

Metodološki aspekti istraživanja iseljeništva web-anketama¹

Stanko RIHTAR

Institut društvenih znanosti Ivo Pilar, Zagreb

Izvorni znanstveni rad

Primljeno: 22. 7. 2022.

UDK 314.151.3.02-054.7:[303.621.3:004.738.5]

[303.621

doi: 10.5559/pi.17.32.04

Anketna istraživanja putem interneta popularnost ponajprije duguju ekonomičnosti i neograničenom teritorijalnom obuhvatu, uključujući mogućnost dosega inače nedostupnih (ili teško dostupnih) populacija. Uz to, važne su i tehničke pogodnosti koje mogu unaprijediti ne samo kvalitetu anketnog postupka, nego i naknadno baratanje podacima i nalazima. Zbog porasta pokrivenosti, usvojenih navika upotrebe interneta, porasta informatičke pismenosti i prikupljenih metodoloških iskustava, reprezentativne ambicije takvih istraživanja sve su opravdane: valjanošću i metrijskim kvalitetama u mnogim su slučajevima dostigla (a u nekim i prestigla) klasične tehnike. Budući da ih velika teritorijalna disperzija ne sputava, anketiranja putem interneta posebno su pogodna (umnogome i jedino pogodna) za istraživanja iseljeništva. No pritom se suočavaju s dvjema glavnim preprekama. Prva je nedovoljna poznatost populacije (njezina opsega i strukture) u cjelini i većine uže ciljanih potpopulacija. Druga su oskudni okviri uzorkovanja. Jedno i drugo bitno otežava ili onemogućuje izbor uzoraka koji bi opravdali širu generalizaciju nalaza. Uže metodološka obilježja, prednosti i nedostaci online anketnih istraživanja u prvom su dijelu rada sagledani iz perspektive ukupne anketne pogreške koja, osim psihometrijske ili pogreške mjerenja, obuhvaća i pogreške autoselekcije, obuhvata i uzorka. U drugom su dijelu prikazane načelne mogućnosti i dosezi istraživanja iseljeništva, posebno vodeći računa o naznačenim poteškoćama koje im sputavaju reprezentativne ambicije.

Ključne riječi: anketiranje, internet, iseljeništvo, anketna pogreška, reprezentativnost

¹ Rad je rezultat istraživanja provedenog u sklopu projekta »Zajedno srce, jedna duša, jedna HRVATSKA« (UP.04.2.1.06.0051) sufinanciranog sredstvima Europske unije iz Europskoga socijalnog fonda na poziv »Tematske mreže za društveno-ekonomski razvoj te promicanje socijalnog dijaloga u kontekstu unapređivanja uvjeta rada«. Dio istraživanja koji je rezultirao ovim radom obavljen je u sklopu dionice 2 projekta, pod nazivom »Suvremeni motivacijski ciklusi migriranja mladih Hrvatske u kontekstu općih migracijskih procesa u Hrvatskoj — empirijska analiza«.

Uvod

Anketiranje upotrebom interneta u znanstvene svrhe započela su krajem osamdesetih godina 20. stoljeća slanjem upitnika elektroničkom poštom (Schaefer i Dillman, 1998). Iako znatno ekonomičnije, prikupljanje podataka na taj način uglavnom dijeli nedostatke povezane s njihovim prikupljanjem »običnom« poštom (Fricker i Schonlau, 2002). Zbog toga nije posebno zaživjelo, no i dalje se upotrebljava za istraživanje užih i poznatijih populacija o socijalno manje osjetljivim temama (Lamza Posavec, 2021).

Drugi način anketiranja putem interneta, postavljanje upitnika na web-stranice ili portale, postalo je znatno popularnije. Glavne su mu prednosti ekonomičnost, neograničen teritorijalni doseg (uključujući varijabilnost i mogućnost obuhvata inače nedostupnih ili teško dostupnih populacija), vremenska neograničenost i jednostavnost postupka, uz velike mogućnosti tehničkog unaprjeđenja svih faza istraživanja.

Dio osmišljavanja, konceptualizacije i operacionalizacije istraživanja koji se odnosi na izradu upitnika i uzorkovanje ne zahtijeva manji trud i vrijeme u usporedbi s ostalim vrstama anketnih istraživanja. Presudne uštede mogu se zahvaliti automatiziranom prikupljanju podataka koje inače izaziva glavne financijske troškove i zahtijeva ulaganje najvećeg istraživačkog truda i vremena (Andrews i sur., 2003; Llieva, Baron i Healey, 2002).

Osim prikupljanja, i baratanje podacima jednostavnije je i brže. Automatizirano bilježenje odgovora jednostavnije je i manje podložno pogreškama. Odvija se u realnom vremenu, pa podatci nisu dostupni samo po završetku, nego i tijekom anketiranja. To podrazumijeva stalnu mogućnost kontrole, tako da uočavanje pogrešaka na vrijeme omogućuje modifikacije postupka i ispravljanje pogrešaka »u hodu«, pa i ponavljanje cijelog istraživanja ako je potrebno (Prasad Nayak i Narayan, 2019). Uz to što cijelo istraživanje može biti kraće, kraće može biti i samo ispunjavanje upitnika (prema nekim provjerama, i više nego dvostruko; Chang i Crosnick, 2010).

Konkurencija na tržištu informacijsko-komunikacijskih tehnologija unaprijedila je brojnost, raznovrsnost i kvalitetu aplikacija i platformi za anketiranje, uključujući sve veću jednostavnost njihove upotrebe na različitim uređajima. Potonje podrazumijeva posebno relevantnu mogućnost obuhvata korisnika oskudnije informatičke pismenosti. No prednosti se, u tom pogledu, ne nude samo potencijalnim ispitanicima, nego i istraživačima. Početkom stoljeća izrada i provedba online anketa zahtijevala je poznavanje HTML koda, programiranja i izrade web dokumenata. S vremenom je to postalo nepotrebno zbog sve šireg i prilagodljivog spektra gotovih softverskih rješenja. Osim što su jednostavna za upotrebu, mogu se birati prema ciljevima, roko-

vima i financijskim mogućnostima. Naprednija među njima nude pomoć pri oblikovanju nacrti i operacionalizaciji istraživanja. Uz to, nakon prikupljanja podataka omogućuju njihovu obradu i analizu, uz prezentaciju i diseminaciju nalaza.

Općenito, tehničke su prednosti upotrebe interneta raznovrsne i brojne, od mogućeg uvođenja virtualnih anketara zbog simulacije personaliziranog postupka do upotrebe audiovizualnih sadržaja. Takvi se sadržaji najčešće primjenjuju u marketinškim istraživanjima, no upotrebljivi su i univerzalno, među ostalim i zbog prevladavanja jezičnih barijera (tzv. *image-based surveys*).

Zbog nedovoljne pokrivenosti opće populacije internetom, nedostatne informatičke pismenosti većine i nedostupnosti korisnika za regrutaciju, primjena web-anketa u znanstvene svrhe u početku je bila ograničena na istraživanje užih i dostupnijih populacija. No porastom pokrivenosti, stasanjem novih, informatički pismenijih naraštaja i navika svakodnevne upotrebe interneta znanstvena istraživanja širih ili općih populacija nailaze na sve manje ograničenja.

