

Róbert Péter

## Kétes megbízhatóság – adatokban és egymásban

---

### Hozzászólás Léderer Pál cikkéhez

Léderer Pál Rejtő Jenő modorában kezdi dolgozatát: a rugások és pofonok röpködnek, de – mint Fülíg Jimmynél – hiányzik a jelzőlámpa. Én legalábbis semmiféle jelzést, eligazítást nem találtam arra, hogy az első bekezdésben szereplő kísérlet, amely az frás tárgyát, a szociológiai mérés validitására vonatkozó megbízható vizsgálódások nehézségét hivatott illusztrálni – s amely példázathoz a szerző játékosan visszatér a tanulmány végén is –, hogyan vezet el (át vagy keresztül) ahhoz az ideges kirohanáshoz a harmadik és negyedik bekezdésben, az empirikus adatkezelésben matematikai (vagy valóban statisztikai, esetleg matematikai-statisztikai) apparátust alkalmazó szociológusok ellen. Így talán jobban is tenném, ha az általánosságnak ezen a szintjén ezt az egészet elereszteném a fülem mellett. Hogy mégsem ezt teszem, annak pusztán az az oka, hogy Léderer Pállal szemben én egészen mást vélek tapasztalni. Nevezetesen éppen az a benyomásom, hazai és külföldi kollégák munkája alapján, hogy az adatok érvényességén és megbízhatóságán inkább azok aggódnak jobban, akik nem elégednek meg a gyakorisági megoszlások és keresztátlák bemutatásával, hanem az adatokban meghúzódó mélyebb összefüggések statisztikai bizonyítására is kísérletet tesznek.

Mind az érvényesség, mind a megbízhatóság kérdése komoly probléma a szociológiai mérésben és kutatásban. Az én számomra azonban ez az egész inkább úgy jelenik meg, hogy vannak szociológusok, akik sem az adatfelvételben, sem az adatkezelésben nem járnak el tudományos igényvel, mások viszont mindkét területen törekszenek az ilyesmire. Mi több, az a kétes megbízhatóság, amivel kapcsolatban Léderer Pálnak kevés használható javaslata van; ("nincsenek állítások" – mondja, s ezen a felettébb sajnálatos látleleten ő

.....

sem tud javítani<sup>1</sup>), többek között éppen olyan többváltozós elemzési módszerekkel kezelhető (nem javítható, de ellenőrizhető, tudatosítható!), amelyekkel szemben ő olyan ellenszenvvel viseltetik. De ennél több szót ezúttal a szociológia és a tudományosság ügyére én sem akarok vesztegetni. Maradjunk most valóban a megbízhatóság problémájánál.

## Az adatfelvételek gondjairól

---

A survey típusú kutatások terén a szociológiai mérés problémái közül az adatfelvétel megbízhatósága az egyik gond. Léderer úgy látja, az adatfelvétel terén az utóbbi fél évszázad során érdemlegesen semmi "fejlődés" nem történt. Én úgy gondolom, számos változás történt, bár igazán új minőségről talán valóban nem lehet beszélni. A survey-k esetében az adatfelvétel az a folyamat, amelynek eredményeképpen számítógépes elemzésre alkalmas adatbázis (*machine readable dataset*) keletkezik. Válasszuk két részre az ennek során felmerülő megbízhatósági kérdéseket: az egyik rész az *adatközlő személyéhez*, a másik az *információk átviteléhez* kapcsolható. Kezdjük az utóbbival.

Ha Magyarországon az adatfelvételek terén semmi lényeges változás nem történt, a kilencvenes évek elején is nagyjából úgy mennek a dolgok, mint húsz évvel korábban, az ilyen típusú kutatások bölcsőjét ringató néhai Tömegkommunikációs Kutatóközpontban (ez valószínűleg ilyen sommásan nem is igaz), akkor annak döntően *pénzügyi okai* vannak. A nagyvilágban ugyanis számos kísérlet történt arra, hogy az adatközlő által kimondott (leírt) szó és a számítógépes adatbázisban megjelenő kód közötti torzulás csökkenjen. De menjünk sorjában.

Elsőként kísérletek történtek annak a torzításnak a csökkentésére, amit a kérdezőbiztos személye visz bele a folyamatba. Gondoljunk most komoly dolgokra, s ne arra, ami a laikusnak elsőként eszébe jut, hogy a kérdező otthon tölti ki a kérdőívet. Ez a probléma is létezik, valószínűleg nem is lehet elkerülni. Az ellenőrzés nem lehet teljes körű, de alapvetően bízunk kell azokban, akiket erre a munkára alkalmazunk: ez is egy megbízhatósági kérdés. Komoly irodalma van azonban annak, hogy a kérdezőbiztos megjelölésével, személyével, a kérdezési szituációban való viselkedésével milyen módon befolyásolja a kapott válaszokat.<sup>2</sup> A megfelelő infrastrukturális ellátottsággal rendelkező országokban (elsősorban Amerikában) a telefonnal lebonyolított interjú egyik jellemzője, hogy a kérdező és a kérdezett közti kapcsolat mechanikusabb, elidegenedettebb. S ezek a "hátrányok" itt a megbízhatóságot segítik: kevesebb személyspecifikus kérdezői elem terheli meg az adatgyűjtést, a válaszoló csak egy telefonhanggal találkozik. Nem beszélve arról, hogy a telefonhoz kapcsolt számítógép (nem ritka, hogy a számítógép "választ" mintát, generál telefonszámokat) ellenőrzi, hogy a telefonhívás megtörtént-e, hogy a beszélgetés időtartama alapján valószínűsíthető az is, hogy a kérdező tisztességesen végigkérdezett minden kérdést.<sup>3</sup> (A részleges önkitöltés, a nehéz vagy kényelmetlen kérdések "elsikkasztása" az interjú során, majd otthoni bepótlása azért is nehezen tette érhető, mert a kérdezők véletlenszerű ellenőrzése általában arra szorítkozik, hogy elmentek-e a megadott címre vagy sem. Hogy bizonyos kérdések bizonyos kérdezők esetében furcsa eloszlást mutatnak, esetleg csak egy jóval későbbi elemzési szakaszban derül ki, ha a kutató egyáltalán megkísérel utánajárni a dolognak, s ha erre megvan a technikai

.....

lehetőség, tehát a kérdező személye is részét képezi az adatbázisban rögzített információknak.)

