

UPORABA EKSPERTNIH SHEM ZA KVALITATIVNO TESTIRANJE ANKETNIH VPRAŠALNIKOV**

Povzetek. V članku obravnavamo ekspertno evalvacijo, ki je ena od pogosteje uporabljanih kvalitativnih metod za testiranja anketnih vprašalnikov. Pogosto se uporablja samoumevno in nekritično v začetnih korakih oblikovanja anketnega vprašalnika. Uporaba ekspertnih shem naj bi zmanjša subjektivnost in nekonsistentnost pri ugotavljanju pomanjkljivosti v anketnem vprašalniku z ekspertno evalvacijo, vendar empiričnih dokazov za to še ni. V prispevku empirično ovrednotimo uporabo dveh ekspertnih shem v več eksperimentih. Izkaže se, da je ena od shem po enem kriterijih kakovosti bolj primerna za uporabo pri ekspertni evalvaciji. Ker imata shemi drugačno zasnovo in deloma odkrijeta druge vrste pomanjkljivosti, priporočamo sočasno uporabo obeh shem.

Ključni pojmi: kognitivni laboratorij, kvalitativne metode, kakovost anketnega vprašalnika, ekspertna evalvacija, ekspertne sheme, Friedmanov test.

Ekspertna evalvacija kot kvalitativna metoda preverjanja kakovosti anketnega vprašalnika

Kakovost anketnega vprašalnika (Snijkers, 2002) je neposredno povezana z odsotnostjo sistematičnih napak, ki lahko nastanejo zaradi uporabe konkretnega anketnega vprašalnika (oblika, zaporedje vprašanj, vpliv konteksta, npr. Uhan, 1998) in vprašanja (besedilo, vrstni red odgovorov, struktura vprašanj). V splošnem lahko razlikujemo dve komponenti kakovosti in vrednotenja kakovosti, ki sta v konkretnem anketnem vprašalniku povezani (Groves et al., 2004). Najprej gre za kakovost, ki se udejanji v procesu anketiranja in je povezana z ljudmi, ki bodo na vprašanja odgovarjali. Pri tem skušamo zadostiti dvema standardoma kakovosti (kognitivni standard kakovosti in standard uporabnosti). Kognitivni standard se meri z visoko stopnjo

* Teja Cerar, Nina Konavec, študentki magistrskega študija Družboslovne informatike na Fakulteti za družbene vede; dr. Valentina Hlebec, izredna profesorica na Fakulteti za družbene vede, Univerza v Ljubljani.

** Izvirni znanstveni članek.

razumevanja vprašanj; temu je zadoščeno, ko anketiranci konsistentno razumejo anketna vprašanja, imajo zadostno količino informacij za oblikovanje odgovora ter zmorejo oblikovati in dati odgovor. Standard uporabnosti se meri z relativno nizko stopnjo težavnosti odgovarjanja, ki jo dosežemo, ko anketiranci (in anketar) zmorejo z zmerno mero truda ustrezno odgovoriti na anketna vprašanja in izpolniti vprašalnik z izbrano metodo anketiranja. Prav tako pa raziskovalca zanima tudi kakovost rezultatov merjenja – kakovost spremenljivk, ki smo jih hoteli meriti (veljavnost in zanesljivost). Raziskovalci skušajo zmanjšati število napak ali povsem odstraniti sistematične napake v anketnem vprašalniku z večkratnim preverjanjem kakovosti oziroma z večkratnim testiranjem anketnega vprašalnika. Različne klasifikacije metod za preverjanje anketnih vprašalnikov (Bergman, 1995; Akkerboom in Dehue, 1997; American Statistical Association – ASA, 1997: 1; Snijkers, 2002; Scheuren, 2004: 45; DeMaio et al., 2006: 1) metode razlikujejo glede na to:

- (1) ali se uporabljajo za preverjanje posameznih vprašanj ali vprašalnika v celoti – specifične in splošne;
- (2) ali se izvajajo v raziskovalni inštituciji in metodološkem laboratoriju ali na »terenu« – »predterenske« oziroma laboratorijske in terenske;
- (3) ali vključujejo anketirance ali ne – reaktivne in nereaktivne;
- (4) glede na metode, ki se uporabljajo pri testiranju – kvalitativne in kvantitativne;
- (5) glede na fazo v anketni raziskavi – pred zbiranjem podatkov in po njem;
- (6) glede na modele kakovosti anketne raziskave – večstopenjski modeli testiranja (1. definicija in izvedljivost raziskave, 2. kvalitativni laboratorijski testi, 3. kvalitativni operativni terenski testi, 4. kvantitativna pilotna študija, 5. implementacija raziskave).

V prispevku obravnavamo metodo ekspertne evalvacije anketnih vprašalnikov, ki spada med predterenske ali laboratorijske, nereaktivne in kvalitativne metode testiranja, ki se uporabijo pred zbiranjem podatkov v prvih treh fazah petstopenjskega modela testiranja anketnega vprašalnika. V nadaljevanju metodo najprej opredelimo, nato izpostavimo njene prednosti in pomanjkljivosti, nazadnje pa z metaanalizo številnih eksperimentov ovrednotimo uporabo dveh ekspertnih shem.

Ekspertna evalvacija je (Forsyth et al., 1999: 145) kvalitativna metoda za pregled in ovrednotenje vprašalnikov. Cilj ekspertne evalvacije vprašalnika je prepoznavanje težav posameznih vprašanj in morebitna predelava vprašanj z namenom izboljšanja vsebine in oblike vprašanja. Kriteriji za ovrednotenje posameznega vprašanja naj bi bili standardizirani (npr. z uporabo ekspertnih shem), kar omogoča, da so rezultati evalvacije kar se da konsistentni. Pri ekspertni evalvaciji sodelujeta dve skupini ekspertov, in sicer eksperti za vsebino, ki preverjajo, ali vsebina (besedilo) vprašanj

meri načrtovano teoretično razsežnost ali koncept, in eksperti za oblikovanje anketnega vprašalnika, ki preverjajo, ali posamezno anketno vprašanje in vprašalnik v celoti dosegajo metodološke standarde kakovosti. Eksperti za obliko anketnega vprašalnika podrobneje pregledajo besedilo vprašanja, strukturo vprašanja, dane odgovore, zaporedje vprašanj, navodila za izpolnjevanje in logično ureditev vprašanj (preskoki).