Zahvaljujući višestrukim prednostima, ta je tehnika prikupljanja podataka razmjerno rano u zemljama najveće pokrivenosti (poput SAD-a, Njemačke, Velike Britanije i skandinavskih zemalja) stala uz bok ostalim načinima anketiranja (Galešić, 2005), uključujući i terensko licem u lice kao »zlatni standard«. Osim istraživačkih ustanova i agencija, i statistički uredi sve češće dugotrajno i skupo terensko prikupljanje podataka zamjenjuju online anketama: primjerice, američki ured (U. S. Census Bureau) godišnje provodi više od 130 takvih istraživanja (Pew Research Center, 2022).

Metodološki aspekti anketnih istraživanja putem interneta

Uže metodološki aspekti (obilježja, prednosti i nedostatci) prikupljanja podataka web anketama u brojnim se pregledima obično sagledavaju u svjetlu ukupne anketne pogreške koja, osim psihometrijske ili pogreške mjerenja, uključuje i pogreške odaziva, obuhvata i uzorka (Couper i Bosnjak, 2010; Dillman, 2007; Frippiat i Marquis, 2010; Groves, 1989). Držeći se naznačene perspektive, izvojit ćemo samo ključna komparativna obilježja, prednosti i nedostatke spomenute tehnike prikupljanja anketnih podataka — detaljniji bi pregled premašio enciklopedijski opseg.

Mjerenje (anketni postupak)

Anketna istraživanja putem interneta gotovo su redovito asinkrona i podrazumijevaju samostalno ispunjavanje upitnika; sinkrono sudjelovanje anketa-

ra putem nekog od servisa znatno je manje ekonomično i iznimno je rijetko. Općenito, istraživanja koja uključuju komunikaciju s anketarom podrazumijevaju jednokratno odgovaranje, i to u vrijeme prvog ili dogovorenog naknadnog kontakta. Za razliku od toga, asinkrono samoispunjavanje podrazumijeva mogućnost da se odgovara koliko se dugo i kad se želi ili može, i to višekratno (ne i višestruko, što se blokira). Takve okolnosti mogu pogodovati primjeni opsežnijih upitnika, no ipak se preporučuje da budu kratki kako bi se mogli ispuniti jednokratno i kako gubljenje motivacije, koncentracije ili zamor ne bi naveli na površno odgovaranje ili odustajanje. Ako su složeniji, uglavnom se preporučuje da ispunjavanje ne traje više od petnaestak minuta. Ispunjavanje jednostavnijih i zanimljivijih može trajati i nešto duže, bez obzira na opseg i formu (primjerice, veličinu i obuhvat stranica, tzv. *scrolling* ili *paging* oblikovanje i slično; Conrad, 2002; Ganassali, 2008).

Pritom treba u obzir i to da boravak na internetu nije primarno namijenjen sudjelovanju u anketama (Galešić, 2005; Tourangeau i sur., 2013). Osim dodatnog utroška vremena, ono može izazvati neplanirane financijske izdatke i uskratiti online vrijeme ili komunikaciju ukućanima (zahvaljujući tehničkom napretku i konkurenciji na tržištu komunikacija, potonje je ograničenje danas ipak marginalno).

Od ostalih prednosti samoispunjavanja najvažnije su veća anonimnost (koju je, uz uobičajena verbalna jamstva, moguće dodatno zaštititi blokiranjem IP adresa i adresa elektroničke pošte) i standardiziranost postupka nego u slučaju anketa koje uključuju komunikaciju s anketarima. Anonimnost pridonosi većem odazivu, spremnosti odgovaranja na socijalno osjetljiva pitanja i iskrenosti (smanjuje tzv. *non-response bias*; Kreuter i sur., 2008; Yaeger i sur., 2011). No istodobno može podrazumijevati i manju odgovornost ili površnije odgovaranje (*satisficing*; De Leeuw, 2005; Heerwegh i Loosveldt, 2008). Nalazi usporedbi između web anketa i ostalih tehnika u tom pogledu nisu jednoznačni jer način odgovaranja ne ovisi samo o izabranoj tehnici prikupljanja podataka, nego i o razlikama između istraživačkih tema i pristupa općenito (Lamza Posavec, 2021; Toepoel i sur., 2009).

Osim izbjegavanja neugodnosti povezanih sa socijalno osjetljivim pitanjima, standardiziranim samoispunjavanjem izbjegavaju se pristranosti anketara i pogreške pri bilježenju odgovora.

No takav (kruti) postupak ima i nedostataka. Sudjelovanje anketara omogućuje prilagođavanje pitanja razini razumijevanja sugovornika, procjenu iskrenosti odgovaranja (primjerice, čitanjem govora tijela), odnosno neposrednu kontrolu kvalitete odgovaranja i anketne situacije općenito. Pokušaji da se izostanak anketara nadomjesti opširnijim uputama uglavnom ne daju rezultate, kako zbog unificiranosti, tako i zbog ignoriranja (pogotovo

ako su objašnjenja duga ili smještena iza pitanja; Redline, 2013). No eksperimentalne provjere (primjerice, Chang i Krosnick, 2010) pokazuju da samostalno ispunjavanje upitnika ipak može dati konkurentno valjanije odgovore i u slučajevima veće kognitivne ograničenosti. Mogući je razlog više vremena na raspolaganju za samostalnu elaboraciju pitanja, bez samoprezentacijskih motiva i želje da se udovolji anketarima, ne samo socijalno poželjnim nego i brzim (i stoga površnim) odgovaranjem.

Kad je riječ o redoslijedu pitanja, osim logičkog, u anketnim je istraživanjima općenito važan i psihološki: potrebno je paziti na halo-efekt i osjetljivija pitanja postaviti na kraju upitnika kako bi se smanjilo osipanje uzorka. Samoispunjavanje upitnika pri anketiranju običnom ili elektroničkom poštom omogućuje uvid u cjelinu, čime se poništavaju prednosti takvog redoslijeda, no prilikom anketiranja putem web stranica moguće ga je kontrolirati softverskim rješenjima: pitanja mogu biti sekvencijalno dostupna, naknadna se mogu uvjetno granati ili automatski preskakati (primjerice, tzv. *skip logic*) bez posebnih uputa (kojih se inače ne mora pridržavati ili se mogu pogrešno shvatiti). Osim toga, ako je potrebno automatski se mogu različito (ili po slučaju) rasporediti unutar svakog upitnika. Na isti se način može neutralizirati i mogući utjecaj redoslijeda ponuđenih odgovora u zatvorenim pitanjima (tzv. *primacy effect*; Malhotra i sur., 2008).

Automatizacijom se, kao i u prethodno spomenutim slučajevima, izbjegavaju i pogreške anketara: sve zajedno naknadnu logičku kontrolu odgovora na vezana pitanja čini nepotrebnom. Osim logičke, u sam se postupak mogu ugraditi i druge vrste kontrole, poput provjere vremena posvećenog svakom pitanju, konzistentnosti, pridržavanja raspona odgovora i slično (uključujući i upozorenja ako se ne odgovara na formalno propisani način; Galešić, 2005).

