

PSZICHOMETRIAI FEJLESZTÉSI LEHETŐSÉGEK AZ ONLINE TESZTELÉSBEN



NAGYBÁNYAI NAGY Olivér

Károli Gáspár Református Egyetem, Pszichológiai Intézet
nagybanyai.oliver@kre.hu

ÖSSZEFOGLALÓ

Háttér és célkitűzések: Az internetes pszichológiai tesztelés révén számos új elemmel bővíthető a kérdőíves módszertan a hagyományos papír-ceruza mérőeszközökhöz képest. Nemcsak maga a tesztfelvétel, a kiértékelés, valamint az eredmények értelmezése válik lényegesen hatékonyabbá, hanem számos pszichometriai, módszertani újításra is lehetőséget teremt az online tesztelés. Ez utóbbiak kapcsán az ún. válaszstílus-jellemzők szélesíthetik a tesztek mérési spektrumát. *Módszer:* Jelen vizsgálat egy kiterjedtebb kutatás részeként ($N = 1026$) a Facet5 online személyiségteszt révén elemzi a válaszszélsőségesség globális pontszámtorzító hatását, emellett egy kisebb mintán ($N = 137$) más személyiség-kérdőíveken kapott eredmények (NEO PI-3, TCI) alapján demonstrálja a jelenség általános jellegét, valamint javaslatot tesz a válaszgyakorisági jellemzők mérésével és kontrollálásával kapcsolatban. *Eredmények:* Egy adott teszten belül a személyiségkálánkénti válaszszélsőségesség-mutatók együttjárása ($r = 0,7$; $p < 0,01$) arra utal, hogy egy tartalomtól független tényező is befolyásolhatja, torzíthatja a tesztpontszámok alakulását. Egy általánosan érvényesülő, globális hatásról van szó, hiszen a különféle személyiségtesztek összesített válaszszélsőségesség-indexei között is közepesen erős ($r = 0,5-0,6$; $p < 0,01$) kapcsolatok mutathatók ki. A teljes kérdőív során mutatott válaszadás alapján egymástól két jól elkülöníthető válaszszélsőségesség-klaszter mutatható ki, melyekhez személyiségkálánként szignifikánsan eltérő saját standard átlagok képezhetők a választendenciából fakadó általános pontszámeltolódások ellensúlyozására. *Következtetések:* Az eredmények megerősítik a választendencia konzisztens, személyre jellemző megjelenésének tényét. Ez az egyéni válaszstílus a mérési témáktól függetlenül megjelenik és a kérdőívek között is időben stabilan kimutatható, így egyfajta globális rárakódó hatásként befolyásolja a teszteredmények alakulását.

Kulcsszavak: Big Five kérdőív, online tesztelés, szélsőséges és mérsékelt válaszstílus

AZ ONLINE TESZTELÉSI MÓDSZER ÉS A VÁLASZJELLEMZŐK DETEKTÁLÁSA A KÉRDŐÍVES KUTATÁSOKBAN

A személyiségtesztelés terén a klasszikus papír-ceruza verzióhoz képest a számítógépes és online kiértékelési és értelmezési módszerek ma már lényegében általánossá váltak. Bár a számítógépes és online tesztelést gyakran egymás mellett emlegetjük, valójában különálló területeknek tekinthetők, miután az internetes módszerek terjedése az utóbbi évtized során számottevőbbé vált. Ez utóbbiak esetén nem számítógéphez kötöttek zajlik a tesztkitöltés és kiértékelés, sőt mindinkább jellemző, hogy a különféle mobil eszközökön is elérhetővé válnak az internetalapú mérőeszközök.

A tesztkiértékelés módszertana is számos új formai és módszertani elemmel bővült. Részben a különféle nyelvi változatok elérhetősége által biztosított nagymértékű rugalmasság nyomán mind a teszt kitöltésének nyelve, mind pedig a pontszám-kiértékelés, illetve az automatikus értelmezések (riportok) nyelve is az adott cél szerint történhet. Ehhez kapcsolódóan a pontszámítás alapjául szolgáló standard norma megválasztása sem csak egy-egy nyelvterületen áll rendelkezésre, hanem azon belül különféle speciális almintákra vonatkozóan is variálható.

A tesztkitöltés és tesztkiértékelés terén a nyelvi, tartalmi alapú jellegzetességek mellett az online tevékenységek közben keletkező, digitális „lenyomatok” révén egyéb másodlagos válaszjellemzők vizsgálatára is mód nyílik. Digitális lábnyomnak („*digital footprint*”) az olyan információkat nevezük, melyeket online tevékenységek során akarva-akaratlanul maguk után hagynak a számítógép, az online felület használói.

Ezeket az internetes szakirodalomban általánosságban a nyilvánosan, vagyis mások által is hozzáférhető adatokra értjük (Grayson, 2011; Madden, Fox, Smith és Vitak, 2007). A szándékosan publikált, direkt információs tartalmak mellett viszont számottevő adatforrásnak számítanak az online tevékenységek közben keletkező, ún. indirekt digitális lenyomatok (például a vásárlási szokásokkal, bizonyos internetes oldalak látogatásával, a mobiltelefonos ügyintézésel összefüggő rekordok). Annyiban speciális ez a vizsgálati terület, hogy ezen a módon ténylegesen rejtett információk, akár az egyén számára sem ismert folyamatok, döntési mechanizmusok, személyes preferenciák válhatnak feltárhatóvá.

A teszthasználó terén – legyen szó akár marketingkutatásokról, vagy éppen pszichológiai tesztekéről – ilyen jellegű „látens” információval elsődlegesen a *válaszadási mintázat* („*response pattern*”) mutatói szolgálhatnak, melynek legszélesebb területét a *válaszgyakoriságok* („*response distribution*”) elemzése képezi. Ez az adott válaszkála (például az ötfokú Likert-skála) egyes fokozatainak választási gyakoriságát, vagyis a statisztikai valószínűségtől eltérő válaszmelegeszlások tendenciózus jellegét vizsgálja (Berg és Collier, 1953). Bár itt számos egymáshoz kapcsolódó fogalommal találkozhatunk, de ezek néhol eltérő meghatározással és jelentéstartalommal bírnak. Jelen esetben a *válaszgyakoriságokat* definíció szerint olyan szisztematikus torzítási tendenciáknak tekintjük, melyek esetén a válaszadók nem (vagy nem csak) az adott tételek tartalma alapján adják a konkrét kérdőívben a válaszaikat, hanem ezt egyéb ráerakódó hatások is befolyásolják. Ezen a területen a *választendencia* („*response tendency*”) fogalma helyett mindinkább a *válaszstílus* („*response style*”)

terminus (Jackson és Messick, 1958) vált elterjedtté (De Beuckelaer, Weijters és Rutten, 2010; Van Herk, Poortinga és Verhallen, 2004), illetve emellett az ettől némileg eltérő, Cronbach (1946) által bevezetett *válaszbeállítódás* („*response set*”) fogalma (Cheung és Rensvold, 2000; Watkins és Cheung, 1995) tekinthető uralkodónak. Miközben a válaszbeállítódás elnevezést legtöbbször csak a válaszgyakoriságok területére szűkítve használjuk, valójában eredetileg Cronbach (1946, p. 476.) számos egyéb válaszjellemzőt is ebbe a körbe sorolt (pl. válaszsebesség, válasz pontosság stb.).