Načeloma naj bi vprašalnik individualno pregledali najmanj trije eksperti in največ deset.. Končni rezultat je skupno poročilo o odkritih napakah in priporočilih glede izboljšanja vprašalnika (Snijkers, 2002: 69), ki ga lahko eksperti naredijo skupaj v procesu diskusije in medsebojnega usklajevanja; skupno poročilo lahko naredi tudi izbrani ekspert. Pri ekspertni evalvaciji je možna uporaba seznamov ali shem mogočih napak. Z uporabo shem naj bi zmanjšali subjektivnost presojanja kakovosti, saj se lahko zgodi, da imajo eksperti različne poglede na idealno obliko posameznega anketnega vprašanja (npr. sodo ali liho število odgovorov, označevanje samo ekstremnih vrednosti ali poimenovanje vseh odgovorov pri ordinalnih lestvicah).

Poznamo več shem, ki se uporabljajo za odkrivanje napak v anketnih vprašanjih. Najenostavnejša shema je povzeta po priporočilih za kakovostno oblikovanje anketnih vprašalnikov, ki so jih pripravili na švedskem statističnem uradu, in je preprost seznam navodil za oblikovanje anketnih vprašalnikov (Designing questionnaires: Checklist, poimenovali jo bomo CHL-shema; navodila so predstavljena v prilogi, Statistics Sweden, 1992). Njena posebnost je, da se ne osredini le na preverjanje kakovosti anketnih vprašanj in odgovorov ter vprašalnika v celoti, temveč ovrednoti celoten postopek anketne raziskave, začenši s preverjanjem teoretskih osnov vprašanj, ki jih vključuje vprašalnik.

Kompleksnejšo shemo za oblikovanje in preverjanje besedila anketnih vprašanj sta oblikovala Saris in Gallhofer (1997). Povzela in dopolnila sta že obstoječe klasifikacije anketnih vprašanj, ki sta jih nadgradila s formalnim lingvističnim pristopom. Njun sistem ima več uporabnih namenov (oblikovanje vprašanj v fazi operacionalizacije, klasifikacija oblikovanih anketnih vprašanj in ugotavljanje nejasno ali slabo ubesedenih anketnih vprašanj). Kogovšek (1998) je shemo klasifikacije uporabila za ugotavljanje kakovosti anketnih vprašanj iz vzorca vprašanj iz dnevnika Delo (97 vprašanj, ki so bila v Delu objavljena med 1. septembrom 1997 in 18. januarjem 1998). Ugotovila je, da lahko razvrsti večino vprašanj. Pri nerazvrščenih je bila polovica takih, ki jih ni bilo mogoče enolično uvrstiti, ker sistem dopušča neenolično klasifikacijo. Pri drugi polovici nerazvrščenih vprašanj ni bilo mogoče razbrati, na katerega od konceptov vprašanje meri, in je verjetno, da so v teh primerih anketiranci odgovarjali na različna vprašanja. Shemo je uporabila tudi Nadrag (2002) ter jo dopolnila z naborom najpogostejših napak.

Forsyth in Lessler (1991) sta za ekspertno evalvacijo razvila shemo

»Kodirni sistem za evalvacijo vprašalnikov« (angl. Questionnaire Appraisal Coding System, v nadaljevanju QACS-shema). Kodirna shema temelji na strukturi procesa odgovaranja (glej npr. Strack in Martin, 1987) in vsebuje šest področij. Najprej se ugotavlja morebitne težave pri razumevanju navodil, vprašanj in odgovorov, nato težavnost interpretacije, spominjanje in ovrednotenje informacij ter probleme pri oblikovanju ali izbiri odgovorov. Akkerboom in Dehue sta QACS-shemo zaradi prevelike kompleksnosti (Snijkers, 2002: 69) poenostavila v t. i. eklektično shemo za odkrivanje napak v anketnem vprašalniku. V uporabi je še več drugih shem, ki pa spominjajo na že omenjene sheme, zato jih ne bomo podrobneje predstavili.

Ekspertna evalvacija ima nekaj pozitivnih značilnosti in nekaj pomanjkljivosti. Po eni strani so raziskovalci prepričani, da je metoda zanesljiva in pomembna za identifikacijo morebitnih težav v vprašalniku, saj odkrije veliko napak (Snijkers, 2002: 13; DeMaio et al., 2006: 5). Prav tako je časovno in stroškovno zelo učinkovita (Forsyth et al., 1999: 145; Snijkers, 2002: 11). V primerjavi z ekspertno evalvacijo brez uporabe shem je ekspertna evalvacija z uporabo shem bolj sistematična (DeMaio et al., 2006: 5). Po drugi strani pa je količina zaznanih napak odvisna od števila ekspertov ter njihovega predznanja in izkušenj. Prav tako metoda ni prepričljiva za potencialne stranke zaradi nereaktivnosti – anketiranci pri tej metodi preverjanja kakovosti vprašalnika ne sodelujejo (Snijkers, 2002: 71). Najpomembnejša pomanjkljivost pa je subjektivnost ugotovitev, ker je metoda sama odvisna od predznanja ekspertov oz. od področja dela, s katerega izhajajo (DeMaio et al., 2006: 13). Tako se lahko zgodi, kljub uporabi ekspertne sheme, da bosta dva eksperta različno ocenila kakovost posameznega anketnega vprašanja (npr. enemu se bo zdelo anketno vprašanje dvoumno, drugemu pa ne, lahko tudi priporočita različno vsebino ali obliko odgovorov). Vseeno pa so take in podobne sheme lahko učinkovit pripomoček za iskanje slabše oblikovanih anketnih vprašanj. Ekspertne sheme se pogosteje uporabljajo za odpravljanje napak pri vprašalnikih za akademske raziskave, kot uveljavljeno metodo pa jo predstavljajo npr. v nizozemskem ali švedskem statističnem uradu ali na raziskovalnih inštitutih v ZDA, npr. Research Triangle Institute (glej Snijkers, 2002).