Na kraju, osim što online prikupljanje podataka uključuje velike tehničke prednosti, kvaliteta odgovaranja tim putem može biti ometena tehničkim poteškoćama: primjerice, ako kvaliteta korisničkih uređaja nije zadovoljavajuća. Iako se to ograničenje svakom novom generacijom smanjuje (uz automatsko prilagođavanje upitnika različitim vrstama uređaja), neka druga ostaju i dalje jednako prisutna. Primjerice, grafička rješenja i popratni sadržaji na stranici mogu biti (nenamjerno) sugestivni (Galešić, 2005), dok se pitanja tehnički složenijeg formata češće preskaču bez asistencije anketara (Gräf, 2002). Iako se pristup može kontrolirati (uključujući uskraćivanje višekratnog ispunjavanja; Konstan i sur., 2005), ni okolnosti u kojima se odgovara nisu poznate pa upitnik mogu ispuniti ili sudjelovati u ispunjavanju i osobe kojima nije namijenjen. No takvi su slučajevi ipak rijetki: osim zlonamjernosti ili zabave, mogu biti potaknuti i »plemenitijim« motivima. Primjerice, ako oso-

ba koja je izabrana u uzorak želi odgovarati, ali nije dovoljno informatički pismena pa traži usluge kompetentnije osobe ili »pomoćnog anketara« (čime ugrožava metodološke prednosti samoispunjavanja). Anketna situacija, u ovom slučaju i općenito, može se donekle kontrolirati upotrebom web-kamera, no to je u većini slučajeva neekonomično i narušava anonimnost postupka.

Odaziv

Odaziv se načelno može podijeliti na apriorni i preostali nakon uvida u sadržaj upitnika, koj se uoče ne moraju ili se mogu samo djelomočno ispuniti (tzv. *jedinični* i *čestični* odaziv; Groves, 1989). Kao i u slučaju ostalih tehnika samoispunjavanja, i u slučaju web-anketa apriorna je autoselekcija izrazito velika. Pregledi i metaanalize pokazuju da je odaziv na jednokratna istraživanja putem interneta najmanji (u usporedbi sa svim ostalim tehnikama; Baker i sur., 2010; Shish, 2008). Osim toga, većina onih koji se odazovu upitnik ispune samo djelomično (Vehovar i sur., 2000). No i odaziv, poput kvalitete odgovaranja, ne ovisi samo o tehnici ili mediju prikupljanja podataka, nego i o istraživačkim temama i pristupima općenito; primjerice, Lozar Manfreda i sur., 2008).

Uzroci autoselekcije mogu biti mnogobrojni, od onih na koje nije moguće utjecati (sociokulturni kontekst, osobna obilježja, aktualna situacija, manjak vremena, čestina i navike upotrebe interneta), do predmeta istraživanja, obilježja i tehničke realizacije upitnika, poput nezainteresiranosti za temu, zasićenosti, nepovjerljivosti, nemogućnosti utjecaja na okolnosti prvog kontakta, dužine upitnika, socijalne osjetljivosti pitanja, manje neugodnog odbijanja ili odustajanja nego u slučaju komunikacije s anketarima ili (sve manje prisutnih) tehničkih poteškoća (spora komunikacija, prosljeđivanje poziva elektroničkom poštom na pogrešne ili nepostojeće adrese, svrstavanje u »smeće« primatelja i slično (Hoogendoorn i Daalmans, 2009).

U slučaju istraživanja putem interneta odaziv se dijelom može potaknuti iskorištavanjem nekih prednosti koje nudi sam medij. Sažeta iskustva pokazuju da je znatno važnije: personalizirati kontakte u najvećoj mogućoj mjeri i naglasiti vjerodostojnost stranica (što povećava povjerljivost, koja ne utječe samo na odaziv nego i na iskrenost odgovaranja); višekratno apelirati na odaziv različitim kanalima i na različite načine, uz ponudene poveznice na upitnik; uputiti kratke i jasne pozive i izbjegavati njihovo svrstavanje u neželjenu poštu, artikulirati jasne upute i omogućiti jednostavan pristup upitniku, odgovarati na upite kontaktiranih osoba, kreirati privlačno, zanimljivo i informativno pristupno sučelje (konkurentno ostalim sadržajima na internetu) i izbjegavati komercijalne *bannere*, ponuditi različite verzije stranica

(uključujući višjejezične ako je potrebno), unaprijed zahvaliti na sudjelovanju, ekstrinzično ga i intrinzično motivirati (nagraditi ga i naglasiti njegovu širu društvenu korist) (Blumenberg, 2019; Dillman i sur., 2009; Reips, 2000). No pritom je nužan oprez kako se ne bi dodatno povećala motivacijska pristranost uzorka (ionako redovito prisutna u svim vrstama anketnih istraživanja; Groves i sur., 2004). Osim toga, kako bi se smanjilo odustajanje tijekom postupka (čestična autoselekcija) prije ispunjavanja korisno je provjeriti ozbiljnost namjere sudjelovanja, upotrijebiti tehniku »visoke prepreke« (teža pitanja, zbog kojih bi se poslije moglo odustati, postaviti na početku) i predvidjeti vrijeme za zagrijavanje (tijekom kojeg se ujedno donosi konačna odluka o sudjelovanju; Reips, 2002).

Obuhvat, uzorak i uzorkovanje

Kako bi istraživanje bilo reprezentativno, uzorak mora proporcionalno obuhvatiti svu relevantnu varijabilnost ciljane populacije prema obilježjima koja mogu utjecati na predmet mjerenja. Osim toga, konačni izbor jedinica u uzorak, bez obzira na prethodne stratifikacije, mora biti slučajan zbog dvaju razloga. Prvo, kako bi se izbjegla pristranost (u tom slučaju svaka jedinica mora imati jednaku vjerojatnost izbora, ili barem unaprijed poznatu kako bi se pristranost mogla izračunati); drugo, kako bi se moglo zaključivati o populaciji na temelju poznate teorijske distribucije statistika (vrijednosti zabilježanih ili izračunatih na uzorku) koja omogućuje procjenu raspona sigurnosti unutar kojeg se, s unaprijed definiranom vjerojatnošću, mogu očekivati populacijske vrijednosti ili parametri.

U idealnim okolnostima, uz slučajni izbor generalizacija je opravdana ako su ciljana populacija, osnovni skup (potpuna ili dio populacije iz koje se izabire uzorak) i okvir uzorkovanja (popis jedinica osnovnog skupa) isti. Kad je riječ o web-anketama na općim populacijama, glavne su prepreke postizanju reprezentativnosti reduciranost osnovnog skupa (pokrivenosti internetom) i nepostojeći okviri uzorkovanja (adresari ili popisi kontakata osnovnog skupa), odnosno pogreške obuhvata i uzorka, uz poteškoće povezane s načinom izbora jedinica.