A két fent említett fogalom – tehát a válaszbeállítódás és a választílus – közötti különbségtétel nem minden esetben egyértelmű a szakirodalomban. A *válaszbeállítódás* terminust preferáló szerzők alapvetően az itemek formájától, tartalmától, illetve szituációs tényezőktől (externális, külső feltételektől) függő jellemzőként tekintenek erre, míg a *választílus* fogalmát előnyben részesítők pedig egy olyan általánosan (is) érvényesülő, az adott személyre specifikus, belső tényezőként (internális diszpozícióként) kezelik, amely nem (vagy nem elsődlegesen) az itemek jellemzőinek (formájának, tartalmának) függvényében alakul. Ez utóbbi álláspont legkorábbi képviselői (Lentz, 1938; Rorer, 1965) között maga Cronbach (1950) is leginkább a helyzet vagy item kétértelműségét, strukturálatlanságát tekinti a hatás kiváltójának, de már nála is megfogalmazódik, hogy esetleg egy mélyebb személyiségvonás kifejeződéseként is lehetne értelmezni a válaszmegoszlás egyéni alakulását, ami azután a fenti helyzeti tényezők hatására válik a viselkedésben markánssá (pl. amikor a személy bizonytalan a válaszaiban az adott itemek nehézsége vagy kétértelműsége miatt).

A két fogalom nem pusztán a definíciós eltérések miatt érdemel figyelmet, hanem az ebben megnyilvánuló két kutatói hozzáállás szétválasztása miatt is. Egyfelől a *válaszbeállítódás* terminus egy olyan *externális* tényezők által meghatározott jellemzőnek tekinthető, amit a tesztelés során egyfajta minimalizálható, kiiktatandó jelenségként kell kezelni. Ehhez képest viszont a *választílus* egy *internális*, személyen belüli, inherens tulajdonság megnyilvánulásáról szól, amelynek kontroll alatt tartására, befolyásolására nincs lehetőség, vagyis itt csupán a detektálásra, számszerűsítésre lehet törekedni. Ennek ellenére azt tapasztalhatjuk, hogy sajnálatos módon a választílus és a válaszbeállítódás közötti distinkció a legtöbb kutatás esetében nem jelenik meg tisztán, a két fogalmat napjaink szakirodalmában általában szinonimaként kezelik.

A válaszgyakorisági eltérések vizsgálatához az indikátorok igen széles spektruma áll rendelkezésre, számos, jól elkülöníthető mérőszáma van jelen a kutatásokban (Baumgartner és Steenkamp, 2001). A két legtöbbször vizsgált mutató a széli, vagy extrém választílus („*extreme response style*” – ERS), illetve a beleegyező választílus („*agreement response style*” – ARS). A széli válaszpóciókra irányuló preferencia egy olyan konzisztensen megjelenő és időben is stabil (Bachman és O’Malley, 1984; Greenleaf, 1992; Naemi, Beal és Payne, 2009) jellemzőnek tekinthető, mely tartalmilag jól leválasztható az egyébként szintén stabil tesztkitöltési jellegzetességként felmerülő egyetértési hajlandóságtól, vagy beleegyező választílustól, illetve a szociális kívánatosság (szociális megfelelési igény, társas kívánatosság) fogalmától.

Ezekhez képest az MRS („*midpoint response style*” – MidRS) egy lényegesen ritkábban vizsgált mutató, részben azért, mert

ez a sok egyéb válaszlehetőség mellett csupán a középső választást detektálja (egy ötfokú Likert-skála esetén például csak a 3-as választ). A középponti választás a válaszkála-opciók jelentéstartalma függvényében sajátos információval szolgálhat. Egyfelől jelölhet egy mérsékelt, átlagos szintet a vizsgált item vonatkozásában, de értelmezhető úgy is, mint a két ellentétes végpont közötti felezőpontot leképező válaszopció, mely ilyen értelemben egy *semleges választási lehetőséget* („*neutral response option*”) jelenít meg (Weijters, Geuens és Schillewaert, 2010). Amennyiben a válaszok sem az MRS (MidRS), sem az ERS révén nem írhatók le (pl. ha egy ötfokú Likert-skála esetén csak a 2-es, illetve a 4-es választások gyakorisága a kiemelkedő), akkor kiegészítő mutatóként a *mérsékelt választási stílus* („*moderate response style*”, ModRS) tehető mérhetővé.

A válaszgyakorisági mutatók stabil, különböző helyzetekben és eltérő itemkészletek, illetve válaszkálák mellett is tetten érhetők. Az *egyéni* előfordulási valószínűség területén longitudinális vizsgálatok alapján (ugyanazon személyek felméréseivel) mind az ARS, mind pedig az ERS esetén igen magas szintű idői stabilitást lehet kimutatni (Bachman és O'Malley, 1984), még akár többéves időkülönbséggel is (Billiet és Davidov, 2008). Figyelemre méltó, hogy lényegében a személyiségjellemzők stabilitásához hasonló szintű „teszt-reteszt” korrelációk adódnak az extrém választendencia vonatkozásában ($r = 0,8$), de még az ARS terén is közepesen erős ($r = 0,5$) az együttjárás (Littvay, 2010). Mindezek alapján joggal feltételezhető, hogy ezen választendencia-jellemzők mögött valamilyen stabil, látens egyéni pszichológiai tulajdon-

ságok húzódnak meg (Austin, Deary és Egan, 2006).

Nagymintás felmérések alapján kimutatható, hogy a választendencia terén a válaszadók kb. 25–30%-a jellemezhető széli választásokkal (Austin et al., 2006; Eid és Rauber, 2000). Vagyis úgy tűnik, hogy a legkülönbélebb kérdőíves tesztelési helyzetekben meglehetősen nagy gyakorisággal megjelenik *populációs* szinten ez a választási stílus. A *kulturális, nyelvi* (nemzetek, országok közötti) aspektus szerint vizsgálódva szembeötlő, hogy a széli választások terén magasabb gyakoriság mutatható ki a nyugati országokban, míg a beleegyező választási stílus a kelet-ázsiai országokban preferált (Gilman et al., 2008; Hamamura, Heine és Paulhus, 2008). Ezek a sajátosságok arra a módszertani buktatóra is felhívják a figyelmet, hogy a kultúrközi összehasonlításokat végző kutatások, vagy akár az akkulturációs aspektusból, szubkultúrák összevetése mentén zajló vizsgálatok szignifikáns eredményeiben a választendenciák kontrollálásának hiányaként akár műtermékek megjelenése is előfordulhat. Különböző szimulációs és valós adatokon végzett elemzések szerint egy statisztikailag szignifikánsnak tűnő skálák közötti gyenge pozitív együttjárás könnyen megváltozhat, sőt ellenkező előjelűvé is átfordulhat, ha nem pusztán a skálák tartalmi alapon képzett pontszámát vesszük számításba, hanem az egyéni válaszkála-torzításokat is beemeljük a modellbe (Johnson, 2003). Pszichometriai vonatkozásban konkrét kutatási eredmények támasztják alá, hogy a választási stílus jelentősen befolyásolhatja a tesztadatok megbízhatóságát, illetve érvényességét. Így például az extrém válaszok preferálása (ERS) egyfelől növelheti a skála reliabilitását, míg másik

oldalról csökkentheti a validitást (Clarke, 2000). De lényegében bármilyen tendenciózus válaszmegoszlási eltolódás, halmozódás (pl. ARS, ModRS) magával hoz egy bizonyos mértékű „tartalomfüggetlen” megosztott varianciát, ami pedig végső soron a skálák belső konzisztenciáját veszélyeztetheti.

HIPOTÉZIS

A válaszbeállítódás, illetve a válaszstílus tehát részben egy egyénre vonatkozó információnak tekinthető, részben pedig a tesztek kialakítása, fejlesztése terén felhasználható jellemzőként kezelhető. Miközben a hagyományos hazai önkítöltős tesztelési eljárások során kizárólag a tartalomra vonatkozó válaszokra építünk mind az egyéni eredmények kiszámításában, mind pedig a skála pszichometriai jellemzőinek, például belső konzisztenciájának meghatározásában, addig a válaszgyakorisági jellemzők számottevő hatását figyelmen kívül hagyva lényeges mérési torzításokat engedhetünk be a tesztelési folyamatba.