V članku bomo ovrednotili uporabo dveh ekspertnih shem (CHL in QACS), ki se pogosto uporabljata za ekspertno evalvacijo anketnih vprašalnikov. Zanimalo nas bo:

- (1) Ali eksperti na konsistenten ali na različen način uporabljajo shemi? Ena od ključnih pomanjkljivosti ekspertne evalvacije je subjektivnost pri ocenjevanju pomanjkljivosti v vprašalniku, zato je soglasnost ekspertov dober pokazatelj kakovosti sheme. Shema je boljša oziroma bolj priporočljiva za uporabo, če jo eksperti uporabljajo na podoben način.

- (2) S katero shemo zaznamo največ vsebinskih ali strukturnih pomanjkljivosti? Domnevamo, da je večje število odkritih pomanjkljivosti pokazatelj večje kakovosti sheme.
- (3) Katere napake v vprašanjih odkrijemo z določeno shemo? Ob številu odkritih pomanjkljivosti je pomembno, katere napake odkrije specifična shema. Shema, v kateri so določene vrste napak izpuščene, pa domnevamo, da bi lahko vplivale na kakovost podatkov, ni dobra za uporabo.

Domnevamo, da je določena shema boljša za uporabo takrat, ko jo eksperti uporabljajo na konsistenten način, s katero odkrijemo več napak in s katero odkrijemo ključne napake. Na področju ekspertne evalvacije ni zaslediti podobnih primerjav, zato bodo rezultati pripomogli k lažjemu in bolj kakovostnemu ocenjevanju in oblikovanju anketnih vprašanj.

Postopek preverjanja enakovrednosti ekspertnih shem

Obe shemi (CHL in QACS) se uporabljata pri predmetu Oblikovanje anketnega vprašalnika na Fakulteti za družbene vede za presojanje kakovosti anketnih vprašalnikov, ki jih raziskovalci ponudijo za učno gradivo. Da bi shemi lažje med seboj primerjali oziroma da bi primerjali kriterije, ki ocenjujejo iste vrste napak, smo pred primerjavo shem izbrali le tiste kategorije napak pri vsaki shemi, ki merijo vsebinsko iste napake. Podane so v Tabeli 1. Še enkrat je potrebno poudariti, da CHL-shema služi kot orodje, ki ocenjevalcu narekuje splošna navodila oziroma korake bodisi že za samo oblikovanje vprašanj bodisi za kasnejše preverjanje le-teh. QACS-shema pa opredeljuje seznam napak v vprašanjih, na katere moramo biti pozorni.

Podobnosti med shemama se torej nanašajo na težave v navodilih ter v jeziku vprašanja (morda zahtevna slovnica, nerazumljivi ali dvoumni tehnični izrazi) in na podrobnosti vprašanja (ali sprašuje le eno stvar), izpostavljanje zavajajočih vprašanj, ujemanje med vprašanji in odgovori ter na razne morebitne težavnosti v podanih odgovorih.

Ugotovili bi radi, ali se eksperti strinjajo glede načina uporabe CHL- in QACS-shem oziroma katera shema je boljša in učinkovitejša za uporabo. Domnevali smo, da je boljša za uporabo tista shema, kjer se eksperti bolj strinjajo med seboj. Na podlagi tega smo izbrali mero skladnosti ekspertov (ang. *intercoder agreement*), s katero bi lahko ocenili, v kolikšni meri eksperti delujejo podobno. Za primerjavo smo izbrali Friedmanov test, ki omogoča primerjavo nominalnih ocen (navzočnost določene napake v vprašanju (kodirano z 1) in odsotnost napake v vprašanju (kodirano z 0)).

Friedmanov test se uporablja v primeru, ko imamo več ponovitev meritev na istih subjektih (Texasoft, 2004). S testom smo merili število odkritih napak ekspertov na dveh anketnih vprašalnikih. Test uporablja range,

v katere so razvrščeni podatki. Da smo lahko izračunali range, je bilo potrebno v analizo vključiti celoten vprašalnik oz. vsa zaporedna vprašanja ter napake ekspertov. Tako smo ocenjevali strinjanje med eksperti na celotnem vprašalniku. Friedmanov test namreč potrebuje več vrednosti, da lahko razvrsti podatke, zato ocena strinjanja samo na enem samem vprašanju ni bila mogoča. Ko vključimo vsa vprašanja, pa lahko nato na podlagi rangov izračunamo stopnjo značilnosti, s katero ocenimo, kakšno je strinjanje med eksperti.