Reduciranost osnovnog skupa

Iako se pokrivenost internetom kontinuirano povećava, još uvijek nije potpuna. Nedostupnost (uglavnom) starijih, siromašnijih ili slabije obrazovanih slojeva dokazano se odražava u razlikama između nalaza dobivenih istraživanjima putem interneta i nalaza istraživanja dobivenih na reprezentativnim uzorcima (Couper, 2000; Lamza Posavec, 2015). Ako su istraživanjem potpuno ispuštene kategorije stanovništva čiji odgovori mogu utjecati na predmet

istraživanja, ili ako su obuhvaćene u apsolutno premalom broju, ponderiranje prema sociodemografskim obilježjima u prvom slučaju nije moguće, a u drugom povećava pogrešku (Dever, 2008).

Osim nedovoljne pokrivenosti, reduciranosti osnovnog skupa pridonose i navike upotrebe interneta (uključujući sredine u kojima je pokrivenost gotovo potpuna; Lamza Posavec, 2021). Istraživanjima je nedostupan ili je teško dostupan udio korisnika koji internet upotrebljavaju rijetko, uglavnom za zabavu ili komunikaciju putem socijalnih mreža, ne posjećuju stranice na kojima se oglašavaju istraživanja ili nisu dovoljno informatički pismeni za aktivniju upotrebu medija.

Manjkav okvir uzorkovanja

Poteškoće s okvirima uzorkovanja (evidencijama korisničkih kontakata) onemogućuju ili otežavaju slučajan izbor. Osim oskudnosti dostupnih adresara, i njihova je upotreba ograničena zbog pravila privatnosti (*netiquette*) koja onemogućuju slanje poziva elektroničkom poštom bez prethodnog pristanka. Taj se nedostatak tradicionalno pokušava ublažiti što približnijim neprobabilističkim izborom, upotrebom probabilističkog kad je god moguće ili njihovom kombinacijom (Couper, 2000).

Neprobabilistički izbor

Neprobabilistički izbor najčešće se provodi na dva načina. Prvi se zasniva na javnom pozivu, koji može biti upućen kako putem stranica institucije ili agencije koja provodi istraživanje, tako i putem drugih medija i na druge načine, od društvenih mreža nadalje, uključujući nadopune uzoraka upotrebom dostupnih adresara elektroničke pošte (Peifer i Garret, 2014).

Uz jednostavnost, ekonomičnost i velik obuhvat, takva regrutacija uključuje mogućnost kontrole (identifikacije odazivanja i onemogućavanja višekratnog sudjelovanja) i bilježenje sociodemografskih obilježja. Zahvaljujući tome struktura realiziranog uzorka može se ponderiranjem (poststratifikacijom) ili postselekcijom uskladiti s populacijskom, no to uglavnom nije dovoljno kako bi se osigurala reprezentativnost. Odaziv na javno oglašavanje podrazumijeva motivacijsku pristranost (zainteresiranost za temu i praćenje medija u kojima je upućen poziv za sudjelovanje). Stoga se i rezultati mogu znatno razlikovati od onih koje bilježe tematski srodna istraživanja na strukturno ekvivalentnim, no probabilistički izabranim uzorcima koji obuhvaćaju i osobe koje se ne služe internetom (Lamza Posavec, 2021).

Drugi je način neprobabilističkog izbora masovna regrutacija putem oglašavanja na stranicama istraživačkih institucija ili agencija, odnosno angažiranje dobrovoljaca iz prethodnih istraživanja. Iz tako osiguranih velikih

baza podataka potom se stratifikacijom i konačnim slučajnim izborom mogu izabrati strukturno reprezentativni uzorci za različita istraživanja. No ni u tom se slučaju ne može govoriti o pravom probabilističkom izboru zato što inicijalna regrutacija nije bila takva, a ni ovaj način ne uključuje osobe koje se ne služe internetom.

Probabilistički izbor

Poput neprobabilističkog, i probabilistički se izbor može provesti na dva načina. Prvim se (jednostavnim ili sustavnim) slučajnim postupkom »presreću« posjetitelji web-stranica kojima se nudi sudjelovanje u istraživanju. Osim slabog odaziva (koji se ipak može precizno izračunati jer se bilježe i ukupni pristupi), takav način izbora uglavnom ne omogućuje generalizaciju nalaza na populacije široj od ukupnog broja posjetitelja.

Bolji je način prikupljanje adresa stanovanja ili telefonskih brojeva putem službenih izvora, nakon čega se drugim načinima (poštom, telefonom) kontaktira slučajno izabrane osobe kojima se nudi sudjelovanje u istraživanju putem interneta (Gfk, 2013). Slično se može učiniti i prikupljanjem kontakata tijekom prethodnih terenskih ili telefonskih anketiranja na reprezentativnim uzorcima. Akumulacija potencijalnih ispitanika u panel-uzorke omogućuje naknadni slučajni izbor adresa na koje će se slati pozivi za sudjelovanje u istraživanjima putem interneta, no i u ovom je slučaju generalizacija ograničena na njegove korisnike. Uz to treba uzeti u obzir i trostruku autoselekciju: onima koji su odbili sudjelovati u prethodnim reprezentativnim istraživanjima treba pribrojiti i korisnike interneta koji nisu bili voljni dati adrese u istim istraživanjima, a potom i odbijanja tijekom samog online istraživanja. Neka iskustva pokazuju da samo posljednja autoselekcija iznosi oko dvije trećine ukupno kontaktiranih (primjerice, DiSogra i sur., 2013). Relativne prednosti opisanog načina izbora svode se na to da je moguće preciznije izmjeriti naknadnu autoselekciju i provjeriti valjanost web anketa s prethodno prikupljenim nalazima reprezentativnih istraživanja. Nedostaci su neekonomičnost i pristranosti povezane s višekratnim ili longitudinalnim istraživanjima na panel-uzorcima općenito (lažno predstavljanje, osipanje, utjecaj naknada za višekratno sudjelovanje, lažne težnje za dosljednošću ili isto takve sklonosti promjenama mišljenja izazvane samom metodom, odnosno »pristranost panela« općenito; Göritz i sur., 2002; Lamza Posavec, 2021).

Zbog poteškoća s regrutacijom rezultati anketa provedenih na općoj populaciji putem interneta mogu se znatno razlikovati od rezultata prikupljenih klasičnim tehnikama na slučajno izabranim uzorcima (unatoč odgovarajućoj poststratifikaciji; Kennedy i sur., 2016), pri čemu su oni koji se odazovu probabilističkom izboru znatno sličniji apstinentima nego ako su izabrani nepro-

babilistički (Elliott i Valliant, 2017). Unatoč tom i ostalim nedostacima, neprobabilistički je izbor znatno češći jer je praktičniji, financijski i vremenski ekonomičniji i može obuhvatiti znatno veće uzorke uz veći odaziv koji se, u slučaju probabilističkog izbora, ionako sve više smanjuje (Beaumont, 2020; Kohut i sur., 2012).

Ovdje treba podsjetiti da su učinci autoselekcije prilikom istraživanja na probabilističkim uzorcima praktično ekvivalentni učincima neprobabilističkog izbora, samo što se u potonjem slučaju subjektivnost premješta s izbora uzorka na odlučivanje o sudjelovanju. U tom smislu autoselekcija ugrožava valjanost svih anketnih istraživanja bez obzira na tehniku provedbe, medij prikupljanja podataka i način izbora uzoraka (posebno ako je povezana s predmetom mjerenja). No statističkim se inovacijama nepovoljni učinci autoselekcije i neprobabilističkog izbora mogu istodobno smanjiti, odnosno neprobabilistički izbor aposteriorno približiti probabilističkom.