A jelen kutatás hipotézise szerint a hazai személyiségmérésekkel kapott teszteredményekben a válaszszelesőségesség stabil, konzisztens egyéni sajátágként jelenik meg. Ennek következtében a válaszszelesőségesség, mint tartalomtól függetlenül érvényesülő konzisztens hibaforrás, a ténylegesen mérendő személyiségdimenzió mutatott pontszámokat esetenként jelentősen torzíthatja. Ennek kiküszöbölésével, korrekciójával a személyiségtesztekben megjelenő manifest teszteredmények végső soron a látens konstruktumok valódi értékéhez közelebbi eredményekkel szolgálhatnak.

MÓDSZER

Mérőeszközök

A vizsgálatban több mintán három különböző önjellemzős személyiségmérő eljárást alkalmaztunk, két Big Five alapú (Facet5, NEO) kérdőív mellett a Temperamentum és Karakter Kérdőív (TCI) hazai változatát töltöttük ki a résztvevőkkel. Formailag mindhárom kérdőív ötfokú válaszkálájú, felvételük pedig internetes felületen történt.

Facet5

A Facet5 (Buckley és Williams, 2002) által mért dimenziók elnevezése némileg eltér a hagyományostól, de tartalmilag jól illeszkedik a hazai személyiségtaxonómiai elemzések eredményeihez (Nagybányai Nagy, 2013). Az *extraverzió* mérésére a *lendület (Energy)* elnevezésű skála, a *lelkiismeretesség* mérésére pedig a *kontroll (Control)* skála szolgál. Az *akarát (Will)* skála az eredeti *kellemesség* faktor alacsony végpontját célozza meg, míg az *emocionalitás (Emotionality)* skála pedig a Big Five terminológia szerinti *érzelmi stabilitás* dimenzióinak az ellentétes oldalát képezi le. Az ötödik, *empátia (Affection)* skála tartalma, elnevezése nem az angolszász faktorstruktúrában megszokott *nyitottság* faktor megfelelője, ellenben tartalmilag jól összhangba hozható a Szirmák-féle magyar vizsgálat (1994) *integritás* faktora által kijelölt fogalmi tengellyel, melyet markerszavai alapján hosszabban „pozitív viszonyulás másokhoz” megjelöléssel is illethetnénk. A tesztben a skálaértékek meghatározása – a magyar standard minta alapján – egy tízfokú Sten-skálán („standard-ten”) 1–10 közötti pontozással történik.

Magában a Facet5 kérdőívben a két ellentétes végpontként megfogalmazott állítás közötti választás egy ötfokú válaszkálán történik, így az 1-es és az 5-ös válaszadás jelenti a széli választások preferálását, míg a 2-es és a 4-es válaszok pedig a mérsékelt válaszopciót jelölik. Mivel a válaszkála mindkét végpontjára adott választások beleértendők a válaszpreferencia-mutatókba – akár széli választásról van szó, akár mérsékelt választásról – ezért az egyes skálákon belüli fordított itemek (melyek egyébként közel kiegyensúlyozottan vannak jelen a pozitív tételekkel) nem befolyásolják ezen válaszgyakoriságok kiszámítási módját.

A Facet5 esetén a középválasztás, vagyis a 3-as válaszok megjelenését a válaszszélsőségesség szempontjából nem tekinthetjük releváns információnak, hiszen itt tulajdonképpen a két ellentétes oldal közötti választás hiányát tapasztaljuk, ami ez esetben jelentését tekintve a döntés elkerülésével ekvivalens (s mint ilyen, tulajdonképpen egyfajta „nem választást” takar). Az ilyen típusú, tartalmilag kevésbé informatív válaszok csökkentése érdekében maga a kérdőív instrukciója is arra figyelmezteti a vizsgálati személyt, hogy a semmitmondó hármast számot igyekezzen kerülni.

NEO

Az egyik legszélesebb körben elterjedt ötfaktoros önkítöltős kérdőívként a NEO, illetve ennek továbbfejlesztett változatai, a NEO PI-R (Costa és McCrae, 1992), a NEO PI-3 (McCrae, Costa Jr és Martin, 2005) tűnhet a legalkalmasabbnak annak vizsgálatára, hogy ugyanúgy kimutatható-e a válaszszélsőségesség stabil megjelenése, mint a Facet5 esetében. A NEO PI-R 240 tétel tartalmazó (Szirmák és Nagy, 2002) személyiség-kérdőív újabb, továbbfejlesztett

változatában, a NEO PI-3 tesztben mindössze 30 tétel szövegének kisebb módosítását hajtották végre a szerzők a jobb érthetőség és a magasabb belső konzisztencia érdekében (részletes felsorolásukat ld. McCrae et al., 2005, pp. 269–270.). Vizsgálatunkban a NEO PI-3 jelenleg is folyó hazai adaptálása során használt tesztváltozatot alkalmaztuk, ahol a kérdőívre adott válaszadás szintén egy ötfokú Likert-skálán történik.

TCI

A válaszszélsőségességi eredmények általánosíthatóságának ellenőrzésére egyéb, nem Big Five alapú személyiségtesztben nyújtott jellemzőkkel is összevetettük a választendenciák alakulását. A felmérésbe bevont TCI az ötfaktoros személyiségmégközelítés egyik alternatívájának tekinthető (Cloninger, Svrakic és Przybeck, 1993). Jelen esetben a 240 tétel, de ötfokú válaszkálát használó hazai változatot (Rózsa, Kállai, Osváth és Bánki, 2005) alkalmaztuk.

A statisztikai elemzések elvégzése IBM SPSS Statistics 19, valamint RopStat szoftverek segítségével történt. A statisztikai próbák előfeltételeinek ellenőrzése minden esetben megtörtént, de hely hiányában ezek egyenkénti bemutatásától eltekintünk.

Minta

A Facet5 tesztkitöltésben részt vevő teljes minta több mint ezerfős ($N = 1026$), melynek nemi megoszlása közel kiegyensúlyozott ($N_{\text{fi}} = 534$; $N_{\text{nő}} = 492$), míg az életkori adatok (18–65 év közöttiek; $M = 32,2$; $SD = 9,2$) heterogénnek tekinthetők. A kutatás ökológiai validitásának növelése érdekében a résztvevők szinte kivétel nélkül valamilyen valós szervezeti HR-alkalmazás kapcsán töltötték ki a személyiség-kérdőívet a legkülönbözőbb munkakörökből, de jellemzően banki, pénz-

ügyi, HR, IT, értékesítési, műszaki területekről. Az online kitöltésre minden esetben csak személyre szóló belépési kód segítségével volt lehetőség.

A NEO, illetve a TCI tesztek eredményeit egy kisebb (N = 137), pszichológia szakos egyetemista vizsgálati minta szolgáltatta, ahol emellett a Facet5 kérdőív felvétele szintén megtörtént. Itt az egyes elemzésekbe bevont résztvevők száma némileg eltérő volt az értékelhető tesztadatok függvényében (a különféle páronkénti korrelációs összehasonlításoknál), de minden esetben meghaladta a száz főt. A minta nemi összetétele női túlsúllyal jellemezhető (70%), életkori átlaga 20,5 év (*SD* = 2,2 év). Ebben a mintában a tesztek kitöltésére egymást követően egy-egy hetes időkülönbséggel került sor.

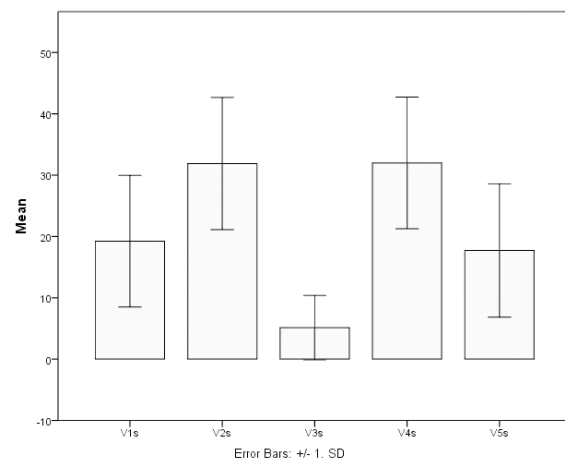
EREDMÉNYEK

Válaszfokokozatok szerveződése a Facet5 tesztben

A rendelkezésre álló széles mintát (N = 1026) elemezve az egyes válaszfokokozatok előfordulási gyakoriságában a mérsékelt választások relatív túlsúlya figyelhető meg, szemben a széli választások alacsonyabb szintjével (1. ábra). Ugyanakkor mindkettőtől lényeges mértékben eltér a középválasztások, vagyis a 3-as válaszok igen alacsony előfordulási gyakorisága, ami a teljes mintában mutatott előfordulásban a legritkább a többi válaszpációhoz viszonyítva.