Ahol az infrastruktúra nem olyan fejlett, egyre szélesebb körben térnek át a postai kérdőívezésre. Itt már semmiféle személy nem ékelődik az adatközlő és az információ közé, ami a kitöltött kérdőívben megjelenik. Ezzel kapcsolatban megemlíthető, hogy a válaszokat nem kell kimondani, "bevallani" egy idegen személy előtt, s ez őszintébb reakciókhoz vezet. Ugyanakkor felvethető, hogy a válaszokat le kell írni, s valamit "írásba adni" megfontolandóbb, tehát lehet, hogy nem így jutunk a legőszintébb véleményekhez. Általában jobbnak tekinthető a kiinduló minta, amelynél egy adott alapsokaságból mindenkinek egyforma az esélye a mintába kerülésre, miközben a személyes kérdés esetében az utazás racionális szempontjai miatt többlépcsős mintavétel történik, s a személyek egyenlő esélye véget ér a település kiválasztásával, a mintába nem került településen élő személy mintába kerülési esélye nulla. Ugyanakkor éppen a legjellegzetesebb kitogás ezen a téren a visszaküldés alacsony gyakorisága. Az pedig nyilvánvaló, hogy a kérdőívet vissza nem küldők eloszlása szignifikánsan eltér a vizsgált alapsokaságtól, amint a személyes kérdés esetében is a válaszmegtadóké. Végső soron tehát lehet, hogy rosszabb mintához jutunk a postai kérdés útján, ami nagyobb mértékű súlyozást igényel később. (A súlyozás szükségessége vagy feleslegessége, a minták reprezentativitásának ellenőrzése külön hosszú történet; még visszatérek rá.)

Megemlíthető még az a kérdés, hogy a postai kérdés esetében mennyivel nagyobb a valószínűsége annak, hogy a kérdőívben szereplő adatok nem egyéni vélemények, hanem a család "kollektív bölcsességének" eredménye. A négy szemközti beszélgetés a személyes kérdésnél sem mindig biztosítható, de ott legalább kiderülhet, ha a kutató interjúszituációs kérdéseket helyez el a kérdőív végén, amelyek aztán akár statisztikai ellenőrzés alapjául is szolgálhatnak később.<sup>4</sup> További gond, hogy milyen hatása van a mérésünknek annak, ha valaki egy kérdőívben szereplő kérdésekre olyan beszélgetés formájában válaszol, ahol egy adott kérdés után nem tudja, mi a következő kérdés, szemben azzal, ha valaki egy kérdőívet előbb végigolvas, elgondolkodik rajta, s azután nekiül kitölteni. (A kettő biztos nem ugyanaz. Kérdés, hogy melyik mérést szeressük jobban az adott esetben.)<sup>5</sup>

Az állítások továbbra is váratnak magukra. Magyarországon telefonos kérdésre Budapesten történtek kísérletek.<sup>6</sup> Itt sem tudok olyan kutatásról, ahol pontosan megegyező kérdőívet személyesen és telefonon is lekérdeztek volna, majd a két adatbázist összehozva az eredményeket az "adatfelvétel módja" változó mentén (a szakirodalomban használt kifejezéssel a *mode effects*re koncentrálna) vizsgálták volna. Ugyanez nem történt meg személyes és postai kérdés esetében sem. (A pontosan megegyező kérdőív is fikció, egy kérdőívnek másképp kell ugyanis kinéznie, ha kérdezőbiztos olvassa fel és ha önkítöltéses.) Az ilyen kutatások hiánya szerintem kevésbé a tudós (vagy nem tudós) kollégák végtelen magabiztosságával, hanem inkább véges anyagi lehetőségeikkel magyarázható. Én a legtöbb fantáziát a postai kérdésre való áttérésben látnám, s el tudnám képzelni, hogy egy ilyen folyamatot támogasson a Magyar Posta abban a reményben, hogy mindaz a pénz, amit a kutató a kérdezőbiztosok munkabérének megspórol, majd postaköltséggént jelenik meg (kérdőívek küldése és visszaküldése, emlékeztető és sokadik felszólító levél küldése azoknak, akik még nem válaszoltak), s így az a posta zsebébe vándorol.<sup>7</sup>

.....

Miután sikeresen megoldottuk azt a problémát, hogy a szociológiai kutatás céljára szolgáló információ az adatközlőtől "elidegenedve" ott van a kérdőívben, ennek eredményképpen ilyen vagy olyan technikai úton kitöltött kérdőívek számai vagy ezrei halmozódnak a kutatóintézetben. A következő feladat, hogy az adatok a papírról a számítógépre kerüljenek, vagy ahogy mondják, a kérdezés után következik a kódolás és a rögzítés szakasza. Ezen a téren is vannak törekvések arra, hogy az információk minél kevésbé torzuljanak (tovább). A magas technika oldaláról kísérleteznek a scannerrel, ami a kérdőívről egyből bevinné a kódot a számítógépbe, bár a siker egyelőre mérsékelt. Megemlíthetők azok a próbálkozások is, ahol a kérdezőbiztos kis hordozható számítógéppel a hóna alatt közlekedik, a kérdőív is a gépen van, és egyből oda kerülnek a válaszok is, miáltal a kódolás, rögzítés, mint olyanok ki is maradnak.<sup>8</sup>

Visszatérve azonban a realitás talajára, olyan problémák merülnek fel, hogy egy standardizált, nagy mintán lekérdezett kérdőívben a kérdések többsége előre kódolt, ún. zárt kérdés legyen, illetve milyen arányban fordulhatnak elő nyitott kérdések. Amelyik kérdőív csupa zárt kérdést tartalmaz, ott a kódolás egyszerű másolás (figyelem: itt is lehet hibát elkövetni!), vagy ez a szakasz a hibaforrások csökkentése érdekében ki is hagyható, s lehet akár a kérdőívből közvetlenül is rögzíteni. Nem szabad értékén alul kezelni mindazokat a törekvéseket, amelyek a kérdőívek tipografálásában testesülnek meg – például, a kódok kerüljenek-e külön előre nyomtatott kódkockába a kérdőív szélén, különböző betűtípusok alkalmazása, táblázatok készítése –, mert ezek a próbálkozások mind olyan munkafeltételek megteremtését szolgálják, hogy a kódolás, rögzítés során csökkenjen a hibák száma. Persze minden ilyen típusú extra a kérdőívben többletköltség, és tudtommal senki nem végzett semmilyen vizsgálatot arra vonatkozóan, hogy a többletköltség arányban van-e a kevesebb hibával, következésképpen a jobb adatokkal. Így a pénztárca vastagsága és kutatói felfogás dönt esetenként mindezekről.

