

A kvanti-kvali áldilemmán túl

Koltai Júlia, Sik Endre, Simonovits Bori

koltajuli@tat.k.elte.hu; sik@tarki.hu; simonovits@tarki.hu

Beérkezés: 2014. 10. 08.

Átdolgozott változat beérkezése: 2015. 02. 12.

Elfogadás: 2015. 02. 22.

ÖSSZEFOGLALÓ: Írásunk célja, egy szempontrendszert kínálni azokra az esetekre, amikor a szociológus módszert választ kutatása tárgyához. A cél megfogalmazása során a belső és külső érvényességre, továbbá a megbízhatóságra törekvést tartjuk elsődlegesen fontosnak. A módszerek közti mérlegelés gyakorlati megvalósulását egy konkrét kutatási témán (a diszkrimináció elterjedtségén) mutatjuk be. A módszerek kiválasztásának dilemmájában külön figyelmet szentelünk a kontrollált kísérletnek és a nem résztvevő megfigyelésnek, mivel úgy véljük, ezek a viszonylag ritkán alkalmazott módszerek jól ötvözik a korábbiakban bemutatott szempontok előnyeit. Mivel úgy tapasztaljuk, a módszerválasztás dilemmáját – véleményünk szerint helytelenül – napjainkban leginkább a kvali-kvanti kevert módszertannal oldják fel a kutatók, írásunk végén érvelünk e gyakorlat ellen; továbbá amellett, hogy a módszerek egy dimenzió mentén történő besorolása helyett, komplexebb osztályozást érdemes alkalmazni.

Kulcsszavak: módszertan, külső érvényesség, belső érvényesség, megbízhatóság, kísérlet, megfigyelés

Bevezetés

Noha szisztematikus kutatást nem végeztünk, de elsősorban hazai és nemzetközi kutatások közreműködőiként, másodsorban oktatóként, harmadsorban számos konferencia és műhelybeszélgetés résztvevőiként szerzett tapasztalataink azt sejtetik, hogy

- (1) a szociológiai kutatások a kutatási design megtervezésekor szinte gondolkodásmentesen (rutinból) választják a surveyt vagy az interjút, s
- (2) mintha egyre gyakoribb lenne (sokszor a megfelelő elemzési tudás vagy a pénzhiány miatt) a kvalitatív módszerek választása a kvantitatívak rovására.

Írásunkban arra teszünk kísérletet, hogy a társadalom megismerésére törekvő szociológus segítsük annak végiggondolásában, hogyan válogasson egy kutatás megtervezése során a szóba jöhető sokféle módszertani megoldás közül. Másként szólva, egy lehetséges szempontrendszert szeretnénk kínálni ahhoz, hogy a kutatási elrendezés (design) kidolgozásakor milyen megfontolások vezessék a társadalom kutatóját annak érdekében, hogy munkája eredménye minél hasznosabb legyen a tudományos megismerés szempontjából.¹

¹ Tudományos megismerés alatt jelen esetben azt értjük, amit a klasszikus szociológia is a társadalomtudomány feladatának tekint: a társadalom folyamatainak megértését oly módon, hogy az így kapott tudás általánosítható és más kutatások eredményeivel összehasonlítható legyen.

Ezt a célt – ahogy a későbbiekben majd részletesebben kifejthetjük – felfogásunk szerint csak úgy lehet elérni, ha a kutató törekszik arra, hogy a kutatás során alkalmazott mérés és a megfogalmazott következtetések a lehető legmagasabb szintű érvényességet (külső és belső), valamint megbízhatóságot érjék el. Erre természetesen csak törekedni lehet, mivel a természettudományi mérés alanyaival ellentétben elemzésünk alanyának, az embernek olyan tulajdonságai vannak, amelyek elvileg is lehetetlenné teszik ennek e két célnak a teljes mértékben és egyidejűleg való elérhetőségét. Az ember ugyanis a Higgs-bozonnal szemben felejt, a nátriumatommal ellentétben hazudhat és a háromszöggel szöges ellentétben jellemző rá a percepció torzítás.²

A tanulmány gondolatmenete a következő: előbb bemutatjuk azt a keretet, amiben egy társadalomkutató össze tudja hasonlítani és a kérdésfelvetése szempontjából értékelni képes a rendelkezésére álló módszereket. Ezután a diszkrimináció elterjedtségének kutatása kapcsán bemutatjuk a módszerek összehasonlító értékelésének egy lehetséges menetét, illetve a kontrollált kísérlet és a nem résztvevő megfigyelés, mint ritkán használt, de szerintünk ígéretes módszerek értékelésének módját. Végül visszatérünk a címben szereplő kérdéshez, s azokat az érveinket soroljuk fel, amelyek miatt elvetendőnek tekintjük a módszerek összehasonlító értékelésére a „kvanti-kvali” dilemmára leegyszerűsítő eljárást, illetve ennek a dilemmának kevert módszerrel való hamis megoldását.

Milyen feltételeknek kell megfelelnie egy szociológiai kutatás módszerének?

Az alkalmazható módszerek összehasonlítását három dimenzió mentén (ezek a megbízhatóság, a belső és a külső érvényesség) látjuk célszerűnek. Kiinduló tételünk az, hogy egy adott szociológiai probléma megoldására lehetőségeinkhez mérten olyan módszert kell választanunk, amely e három dimenzióban egyaránt és egyidejűleg a lehető legmagasabb szintet biztosítja.³

A mérés *megbízhatósága* alatt jelen esetben a mérés konzisztenciáját (Drost 2011: 106) értjük, vagyis hogy egy megismételt mérés mennyire adja ugyanazt az eredményt. Ez a kutatások összehasonlíthatósága szempontjából kiemelkedően fontos, hiszen – bár erről egy kvantitatív kutató nem szívesen beszél – erre a gondolatra alapozzuk a nemzetközi (vagy/és idősoros) kutatássorozat során az azonos kérdésekre adott válaszok összehasonlítását. Ha ugyanis eltéréseket találunk, azt a két ország közötti eltérésekkel (illetve az időközben végbement változásokkal) magyarázzuk, remélve, hogy mérésünk megbízható volt (vagyis a megkérdezettek mindenhol és mindig ugyanazt értették a kérdésen).

2 Ugyanakkor vegyük észre, hogy ha olyan adatbázisokra tudunk szert tenni, amelyek nagy tömegű ember viselkedéséről kérdőívvezés vagy interjúzás nélkül nyújtanak tudást, akkor a társadalomtudomány megkerülheti (csökkentheti) a fenti torzításokból fakadó korlátokat. Lásd a készen kapott kísérleteket (Horzsa–Sik 2012), illetve a Big Data adatbázisokat (pl. Barabási 2010).

3 Itt kell hangsúlyoznunk, hogy az érvényesség elemzése nem csupán a mérés (ahogy elemzésünk során tenni fogjuk), hanem a következtetések esetében (amivel itt nem foglalkozunk) is fontos. Az érvényesség e két típusa általában összefügg: azoknak a méréseknek, melyeknek magas a külső érvényessége (pl. reprezentatív survey), a következtetések szintjén is magas lesz a külső érvényessége. Ugyanez igaz a belső érvényesség esetében is: a mért változók belső érvényessége a következtetések esetében is növeli annak esélyét, hogy a változók között oksági összefüggést kereshessünk.

A *belső érvényesség* alatt (amit e cikkben érvényességként emlegetünk) azt értjük, hogy a kutatás során ténylegesen azt mérjük-e, amit mérni akarunk (Shadish et al. 2001: 20). Egy mérés érvényességét sok szempont alapján értékelhetjük. Ilyen például a *gyakorlati érvényesség*, amelynek lényege, hogy a kidolgozott mérési módszerünk megfelel-e a mérni kívánt fogalom valódi jelentésének, vagyis valóban azt méri-e, amit mérni akartunk.⁴ A *szerkezeti érvényesség* azt vizsgálja, hogy elméleti feltevéseink szerint viselkedik-e a mérőszám akkor, amikor azt vizsgáljuk, hogyan függ össze más változókkal. A *terjedelmi érvényesség* az érvényességnek azon aspektusát teszteli, hogy vajon mérésünk a mérni kívánt fogalomnak minden részét lefedi-e.⁵ Bármely szempont szerint vizsgáljuk tehát az érvényességet, mindig arra keressük a választ, hogy sikerült-e hitelesen megragadnunk a mérés során annak a jelenségnek a lényegét, amit mérni eredetileg célunk volt.

Végül a *külső érvényesség* azt mutatja, hogy mérésünk eredménye mennyire általánosítható (Shadish et al. 2001: 21–22), amit itt kettős értelemben használunk. Egyrészt azt vizsgáljuk, hogy a mérés eredménye a társadalom mely részére általánosítható s melyre nem. Ebben a megközelítésben a külső érvényességet a minta összetétele korlátozza. Másrészt vizsgáljuk, hogy eredményeink mennyire alkalmasak a viselkedés *előrejelzésére* a való életben.