A Facet5 kérdőívet az egymást követő 10-10 item mentén felbontva is érdemes megvizsgálni, hogy a kérdőív lefutásában miként alakul a válaszpációk előfordulási gyakorisága. A grafikonokon (2. ábra) látható, hogy 10-es itemblokkokban (tehát tar-

talmilag vegyes tételcsoportokon) elemezve a mérsékelt, illetve a széli választások átlagos előfordulása és kumulatív gyakorisága páronként együtt, de jól érzékelhetően a többitől eltérő meredekségben fut. A középválasztások ezzel szemben elkülönült lefutást mutatnak, a 10-es itemblokkokban ez marad a legritkább a többi válaszpációhoz viszonyítva.

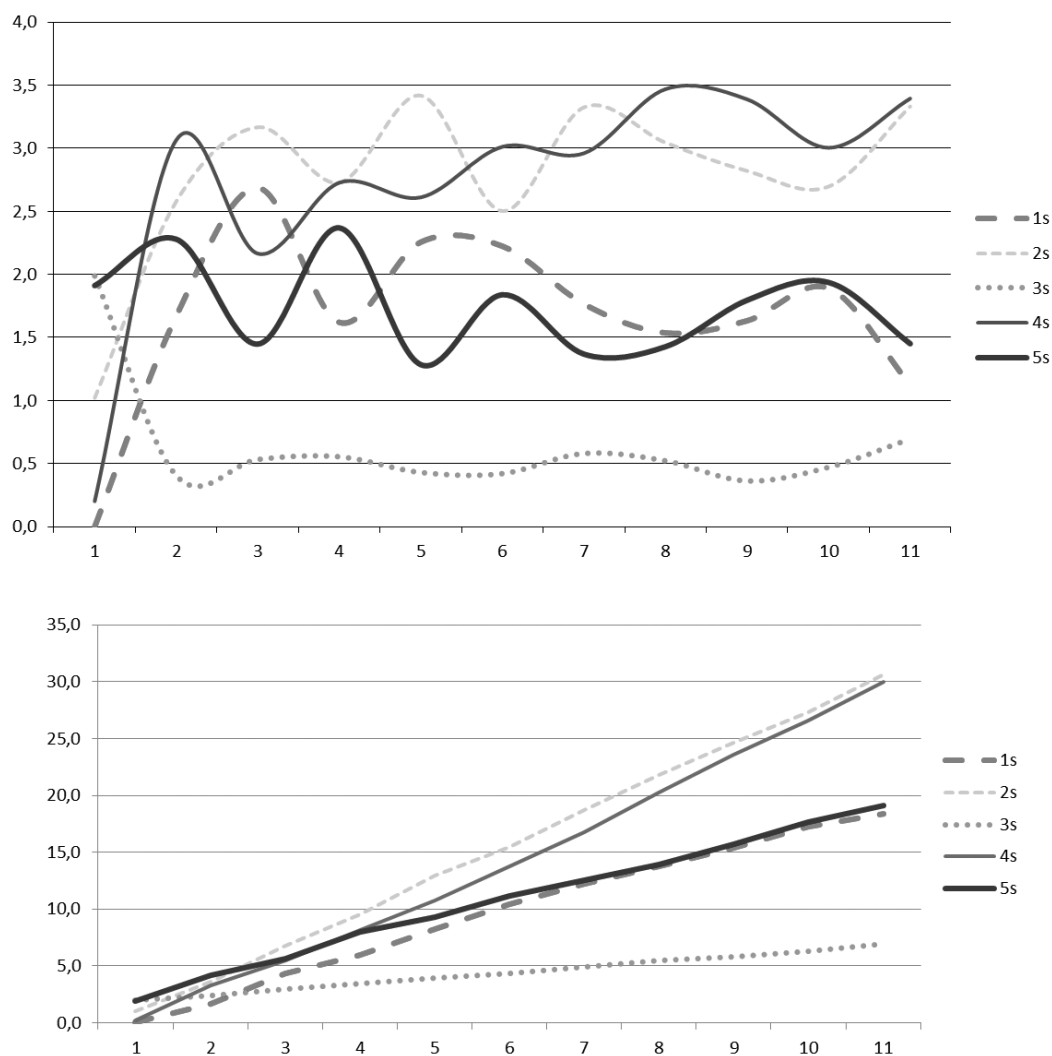


1. ábra. A Facet5 kérdőíven (106 itemen) egyénileg összegzett válaszpációk (1–5 fokozat) előfordulásának átlaga és szóródása

A széli, illetve mérsékelt választások korrelációs elemzése szintén – az előfordulásokban is tapasztalt – páronkénti összekapcsolódásról tanúskodik (1. táblázat). Ezen válaszpárok nem csupán hasonló gyakorisággal bírnak, hanem az egyéni válaszadásokban páronként erős együttjárást is mutatnak, miközben egymással pedig komplementer viszonyban állnak.

1. táblázat. A széli, illetve mérsékelt válaszfokokozatok gyakoriságának együttjárásai

Pearson-korreláció	V5	V4
V1	0,843	-0,849
V2	-0,845	0,715
(p < 0,01)		



2. ábra. Az öt válaszfokozat átlagos gyakoriságának (fent), illetve kumulatív gyakoriságának (lent) alakulása 11 blokkra felbontva a teljes kérdőívet (az egymást követő 10-10 item mentén)

A válaszkategóriák előfordulása ugyanakkor nem tekinthető függetlennek egymástól, hiszen bármelyik válaszopció megjelenése automatikusan lecsökkenti minden más válaszlehetőség jelenlétét. Egy ugyanilyen terjedelmű (106 tételes) kérdőívnek egy ugyanilyen méretű (1026 esetet tartalmazó) mintára generált véletlenszerű tesztválaszai esetén szintén adódnak gyenge negatív együttjárások minden válaszfokozat előfordulási gyakorisága között ($r = -0,26$; $p < 0,001$; $r = -0,2$; $p < 0,001$). Vagyis minden ilyen típusú Pearson-féle korrelációs értéket ehhez az „alapszinthez” képest célszerű viszonyítani:

a korrelációban bekövetkező kb. 0,1-es változás ettől a gyenge negatív kapcsolattól már szignifikánsan különböző kapcsolatnak tekinthető, vagyis lényegében minden pozitív irányú együttjárást szignifikánsan eltérőnek lehet tekinteni. Esetünkben a Facet5 kérdőívben az 1-es és az 5-ös válaszok esetén, valamint a 2-es és a 4-es válaszfokozatok esetén tapasztalható magas pozitív együttjárások erősen szignifikánsan különböznek a fenti random adatokon tapasztalt negatív korrelációs eredményektől ($Z = 33,01$, $p < 0,001$; $Z = -24,41$; $p < 0,001$). Főkomponens-elemzést végezve mind az öt válaszfokozat egyéni

mennyiségén kétfaktoros megoldás születik, amiből az elsőre 0,9 feletti töltéssel terhelnek az 1-es és az 5-ös válaszok pozitív előjellel, míg ugyanilyen erősséggel, de negatív előjellel szerepelnek rajta a 2-es és a 4-es válaszopciók. Ez a faktor a varianciák 70%-át fedile, a második pedig – melyre csak a 3-as válaszok kerülnek – 21%-os magyarázott hányaddal bír. Ehhez képest a fenti módon generált random adatokon elvégzett hasonló elemzésben azt tapasztalhatjuk, hogy minden válaszopció külön-külön faktorra rendeződik, vagyis ez esetben nem mutatható ki ugyanilyen egységes válaszszélsőségességi szerveződés.