Nyitott kérdések alkalmazása esetében a kódolás nem úszható meg, s jelentős többletmunkát jelent. Olyan érvek szólnak mégis mellette, hogy a kérdezett valódi álláspontját fogalmazhatja meg, s nem kényszerül arra, hogy véleményét "jobb híján" besorolja valamelyik adott válaszkategóriába. Vagy másképp fogalmazva ugyanezt: a kutató kívánja eldönteni, hogy az elhangzott (és ideális esetben szó szerint lejegyzett) választ miként értékeli, s nem bizza ezt a feladatot a kérdezőbiztosra azzal, hogy az karikázzon egy kódot a kérdőívben. A nyitott, illetve zárt kérdések alkalmazására persze nem lehet aranyszabályt megfogalmazni, téma, illetve a választott adatfelvételi és mérési módszer kérdéséről van szó.<sup>9</sup> Amivel tisztában kell lenni: zárt kérdések esetében vállaltuk azt a veszélyt, hogy az elhangzott válaszhoz képest kisebb-nagyobb mértékben inadekvát információ került a kérdőívbe; nyitott kérdések esetében vállaltuk azt a veszélyt, hogy utólag, sokkal lassabban, igen sok munka árán jutunk el az információhoz, miközben esetleg pusztán a kódolás szakaszára tolódik el az a torzulás, amit a kérdezés szakaszában kiküszöböltünk. Érdeemes ezzel összefüggésben utalni arra, hogy a pszichológusok rendszeresen kódolnak duplán, egymástól független kódolókkal, majd statisztikai próbának vetik alá az eredményeket. Ez az eljárás a szociológia gyakorlatára ismereteim szerint kevésbé jellemző – tétellezzük fel, hogy megintcsak idő és pénzhiány miatt. Hogy ezzel adott esetben milyen hibát követünk el – megintcsak nem tudható.<sup>10</sup>

Végül, az adatok megbízhatóságának javítása érdekében tett számos próbálkozásról lehet beszámolni a rögzítés terén is. Ezek egyike a számítógépes rögzítőfile készítése, ami kiszűri

.....

a lehetetlen kódokat, vagy akár logikai ellenőrzések beépítését is lehetővé teszi. (Lehetetlen kód esetében a program leáll, a rögzítő észreveszi, hogy hibázott.) Ezzel a módszerrel elkerülhető, hogy mintánkból utóbb ki kelljen dobni azt a néhány embert, akinek a neme 3-as vagy 4-es, továbbá azokat, akik az ötfokú skálán megítélhető kérdésre 6-os vagy 7-es válaszkóddal szerepelnek.<sup>11</sup> Azt persze rögzítőprogram alkalmazása mellett sem tudjuk, hogy az ötfokú skálán megítélhető kérdések rögzítésekor hányszor sikerült például 3-as helyett 4-est ütni, ami azért profi rögzítővel is megeshet. Ezt akkor tudnánk, ha minden kérdőívet kétszer és egymástól függetlenül rögzítenénk, s a kapott eredményt összevetnénk, ami idő- és pénzszerűe miatt szintén nem jellemző. Lehet, hogy olyan kevés hiba adódna, hogy meg sem érné ilyesmit csinálni, de ma ennek eldöntése is jobbára hit és külső körülmények kérdése. Ebben a helyzetben pedig nem marad más hátra, mint a megbízhatóság, ezúttal a rögzítőben, mármint hogy csak olyan gyorsan dolgozik, amilyen sebességgel hibátlanul (minimális hibával) képes dolgozni, s nem olyan gyorsan, amilyen órabért a vállalt munkáért el szeretne érni.

Míndezeneket csak a teljesség kedvéért írtam le, azt illusztrálandó, hogy a *test-retest* problémán túlmenően mi mindenben aggódhat még a tudós szociológus, mennyi más oka is lenne még kételkedni abban, amit kutatási eredményei mutatnak, s mit lehet(ne) tenni, mivel kéne próbálkozni az adatfelvétel terén annak érdekében, hogy adataink valamivel megbízhatóbbak legyenek. Könnyen lehet persze, hogy most Léderer Pál erre azt mondja, mindez sok pénzért (és még több pénz igényének emlegetésével!) csupán egy nagy pofon a szarnak, mert az igazi gondok azokkal a megbízhatósági kérdésekkel vannak, amelyek az adatközli személyéhez kapcsolódnak. Lássuk tehát ezeket is!

### Azok a rohadt emberek...

.....

Nyilvánvaló, hogy a szociológiai adatfelvétel során a legtöbb problémát a válaszolók okozzák. Listába szedve ezeket:

1. Nem ülnek otthon és nem várják a kérdezőbiztost, mi több, képesek sorozatosan nem otthon tartózkodni, munkát, szórakozást stb. fontosabbnak tekinteni, mint egy szociológiai kérdőívre válaszolni – helyettük ezért mást kell megkérdezni, amivel torzítják a mintát.

2. Ha otthon is vannak, nem állják komolyan venni a kérdőívre nyomtatott szöveget, miszerint a "Válaszadás önkéntes" – őhelyettük is mást kell megkérdezni, amivel tovább torzul a minta.

3. Ha kötélnak is állnak végül, egyes kérdésekre nem tudnak vagy nem akarnak válaszolni, ez pedig adathiányt eredményez, ami nehezíti majd a későbbi elemzést.

4. Füllentenek: például letagadják valós jövedelmeiket, s ki tudja még mi mindent.

5. Ha legjobb szándékuk szerint és őszintén is akarnak válaszolni, akkor is összevissza beszélnek; ha egyik nap kérdezik őket, ezt mondják, egy másik nap meg valami mást.

Míndezenek sokak számára elégséges indok lehet arra, hogy ne is foglalkozzunk survey jellegű szociológiai adatgyűjtésből származó megbízhatatlan információk – mégoly tudományos(?) – elemzésével. De ha valaki mégis ilyenre adná a fejét, hogyan nézhet szembe egyáltalán a fenti problémákkal? Valóban olyan tehetetlenek lennénk, amint azt Léderer cikke sugallja? Tehát próbálkozhatunk éppen, de a végén úgyis jön a "marhanagy füles". Kezdjük a mintával!

