögnunum sem matsmennirnir fengu. Það var því ómögulegt að rekja svör til þátttakenda í stöðlun WASI^{IS}. Matsmenn voru hvattir til að hafa samband við umsjónarmann verkefnis ef einhverjar spurningar vöknudu.

Úrvinnsla gagna

Þegar matsmenn höfðu gefið stig fyrir öll svör sem þeim var úthlutað og skilað til rannsakenda voru gögnin færð í SPSS. Öll úrvinnsla gagnanna fór síðan fram í forritinu.

Við úrvinnslu gagnanna var sama aðferð notuð og kynnt er í rannsókn Önnu Sigríðar Jökulsdóttur og félagar (2004). Til að meta samræmi matsmanna var notaður áreiðanleikastuðull fyrir samkvæmi margra matsmanna (e. *intraclass correlation coefficient*). Þessi stuðull gefur til kynna dreifingu vegna matsmanna á móti dreifingu vegna matsmanna og dreifingu vegna annarra þátta (Endicott, Tracy, Burt, Olson og Coccato,

2002; ShROUT og Fleiss, 1979). Gildi nálægt einum gefur til kynna mikið samræmi í mati margra matsmanna en gildi nálægt núlli gefur til kynna lítið samræmi.

Gögnin voru lítillega neikvætt skekkt en ekki var talið að það hefði afgerandi áhrif á niðurstöður. Þar sem allir próftakar voru metnir af öllum matsmönnum er hægt að greina áhrif matsmanna frá áhrifum á samvirkni matsmanna og próftaka sem og tilviljunarvillu. Í þessari rannsókn er miðað við mat stakra matsmanna.

Niðurstöður

Í 1. töflu kemur fram meðaltal og staðalfrávik sex matsmanna sem mátu Orðskilning og sjö matsmanna sem mátu Líkingar. Í Orðskilningi er það matsmaður þrjú (sálfræðingur) sem hefur lægst meðaltal stiga (37,23) og auk þess lægsta staðalfrávik (7,85). Matsmaður fjögur (sálfræðinemi) hefur hæst meðaltal stiga eða 45,14 en matsmaður tvö (sálfræðingur) er með hæsta staðalfrávik eða 10,43. Niðurstöður einhliða dreifigreiningar í Orðskilningi

benda til þess að munur sé á meðaltölum matsmanna ($F(5,414) = 7,25, p < 0,001$). Samkvæmt eftiráprófi Scheffe er meðaltal matsmanns þrjú marktækt lægra en meðaltal allra annarra matsmanna nema matsmanns eitt (sálfræðingur). Reiknaður var munur milli hópa með t -prófi. Samkvæmt t -prófi er munur á meðaltali summu stiga eftir því hvort matsmaður er sálfræðingur eða sálfræðinemi ($t(418) = -3,87, p < 0,001$). Sálfræðinemar gefa að meðaltali fleiri stig heldur en sálfræðingar.

Í Líkingum er matsmaður sex (sálfræðinemi) með lægst meðaltal stiga eða 29,29 og auk þess lægsta staðalfrávik (5,12). Matsmaður fimm (sálfræðinemi) er með hæst meðaltal stiga eða 31,94 og auk þess hæsta staðalfrávik (5,57). Niðurstöður einhliða dreifigreiningar benda ekki til að munur sé á meðaltölum matsmanna ($F(6,483) = 1,72, p = 0,114$). Reiknaður var munur milli hópa með t -prófi. Samkvæmt t -prófi er ekki munur milli hópa í undirprófinu Líkingar ($t(488) = -0,987, p = 0,324$).

Í 2. töflu má sjá meðaltal og staðalfrávik mælitalna fyrir allar úrlausnirnar. Í Orðskilningi

1. tafla. Meðaltal og staðalfrávik stiga á undirprófunum Orðskilningur og Líkingar í WASIIS.