Tabela 1: PODOBNOSTI KATEGORIJ NAPAK MED SHEMAMA CHL IN QACS

	CHL	QACS
NAVODILA	Enostavna, splošna, nezavajajoča	Nasprotujoča, nenatančna, skrita
JEZIK	Enostaven jezik	Zahtevna slovnica v vprašanju, tehnični, nedefiniran, nejasen/dvoumen izraz
JASNOST VPR.	Jasno vprašanje	Nejasen problem, tehnični izrazi v vprašanju, več definicij v enem vprašanju, nejasen cilj vprašanja
ENA STVAR VPR.	Vprašanje sprašuje le eno stvar	Več vprašanj v enem, vsebuje več različnih trditvev
VPRAŠANJE	Zavajajoče vprašanje	Zakrito vprašanje
VSEBINSKI SKLOP VPR.	Vprašanje postavljeno v logično in vsebinsko povezan sklop	Ujemanje med vprašanjem in odgovorom
ODGOVORI	Odgovori vsebujejo vse možnosti, vse vrednosti (različnost ekstremov), odg. tipa da-ne je omejujoč, nevtralni odgovor »Ne vem«	Manjkajoče kategorije
ODGOVORI	Odgovori se med seboj izključujejo	Prekrivajoče kategorije
ODGOVORI	Zaporedje odgovorov	Zaporedje odgovorov
ODGOVORI	Preveliko število odgovorov	Več možnih odgovorov kljub navodilom za obkrožitev enega

Test primerja tri ali več parov skupin. Ekspert predstavlja eno skupino. Friedmanov test nato vrednosti v vsaki vrsti razvrsti najprej od najmanjše do največje, vsaka vrsta pa je rangirana še posebej. Nato v vsakem stolpcu oziroma skupini, ki predstavlja eksperta, sešteje vse range. Če so si vsote med seboj različne, potem bo stopnja značilnosti majhna in bo zadovoljevala kriterij 5-odstotnega tveganja ob zavrnitvi ničelne domneve, s katero trdimo, da eksperti pri ocenjevanju anketnih vprašalnikov uporabljajo shemi na podoben način. Tako smo lahko s 95-odstotno gotovostjo ugotavljali, da obstaja razlika med vključenimi eksperti pri ocenjevanju vprašalnikov z obema shemama.

Opis eksperimentov

V članku bomo uporabili evalvacijo dveh anketnih vprašalnikov (Dejavniki razvoja kohezivnosti in sodelovanja v soseskah (Filipovič Hrast, 2008) in Avtoritetni odnosi v družini (Kuhar, 2008)), ki so bili ovrednoteni v letu 2007/08 in 2008/09. Na obeh vprašalnikih smo najprej izvedli analizo po ekspertih. Na podanih ocenah oziroma na številu napak, ki jih je posamezni ekspert opazil na vprašalniku, smo izvedli Friedmanov test glede na obe uporabljeni shemi. Nato smo pregledali število napak med eksperti pri obeh shemah na obeh vprašalnikih.

Prvi vprašalnik je sestavljen iz 57 vprašanj in 4 sklopov: stanovanje in soseska, delovanje v soseski, odnosi v soseski, demografija. Ocenjen je bil z uporabo CHL- in QACS-sheme, ocenjevalo pa ga je 6 ekspertov (E1E2,¹ E17, E18, E19, E20). Drugi vprašalnik je sestavljen iz 32 vprašanj. Prvi del vprašalnika se nanaša na odnose med staršem in otrokom, drugi del sprašuje o odnosih v družini, zadnji del vprašalnika pa sestavlja demografija. Ocenjen je bil z uporabo CHL- in QACS-sheme, ocenjevalo pa ga je 5 ekspertov (E11, E12, E13, E14, E2). Vsi eksperti, razen ekspert E2, ki je ocenil celoten vprašalnik, so ocenili vprašalnik do vključno 19. vprašanja.

Tabela 2: PREGLED POVPREČNIH RANGOV PO EKSPERTIH

Dejavniki kohezivnosti v soseskah			Avtoritetni odnosi v družini		
Schema	Ekspert	Povprečni rang	Schema	Ekspert	Povprečni rang
CHL	E1E2	3.27	CHL	E2	4.00
	E17	2.89		E11	3.26
	E18	2.53		E12	1.92
	E19	3.20		E13	2.82
	E20	3.11		E14	3.00
QACS	E1E2	2.82	QACS	E2	4.58
	E17	3.15		E11	2.50
	E18	2.22		E12	1.92
	E19	3.97		E13	2.50
	E20	2.84		E14	3.50

Na podlagi rezultatov v Tabeli 2, ki prikazuje porazdelitev povprečnih rangov, ključnih za zavrnitev ali sprejetje zastavljenih domnev, in v Tabeli 3, kjer so prikazane vrednosti Friedmanovega testa, ugotavljamo, da

¹ Eksperta E1 in E2 sta v tem primeru zaradi lažje primerjave (ne)strinjanja med vsemi eksperti združena v enega (E1E2), saj sta vsak pregledala pol vprašalnika.

eksperti uporabljajo shemi na različen način. Kot omenjeno, Friedmanov test namreč deluje tako, da primerja več parov skupin. V Tabeli 2 je posamezen ekspert ponazorjen kot ena skupina, za katero primerjamo število napak v posameznem vprašanju oz skupno v celotnem vprašalniku. Friedmanov test vrednosti v vsaki vrsti razvrsti najprej od najmanjše do največje, vsaka vrsta pa je rangirana še posebej. Nato v vsaki skupini, ki predstavlja eksperta, sešteje vse range. Če so si vsote med seboj različne, potem je stopnja značilnosti majhna in bo zadovoljevala kriterij 5-odstotnega tveganja ob zavrnitvi ničelne domneve, ki jo prikazuje tabela 3.

Tabela 3: PREGLED VREDNOSTI FRIEDMANOVEGA TESTA

Vrednosti Friedmanovega testa				
Dejavniki kohezivnosti v soseskah				
Shema	Št. enot	Hi-kvadrat	Prostostne stopnje	Stopnja značilnosti
CHL	57	10.347	4	0.035
QACS	57	46.804	4	0.000
Avtoritetni odnosi v družini				
Shema	Št. enot	Hi-kvadrat	Prostostne stopnje	Stopnja značilnosti
CHL	19	21.329	4	0.000
QACS	19	48.395	4	0.000

Stopnja značilnosti Friedmanovega testa pri obeh shemah v obeh vprašalnikih namreč zadovoljuje 5-odstotni kriterij stopnje značilnosti ob zavrnitvi ničelne domneve (eksperti uporabljajo shemi na značilno različen način). Razlikovanje med eksperti nam omogoča nadaljnjo analizo podatkov po shemah. Pri tem moramo poudariti, da večje nestrinjanje med eksperti pomeni, da shemi nista bili enostavni za uporabo. V primeru statistično neznačilnega Friedmanovega testa, kjer bi šlo za večje strinjanje med eksperti pri uporabi shem, bi pomenilo, da je določena shema boljša za uporabo, ker jo eksperti uporabljajo na enak način in bolj konsistentno. Na osnovi Friedmanovega testa torej ne moremo reči, da je ena od shem boljša za uporabo glede na kriterij konsistentne (podobne) uporabe sheme z več eksperti.

