

Nekatere dileme spletnega zbiranja podatkov v luči pedagoških raziskav

Znanstveni članek

UDK 303.6:004.77

KLJUČNE BESEDE: spletno zbiranje podatkov, metodologija, vzorec, odzivnost, etičnost, vprašalnik

POVZETEK – Prvo spletno zbiranje podatkov v raziskovalne namene je bilo izvedeno v devetdesetih letih prejšnjega stoletja. Z vse večjo dostopnostjo do interneta pa danes raziskovalcem nudi enostaven ter časovno in cenovno privlačen način zbiranja podatkov. S spletnim anketiranjem kot zelo kompleksnim pojavom se ukvarjajo predvsem družboslovni informatiki, ki izpostavljajo njegove prednosti ter opozarjajo na morebitne slabosti. Spletno anketiranje pa (predvsem zaradi svojih prednosti) vse pogostejše uporabljajo tudi v raziskavah na področju vzgoje in izobraževanja. Zato v prispevku nekoliko podrobneje opozarjamo predvsem na metodološke dileme (vprašanje vzorčenja, vprašanje odzivnosti ter vprašanje etičnosti) spletnega zbiranja podatkov, ki jih skušamo s pomočjo primera raziskave s pedagoškega področja kritično presoditi tudi za področje vzgoje in izobraževanja.

Scientific paper

UDC 303.6:004.77

KEYWORDS: Web-based survey, methodology, sample, responsiveness, ethics, questionnaire

ABSTRACT – The first web-based survey was conducted in the 1990s. Today, with the increasing accessibility of the internet, a web-based survey presents an easy, time and cost-saving way of collecting data for the researchers. Web-based surveys are most commonly the subject of social informatics research and they are considered to be a very complex phenomenon having its advantages, as well as disadvantages. But, mainly due to their advantages, web-based surveys have been increasingly used in the field of educational research as well. This is why we also pay attention to some, mainly methodological issues (sampling, responsiveness, and ethics), which are further considered in the context of educational research.

1. Uvod

Zbiranje podatkov je ena izmed ključnih faz (Sagadin, 1993, str. 21; Cencič, 2009, str. 44) raziskovalnega procesa, saj so prav zbrani podatki podlaga za oblikovanje zaključkov, sklepov raziskave. Za zbiranje podatkov lahko uporabimo različne tehnike, ki jih razvrščamo glede na vrste podatkov, ki jih želimo zbrati (tj. kvalitativne ali kvantitativne), glede na strukturiranost (strukturirane ali nestrukturirane) ter glede na dostop do podatkov (posredne ali neposredne) (Cencič, 2009, str. 42), za kvantitativno raziskovanje pa velja, da podatke zbirajo pretežno z vprašalnikom (kot instrumentom zbiranja podatkov). Pri sestavljanju vprašalnika moramo biti poleg vsebinskih in zaradi vsebinskih vprašanj pozorni še na naslednja vprašanja: kje bomo zbirali podatke, kdaj jih bomo zbirali, kdo jih bo zbiral in kdo izpolnjeval. Vsi ti dejavniki namreč vplivajo na veljavnost podatkov (Cencič, 2009, str. 44). Ta vprašanja se ne-

dvomno navezujejo tudi na načine zbiranja podatkov (Sagadin, 1993, str. 159–160), ki so: osebni (z “anketarjem” ali brez njega), telefonski, preko pošte ali preko interneta (Cohen, Manion in Morrison 2007, str. 344; Cencič, 2009, str. 58) – prav slednji, ki je v zadnjem desetletju postal zelo priljubljen in razširjen način zbiranja podatkov, je središče zanimanja pričujočega prispevka. Kot drugi načini zbiranja podatkov ima namreč tudi spletno zbiranje podatkov mnoge prednosti, pa tudi slabosti. V prispevku pa ne želimo le strniti različne prednosti, slabosti ter značilnosti spletnega anketiranja, raziskovalce želimo tudi informirati o dilemah in omejitvah, na katere naj bi bili pozorni in naj bi jih upoštevali, ko načrtujejo in izvajajo raziskave ter interpretirajo rezultate raziskav, ki temeljijo na spletnem zbiranju podatkov.

2. Spletno zbiranje podatkov v raziskovalne namene

2.1. Razvoj spletnega zbiranja podatkov

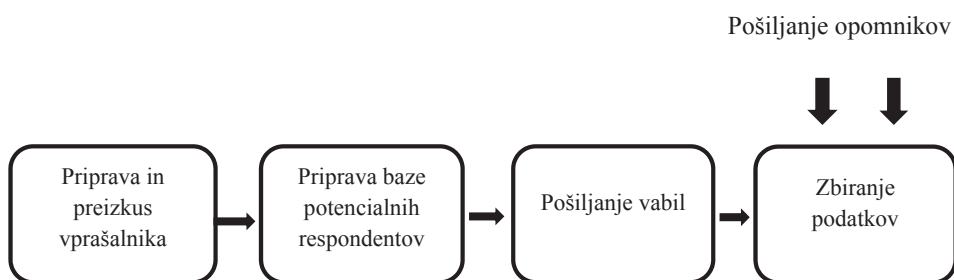
Začetki spletnega zbiranja podatkov v raziskovalne namene segajo v osemdeseta leta prejšnjega stoletja, ko je t.i. internetno zbiranje podatkov potekalo preko elektronske pošte. Prvo spletno anketiranje, kot ga poznamo danes, je bilo izvedeno leta 1994 (Pitkow in Colleen, 1997), njegova raba pa se je nato s hkratnim razvojem ustreznih podpornih tehnologij, programov in orodij ter posledično z možnostjo oblikovanja vprašalnikov na različnih spletnih straneh le še povečevala. Do danes je ta način prešel v enega izmed glavnih načinov pridobivanja podatkov (Lozar Manfreda, Vehovar in Batagelj, 2000, str. 1035; Cohen, Manion in Morrison, 2007, str. 230) oziroma kot ugotavljajo Silber, Lischewski in Leibold (2013, str. 121), je spletno anketiranje skorajda spodrinilo tradicionalne načine anketiranja preko telefona, osebno ter preko pošte. To stanje lahko pripišemo predvsem prednostim spletnega zbiranja podatkov pa tudi vedno večjemu deležu uporabnikov svetovnega spleta (Roster, Rogres, Albaum in Klein, 2004; Silber, Lischewski in Leibold, 2013 str. 121). Spletno zbiranje podatkov poteka v več zaporednih fazah, ki jih v nadaljevanju predstavljamo.