.....

A szociológusnak lehetősége van arra, hogy a lekérdezett kérdőívekből adódó "mintáját" összevesse a népszámlálással (mikrocenzussal), és így legalább bizonyos változók mentén ellenőrizze azt, hogy adatai mennyire tekinthetők majd reprezentatívnak. Ezek a változók általában a nem, a lakóhely településtípusa, a korcsoport és az iskolázottság. Magam nem szoktam megelégedni azzal, hogy a mintából adódó vonatkozó "primer" eloszlások és a népszámlálási eloszlások "blickre" nem térnek el durván, jobban szeretem, ha egy  $\chi^2$  próba is alátámasztja a "szemmértéket". De még ezt sem tartom elégségesnek. Igazából a fenti négy változó "kompozit" (együttes) eloszlására vagyok kíváncsi, tehát például arra, hogy a mintában a 30–40 év közötti szakmunkás végzettségű városi férfiak aránya illeszkedik-e ahhoz az arányhoz, ami a népszámlálásban szerepel. (Én is belátom persze, hogy ez nem jelent garanciát arra, hogy más változók közötti "kompozit" eloszlások is reprezentatív adatoknak tekinthetők, hanem ez pusztán *vélelmezés*.) Amennyiben aztán gond merül fel a minta reprezentativitásával, ami nem ritka, lehetőség van a súlyozásra.<sup>12</sup>

A kiesett címek (válaszmegtagadás, háromszori sikertelen megkeresés) ugyanis szisztematikusan torzítják a mintát.<sup>13</sup> A soha otthon nem lévők, a nem válaszolók nem véletlenszerűen oszlanak el; a pótcímek használata pedig elsősorban azt biztosítja, hogy a tervezett kérdőívszám elkészüljön. Még az sem mindig megoldható, hogy a pótcím a főcím azonos nemű, hasonló életkorú, azonos típusú településen élő "alteregója" legyen. De ha ez meg is valósul, vajon biztos-e, hogy a pótcím a kiesett főcímmel azonos vallású, ugyanarra a pártra szavazna, apja ugyanabba a foglalkozási csoportba tartozik stb? Ha a szociológus kicsit mélyebben akar meggyőződni adatai ilyen szempontból vett megbízhatóságáról, megteheti, hogy azokat aszerint vizsgálja meg, hogy eredetileg a mintába került személytől (főcím) származnak, vagy kiesett személy helyett megkérdezett válaszolótól (pótcím).

Kutatási eredményeink megbízhatóságát veszélyeztető következő gond az ún. *non-response error*, vagy válaszhiány, ami abból keletkezik, hogy a válaszolók egy része bizonyos kérdésekre nem tud vagy nem akar felelni.<sup>14</sup> Ha azután a különböző kérdések közötti kapcsolatra vagyunk kíváncsiak, elegendő, hogy akár egy szóban forgó kérdés esetében is válaszhiány merüljön fel, máris nem tudunk az összefüggésre vonatkozó következtetést levonni az adott esetben. Ha pedig az ilyen eseteket az elemzésből kiiktatjuk, akkor a kapott eredmény végül is nem a teljes mintán alapul. Ez különösen azért probléma, mert a válaszhiányok valószínűleg szintén nem véletlenszerűen oszlanak el a mintában; az ilyen esetek kihagyása szisztematikusan torzít.

E probléma kezelésére is kínálkoznak – ha nem is tökéletes – megoldások. A legegyszerűbb, ha a válaszhiányt helyettesítjük a mintaátlaggal (mean substitution). Ezzel sajnos az átlagos esetek számát növeljük mesterségesen, s torzítjuk az adott jelenséggel kapcsolatban a minta eloszlását. Tehetjük azt, hogy a válaszhiányt nem a mintaátlaggal helyettesítjük, hanem egy olyan átlaggal, amely a minta egy bizonyos szeletére, alcsoportjára jellemző. Ha a válaszhiány egy idős falusi asszonynál fordul elő, azt a mintában lévő idős falusi asszonyok válaszainak átlagával helyettesítjük. De az átlaggal való helyettesítés helyett választhatjuk azt a megoldást is, hogy a hiányzó adatot más meglévő adatok alapján egy regressziós függvénnyel becsüljük.<sup>15</sup> Valamennyi megoldás pótlólagos (ráadásul időigényes!), hiszen az lenne az ideális, ha nem merülne fel a válaszhiány problémája, de mégis meglehet az a jó érzésünk, hogy nemcsak sopánkodtunk, hanem próbálkoztunk valamivel a lehetőségek határán belül.

.....

Az érdemi elemzés megkezdése előtt azonban nemcsak a válasziányos eseteket kell valamilyen módon rendezni, hanem célszerű az adatok *konzisztenciájáról* is meggyőződni. Saját adatfelvételeim esetében – szívem szerint – heteket, olykor hónapokat töltök az adatok tisztításával (gyakorisági megoszlások és keresztátlák futtatásával), mielőtt az érdemi elemzéshez kezdek. A gyanús kérdőívet visszakeresem, ha tehetném – nem tehetem! –, sokszor visszaküldéném a kérdezőt egy-egy kérdés tisztázására. Bár úgy gondolom, a magyar társadalomban sok az inkonzisztencia, mégsem árt utánanézni például annak, ha diplomához kötött foglalkozású embereknek nincs felsőfokú végzettsége, vagy ha fizikai munkásoknak diplomájuk van. Jó lenne a létező státusinkonzisztencia mértékét nem megfélemleni mérési hibákkal. Az is gyanús, ha hagyományos mezőgazdasági foglalkozást űző kérdeztettek az iparban vagy a kereskedelemben szerepelnek az ágazatra vonatkozó kérdésnél. S egy utolsó példaként: bár sokat hallottam arról, hogy férfiak idősebb feleségüket fiatalabbra cserélték, mégis elbizonytalanodom, ha adatbázisomban tucatszámra találok olyan eseteket, ahol a kérdeztett és az anyja születési éve között nincs meg legalább 15 év különbség.