	Matsmaður	Meðaltal	Staðalfrávik	Fjöldi úrlausna
Orðskilningur				
Matsmaður 1	Sálfræðingur	40,93 ^{ab}	8,52	70
Matsmaður 2	Sálfræðingur	44,23 ^a	10,43	70
Matsmaður 3	Sálfræðingur	37,23 ^b	7,85	70
Matsmaður 4	Sálfræðinemi	45,14 ^a	10,17	70
Matsmaður 5	Sálfræðinemi	44,56 ^a	9,28	70
Matsmaður 6	Sálfræðinemi	43,50 ^a	9,65	70
Líkingar				
Matsmaður 1	Sálfræðingur	31,57	5,42	70
Matsmaður 2	Sálfræðingur	30,89	5,30	70
Matsmaður 3	Sálfræðingur	30,99	5,28	70
Matsmaður 4	Sálfræðinemi	30,54	5,51	70
Matsmaður 5	Sálfræðinemi	31,94	5,57	70
Matsmaður 6	Sálfræðinemi	29,29	5,12	70
Matsmaður 7	Sálfræðinemi	30,87	5,43	70

Aths. Munur er á meðaltölum með ólíka bókstafi skv. Scheffe prófi ($p < 0,05$).

er það matsmaður þrjú (sálfræðingur) sem hefur lægst meðaltal mælitalna eða 37,97 og auk þess lægsta staðalfrávikkið eða 7,67. Matsmaður fjögur (sálfræðinemi) er með hæsta meðaltal mælitalna eða 45,00 og hæsta staðalfrávikkið eða 10,01. Niðurstöður einhliða dreifigreiningar benda til þess að munur sé á meðaltölum mælitalna matsmanna ($F(5,414) = 6,30, p < 0,001$). Samkvæmt eftiraprófi Scheffe er meðaltal matsmanns þrjú (sálfræðingur) marktækt lægra en meðaltal allra annarra matsmanna nema matsmanns eitt (sálfræðingur). Reiknaður var munur milli hópa með t -prófi. Munur er á meðaltali mælitalna eftir því hvort matsmaður er sálfræðingur eða sálfræðinemi ($t(418) = -3,61, p < 0,001$). Sálfræðinemar virðast að öllu jöfnu meta herra heldur en sálfræðingar.

Í Líkingum er lægsta meðaltalið hjá matsmanni sex (sálfræðinemi) eða 37,00 og auk þess lægsta staðalfrávikkið, 5,49. Matsmaður fimm (sálfræðinemi) hefur hæsta meðaltalið eða 40,93 og auk þess hæsta staðalfrávikkið, 7,13. Niðurstöður einhliða dreifigreiningar benda til að munur sé á meðaltölum matsmanna

($F(6,483) = 2,60, p = 0,017$). Samkvæmt eftiraprófi Scheffe er meðaltal matsmanns sex (sálfræðinemi) marktækt lægra en meðaltal matsmanns fimm (sálfræðinemi). Ekki er munur milli annarra matsmanna. Reiknaður var munur milli hópa með t -prófi. Ekki var munur milli hópa ($t(488) = -1,14, p = 0,257$).

Samkvæmni í stigagjöf eða summu stiga milli matsmanna (e. *intraclass correlation coefficient*) á undirprófinu Orðskilningur er 0,98 og á mælitölum 0,99. Í undirprófinu Líkingar er samkvæmni á summu stiga milli matsmanna 0,99 og á mælitölum 0,98. Samkvæmni á báðum undirprófunum er því viðunandi.

Samkvæmni matsmanna var athuguð fyrir hvert atriði í Orðskilningi og Líkingum. Í Orðskilningi er samkvæmni milli matsmanna á bilinu 0,76 (atriði númer 8 í WASIS) til 0,99 (atriði númer 11). Af 36 atriðum í Orðskilningi er samkvæmni í 25 atriðum 0,90 eða hærri. Í þremur atriðum er áberandi minnst samkvæmni milli matsmanna (atriði númer 8 (0,76), 17 (0,80) og 29 (0,83)).

