

VÁLASZSTÍLUS-ALAPÚ NEMZETKÖZI STANDARDOK AZ ÖNJELLEMZŐS SZEMÉLYISÉG- MÉRÉS KORRIGÁLÁSÁRA

NAGYBÁNYAI NAGY Olivér

Károli Gáspár Református Egyetem, Pszichológiai Intézet¹

(nagybanyai.oliver@kre.hu)

Absztrakt

Háttér és célkitűzések: Számos korábbi vizsgálat igazolta a válaszadási stílus általános jelenlétét az önkitöltős kérdőíveknél. A stabilan megjelenő egyéni válaszszélsőségségi stílus kimutatása azért fontos, mivel torzíthatja a tesztpontszámok standard átlagait is. Nagymintás kutatásunk célja, hogy kimutassuk az egyénre jellemző stílus mellett csoportszinten, nyelvtérületenként is a válaszszélsőségség lehetséges torzító hatásait a mérés pontosságára, a változók közötti összefüggésekre, illetve az azokból levonható következtetésekre is.

Módszer: A Facet5, mint nemzetközileg elterjedt, Big Five alapokra épülő személyiség kérdőív, lehetővé teszi független személyiségdimenziók mérése révén a választílusok differenciált azonosítását is (maradék-válaszszélsőségség arány). Az elemzett mintát összesen 28 (köztük magyar) nyelvből származó online tesztkitöltés képezte ($N=261.592$).

Eredmények: A Facet5 személyiségdimenziók mindegyikétől függetlenül kiszámítható öt válaszszélsőségség arány egymással markáns korrelációt mutat ($r=0,7$). Ez a válaszszélsőségség kb. 5-15%-nyi zaj forrása, ami kiszűretlenül maradvá, jelentősen megnehezítheti a megcélzott személyiségdimenzió pontos mérését. A teljes kérdőívben miert válaszszélsőségség arány alapján beazonosítható válaszklaszterek lehetőséget teremtenek a választílus szerinti standard normák meghatározására. Az újonnan létrehozott standardok alapján, újstandardizálások segítségével korrigálhatók az eredeti tesztpontszámok torzításai. Mivel a válaszszélsőségség arány és a válaszklaszterek előfordulásában jelentős különbségek mutathatók ki a nyelvtérületek összehasonlításában, így az újstandardizálások is nyelvtérület-specifikusan és eltérő mértékű korrekciók révén valósulhatnak meg. Például a Big Five „normaszegés – szabályok betartása” témájában a svéd – spanyol összehasonlításában így már ellentétes irányba fordul át az átlagok közti különbség.

Következtetések: Az önjellemzős személyiségmérésben a válaszszélsőségség miatti torzítás jelentős hatással bírhat az interkulturális összehasonlítások eredményeire, ideértve a műtermékek megjelenését, illetve a valós különbségek elmaszkolódását. A személyiségdimenziók pontszámainak nyelvi és válasz-klaszterek szerinti újstandardizálása egy viszonylag egyszerű módszer kínál a visszakorrigálására.

Kulcsszavak: szélsőségség és mérsékelt választílus ■ Big Five ■ online tesztelés

Summary

Response-style based international standards for the correction of self-reported personality measurement

Background and aims: Many former researches have proved the general presence of response style in self-reported questionnaires. Detection of individual extreme re-

¹ A kutatás a Károli Gáspár Református Egyetem kutatási finanszírozásának köszönhetően jött létre (Témaszám: 20649B800)

sponse-style is highly important because it might distort the standard averages so the conclusions of a study as well. The purpose of our large sample research is to show the effect of response-extremity at individual level and at group level in each language area.

Methods: Facet5, as an internationally known personality questionnaire built on the bases of Big Five theory, can identify the content independent “remnant” response extremity ratio index, which is calculated regardless the given personality dimensions of Big Five. The examined sample ($N=261.592$) is derived from online questionnaires filled in 28 (including Hungarian) languages.

Results: There is a strong correlation between the five “remnant” response extremity ratio indices ($r=0,7$). This response extremity is roughly the source of 5-15% noise, which can complicate the exact measurement of the targeted personality dimension, if remained unfiltered. We can identify 3 response clusters with the measured response-extremity in the whole questionnaire, and we can define clusters’ standard norms based on the response style. The newly created standards of Big Five scales allow the re-standardization and the corrections of the distortion of the original test scores. When comparing different language areas, significant differences can be observed between the occurrence of response extremities and response clusters, therefore re-standardizations must be language area-specific and corrections made with varying degrees. For example, the Swedish – Spanish comparison in the „compliance / rule-breaking” theme of Big Five, results opposite direction differences between the averages.

Discussion: Response extremity might distort self-reported personality measures both individually and per language area. All this might have a significant effect on the results of intercultural comparisons, including the appearance of artificial products and obscuration of real differences. Re-standardization of the personality-dimension scores is a relatively simple solution for correcting this distortion.

Kulcsszavak: extreme and moderate response style ▪ big five ▪ online testing

ELMÉLETI HÁTTER

A személyiségtesztelés terén a klasszikus papír-ceruza verzióhoz képest a számítógépes és online kiértékelési és értelmezési módszerek ma már lényegében általánossá váltak. Ehhez kapcsolódóan a tesztkiértékelés módszertana is számos új formai és módszertani elemmel bővült. A különféle nyelvi változatok elérhetősége által biztosított nagymértékű rugalmasság nyomán a teszt kitöltésének vagy éppen az eredmények automatikus értelmezésének (riportok) nyelve is könnyedén változtatható, de hasonlóan nagy mozgásteret jelenthet a pontszámítások alapjául szolgáló standard norma megválasztása, akár nyelvenként, akár egyéb speciális szempontokkal képzett alminták szerint.

Az önjellemzős tesztek („self-report data”) alapvetően a válaszok tartalmi aspektusára fókuszálva igyekeznek számszerűsíteni a célváltozót, ugyanakkor a számítógépes és online metódusok olyan „látens” információk rögzítését is lehetővé teszik, mint a válaszadás itemenkénti ideje, vagy éppen a válaszkálán (például az ötfokú Likert-skálán) produkált *válaszgyakoriságok* („response distribution”), amellyel a statisztikai valószínűségtől eltérő tendenciózus jelleg vizsgálható (Berg & Collier, 1953). Bár ezen a területen számos egymáshoz kapcsolódó fogalommal találkozhatunk, azonban ezek néhol eltérő meghatározással és jelentéstartalommal bírnak. A *választendencia* („response tendency”) fogal-