Mindezek alapján úgy tűnik, hogy indokolt a széli választások, illetve a mérsékelt választások gyakoriságának egymáshoz viszonyított fölényét egy közös értékben, a *válaszszélsőségesség-mutatóban* megragadni. Ez mind a teljes kérdőívre, mind pedig mérési területenként az egyes személyiségdimenziókra kiszámítható. A további korrelációs elemzésekből az is kitűnik, hogy a személyiségskálánkénti válaszszélsőségesség-mutatók sem tekinthetők függetlennek egymástól, mindegyik erős, szignifikáns lineáris kapcsolatban ($r = 0,69-0,74; p < 0,01$) áll a többivel.

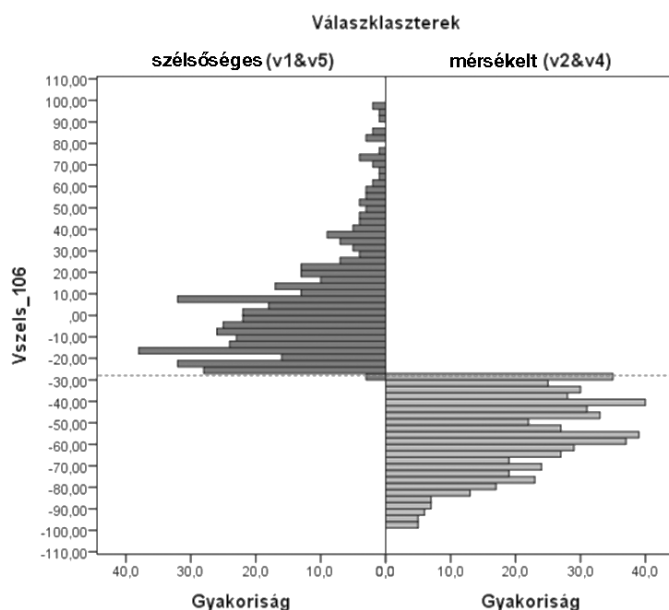
Módszertani szempontból lényeges különbséget tenni a személyiségskála extrémítása és a válaszkálán mutatott extrémítás között. Fontos leszögezni, hogy az itemekre adott válaszok szélsőségessége önmagában nem feltétlenül kell, hogy a skálapontszámok extrém magas vagy alacsony értékét eredményezze. Mivel egy konkrét személyiségskála itemeinél a válaszkála szélsőséges használata egyaránt eredményez szélsőséges skálapontszámot és magas válaszsélsőségesség-indexet, ezért a személyiségdimenzió tételeire adott *válaszsélsőségesség* he-

lyett esetenként a teljes kérdőívben mutatott *„total”-szélsőségességet*, illetve máskor pedig a vizsgált skála itemei nélkül mutatott *„maradék”-szélsőségességi* mutatót célszerű alkalmazni (az item-totál, illetve item-maradék terminusok analógiájára), hogy ezzel a válaszadás tartalmi, illetve formai jellege szétválasztható maradjon.

A Facet5 adatbázison lefuttatott öt regresszióelemzés révén szignifikáns kvadraticus kapcsolatok mutathatók ki egyesével mind az öt maradék-válaszsélsőség változó, valamint a hozzá tartozó személyiségdimenzió nyerspontszáma között ($R = 0,26-0,29; p < 0,01$), amelyek kb. 7–9%-os torzítatlan becslésű megmagyarázott hányadot képviselnek (*adjusted R²*). Összességében ez azt jelenti, hogy a személyiségvonások pontszáma és az összes többi (nem az adott skálához tartozó, tehát maradék) item válaszsélsőségessége között is felfedezhető kapcsolat – függetlenül a ténylegesen mérni kívánt személyiségvonásra (tehát a konkrét dimenzióhoz tartozó itemekre) adott választóktól.

Válaszsélsőségességi klaszterek a Facet5 tesztben

A válaszsélsőségességi mutatók közötti kapcsolatok elemzésén túl az egyének közötti variabilitás megragadására célszerűnek tűnik bizonyos válaszsélsőségességi mintázatot mutató alcsoportok elkülönítésére törekedni. Amennyiben a skálahasználat dimenzionális megközelítésén túllépve típusok elkülönítésére fókuszálunk, akkor érdemes megpróbálni különféle válaszklaszterek feltárását megcélózni. A jelenlegi vizsgálati mintát felhasználva ezt a fajta kategorizálást a *„Two Step Cluster”* módszer alkalmazásával megkísérelve az eredmények alapján úgy tűnik, hogy a széli és mérsékelt választások teljes

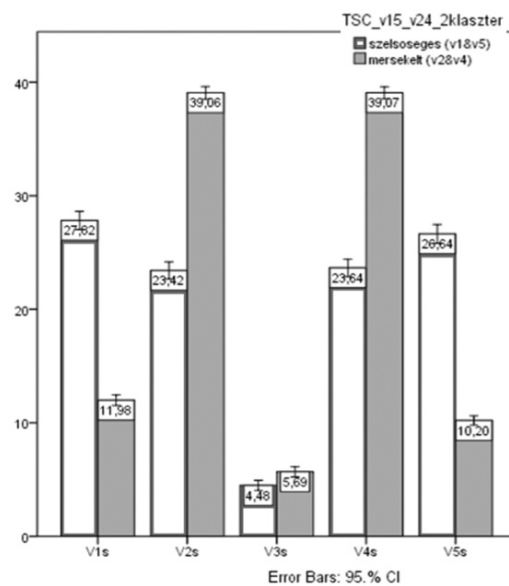


3. ábra. A két válaszklaszter a total-válaszszélsőségesség-értékek megoszlása alapján

kérdőív során mutatott előfordulása (total-szélsőségesség) alapján egymástól két jól elkülöníthető válaszszélsőségesség-klaszter adódik. A klaszterelválas, tehát az interklaszter távolságok, valamint a klaszterkoherencia, vagyis az intraklaszter távolságok viszonyát leíró sziluett-együttható értéke = 0,7, ami kifejezetten jónak mondható. Diszkriminanciaanalízis alkalmazásakor (*Wilks-féle lambda* = 0,385; $kh^2 = 976,53$; $p < 0,001$) az ötféle válaszfokozat változója alapján a klaszterekbe történő besorolás a mérsékelt választílusú csoport esetén 100%-os, míg a szélsőséges választílusú klaszterbe 93%-os. A klaszterbe tartozás legegyszerűbben tulajdonképpen a teljes kérdőív válaszszélsőségesség-mutatója mentén definiálható, ahol nem az elméleti középértéket jelölő nulla pont, hanem egy enyhén mérsékelt választási túlsúly képezi a vágópontot (-28-as érték) a +/-106 intervallumon belül (3. ábra).

Az egyes válaszopciók gyakorisági átlaga jól láthatóan különbözik a két válaszklaszter között (4. ábra). Mivel az előfordu-

lasi gyakoriság terén a mérsékelt választások aránya összességében magasabb, így a két klaszter közül a szélsőséges választási preferenciával rendelkezők kisebb csoportot alkotnak ($N = 470$; 45,8%), a mérsékelt válaszklaszterbe tartozók pedig nagyobbat ($N = 556$; 54,2%).