Spomenute inovacije gotovo redovito, osim sociodemografskih, uzimaju u obzir i tzv. »webografske« varijable (vjerojatnosti povezane s upotrebom interneta; Lee i Valliant, 2009). Najpopularnije su različite varijante inklinacijskog ponderiranja (*propensity adjustment*), preuzetog iz eksperimentalnih istraživanja (Rosenbaum i Rubin, 1983) kojim se izračunava vjerojatnosti podpopulacijskih pripadnosti (primjerice, korisnicima i nekorisnicima interneta). Pritom se poseže za pomoćnim (*auxiliary*) varijablama dostupnim iz vanjskih izvora (popisni podatci, nalazi reprezentativnih ili priručnih istraživanja na slučajnim uzorcima; Schonlau i Couper, 2017) kako bi se omogućila generalizacija dopunjavanjem inklinacijskog ponderiranja iterativnim (*raking*) ili simultanim (*matching*) postupcima (Castro-Martín i sur., 2020; Chen i sur., 2020; Ferri-García i Rueda, 2018; Ferri-García i sur., 2021; Valliant, 2020).

Zahvaljujući metodološkim unaprjeđenjima poput opisanih, web-anke te konkurentnom valjanošću u velikoj mjeri dostižu klasične tehnike. Neri-jetko ih i nadmašuju, pogotovo ako pojave koje se ispituju nisu povezane s dostupnošću i načinom upotrebe interneta ili ako se provode na adekvatno popisanim i dostupnim posebnim populacijama (vidjeti primjere u Galešić, 2005). U mnogim slučajevima više ne zaostaju ni pri istraživanjima općih populacija (Couper, 2017), a zbog ključnih su prednosti (ekonomičnosti i neograničenog teritorijalnog dosega) posebno ili jedino pogodne i za istraživanja iseljeništa.

Mogućnosti i ograničenja anketnih istraživanja iseljeništa

Anketna istraživanja iseljeničke populacije u cjelini i njezinih užih dijelova bila su teško provediva klasičnim tehnikama. Upotreba terenskih anketa

zbog velike teritorijalne disperzije gotovo da i nije bila moguća, no razvojem interneta to je ograničenje praktički ukinuto. Iako sva prethodno opisana obilježja anketnih istraživanja putem interneta vrijede i u ovom slučaju i ne treba ih ponavljati, mogu se naznačiti dvije moguće prednosti.

Prvo, može se očekivati da je pogreška obuhvata u inozemstvu manja s obzirom na to da se iseljavalo u razvijenije sredine u kojima je pokrivenost internetom veća (uz vjerojatno ranije stjecanje navika i vještina upotrebe, kako zbog tamošnjeg socijalnog i administrativnog okoliša, tako i zbog komunikacije s domovinom). Drugo, kad je riječ o mlađem iseljeništvu i novom iseljavanju, osim što je pokrivenost praktično potpuna, za pretpostaviti je i zadovoljavajuću informatičku pismenost. Budući da je aktivnije služenje internetom bilo nužno i za ispunjavanje raznovrsnih zahtjeva povezanih s iseljavanjem, vještine koje su za to bile potrebne vjerojatno su dovoljne i za ispunjavanje online upitnika.

Unatoč spomenutim prednostima, reprezentativna anketna istraživanja iseljeništva putem interneta suočavaju se s teško premostivom preprekom: nepoznatošću (nedovoljnom poznatošću) brojnosti i strukture užih dijelova populacije i populacije u cjelini koja je, uz to, većinom nedostupna zbog okvira uzorkovanja (popisa kontakata) zanemarivog opsega.

Nepoznatost populacije

Procjene opsega, strukture i disperzije iseljeničke populacije postoje, uključujući i službene evidencije kojima raspolažu nadležne institucije, organizacije i udruge u Hrvatskoj i inozemstvu. No te evidencije (kako populacije u cjelini, tako i njezinih užih dijelova) nisu potpune, ne zna se koliko su precizne, a u slučaju novog iseljavanja razmjerno brzo i zastarijevaju zbog njegeve dinamike.

Strogo gledano, reprezentativna istraživanja (opsegom i strukturom) nepoznatih populacija nisu moguća. Nepoznavanje objektivnih obilježja, odnosno populacijskih udjela statusnih kategorija relevantnih za predmet mjerenja, onemogućuje apriornu i aposteriornu stratifikaciju (planiranje uzoraka i naknadno usklađivanje sa strukturom populacije).

Podsjetimo, kvantitativna anketna istraživanja obuhvaćaju objektivna ili statusna (sociodemografska, profesionalna i slično) i subjektivna obilježja, odnosno trajnija ili manje trajna psihološka obilježja i (namjere) ponašanja. Objektivna obilježja, osim što služe za kontrolu realizacije uzorka, obično služe i kao odrednice ili prediktori (»nezavisne« varijable) subjektivnih (njihovog opsega, spektra ili hijerarhije).

Načelno, kad je riječ o strukturno nepoznatim ciljanim populacijama generalizacija bi bila opravdana samo ako su subjektivno homogene unatoč

objektivnoj varijabilnosti. No kako to redovito nije slučaj, preostaje uzorkom obuhvatiti svaku relevantnu statusnu kategoriju u asolutno dovoljnoj mjeri, uključujući i njezinu unutrašnju ili ukrižanu varijabilnost. Zadovoljenjem tih uvjeta nude se dvije mogućnosti.

Prvo, ako se može samo okvirno, ali utemeljeno pretpostaviti populacijski udio svake od spomenutih kategorija, moguća je i procjena populacijske zastupljenosti istraživanjem zabilježenih proporcija i prosjeka subjektivnih varijabli. No na istoj, okvirnoj razini preciznosti, kondicionalno i u obliku argumentirane spekulacije. Konkretnije, eventualno se može govoriti o manjoj ili većoj općoj zastupljenosti pojedinih pojava (ordinalno), ali ne i preciznije: za koliko ili koliko puta manjoj ili većoj (intervalno ili omjerno kvantificirano).

Drugo, ako su početni uvjeti zadovoljeni i ako se uz to pokaže da su kategorije subjektivno homogene unutar sebe, moguće je pretpostaviti da međusobne usporedbe proporcija i prosjeka vrijede i šire, i to uz manje rezervi nego u prvom slučaju.

Slično vrijedi i kad su u pitanju međusobna povezanost i utjecaji istraživanih pojava, što se obično provjerava korelacijskim i determinacijskim analizama. Ako se utvrdi da su međusobni odnosi subjektivnih varijabli unutar statusnih kategorija slični te ako su, općenito, linearni, opravdano je pretpostaviti da su i koeficijenti kojima se mjere razmjerno univerzalni (neovisno o stvarnoj ili populacijskoj zastupljenosti pojedinih kategorija).