A megbízhatóság, a belső és a külső érvényesség egyidejű megvalósulását a céllovás segítségével szemléltetve⁶ azt mondhatjuk, hogy a belső érvényesség maximuma az, amikor a céltábla közepét találjuk el (vagyis azt érzük el, amit akartunk), a megbízhatóság maximuma az, amikor mindig ugyanoda lövünk (vagyis az ismétlés mindig sikeres), végül a külső érvényesség akkor maximális, ha a jó céllovó egyben kiváló vadász is. Ebben a hasonlatban tehát az a jó mérés, amikor Robin Hood mindegyik lövése a szarvas szívének közepébe talál.

Vegyük észre, hogy ha egy mérés az egyik dimenzió alapján mégoly tökéletes is, ez nem jelenti azt, hogy a mérés eredménye torzítatlan lesz, hiszen a másik két dimenzióban ettől még (sőt nagy valószínűséggel) rosszul teljesíthet. Éppen ezért tartjuk fontosnak a három dimenzió szerint együttesen értékelni a méréseket. Azt állítjuk, hogy a kutatási elrendezés kialakítása során helyesebb olyan módszert választani, amely mindhárom dimenzió esetében „elég jó”, mint egy olyan módszert, amelyik egyik dimenzió mentén tökéletes ugyan, de a másik két metszetben rosszul teljesít.

A három értékelési szempont összefüggése

Az 1. ábrán az eddig tárgyalt három dimenzió összefüggését mutatjuk be két szélsőséges példa segítségével.⁷ Az ábrán „B”-vel jelölt mérés különlegessége, hogy elvben mindhá-

4 Például az antiszemitizmus mérésekor nem a kért igazságszággal kapcsolatos attitűdjeire kérdezzük-e rá.

5 Például az antiszemitizmus mérésekor ennek minden dimenziójára (pl. politikai és vallási antiszemitizmus) rákérdeztünk-e, illetve lefedi-e mérésünk az antiszemitizmus kognitív, affektív, illetve diszkriminatív dimenzióit egyaránt (Enyedi et al. 2004). A szerkezeti és terjedelmi érvényességről lásd Babbie (2008).

6 Babbie (2008: 167) példáját továbbfejlesztve.

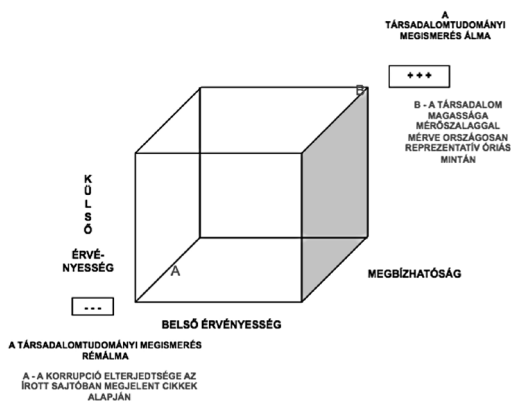
7 Elemzésünkben mindig az egyes módszerek alkalmazása során célként kitűzhető (vagyis a módszer lényegéből következő) lehetséges kimenetekre figyeltünk, arra, hogy mire lehet (s nem lehet) alkalmas az adott módszer használata. Minden esetben elvonatkoztatunk a konkrét alkalmazás mikéntjétől: sosem vettük a módszerek értékelésénél tekintetbe, hogy lehet rosszul is elemezni, olvashatatlanul írni vagy a mégoly tökéletesen is választott módszer alkalmazása során hibázni. Másként, kockáinkban a módszerek ideáltípusait helyeztük el.

rom dimenzió mentén magas értéket kaphat, mintegy megvalósítva a szociológus álmát. Erre lehet példa a társadalom magassága, amely olyan releváns szociológiai kérdések megválaszolása során lehet fontos, mint a munkahelyi diszkrimináció vagy a siker esélye, esetleg a férfi-nő kapcsolat alakulása (lásd Steckel [2008] áttekintését).

Mivel a magasság mérésének belső érvényessége maximális (egy centiméter mindig egy centiméter hosszú),⁸ és a nagymintás survey megismétlése (a mintavétel hibáján túl) a megbízhatóság lehetséges maximumát jelenti, ezért csak a külső érvényesség magas értéke mellett kell érvelnünk.⁹ Úgy véljük, hogy az ilyen módon kapott eredmények jól általánosíthatók a teljes népességre, amire a minta vonatkozott, és a kapott összefüggések (a férfiak magasabbak a nőknél, a magasabb társadalmi osztályba tartozók az alacsonyabb osztályba tartozóknál, a magasság előny a munkaerőpiacon stb.) is általában előrejelzik a magasság várható társadalmi hatásait.¹⁰

Az előbbi mérés ellentéte – a szociológus rémálma (az 1. ábra „A”-val jelölt esete) – olyan példa lenne, amelyik a három dimenzió egyikében sem alkalmas a kutatásra. Egy ilyen rémálomhoz talán legközelebb a korrupció elterjedtségének a nyomtatott sajtóbeli hírekre támaszkodó kutatások állnak (Szántó et al. 2011). Ebben az esetben figyelembe véve a korrupció fogalmának meghatározhatatlanságát (továbbá a cikkek szerzőinek és interjúalanyainak ehhez való ismeretlen viszonyát), valamint a látencia ismeretlen, de nyilvánvalóan el nem hanyagolható mértékét, azt gondoljuk, hogy egy ilyen kutatás belső és külső érvényessége elhanyagolható, így legfeljebb megbízhatónak tekinthető.¹¹

1. ábra: A társadalomtudományi megismerés „álma” és „rémálma”

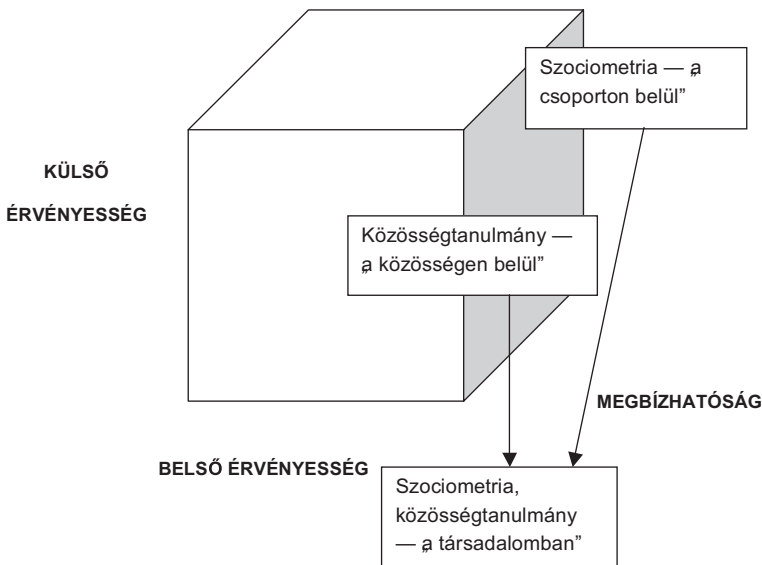


- 8 Természetesen tudjuk, hogy a testmagasság sem állandó (reggel, egészségesen és az első randevún magasabbak vagyunk, mint este, beteg és a házasság hetedik évében), így a felhozott példa inkább csak közelít az elméletileg maximális belső érvényességhez.
- 9 A félreértések elkerülése végett hangsúlyozzuk, hogy nem minden testi tulajdonságra igaz ez a jellemző: a szépség „mértéke” például nagyon sok társadalmi tényezőtől függ (hiszen időről időre másmilyen a szépségideál), de a kövérséget sem mérhetjük a magassághoz hasonló megbízhatósággal.
- 10 Természetesen nem feltételezzük, hogy mindenütt s azonos erővel érvényesülnek a fenti összefüggések, de amikor ezekkel ellentétes eredményekre jutunk, azt sosem a módszer külső érvényességének hiányával, hanem a társadalmi jelenségek összetettségével magyarázzuk.
- 11 Korábban is érveltünk már amellett, hogy vannak olyan tanulmányok, amelyek eredményeit figyelmen kívül kell hagyni, mert szerintünk komolytalan módszereken alapulnak. Ezért vetettük el például az informális gazdaság terjedelme mérésének egyes elemzéseit (Sik 1996), illetve „szóltunk be” a korrupció nagyságának mérésre alkalmazott (és széles körben elfogadott), de módszertanilag megkérdőjelezhető módszerek alkalmazóinak (Sik 2000).

Az is belátható, hogy a három dimenzió közötti kapcsolatnak vannak olyan szabályosságai, amelyek ritka kivételként teszik lehetővé együttes maximalizálásukat. A megbízhatóság és a külső érvényesség párosa, illetve a belső érvényesség jó eséllyel kizárják egymást: ha egyetlen jelenséget akarunk nagyon alaposan megismerni, akkor a „semmiről fogunk megtudni mindent”, ha pedig egy jelenség rengeteg változatát akarjuk megismerni, akkor „mindenről fogunk semmit tudni”. Ugyanakkor a megbízhatóság a külső érvényességgel gyakran együtt jár: bizonyos értelemben ez a két dimenzió át is fed egymással, amennyiben mind a kettő „kifelé irányul”, az elemzés általánosíthatóságára vonatkozik.