Í Líkingum er samkvæmni matsmanna á bilinu 0,79 (atriði númer 3) til 0,99 (atriði

2. tafla. Meðaltal og staðalfrávik mælitalna á undirprófunum Orðskilningur og Líkingar í WASIS.

	Matsmaður	Meðaltal	Staðalfrávik	Fjöldi úrlausna
Orðskilningur				
Matsmaður 1	Sálfræðingur	41,21 ^{ab}	8,20	70
Matsmaður 2	Sálfræðingur	44,31 ^a	9,93	70
Matsmaður 3	Sálfræðingur	37,97 ^b	7,67	70
Matsmaður 4	Sálfræðinemi	45,00 ^a	10,01	70
Matsmaður 5	Sálfræðinemi	44,64 ^a	9,20	70
Matsmaður 6	Sálfræðinemi	43,61 ^a	9,19	70
Líkingar				
Matsmaður 1	Sálfræðingur	40,27 ^{ab}	6,61	70
Matsmaður 2	Sálfræðingur	39,36 ^{ab}	6,25	70
Matsmaður 3	Sálfræðingur	39,34 ^{ab}	6,16	70
Matsmaður 4	Sálfræðinemi	38,93 ^{ab}	6,61	70
Matsmaður 5	Sálfræðinemi	40,93 ^a	7,13	70
Matsmaður 6	Sálfræðinemi	37,00 ^b	5,49	70
Matsmaður 7	Sálfræðinemi	39,10 ^{ab}	6,31	70

Aths. Munur er á meðaltölum með ólíka bókstafi skv. Scheffe prófi ($p < 0,05$).

númer 8 og 17) í 22 atriðum undirprófsins. Af 22 atriðum í Líkingum er samkvæmni 0,90 eða hærrí í 20 atriðum. Í tveimur atriðum í Líkingum er samkvæmni lægri en 0,90 (atriði númer 3 (0,79) og 19 (0,85).

Umræða

Samkvæmni matsmanna á undirprófunum Orðskilningur og Líkingar í WASI^{IS} er viðunandi hvort sem miðað er við summu stiga eða mælitölur. Samkvæmni matsmanna í stigagjöf er 0,98 fyrir Orðskilning og 0,99 fyrir Líkingar. Samkvæmni mælitalna er 0,99 fyrir Orðskilning og 0,98 fyrir Líkingar. Ekki er munur á mati sálfræðinga og sálfræðinema í Líkingum en sálfræðinemar gefa fleiri stig en sálfræðingar í Orðskilningi. Almennt er samkvæmni matsmanna ásætlanleg á munnlegu undirprófunum Orðskilningur og Líkingar í WASI^{IS}.

Niðurstöður rannsóknarinnar eru sambærilegar við rannsóknir á samkvæmni matsmanna á sömu undirprófum í bandarískri útgáfu WASI (Wechsler, 1999). Niðurstöðunum svipar einnig til niðurstaðna Önnu Sigríðar og félaga (2004) á samkvæmi milli matsmanna á greindarprófinu WPPSI-R^{IS} en í þeirri rannsókn var samkvæmni í mati á Orðskilningi og Líkingum einnig viðunandi. Það kemur ekki á óvart þar sem sama aðferðafræði var notuð við þýðingu og staðfærslu matsreglna í WASI^{IS} og WPPSI-R^{IS}. Loks staðfesta niðurstöður í þessari rannsókn niðurstöður þeirra Hafðísar Rósu Sæmundsdóttur (2013) og Emanúels Geirs Guðmundssonar (2013) á samkvæmni matsmanna á undirprófunum Orðskilningur og Líkingar í WASI^{IS}. Þó svo að munur færri svör hafi verið metin í þeirra rannsóknunum og samkvæmni athuguð hjá færri matsmönnum en í þessari rannsókn eru niðurstöður svipaðar. Í þeirra rannsóknunum var samkvæmni matsmanna í Orðskilningi 0,97 og í Líkingum 0,98.