ma helyett inkább a *válaszstílus* („*response style*”) (Jackson & Messick, 1958) az elterjedtebb terminus (De Beuckelaer, Weijters, & Rutten, 2010; Van Herk, Poortinga, & Verhallen, 2004), de szintén ide kapcsolódó a Cronbach (1942; 1946) által bevezetett *válaszbeállítódás* („*response set*”) fogalma is (Cheung & Rensvold, 2000; Watkins & Cheung, 1995). Miközben a *válaszbeállítódás* elnevezést legtöbbször csak a válaszgyakoriságok területére szűkítve, azzal szinonimaként használjuk, valójában eredetileg Cronbach (1946, p. 476) számos egyéb válaszjellemzőt is ebbe a körbe sorolt (pl. válaszsebesség, válaszpontosság stb). A fogalmi különbségtétel korántsem egyértelmű a szakirodalomban, de a *válaszbeállítódás* terminust preferáló szerzők alapvetően az itemek formájától, tartalmától, illetve szituációs tényezőktől (külső feltételektől) függő jellemzőként tekintenek erre, míg a *válaszstílus* fogalmát előnyben részesítők pedig egy olyan általánosan (is) érvényesülő, az adott személyre specifikus, inherens tényezőként (belső diszpozícióként) kezelik, ami nem (vagy nem elsődlegesen) az itemek jellemzőinek (formájának, tartalmának) függvényében alakul. Maga Cronbach (1950) is leginkább a helyzet vagy item kétértelműségét, strukturátlanságát 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 markánsná, (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, cél szétválasztása miatt is – bár a legtöbb kutatás esetében keverten jelenik meg mindkettő. Egyfelől a *válaszbeállítódás* terminus egy olyan *externális* tényezők által meghatározott jellemzőnek tekinthető, amit a tesztelés során egyfajta minimalizálendő, kiiktatandó jelenségként kell kezelni. Ehhez képest viszont a *válaszstílus* egy *internális*, személyen belüli tulajdonság megnyilvánulásáról szól, aminek kontroll alatt tartására, befolyásolására nincs lehetőség, vagyis itt csupán a detektálásra, számszerűsítésre lehet törekedni. Jelen kutatásban 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 ráarakódó hatások is befolyásolják. A válaszgyakorisági eltérések vizsgálatához az indikátorok igen széles spektruma áll rendelkezésre, számos, jól elkülöníthető mérőszáma létezik a különféle kutatásokban (Baumgartner & Steenkamp, 2001).

A két legtöbbször vizsgált mutató a széli, vagy extrém válaszstílus („*extreme response style*” – ERS), illetve a beleegyező válaszstílus („*agreement response style*” – ARS). A széli válaszopciókra irányuló preferencia egy olyan általánosan és konzisztensen, időben is stabilan megjelenő (Bachman & O’Malley, 1984; Greenleaf, 1992; Naemi, Beal, & Payne, 2009) jellemző, mely tartalmilag jól leválasztható az egyébként szintén stabil egyetértési hajlandóságtól, vagy

beleegyező választílustól, illetve a szociális kívánatosságra törekvés (szociális megfelelési igény) fogalmától. Ezekhez képest az MRS (vagy MidRS – „*midpoint response style*”) egy lényegesen ritkábban vizsgált mutató, részben azért, mert ez a sok egyéb válaszlehetőség mellett csupán egyetlen, középső választás előfordulását detektálja (az ötfokú Likert-skála esetén például csak a 3-as választ). Azonban a középponti választás a válaszkála opciók jelentéstartalma függvényében fontos információval szolgálhat. Egyfelől jelölhet egy mérsékelt, átlagos szintet a vizsgált item (vagy a mérendő látens konstruktum) vonatkozásában, de értelmezhető úgy is, mint a két ellentétes végpont közötti felező pontot leképező *semleges választási lehetőség* („*neutral response option*”) (Weijters, Geuens, & Schillewaert, 2010). Ennek emelkedett előfordulása utalhat az önjellemzés bizonytalanságára, sőt bizonyos szint felett egyszerűen értékelhetetlenné teheti a tesztet a jelentős információ hiány. Amennyiben a válaszok sem az MRS (vagy MidRS), sem az ERS révén nem írhatók le (például, 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* (ModRS – „*moderate response style*”) szolgálhat hasznos mutatóként (Weijters, Geuens, & Schillewaert, 2010).

A válaszgyakorisági mutatók stabil, különféle helyzetekben és eltérő item-készletek, illetve válaszkálák mellett is stabilan tetten érhetők. Az *egyéni szintű* előfordulási valószínűségben longitudinális vizsgálatokkal (ugyanazon személyek felméréseivel) mind az ARS, mind pedig az ERS esetén igen magas szintű idői stabilitást lehet kimutatni (Bachman & O’Malley, 1984), még akár több éves időkülönbséggel is (Billiet & 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 ($r=0,5$) a megfigyelt együttjárás (Littvay, 2010). Mindezek alapján az a magyarázat is felvethető, hogy ezen választendencia jellemzők mögött valamilyen stabil, látens egyéni pszichológiai tulajdonságok is meghúzódhatnak (Austin, Deary, & Egan, 2006), tehát lényegében választílus-ként is értelmezhetők.

Nagymintás felmérések alapján az egyéni szintű stabilitás mellett kimutatható, hogy a választendencia terén a válaszadók kb. 25-30%-a jellemezhető széli választásokkal (Austin, et al., 2006; Eid & Rauber, 2000). Vagyis úgy tűnik, hogy a legkülönbélebb kérdőíves tesztelési helyzetekben meglehetősen nagy gyakorisággal megjelenik *populációs szinten* is ez a választílus. 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 választílus pedig a kelet-ázsiai országokban preferált (Gilman et al., 2008; Hamamura, Heine, & 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, a választendenciák kontrollálá-

sának hiányából fakadó műtermékek megjelenéséhez is vezethetnek. Különbéféle szimulációs, illetve 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; Plieninger, 2017).

Ilyen álszignifikanciákon nyugvó különbségek mellett általános pszichometriai problémák is felvetődnek, hiszen a választílus jelentősen befolyásolhatja a tesztadatok megbízhatóságát, illetve érvényességét. Így például az extrém válaszok preferálása (ERS) megnövelheti a skála reliabilitását, míg másik oldalról pedig csökkentheti a validitást (Clarke, 2000). De lényegében bármilyen tendenciózus válaszmegoszlási eltolódás, halmozódás (pl. ARS, ModRS) magával tud hozni egy olyan mértékű „tartalom-független” megosztott varianciát, ami már a skálák belső konzisztenciáját veszélyeztetheti.

HIPOTÉZIS

A válaszbeállítódás, illetve a választílus, mint az egyénre vonatkozó adat a tesztek fejlesztése, illetve ez eredmények értelmezése terén felhasználható információ forrásként szolgálhat. Miközben a hagyományos önkitöltős 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. Hipotézisünk szerint az önjellemzős személyiségtesztelésben a válaszszelesőségesség, mint tartalomtól függetlenül érvényesülő konzisztens hibaforrás a ténylegesen mérendő személyiségdimenzió mutatott pontszámokat torzíthatja. Ennek kiküszöbölésével, korrekciójával pedig a személyiségtesztben megjelenő manifeszt teszteredmények végső soron a látens konstruktumok valódi értékéhez közelebbi eredményekkel szolgálhatnak.