4. ábra. A két válaszklaszter közötti eltérés a válaszopciók gyakorisági átlaga alapján

2. táblázat. A személyiségskálák válaszklaszterenkénti alapstatisztikái

Eredeti Sten-pontszámok	Akarat	Lendület	Empátia	Kontroll	Emoc.
Szélsőséges válaszklaszter (N = 470)					
Átlag	5,69	5,96	5,8	5,72	5,31
Szórás	2,28	2,18	2,17	2,16	2,32
Minimum	1	1	1	1	1
Maximum	10	10	10	10	10
Mérsékelt válaszklaszter (N = 556)					
Átlag	5,38	5,27	5,28	5,15	5,66
Szórás	1,63	1,69	1,61	1,69	1,64
Minimum	1	1	1	1	1,4
Maximum	10	9,4	9,2	9	10

A Facet5 teszteredményeket vizsgálva feltűnő, hogy a mérsékelt klaszterben az elméletileg kijelölt standard tízfokú (Sten) skálapont-tartományokhoz képest az egyes személyiségdimenziókon mért maximumok és minimumok fél–egy ponttal szűkebb övezeteket jelölnek ki (2. táblázat). Vagyis a mérsékelt válaszklaszter tagjainál bizonyos pontszámok egyáltalán nem is „tudnak” megjelenni, egész egyszerűen kiesnek a mérési tartományból, miközben a szélsőséges klaszterbe tartozók eloszlása megfelelően „kihasználja” a skálák teljes terjedelmét (1–10). Az eredeti Sten-pontszámok válaszklaszterenkénti átlaga minden személyiségskálában 1%-os szinten szignifikánsan különbözik egymástól. Valamint a Sten-skáláknál elméletileg elvárt alapstatisztikai jellemzőik is több tekintetben eltérnek az 1–10 pontos terjedelemtől, illetve az 5,5 pontos átlagtól, valamint a 2 pontos szórástól.

Az eredmények alapján úgy tűnik, hogy a Facet5 tesztben a választendencia egyénre jellemző, konzisztens jelenléte egyfajta glo-

bális ráarakódó hatásként befolyásolja a teszteredmények alakulását. Ennek kiküszöbölésére válaszklaszterenkénti standardokat célszerű kialakítani, a választendenciából fakadó általános pontszámeltolódások ellensúlyozására, vagyis a választorzítási hatástól megtisztított ún. korrigált Sten-pontszámok létrehozására. A Facet5 személyiségdimenziók korrigált Sten-pontszámai a saját (mérsékelt vagy szélsőséges) válaszklaszter szerinti standard normák alapján számíthatók (szemben a teljes mintát felölelő eredeti normára épített standardizált pontokkal).

A pontszámkorrekciók mértéke (y tengely) az eredeti Sten-pontszámok (x tengely) függvényében kb. –1 pont és +1 pont között változik mind az öt személyiségdimenzió mentén (5. ábra). Számszerűleg nem tűnnek nagynak ezek a pontszámeltolódások, de fontos hangsúlyozni, hogy éppen a Sten-skála extrémebb pontszámövezeteiben a legjelentősebbek, vagyis abban a skálatartományban, amely a legérzékenyebb a kis eltérésekre.

Például előfordulhat, hogy a korrekció révén az 1%-os előfordulási gyakoriság helyett a 8%-nak megfelelő populációs övezetbe tarthat ugyanaz a személy. Ilyen értelemben tehát a pontszámtorzulások lényeges értelmezésbeli eltérésekhez vezethetnek mind egyéni szinten, mind pedig a csoportos eredmények interpretációja esetén. Bár azt tapasztalhatjuk, hogy az extrém tartományok felől a középső pontszám tartományok irányába egyre csökken a korrekció mértéke, ugyanakkor még ezek sem tekinthetők elhanyagolhatóknak, hiszen például az 5,5 pontos elméleti középértéktől történő fél pontos elmozdulás nagyjából 9–10%-os populációs gyakoriságbeli különbségnek feleltethető meg.

Az összesített eredmények szerint ezek a Facet5 pontszámkorrekciók meglepően nagy valószínűséggel fordulnak elő a mintában. A szélsőséges választípus esetén a vizsgálati személyek közel fele (45%) érintett legalább minimális, fél Sten-pontos pontszám torzulásokban, míg a mérsékelt választípusúaknak nagyjából egyharmadánál (34%) fordul elő ilyen vagy ezt meghaladó szintű pontszámeltolódás. Az öt mérési dimenzióban összesített Sten-korrekciók egy-pontos értéket meghaladóan a teljes minta 40%-ában kimutathatók.

Ehhez képest a korrigált Sten-pontszámok esetén már egyik személyiségskálán sem jelentkeznek szignifikáns átlagok közötti különbségek a két válaszklaszter között.

Válaszfokozatok szerveződése más személyiségtesztekben

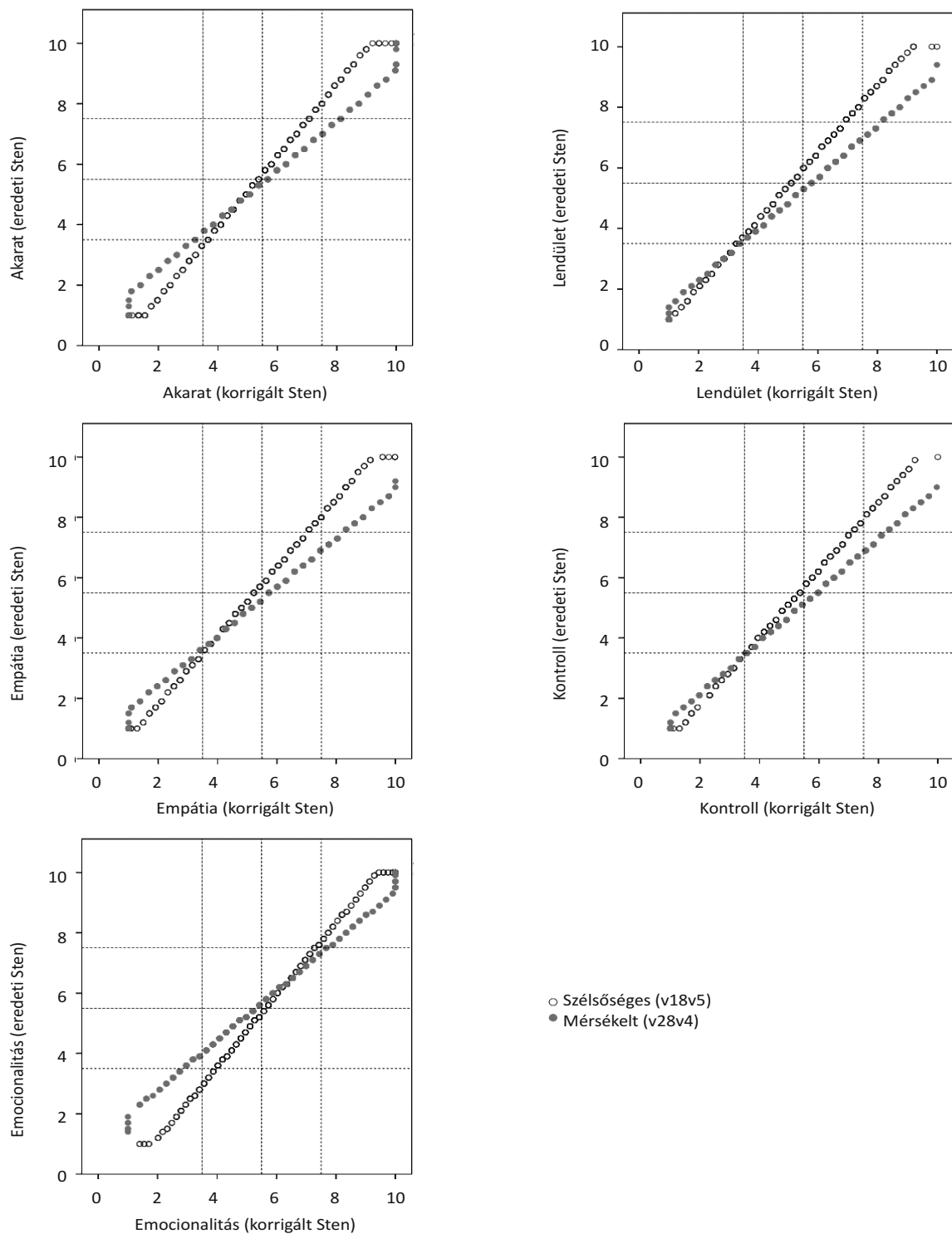
A következőkben a Facet5 teszt kitöltésben megfigyelhető válaszszelesőségességet összevetettük a másik két személyiség-kérdőív hasonló mutatóival egy olyan mintán, ahol ugyanazon vizsgálati személyek mindhárom