Í Orðskilningi kemur fram munur á stigagjöf sálfræðinga og sálfræðinema þegar mat þeirra er tekið saman. Í Líkingum er aftur á móti ekki munur á stigagjöf þeirra.

Starfandi sálfræðingar gáfu að meðaltali færri stig heldur en sálfræðinemar í Orðskilningi. Einn af þremur sálfræðingum gefur áberandi færri stig en allir sálfræðinemarnir. Það skýrir mun á hópunum tveim en tilhneiging er þó til staðar hjá sálfræðinimum að gefa herra en sálfræðingar.

Almennt á við að samkvæmni í mati í Orðskilningi og Líkingum er viðunandi í WASI^{IS}, bæði fyrir undirprófin í heild og einstök atriði undirprófanna. Matsreglur í WASI^{IS} fyrir Orðskilning og Líkingar eru því hlutlægar og stuðla að því að sérkenni þess sem metur svör eftir matsreglunum hafa engin eða hverfandi áhrif á greindartölur prófsins.

Interscorer agreements on the Vocabulary and Similarities subtests of the WASI^{IS}

The accuracy of IQ scores depends on the objectivity of subtests' scoring criteria. The aim of this study is to investigate the interscorer agreements on the Vocabulary and Similarities subtests of the WASI^{IS}. A sample of 70 cases was randomly selected from the Icelandic standardization sample of WASI. Three psychologists and three psychology graduate students independently scored each of the 70 Vocabulary protocols and three psychologists and four psychology graduate students independently scored each of the 70 Similarity protocols. On average the psychology graduate students' raw scores are higher than psychologists' on the Vocabulary subtest but the two groups do not differ on the Similarities subtest. The interrater reliability coefficients across the raters were 0.98 and 0.99 for Vocabulary and Similarities raw scores, respectively. The interrater reliability coefficients across the raters were 0.99 and 0.98 for Vocabulary and Similarities scale scores, respectively. These results show that the Vocabulary and Similarities subtests can be scored reliably in the Icelandic standardization of WASI.

Keywords: WASI^{IS}, interscorer agreements, vocabulary, similarities, intelligence tests

Arna Rún Oddsdóttir cand. psych. is a psychologist and Dr. Einar Guðmundsson is a professor at the University of Iceland, Department of Psychology. Correspondence to: Arna Rún Oddsdóttir (Email: arnarun@hotmail.com) or Dr. Einar Guðmundsson, Faculty of Health Sciences, Department of Psychology, Oddi, Sturlugata, 101 Reykjavík (Email: eing@hi.is).