Az előzőekben felhozott példák *ténykérdésekre* vonatkoztak, nem véleményekre. Ha a visszakeresett kérdőívben kiderül, hogy kódolási, rögzítési hibáról van szó, a javítás egyszerű. A nagyobb probléma akkor van, ha a kérdőívben valóban az az adat (kód) szerepel, ami esetleg valószínűtlennek tűnik – például, hogy valaki előbb kezdett el dolgozni, semmint befejezte volna nappali tagozatos iskoláit. Mivel a kérdező többnyire nem küldhető vissza, hogy tisztázza a – látszólagos – ellentmondást, az eset vagy mégis kimarad az elemzésből, vagy a meglévő információk alapján kell vélelmezni és korrigálni az inkonzisztens adatot. Ez végül is elég önkényes megoldásnak minősíthető.

Úgy gondolom, hogy a ténykérdések megbízhatósága kapcsán a kérdezők felelőssége különösen nagy, és érzésem szerint ebben a vonatkozásban nincsenek is megfelelően kiképezve. A problémát abban látom, hogy a közvéleménykutatások általánosabb elterjedtsége miatt a kérdezők döntően arra vannak "szocializálva", hogy "minden válasz jó". Ha egy kérdező kedden becsönget egy lakásba és a kérdőívben szerepel egy olyan felteendő kérdés, hogy "Milyen nap van ma?", s a kérdeztett erre azt feleli, hogy "szerda", akkor a kérdező ezt a választ szemrebbecs nélkül bekarikázza. Ez a magatartás illeszkedik a kutatási célkitűzésekhez, ha a kutató arra kíváncsi, hogy az emberek *véleménye* vagy *tudomása* szerint milyen nap van. A vélemény- vagy ismeretszint-kérdésekkel szemben a ténykérdések esetében viszont a kérdező szerintem akkor járul hozzá az adatok nagyobb megbízhatóságához, ha észreveszi a válaszok esetleges ellentmondásait, ha hozzásegíti a kérdeztettet ahhoz, hogy minél konzisztensebben válaszoljon.<sup>16</sup> Ez a kérdezői magatartás persze nehezebben *standardizálható*, de nekem amúgy is az az érzésem, hogy a standardítás követelménye kissé mitikussá vált az adatfelvételekben. Attitűdkérdések itemjei esetében persze lényeges, hogy azokat mindig minden kérdező szó szerint ugyanúgy olvassa fel. Kétségesebb azonban számomra, hogy valóban szó szerint elő kell-e írni a kérdező számára, hogyan szerzi meg azt az információt, hogy a kérdeztett milyen gyakran jár színházba, vagy hogy milyen háztartási eszközök vannak a lakásában, s ezeket mondjuk mikor vásárolta. Lehet, hogy az ilyen kérdések esetében a *Kérdező, derítsd ki!* utasítás vezet megbízhatóbb válaszokhoz.

A válaszok megbízhatósága természetesen nem kizárólag a kérdezők kérdezési stílusán múlik. Korábbi példám arra vonatkozóan, hogy valaki előbb állt munkába, semmint kijárta

.....  
volna iskoláit, származhat abból is, hogy a *kutató* nem határozta meg pontosan, mit is ért első foglalkozás alatt, s a kérdező a nyári vakáció alatt végzett pénzkereső munkát is számításba vette.

Más a helyzet akkor, ha véleménykérdések konzisztenciájáról van szó. Itt kétségtelenül nehezebb elbírálni, hogy hol a határ, meddig tekinthetők – bizonyos ellentmondásokkal terheltlen is – az adatok konzekvensnek s mettől nem. Én azt az esetet, ha valaki egy itemsor valamennyi, különféle és egymás kontrolljaként értelmezhető állításával "teljes mértékben egyetért". valóban hajlamos lennék kidobni a vizsgálatból. Emellett, ha bizonyos tudattartalmakról és véleményegyüttesekről megbízható állításokat akarok tenni, akkor szembe kell nézmem a "rejtett véleményélküliség" problémájával is, például azon a módon, ahogy Angelusz Róbert javasolja.<sup>17</sup>

Mindezekon túl a válaszok megbízhatóságára vonatkozóan – a legelemibb kérdések esetében is – a legmegdöbbentőbb és legelbizonytalanítóbb tapasztalatokat szerezheti a kutató. Marton Ádám számol be arról, hogy az 1988-as KSH preszúfsvizsgálat esetében egy két-három héttel később végzett utóvizsgálat eredményei szerint a megkérdezettek 20%-ánál a két felvételből származó iskolaivégzettség-kód nem felelt meg egymásnak. Ezek után az már nem is csoda, hogy a két megkérdezés alkalmával 30% másképp ítélte meg, hogy a szülei befolyásolták-e a pályaválasztáskor, s 36% máshogy látta, hogy jelenlegi foglalkozása megegyezik-e azzal, ami eredetileg szeretett volna lenni. (Mindkét utóbbi kérdés esetében a legegyszerűbb igen–nem válaszalternatívák szerepeltek a mérésben.)<sup>18</sup>

Az olvasó ezen a ponton megkérdezheti: én most vitatkozom-e Léderer Pállal, avagy mellette érvelek? A válaszom: semmiképp sem kívánom kisebbiteni az általa felvetett problémát, de – mint ez az eddigiekből is remélhetőleg kiderült – *más értelmezési keretbe szeretném helyezni, s így részben más végkövetkeztetésre kívánok jutni*. A következő jogos kérdés: ha a megbízhatósággal kapcsolatban ilyen sok a gond, (tulajdonképpen jóval több, mint amennyit a vitaindító cikk felvetett), hogyan lehet egyáltalán survey-n alapuló empirikus szociológiát művelni? A válaszom egyszerű, s talán valakik számára cinikusnak is hangozhat, bár szerintem éppenséggel konform a Léderer által kívánatosnak tartott attitűddel: *ennek tudatában*.