Ker se uporaba obeh shem glede na vse obravnavane eksperte v vprašalnikih razlikuje, je potreben kvantitativni pregled števila pomanjkljivosti v vprašalnikih med eksperti pri obeh shemah. Iz tega lahko razberemo, kako je posamezen ekspert uporabil shemi in s katero je odkril največ pomanjkljivosti. To je prvi korak pri odgovarjanju na ključno vprašanje o tem, katera shema je učinkovitejša in bolj preprosta za uporabo.

Tabela 4: PREGLED ŠTEVILA POMANJKLJIVOSTI

DEJAVNIKI KOHEZIVNOSTI V SOSESKAH			AVTORITETNI ODNOSI V DRUŽINI		
Ekspert	CHL	QACS	Ekspert	CHL	QACS
E1E2	103 (1,8)	40 (0,7)	E2	33 (1,0)	41 (1,3)
E17	71 (1,2)	70 (1,2)	E11	14 (0,4)	8 (0,3)
E18	52 (0,9)	24 (0,4)	E12	3 (0,09)	1 (0,03)
E19	78 (1,4)	106 (1,9)	E13	13 (0,4)	8 (0,3)
E20	88 (1,5)	52 (0,9)	E14	15 (0,5)	18 (0,6)
Skupaj	392 (6,9)	292 (5,1)	Skupaj	78 (2,4)	76 (2,4)

V Tabeli 4 je predstavljeno absolutno in relativno (v oklepajih) število pomanjkljivosti, ki jih je posamezen ekspert zaznal ob pregledu vprašalnika. Relativno število dobimo tako, da število pomanjkljivosti delimo s številom vprašanih.

S kvantitativnega vidika je bil v prvem vprašalniku z uporabo QACS-sheme pri ocenjevanju najbolj kritičen ekspert E19², z uporabo CHL-sheme pa ekspert E1E2. S CHL-shemo je posamezen ekspert zaznal več pomanjkljivosti (skupno 392) v primerjavi s QACS-shemo (skupno 292). Razen pri ekspertu E19 je uporaba CHL-sheme pokazala več pomanjkljivosti kot uporaba QACS-sheme. V drugem vprašalniku pa je bil z uporabo obeh shem najbolj kritičen ekspert E2. Odkril je tudi največ pomanjkljivosti. Tudi v tem vprašalniku so eksperti s CHL-shemo skupno zaznali več pomanjkljivosti (78) kot z uporabo QACS-sheme (76). Razen pri dveh ekspertih (E2 in E14) je uporaba CHL-sheme pokazala več pomanjkljivosti kot QACS-shema. Iz teh podatkov lahko sklepamo, da je CHL-shema pri ocenjevanju nekoliko bolj učinkovita.

Poleg količine zaznanih pomanjkljivosti po posameznem ekspertu z uporabo obravnavanih shem nas je zanimalo tudi, katera so vprašanja, ki so se na skupno oceno ekspertov izkazala za najbolj problematična. Ob tem smo si v vprašalniku *Dejavniki kohezivnosti v soseskah* zastavili kriterij, s katerim smo bili pozornejši na tista vprašanja, katerim so eksperti skupno pripisali deset pomanjkljivosti ali več. V vprašalniku *Avtoritetni odnosi v družini* pa je bil kriterij določen pri petih vprašanjih ali več. Gre za določanje kriterijev problematičnih vprašanj glede na posamezen vprašalnik, saj so tematsko in po obsegu različni.

² Število pomanjkljivosti je lahko zelo veliko, saj sheme vključujejo veliko število kriterijev oziroma potencialnih pomanjkljivosti v vprašalniku. Odvisno je tudi od natančnosti in izkušenosti ekspertov, npr. izkušen ekspert bo »pomanjkljivosti« razvrstil po pomembnosti in kot napake označil le take pomanjkljivosti, ki bi lahko vplivale na kakovost podatkov. Manj izkušen ekspert pa lahko kot napako označi tudi malenkosti, ki ne vplivajo na nobenega od standardov kakovosti. Omeniti je tudi treba, da je šlo za delovni verziji vprašalnikov, ki nista bili še pripravljena za uporabo na terenu. Prav tako iz števila napak ne smemo sklepati o kakovosti teh dveh vprašalnikov, ker so jih ocenjevali različni eksperti. Prav mogoče je, da so bili eksperti, ki so vrednotili prvi vprašalnik, bolj kritični oziroma so svojo nalogo bolj zauzeto opravili.