2.2. Proces spletnega zbiranja podatkov

Izhodišče za spletno zbiranje podatkov je vprašalnik, pri sestavi katerega je treba upoštevati vsa načela in priporočila (npr. Sagadin, 1993, str. 145–165; Cencič, 2009, str. 85–89), ki veljajo tudi za vprašalnike v papirnati obliki. Vprašalnik je zelo priporočljivo tudi preizkusiti in na ta način zagotoviti, da ustreza bistvenim merskim značilnostim: veljavnosti, zanesljivosti in objektivnosti. Pri spletnem vprašalniku (Fox, Murray in Warm, 2003, str. 71) pa je treba premisliti še o vizualnih, avditivnih oziroma interaktivnih vidikih vprašalnika. Nadalje je treba razmisliti o samih potencialnih respondentih ter načinih dostopanja do njih (prav tam, str. 72). Pri tem lahko preteh-

tamo o možnostih dostopa do že obstoječih baz, oblikovanja svoje baze, glede na cilje raziskave pa tudi o morebitnih že obstoječih spletnih skupinah. Ko je baza oblikovana, potencialnim respondentom pošljemo povabilo k sodelovanju, pri tem pa tudi že načrtujemo opomnike – praviloma naj ne bi poslali več kot tri opomnike. Različne analize (Lozar Manfreda, Vehovar in Batagelj, 2000, str. 1039) namreč kažejo, da je največji odziv prav na dan ali nekaj dni po tem, ko respondenti prejmejo opomnike.

Slika 1: Proces spletnega zbiranja podatkov



Načela ravnanja pri spletnem zbiranju podatkov so torej dokaj enostavna, ta način zbiranja podatkov pa odlikujejo tudi druge prednosti, ki jih opredeljujemo v nadaljevanju.

2.3. Prednosti spletnega zbiranja podatkov

Raziskovalci (Lozar Manfreda, 2001, str. 54; Wright, 2005; Cohen, Manion in Morrison, 2007) opredeljujejo predvsem naslednje prednosti spletnega anketiranja:

Nizki stroški

V primerjavi z drugimi načini zbiranja podatkov (kot sta npr. osebno, telefonsko) spletno anketiranje odlikujejo nizki stroški. Pri tem načinu zbiranja podatkov odpadejo stroški papirja, tiskanja, pošiljanja (kuverte, poštnina), prevoza do anketirancev ipd.

Časovni prihranek

Kot navaja Lozar Manfreda (2001), so podatki zbrani hitro, lahko jih tudi takoj oziroma sproti pregledujemo. Prav tako se izgubijo časovne omejitve: podatki se zbirajo 24 ur na dan, 7 dni v tednu, kar je v primerjavi s tradicionalnimi načini anketiranja velika prednost. Vendar pa, kot opozarjajo Fox, Murray in Warm (2003), je na drugi strani časovno zamudna faza pred samim zbiranjem podatkov, ko je treba pridobiti oziroma oblikovati baze e-naslovov, ki bodo kasneje uporabljeni z name-

nom pošiljanja vabil k izpolnjevanju vprašalnika. Nedvomno pa raziskovalci prihranijo veliko časa, ker pridobljenih podatkov ni treba nato ročno vnašati v programe za statistično obdelavo podatkov (Lozar Manfreda, 2001; Cohen, Manion in Morrison, 2007), saj prenos iz spletne baze v program navadno opravijo zelo hitro.

Geografska neomejenost

S spletnim anketiranjem se izbrišejo tudi geografske omejitve, saj je vprašalnik na voljo povsod po svetu, kjer ima posameznik dostop do interneta. To bistveno olajša npr. mednarodne študije (Lozar Manfreda, 2001; Wright, 2005).

Prednosti za respondente (ki so posledično lahko tudi prednosti za raziskovalce)

Zaradi možnosti izkoriščanja raznih vizualnih, zvočnih in drugih multimedijskih učinkov so spletni vprašalniki bolj zanimivi (Lozar Manfreda, 2001, str. 64). Cohen, Manion in Morrison (2007, str. 230) kot prednosti opredeljujejo tudi to, da respondenti lahko izpolnjujejo vprašalnik v njim ustreznem okolju (npr. doma) in ob njim ustreznem času.

Na tem mestu je treba poudariti, da se lahko prednosti spletnega zbiranja podatkov realizirajo le v primeru, če je vprašalnik, ki je osnova takega zbiranja, tako vsebinsko, jezikovno kot tudi metodološko ustrezno in korektno oblikovan, torej če pri sestavi in oblikovanju sledijo že uveljavljenim in preizkušenim načelom (npr. Sagadin, 1993; Cencič, 2009).

2.4. Slabosti/dileme spletnega zbiranja podatkov

Ob prednostih pa se pojavljajo tudi nekatere, predvsem metodološke dileme, pa tudi neenotni pogledi na posamezna vprašanja, vezana na spletno zbiranje podatkov.

Vprašanje vzorca/vzorčenja

Nekateri (Frippiat, Marquis in Wiles-Poriter, 2012) so mnenja, da je pri spletnem zbiranju podatkov zelo težko zagotoviti slučajnostni vzorec. Pojavlja se namreč (kot to imenujejo Lozar Manfreda, Vehovar in Batagelj, 2000, str. 1038) napaka vzorčenja. Na vabila k izpolnjevanju se praviloma ne odzovejo vsi povabljeni, pač pa samo zainteresirani (prav tam; Fox, Murray in Warm, 2003; Wright, 2005), kar lahko pomeni, da se značilnosti vzorca razlikujejo od značilnosti ciljne populacije. Dejstvo je tudi, da so iz vzorca takoj izvzete osebe, ki nimajo dostopa do interneta, slabše informacijsko pismeni ter starejši. Raziskave (Kwak in Radler, 2002; Roster, Rogers, Albaum in Klein, 2004, str. 371) namreč kažejo, da internet pogosteje uporabljajo mlajši, višje izobraženi, moški, znatno manj pa se interneta poslužujejo socialno šibkejši. Problem dostopnosti do interneta je v povezavi z vzorcem sicer zmanjšan ali odpravljen v populacijah, kjer večina članov uporablja internet (Lozar Manfreda, Batagelj in Vehovar, 2002).