MÓDSZER

Mérőeszköz

A vizsgálatban a Facet5 (Buckley & Williams, 2002) Big Five alapú online személyiségteszt hazai és nemzetközi alkalmazásokban kapott adatait használtuk fel. A teszt által mért dimenziók tartalmilag jól illeszthetők a hazai személyiségtaxonómiai elemzések sajátosságaihoz is (Nagybányai Nagy, 2013). Az öto-

dik, *Empátia (Affection)* skála tartalma, elnevezése az angolszász faktorstruktúrában megszokott *Nyitottság* faktorról nem azonosítható egyértelműen, inkább a Szirmák-féle magyar vizsgálat (1994) terminológiájával hozható kapcsolatba, melyben az *Integritás* faktor markerszavai alapján hosszabban „Pozitív viszonyulás másokhoz” megjelöléssel is illethetnénk. Hasonló eredményeket korábban más Big Five kutatások is megemlítenek, így például az olasz személyiségstruktúrában külön IV. faktorként jelenik meg az altruizmus („szenzitív”, „nagylelkű” terminusok) tengelye, valamint külön egy V. faktor („forradalmi”, „progresszív” jelzőkkel) a non-konformitást képezi le (De Raad, B., M. Perugini, & Zs Szirmák, 1998). A Facet5 tesztben a további Big Five dimenziók közül az *Extraverzió* mérésére a *Lendület (Energy)* elnevezésű skála, a *Lelkiismeretesség* mérésére pedig a *Kontroll (Control)* skála szolgál. Az *Akarat (Will)* skála az eredeti *Kellemesség* faktor alacsony végpontját célozza meg, míg az *Emocionalitás (Emotionality)* skála pedig a Big Five terminológia szerinti *Érzelmi stabilitás* dimenzióknak az ellentétes oldalát képezi le.

A Facet5 tesztben a dimenzió pontszámok meghatározása – az adott nyelvtérület standard normája alapján – egy tízfokú („standard-ten”) Sten-skálán 1-10 közötti pontozással történik. A teszt elterjedt nemzetközi alkalmazása alapján 28 nagymintás nyelvi standard norma áll rendelkezésre a nyerspontok átszámításához. 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. A Facet5 esetén a középválasztás, vagyis a 3-as válaszok megjelenését a válaszszélsőségeség szempontjából nem tekinthetjük releváns információnak, hiszen ebben a két ellentétes oldal közötti választás hiányát tapasztaljuk, ami lényegében az itemmel kapcsolatos 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. A kutatás fókuszára miatt lényeges azt is leszögezni, hogy az egyes skálákon belüli fordított itemek (melyek egyébként közel kiegyensúlyozottan fordulnak elő a pozitív tételekkel) nem befolyásolják a válaszszélsőségeség kiszámítási módját, mivel a széli választások (1-es & 5-ös) vagy a mérsékelt választások (2-es & 4-es) előfordulása a tételfordítások után is ugyanúgy jelennek meg.

A statisztikai elemzések elvégzése IBM SPSS Statistics 26, valamint RopStat szoftverek segítségével történt. A statisztikai próbák előfeltételeinek ellenőrzése minden esetben megtörtént, de hely hiányában ezek egyenkénti bemutatásától eltekintünk.

Minta

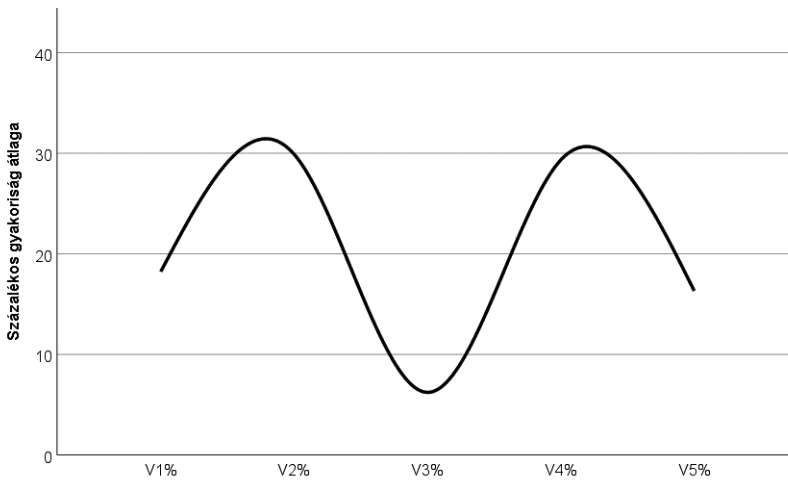
A Facet5 teszt kitöltésében résztvevők almintái összesen 28 nyelvterületről származnak, minden al minta 400 fő feletti, így a teljes mintaméret meghaladja a negyed millió főt ($N=261.592$). A kutatás ökológiai validitásának növelése érdekében a résztvevők szinte kivétel nélkül valamilyen valós szervezeti HR-alkalmazás kapcsán töltötték ki a személyiségkérdőívet a legkülönbözőbb munkakörökből, de jellemzően banki, pénzügyi, HR, IT, értékesítési, műszaki területekről. A rendszer az online kitöltésre minden esetben csak személyre szóló belépési kód segítségével adott lehetőség, így a random vagy többszöri próbálgatással történő kitöltés minimalizálható volt.

A teljes minta 0,8%-át kizselektáltuk annak érdekében, hogy az extrém egyoldalú válaszhalmozódásokat, az értelmetlen válaszkála-használatot kiszűrjük. Az 1-5-ig terjedő válaszopciókban maximalizáltuk az elérhető gyakoriságokat, ezzel az összes itemen mutatott válaszsóródás elérte minimálisan a 0,5-es szintet (így vált kiküszöbölhetővé, hogy például a vsz. csupán egy vagy két opció használjon, vagyis pl. csak 4-es és 5-ös válaszokat adjon a teljes kérdőívben).

EREDMÉNYEK

A válaszkála-opciók differenciált előfordulása

Globálisan vizsgálva a jelenleg rendelkezésre álló széles mintát ($N=261.592$) az egyes válaszfokozatok előfordulási gyakoriságában a mérsékelt (2-es és 4-es) választások relatív túlsúlya figyelhető meg, szemben a széli választások (1-es és 5-ös) alacsonyabb szintjével (1. ábra). Ugyanakkor mindkettőtől lényeges mértékben eltér a középválasztások, vagyis a 3-as válaszok igen alacsony előfordulási gyakorisága, ami a legkritikább a többi válaszopcióhoz viszonyítva. Ez utóbbi háttérben az áll, hogy a kérdőív instrukciója kifejezetten 3-as válaszok kerülésére figyelmeztet.



1. ábra. A Facet5 kérdőívben az ötfokozatú válaszok (v1 – v5) egyéni százalékos előfordulásának átlaga. (Az ábrán az összekötő vonal a vizuális összehasonlítást segítő szerepet játszik, statisztikai értékek csak az öt kategória feletti pontokban értelmezhetők.)

A középválasztások figyelmen kívül hagyásával a széli-, illetve mérsékelt- választások előfordulásaira fókuszálva erős páronkénti összekapcsolódás figyelhető meg. (1. táblázat). Ezen válasz párok tehát nem csupán hasonló gyakorisággal bírnak globálisan a mintában, de egyéni szinten a válaszadásokban páronként masszív együttjárást is mutatnak, miközben pedig a széli-, illetve mérsékelt- választások megjelenése egymással komplementer viszonyban áll.

1. táblázat. A széli- (v1 – v5), illetve mérsékelt- (v2 – v4) válasz fokozatok előfordulása közötti együttjárások

Pearson-korreláció	V5	V4
V1	0,862**	-0,856**
V2	-0,866**	0,767**

**($p < 0,01$)

Ennek az eredménynek az értelmezéséhez fontos kiegészítésként megjegyezni, hogy a válaszkategóriák választási előfordulása egymástól nem teljesen független, hiszen valamely válaszopció megjelenése automatikusan lecsökkenti az egyéb válaszlehetőségek jelenlétét. Ha egy Facet5 teszthez hasonló méretű, 106 item terjedelmű kérdőívben egy ekkora, kétszáz-hatvanezres mintára véletlenszerű tesztválaszokat generálnánk, akkor szintén adódnának gyenge negatív együttjárások a válaszfokozatok előfordulási gyakoriságai között (*Pearson-féle* $r = -0,247 - -0,252$; $p < 0,001$; *Spearman-féle rho* $= -0,236 - -0,241$; $p < 0,001$).