Hogy érvelésünk ne legyen áttekinthetetlen, ezért a tanulmányban nem térünk ki minden, a társadalomtudományokban használatos módszerre. Nem foglalkozunk például a közösségkutatással vagy a szociometriával (s általában az antropológia és a hálózatkutatás sokféle módszerével), mert ezeknek a kutatásoknak jobbra nem célkitűzésük, hogy általuk megismerhető legyen a társadalom egésze. Ebből pedig az következik, hogy nem is helyezhetők el egyszerűen a mi tipológiánkban, mert helyzetük más akkor, ha a maguk elé kitűzött célt saját közegükre, s más, ha a társadalom egészére értelmezzük (2. ábra).

2. ábra: A közösségkutatás és a szociometria helyzetének változatai a kutatás céljának függvényében



Míg a két módszer az adott közösség vagy csoport megismerése szempontjából magas belső és külső érvényességet,¹² illetve a hálózatkutatások sztenderdizált technikai mi-

¹² Mi jellemzi az adott egységen belül a nem vizsgált folyamatokat, hogyan változik a struktúra stb.

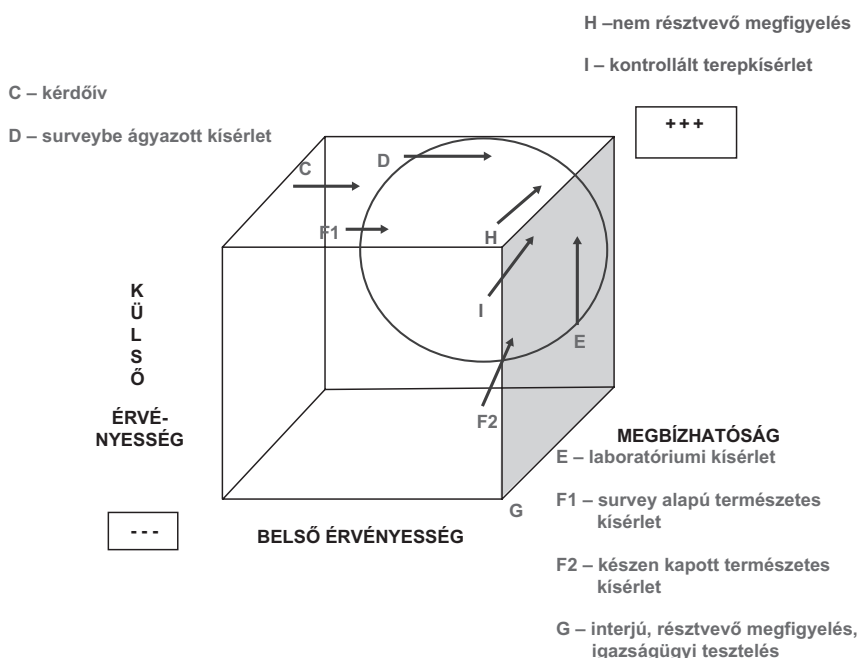
att magas megbízhatóságot jelent, addig ezen módszerek mégoly tökéletes alkalmazása sem alkalmas a társadalmi folyamatok megbízható és általánosítható megismerésére.¹³

Hogyan válasszunk módszert egy adott kérdés megválaszolásához? – A diszkrimináció elterjedtségének esete

Ebben a részben azt mutatjuk be, hogy a diszkrimináció elterjedtségének mérésére milyen módszerek állnak rendelkezésre, s ezek melyike és hogyan finomítható annak érdekében, hogy a belső és külső érvényesség, továbbá a megbízhatóság egyidejű növelésével a lehető legjobban megnöveljük előrejelzésünk minőségét.

A 3. ábra a diszkrimináció mérésére szolgáló módszereket mutatja be aszerint, hogy hogyan sorolhatók be e módszerek a mérés értékelési kritériumának három dimenziója alapján.

3. ábra: A diszkrimináció elterjedtségének mérésére szolgáló módszerek a három mérési kritérium szerint



Az ábrán látható kör (ami valójában persze gömb) a diszkrimináció elterjedtségének mérésére a három dimenzió mentén egyaránt jól alkalmazható megoldások halmaza. A nyilak az adott módszeren belüli lehetséges változatok, amelyek, noha az adott

¹³ Sőt paradox módon minél tökéletesebb egy közösségkutatás, annál kevésbé lehet az ott alkalmazott eljárást máshol alkalmazni, hiszen minden közösség megismeréséhez más módszertani ötvözetet kell alkalmazni.

módszer alapvető korlátait nem tehetik semmissé, de olykor nem elhanyagolható mértékben javíthatják a választott módszer minőségét.

Jól látható, hogy a körön belül helyeztük el azokat a módszereket, melyek mindhárom kritérium szerint „elégé” jól mérik a diszkrimináció elterjedtségét, s e körhöz közel helyeztük el azokat, amelyek oly módon finomíthatók, hogy mindhárom dimenzió mentén jónak tekinthessük őket. E módszerek közé soroltuk a nem résztvevő megfigyelést (H) és a kontrollált kísérletet (I), mivel e módszerek esetében képzelhető el a magas szintű belső érvényesség megvalósítása mellett a külső érvényesség és a megbízhatóság egyidejű növelése. A többi módszer véleményünk szerint e körön kívül esik:

vagy azért, mert belső érvényességük alacsony (a diszkrimináció elterjedtségét mérő kérdőíves kutatás (C) jelentősen torzít, mivel nem a diszkrimináció valós szintjét méri, hanem attól függően, hogy az áldozat vagy az elkövető tapasztalatára, illetve percepciójára hagyatkozunk, jelentősen alul-, illetve felülbecsli a diszkrimináció elterjedtségét (Sik–Simonovits 2009);

vagy azért, mert belső érvényességük ugyan magas (interjú, résztvevő megfigyelés, igazságügyi tesztelés, lásd G), de megbízhatóságuk és külső érvényességük elhanyagolható, hiszen céljuk nem a jelenség elterjedtségének, hanem egy-egy esetnek a lehető legmélyebb megismerése. Ennek tudható be az is, hogy a 3. ábrán e módszerek mellett nem szerepel a jobb felső irányba mutató nyíl sem;

a kimaradt módszerek közül a kérdőívbe ágyazott (D), a laboratóriumi (E) és a természetes kísérleteket (F1, F2) pedig „határeseteknek” tartjuk. Az elsőt korlátozott belső érvényessége, a laboratóriumi kísérletet alacsony szintű külső érvényessége¹⁴ miatt. A természetes kísérleteket két csoportra bontottuk aszerint, hogy egy korábban végzett survey adatainak átrendezése (F1, például Alessina–Fuchs–Schündeln 2005), vagy nem kutatási célból felvett (tehát készen kapott) adatbázis (F2, például Edward et al. 2008, illetve a Horzsa–Sik-tanulmányban [2012] szereplő három diszkrimináció elterjedtségét becslő természetes kísérlet szolgál) a kísérleti elrendezés alapjául. E kísérleti módszereknek a 3. ábrán való elhelyezkedése, illetve a nyílak azt jelzik, hogy ezek olyan mérési megoldások, amelyek kivitelezése során egy alkalmas kutatási elrendezés érdemben javíthat az alkalmazott módszer gyengeségein, miközben megőrizhető eredendő előnyük.¹⁵

Az alábbiakban a „körön belül” elhelyezkedő két mérési módszert alkalmazását mutatjuk be a diszkrimináció elterjedésének területén, majd azt vizsgáljuk, hogyan finomíthatók ezek a módszerek annak érdekében, hogy megbízhatóságuk és külső érvényességük a lehető legmagasabb legyen, s ez ne menjen a belső érvényesség rovására.

14 A laboratóriumi kísérletek külső érvényességét ugyanakkor a kísérleti elrendezés javításával (pl. az alanyok reprezentatív beválogatásával), illetve surveyvel és kontrollált kísérletekkel való kombinálásával erősen növelni lehet.

15 A survey alapú természetes kísérletek (F1) elemzésének általánosíthatóságát növelni lehet (a survey belső érvényességét növelő mindenféle megoldáson felül), ha a különféle surveyket több országban és időpontban megismételjük, s ugyanez tehető meg a készen kapott természetes kísérletek esetében is (F2), amennyiben ezekhez hasonló adatbázisok más országokban is fellelhetőek, illetve ha az elemzést más jelenségek adatbázisain is lehet alkalmazni.

A kontrollált terepkísérlet alkalmazása a diszkriminatív viselkedés elterjedtségének becslésére

A szakirodalomban *diszkriminációtesztelésnek* (az angolszász irodalomban *situation testing* vagy *discrimination-audit*) nevezett módszer lényege a mindennapi élet legkülönbözőbb területein – állásszerzés, lakásbérlés, igazságszolgáltatás, rendőrségi igazoltatás, szolgáltatás stb. – felmerülő diszkrimináció előfordulási gyakoriságának vizsgálata tesztelő párok segítségével¹⁶. Tehát ha például egy roma és egy nem roma álláskereső jelentkezik ugyanarra az állásra, állásszerzésük esélyét sokszor és azonos módon ismételt kísérleti helyzetek összehasonlításával mérjük.