Tabela 5: PREGLED VPRAŠANJ Z NAJVEČJIM ŠTEVILOM POMANJKLJIVOST

DEJAVNIKI KOHEZIVNOSTI V SOSESKAH			AVTORITETNI ODNOSI V DRUŽINI		
Shema	Vprašanje	Št.	Shema	Vprašanje	Št.
CHL	51, 52, 53, 54, 55, 56, 57	14	CHL	19	10
	23	12		10	8
	22	11		4	6
	20	10		11, 16, 18	5
QACS	16	13	QACS	19	8
	23	12		13, 16	7
	22, 48	10			

Eksperti so v vprašalniku *Dejavniki kohezivnosti v soseskah* z uporabo CHL-sheme odkrili največ 14 pomanjkljivosti pri večjem številu vprašanj (7), z uporabo QACS-sheme pa največ 13 napak pri samo enem vprašanju. V drugem vprašalniku so odkrili največ 10 napak s CHL-shemo in največ 8 napak s QACS-shemo. Zanimivo je, da so eksperti v obeh vprašalnikih z uporabo CHL-sheme v glavnem izpostavili različna, najbolj problematična vprašanja kot pa z uporabo QACS-sheme. Prav tu smo priča dejstvu, da se shemi že v osnovi razlikujeta med seboj. To pa je najverjetneje glavni razlog za tovrsten rezultat. Kljub temu pa so se štiri vprašanja v obeh vprašalnikih izkazala kot problematična po kriterijih CHL- in QACS-sheme. Skupno so eksperti tako v vprašalniku *Dejavniki kohezivnosti v soseskah* kot v vprašalniku *Avtoritetni odnosi v družini* izpostavili po dve vprašanji. Glede na CHL-shemo imajo omenjena vprašanja večje število napak kot glede na QACS-shemo.

Tabela 6: SKUPNO ŠTEVILO NAPAK PO OBEH SHEMAH IN DODATNO ŠTEVILO NAPAK GLEDE NA POSAMEZNO SHEMA V NAJBOLJ PROBLEMATIČNIH VPRAŠANJIH V VPRAŠALNIKI

VPRAŠALNIK	Skupno št. napak, ki so jih eksperti zaznali tako s CHL- kot QACS-shemo v najbolj problematičnih vprašanjih	Dodatno št. napak po CHL-shemi	Dodatno št. napak po QACS-shemi	Skupaj
Dejavniki kohezivnosti v soseskah	5	2	4	11
Avtoritetni odnosi v družini	5	3	6	14
Skupaj	10	5	10	25

Ugotovili smo tudi ključno razliko v uporabi shem CHL in QACS. Eksperti so s CHL-shemo v primerjavi s QACS-shemo dodatno ocenjevali tudi

zasnovno odgovorov v vprašanju (preveliko število odgovorov, odgovori ne vsebujejo vseh vrednosti oz. možnosti, manjka nevtralni odgovor »Ne vem«), medtem ko so z uporabo QACS bili pozornejši na navodila, slovnico (tehnični, nedefinirani, nejasni izrazi), cilj ali nejasnost vprašanja. Katere dodatne kategorije so bile poudarjene glede na shemo, pa je bilo odvisno od posameznega vprašanja.

Kot prikazuje tabela 7, je bilo s QACS-shemo največkrat poleg že omenjenih kategorij ocenjeno, da se porajajo napake v izrazih in navodilih, prihaja do neujemanja med vprašanjem in odgovorom, prihaja do nejasnosti v cilju in do problema v vprašanju (ali pa le-to ponuja več odgovorov, kot je zaželeno), da je zaporedje odgovorov neprimerno, slovnica zahtevna ter da se kategorije v odgovorih prekrivajo. S CHL-shemo pa so eksperti dodatno ocenili še kategorijo napak glede nejasnosti vprašanja, nejasna in zavajajoča navodila, neprimerno ali preveliko število odgovorov ter pomanjkanje vrednosti v njih ter da vprašanje ni postavljeno v logičen vsebinski sklop in ne sprašuje le po eni stvari.

V tabeli 6 je razvidno, da gre bistvu za podobne kategorije napak, le da so eksperti določeno napako pri istem vprašanju ocenili le z eno shemo.

Tabela 7: DODATNE KATEGORIJE NAPAK NAJDENE S POSAMEZNO SHEMO

Shema QACS	<p><i>Dodatne kategorije napak</i></p> <ul style="list-style-type: none"> - nedefiniran in nejasen oz. dvoumen izraz - neujemanje med vprašanjem in odgovorom - skrita, nenatančna navodila - nejasen cilj - nejasen problem - vprašanje ima več možnih odgovorov kljub navodilom, da se obkroži eno - več vprašanj v enem - neprimerno zaporedje odgovorov - nasprotujoča si navodila - nenatančna navodila - prekrivajoče se kategorije - vprašanje vsebuje zahtevno slovnico
Shema CHL	<p><i>Dodatne kategorije napak</i></p> <ul style="list-style-type: none"> - nejasno vprašanje - preveliko število odgovorov - ne sprašuje le eno stvar - zavajajoče vprašanje - zapletena ter zavajajoča navodila - neprimerno zaporedje - preveliko število odgovorov - odgovori ne vsebujejo vseh vrednosti - jezik ni enostaven - vprašanje ni postavljeno v logično in vsebinsko povezan sklop - manjka nevtralni odgovor »Ne vem«

Ugotovili smo, da prinaša uporaba CHL- in QACS-sheme na obeh vprašalnikih različne rezultate. S CHL-shemo je bilo moč ugotoviti večje število napak, medtem ko je bila z uporabo QACS-sheme poudarjena vsebinska plat napak v vprašanju. Prav zaradi tega bi s težavo sklenili, da ena shema omogoča bolj kakovostno ekspertno evalvacijo v primerjavi z drugo. Posebej je potrebno poudariti, da je bilo večina sodelujočih ekspertov neizkušenih, z malo predznanja, saj so vprašalnik ocenjevali študenti dodiplomskega študija, ki se poprej niso srečali z metodo ekspertne evalvacije. Tako moramo v ocenah upoštevati tudi subjektivnost, ki pa je prisotna tudi pri izkušenih ekspertih. To pa je tudi ključna kritika metode ekspertne evalvacije. Omejena pomanjkljivost pa ne vpliva na veljavnost ugotovitev o kakovosti primerjanih ekspertnih shem, saj v eksperimentalnem načrtu preverjamo vpliv ene same spremenljivke (tj. vrste sheme). V raziskavah, kjer se preverja vpliv metode na variabilnost rezultatov, so npr. kvotni vzorci ali vzorci na študentski populaciji povsem sprejemljivi, dokler gre za slučajno dodelitev uporabljenih metod v dve ali več skupin. Res pa je, da izkušeni eksperti z dolgotletnimi izkušnjami najbrž z večjo suverenostjo uporabljajo bolj kompleksne sheme, v našem primeru je to QACS-shema.