Elhiszem például, hogy a szociológiában (mondjuk a pszichológiához képest) valóban kevesebb a kimunkált, ugyanannak a jelenségnek a mérésére visszatérően ugyanúgy alkalmazott skála, index, itemsor, s könnyen lehet, hogy "a bűnözés és a deviancia mérésére 64 különböző vizsgálatban 47 mértéket számolhatunk össze". De ha véletlenül azt tapasztalnánk, hogy ezeknek a méréseknek mindegyike mondjuk hasonlóan szignifikánsan negatívul korrelál a kedvező családi háttérrel, akkor talán mégis megengedhető egy csipetnyi önelégültség, talán mégis megbízhatónak tekinthető egy olyan tudományos állítás, hogy a bűnözés és a deviancia magasabb mértéke és a rosszabb családi háttér összefüggnek egymással. Hátha egy állítás megbízhatóságára vonatkozóan nemcsak azonos mérések egyezése alapján, hanem különböző (bár egymástól valószínűleg nem gyökeresen eltérő) mérések hasonló *működése* alapján is lehet következtetést levonni.<sup>19</sup>

A hazai szociológiai adatfelvételekre valóban nem jellemző a test–retest módszer alkalmazása. Elűthetném a problémát azzal, hogy ez is csak pénz kérdése, telik-e a kutatási keretből még egyszer kiküldeni a kérdezőt az emberekhez, (vagy legalább a minta egy –

.....

szintén reprezentatív – almintájához), és újra lekérdeztetni a kérdőívet (vagy legalább egy részét), de nyilván többről van szó. Valóban nem tudjuk, hogy milyen idő múlva ismételjük meg a mérést, hogy a válaszok megbízhatóságát mérjük vele. Azt sem tudjuk, hogyan kellene egyáltalán "eladni" az interjúszituációban azt a helyzetet, hogy a kérdező megint hecsönget az ajtón. Mit tegyünk, hogy a kérdező ne kezdjen óhatatlanul gyanakodni, hogy valójában az emlékezetét vizsgálztatják, miáltal törekedni fog arra, hogy minél inkább *ugyanúgy* válaszoljon. Vagy hogy ne higgye azt, hogy korábban "rossz" válaszokat adott, de most felkínálták neki a korrekció lehetőségét, miáltal törekedni fog arra, hogy minél inkább *másképp* válaszoljon. Úgy gondolom azonban, hogy ezeknek a kérdéseknek módszertani kutatásokat kell szentelni, talán éppen az OTKA támogatásával.

A következő probléma ugyis abban a pillanatban kezdődik, amikor a test–retest módszer alkalmazásának következtében ugyanarra a jelenségre többféle és eltérő eredményt mutató mérés áll rendelkezésre. Mi a teendő ebben az esetben? A válaszom valószínűleg gyökeresen eltérő attól, ahogy Léderer közelít a problémához. Szerintem ugyanis éppen itt érkezünk el a matematikai statisztika *szükségszerű* alkalmazásához. Lehet ugyanis, hogy az emberek kétszer egymás után nem tudják megmondani, mi is valójában az iskolai végzettségük. Még kevésbé tudják megmondani, hogy mondjuk hol nyaraltak az elmúlt öt évben. Én azonban elsődlegesen nem is erre vagyok kíváncsi, nem a társadalom iskolázottságáról akarok állításokat megfogalmazni, se nem arról, hogy az emberek hová járnak nyaralni Magyarországon. Engem mondjuk inkább az érdekel, hogy az iskolai végzettség hogyan határozza meg egy oksági modellben azt, hogy kinek hová van lehetősége elmenni nyaralni. S mit tesz Isten, mindkét adatbázison, a primer mérések *eltérése ellenére* azt tapasztalom, hogy az iskolai végzettség alapján egy *hasonló regressziós egyenlettel* tudom becsülni mondjuk annak a valószínűségét, hogy valaki elment-e egyáltalán nyaralni az elmúlt öt évben, s ha igen, járt-e ez alatt az idő alatt külföldön! Itt ugye nem arról van szó, hogy primer adatok azonos eloszlása alapján feltételeznék valamit a kompozitadatok eloszlásának megbízhatóságáról. Itt éppen eltérő primer adatok ellenére jutok statisztikailag hasonlóan alátámasztott eredményre. Természetesen a statisztikai becslések megbízhatósága egy bizonyos *konfidencia-intervallumon* belül érvényes. A jelen fiktív példa esetében éppen az történt, hogy *mindkét* mérés belül volt egy bizonyos maximum és minimum érték közötti *range-en*, s *mindkét* regressziós becslés belül volt ugyanazon a konfidencia-intervallumon: hasonlóság alatt ezt értettem. Akkor most mit szabad nekem a nyaralási lehetőségek és az iskolázottság társadalmi összefüggéseiről kijelentennem?

Összefoglalva a véleményemet Léderer Pál tanulmányáról: Elfogadom, hogy az általa felvetett probléma létezik, de vitatom kritikai attitűdjének *irányát*. Nem osztom *indulatát*, hiszen amit állít, azt előtte már jóval tárgyilagosabban megfogalmazták mások, *ímgyén*: "a társadalomtudományok elértek arra a pontra, ahol a modellek minősége messze meghaladja maguknak a méréseknek a minőségét".<sup>20</sup> Az én olvasatomban Léderer írása úgy hat, hogy miközben a szerző a mérések minőségével kapcsolatban "ésszerű kételyeknek" ad hangot, a matematikai statisztikai módszerek alkalmazásának elutasításával a modellek minőségét *leszállítaná* a mérések minőségének szintjére. Érzésem szerint nem ez az út vezet a megoldáshoz. Úgy gondolom, amíg méréseink megbízhatóságát nem tudjuk jelentősen javítani, éppen a statisztikai eszközök fokozott alkalmazása visz közelebb a tudományosan megbízhatóbb eredményekhez.<sup>21</sup>

.....

Léderer végső soron, írásának utolsó bekezdésében, egy attitűd hiányát kifogásolja, a "saját tevékenységünkre való szüntelen kritikai reflexió"-t. Ezt még el tudnám fogadni, ha ez a hiány nem minősülne a következő mondatban mindjárt *műfajhoz kötődő mentalitásnak*, s nem kapcsolódna a tudományosság kisajátításának vádjához. A "saját gyöngé pontjaink" vizsgálatának hiányát ugyanis, amennyiben ez valóban fennáll, én Lédererrel szemben kevésbé tudom műfajokhoz és stílusokhoz (vagy kevésbé rébuszokban beszélve: statisztikai modellekkel dolgozó vagy azokat elvető kutatókhoz) kötni. Véleményem szerint akkor már ennél *általánosabban elterjedt* jelenségről van szó. Bár felvetném azt is: azért mert valaki tanulmányában nem szentel mindig két bekezdést annak, hogy munkája során milyen megbízhatósági problémákba ütközött, ez már bizonyíték arra, hogy semmiféle kétség nem merült fel benne saját eredményeivel kapcsolatban?