A *tesztelés lényege* a kontrollált kísérleti elrendezés, amely a *homogenizálás* és a *randomizálás* technikájának együttes alkalmazásával valósulhat meg. Fontos hangsúlyozni, hogy a homogenizálás és a randomizálás követelményének nemcsak a tesztelők, hanem a teszthelyzet esetében is meg kell felelni ahhoz, hogy a kontrollált elrendezés megvalósulhasson. A homogenizálás és a randomizálás a gyakorlatban egymásra épülő folyamat, ahol első lépésként a homogenizálást kell megvalósítanunk (ez a gyakorlatban azt jelenti, hogy kijelöljük a kutatási témánk szempontjából releváns populációt és mintavételi keretet, ahol a diszkrimináció elvileg előfordulhat), majd ezt követi a randomizálás (ezen belül a teszthelyzetek véletlenszerű kiválasztása). A véletlen elrendezésre épülő kontrollált kísérleti elrendezés biztosítja, hogy az esetlegesen tapasztalt diszkriminációt nagy valószínűséggel a kísérleti változóval magyarázhatjuk.

A diszkriminációtesztelés módszerét a megvalósítás *típusa* szerint három csoportba soroljuk aszerint, hogy a tesztelés során a tesztelő (1) személyesen; (2) a „hangjával” (telefonos tesztelés); vagy (3) virtuálisan, egy szöveg által megszemélyesítve (írásos tesztelés) vesz részt.

A diszkrimináció elterjedését vizsgáló *kontrollált kísérlet* lényege, hogy nem az egyes esetek minden részletre kitérő megismerésére szolgál – szemben az igazságügyi teszteléssel –, hanem a nagyszámú és kontrollált helyzetben megismert viselkedés valószínűségei alapján teszi becsülhetővé a diszkrimináció bekövetkezésének valószínűségét. Tehát annak ellenére, hogy az egyes esetekben nem tudhatjuk pontosan, történt-e diszkrimináció, s ha igen, akkor ez ízlés alapú vagy statisztikai diszkrimináció volt-e (Sik–Simonovits 2012), a nagyszámú és jól elrendezett mérés statisztikai tesztje ezeket az eseti bizonytalanságokat kiszűri. Másként, a tesztelés sikeres megvalósításának kulcskérdése a kontroll, ami a homogenizáció és a randomizáció együttes alkalmazásával érhető el.

A különböző típusú kísérletekben eltérő módon biztosítható a kontroll, továbbá a homogenizáció és a randomizáció megvalósítása is eltérő módszertani problémákat vet fel. A személyes, telefonos és írásos tesztelés megkülönböztetése a védett tulajdonság megjelenítése szempontjából kulcsfontosságú, mert mindhárom tesztelési típus más és más megbízhatósági, illetve belső és külső érvényességi korlátokat hordoz magában.¹⁷

¹⁶ A hazai szakirodalomban lásd Sik – Simonovits (2008), Simonovits (2010, 2012a, 2012b).

¹⁷ Mindez függ természetesen a védett tulajdonság jellemzőitől is: míg a férfiak és a nők megkülönböztetése könnyen és egyszerűen megvalósítható, addig például a származás megjelenítése korántsem ilyen egyszerű feladat.

A tesztelő védett tulajdonságának megjelenítése (vagyis hogy a „kísérleti inger” a beszélgetőpartner tudomására jut-e) egyértelműen a személyes tesztelésnél a leg-sikeresebb: itt ugyanis a tesztelő a neve, hangja és beszédstílusa mellett különböző nonverbális jegyekkel, valamint teljes külső megjelenésével (testi jegyek, ruházkodás stb.) is közvetítheti a tesztelői profilt. A telefonos tesztelésnél a védett tulajdonság felismerhető a tesztelő nevéen és hangján keresztül: ebben az esetben, mivel nem látja egymást a két beszélgetőpartner, a nyelvezet és a beszédstílus mellett nagyobb hangsúlyt kaphat a név elhangzása, a hangszín, a beszéd sebessége és nem utolsósorban az érthetősége, viszont a nonverbális kommunikáció elemei csak korlátozottan tárhatók fel.¹⁸ Az írásos tesztelésnél a védett tulajdonságot a jelentkező neve és írása jeleníti meg. Az írás alapján számos tulajdonságot bemutatathatunk. Operálhatunk a korlátozott és kidolgozott kóddal (Bernstein 1971): ha társadalmi státust akarunk tesztelni, szándékosan elhelyezett nyelvi és helyesírási hibákkal tarkíthatjuk a leveleket, illetve dönthetünk a kézzel írt levélforma mellett is, amely ismét szélesíti a beavatkozási lehetőségeinket. Az önéletrajzos teszteléssel lehetséges az iskolai végzettség, de sok más személyes jellemző (lakcím, hobbi, iskola neve stb.) felhasználása is a védett tulajdonság lehető legpontosabb azonosítására (Sik–Simonovits 2012).

A teszthelyzet kontrollálása szempontjából a legfontosabb egyrészt maguknak a helyzeteknek a megfelelő kiválasztása (a mintavétel), másrészt a kísérlet menetének (egységes megfigyelési szabályok, tesztkérdőív, beszédpanelek stb.) részletes kidolgozása. Az első esetben a legfontosabb elv, hogy olyan helyzetekben teszteljünk, ahol a diszkrimináció valós és elterjedt probléma lehet, valamint hogy az általunk létrehozott helyzet „életszerű” legyen.¹⁹ Fontos továbbá, hogy a tesztelés típusa illeszkedjen a tesztelt helyzethez: azaz például ha írásban teszteljük a munkaerő-piaci diszkriminációt, akkor olyan munkakörökre szűkítsük a mintavételi keretet, ahol tipikus az önéletrajzküldés (homogenizálás), és ezután valósítsuk meg a véletlen mintavételt (randomizálás).

Ha már kiválasztottuk a tesztelendő helyzetek körét, második lépésben a helyzetet magát kell kontrollálnunk, amiben tesztelünk. Az egységes megfigyelési szabályok minden esetben kiemelten fontosak, de talán a személyes tesztelésnél van a legnagyobb szerepük: itt akár külső megfigyelőket is használhatunk a még pontosabb dokumentáció érdekében (pl. szórakozóhelyek beengedési gyakorlatának tesztelésénél). A beszédpanelek kidolgozása és standardizálása a személyes és telefonos teszteléseknél egyaránt fontos és nagy jelentőségű például a bemutatkozás vagy a védett tulajdonság közlése szempontjából.

Összehasonlítva a három tesztelés típus sajátosságait, elmondható, hogy míg a *belső érvényesség* a személyes tesztelés esetében a legmagasabb szintű (mivel itt a leg-

¹⁸ Ugyanakkor lehetséges a beszélgetéspanelbe illesztett információk segítségével fokozni a védett tulajdonság megjelenítésének egyértelműségét (Otlakán 2012).

¹⁹ Ha például az asztalfoglalás esetén felmerülő diszkriminációt akarjuk tesztelni, akkor olyan éttermekből vegyünk mintát, ahol az asztalfoglalás tipikus jelenség. Ezzel természetesen korlátozzuk a mintavétel körét és kizárjuk a legalacsonyabb kategóriájú vendéglátó-ipari helyeket a mintavételi keretből, ami csökkenti a tesztelés eredményeinek külső érvényességét.

esélyesebb a védett tulajdonságnak a tesztalany számára egyértelmű megjelenítése) és az írásos tesztelés esetében a legalacsonyabb (hiszen bizonytalan, hogy sikerült-e a védett tulajdonságot minden résztvevő számára azonos módon megjeleníteni), addig a *megbízhatóság* és a *külső érvényesség* ezzel ellentétesen alakul: legmagasabb szintű az írásos tesztelés és legalacsonyabb a személyes tesztelés esetében. A megbízhatóság növelésének az szab határt, hogy a teszthelyzetek kontrollálhatósága véges, s nem minden esetben valósítható meg azonos módon. A külső érvényességbeli eltéréseknek is gyakorlati oka van: egyrészt korlátozott a megvalósítható tesztelések száma, másrészt nem egyszerű a reprezentativitás elérése. Ezek a feltételek ugyanakkor az írásos teszteléseknél valósíthatók meg leginkább, a személyes tesztelések esetében a legkevésbé, amikor a kutatás megvalósítása során nagyon sok kompromisszumra kényszerül a kutató.