Vagyis minden válasz-összekapcsolódási korreláció értelmezésekor ehhez az enyhe negatív „alapszinthez” képest célszerű viszonyítani: így nagyjából már egy 200 fős minta esetén is azt mondhatjuk, ha a Pearson-féle korrelációnak bármilyen pozitív értékét detektáljuk, hogy ez az alapszinttől szignifikánsan különböző kapcsolatnak tekinthető. Esetünkben a Facet5 kérdőívben is erről van szó: az 1-es és az 5-ös válaszokban, valamint a 2-es és a 4-es válaszfokozatoknál tapasztalható igen magas pozitív együttjárások ($r=0,767 - 0,862$; $p<0,001$) nagyon erős eltérést jeleznek a random adatokon tapasztalt alapszintű enyhe negatív korrelációs eredményektől ($Z=574,72 - 576,66$; $p<0,001$).

Főkomponens elemzést végezve a Facet5 mintán mind az öt válaszfokozat egyéni mennyiségein egy kétfaktoros megoldás születik, amiből az elsőre 0,9 feletti töltéssel, pozitív előjellel terhelnek az 1-es és az 5-ös válaszok, míg ugyanilyen erősséggel, de negatív előjellel szerepelnek rajta a 2-es és a 4-es válaszopciók (2. táblázat). Ez a faktor a varianciák 70%-át fedi le, míg a második főkomponensre pedig csak a 3-as válaszok változója kerül és 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 ott nem mutatható ki semmilyen egységes válaszszélsőségségi szerveződés.

2. táblázat. A válaszopciók rotált (varimax) főkomponens mátrixa

Válasz-opciók	1	2
v1	-0,941	-0,214
v2	0,940	-0,043
v5	-0,940	-0,218
v4	0,937	-0,066
v3	0,064	0,998

Mindezen elemzésekre azért volt szükség, hogy kimutassuk a széli választások, illetve a mérsékelt választások összevont gyakorisági indexének indokoltságát, másrészt pedig továbblépve, ez alapján a kettő egymáshoz viszonyított fölényét egy közös értékben, a *válaszszélsőségség arány mutatóban* tegyük megragadhatóvá – ötfokú válaszkála esetén $[V_1+V_5] - [V_2+V_4]$ képlettel. Ez az összevont válaszstílus mutató mind a teljes kérdőívre, mind pedig Big Five dimenzióként kiszámítható. Ami azért is lényeges, mert ezzel a személyiségskálánkénti válaszszélsőségség arány mutatók függetlenségét, illetve kapcsolatának erősségét is vizsgálni tudjuk.

A korrelációs elemzések szerint mind az öt Big Five dimenzió válaszszélsőségség arány mutatója között erős, szignifikáns lineáris kapcsolat mutatható ki ($r=0,74 - 0,79$; $p<0,01$). Vagyis az egyéni szélsőségségi válaszadási trend jól

láthatóan a kérdőív egészen végigvonul, a mérési területek mindegyikén hasonló módon kirajzolódik. Ez az eredmény azért is fontos adalék, mivel itt a Big Five pontszámok egyéni sokszínűsége mellett megjelenő markáns individuális hatásról van szó. Tehát miközben az egyes személyiségskálákon akár nagyon eltérő pozíciót is mutathatnak az egyének (hiszen a Big Five faktorok definíció szerint korrelálatlanok), ami nyilvánvalóan befolyásolhatja az önjellemzésben adott válaszok extrém vagy mérsékelt jellegét, addig ezzel párhuzamosan a válaszadás stílusa a skálák (itemek) tartalmától független is leképződik, arra mintegy ráakódik.

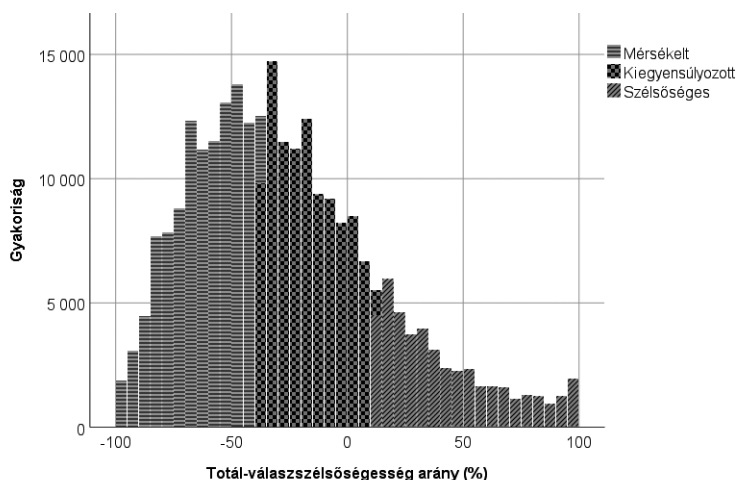
Módszertani szempontból éppen ez jelent komoly problémát, hiszen egy konkrét személyiségskála itemeinél a válaszkála szélsőséges használata egyaránt vezethet szélsőséges skálapontszámhoz és ugyanakkor magas válaszzélsőségeség arányhoz is. Ezen két hatás keveredésének kivédésére, a válaszadás tartalmi, illetve formai jellegének szétválasztása érdekében a teljes kérdőívben mutatott „*totál*”-válaszzélsőségeség arány mellett a vizsgált személyiségskála itemei nélkül mutatott „*maradék*”-válaszzélsőségeség arány is alkalmazni fogjuk (a pszichometria item-totál, illetve item-maradék terminusainak analógiájára). A tartalmi – formai szétválasztás nyilvánvalóan nem lehet tökéletes, hiszen láttuk, hogy ez a választílus minden személyiségdimenzió végighúzódik. Ennek tudható be, hogy a teljes Facet5 adatbázison egy-egy maradék-válaszzélsőségeség arány és a hozzá tartozó személyiségdimenzió lineáris és kvadratikusan értéken képzett regressziós modell mind az öt esetben szignifikáns eredményt hoz ($R=0,21 - 0,36$; $p<0,01$), amelyek kb. 5-13%-os megmagyarázott hányadot (R^2) képviselnek – ami a torzítatlan becsléssel (*adjusted R²*) is azonos értékű jelen mintaméret mellett. Mivel tehát ez a zaj nagyjából 10% körüli értéket is elérhet, így kiszűretlenül maradván megnehezítheti a megcélzott személyiségdimenzió pontos mérését.

A válaszzélsőségeség arány klaszterei

A válaszzélsőségeség arány mutatók közötti kapcsolatok elemzésén túl, az egyének közötti variabilitás megragadására, célszerűnek tűnik az azonos válaszzélsőségeségi mintázatot mutató alcsoportok elkülönítésére is kísérletet tenni. Amennyiben a skálahasználat dimenzionális megközelítésén túllépve a választílus típusok elkülönítésére fókuszálunk, akkor különféle klaszterek feltárását érdemes megcélozni. Erre az ún. kétlépéses klaszterezést (Two Step Cluster, TSC), a hierarchikus eljárást egyik változatát alkalmaztuk, amely abban speciális, hogy a klaszterek száma menet közben „adódik ki”, a modell nem igényel célváltozót („target field”). A TSC módszer első lépésként a nyers adatok alapján létrehozza a „részklasztereket” („sub-clusters”) egy kezelhető számú körét, majd második lépésben pedig ezeknek a részklasztereknek az egyre nagyobb „gyűj-

tóklaszterekbe” történő összevonása történik meg (de már az elemi adatok újrafeldolgozása nélkül).