Nem tudom persze azt sem, hogy Léderer, amikor kifogásol bizonyos típusú mentalitásokat és hiányol bizonyos attitűdöket, szaktudományos publikációkra gondol, vagy olyan megnyilatkozásokra, amelyek nem a tudományos közösség számára, hanem a média fogyasztóinak, illetve a politikacsinálóknak készültek. Ez utóbbi esetben ugyanis figyelembe kell venni a *társadalmi közeget* is. Gondoljuk meg: ha egy riporter egy szociológussal beszélget, habozás nélkül vitába száll vele és személyes tapasztalataira hivatkozik, arra, hogy mit lát például az utcán járva. Előfordulhat az, hogy egy riporter egy csillagással beszélgetve úgy kezdjen érvelni, hogy ha ő kinéz éjjel az ablakon, úgy látja az égen, hogy...? Nem valószínű! Pontosan kifejeződik ebben az, hogy a csillagászat olyan tudomány, amelyről feltételezhető, hogy műszerei és mérései révén olyasmit és úgy "lát", ahogy az "szabad szemmel" nem lehetséges. Ez a feltételezés nem áll fenn a szociológiával szemben. Ez 2021-t persze nem a riporter a felelős. Hiszen Léderer cikke éppen azt sugallja, hogy a szociológia műszerei és mérései valóban nem állják meg a helyüket, vagy legalábbis az egész sokkal vitathatóbb, mint a csillagász távcsöve, illetve számításai. Lehet, persze, hogy majd valamikor a mai csillagászati műszerek és mérések is megkérdőjeleződnek. Az is nagyon valószínű, hogy a csillagásztól – csakúgy, mint a szociológustól – elvárható, hogy kételkedjen saját kutatási eredményeiben, hogy folyamatosan keresse az újabb és újabb bizonyítékokat azok megbízhatóságára.

A társadalmi közeg tehát visszahat a kutatói magatartásra és mentalitásra. A hazai szociológia professzionalizálódása, amennyiben megindult, kevésbé szaktudományos, hanem inkább *szakértői, tanácsadói* formát öltött az elmúlt évszázadokban. Egyre többen ambicionálják azt, hogy kevésbé a tudományos közösség, inkább a politikacsinálók számára dolgozzanak, nekik szolgáltassanak érvanyagot. Az ilyen munkákban a tudományos megbízhatóság könnyen háttérbe szorul a gyorsaság és a hasznosság szempontjai mögött. Az újságcikk, a gyorsjelentés, a kormányzati megrendelőnek írt tanulmány műfajától idegen a Léderer által emlegetett kritikai reflexió attitűdje. A média nem kíváncsi a kutatói kétségekre, a tudományos piac nem vevő bizonytalan válaszra, csak a kérdésekre megfogalmazott határozott és kétséget nem ismerő állításoknak van értékük.

Mindez rendjén is lenne, s valószínűleg a fenti láttelelet sem hazai specifikum. A baj ott kezdődik, amikor az alkalmazott szociológusi, kutatói szerepnek egyre kevesebb alternatívája kínálkozik. Nálunk gazdagabb országokban ugyanis – még ha a mostani recesszió idején a helyzet sokat romlott is – a tudományos kutatóknak is van presztízse, finanszírozója és – ebből következően – megfelelő egzisztenciát biztosító jövedelme. Ennek hiányában lehet, hogy továbbra is várhat magára az a bizonyos, a "tudomány lényegétől elválasztha-

.....

atlan attitűd". Helyette marad a *kétes megbízhatóság* az adatokban és – úgy tűnik – egymásban: kérdezőben, kódolóban, rögzítőben és statisztikai módszerekkel "bűvészkedő" szociológus kollégában.

## Jegyzetek

1. Bizonyos állítások azért talán mégis vannak. Létezik például egy OTKA által támogatott kutatás a társadalomtudományi mérésről. Ennek keretében jelent meg *A mérés problémája a társadalomtudományokban*, OMIKK–TÁRKI, Bp. 1993. című kötet.

2. Az egyik klasszikus munka a témában Herbert Hyman et al.: *Interviewing in Social Research*, Chicago, The University Press of Chicago. 1954.

3. Meglehetősen széles áttekintést nyújt minderről a *Telephone Survey Methodology* című munka, R. M. Groves et al. szerkesztésében (John Wiley & Sons, New York, 1988). A kötet megtalálható a TÁRKI könyvtárában.

4. Az egykori Tömegkommunikációs Kutatóközpont Módszertan-sorozatának több darabja foglalkozott olyan kérdésekkel, amelyek itt felmerülnek. (Ezek az írások jelenleg az MTA–ELTE Kommunikációelméleti Kutatócsoportjának Könyvtárában találhatóak meg.) Közülük az egyik éppen a válaszok megbízhatóságát vizsgálta interjúszituációs kérdések és kérdezői jelentések alapján. (G. Csilléri Klára: *A válaszok megbízhatósága a közvéleménykutatásban*, TK Módszertan, Bp. 1976. 5. szám.)

5. A *szociológiai felvétel módszerei* című könyvben (Szerk. Cseh-Szombathy László és Ferge Zsuzsa, KJK. 1975) olvasható a postai kérdezőssel kapcsolatban: "A kérdezők tevékenységében több hibaforrással kell számolni, ami meglehetősen aláássa a felmérési eredmények megbízhatóságát és érvényességét, és megnyugtató az a tudat, hogy ilyesmire nincs gondunk" (179. old.). Ugyanitt, illetve a következő oldalakon szerepel az az állítás, hogy bizonyos ténykérdések esetében éppen akkor lesz megbízhatóbb a válasz, ha az családi megbeszélésen alapul, vagy ha több a gondolkodási idő, lehetőség van dokumentumok előkeresésére. Nyilvánvaló, spontán véleménykérdések esetében ez nem áll fenn.