Nadalje se lahko pojavijo tudi težave pri zbiranju podatkov pri respondentih, ki sicer imajo dostop do interneta. Dejstvo je, da ljudje pogosteje menjujemo elektronske poštne naslove kot pa, na primer, domači naslov ali telefonsko številko (Lozar Manfreda, 2001), kar lahko poveča število nedostavljenih vabil k izpolnjevanju. Težava so lahko tudi nepopolni sezname naslovov elektronskih pošt (angl. mailing list) ali neaktivni elektronski poštni naslovi (Wright, 2005). Največ zadržkov glede spletnega anketiranja je povezanih prav z neslučajnostnimi vzorci in s tem povezano neupravičenostjo posploševanja, vendar pa, kot opozarjajo Cohen, Manion in Morrison (2007, str. 237), je to pogosta težava tudi drugih raziskav, ne le raziskav, ki temeljijo na spletnem zbiranju podatkov. Je pa nesporno, da se pri vzorcu vedno sprašujemo, ali se prostovoljci, ki so se vključili v raziskavo, in njihova mnenja ter ocene razlikujejo od tistih, ki se v raziskavo niso vključili (prav tam).

Vprašanje odzivnosti

Raziskava (Vehovar, Lozar Manfreda in Koren, 2008), v kateri so primerjali stopnjo odzivnosti v 24 različnih raziskavah, kaže, da je pri spletnem zbiranju podatkov stopnja odzivnosti v povprečju 11 odstotkov nižja v primerjavi z odzivnostjo pri zbiranju podatkov na druge načine. Nižja pa ni le odzivnost (Cohen, Manion in Morrison, 2007), pač pa tudi popolna izpolnitev vprašalnika, pri čemer nekateri (Manfreda Lozar, Vehovar, Batagelj, 2002) ugotavljajo, da je dolžina vprašalnika pomemben dejavnik izpolnjevanja, drugi (Cohen, Manion in Morrison, 2007), pa trdijo, da dolžina vprašalnika ne vpliva na njegovo (ne)popolno izpolnitev. Ugotovili so (Silber, Lischewski in Leibold, 2013), da od začetka do konca vprašalnika izpolnjevanje opusti približno tretjina tistih, ki se odzovejo, pri čemer jih večina odneha po prvi tretjini vprašanj. Respondenti v procesu izpolnjevanja izgubijo interes in lažje opustijo izpolnjevanje spletnega vprašalnika, kot pa oddajo prazen papirnat vprašalnik. Veliko je tudi takih, ki preberejo uvodni nagovor in se sploh ne odločijo za izpolnjevanje (Cohen, Manion in Morrison, 2007).

Med dejavnike, ki vplivajo na stopnjo odgovorov (Vehovar, Lozar Manfreda in Koren 2008; Cohen, Manion in Morrison 2007), uvrščamo:

- značilnosti anketiranja (tj. sociodemografske značilnosti, izkušnje z raziskavami, zanimanje za raziskovalno temo in razpoloženje anketiranja);
- družbeno okolje (splošno javno mnenje in tradicija raziskav);
- tehnološko okolje (zmogljivost računalnika, dostopnost do interneta, alternativne naprave, IKT pismenost);
- zasnova raziskave (vabilo, opomniki, nagrade, dolžina vprašalnika, oblika vprašalnika).

Upoštevati pa je treba tudi argument, da ni toliko pomembna sama odzivnost, pač pa je pomembnejša reprezentativnost tistih, ki se odzovejo ali, povedano drugače, manjši vzorci so lahko bolj reprezentativni kot večji (Cook, Heath in Thompson, 2000, str. 821).

Etične dileme

Na eni strani se zdi, da so v spletnem okolju družbene norme manj pomembne in je na ta način manj družbeno zaželenih odgovorov, na drugi strani pa se postavlja vprašanje, ali je na ta način v odgovorih več neiskrenosti, namišljenih, virtualnih identitet (Lozar Manfreda in Vehovar, 2002), kar lahko vpliva tudi na avtentičnost podatkov (Fox, Murray in Warm, 2003, str. 173). Dvom v avtentičnost se zdi upravičen tudi z vidika dejstva (Lozar Manfreda in Vehovar, 2002), da je zelo verjetno, da respondenti, medtem ko izpolnjujejo vprašalnik, počno še druge stvari, kot npr. pregledujejo e-pošto, berejo novice, iščejo informacije, uporabljajo različne klepetalnice in si izmenjujejo informacije na socialnih omrežjih, kar seveda vodi v netočno in nenatančno branje in odgovarjanje. Ob tem ni odveč opozoriti tudi, da elektronsko gradivo že v izhodišču beremo z manjšo mero koncentracije in hitreje ter bolj površno.

Nekateri raziskovalci (npr. Dillman, Smyth, Christian in Stern, 2003; Roster, Rogers, Albaum in Klein, 2004; Heerwegh in Loodsveldt, 2008, str. 7–22) so ugotavljali tudi tendence pri odgovarjanju na različne tipe vprašanj ter iskali razlike v odgovorih med spletnimi in papirnati vprašalniki.

Tako so za anketna vprašanja zaprtega tipa pri spletnem anketiranju z možnima odgovoroma da in ne (npr. Dillman, Smyth, Christian in Stern, 2003, str. 7) ugotovili povečan delež pritrdilnih odgovorov (11,3 odstotka več pritrdilnih), saj naj bi v osnovi neopredeljeni respondenti raje izbirali trdilno obliko. Ko pa imajo respondenti tudi možnost izbire odgovora "ne vem", prevladuje prav ta odgovor (Heerwegh in Loosveldt, 2008). Pri vprašanjih z več možnimi odgovori pa je značilno, da so najpogosteje izbrani tisti odgovori, ki so višje na seznamu (Dillman, Smyth, Christian in Stern, 2003). Pri spletnem zbiranju podatkov je bil prav tako zabeležen večji delež odgovorov na anketna vprašanja odprtega tipa, ki so bili tudi daljši (Dillman, Smyth, Christian in Stern, 2003, str. 8). Ugotovljeno je bilo tudi (Roster, Rogres, Albaum in Klein, 2004, str. 371; Heerwegh in Loosveldt 2008), da so odgovori bolj enotni, ize-načeni, med respondenti se ne pokažejo večje razlike.

Izpostavlja se tudi vprašanje zagotavljanja anonimnosti (Cohen, Manion in Morrison 2007). Dvom v anonimnost se pojavlja predvsem zaradi vabila preko elektronske pošte pa tudi zaradi možnosti sledenja "ip" številki.

3. Predstavitev primera spletnega zbiranja podatkov za raziskavo s področja vzgoje in izobraževanja

Na tem mestu z vidika spletnega anketiranja, predvsem pa z vidika izpostavljenih metodoloških dilem, predstavljamo in analiziramo primer raziskave, ki je temeljila na spletnem zbiranju podatkov, vključevala pa je vzgojitelje in pomočnike vzgojiteljev v javnih vrtcih v Sloveniji.