Meg kell jegyezni azonban, hogy a mintavételi keret torzító hatása majdnem minden tesztelési típusnál felmerül, így a mérés külső érvényessége gyakran gyengül. Akármilyen alaposan veszünk mintát például az elektronikus vagy a nyomtatott sajtóban rendelkezésre álló álláshirdetésekből, a tesztelés külső érvényessége korlátozott lesz, mivel azokat az állásokat, melyekhez informális úton keresnek csak jelentkezőket, ezen a módon nem tudjuk elérni.²⁰

A nem résztvevő megfigyelés alkalmazása a diszkriminatív viselkedés feltérképezésére

Bár egyes szerzők a megfigyelést alapvetően a kulturális antropológia módszereként mutatják be (pl. Crapo 1990: 7), a módszertani tankönyvek általában a különböző társadalomtudományos diszciplínák – az antropológia, a néprajz és a szociológia – határterületén helyezik el (pl. Héra és Ligeti 2005: 116–118), illetve a terepkutatások között tárgyalják a megfigyelések különböző típusait (Babbie 2008). Babbie a terepkutatás gyűjtőfogalmát használja a résztvevő megfigyelés, a nem résztvevő megfigyelés és az esettanulmány módszereinek összefoglalására, melyeknek legfőbb előnye, hogy természetes közegükben figyelhetjük meg a társadalmi jelenségeket (uo.: 304–310).

Tanulmányunkban azokkal a nem résztvevő megfigyelésekkel foglalkozunk, melyek hipotézis tesztelésére szolgálnak, azaz konfirmatív célúak.²¹ Az általunk bemutatott kutatások célja a diszkrimináció szempontjából lényeges elemek strukturált megfigyelése az elkövető és az áldozat interakciója során. Ez a három módszertani ismérv értelmében azt jelenti, hogy a lehetséges megfigyelési megoldások közül azokat használjuk, melyek a belső érvényesség rovására feladnak valamennyit a megbízhatóság és a külső érvényesség növelése céljából.

20 Csapán feltételezhetjük, hogy az ilyen, informális elemekben dús helyzetekben nagy lehet a szerepe a szubjektív (tehát diszkriminációra hajlamosító) munkáltatói döntéseknek. Kutatói döntéstől függ az is, hogy hogyan kezeljük azokat a hirdetéseket, melyek eleve diszkriminatív kitételeket tartalmaznak (pl. „fiatal, csinos titkárnőt keresünk”), ugyanis ha kihagyjuk az ilyeneket, akkor alulbecsüljük a diszkrimináció mértékét.

21 Nem foglalkozunk az antropológiában használatos megfigyelésekkel, melyek célja egy-egy eset, közösség alapos megismerése, vagy az exploratív, feltáró jellegű, mintázatok kereső megfigyelésekkel, melyek előzetes elképzelés nélkül vizsgálják a kutatás tárgyát.

A diszkriminációmérés céljából alkalmazott megfigyeléses módszerre jó példa a moszkvai metrókutatás (OSI 2005), melynek során a kutatás fő célja az igazoltatáskor tetten érhető etnikai alapú kiválasztás (*ethnic profiling*)²² gyakorlatának feltérképezése volt; az aránytalan igazoltatási gyakorlatot a nemzetközi és az európai jog egyaránt tiltja.²³ A kutatás során öt hónapon keresztül rögzítették 15 moszkvai metróállomáson áthaladó csaknem harmincnégyezer utas etnikai hovatartozását, korát és nemét, valamint a kijáratoknál megállított több mint 1500 fő igazoltatásának körülményeit. Mivel a megfigyelőknek az etnikai csoportok biztos azonosítására nem volt lehetőségük, ezért a szláv/nem szláv egyszerűsítést alkalmazták a kódolás során. A kutatás legfontosabb megállapítása, hogy a nem szláv származásúaknak több mint 21-szer nagyobb volt az esélyük arra, hogy igazoltassák őket a szláv származásúakhoz – vagy legalábbis azokhoz, akik annak néznek ki – képest.

Látható, hogy a nem résztvevő megfigyeléses módszer megfelelő alkalmazásával, az esélyhányados segítségével becsülhető és más kutatások eredményeivel összehasonlítható az etnikai szempontú kiválasztási gyakorlat elterjedtsége. A megfigyelések számszerű adatainak értelmezésekor rendkívül fontos a megfelelő viszonyítási alap megtalálása (*benchmarking*).²⁴ Azaz annak tisztázása, hogy ha például egy kisebbségi csoport aránytalan kiválasztási gyakorlatát akarjuk vizsgálni, akkor azt kell meghatároznunk, hogy milyen népességhez viszonyítsuk a megfigyelés során kapott adatokat.²⁵

Ezek után összehasonlítva a diszkrimináció mérésére szolgáló különböző kísérlettípusokat és a nem résztvevő megfigyelést, a belső, a külső érvényesség és a megbízhatóság szerint azt találjuk, hogy a módszerek közötti legfőbb különbség a valóságközeliségre törekvés mentén ragadható meg: ebből a szempontból a kontrollált terepkísérlet, a természetes kísérlet és a nem résztvevő megfigyelés *természetes közegben*,²⁶ a másik két kísérlettípus (a laboratóriumi és a surveybe ágyazott) *mesterséges közegben* valósul meg.²⁷

22 „Az etnikai profilalkotás, azaz az etnikai vagy faji hovatartozás ismérvein alapuló kiválasztás, akkor valósul meg a rendészeti gyakorlatban, ha az intézkedő rendőrök külső jegyek alapján határozzák meg, kit igazoltatnak, illetve kívül szemben fogantatásnak valamilyen intézkedést” (Kádár et al. 2008: 3).

23 A Rasszizmus és Intolerancia Elleni Európai Bizottság (European Commission against Racism and Intolerance, ECRI) 11. számú általános ajánlásában megfogalmazza, hogy a tagállamok kormányai „az etnikai alapú kiválasztás tekintetében egyértelműen definiálják és törvényben tiltják meg az etnikai ismérveken alapuló kiválasztást”.

24 A viszonyítási alap a moszkvai metrókutatás során az utazóközönségben megfigyelt, nem szlávként azonosítottakhoz képest volt.

25 A rendőrség közúti ellenőrzési gyakorlatának elemzése során például a probléma az, hogy még ha van is információnk az egyes etnikai csoportok népességen belüli arányáról, nem ismerjük sem az egyes etnikai csoportok közlekedési szokásait, illetve autótulajdonlási és használati eloszlásait, sem a valós szabálysértési arányokat, ahogy annak arányát sem, hogy külső jegyek alapján hány embert sorolnának az adott csoport tagjának a teljes társadalmon belül. Lásd Engel és Calnon (2004) írását, mely a rendőrségi igazoltatási gyakorlatokban az etnikai profilalkotás igazolásához (vagy cáfolásához) használható különböző adatgyűjtési módszerek hasonlítja össze.

26 Vegyük észre továbbá, hogy a kontrollált terepkísérlet és a nem résztvevő megfigyelés között további átfedések is vannak: a kontrollált kísérletnél (ha erre mód van) is lehet nem résztvevő megfigyelőt alkalmazni (pl. amikor egy külső megfigyelő követi az állásinterjú vagy a diszkóba beengedés folyamatát), ugyanakkor a nem résztvevő megfigyelés lényeges eleme, hogy a szituáció egyes elemeit kontrollálni lehet.

27 A mesterséges szituációban végzett mérés a természetes helyzetűnél markánsabb eredményeket mutathat. Egy kutatás azt bizonyította, hogy azonos inger hatására a surveybe ágyazott kísérletekben résztvevők véleménye jelentősebb mértékben mozdult el, mint a természetes kísérlet alanyaié (Barabas–Jerit 2010). A különbség oka a szerzők szerint az, hogy míg a kutató által megtervezett kísérleti szituációban mesterségesen tiszta körülmények között kapja meg a kísérleti alany az ingert, a valóságban ez a fajta sterilitás nincs jelen. Fontos különbség továbbá, hogy míg a mesterségesen létrehozott kísérleti szituációban a véletlen elrendezés dönt arról, hogy kit érjen az inger, addig a valóságban csak azok figyelnek oda egy bizonyos hírre, akiket érdekel az adott téma.

Mivel a kísérleti, illetve kontrollcsoportba sorolódás a természetes kísérletekben (éppúgy, mint a valóságban) nem véletlenszerűen történik, ebből adódóan jelentős torzítással (*selection bias*) kell számolnunk. Ráadásul a természetes közegben sokkal több a zavaró körülmény, amely elterel(het)i a kísérleti alany figyelmét, továbbá az egyénre ható ingerek sokaságából nem tudhatja biztosan a kutató, hogy a kísérleti alanyt mely ingerek érték el. Tehát a természetesség „ára” a csökkenő megbízhatóság és belső érvényesség.

Ugyanakkor a surveybe ágyazott kísérlet és a laboratóriumi kísérlet megbízhatósága egyaránt magas, de az előbbi legnagyobb problémája, hogy ez a módszer nem képes a valóság leképezésére, ezért a belső érvényessége alacsony. Ennek ellentéte a laboratóriumi kísérlet, ahol a belső érvényesség magas (hiszen a kutató „uralja” a maga által létrehozott „pseudeovalóságot”), de a külső érvényessége alacsony.²⁸

A sterilitás mellett a bemutatott módszerek közötti másik fontos különbség nem a méréshez, hanem a következtetések természetéhez kapcsolódik. Míg a megfelelően megvalósított kontrollált kísérletek esetében a kísérleti változó, illetve annak időbeli hatása a többi tényező hatásától elkülöníthető, és éppen ezért a kísérleti és a kontrollcsoport közötti eredménykülönbségekből *okási* következtetéseket vonhatunk le, a nem résztvevő megfigyelés és a többi módszer esetében nem teljesülnek az okási következtetéshez szükséges feltételek, ezért ezeknél csupán együttjárásra épülő következtetésről beszélhetünk.^{29, 30}

Mivel a fejezet központi állítása, hogy a kontrollált terepkísérlet és a nem résztvevő megfigyelés alkalmasak leginkább a diszkrimináció elterjedtségének előrejelzésére, ezért a továbbiakban e két módszer legfőbb előnyeit és hátrányait foglaljuk össze a többi módszerhez képest, és arra keressük a választ, hogyan finomíthatók ezek a módszerek a külső és belső érvényességük, illetve megbízhatóságuk növelése érdekében.