6. Lásd Tarján József tanulmányát: "Lehet-e telefonon közvéleményt kutatni?"

7. A telefonos adatfelvétellel szemben akkor is több fenntartásom lenne, ha Magyarországon ezen az úton is elfogadható, országosan reprezentatív mintához lehetne jutni, mivel a külföldi mode effects vizsgálatok valamivel magasabb válaszmegtagadói arányt, illetve félbeszakadt adatfelvételt jeleznek. Az adatfelvételek során elterjedt válaszlapok, illetve kártyák is alig-alig alkalmazhatók telefonos kérdezés esetén. Seymour Sudman és Norman M. Bradburn (*Asking Questions. A Practical Guide to Questionnaire Design*, Jossey-Bass Inc., Publishers, London. 1982.) javaslata egy telefonos kérdőív egyik kérdésére vonatkozóan szinte komikusan hangzik: Tekintse a telefonkészülékén látható számokat egy skálának! Ha a 9-es vonatkozik azokra, akiknek nagyon jól megy és az 1-es azokra, akiknek egyáltalán nem megy jól, hová helyezné ön saját magát? (271. old.).

8. A számítógépes adatgyűjtésről és annak különböző válfajairól, mint például a computer-assisted telephone interviewing (CATI), a computer-assisted personal interviews (CAPI) lásd Willem E. Saris: *Computer-Assisted Interviewing*, Sage, London. 1991. A könyv megtalálható a TÁRKI könyvtárában.

9. Az említett *Telephone Survey Methodology* című munka egyik tanulmánya szerint a telefonos kérdezésben a nyitott kérdésekre adott válaszok rendszerint rövidebbek, mint a kérdőíves kérdezésnél, aminek oka, hogy a telefonon keresztüli kommunikáció általában egyszerűbb és lényegre törőbb információátadást jelent. Valószínű, hogy a nyitott kérdésekre az önkéntes kérdőívben is rövidebben válaszolnak az emberek, mert nem szeretnek annyit írni, amennyit esetleg előszóban elmondának (Seymour Sudman és Norman M. Bradburn: *i. m.*)

10. A nyitott és zárt kérdések problémájával, de számos más idcvonatkozó kérdéssel is foglalkozik Jean. M. Converse és Stanley Presser: *Survey Questions* című (Sage, London, 1986) könyve. Megtalálható a TÁRKI könyvtárában.

11. A data entry programokról lásd Linda B. Bourque és Virginia A. Clark: *Processing Data*, Sage, London, 1992. Megtalálható a TÁRKI könyvtárában.

12. Egy, az 1990-es népszámlálás adataira építő súlyozó program készült a TÁRKI-ban.

13. Csak megjegyzem, hogy a mintavételből, mintatorzulásból származó hibákról írva Graham Kalton: *Introduction to Survey Sampling* (Sage, London, 1983) című könyvében a *legalább négyeszeri* megkeresést javasol azoknál a személyeknél, akiket nehezen lehet otthon találni, szemben a hazai háromszori megkeresés gyakorlattal.



14. Tapasztalataim szerint igen gyakran eredményez válaszhányt a kérdőívbe beépített *filter-* vagy szűrőkérdés, aminek az lenne a célja, hogy egy adott kérdést csak a releváns válaszolóknak tegyenek fel. Például a "Milyen típusú autója van Önnek?" vagy a "Mi a véleménye a tv-híradóról?" kérdéseket logikusan megelőzhetik a "Van-e autója?", illetve "Nézi-e a tv-híradót?" kérdések. A gyakorlatban sajnos a kérdező sokszor nem csak akkor "ugorja át" az érdemi kérdést, amikor tagadó választ kap a technikai filterkérdésre, hanem akkor is, amikor azt fel kellene tennie. Ez az információvesztés az adatfelvétel során nagyobb baj szerintem, mintha "irreleváns" kérdést tennénk fel, tehát valakinek az autójáról érdeklődnénk "bután", mielőtt tisztáoznánk, hogy van-e neki egyáltalán olyan.

15. Lásd ezekről a módoszatokról, illetve a sampling és a non-response error problémáiról Graham Kalton, *i. m.* A könyv megtalálható a TÁRKI könyvtárában.

16. A German Life History kutatás esetében például éppen úgy instruálták a kérdezőket, hogy az élettörténet eseményeit igyekezzenek a kérdezettekkel *együtt rekonstruálni*, s kérdezzenek vissza, ha ellentmondást tapasztalnak bizonyos válaszok között.

17. Angelusz Róbert: *Kommunikáló társadalom*, Gondolat, Bp. 1983, 165. és következő oldalak.

18. Marton Ádám: *A válaszadási hiba vizsgálata*. ELAR Módszertani tanulmányok, KSH, Bp., 1990. Megjegyzendő, hogy az utóvizsgálat kérdezője *más személy* volt, mint az eredeti kérdező. Az esettanulmány ebben a tekintetben tehát nem pontosan ugyanazzal a mérőműszerrel végzett teszten alapul (amennyiben a műszer valóban az "ember + kérdőív"). Az eredmények optimista interpretációja így a mérési hibák bizonyos hányadát a kérdezői hatásnak tulajdoníthatja. (A KSH összeírói ezen kívül gyakran nem is kapnak olyan típusú kiképzést, mint a szociológiai felvételek kérdezői.)

19. Lehet, hogy a Bonjean és munkatársai által feltárt helyzetkép (Charles M. Bonjean, Richard J. Hill és S. Dale McLemore: *Sociological measurement: An inventory of scales and indices*, Chandler Publishing Co., San Francisco, 1967) sok tekintetben nem változott, bár a korábban említett OTKA-kutatás a társadalomtudományi mérésről ezt nem minden téren támasztja alá. Bizonyos vonatkozásban, például éppen a rétegződés- és mobilitáskutatásokban a foglalkozás mérésénél, megfigyelhető egyfajta kikristályosodás a presztízs-pontszámok, a SEI típusú skálák, vagy az Erikson–Goldthorpe–Portocarero osztályozás alkalmazása révén. (Lásd Fábrián Zoltán: *Sociological measurement: Individual variables in AJS and ASR*, 1986–1990, TÁRKI, Bp. 1993)

20. G. W. Bohrnstedt és E. F. Borgetta: *Social measurement*, Sage, London, 1981. Idézi Kolosi Tamás és munkatársai: *A mérés problémája a társadalomtudományokban*, *i. m.* 1. old.

21. Lásd erre Bornemisza Eszter több tanulmányát a hibásan besorolt esetek statisztikai elemzéséről. Például: "Kötés mintavételi eljárás hibásan besorolt adatok elemzésére", In *A mérés problémája a társadalomtudományokban*, *i. m.* 58–79. old.