teszt kitöltését elvégezték ($N = 137$). A NEO PI-3 itemekre adott válaszok alapján kiszámított, valamint a Facet5 tesztben összesített válaszszelesőségesség-mutatók között egy közepesen erős és szignifikáns ($r = 0,65$; $p < 0,01$) együttjárást lehetett kimutatni a teljes kérdőív mért (totál-) válaszszelesőségesség tekintetében (3. táblázat). (Amennyiben egy random teszt kitöltésből származó mintán végeznénk ezt a kapcsolatvizsgálatot, úgy nem adódhatna sem az egyes válaszfokozatok, sem a válaszszelesőségesség-mutatók vonatkozásában szignifikáns kapcsolat, tehát itt a teljes függetlenség tekinthető viszonyítási „alapszintnek”).

3. táblázat. A Facet5, TCI és NEO kérdőívben kapott válaszszelesőségesség-indexek közötti korrelációs együtthatók

<i>Pearson/ Spearman korreláció</i>	Facet5 (vszels)	TCI (vszels)	NEO (vszels)
Facet5 (vszels)	1,000	0,550	0,640
TCI (vszels)	0,547	1,000	0,603
NEO (vszels)	0,649	0,643	1,000
** ($p < 0,01$)			

Az itt kapott eredménynek az értelmezése során érdemes figyelembe venni azt is, hogy a két teszt válaszkálája némileg eltérő struktúrájú. Mindkét mérőeszköz ötfokú válaszkálát használ, de a hagyományos, Likert-típusú formátumhoz képest a Facet5 a két, kvázi el- lentétes tartalmú állítás-pár közötti választásban a középső, 3-as válaszok kerülésére ösztönöz, már a teszt instrukciójából fakadóan is. Ennek következtében az egyes válaszfokozatok előfordulási valószínűsége sem azonos, így a NEO PI-3 esetén a 3-as válaszok aránya nem is tér el olyan kiugró módon a többi válaszpóció előfordulási gyakorisá-



5. ábra. A Facet5 személyiségskálák eredeti Sten-pontjai (x tengely), valamint azok két válaszklaszter szerint korrigált Sten-értékei (y tengely)

gától. Hasonló módon a TCI és a Facet5 válaszszélsőségesség-mutatók közötti együttjárás is szignifikánsnak bizonyult ($r = 0,55$;

$p < 0,01$), mégpedig nagyjából azonos mértékben, mint azt a NEO PI-3 esetén láthattuk (3. táblázat).

A középválasztás gyakorisága tekintetében lényeges kiegészítő eredmény, hogy az azonos felépítésű Likert-skálával operáló személyiségtesztekénél hasonlóan alakul ezek egyéni gyakorisága. Mivel a Facet5 esetén lényegesen különböző jelentéstartalommal bírnak a 3-as válaszok, így csak a NEO-PI-3, valamint a TCI személyiségteszt 3-as választási gyakorisági értéke vethető össze: a páronkénti korrelációs összehasonlítások eredménye szerint ezek viszont egymással közepesen erős ($r = 0,67$; $p < 0,01$) együttjárást mutatnak (4. táblázat).

4. táblázat. A Facet5, a TCI és a NEO kérdőív esetén megjelenő középválasztások előfordulásának együttjárásai

Pearson/ Spearman korreláció	Facet5 (v3)	TCI (v3)	NEO (v3)
Facet5 (v3)	1,000	0,040	0,071
TCI (v3)	0,026	1,000	0,648**
NEO (v3)	0,098	0,674**	1,000
** ($p < 0,01$)			

Ez utóbbi eredmény megerősíti a különféle választendencia-mutatók konzisztens – mind a mérési terület között, mind pedig időben stabil –, személyre jellemző megjelenésének tényét. Ugyanakkor rávilágít a választendencia mérésének azon kulcsfontosságú módszertani aspektusára, hogy a válaszgyakorisági elemzésekben az összehasonlítandó kérdőíveknek alapvetően ekvivalens válaszformákkal, illetve ami ugyanilyen fontos, azonos válaszfokozatokkal és azonos jelentéstartalommal célszerű rendelkezniük.

MEGVITATÁS

A jelen vizsgálatban kapott eredmények alapján úgy tűnik, hogy a választendencia egyénekre jellemző, de a válaszadás tartalmától részben függetlenül is érvényesülő jellegzettségként számottevő hatással bírhat a kérdőív minden egyes tételére adott válaszban. Ezek az eredmények jelentős gyakorlati következménnyel járhatnak a kérdőíves mérések hatékonyságát illetően. Nem arról van szó, hogy az egyes egyének azért mutatnak gyakrabban szélsőséges válaszadást a kérdőívek kitöltésekor, mert bizonyos személyiségskálákon szélsőséges pozíciót foglalnak el, hanem itt egy rejtett, másodlagos hatás érvényesül. A probléma abban áll, hogy ez a válaszpreferencia más, akár tartalmilag teljesen független mérési területeken és kérdőívekben mutatott válaszszélsőségességre is kihat, és így módon bármilyen személyiségdimenzió elért pontszámot képes lehet torzítani. Vagyis ez az egyénre jellemző választendencia pusztán önmagában egyfajta globális torzító hatásként a különféle mérési területekre tartalomtól függetlenül „rárakódva” meghatározza a válaszok szélsőségességének mértékét, majd pedig ezen keresztül maguknak a személyiségskála-pontszámoknak az alakulását is, és összességében végső soron a látens személyiségjellemzők mérésének hatékonyságát befolyásolja. Márpedig ha ez a torzítási effektus általánosnak, minden mérési területre kiterjedő hatásnak tekinthető, akkor olyan teszteredmények megjelenésével is számolni kell, melyekben nem valós tulajdonságok manifesztálódnak (1. fajú hiba), illetve másik oldalról pedig esetenként ténylegesen létező látens tulajdonságok maradhatnak rejtve (2. fajú hiba) a választendencia befolyásoló hatása miatt.

A jelen vizsgálat legfontosabb eredményei a következőkben összegezhetők:

1. Az önjellemzős tesztkitöltésekben hazai mintán is kimutatható egy tartalomtól többé-kevésbé függetlenül érvényesülő válaszszélsőségességi preferencia, melynek egyénenkénti jellege a különféle tartalmú személyiségmérési dimenziók eredményeiben is megnyilvánul. Ráadásul ez nem csak egyetlen teszt esetén jelenik meg, továbbá néhány hetes intervallumokban vizsgálva is állandónak mutatkozik, vagyis nagyrészt időben is stabilnak tekinthető.

2. A válaszszélsőségesség egyéni szintje alapján két, egymástól markánsan elkülönülő válaszadói klaszter képezhető, tehát a válaszszélsőségesség egyéni különbségei a válaszszélsőségességi klaszterekkel jól megragadhatók.

3. Mivel a válaszszélsőségesség mint tartalomtól függetlenül érvényesülő hibaforrás

a látens mérési dimenziókon mutatott pontszámokat torzíthatja, ezért ennek kiküszöbölése érdekében olyan válaszklaszterek szerinti standard normák kialakítása szükséges, melyek a korrigált manifeszt értékek tekintetében valódi összehasonlításokat tesznek lehetővé. Ezen standard normák alkalmazása révén a korrigált Sten-pontszámokban már egyik személyiségskálán sem jelentkeznek átlagok közötti különbségek a két válaszklaszter között.