3.1. Vprašalnik

Vprašalnik, ki smo ga uporabili za spletno zbiranje podatkov, sestavljajo trije sklopi:

- sklop 5 anketnih vprašanj zaprtega tipa,
- sklop 45 petstopenjskih Likertovih lestvic stališč in
- sklop 45 petstopenjskih ocenjevalnih lestvic.

Vprašalnik ustreza kriteriju (konstruktne) veljavnosti, saj rezultati factorske analize kažejo, da prvi (izmed dobljenih 10) faktorji pojasnjuje 39,7 odstotka, kar je več od predpostavljene spodnje meje veljavnosti, to je 20 odstotkov (Čagran, 2004, str. 3). Zanesljivost vprašalnika potrjujeta tako koeficient zanesljivosti ($r_{tt} = 0,847$) kot tudi Cronbachov α koeficient ($\alpha = 0,938$). Objektivnost vprašalnika smo zagotavljali predvsem z izbranimi tehnikami zbiranja podatkov (anketna vprašanja zaprtega tipa, Likertova lestvica stališč, ocenjevalna lestvica), ki omejujejo subjektivno interpretacijo. Na podlagi zgoraj zapisanih dejstev lahko trdimo, da je šlo za metodološko ustrezen instrument.

3.2. Postopek zbiranja podatkov

Za raziskavo smo načrtovali uporabo večstopenjskega slučajnostnega vzorca. V ta namen smo na spletni strani Ministrstva za izobraževanje, kulturo in šport (<https://krkal.mss.edus.si/resgistriweb/SeznamVrtci.aspx>, 2012) pridobili seznam vrtcev po regijah, pri čemer je bilo v bazi 383 vrtcev, ki so bili razvrščeni v 12 regij. K izpolnjevanju vprašalnika smo povabili po tri vrtce (vsakega prvega mestnega, primestnega in vaškega v regiji) iz vsake regije, skupaj torej 36 vrtcev. Elektronski naslovi posameznih vrtcev so sicer objavljeni na strani Ministrstva, a se je po pregledu izkazalo, da so nekateri neveljavni, zato smo vabila poslali na elektronske naslove, ki smo jih pridobili na uradnih spletnih straneh izbranih vrtcev. Razmišljali smo tudi o neposrednih vabilih (vsakemu posameznemu respondentu), a na spletnih straneh vrtcev niso bili objavljeni posamični elektronski naslovi vzgojiteljev in pomočnikov, zato smo se, kot smo že zapisali, odločili, da vabila pošljemo na elektronske naslove vrtcev in ne posameznikov. V vabilu smo prejemnika pošte zaprosili, da poveza do vprašalnika in vabilo posredujejo vsem vzgojiteljem/vzgojiteljicam ter njihovim pomočnikom/pomočnicam. V vabilu smo se predstavili, predstavili smo tudi cilje raziskave, zapisali smo poveza do spletnega vprašalnika ter opredelili predviden čas zbiranja podatkov (dva meseca), se zahvalili za sodelovanje ter ponudili možnost seznanitve z rezultati. Po preteku prvih desetih dni zbiranja podatkov smo povabljenim poslali prvi opomnik, po preteku enega meseca pa še drugi opomnik (gl. tudi graf 2).

3.3. Odzivnost

Kot je razvidno iz tabele 1, je na nagovor kliknilo 732 potencialnih respondentov, od tega jih je le 75 odstotkov nadaljevalo s pregledom začetne (prve) strani vprašalnika, vprašalnik pa je dejansko začelo izpolnjevati 482 anketirancev. Le 44 odstotkov, (tj. 323), torej manj kot polovica vseh, ki so se na vabilo odzvali, pa je dejansko izpolnilo vprašalnik. Potrdila se je trditev, da je veliko takih, ki preberejo uvodni nagovor in se nato ne odločijo za izpolnjevanje (Cohen, Manion, Morrison in 2007). Same stopnje odzivnosti so pri spletnih vprašalnikih sicer zelo različne in odvisne od mnogih dejavnikov (Vehovar, Lozar Manfreda in Koren, 2008; Cohen, Manion in Morrison, 2007).

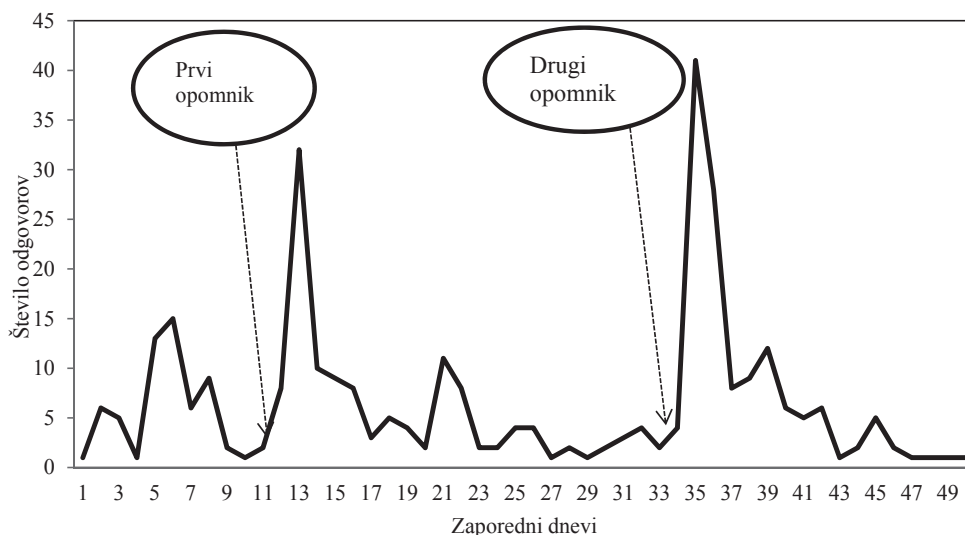
Tabela 1: Pregled osipa pri izpolnjevanju spletnega vprašalnika

<i>Stopnje</i>	<i>Frekvenca</i>	<i>Stopnja (%)</i>
Klik na nagovor	732	100
Klik na vprašalnik	550	75
Začel izpolnjevati	482	66
Delno izpolnjeno	445	61
Izpolnjeno	323	44

Zbiranje podatkov je potekalo 50 dni, pri čemer je z grafa 1 razvidno, da je bila najvišja dnevna frekvenca po prvem vabilu dosežena po enem tednu. Kot smo že omenjali, smo vabila pošiljali na naslove vrtcev in nato so nosilci teh elektronskih naslovov posredovali vabilo neposredno vzgojiteljem in pomočnikom vzgojiteljev. Deset dni po poslanem vabilu smo poslali prvi opomnik in učinek le-tega je viden tudi iz krivulje na grafu 1, saj se je nekaj dni po tem opomniku frekvenca odgovorov zvišala. Še bolj učinkovit je bil drugi opomnik, ki smo ga poslali mesec dni po vabilu, kar se kaže tudi v najvišji doseženi dnevni frekvenci v celotnem obdobju zbiranja podatkov – to je 42 odgovorov.