A példaként hozott módszerek három dimenzió alapján történő értékelésének összefoglalása

A kontrollált terepkísérletek legfőbb előnye egyrészt, hogy mivel a kísérleti hatást a kutató tervezi meg, ezáltal olyan hatások is tesztelhetők, melyek természetes közegben

Továbbá nem feledkezhetünk meg az ún. keretkezési hatásokról (framing effect), melyek a döntési probléma különböző megfogalmazásából erednek, és mind a laboratóriumi, mind a surveybe ágyazott kísérleti módszereknél jelentős belső érvényességi dilemmákat vetnek fel (Szántó–Tóth 2003).

28 A sterilitás problémájával például a játékelméleti dilemmák kérdőívbe ágyazásánál is számolni kell. Kérdéses például, hogy a laboratóriumi körülmények között definiált kockázattal kapcsolatos attitűdöt a kérdőívben kockázatvállalási hajlandóság által mérve mennyiben képes a kutató a tényleges viselkedés előrejelzésére, tehát hogy valaki rendelkezik-e bizottsággal, illetve milyen feltételek között kötne biztosítást.

29 Részben kivétel ez alól a panelvizsgálat, azonban az előzetes és az utómérés között eltelt idő alatt történő események kétségessé teszik okási következtetés lehetőségét (confounding factors).

30 A természetes kísérletek esetében elrendezésfüggő az, hogy lehetséges-e okási modell létrehozása: ha megvalósul a kísérleti és kontrollcsoportba sorolás randomizált elrendezése, akkor okási következtetéseket tehetünk, ha viszont nem véletlenül sorolódtak be a kísérleti és kontrollcsoportba az alanyok, akkor – hasonlóan a megfigyelésekhez – statisztikai együttjárásra épülő következtetéseket tehetünk csupán (Gerber–Green 2012).

nem figyelhető meg, ráadásul a kísérleti hatás elemei sokféleképpen variálhatók.³¹ Másrészt, a kontrollált kísérlet esetében ki lehet szűrni a kísérleti elrendezést zavaró és a hipotézis tesztelése szempontjából irreleváns hatásokat.³² Emiatt e módszer belső érvényessége a társadalmi viselkedés előrejelzése szempontjából kiemelkedő.

A diszkrimináció elterjedtségének mérésére irányuló *nem résztvevő megfigyelés* legfontosabb *előnyei* egyfelől a megfelelő viszonyítási alap megtalálása esetén a standardizált mérőszámmal (esélyhányados) lehetséges becslés nagy megbízhatósága és külső érvényessége. Másfelől, mivel a kutatás valós körülmények között vizsgálja a diszkrimináció előfordulását, ezért a többi módszerhez képest árnyalt képet kaphatunk a diszkriminációt magában foglaló interakciók közegéről.

A kontrollált terepkísérletek és a megfigyelések legfőbb *hátrányai* az alábbiakban foglalhatók össze:

korlátozott azon jelenségek köre, amelyeket tesztelhetünk, illetve megfigyelhetünk (általában a nyilvános helyen és rövid idő alatt zajló jelenségek ilyenek), és limitált a tesztelhető ingerek köre is (főként megvalósíthatósági és etikai okok miatt);

a kontrollált terepkísérletek esetében további probléma, hogy nem lehetünk biztosak abban, hogy egy adott szituációban a kísérlet alanyát valóban elérte-e az a hatás, amit a tesztelés mérni hivatott;³³

torzítási problémát okozhat továbbá a kontrollált kísérletek esetén az összetétel és a lemorzsolódás torzító hatása, vagyis hogy akik valamilyen okból nem vesznek részt (végig) a kísérletben, azok esetleg a tesztelés releváns jellemzőiben eltérnek azoktól, akik (végig) résztvevői a kísérletnek.³⁴

A kontrollált kísérlet és a nem résztvevő megfigyelés előnyeinek és hátrányainak a figyelembevétele mellett fontos még felhívni a figyelmet arra, hogy a 3. ábrában e két megoldás mellett szereplő nyilak arra utalnak, hogy e két módszer hiányosságai legalább részben csökkenthetők anélkül, hogy belső érvényességük sérülne, mégpedig a következő módokon:

az adott korlátok közötti „véletlenebb” mintavétel növeli annak esélyét, hogy a kapott eredmények külső érvényessége magasabb legyen, s ugyanígy hat a nagyobb elemszám is. Mindez még tovább fokozható, ha hasonló, de nem azonos helyzeteket figyelünk meg, illetve tesztelünk;³⁵

31 Például abban a kutatásban, amikor a fekete és fehér bőrszint sejtető neveket a szerzők véletlenszerűen rendelték hozzá a különböző típusú fiktív életrajzokhoz (Bertrand–Mullainathan 2003).

32 Ilyen lehet például a személyes tesztelés esetében egyes kisebbségi csoportoknak a kisebbség más csoportjaitól való eltérő öltözködésének hatása. Például ha a muszlim nőkkel szembeni diszkriminációt teszteljük, ne homogenizáljunk fejkendőviselésre, mert ezzel a muszlim nők közül csak a vallásosabbakat jelentjük meg.

33 Például egy munkaerő-piaci diszkrimináció tesztelése során érzékelték-e a munkáltató, hogy a jelentkező roma, s hogy ez az inger szignifikánsan erősebb volt-e az adott szituációra ható egyéb ingereknél.

34 Hasonlóan a panelkutatáshoz, illetve a surveyhez: ha feltételezzük, hogy akik elzavarják a kérdezőt, illetve akik megszakítják az interjút, azok összetétele eltér az engedelmes válaszadóképtől olyan dimenziókban, amelyek semmilyen súlyozással nem javíthatók.

35 Például: nemcsak alacsonyan, hanem magasabban kvalifikáltak számára kiírt állásokat is tesztelünk, továbbá nemcsak a metrókijáratokban, hanem más nyilvános helyszíneken, illetve nemcsak Moszkvában, hanem New Yorkban és Párizsban is végzünk nem résztvevő megfigyelést.

- a nagyobb és „véletlenebb” minta eközben nem rontja a belső érvényességet, sőt: ugyan továbbra sem fogjuk azt tudni, hogy az egyes teszthelyzetekben pontosan mi befolyásolta a (nem) diszkrimináló munkaadókat, illetve az egyes megfigyelések mennyire voltak torzítottak a megfigyelő tévedései miatt, a nagyobb és „véletlenebb” minta az ebből fakadó belső érvényesség bizonytalanságát csökkenti, az eredmény általánosíthatóságát növeli;
- a megfelelően előkészített – melybe beletartozik a tesztelők, illetve megfigyelők tréningje, a tesztkérdőív, illetve megfigyelési napló kidolgozása és kipróbálása – és kivitelezett terepkutatással a mérés belső érvényessége fokozható, ami a kutatás sokszor történő megismétlése esetén a megbízhatóság növelésének záloga.³⁶

Mi tehát a megoldás?

Kezdjük azzal, hogy egy példa segítségével érzékeltetjük: a minket foglalkoztató kérdés nem társadalomtudomány-specifikus.

Tételezzük föl, hogy egy gazda meg akarja ismerni frissen vett földje termelési potenciálját. Hogyan tegye ezt? Fúrhat egy mély lyukat, hogy az abból vett minta alapján megismerje a föld kémiai összetételét. Ha jó mélyen fúr, még az altalaj összetételét is megismerheti. De megteheti azt is, hogy sok és eltérő helyen megkapargathatja a földjét, s az abból vett minták összetételét statisztikai módszerrel vizsgálja. Megteheti továbbá, hogy kivesz valamennyi talajt a földjéről, azt beviszi egy talajlaboratóriumba, s ott mindenféle kezelésnek veti alá, hogy lássa, miként reagál földje a különféle kezelésekre. De megteheti azt is, hogy a földjét négyzetekre osztja, s a különféle négyzetekben pár éven keresztül eltérő termelési megoldásokat alkalmaz, majd e kezelések eredményeit hasonlítja össze.