Heimildir

- Alfonso, V.C., Johnson, A. og Patinella, L. (1998). Common WISC-III examiner errors: Evidence from graduate students in training. *Psychology in the Schools*, 35, 119-124.
- Anna Sigríður Jökulsdóttir, Einar Guðmundsson, Gyða Haraldsdóttir, Rúnar Helgi Andrason og Ævar Árnason (2004). Samkvæmni í mati á munnlegum undirprófum WPPSI-R^{IS}. *Sálfræðiritið – Tímarit Sálfræðingafélags Íslands*, 9, 77-89.
- Belk, M. S., LoBello, S. G., Ray, G. E. og Zachar, P. (2002). WISC-III administration, clerical, and scoring errors made by student examiners. *Journal of Psychoeducational Assessment*, 20, 290-300.
- Einar Guðmundsson (2006). Þýðing og staðfærsla sálfræðilegra prófa. *Sálfræðiritið – Tímarit Sálfræðingafélags Íslands*, 10-11, 23-40.
- Einar Guðmundsson (2011). Erlend norm og túlkun sálfræðilegra prófa. *Sálfræðiritið – Tímarit Sálfræðingafélags Íslands* (rafrænt fylgirit), 16, 94-97.
- Einar Guðmundsson (2015). *Mat á greind fullorðinna*. Reykjavík: Menntamálastofnun.
- Einar Guðmundsson (2016). Staðlað greindarpróf fyrir fullorðna á Íslandi: WASI^{IS}. *Sálfræðiritið – Tímarit Sálfræðingafélags Íslands*, 20-21, 7-22.
- Einar Guðmundsson og Hólmfríður Ólafsdóttir (2003). *Greindarpróf David Wechsler handa börnum á leikskóla- og grunnskólaaldri. Íslensk staðfærsla og stöðlun. Handbók*. Reykjavík: Námsmatsstofnun.
- Einar Guðmundsson, Áslaug Kristinsdóttir, Haukur Pálmason, Sigurður Levy og Þuríður Pétursdóttir (2004). Samkvæmni í mati á Myndflötum og Völundarhúsum í WPPSI-R^{IS}. *Sálfræðiritið – Tímarit Sálfræðingafélags Íslands*, 9, 91 – 100.
- Emanúel Geir Guðmundsson (2013). *Undirbúningur stöðlunar WASI fyrir 17 til 64 ára Íslendinga og samkvæmni matsmanna á undirprófinu Líkingar: Úrtak utan höfuðborgar*. Óbirt cand. psych. ritgerð. Háskóli Íslands, Heilbrigðisvísindasvið.
- Endicott, J., Tracy, K., Burt, D., Olson, E. og Cocco, E.F. (2002). A novel approach to assess inter-rater reliability in the use of the overt aggression scale-modified. *Psychiatry Research*, 112, 153-159.
- Hafðís Rósa Sæmundsdóttir (2013). *Undirbúningur stöðlunar á WASI fyrir 17-64 ára og samkvæmni matsmanna á undirprófinu Orðskilningur*. Óbirt cand. psych. ritgerð. Háskóli Íslands, Heilbrigðisvísindasvið.
- Helgi Sigurður Karlsson og Karl Fannar Gunnarsson (2008). *Samkvæmni í mati á þremur munnlegum undirprófum WAIS-III*. Óbirt BA ritgerð. Háskólinn á Akureyri, Hug- og félagsvísindasvið.
- Loe, S.A., Kadlubek, R.M. og Marks, W.J. (2007). Administration and scoring errors on the WISC-IV among graduate student examiners. *Journal of Psychoeducational Assessment*, 25, 237-247.
- Shultz, K.S., Whitney, D.J. og Zickar, J. (2014). *Measurement theory in action. Case studies and exercises* (2. útgáfa). Los Angeles: Sage.
- Shrout, P.E. og Fleiss, J.L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin*, 86, 420-428.
- Slate, J.R. og Chick, D. (1989). WISC-R examiner errors: Cause for concern. *Psychology in the Schools*, 26, 78-84.
- Slate, J.R. og Jones, C.H. (1989). Examiner errors on the WAIS-R: A source of concern. *The Journal of Psychology*, 124, 343-345.
- Slate, J.R., Jones, C.H., Coulter, C. og Covert, T.L. (1992). Practitioners administration and scoring of the WISC-R: Evidence that we do err. *Journal of School Psychology*, 30, 77-82.
- Slate, J.R., Jones, C.H., Murray, R.A. og Coulter, C. (1993). Evidence that practitioners err in administering and scoring the WAIS-R. *Measurement and Evaluation in Counseling and Development*, 25, 156-161.
- Schwartz, M.L. (1966). The scoring of WAIS comprehension responses by experienced and inexperienced judges. *Journal of Clinical Psychology*, 22, 425-427.
- Wechsler, D. (1989). *Manual for the Wechsler Preschool and Primary Scale of Intelligence-Revised*. New York: Psychological Corporation.
- Wechsler, D. (1997). *Wechsler Adult Intelligence Scale—Third Edition*. San Antonio, TX: The Psychological Corporation.
- Wechsler, D. (1999). *Wechsler Abbreviated Scale of Intelligence (WASI). Manual*. San Antonio, Texas: Psychological Corporation.