404

Ključne ugotovitve in diskusija

V članku smo hoteli ugotoviti, kako učinkovite so ekspertne sheme pri odkrivanju pomanjkljivosti v anketnih vprašalnikih ter s pomočjo katere od dveh uporabljenih shem eksperti najbolj kakovostno ocenjujejo dva anketna vprašalnika. Zastavili smo si tri raziskovalna vprašanja. Ali eksperti na konsistenten način uporabljajo shemi? S katero zaznamo največ vsebinskih ali strukturnih napak? Katere napake v vprašanjih odkrijemo z določeno shemo? Zanimalo nas je, katere so tiste kategorije napak, ki jih zaznamo z obema shemama hkrati, ter katere so tiste kategorije, ki jih dodatno izpostavimo z eno ali z drugo shemo. Shemi se v zasnovi razlikujeta. CHL-shema je sestavljena iz navodil, kako oblikovati vprašalnik, QACS-shema pa služi kot seznam kod oziroma napak. V začetni fazi smo za nadaljnje raziskovanje in za lažjo primerljivost izločili le tiste kategorije napak, ki vsebinsko klasificirajo identične napake.

Ključna ugotovitev je, da se eksperti pri uporabi shem značilno razlikujejo oziroma jih uporabljajo na različen način. Glede na kriterij konsistentne uporabe med eksperti ne moremo ugotoviti, katera shema je boljša za uporabo. S CHL-shemo so eksperti večinoma našli znatno več pomanjkljivosti kot z QACS-shemo. Glede na kriterij količine pomanjkljivosti je shema CHL bolj učinkovita. Glede na dejstvo, da so vsi eksperti uporabili obe shemi, je bila mogoča primerljivost najbolj kritičnih ekspertov. V obravnavi obeh vprašalnikov se je izkazalo, da sta eksperta E19 oziroma E2 z uporabo

obeh shem odkrila največje število napak. Ob tem je treba upoštevati, da je pomemben dejavnik pri ocenjevanju vprašalnikov subjektivna zaznava posameznega vprašanja. Slednje vpliva na količino ocenjenih napak pri posameznem ekspertu, neodvisno od uporabe sheme.

Poleg tega je zanimivo tudi, da so eksperti z uporabo obravnavanih shem kot najbolj problematična vprašanja izpostavili različna vprašanja pri istem vprašalniku. Tako smo lahko že na podlagi tega dognali, da z uporabo ene ali druge sheme opazimo različne probleme v določenem vprašanju.

Eksperti so z vsako shemo poudarili še dodatne kategorije napak. S QACS-shemo so eksperti glede na vse obravnavane vprašalnike našli največ napak, a vseeno ne moremo trditi, da je iz kvantitativnega vidika uporaba QACS-sheme bolj učinkovita. Kot smo že opozorili, je tovrstna primerjava omejena, saj je v vsakem vprašalniku sodelovalo različno število ekspertov in vprašalniki vsebujejo različno število vprašanj. Tako je bilo število napak večje pri tistem vprašalniku, ki ga je ocenjevalo večje število ekspertov, kot tudi glede na obseg vprašalnika. Ena od omejitev raziskave je tudi, da so vprašalnika ocenjevali različni eksperti. Iz ugotovitev torej ne moremo povsem izločiti tega vpliva.

Vseeno pa potrjujemo, da je v samem testiranju kakovosti anketnih vprašalnikov v fazi pretestiranja ekspertna evalvacija z uporabo kodirnih shem učinkovita metoda, ker se z njeno uporabo zaznajo zahtevnosti napake vprašalnika na različnih ravneh. Na podlagi analize (ne)strinjanja ekspertov v ocenjevanju s CHL- in QACS-shemo poudarjamo pomembnost uporabe različnih kodirnih shem za evalvacijo anketnih vprašanj, saj v kombinaciji več shem pridobimo boljše, podrobnejšo oceno možnih napak bodisi v sami zasnovi vprašanj bodisi celotnega vprašalnika. Uporaba shem v fazi testiranja anketnih vprašalnikov je torej ključna, če hočemo oblikovati kar se da dober anketni vprašalnik, s katerim bi dosegli najmanjšo možno mero neželenih napak, katere izvirajo že iz samega ogrodja oz. zgradbe vprašanj in vprašalnika v celoti.

Priloge

ŠVEDSKA SHEMA CHECK LIST

Koraki pri oblikovanju AV

Koraki pri oblikovanju anketnega vprašalnika so v shemi, ki so jo oblikovali na Švedskem (povzeto po: Designing questionnaires: Checklist. SCB, Statistics Sweden).

Ta navodila se lahko uporabijo za:

1. Pred (med) oblikovanjem anketnega vprašalnika.
2. Po oblikovanju anketnega vprašalnika (za zagotavljanje osnovnih pravil oblikovanja).
3. Za preverjanje vprašalnikov, ki so jih oblikovali drugi.
4. Za vrednotenje že zbranih anketnih podatkov.

Uvodni koraki

1. Oblikuj teoretične (koncept) in empirične spremenljivke.
 - a) Razišči vsebino glede na raziskovalni problem. (Začni na začetku!)
 - b) Dobro definiraj teoretične koncepte in spremenljivke ter empirične spremenljivke.
 - c) Izberi vzorec.
 - č) Naredi seznam spremenljivk (ime; anketno vprašanje, s katerim smo to spremenljivko merili; izčrpna definicija; vir; teoretični konstrukt, ki ga spremenljivka meri; enote merjenja; kakovost merjenja).
 - d) Pri demografskih spremenljivkah uporabi standardizirana vprašanja.
 - e) Poišči podobne raziskave in jih uporabi.
 - f) Grafično prikaži povezave med spremenljivkami.
 - g) Za isto teoretično spremenljivko uporabi več anketnih vprašanj.