A jelenlegi vizsgálati mintát felhasználva a széli- és mérsékelt választásoknak a teljes kérdőív során mutatott előfordulási aránya (ez a totál-válaszszélsőségesség arány változó) alapján három egymástól jól elkülöníthető válasz-klaszter adódik (2. ábra). A klaszter szétválás, tehát az inter-klaszter távolságok, valamint a klaszter koherencia, vagyis az intra-klaszter távolságok viszonyát leíró Sziluett-együttható értéke = 0,7 – ami már 0,5 felett is kifejezetten hatékonynak mondható. Logisztikus regresszió alkalmazásakor a klaszterekbe történő viszszaosztás (keresztvalidálás) összességében 99,8%-os sikerességű az ötféle válaszfokozat változói alapján ($Khi^2=508967623,215$; $p<0,001$; *Nagelkerke pseudo $R^2=0,994$*), illetve 100%-os a klaszter kialakításhoz eleve használt egyesített totál-válaszszélsőségesség arány index alapján.



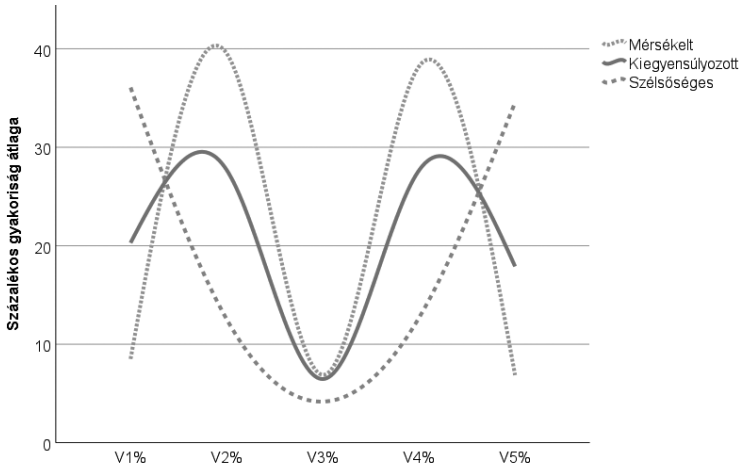
2. ábra. A szélsőséges- és mérsékelt-válaszopciók százalékában kifejezett totál-válaszszélsőségesség arány gyakorisági megoszlása a három válaszklaszter szerint

3. táblázat. A válaszszélsőségességi klaszterek mérete a teljes mintában

Klaszter-név	Fő	%
Mérsékelt	110.605	42,3
Kiegyensúlyozott	102.552	39,2
Szélsőséges	48.435	18,5

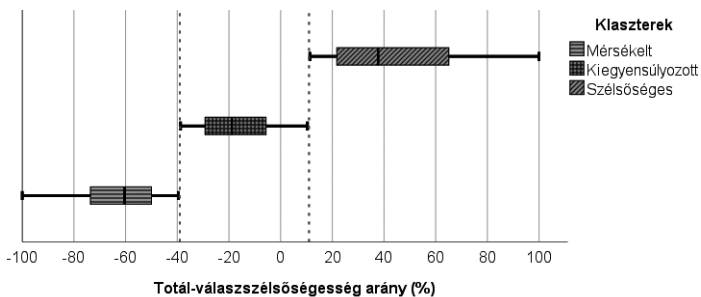
A klaszterek közül a Szélsőséges választási preferenciával rendelkezők alkotják a legkisebb, nagyjából 20%-nyi csoportját a mintának, míg a Mérsékelt válaszklaszterbe tartozók és a Kiegyensúlyozott csoport közel azonos méretű, kb. 40-

40%-ot képvisel (3. táblázat). Az egyes válaszopciók gyakorisági átlaga jól láthatóan különbözik a három válaszklaszter esetén (3. ábra).



3. ábra. Az ötfokozatú válaszok (v1 – v5) egyéni százalékos előfordulásának átlaga klaszterenként bontva. (Az ábrán az összekötő vonal a vizuális összehasonlítást segítő szerepet játszik, statisztikai értékek csak az öt kategória feletti pontokban értelmezhetők.)

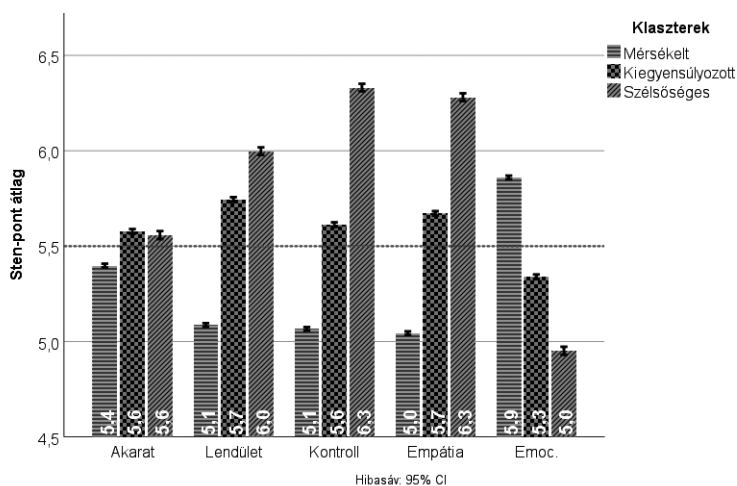
A klaszterek nem lógnak össze, a válaszszélsőséges arány határértékei -40%, illetve +10% körül találhatóak (százalékos értékben kifejezve), míg középpértékük a Mérésékelttől a Szélsőséges klaszter felé haladva rendre -60%, -20% és +40% környékén alakul (4. ábra).



4. ábra. A klaszterek határértékei és középpértéke a válaszsélsőséges arány százalékos értékben kifejezve

Újrastandardizálás és pontszám-korrekcó

A Facet5 teszteredményeket vizsgálva az eredeti Sten-pontszámok válaszklaszterenkénti átlaga egyedül az Akarat dimenzióban (Big Five Kellemesség faktórának inverz skálája) nem tér el a Mérsékelt és a Kiegyensúlyozott klaszter között, de minden más összevetésben 1%-os szinten szignifikánsan különböznek a skálaátlagok egymástól. Ráadásul a Sten-skáláknál az elméletileg elvárt 5,5 pontos értéktől is 1%-os szinten szignifikáns az összes skálaeltérés – még a Mérsékelt klaszter pontszámainál is, bár a másik két klaszternél jóval nagyobb mértékűek ezek (4. ábra). Mindez azt a problémát világítja meg, hogy a választílus (vagyis a kitöltés formai jellege) alapján történő csoportosítás a Big Five átlagok (vagyis a kitöltés tartalmi jellege) terén is jelentős differenciákhoz vezet.