Vegyük észre, hogy a példában szereplő megoldások társadalomtudományi megfelelői (rendre a mélyinterjú, a kérdőívezés, a laboratóriumi és a kontrollált kísérlet) ugyanarra jók, illetve nem jók, mint az *1. ábrában* szereplő példák:

- a mélyfúrással mindent megtud a gazda a talajáról, de csak egy helyen, s ebből következtetni a teljes területre igencsak hibás lehet;
- a kapargatással (ha jó az elrendezés) a teljes földterületről tud meg valamit keveset, de nem többet, mint ami a felülethez közel található;
- a laboratórium mesterséges körülményeit sosem tudja majd a valóságban reprodukálni.

Ugyanakkor a kontrollált kísérlet (persze, ha a kezelések helyesen vannak kialakítva és végrehajtva, s van elég pénz és idő az ismétlésükre) mindezen hibáktól, korlátokról mentes lehet.

Visszatérve a társadalomtudomány területére, mint azt a tanulmány elején felvettük, a szociológus a kutatás módszerének kiválasztása, illetve a kutatási elrendezés

³⁶ Elsősorban, hogy miképpen viselkedjenek a kontrollált kísérlet tesztelői, illetve mi alapján azonosítsák a védett tulajdonságot a megfigyelés során (továbbá a megfigyelések szabályai betartásának ellenőrzése és pontos dokumentációja).

részleteinek kidolgozása során könnyen elmulaszthatja a lehető legjobb megoldás kiválasztását. Mi több, könnyen belecsúszhat a rutinszerű kvanti-kvali dilemma közötti kényszerválasztásba. Másként szólva a problémáról: ha egyáltalán gondol erre a kutató, akkor a háromféle kritériumnak egyszerre való megfelelés nehézségét a kvantitativ-kvalitativ megoldás közötti választással próbálja megoldani. Ez a gordiuszi megoldás ugyanakkor éppen olyan, ahogy azt Nagy Sándor elképzelte – egyszerű, gyors és durva. Ami azonban jó egy macedón hadvezérnek ahhoz, hogy király legyen, biztos, hogy nem jó egy kutatás megtervezése során.

Sok társadalomtudós a gordiuszi csomó átvágását (a kvanti-kvali dilemma megkerülését) az úgynevezett „mixed method” választásával látja lehetségesnek. Ez a megoldás azonban semmivel sem elegánsabb, mint az újgazdag bankár által alkalmazott megoldás a fürdővíz milyenségére utaló kérdésre (vö. Hyppolit, a lakáj) adott válasznál: „egy kicsit ebből, egy kicsit abból”. Mint ahogy a langyos fürdővíz sem jobb a forrónál és a hidegnél, ugyanúgy rossz lehet a módszertan megválasztása során a módszerek keverése, mert olyan megoldás születhet, amely mindkét módszer hátrányait egyesíti. A hibrid megoldás előnyben részesítése részben persze már csak következmény, amennyiben e legtöbbször rossz válasz a dichotóm módszertani klasszifikációból fakad. Belátható ugyanis, hogy ha a kvantitativ-kvalitativ dichotómia uralja a módszertani diskurzust, akkor a hibrid megoldások sokszor jobbra csupán a „csepp a tengerben” vagy a „tenger a cseppben” jellegű álmegoldások alkalmazását jelenthetik.

A „cseppben a tenger” a kvalitativ kutatók gyakran alkalmazott érve, amely abból a feltételezésből fogan, hogy egy eset alapos elemzése lehetővé teszi a társadalmi megismerést: elég feltételezni, hogy minden más eset is az alaposan leírthoz hasonló. Ez azonban az esetek egy jelentős részében csak a kvalitativ módszerek alacsony külső érvényességének és szintén kismértékű megbízhatóságának elrejtése, továbbá a kvalitativ adatok számszerűsítésével való próbálkozás is csupán kompenzáció, amely azonban nem segít a tényleges módszertani problémákon.

Ennek ellentéte a „tengerben a csepp”, amikor a kvantitativ kutató a survey-kutatás alacsony belső érvényességének tudatában a kvantitativ elemzésbe kvalitativ kutatások eredményeit keveri bele, azt remélve, hogy elfedi a belső érvényesség alacsony voltát.³⁷

Úgy véljük, ezek a „hibridek” nemhogy segítenék, de rontják a megismerés esélyét azért, hogy elfedik a módszerek korlátait, s ezzel gátolják az igazi megoldások keresésének szükségességét.

A módszerek álságos elegyítése helyett tehát hozzunk egyértelmű döntést a kövendő módszerről.

- (1) Válasszunk olyan módszert, amely a legalkalmasabb a kérdésünk megválaszolásához: minden esetben az 1. ábrán szereplő kocka jobb felső sarkának a közelítése legyen a cél. Tehát ne válasszunk a kvalitativ módszert, hiszen ekkor

³⁷ Ez persze kisebb baj, mint az előző, amennyiben ily módon nem kerül sor hamis általánosításra, de azért ilyenkor is helyesen teszi a kutató, ha hangsúlyozza a színesítés korlátait.

eleve lemondunk a szerzett tudás külső érvényességének és megbízhatóságának növeléséről. Másként, hiába csinálunk száz „mélységesen mély” interjút, ez továbbra is csupán száz csepp lesz a tengerből, amiből továbbra nem tudunk meg semmit a tengervíz összetételéről, a tenger áramlatairól. S nem (csupán) a kis elemszám miatt, hanem mert e módszer célja (a csepp lehető legalaposabb elemzése) és a vizsgálat alanyának kiválasztása (nem volt cél a reprezentativitás) miatt a száz egyedi eset eredményei nem lesznek aggregálhatók.

- (2) Ha ezután bármely más módszert választunk, ezt oly módon tegyük, hogy a kutatási design a lehetséges legmagasabb értéket érje el mindhárom mérési dimenzióban, vagyis a választott módszer korlátai között törekedjünk az 1. ábrán minél inkább jobbra és felfelé jutni.³⁸
- (3) Ne keverjük, hanem ötvözzük a módszereket! Ezalatt azt értjük, hogy alkalmazunk minden olyan módszert, amely a végső eredményt segíti, de ne keltjük annak látszatát, hogy ezek elegyítése önmagában elégséges a társadalmi tudás növeléséhez.³⁹ Így például a közösségtanulmány sokféle módszer ötvözetének (s nem keverésének) tekinthető, amennyiben az egymásra épülő, egymás eredményeit oda-vissza támogató módszertani egyvelegről lehet ebben az esetben beszélni – éppen ezért nem is lehet tipológiánk egyetlen pontján elhelyezni.

A módszerek együttes alkalmazása során minden módszert pontosan arra (és csak arra) használjunk, amire alkalmas, s éppen ezért ezek ötvözésével legyünk óvatosak (Bazeley 2004). Vagyis a kvantitatív kutatás mellé rendelhetünk kvalitatív kutatást az előkészítés fázisában (javuljon a belső érvényesség), de a végén ezek elemzésének eredménye (érzékeltetve a kvantitatív megállapítások lehetséges mechanizmusait, a létrehozott típusok finomszerkezeteit) ne adódjon hozzá a kvantitatív alapú következtetésekhez. A kvalitatív kutatás során pedig mellőzni kell minden utalást, ami alapján hamis következtetést lehetne levonni a külső érvényesség szempontjából. S ami a szakma számára különösen fontos: ne ringassuk magunkat abba az illúzióba, hogy a módszerek keverése a módszertani gondok megoldásának csodagyógyszere.

Abstract: The aim of our paper is to offer a system of viewpoints for sociologists having to choose a method for their research. In the course of formulating their goals, researchers should give primary importance to reaching internal and external validity, as well as reliability. The practical realization of weighing between different methods is shown by a specific research topic (the prevalence of discrimination). In deciding between methods we put a special emphasis on controlled experiments and non-participant observation as

38 A társadalomtudományokban leggyakrabban alkalmazott módszer, a survey esetében például alkalmazható a belső érvényesség növelése érdekében a rugalmas kérdésés (Letenyei–Nagy 2007), illetve a kognitív interjúzás (Caspar et al. 1999). De ilyenek tekinthető a migrációpotenciál-kutatás során alkalmazott tisztítási megoldások sokféle változata is, amelyek célja az, hogy a migrációról álmódokat megkülönböztessék a migrációt tervezőktől. Továbbá az elemzésnek a releváns népességre való szűkítése is fontos javítás lehet, mert ezzel a várható viselkedés előrejelzésének pontossága (no nem nagyon, de elvben) növelhető (Sik 2003).

39 Például ahogy az informális gazdaság nagyságának megbecslésére tett javaslat (Sik 1996) az antropológiai feltáró megfigyelést és az anonim interjú antropológiai módszereit kombinálta statisztikai vagy/és surveyadatokkal. Ettől az ötvözetből azt lehetett remélni, hogy az előbbi a becslés belső érvényességének (az „őszinteségi momentumnak” – u.o.: 286), az utóbbi az így kapott adatok általánosíthatóságának (vagyis a külső érvényességnek) együttes érvényre jutását biztosíthatta volna.

we believe that these relatively rarely used methods approach the advantages of methods described in the first part of our paper. These days the dilemma of choosing between different methods is usually resolved by researchers resorting a mixed—both qualitative and quantitative—methodology; which the authors of this paper believe is a mistake. At the end of our paper we make our case against such practices, as well as propose a more complex approach to the problem, i.e. amalgamation instead of mixing.