Povzemimo, tudi v primeru raziskave na pedagoškem področju se je – kot so izpostavljali že drugi (npr. Lozar Manfreda, 2001; Wright, 2005) – izkazalo, da se lahko kot problem spletnega zbiranja podatkov izkaže že sam dostop do potencialnih respondentov, kar se je v našem primeru potrdilo z neobjavljenimi oziroma tudi neveljavnimi e-poštnimi naslovi. Prav tako se je potrdil problem odzivnosti, ki je neposredno po opomnikih sicer narasla (prim. Lozar Manfreda, Vehovar in Batagelj, 2000, str. 1039), še bolj pa samega osipa, saj je manj kot polovica vseh, ki so prebrali uvod v vprašalnik, dejansko izpolnila vprašalnik v celoti.

Graf 1: Prikaz frekvence izpolnjevanja vprašalnika



Povzemimo, tudi v primeru raziskave na pedagoškem področju se je – kot so izpostavljali že drugi (npr. Lozar Manfreda, 2001; Wright, 2005) – izkazalo, da se lahko kot problem spletnega zbiranja podatkov izkaže že sam dostop do potencialnih respondentov, kar se je v našem primeru potrdilo z neobjavljenimi oziroma tudi neveljavnimi e-poštnimi naslovi. Prav tako se je potrdil problem odzivnosti, ki je neposredno po opomnikih sicer narasla (prim. Lozar Manfreda, Vehovar in Batagelj, 2000, str. 1039), še bolj pa samega osipa, saj je manj kot polovica vseh, ki so prebrali uvod v vprašalnik, dejansko izpolnila vprašalnik v celoti.

3.4. Vzorec

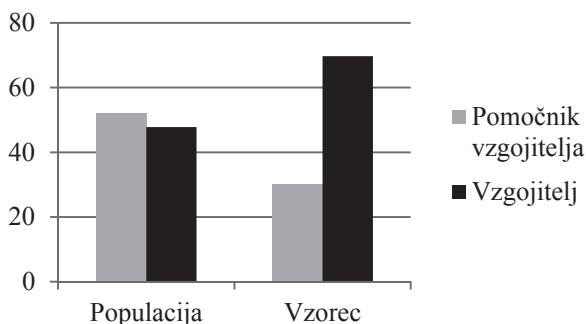
Kot smo že opredelili, je bila za raziskavo predvidena uporaba slučajnostnega vzorca, vendar pa je vprašanje (Lozar Manfreda, Vehovar in Batagelj, 2000; Frippiat, Marquis in Wiles–Poriter, 2012), ali je pri spletnem zbiranju sploh mogoče zagotoviti slučajnostni vzorec. Ko smo analizirali zbrane podatke, so se nam porajala številna vprašanja glede vključenih oziroma tudi nevključenih, kot npr.:

Ali so vabila dejansko dosegla vse vzgojitelje/pomočnike v izbranih vrtcih? Ali imajo vsi elektronsko pošto oziroma veljaven naslov? Kako pogosto pregledujejo svoj elektronski predal? Glede na prostovoljnost udeležbe, kateri povabljeni so izpolnili vprašalnik? Se mnenja vključenih razlikujejo od mnenj tistih, ki se za sodelovanje niso odločili? ipd.

Da bi dobili nekatere odgovore oziroma se jim vsaj približali, v nadaljevanju prikazujemo primerjavo nekaterih značilnosti med osnovno statistično množico in vzorcem.

Osnovno statistično množico so predstavljali vzgojitelji in pomočniki vzgojiteljev, zaposleni v javnih vrtcih v Sloveniji v šolskem letu 2011/2012. Po podatkih Statističnega urada RS je bilo v slovenskih javnih vrtcih skupno zaposlenih 9887 vzgojiteljev in pomočnikov <http://pxweb.stat.si/pxweb/Dialog/Saveshow.asp> (Zaposleni na delovnem mestu vzgojitelja po izobrazbi in spolu, 2012).

Graf 2: Primerjava populacije in vzorca po zastopanosti vzgojiteljev in pomočnikov vzgojiteljev (v odstotkih)