5. ábra. Az eredeti Sten-pontszámok szerint elvárt ($M=5,5$; $SD=2$) személyiségskála értékektől mutatott klaszterenkénti eltérések átlaga és 95%-os konfidencia intervalluma (CI)

Az eredmények alapján úgy tűnik, hogy a Facet5 tesztben a választendencia egyénre jellemző, konzisztens jelenléte globális, ráarakódó hatásként befolyásolja, torzítja a Big Five skálapontszámok alakulását. Ennek kiküszöbölése érdekében az egyik legegyszerűbb módszernek a válasz-klaszterenkénti normatív átlagok használata kínálkozik – az eddigi nyelvenkénti, országokkénti standardizálások mellett – ún. újrastandardizált pontok kialakítására. Hasonlóan ahhoz, mint ha mondjuk a nemek szerinti erős átlageltolódások miatt külön férfi és női standardhoz történő viszonyítással számítanák ki az egyéni pontértékeket.

A választílus-klaszterek és nyelvi csoportok szerinti együttes standardizálás révén (tehát ahol egymás után egy-egy nyelvnek egy-egy válasz-klaszterére történik a nyers értékek standardizálása, majd ezeket egy közös Sten-skálában

egyesítjük) kaphatnánk a legprecízebb eredményt, hiszen akkor a teljes mintánk minden Big Five skáláján 5,5 pont átlagú és 2 pont szórású értékkel rendelkeznenék. Ehhez képest lényegesen egyszerűbb módszert jelent a már rendelkezésre álló nyelvi standard-értékek három klaszter szerinti újstandardizálása. Bármelyik úton is indulunk, a választendenciából fakadó általános pontszám eltolódásokat ellensúlyozó korrekciókat végzünk, amivel a válaszszélsőségességi-hatástól megtisztított korrigált Sten-pontszámokat hozunk létre. Ennek a válasz-klaszterekre végzett újstandardizálással kapott pontszámoknak az alap, nyelvi-standard szerinti Sten-ponttól való eltérése (vagyis a Sten-pont korrekciók mértéke) minden klaszterben és minden személyiségskálán 1%-os szinten szignifikáns változást jeleznek, és jól láthatóan a Kiegyensúlyozott klasztert érintik a legkevésbé (5. ábra és 4. táblázat).

4. táblázat. A Facets Big Five dimenziók újstandardizálás miatti Sten-pont korrekciói választaszterenként ($H_0=0$ értékhez viszonyítva)

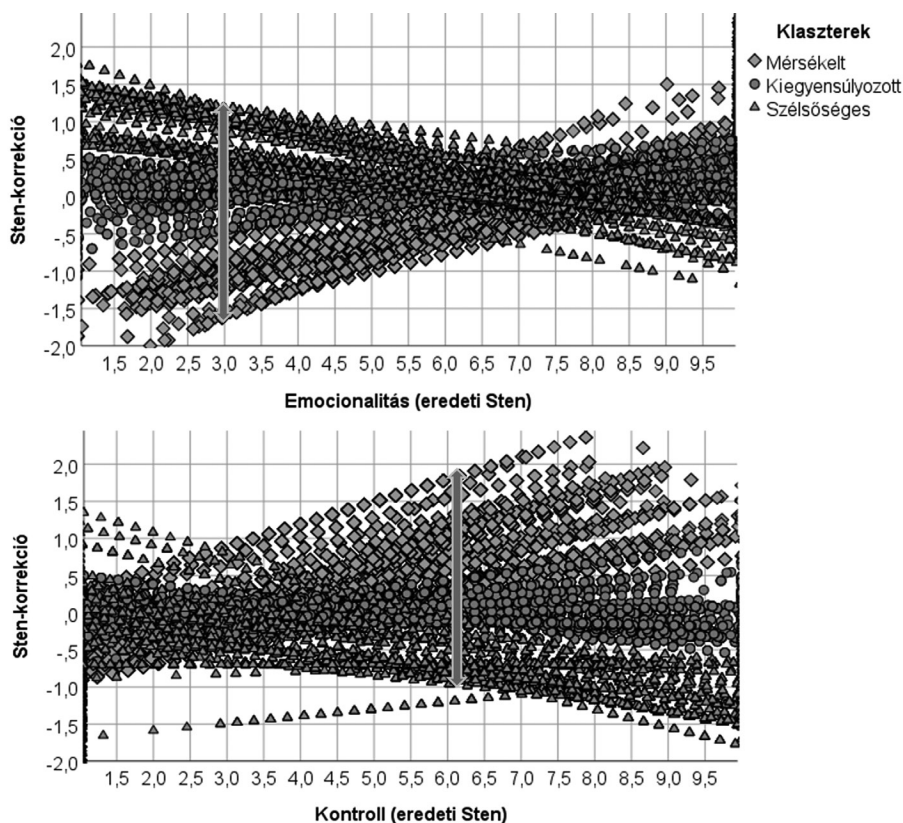
Sten-pont korrekció	Mérésékelt ($df=110.604$)			Kiegyensúlyozott ($df=102.551$)			Szélsőséges ($df=48.434$)		
	M	SD	t	M	SD	t	M	SD	t
Akart korrekció	0,10	0,36	93,63**	-0,08	0,16	-159,53**	-0,06	0,47	-27,42**
Lendület korrekció	0,41	0,37	375,67**	-0,24	0,18	-443,85**	-0,50	0,38	-288,76**
Kontroll korrekció	0,43	0,47	309,43**	-0,11	0,21	-167,92**	-0,83	0,41	-442,94**
Empátia korrekció	0,46	0,48	315,52**	-0,17	0,24	-224,31**	-0,78	0,49	-351,06**
Emoc. korrekció	-0,36	0,39	-305,63**	0,16	0,17	300,02**	0,55	0,45	265,97**

(** $p < 0,001$)

A Sten-korrekciók iránya nem minden esetben triviális, például az Emocionalitás esetén a Mérésékelt választílus korrigálása inkább pozitívabb pontszámokat eredményez, míg a Szélsőséges klaszter pedig negatív irányú korrekciót igényel (a Kiegyensúlyozott pedig keverten), ezzel szemben az összes többi személyiségdimenzió éppen fordított irányúak a torzulások, így a korrekciók iránya is (5. ábra és 6. ábra).

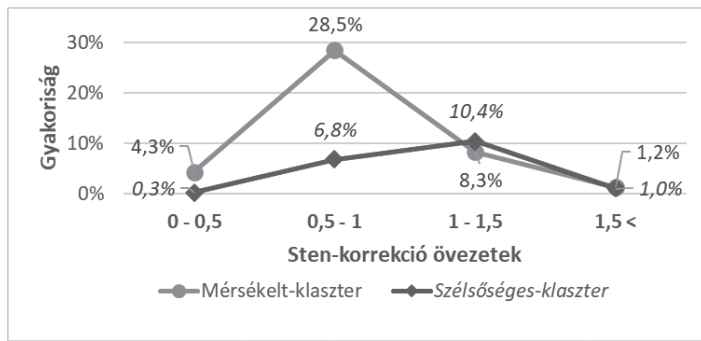
Ezen pontszámeltolódások elsősre nem tűnnek nagyoknak az 1-10-es terjedelmű Sten-skálán, de tegyük hozzá, hogy itt több tízezres mintákon megjelenő átlagos torzulásokról van szó, és ezeken belül az egyéni korrekciók jóval markánsabbak is lehetnek, ráadásul a magas, illetve alacsony Sten-pont tartományoknál nem ugyanolyan módon érvényesülnek. A pontszám torzulások a Sten-skála extrémebb pontszámövezeteiben a legjelentősebbek, éppen abban a skála tartományban, ami a pontszám értelmezésben legérzékenyebb a kis eltérésekre. Másik oldalról értelmezve fontos azt is látni, hogy ugyanaz a személyiség-skála Sten-pontszám hozhat +1,5 vagy -1,5 pont korrekciót is egy-egy átlagos egyén szintjén szintjén, attól függően, hogy azok melyik klaszterbe tartoznak. Vagyis

egy ilyen példában (6. ábrán nyilakkal jelölve) összességében már 3 Sten-pontos differencia is létrejöheth két személy között egyszerűen a különböző választílus-ba tartozásuk miatt, ami pedig számottevő, már kb. 30-50%-os gyakoriságbeli (percentilis) eltérést jelent.