Okvir uzorkovanja i način izbora uzorka

U slučajevima nepopisanih, nepoznatih ili ograničeno dostupnih populacija često se poseže za neslučajnim uzorkovanjem tzv. snježnom grudom (*snowball*). Ne računajući istraživanja društvenih mreža, primjena te tehnike u anketnim istraživanjima putem interneta razmjerno je nova, pa se i širina spektra metodoloških prednosti i nedostataka u tom području primjene još uvijek intenzivnije provjerava (Lu i sur., 2012; Schonlau i Couper, 2017).

Postupak se svodi na to da se inicijalno kontaktiraju poznati članovi populacije, koji potom sami regrutiraju ili upućuju na druge (tzv. *respondent-driven sampling*). Postupak pretvaranja grude u lavinu ponavlja se dok se ne postigne željena veličina uzorka linearnom, eksponencijalnom nediskriminativnom ili eksponencijalno diskriminativnom regrutacijom.

Budući da se takvim načinom izabiru osobe koje dijele neko zajedničko svojstvo i međusobno komuniciraju, i realizirani će uzorak biti homogeniji od populacije u cjelini. Stoga je jedan od glavnih ciljeva primjene *snowball* tehnike u kvantitativnim istraživanjima postizanje heterogenosti (Kirchherr i Charles, 2018). Iako se ona polako povećava svakim korakom poželj-

no ju je ubrzati u okviru regrutacije koja se, načelno, može provesti na dva načina.

Prvi je način da svaki izabrani član uzorka u svakom koraku dostavi kontakte poznatih osoba koje bi mogle sudjelovati u istraživanju, uz navođenje njihovih obilježja relevantnih za predmet istraživanja. Prikupljeni kontakti mogu poslužiti kao okvir iz kojeg se će se birati novi uzorci (koji, po potrebi, već inicijalno mogu biti heterogeniji budući da je poznata struktura mogućih kontakata). Drugi je način da inicijalno izabrani ispitanici sami šalju poveznice na upitnik poznatim osobama, pri čemu ih se može zamoliti da biraju osobe koje se od njih, prema zadanim obilježjima, što više razlikuju.

Svaki od spomenutih načina ima prednosti i nedostatke. Prednost prvog je kontrola, a nedostaci su dugotrajnost, uz potencijalno veću autoselekciju. Obratno, nedostatak drugog načina je izostanak kontrole, a prednosti su kraće vrijeme potrebno za istraživanje i potencijalno veći odaziv zbog povjerljivosti i veće skrupuloznosti prema odbijanju poznatih osoba. Povjerljivost pri tom ne potiče samo sudjelovanje, nego i prikupljanje novih kontakata (Noy, 2009; Shaghaghi i sur., 2011).

U svakom slučaju, cilj je obuhvatiti svu relevantnu varijabilnost populacije sve dok svaka potrebna statusna kategorija ne bude apsolutnim brojem dovoljno velika. Kako bi se to lakše ostvarilo, poželjno je da već i inicijalni kontakti (*seed*) budu što raznolikiji, pa ih je korisno prikupiti iz što više izvora (iako to, samo po sebi, bezuvjetno ne jamči i konačnu heterogenost; Etikan i sur., 2016; Kirchherr i Charles, 2018; Morgan, 2008; Waters, 2015).

Budući da je dijelom ipak evidentirano, u slučaju istraživanja iseljenišva inicijalno se uzorkovanje može provesti upotrebom javno (institucionalno) i privatno dostupnih kontakata i javnim pozivom različitim kanalima (uključujući društvene mreže), i potom nastaviti postupkom snježne grude. Nakon toga, prikupljanje podataka moguće je provesti u više valova kako bi se pristup u svakom novom valu mogao modificirati na temelju iskustava s prethodnim (Shafie, 2010; Robins, 2015).

Svaki je novi korak dobrodošao kako bi se povećala heterogenost i što približnije simulirao slučajni izbor. No iako ne mora izazvati posebno velike financijske troškove, uključivanje većeg broja koraka ili novih valova istraživanja nakon evaluacije prethodnih troši vrijeme, pa se postavlja pitanje koliko je koraka potrebno. U tom pogledu mogu okvirno poslužiti dva oslonca.

Prvi je stabilizacija. Ako objektivna heterogenost i dalje raste nakon što je postignuta u teritorijalno i statusno nužnoj mjeri, a subjektivnost stagnira (prosječni se odgovori na pitanja u upitniku više značajnije ne mijenjaju), za

očekivati je da daljnja regrutacija (iako će povećati preciznost) više neće bitnije promijeniti odgovore ili smanjiti pristranost. Osim toga, ako nisu ispuštene po istraživanje relevantne kategorije moglo bi se pretpostaviti da struktura uzorka proporcijama, barem grubo (i samo u nekim slučajevima), naznačuje i populacijsku. Stabilnost se može provjeriti i formalno, testiranjem hipoteza po koracima ili valovima (uobičajenim u tzv. sekvencijalnim analizama prema Waldovom inicijalnom prijedlogu; Dragan i Alexandru, 2013; Etikan i sur., 2015; Johnson, 2005; Sagalnik, 2006; Wald, 1947). Konačno, ako se premaši potrebna veličina uzorka, moguća je i postselekcija izostavljanjem najhomogenijih slučajeva, no samo prema *ad hoc* kriterijima s obzirom na to da univerzalnih preporuka nema.

Kao drugi oslonac mogu poslužiti rezultati reprezentativnih istraživanja u Hrvatskoj koja uključuju pitanja o iseljavanju. Za pretpostaviti je da će se struktura dijela domicilne populacije koji namjerava iseliti reproducirati u iseljeništvu, pa bi ju i u uzorcima na kojima se ono istražuje trebalo preslikati. Za razliku od prvog oslonca koji je univerzalniji, ovaj može poslužiti samo u slučaju novijeg iseljavanja. No i taj je kriterij razmjerno grub. Osim autoselekcije, reprezentativnost zabilježene strukture smanjuju i razlike između iskazanih namjera i ponašanja: oni koji su iskazali namjeru iseljavanja ne moraju to redovito i učiniti, i obratno, dio onih koji to nisu namjeravali na kraju ipak iseli.

Sagledano u cjelini, zbog svih navedenih ograničenja istraživanja iseljenništva mogu biti samo orijentacijska, uz iznimku onih koja se zasnivaju na uzorcima specifičnih, dovoljno evidentiranih i dostupnih ciljanih populacija.

Zaključak

Osim što nedovoljno poznata struktura iseljeničke populacije u cjelini ili uže ciljanih potpopulacija onemogućuju precizniju generalizaciju istraživačkih nalaza, ni nedostaci neprobabilističkog izbora uzoraka ne mogu se bitno smanjiti zbog oskudnog izbora i kvalitete pomoćnih varijabli iz vanjskih izvora. Unatoč tome, ako su metodološki rigorozno provedena i interpretativno korektna (uzimaju u obzir ograničenja), istraživanja iseljenništva mogu pokazivati i u reprezentativnom smjeru. Interpretativna korektnost nužna je i prilikom šire diseminacije nalaza budući da brožčani pokazatelji često poprimaju mitološki status u javnosti (tim više što su precizniji), koja se uglavnom ne upušta u propitivanje njihovog podrijetla i valjanosti.