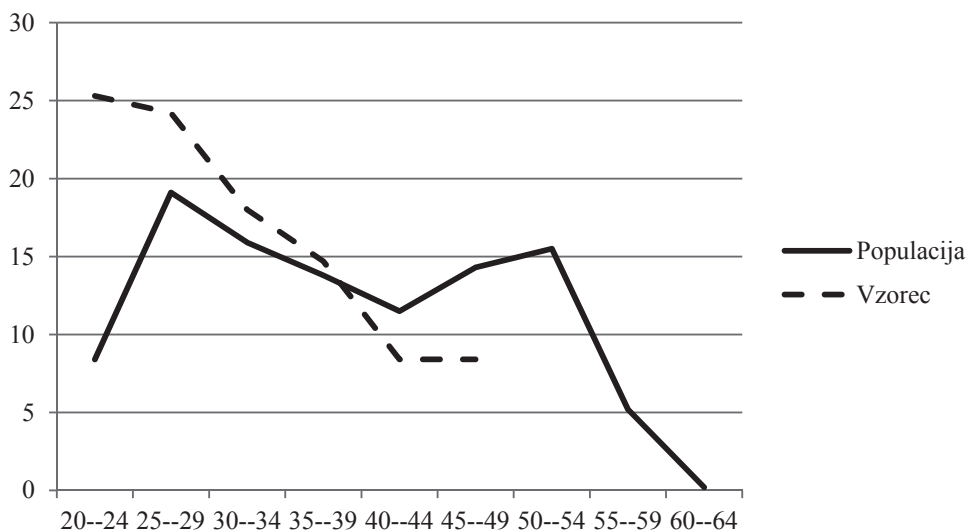


V osnovni statistični množici je bilo 4.725 (47,8%) vzgojiteljev in 5.162 (52,2%) pomočnikov vzgojiteljev (Zaposleni na delovnem mestu vzgojitelja po izobrazbi, 2012), v vzorcu pa je bila zastopanost nekoliko drugačna. Kot je razvidno z grafa 2, so v vzorcu prevladovali (69,7%) vzgojitelji, pomočniki vzgojiteljev so bili zastopani v manjši meri (30,03%). Razliko v strukturi populacije in vzorca glede na delovno mesto bi nemara lahko pripisali prav načinu zbiranja podatkov. Kot so namreč že ugotavljali (npr.: Kwak in Radler, 2002; Roster, Rogers, Albaum in Klein, 2004), internet manj pogosteje uporabljajo nižje izobraženi. Dejstvo pa je, da morajo pomočniki vzgojiteljev za opravljanje svojega poklica doseči nižjo (tj. peto) stopnjo izobrazbe kot vzgojitelji, ki se lahko na tem delovnem mestu zaposlijo ob predpostavki, da so zaključili visokošolsko strokovno izobraževanje. Sklepamo torej lahko, da se pomočniki vzgojiteljev v manjši meri uporabljajo internet, ob tem pa se postavlja vprašanje o njihovi informacijski pismenosti.

Kot je razvidno z grafa 3, je v osnovni statistični množici največ (19,1%) vzgojiteljev in pomočnikov, starih med 25 in 29 let. V vsaki naslednji starostni skupini je manj vzgojiteljev in vzgojiteljic, nato pa se poveča zastopanost tega kadra v starostih skupinah od 45 do 49 (14,3%) ter 50–54 (15,5%) let. Zelo malo (5,4%) je vzgojiteljev, ki so stari 55 in več let. V vzorcu so sicer najmočnejše (25,3%) zastopani vzgo-

jitelji in pomočniki, ki so stari med 20 in 24 leti, vendar pa jim takoj sledijo stari od 25 do 29 let (24,2%), ki so najbolj zastopani v populaciji. Nato pa s starostjo upada zastopanost respondentov, razberemo lahko tudi, da se v raziskavo ni vključil noben vzgojitelj ali pomočnik, ki je star 50 ali več let. Na podlagi enega primera sicer ne moremo posploševati, vendarle pa si lahko dovolimo zapisati, da analiza podatkov na grafu 3 potrjuje (npr. Kwak in Radler, 2002; Roster, Rogers, Albaum in Klein, 2004, str. 371), da se v spletno anketiranje vključujejo predvsem mlajši in da se na ta način v raziskave vključi le malo starejših respondentov.

Graf 3: Primerjava populacije in vzorca po zastopanosti glede na starost (v odstotkih)



Kot je razvidno z grafa 3, je v osnovni statistični množici največ (19,1%) vzgojiteljev in pomočnikov, starih med 25 in 29 let. V vsakem naslednji starostni skupini je manj vzgojiteljev in vzgojiteljic, nato pa se poveča zastopanost tega kadra v starostih skupinah od 45 do 49 (14,3%) ter 50–54 (15,5%) let. Zelo malo (5,4%) je vzgojiteljev, ki so stari 55 in več let. V vzorcu so sicer najmočnejše (25,3%) zastopani vzgojitelji in pomočniki, ki so stari med 20 in 24 leti, vendar pa jim takoj sledijo stari od 25 do 29 let (24,2%), ki so najbolj zastopani v populaciji. Nato pa s starostjo upada zastopanost respondentov, razberemo lahko tudi, da se v raziskavo ni vključil noben vzgojitelj ali pomočnik, ki je star 50 ali več let. Na podlagi enega primera sicer ne moremo posploševati, vendarle pa si lahko dovolimo zapisati, da analiza podatkov na grafu 3 potrjuje (npr. Kwak in Radler, 2002; Roster, Rogers, Albaum in Klein, 2004, str. 371), da se v spletno anketiranje vključujejo predvsem mlajši in da se na ta način v raziskave vključi le malo starejših respondentov.

Analiza vključenih v vzorec potrjuje navedbe prejšnjih analiz (npr. Kwak in Radler, 2002; Fox, Murray in Warm 2003; Roster, Rogers, Albaum in Klein, 2004; Wright, 2005), da se značilnosti ciljne populacije razlikujejo od značilnosti vzorca, kar posledično pomeni, da vzorca ne moremo označiti za reprezentativnega in na tej osnovi tudi težko posplošujemo rezultate raziskav, ki temeljijo na spletnem zbiranju podatkov.

4. Sklep

Spletno zbiranje podatkov je zaradi svojih prednosti postalo pogosto uporabljen način zbiranja podatkov, vendar pa je, zlasti za namene znanstvenega raziskovanja, pri uporabi spletnega anketiranja potrebna tudi določena mera kritičnosti in previdnosti. Posebno pozornost je treba nameniti vprašanju vzorca, saj se v njem pravzaprav zrcalita tako dilema odzivnosti (kdo se odzove) kot tudi dilema etičnosti (npr. vprašanje prevzemanja virtualne identitete, večkratnega izpolnjevanja vprašalnika). Temeljno vprašanje pri vzorcu namreč je, ali se lahko na podlagi spletnih vzorcev, ki so, kot menijo nekateri (Frippiat, Marquis in Wiles-Poriter, 2012), predvsem neslučajnostni, in je potemtakem vprašljiva tudi njihova reprezentativnost (Lozar Manfreda, Vehovar in Batagelj, 2000; Fox, Murray in Warm, 2003; Wright, 2005), sploh lahko sklepa in posplošuje. Zapišemo lahko, da podobna dilema velja tudi za raziskovanje na področju vzgoje in izobraževanja.