Gyakorlati oldalról nézve tehát úgy tűnik, hogy a személyiségdimenziók pontszámait torzító általános válaszgyakorisági hatás ezen a módon kiküszöbölhetővé válik. Az online tesztrendszerekben pedig kifejezetten egyszerűen megvalósítható a standard normák ilyen irányú bővítése, így a tesztkitöltést követően a nyerspontszámokat a megfelelő válaszszélsőségesség szerinti normaátlagok és szórások alapján lehet korrigált standard skálapontszámokká konvertálni.

SUMMARY

DEVELOPING PSYCHOMETRIC METHODS IN ONLINE TESTING

Background and aims. Questionnaire methodology can be broadened with several new elements by online testing, as compared with traditional paper-pencil assessment procedures. Not only completion, evaluation and interpretation of the results might become significantly more efficient, but online testing enables the introduction of numerous additional psychometric and methodological innovations. Related to these, the so called response style characteristics may broaden the scope of measurement of the tests. *Methods.* Present study analyses the global score-biasing effect of response-extremity by the Facet5 online personality questionnaire, as part of a more extensive research (N = 1026). In addition, the general characteristics of this phenomenon is demonstrated based on other personality questionnaire results (NEO PI-3, TCI) on a smaller sample (N = 137), and suggestions regarding the measurement and control of response-frequency characteristics are provided as well. *Results.* The correlations of response-extremism indicators of the personality scale within one test ($r = 0.7$; $p < .01$) suggest that a thematically independent factor can influence or distort the evolution of test scores.

This can be considered as a general, global effect, as moderate correlations can be shown between the overall response extremism indices of the different personality tests ($r = 0.5$ to 0.6 ; $p < .01$). Based on the complete response pattern of the questionnaire two distinct response-extremism clusters can be detected, for which significantly different standard norms can be created to offset their overall score shifts resulting from the response tendency. *Discussion.* The results confirm that there is a consistent, individually characteristic response tendency. This individual response style appears deliberately from the measured themes and can be detected as a stable characteristic of the different questionnaires, so it influences the test scores by a global, additive effect.

Keyword: Big Five questionnaire, online testing, extreme and moderate response style

IRODALOM

- AUSTIN, E. J., DEARY, I. J., EGAN, V. (2006): Individual differences in response scale use: Mixed Rasch modelling of responses to NEO-FFI items. *Personality and Individual Differences*, 40(6), 1235–1245.
- BACHMAN, J. G., O'MALLEY, P. M. (1984): Yea-saying, nay-saying, and going to extremes: Black-white differences in response styles. *Public Opinion Quarterly*, 48(2), 491–509.
- BAUMGARTNER, H., STEENKAMP, J. B. E. M. (2001): Response styles in marketing research: A cross-national investigation. *Journal of Marketing Research*, 38(2), 143–156.
- BERG, I. A., COLLIER, J. S. (1953): Personality and group differences in extreme response sets. *Educational and psychological measurement*, 13(2), 164–169.
- BILLIET, J. B., DAVIDOV, E. (2008): Testing the stability of an acquiescence style factor behind two interrelated substantive variables in a panel design. *Sociological Methods & Research*, 36(4), 542–562.
- BUCKLEY, N., WILLIAMS, R. (2002): Testing on the web – Response patterns and image management. *Selection & Development Review*, 18, 3–8.
- CHEUNG, G. W., RENSVD, R. B. (2000): Assessing extreme and acquiescence response sets in cross-cultural research using structural equations modeling. *Journal of Cross-Cultural Psychology*, 31(2), 187–212.
- CLARKE, I. (2000): Extreme Response Style in Cross-Cultural Research: An Empirical Investigation. *Journal of Social Behavior & Personality*, 15(1), 137–152.
- CLONINGER, C. R., SVRAKIC, D. M., PRZYBECK, T. R. (1993): A psychobiological model of temperament and character. *Archives of general psychiatry*, 50(12), 975.
- COSTA, P. T., MCCRAE, R. R. (1992): Professional manual: Revised NEO personality inventory (NEO-PI-R) and NEO five-factor inventory (NEO-FFI): *Psychological Assessment Resources*, Odessa, FL.
- CRONBACH, L. J. (1946): Response sets and test validity. *Educational and Psychological Measurement*, 6(4), 475–494.
- CRONBACH, L. J. (1950): Further evidence on response sets and test design. *Educational and Psychological Measurement*, 10, 3–31.

- DE BEUCKELAER, A., WEIJTERS, B., RUTTEN, A. (2010): Using ad hoc measures for response styles: a cautionary note. *Quality and Quantity*, 44(4), 761–775.
- EID, M., RAUBER, M. (2000): Detecting measurement invariance in organizational surveys. *European Journal of Psychological Assessment*, 16(1), 20.
- GILMAN, R., HUEBNER, E. S., TIAN, L., PARK, N., O'BYRNE, J., SCHIFF, M., SVERKO, D., LANGKNECHT, H. (2008): Cross-National Adolescent Multidimensional Life Satisfaction Reports: Analyses of Mean Scores and Response Style Differences. *Journal of Youth and Adolescence*, 37(2), 142–154.
- GRAYSON, R. (2011): *Managing your digital footprint*. Rosen Central.
- GREENLEAF, E. A. (1992): Measuring extreme response style. *Public Opinion Quarterly*, 56(3), 328–351.
- HAMAMURA, T., HEINE, S. J., PAULHUS, D. L. (2008): Cultural differences in response styles: The role of dialectical thinking. *Personality and Individual Differences*, 44(4), 932–942.
- JACKSON, D. N., MESSICK, S. (1958): Content and style in personality assessment. *Psychological Bulletin*, 55(4), 243.
- JOHNSON, T. R. (2003): On the use of heterogeneous thresholds ordinal regression models to account for individual differences in response style. *Psychometrika*, 68(4), 563–583.
- LENTZ, T. F. (1938): Acquiescence as a factor in the measurement of personality. *Psychological Bulletin*, 35(9), 659.
- LITTVAY, L. (2010): The Genetic Heritability of Survey Response Styles. Survey Research and Methodology program (SRAM) – Dissertations & Theses. University of Nebraska, Lincoln.
- MADDEN, M., FOX, S., SMITH, A., VITAK, J. (2007): *Digital footprints: Online identity management and search in the age of transparency*. Pew Internet & American Life Project Washington, DC.
- MCCRAE, R. R., COSTA Jr, P. T., MARTIN, T. A. (2005): The NEO-PI-3: A more readable revised NEO personality inventory. *Journal of Personality Assessment*, 84(3), 261–270.
- NAEMI, B. D., BEAL, D. J., PAYNE, S. C. (2009): Personality Predictors of Extreme Response Style. *Journal of Personality*, 77(1), 261–286.
- NAGYBÁNYAI NAGY O. (2013): Online személyiségmérés a hazai Big Five struktúra mentén: a Facet5 teszt magyar adaptációja. *Pszichológia*, 33(1), 37–59.
- RORER, L. G. (1965): The great response-style myth. *Psychological Bulletin*, 63(3), 129–156.
- RÓZSA S., KÁLLAI J., OSVÁTH A., BÁNKI M. (2005): Temperamentum és Karakter: Cloninger pszichobiológiai modellje. A Cloninger-féle Temperamentum és Karakter Kérdőív felhasználói kézikönyve. *Medicina, Budapest*.
- SZIRMÁK Z., DE RAAD, B. (1994): Személyiségtaxonómia: a magyar nyelv személyleíró szókincse. *Magyar Pszichológiai Szemle*, 50(1–2). 39–65.
- SZIRMÁK, Z., NAGY, J. (2002): *The Hungarian adaptation of the NEO Personality Inventory, Revised, Form S: Progress report on the reliability and validity of the first translation (NEOPIR-H1) and the item corrections (NEOPIR-H2) for the final version*. Unpublished manuscript, Free University of Berlin, Germany.