6. ábra. Az Emocionalitás-, illetve a Kontroll-dimenzió eredeti (nyelvi) standard szerinti pontszámához szükséges válaszklaszterenkénti Sten-korrekcók (nyilakkal jelölve a legalább hárompontos egyéni differenciák határa)

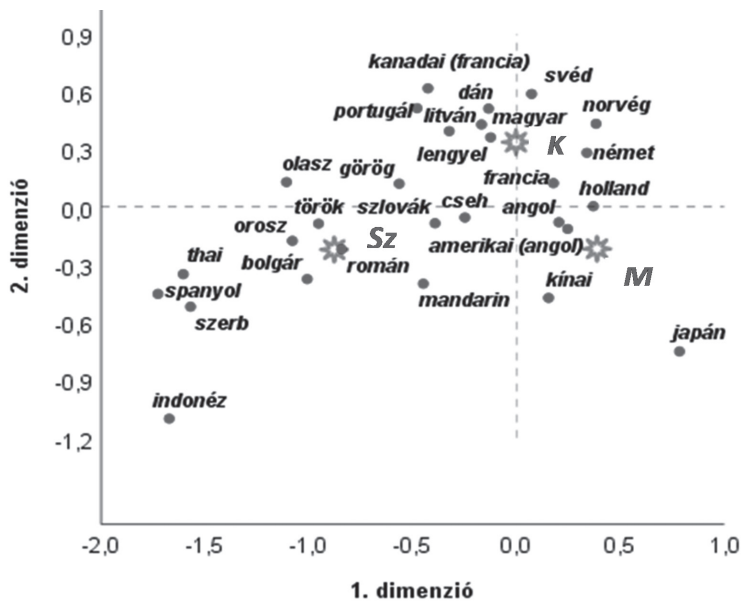
A skálapontszám korrekciók meglepően nagy valószínűséggel fordulnak elő a teljes mintában. Legalább egy Big Five területen valamilyen minimális, 0,5 Sten-pontot elérő pontszámtorzulást a teljes mintánk közel 60%-a mutat (a Mérsékelt választílusúakból 38%, a Szélsőséges válaszklaszterből 18%, míg a Kiegyensúlyozott csoportból csupán 3% származik), de még az 1 Sten-pontos értéket meghaladó korrekciók is a teljes minta ötödénél (22%) előfordulnak (7. ábra).



7. ábra. Sten-korrektciók nagyság szerinti övezeteinek gyakorisága a teljes mintán belül Mérésékelt-, illetve Szélsőséges-klaszterek mentén szétbontva

A Big Five pontszámok nyelvterületek közötti valódi- és fals-különbségei

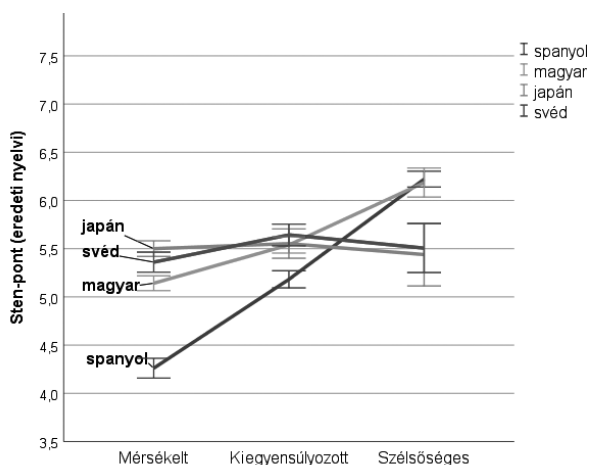
A 28 nyelvi csoportra külön-külön standardizált alap Big Five értékek, tehát az eredeti nyelvi-standard pontok egységesen 5,5 pont átlagúak és 2 pont szórásúak minden Facet5 skála esetén. Ehhez képest látványos széttartás következik be, ha a három válaszklaszter szerinti bontásban vizsgálódunk tovább.



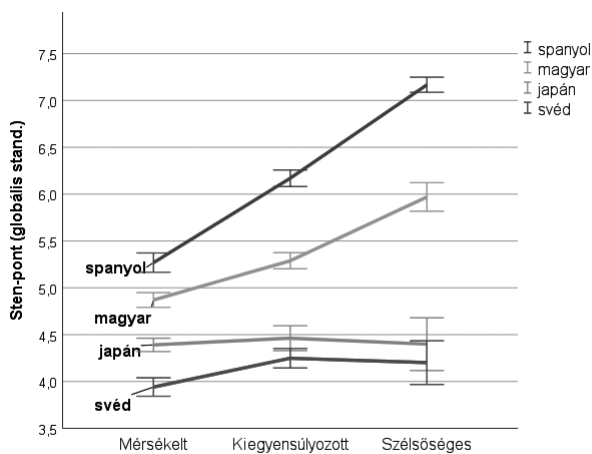
8. ábra. A Mérésékelt- (M), Kiegyensúlyozott- (K), Szélsőséges- (Sz) klaszterközéppont és a 28 nyelvi csoport fordított-U alakú kétdimenziós elhelyezkedése (válaszszélsőségesség-arány korrespondencia-analízise)

Mivel a nyelvterületek válaszszelesőségesség arány mintázata között jelentős különbségek vannak – emiatt is tagolódnak csoportokba (8. ábra) – így az eredeti Sten-pontok is más-más mértékben és irányban válnak el egymástól klaszterenként. Annak érdekében, hogy most már a válaszklaszterek szerint is 5,5-ös Sten-átlagú értékeink lehessenek – nyelvenként eltérő mértékű korrekcióra lesz szükség. Bizonyos nyelveknél (pl. japán, holland) alig, míg másoknál (pl. portugál, mandarin) viszont komolyabb, akár +/-1 pontos eltolást is végre kell hajtani – amiatt, hogy egy szélsőséges vagy egy mérsékelt klaszterből származó személyről van szó. Ez utóbbi nyelvi csoportokban pedig már nem ritka az sem, hogy a korrigálás kb. 1,5 – 2,5 pontos egyéni különbségeket eredményez ugyanazon kezdeti Sten-értékű (azonos nemzetiségű) személyeket összevetve. Nézzünk erre egy konkrét és extrém példát! Amennyiben a Big Five Lelkiismeretesség faktorát leképező Facet5 Kontroll-skála nyers pontjait (amit most a könnyebb értelmezhetőség érdekében a teljes, globális mintánkon Sten-pontokra standardizáltunk) összevetjük a különféle nyelvterületek között, akkor számos szignifikáns és jelentős átlagbeli eltérést találunk – így például a spanyol átlag messze magasabb, mint a svéd középérték, de persze a nyelvenként standardizált pontjaik ($M=5,5$; $SD=2$) természetesen nem térnek el egymástól (9. ábra). Ha azonban az eredeti (nyelvenként standardizált) Sten-pontok válaszstílus-klaszterenkénti újraprofilozált értékeit vetjük össze, akkor már a svédek és spanyolok viszonylatában kisebb különbségekkel, ám ezzel éppen ellentétes mintázatra és végkövetkeztetésre fogunk jutni a „normaszegés – szabályok betartása” Big Five témáját illetően, mivel egy nagy negatív, illetve egy kis pozitív irányú hatásmérték ellentmondása áll egymással szemben (5. táblázat). Tehát a válaszklaszterek figyelembevétele nélkül egy igen erős interkulturális hatásról ($Cohen-d=1,26$) számolnánk be, míg a korrekció elvégzése után csupán egy alacsony szintű hatásról ($Cohen-d=0,22$), de ami még fontosabb, az eddigiekkel szögesen ellentétes irányú differenciáról értekezhetnénk.