Svako istraživanje koje zadovoljava spomenute kriterije može, uz nalaze koje nudi, biti i od dodatne pomoći: osim moguće akumulacije u panele, ako se u upitnike uključe i pitanja o široj situaciji u iseljeništvu, za očekivati je da se mogu prikupiti korisni podatci kojima inače ne raspolažu službeni i

druge vrste izvora. Time i populacija postaje dostupnija i poznatija, što može pridonijeti boljem planiranju i unaprjeđenju opće kvalitete naknadnih istraživanja.

Literatura

- Andrews, D., Nonnecke, B., & Preece, J. (2003). Electronic survey methodology: A case study in reaching hard-to-involve Internet users. *International Journal of Human-Computer Interaction*, 16 (2), 185-210.
- Baker, R., Blumberg, S. J., Brick, J. i sur. (2010). *Research synthesis. AAPOR report on online panels*. <https://pprg.Stanford.edu>
- Beaumont, J. F. (2020). Are probability surveys bound to disappear for the production of official statistics? *Survey Methodology*, 46 (1): 1-28.
- Blumenberg, C., Menezes, A. M. B., Gonçalves, H, i sur. (2019). The role of questionnaire length and reminders frequency on response rates to a web-based epidemiologic study: a randomised trial. *International Journal of Social Research Methodology*, 22: 625-635.
- Castro-Martín, L., Rueda, M., Ferri-García, R. (2020). Inference from non-probability surveys with statistical matching and propensity score adjustment using modern prediction techniques. *Mathematics*, 8 (6):879. <https://doi.org/10.3390/math8060879>
- Chang, L., Krosnick, J. A. (2010). Comparing oral interviewing with self-administered computerized questionnaires. An experiment. *Public Opinion Quarterly*, 74 (1): 154-167.
- Chen, Y., Li, P., Wu, C. (2020). Doubly robust inference with nonprobability survey samples. *Journal of American Statistical Association*, 115 (532):2011—2021. <https://doi.org/10.1080/01621459.2019.1677241>
- Conrad, F. G. (2002.). *Interactive aspects of web surveys: Lack of use, ease of use and user models*. International Workshop on Websurveys, Mannheim, Germany.
- Couper, M. P. (2017). New developments in survey data collection. *Annual Reviews in Sociology*, 43: 121-145.
- Couper, M. P. (2000). Web surveys: A review of issues and approaches. *Public Opinion Quarterly*, 64 (4): 464-494.
- Couper, M. P., Bosnjak, M. (2010). Internet Surveys. U P. V. Marsden, J. D. Wright (ur.), *Handbook of Survey Research* (str. 527-550). Somerville: Emerald Group Publishing.
- De Leeuw, E. (2005). To mix or not to mix data collection methods in surveys. *Journal of Official Statistics*, 21 (2): 233-255.
- Dever, J. A., Rafferty, A., Valliant, R. (2008). Internet surveys: Can statistical adjustments eliminate coverage bias? *Survey Research Methods*, 2 (2): 47-60.
- Dillman, D. A. (2007). *Mail and internet surveys: The Tailored Design Method 2007 update with new internet, visual, and mixed-mode guide*. Hoboken: John Wiley & Sons, Inc.

- Dillman, D. A., Smyth, J. D., Christian, L. M. (2009). *Internet, Mail, and Mixed-Mode Surveys: The Tailored Design Method*. Hoboken: Wiley & Sons.
- DiSogra, C., Cobb, C., Chan, E., Dennis, J. (2011). *Calibrating non-probability Internet samples with probability samples using early adopter characteristics*. Paolo Alto: Knowledge Networks.
- Dragan, I. M., Alexandru, I. M. (2013). Snowball Sampling Completion. *Journal of Studies in Social Sciences*, 5 (2): 160-177.
- Elliott, M. R., Valliant, R. (2017). Inference for nonprobability samples. *Statistical Science*, 32 (2):249-264.
- Etikan, I., Alkassim, R., Abubakar, S. (2016). Comparision of snowball sampling and sequential sampling technique. *Biometrics & Biostatistics International Journal*, 3 (1): 6-7.
- Ferri-García, R., Beaumont, J. F., Bosa, K., Charlebois, J., Chu, K., (2021). Weight smoothing for nonprobability surveys. *TEST*, 10.1007/s11749-021-00795-7.
- Ferri-García, R., Rueda, M. (2018). Efficiency of propensity score adjustment and calibration on the estimation from non-probabilistic online surveys. *SORT*, 42 (2): 159-182.
- Fricker, R. D., Schonlau, M. (2002). Advantages and disadvantages of Internet research surveys: Evidence from the literature. *Field Methods*, 14 (4): 347-367.
- Fripiat, D., Marquis, N. (2010). Web surveys in the social sciences: An overview. *Population*, 65 (2): 285-311.
- Galešić, M. (2005). Anketna istraživanja putem interneta: mogući Izvori pogrešaka. *Društvena istraživanja*, 14 (1-2): 297-320.
- Ganassali, S. (2008). The influence of the design of web survey questionnaires on the quality of responses. *Survey Research Methods*, 2 (1): 21-32.
- GfK (2013). *KnowledgePanel. Design summary*. <https://www.gfk.com>
- Görizt, A. S., Reinhold, N., Batinic, B. (2002.). Online Panels. U B. Batinic, U. Reips i M. Bosnjak (ur.), *Online Social Sciences* (str. 27-47). Seattle: Hogrefe & Huber.
- Gräf, L. (2002.), Assessing Internet questionnaires: The online pretest lab. U: B. Batinic, U. Reips i M. Bosnjak (ur.), *Online Social Sciences* (str. 49-68). Seattle: Hogrefe & Huber.
- Groves, R. M. (1989.). *Survey errors and survey costs*. New York: Wiley.
- Groves, R. M., Presser, S., Dipko, S. (2004.). The role of topic interest in survey participation decisions. *Public Opinion Quarterly*, 68 (1): 2-31.
- Heerwegh D., Loosveldt, G. (2008). Face-to-face versus web surveying in a high-internet-coverage population differences in response quality. *Public Opinion Quarterly*, 72 (5): 836-846.
- Hoogendoorn, A. W., Daalmans, J. (2009). Nonersponse in the recruitment of an internet panel based on probability sampling. *Survey Research Methods*, 3 (2): 59-72.
- Johnson, T. P. (2005). *Snowball Sampling. Encyclopedia of Biostatistics*. New York: Wiley.