Prva težava pri zagotavljanju slučajnostnega vzorca se pojavi že pri elektronskih poštnih naslovih, saj vsi zavodi nimajo objavljenih naslovov e-pošte posameznih strokovnih delavcev. Zato vabilo navadno pošljemo vodji in vodja naj bi ga nato preposlal posameznim strokovnim delavcem. Tudi v primeru take rešitve pa se poraja vprašanje pogostosti pregledovanja e-pošte, oziroma sploh same aktivnosti tega orodja. Nadalje se pojavi vprašanje odzivnosti, na katero vplivajo mnogi dejavniki (gl. npr. Cohen, Manion in Morrison, 2007; Vehovar, Lozar Manfreda in Koren, 2008). Kot so že ugotavljali (Kwak in Radler, 2002; Roster, Rogers, Albaum in Klein, 2004), se za izpolnjevanje spletnega vprašalnika pogosteje odločijo mlajši in više izobraženi. To ugotovitev lahko prenesemo tudi na pedagoško področje – pri analizi vzorca se je namreč pokazalo, da se je kljub drugačni zastopanosti v populaciji v raziskavo vključilo več vzgojiteljev kot pomočnikov (ki imajo načeloma nižjo stopnjo formalne izobrazbe) ter da so v prevladovali predvsem mlajši respondenti, medtem ko se vzgojitelji in pomočniki, ki so stari 50 in več let, za izpolnjevanje spletnega vprašalnika niso odločili. Izpostavimo lahko tudi dejstvo, da bi morebiti bila struktura vzorca drugačna, če bi v raziskavo vključili učitelje, katerih delo je, v primerjavi z vzgojitelji, v večji meri vezano na uporabo računalnika in interneta (npr. spletne redovalnice, dnevniki, zbornice, e-asistent ipd.).

Sklenemo lahko, da opozorila v zvezi z metodološkimi dilemami spletnega načina zbiranja podatkov veljajo tudi za raziskovanje področja vzgoje in izobraževanja.

Je pa pri tem treba pripomniti tudi, da bi bilo v zvezi s spletnim zbiranjem podatkov smiselno in koristno opraviti tudi poglobljene analize glede veljavnosti in zanesljivosti takih raziskav, pri čemer bi kazalo uporabiti npr. t.i. “split-samples”, torej bi isto raziskavo izvedli s pomočjo dveh različnih, vendar primerljivih skupin, lahko pa bi se odločili tudi za zaporedni raziskavi, pri čemer bi pri prvi anketiranci izpolnjevali spletni, pri drugi pa klasični (papirnati) vprašalnik. Na ta način bi morebiti tudi poiskali in predlagali morebitne pristope, s katerimi bi lahko ublažili slabosti spletnega zbiranja podatkov, saj gre navsezadnje za zelo razširjen način zbiranja podatkov, ki tudi na področju pedagoškega raziskovanja vedno bolj izpodriva že dolgo uveljavljene, klasične načine zbiranja podatkov.

Tina Štemberger, PhD

Some dilemmas of web-based research in an educational research context

Data collection is one of the most important phases in various research processes (Sagadin 1993, p. 21; Cencič 2009, p. 44). The fact is that the collected data are the basis for drawing conclusions. In quantitative research, data are most commonly collected by means of a questionnaire (Cencič 2009). When designing a questionnaire, one has to pay attention to and decide upon many issues, including the mode of data collection: in person, on the telephone, by post or via the Internet. However, web-based data collection has become fairly widely used over the last ten years and this is why we present the basic characteristics, the process and the advantages of web-based surveys. Our aim is also to inform researchers about the dilemmas and limits of such surveys that they should take into account when planning research prospectively based upon a web-based questionnaire.

Web-based questionnaires date back to the 1990s. Because of the advantages they offer, they have almost completely supplanted other modes of data collection. The main advantages seem to be that they save money and time, are not geographically restricted and can draw upon various multimedia effects to render them more attractive. The basis of web-based data collection is a questionnaire, which has to be substantively and methodologically sound; at the same time, various visual, audial and other effects need to be considered. Subsequently, the database of potential respondents has to be prepared. The next phase encompasses the despatch of invitations and any reminders required to complete web questionnaires.

However, when using web-based questionnaires, some important issues need to be considered, i.e. sampling, response and ethics. Some researchers claim that the web-based survey cannot provide a random sample, which consequently means that the results cannot be generalised. Web samples do not include those who do not have access to the Internet, nor do they cater for groups of people who tend to use the Internet to a

limited extent, such as older people, poorly educated and socially disadvantaged people and those who are not very skilful with computers. Another significant problem consists of inappropriate or invalid e-mail addresses.

The response can also be a significant problem. The fact is that the response rate is, on average, 11% lower for web-based questionnaires than it is for other modes of data collection. Another problem is the drop-out rate: many respondents stop completing the questionnaire before they reach the last question. The response rate is influenced by many factors and can be higher if one considers all the circumstances when planning research.

Ethical dilemmas also arise. On one hand, it seems that social norms are less important in the virtual world and that answers do not have to confirm to such norms, but on the other hand, questions of insincerity, virtual identities and lack of authenticity arise. Furthermore, while respondents are completing the questionnaire, it is fairly likely they are doing other things (e.g. surfing other web pages, searching for information, reading the news, chatting online, etc.) at the same time, which can lead to inaccurate reading and answering. In addition to these dilemmas, there is also some mistrust about anonymity (due to traceability via the computer's IP address).

Based on the emphasized dilemmas of web-based data collection, we analysed the results of a web-based survey. A valid and reliable questionnaire was used, and objectivity was prioritised. The respondents included were pre-school teachers and their assistants.

We decided to contact our possible respondents via e-mail. An invitation to complete the questionnaire was sent to selected Slovenian kindergartens (one urban and one rural in any given region). The e-mail was sent to the head teachers or their assistants and they were requested to distribute it further to pre-school teachers and the assistants in their kindergartens. The first reminder was sent ten days after despatch of the invitation, and a second reminder was sent after a month.

A total of 732 pre-school teachers and their assistants responded to the invitation (they read the introduction to the questionnaire), 482 of them actually started completing it, whereas only 323 completed it fully. These results confirm the previous claims (Cohen, Manion & Morrison 2007) that many respondents read the introduction to a questionnaire, but not all of them proceed to complete it. We also found that the highest daily response rates were recorded immediately after issuing the two reminders. Based upon this analysis, we can state that the general difficulties of web-based questionnaires can be applied to educational research as well: invalid e-mails, low response rate, drop-out phenomena.

The question of representativeness of the sample was also taken into consideration. We were interested in the structural demographics of the respondents – specifically their professional role and their age. Population statistics reveal that there are a total of 9,887 pre-school teachers and their assistants, 47.8% being pre-school teachers and 52.2% being pre-school teachers' assistants. However, in the sample, the ratio differs: 69.7% are pre-school teachers and 30.3% are pre-school teachers' assistants. The difference between the population and the sample can be explained by the fact that, as a rule, assistants tend to be less educated than pre-school teachers, which reinforces the

findings about Internet usage (Kwak & Radler 2002; Roster, Rogers, Albaum in Klein 2004) which showed that the Internet is more often used by more educated people. We can claim that it was the mode of data collection that caused the difference between the demographic structure of the population and that of the sample.