Irodalom

- Alessina, A. – Fuchs-Schündeln, N. (2005): *Good-by-Lenin (or not?): The effect of Communism on people's preferences*. NBER Working Paper 11700. Cambridge, MA. <http://www.nber.org/papers/w11700.pdf>
- Babbie, E. (2008 [1985]): *A társadalomtudományi kutatás gyakorlata*. Budapest: Balassi Kiadó.
- Barabas, J. – Jerit, J. (2010): Are survey experiments externally valid? *American Political Science Review*, 104(2): 226–242. http://www.jjerit.com/images/BarabasJerit_APSR_2010.pdf
- Barabási A. L. (2010): *Villanások – a jövő kiszámítható*. Budapest: Nyitott Könyvműhely.
- Bazeley, P. (2004): Issues in mixing qualitative and quantitative approaches to research. In Buber, R. – Gadner, J. – Richards, L. (eds.): *Applying Qualitative Methods to Marketing Management Research*. UK: Palgrave Macmillan, 141–156. <http://www.researchsupport.com.au/MMIssues.pdf>
- Bernstein, B. (1971): *Class, Codes and Control. I–III*. London: Routledge and Kegan Paul.
- Bertrand, Marianne and Sendhil Mullainathan (2003) Are Emily and Greg More Employable than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination, NBER Working Paper Series, www.nber.org/papers/w9873
- Campbell, D. – Stanley, J. (1963): *Experimental and Quasi-Experimental Designs for Research*. Chicago: Rand McNally.
- Caspar, R. A. – Lessler, J. T. – Willis, G. B. (1999): *Reducing Survey Error through*. Research on the Cognitive and Decision Processes in Surveys. Short course presented at the Meeting of the American Statistical Association
- Crapo, R. H. (1990): *Cultural Anthropology, Understanding Ourselves & Others*. Guilford, CT: The Dushkin Publishing Group.
- Drost, E. A. (2011): Validity and reliability in social science research. *Education and research perspectives*, 38(1): 105–123.
- Edward M. – Saiegh, S. M. – Satyanath, S. (2008): *National Cultures and Soccer Violence*. NBER Working Paper 13968. Cambridge, MA. <http://www.nber.org/papers/w13968>
- Engel, R. – Calnon, J. (2004): Comparing benchmark methodologies for police-citizen contacts: Traffic stop data collection for Pennsylvania State Police. *Police Quarterly*, 7(1): 97–125. http://www.uc.edu/content/dam/uc/ccjr/docs/articles/engel_articles/Comparing_Benchmarks.pdf
- Enyedi Zs. – Fábrián Z. – Sik E. (2004): Nőtt-e az előítéletek Magyarországon? In Kolosi T. – Tóth István Gy. – Vukovich Gy. (szerk.): *Társadalmi riport 2004*. Budapest: TÁRKI, 375–399. <http://www.tarki.hu/adatbank-h/kutjel/pdf/a809.pdf>

- Gerber, A. S. – Green, D. P. (2012): *Field Experiments: Design, Analysis, and Interpretation*. New York: WW Norton.
- Héra, G. – Ligeti, Gy. (2005): *Módszertan. Bevezetés a társadalmi jelenségek kutatásába*. Budapest: Osiris.
- Horzsa G. – Sik E. (2012): A diszkrimináció mérése készen kapott kísérletek (KKK) segítségével. In Sik E. – Simonovits B. (szerk.): *A diszkrimináció mérése*. Budapest: ELTE TáTK, 186–201. http://www.tarki.hu/hu/about/staff/sb/Diszkriminacio_merese.pdf
- Kádár A. – Körner J. – Moldova Zs. – Tóth B. (2008): *Szigorúan ellenőrzött iratok. A magyar igazoltatási gyakorlat hatékonyságáról*. Kutatási jelentés. Budapest: Magyar Helsinki Bizottság. <http://helsinki.hu/wp-content/uploads/books/hu/Szigoruan-ellenorzott-iratok.pdf>
- Letenyei L. – Nagy G. D. (2007): A rugalmas kérdőív. *Szociológiai Szemle*, 17(1–2): 29–46. http://www.tarsadalomkutatatas.hu/kkk.php?TPUBL-A-837/szoc_szemle_2007_1/TPUBL-A-837.pdf
- OSI – Open Society Institute (2005): *Ethnic Profiling in the Moscow Metro*. New York: Open Society Institute. http://www.opensocietyfoundations.org/sites/default/files/metro_20060613.pdf
- Otlakán K. (2012): Diszkriminálnak-e a hazai munkáltatók? In Sik E. – Simonovits B. (szerk.): *A diszkrimináció mérése*. Budapest: ELTE TáTK, 265–277. http://www.tarki.hu/hu/about/staff/sb/Diszkriminacio_merese.pdf
- Pálosi É. – Sik E. – Simonovits B. (2007): Diszkrimináció a plázában. *Szociológiai Szemle*, 17(3–4): 135–148. <http://www.szociologia.hu/dynamic/070304palosi.pdf>
- Shadish, W. R. – Cook, T. D. – Campbell, D. T. (2001): *Experimental and Quasi-Experimental Designs for Generalized Causal Interference*. Boston, New York: Houghton Mifflin Company. <https://depts.washington.edu/methods/readings/Shadish.pdf>
- Sik E. (1996): Traktátus az informális gazdaság és a feketemunka terjedelmének megismerhetetlenségéről általában és a posztoszocialista átalakulás során különösen. In Várnai Gy. – Tausz K. (szerk.): *Rejtőzködő jelen*. Tanulmányok Ferge Zsuzsának. Budapest: Hirschler Rezső Szociálpolitikai Egyesület, 244–293.
- Sik E. (2000): A rossz, a rosszabb és a legrosszabb. *Café Babel*, 1: 101–110.
- Sik E. (2003): A migrációs potenciál kutatásának alapfogalmai. In Örkény A. (szerk.): *Men+ni vagy maradni? Kedvezménytörvény és migrációs várakozások*. Budapest: MTA Kisebbségkutató Intézet – Nemzetközi Migrációs és Menekültügyi Kutatóközpont, 15–18.
- Sik E. – Simonovits B. (2008): Egyenlő bánásmód és diszkrimináció. In Kolosi T. – Tóth I. Gy. (szerk.): *Társadalmi Ríport 2008*. Budapest: TÁRKI, 363–386. http://www.tarsadalomkutatatas.hu/kkk.php?TPUBL-A-822/publikaciok/tpubl_a_822.pdf
- Sik E. – Simonovits B. (2009): A diszkrimináció mértéke a munkaerőpiacon. Közelpép. In Fazekas K. – Lovász A. – Telegdy Á. (szerk.): *Munkaerő-piaci Tükör 2009*. Buda-

- pest: MTA Közgazdaságtudományi Intézet – Országos Foglalkoztatási Közalapítvány, 118–133. <http://real.mtak.hu/10473/1/egyben.pdf>
- Sik E. – Simonovits B. (szerk.) (2012): *A diszkrimináció mérése*. E-tankönyv. Budapest: ELTE TáTK. http://www.tarki.hu/hu/about/staff/sb/Diszkriminacio_merese.pdf
- Simonovits B. (2010): A hátrányos megközelítés észlelése, tapasztalása és a diszkriminatív viselkedés. A diszkrimináció mérési módszereinek egy lehetséges tipológiája. In Némédi D. – Szabari V. (szerk.): *Kötő-jelek 2009*. ELTE Szociológia Doktori Iskola Évkönyve. Budapest: ELTE TáTK Doktori Iskolája, 33–69.
- Simonovits B. (2012a): A diszkriminációtesztelés etikai dilemmái. *Esély*, 2: 54–65. http://www.esely.org/kiadvanyok/2012_2/simonovics.pdf
- Simonovits B. (2012b): Ali, Chen, Fatima és Sára esélyei a magyar munkaerő- és lakáspiacon. I: Sik E. – Simonovits B. (szerk.): *Migráns esélyek és tapasztalatok*. Budapest: TÁRKI, 68–117. http://www.tarki.hu/hu/research/migrans/tarki_eia_tanulmanyok_2011.pdf
- Steckel, R. H. (2008): Heights and human welfare: Recent developments and new directions. NBER Working Paper, No. 14536. <http://cid.bcrp.gob.pe/biblio/Papers/NBER/2008/diciembre/w14536.pdf>
- Szántó Z. – Tóth I. J. – Varga Sz. – Cserpes T. (2011): A korrupció típusai és médiareprezentációja Magyarországon 2001–2009. *Belügyi Szemle*, 59(11): 53–74.
- Szántó Z. – Tóth I. Gy. (2003): Dupla vagy semmi, avagy Kockáztassuk-e a talált pénzt? Kísérlet a kockázattal szembeni attitűd mérésére kérdőíves adatfelvételi módszerrel. In Szántó Z. – Gál R. I. (szerk.): *Cselekvéstudomány és társadalomkutatás: Im memoriam Csontos László*. Budapest: Közgazdasági Szemle Alapítvány, 13–47.