VÁLASZTÍLUS-ALAPÚ NEMZETKÖZI STANDARDOK AZ ÖNJELLEMZŐS SZEMÉLYISÉGMÉRÉS...



Hibasáv: 95% CI



Hibasáv: 95% CI

9. ábra. A Facet5 Kontroll-skála eredeti nyelvi standardizált pontjainak (fent), globális mintán standardizált nyers pontjainak (középen), illetve a klaszterenként újstandardizált pontjainak (lent) átlaga és 95%-os konfidencia intervalluma (CI) négy nyelvterületen

5. táblázat. A svéd – spanyol minta eltérése a Facet5 Kontroll-skálán: az újstandardizált pontszámítás eltérő mértékű és ellentétes előjelű Cohen-d hatásmértéket eredményez a nyers-pontból standardizáltak eltéréseihez képest

	Svéd (N=2524)		Spanyol (N=4421)		Cohen-d	CI95
	M	SD	M	SD		
Nyers-pont globális standardizáltja	4,12	1,82	6,49	1,92	-1,26**	-1,33 -1,2
Eredeti-pont (nyelvi-standard) klaszterenkénti újstandardizáltja	5,54	2,05	5,13	1,77	0,22**	0,14 0,27

MEGVITATÁS

A jelen vizsgálatban azt a korábban többször is kimutatott hatást vizsgáltuk, miszerint a választendencia egyénekre jellemző módon (pl. Bachman & O'Malley, 1984), a mérendő dimenzió tartalmától részben függetlenül is érvényesülve, számottevő hatással bírhat a kérdőív minden egyes tételére adott válaszban, ami az önjellemzős kérdőívek mérési hatékonyságát, torzítottságát is érdemben befolyásolhatja. A probléma lényegét az adja, hogy az egyéni válaszszélsőségeség nem tisztán az itemek tartalmi megítéléséből származik, tehát nem csupán a mért személyiségskálákon elfoglalt szélsőséges pozíció állhat a háttérben, hanem egy rejtett, másodlagos torzító hatás is érvényesül benne. Mivel a válaszpreferencia egymástól tartalmilag teljesen független mérési területeken is képes torzítani, így globális befolyásoló hatása miatt 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 válaszszélsőségeségi torzítások mind az egyéni, mind pedig a csoportos eredmények interpretációjában, komoly értelmezésbeli eltérésekhez, akár ellentmondó következtetések levonásához is vezethetnek.

A konkrét empirikus elemzésünk a Facet5 által mért Big Five személyiségdimenziók széleskörű nemzetközi adatbázisa köré szerveződött, ami több mint negyedmillió személy teszteredményeinek felhasználását tette lehetővé. A választílus miatti torzítás kontrollálásához az első fontos lépést a válaszszélsőségeség arány indexének kialakítása, valamint ez alapján pedig három, egymástól markánsan elkülönülő Mérsékelt, Kiegyensúlyozott és Szélsőséges válaszadói klaszter meghatározása jelentette. Ez volt szükséges ahhoz, hogy minden Big Five dimenzióhoz meghatározhatóvá váljanak a választílus-klaszterek speciális átlagértékei, illetve az így definiált standard normák segítségével pedig újstandardizálhatókká váljanak az eredeti tesztpontszámok.

Ahogy más kutatások is kimutatták, a nyelvterületek válaszszelesőségességi mintázata között jelentős különbségek vannak (Gilman et al., 2008; Hamamura, Heine, & Paulhus, 2008), így az újstandardizálási pontszám-korrekciók mértéke és iránya is eltérően alakul az egyes személyiségdimenziókon. Mivel a választílus hatás az empirikus minták jelentős részét érintheti (Austin, et al., 2006; Eid & Rauber, 2000), ezért a bemutatott korrekciók elvégzésével, jelentősebb volumenű pontszámváltozások is történhetnek, így pedig az eredmények elemzése és értelmezése terén egymással ellentétes végkövetkeztetések is születhetnek. Jelen példánkban bizonyos nyelvterületi differenciák terén (pl. spanyol / svéd) a korábbi műtermék jellegű eredményekkel szemben a korrigált értékek kifinomultabb valódi összehasonlításokat tesznek lehetővé.

Gyakorlati oldalról nézve úgy tűnik, hogy a személyiségdimenziók pontszámait torzító általános válaszszelesőségességi hatás ezen klaszter-specifikus újstandardizálási módszer révén egyszerűen kiküszöbölhető. Mivel az egyre elterjedtebb online tesztrendszerekben kifejezetten egyszerűen megvalósítható a standard normák ilyen irányú bővítése, így a tesztkitöltést követően a megfelelő válaszszelesőségességi standard normákat (is) figyelembe véve lehet a korrigált standard skálapontszámokat kiszámítani.

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 inter-related 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., & Rensvold, 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.
- Cronbach, L. J. (1942). Studies of acquiescence as a factor in the true-false test. *Journal of Educational Psychology, 33*, 401-415.

- 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.
- De Raad, B., M. Perugini, & Zs Szirmák (1998). In Pursuit of a Cross-lingual References Structure of Personality Traits: Comparisons among Five Languages. *European Journal of Personality* 11(3):167-85.
- 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., . . . 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.
- 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.
- Littvay, L. (2010). The Genetic Heritability of Survey Response Styles.
- 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.
- 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.
- Plieninger, H. (2017). Mountain or molehill? A simulation study on the impact of response styles. *Educational and Psychological Measurement*, 77(1), 32-53.
- Szirmák, Z., & De Raad, B. (1994). Személyiségtaxonómia a magyar nyelv személyleíró szókincse. *Magyar Pszichológiai Szemle*, 50(1-2), 39-65.
- Van Herk, H., Poortinga, Y. H., & Verhallen, T. M. M. (2004). Response styles in rating scales. *Journal of Cross-Cultural Psychology*, 35(3), 346.
- Watkins, D., & Cheung, S. (1995). Culture, Gender, and Response Bias: An Analysis of Responses to the Self-Description Questionnaire. *Journal of Cross-Cultural Psychology*, 26(5), 490-504.
- Weijters, B., Geuens, M., & Schillewaert, N. (2010). The stability of individual response styles. *Psychological Methods*, 15(1), 96-110.