- Kirchherr, J., Charles, K. (2018). Enhancing the sample diversity of snowball samples: Recommendations from a research project on anti-dam movements in Southeast Asia. *Plos One*, 13 (8): e0201710. <https://doi.org/10.1371/journal.pone.0201710>
- Kohut, A., Keeter, S., Doherty, C., Dimock, M., Christian, L. (2012). *Assessing the representativeness of public opinion surveys*. Washington: Pew Research Center.
- Konstan, J. A., Rosser, B. R. S., Ross, M. W., Stanton, J., Edwards, W. M. (2005). The story of subject naught: A cautionary but optimistic tale of Internet survey research. *Journal of Computer-Mediated Communication*, 10 (2): article 11.
- Kennedy, C., Mercer, A., Keeter, S., Hatley, N., McGeeney, K., Giménez, A. (2016). Evaluating online nonprobability surveys. Washington, DC: Pew Research Center.
- Kreuter, F., Presser, S., Tourangeau, R. (2008). Social desirability bias in CATI, IVR and web surveys. The effects of mode and question sensitivity. *Public Opinion Quarterly*, 72 (5): 847-865.
- Lamza Posavec, V. (2015). *Mjerenje javnosti. Metodologijski i analitički pristupi istraživanju javnoga mnijenja*. Zagreb: Institut društvenih znanosti Ivo Pilar.
- Lamza Posavec, V. (2021). *Metodologija društvenih istraživanja. Temeljni uvidi*. Zagreb: Institut društvenih znanosti Ivo Pilar.
- Lee, S., Valliant, R. (2009). Estimation for volunteer panel web surveys using propensity score adjustment and calibration adjustment. *Sociological Methods & Research*, 37 (3): 319-343.
- Llieva, J., Baron, S., & Healey, N. M. (2002). Online surveys in marketing research: Pros and cons. *International Journal of Market Research*, 44 (3), 361-367.
- Lozar Manfreda, K., Bosnjak, M., Brezalek, J., Haas, L., Vehovar, V. (2008). Web surveys versus other survey modes: A meta analysis comparing response rates. *International Journal of Market Research*, 50 (1): 79-104.
- Lu, X., Bengtsson, L., Britton, T., Camitz, M., Kim, B. J., Thorson, A., Liljeros, F. (2012). The sensitivity of respondent-driven sampling. *Journal of Royal Statistical Society, Serie A* (175): 191-216.
- Malhotra, N. (2008). Completion time and response order effects in web surveys. *Public Opinion Quarterly*, 72 (5): 914-934.
- Morgan, D. L. (2008). Snowball sampling. *The SAGE Encyclopedia of Qualitative Research Methods*. Thousand Oaks: SAGE Publications.
- Noy, C (2009). Sampling knowledge: the hermeneutics of snowball sampling in qualitative research. *International Journal of Social Sciences Research Methodology*, 11 (4): 327-344.
- Peifer, J., Garret, K. (2014). *Best practices for working with opt-in online panels*. <https://www.comm.ohio-state.edu>
- Pew Research Center (2022). *Key facts about the quality of the 2020 census*. <https://pewresearch.org>
- Prasad Nayak, M. S. D., Narayan, K. A. (2019). Strengths and weaknesses of online surveys. *IOSR Journal of Humanities and Social Sciences*, 24 (5): 31-38.

- Redline, C. (2013). Clarifying categorical concepts in a web survey. *Public Opinion Quarterly*, 77 (Special Issue): 89-105.
- Reips, U. D. (2000). The Web experiment method: Advantages, disadvantages, and solutions. U M. H. Birnbaum (ur.), *Psychological experiments on the Internet* (str. 89-114). San Diego: Academic Press.
- Reips, U. D. (2002). Standards for Internet-based experimenting. *Experimental Psychology*, (49): 4, 243-256.
- Robins, G. (2015). *Doing social network research: Network-based research design for social scientists*. London: SAGE Publications.
- Rosenbaum, P. R., Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70 (1): 41-55.
- Salganik, M. (2006). Variance estimation, design and sample size calculations for respondent-driven sampling. *Journal of Urban Health*, 83 (Suppl. 1): 98-112.
- Schaefer, D. R., Dillman, D. A. (1998). Development of standard e-mail methodology: Results of an experiment. *Public Opinion Quarterly*, 62 (3): 378-397.
- Schonlau, M., Couper, M. P. (2017). Options for conducting web surveys. *Statistical Science*, 32 (2): 279-292.
- Shafie, T. (2010). Design-based estimators for snowball sampling. SSRN, <https://ssrn.com/abstract=2471006>
- Shaghghi, A, Bhopal, R. S, Sheikh, A. (2011). Approaches to recruiting 'hard-to-reach' populations into research: A review of the literature. *Health Promotion Perspectives*, 1 (2): 86-94.
- Shih, T. H. (2008). Comparing response rates from web and mail surveys: a meta-analysis. *Field Methods*, 20: 249-271.
- Toepoel, V., Corrie, V., Das, M., Van Soest, A. (2009). Design of web questionnaires: An information-processing perspective for the effect of response categories. *Sociological Methods Research*, 37 (3): 371-392.
- Tourangeau, R., Conrad, F. G., Couper, M. P. (2013). *The Science of Web Surveys*. New York: Oxford University Press.
- Valliant, R. (2020). Comparing alternatives for estimation from nonprobability samples. *Journal of Survey Statistics and Methodology*, 8 (2): 231-263.
- Van Meter, K. M. (1990). Methodological and design Issues techniques for assessing the representatives of snowball samples. *NIDA Research Monographs*, 98: 31-44.
- Vehovar, V., Batagelj, Z., Lozar Manfreda, K. (2000), *Participation in solicited web surveys: Who comes farthest?* Fifth International Conference on Social Science Methodology, Cologne.
- Yaeger, D. S., Krosnick, J. A., Chang, L. C. i sur. (2011). Comparing the accuracy of RDD telephone surveys and internet surveys conducted with probability and non-probability samples. *Public Opinion Quarterly*, 75 (4): 709-747.
- Wald, A. (1947). *Sequential Analysis*. New York: Wiley.
- Waters, J. (2015). Snowball sampling: a cautionary tale involving a study of older drug users. *International Journal of Social Research Methodology*, 18 (4):367-380.

Stanko Rihtar:

Methodological Aspects of Diaspora Research Through Web-Surveys

Online survey research owes its popularity primarily to cost-effectiveness and unbridled territorial coverage, including the possibility of reaching otherwise inaccessible (or difficult to reach) populations. Additionally, there are also certain important technical advantages, which can improve not only the quality of the survey procedure, but also the subsequent handling of data and findings. Due to the increase in coverage, the adopted Internet-use habits, the increase in computer literacy and the accumulated methodological experiences, the representative ambitions of such research are becoming justified: In terms of validity and metric qualities, such research has in many cases reached (and in some cases even surpassed) classical techniques. Since they are not hindered by a large territorial dispersion, as other methods may be, online surveys are particularly suitable (in many ways the only suitable method) for diaspora research. However, two main obstacles need to be faced. The first is insufficient knowledge of the population (its size and structure) as a whole and of the majority of the narrowly-targeted subpopulations. The second is the lacking sampling framework. Both of these obstacles make it difficult or even impossible to choose samples that would justify a wider generalization of the findings.

The methodological features, advantages and disadvantages of online survey research are examined in the first part of this paper from the perspective of total survey error, which, in addition to the psychometric or measurement error, also includes self-selection, coverage and sample errors. In the second part of this paper, the main possibilities and scope of diaspora research are presented, especially taking into account the indicated limitations that hinder their representative ambitions.

Key words: *survey, internet, diaspora, survey error, representativeness*