Furthermore, we also found differences in the demographic structure of the population and the sample in terms of respondents' age. In the sample, there are no pre-school teachers or pre-school teachers' assistants aged 50 and above. Yet, this demographic group constitutes as much as 20% of the population. We can assume that the difference is due to the fact that Internet usage is more prevalent among younger users and that it is much harder to involve older people in completing web-based questionnaires.

The characteristics of the target population are, thus, different from the characteristics of the sample, thereby rendering the sample fairly non-representative; one needs to be cautious when generalising results based upon any web-based survey.

Despite the advantages of web-based data collection, it is necessary – especially in the instance of data collection for scientific purposes – to be cautious and critical when interpreting the results. Special attention needs to be paid to the sample, also reflecting the ethical dilemmas and the response dilemma. These generally acknowledged dilemmas can, by all means, be applied to the field of educational research as well. As presented in the example above, the web-based survey primarily encompassed younger respondents as well as those with a higher than average level of education. But it also seems important to pose a question about the – possibly – less appreciable difference between the structure of the population and the structure of the sample invited to complete the web-based questionnaire. The work of teachers, many of whom use so-called e-registers and e-assistants etc. on a daily basis, relies to a considerable extent upon being connected via the Internet.

In conclusion, there is the need for some in-depth research, which should be carried out by using split-sample research or sequential research processes. Such research might assist in finding some ways or approaches to mitigate the drawbacks of web-based surveys. After all, this is a very common way of collecting data and has almost entirely substituted all other traditional modes of data collection.

LITERATURA

1. Cencič, M. (2009). Kako poteka pedagoško raziskovanje. Primer kvantitativne empirične neeksperimentalne raziskave. Ljubljana: Zavod republike Slovenije za šolstvo.
2. Cohen, L., Manion, L., Morrison, K. (2007). Research Methods in Education. London in New York: Routledge.
3. Cook, C., Heath, F., Thompson, R.L. (2000). A Meta-Analysis of Response Rates in Web-or Internet-Base Surveys. Educational and Psychological Measurement, 60, št. 6, str. 821–836.
4. Čagran, B. (2004). Univariatna in multivariatna analiza podatkov: zbirka primerov uporabe statističnih metod s SPSS. Maribor: Pedagoška fakulteta.
5. Dillman D.A., Smyth, J.D., Christian, L.M., Stern, M.J. (2003). Multiple answer questions in self-administered surveys: the use of check-all-that-apply and force-choice question formats.

- Pridobljeno dne 30.03.2014 s svetovnega spleta: http://www.websm.org/uploadi/editor/File/Dillman_2002_Multiple_Answer_Questions.pdf.
6. Frippiat, D., Marquis, N., Wiles–Poriter, E. (2012). Web Surveys in the Social Sciences: an Overview. *Population* (English Edition), 65, št. 2, str. 285–311.
 7. Fox, J., Murray, C., Warm, A. (2003). Conducting research using web–based questionnaires: practical, methodological and ethical considerations. *International Journal of Social Research Methodology*, 6, št. 2, str. 167–180.
 8. 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, št. 5, str. 836–846.
 9. Kwak, N., Radler, B. (2002). A Comparison Between Mail and Web Surveys: Response Pattern, Respondent Profile, and Data Quality. *Journal of Official Statistics*, 18, št. 2, str. 257–273.
 10. Lozar Manfreda K. (2001). Web Survey Errors. Doktorska disertacija. Ljubljana: Fakulteta za družbene vede.
 11. Lozar Manfreda, K., Vehovar, V. (2002). Do Mail and Web Surveys Provide Same Results? Metodološki zvezki, 18, 149–169.
 12. Lozar Manfreda, K., Vehovar, V., Batagelj, Z. (2000). Veljavnost interneta kot anketnega orodja. *Teorija in praksa*, 37, št. 6, str. 1035–1051.
 13. Lozar Manfreda, K., Batagelj, Z., Vehovar, V. (2002). Design of Web Survey Questionnaires: The Three Basic Experiments. *Journal of Computer–Mediated Communication*, 7, št. 3, str. 1–24.
 14. Pitkow, J.E., Colleen, K. (1997). GUV's 7th WWW User Survey Online. Pridobljeno dne 15.03.2014 s svetovnega spleta: www.guv.gatech.edu/user_surveys/survey-1997-04.
 15. Roster, C.A., Rogres, R.D., Albaum, G., Klein, D. (2004) A Comparison of response characteristics form web and telephone surveys. *International Journal of Market Research*, 4, št. 6, str. 359–373.
 16. Zagadin, J. (1993). Poglavja iz metodologije pedagoškega raziskovanja. Ljubljana: Zavod republike Slovenije za šolstvo in šport.
 17. Seznam vrtcev (2012). Pridobljeno dne 28.10.2012 s svetovnega spleta: <https://krka1.mss.edus.si/resgistrirweb/SeznamVrtci.aspx>.
 18. Silber, H., Lischewski, J., Leibold, J. (2013). Comparing Different Types of Web Surveys: Examining Drop–outs, Non–Response and Social Desirability. *Metodološki zvezki*, 10, št. 2, 121–143.
 19. Status ankete Pomen ustvarjalnosti vzgojitelja v njegovi vzgojno–izobraževalni praksi (2012). Pridobljeno dne 25.04.2014 s svetovnega spleta: <https://www.1ka.si/admin/survey/index.php?anketa=18814&a=reporti>.
 20. Vehovar, V., Lozar Manfreda, K., Koren, G. (2008). Internet Surveys. V: Donsbach, W. in Traugott, M. W. (ur.). *The Sage Handbook of Public Opinion Research*. Los Angeles: Sage, str. 271–283.
 21. Zaposleni na delovnem mestu vzgojitelja po izobrazbi in spolu. (2012). Statistični urad Republike Slovenije. Pridobljeno dne 15.09.2014 s svetovnega spleta: <http://pxweb.stat.si/pxweb/Dialog/Saveshow.asp>.
 22. Wright, K.B. (2005). Researching Internet–based populations: Advantages and disadvantages of online survey research, online questionnaire authoring software packages, and web survey services. *Journal of Computer–Mediated Communication*, 10, št. 3. Pridobljeno dne 09.03.2014 s svetovnega spleta: <http://jcmc.indiana.edu/vol10/issue3/wright.html>.

Dr. Tina Štemberger (1980), docentka za področje pedagoške metodologije na Pedagoški fakulteti Univerze v Kopru.

Naslov: Podgraje 27, 6250 Ilirska Bistrica, Slovenija; Telefon: (+386) 041 782 741

E-mail: tina.stemberger@pef.upr.si