Oblikovanje anketnih vprašanj

- a) Ne dodajaj vprašanj le zato, ker so zanimiva.
- b) Jezik naj bo preprost.
- c) Vprašanja naj bodo jasna in kratka.
- č) Definiraj dvoumne izraze.
- d) Postavi vprašanja v logično in vsebinsko povezane skupine.
- e) Premisli zaporedje vprašanj.
- f) V vsakem vprašanju vprašaj le eno stvar.
- g) Opredeli čas in prostor.
- h) Pazljivo oblikuj hipotetična vprašanja (npr. o načrtih v prihodnosti) in retrospektivna vprašanja (vprašanja o preteklosti).
- i) Izogibaj se vprašanj, ki odgovor vsiljujejo (npr. se strinjam).
- j) Izogibaj se čustveno obremenjenih besed in zavajajočih vprašanj.
- k) Vprašanja o stališčih vedno uravnovesi.

- l) Zagotovi, da vprašanja ne delujejo na anketiranca zastrašujoče ali žaljivo.

Vsako anketno vprašanje se lahko razume na različne načine!

Testiraj svoja vprašanja! Spremeni jih in začni znova!

Odgovori

a) Oblikovanje odgovorov

- i. Odgovori morajo vsebovati vse možnosti.
- ii. Odgovori naj vsebujejo vse možne vrednosti (ekstremi se morajo dovolj razlikovati).
- iii. Razmisli o navajanju sidrišča.
- iv. Uporabi lestvico, če je mogoče.
- v. Odprta vprašanja uporabi le, ko so nujno potrebna.

b) Vnaprej oblikovani odgovori

- i. Odgovori naj bodo izčrpní (ne pozabi na odgovor: »Drugo«).
- ii. Odgovori naj bodo medsebojno izključujoči.
- iii. Odgovorov naj ne bo preveč.
- iv. Odgovori tipa »da-ne« se lahko zdijo anketirancem preveč omejujoči.
- v. Anketiranca ne sili k odgovorom na vprašanja, o katerih nič ne vedo (dodaj nevtralni odgovor ali odgovor »Ne vem«).
- vi. Preveri zaporedje odgovorov.

Oblika vprašalnika

- a) Predstavi raziskovalno skupino.
- b) Imenuj kontaktno osebo.
- c) Predstavi raziskavo in vprašalnik.
- č) Dobro oblikuj vezna vprašanja med posameznimi vsebinskimi sklopi vprašanj.
- d) Vprašanja naj ne bodo natlačena (pusti dovolj prostora za odgovore).
- e) Odgovarjanje na vprašanja naj bo enostavno.
- f) Dodaj navodila za izpolnjevanje vprašanj (primeri naj bodo preprosti, splošni in ne smejo zavajati).
- g) Navodila za preskoke med vprašanji na bodo čim bolj enostavna. Naj jih bo čim manj.
- h) Uporabi ilustracije.
- i) Ne pozabi na logotip, papir vprašalnika je lahko obarvan. Tisk naj bo kakovosten.

Testiraj vprašalnik v celoti. Popravi ga in ga znova testiraj.

Forsyth in Lesser QACS-shema

Razumevanje		Priklic	Presoja	Referenčni okvir odgovorov
Omejitev vsebine Omejitev časa		Priklic informacij	Evalvacija priklicanih informacij	Ustvariti meje odgovorov Oblikovanje odgovorov
Referenčni okvir	Referenčni okvir			
Zahteve naloge	Možni problemi (cont'd)	Zahteve naloge	Zahteve naloge	Zahteve naloge
Tipi referenčnega okvira Sedanje značilnosti/ obnašanje Pretekle značilnosti/ obnašanje Sedanje stališče/ mnenje Preteklo stališče/ mnenje Splošno znanje	Vprašanja (Questions cont'd) Struktura vprašanj: Zakrita vprašanja Zahtevna slovnica Implicitne domneve V enem vprašanju več vprašanj V enem vprašanju več definicij Nejasen cilj Ujemanje med vprašanjem in odgovorom	Naloga priklica Priklic dogodka Priklic vrste dogodkov Priklic splošne informacije Priklic prejšnjega odgovora Priklic stališča	Tipi procesa (presojanja) Celota (vsota) ocenjevanja Določi dogodke +/- Določi ujemanja +/- Določi datum in začetek Določi starost Oceni trajanje Oceni povprečje Sestavljeno ocenjevanje	Opis odgovora Da/Ne Kvalitativni: nominalna Kvalitativni: ordinalna Kvalitativni: odprt Kvantitativni: številka Kvantitativni: kompleksna Trajanje Časovna točka Starost Odgovor na skrita vprašanja
Osredotočenost Poročilo o sebi Nadomestno poročilo	Povzetek problemov Nenatančen referenčni okvir Kompleksen ref. okvir Nepomembno stališče	Postopek spomina Po spominu Prepoznanje Posploševanje (hevristično) Mešano	Združevanje informacij Seštejeemo Kvalitativna presoja Kvantitativna presoja	Možne težave (problemi)
Nivo referenčnega okvirja Osnovni Podrejen Nadrejen Več nivojski	Referenčno obdobje	Kontekst priklica Splošno znanje o sebi Splošno znanje o svetu Posebno obnašanje (ali poskusi) Vrsta vedenja Čustvo/Stališče Časovna točka/ Obdobje	Možne težave	Navodila Zakrita navodila Zakrite definicije
Sprememba v referenčnem okvirju Področje spremembe Spremembe nivoja Nenadna sprememba Premestitev referenčnega okvirja	Zahteve naloge	Možni problemi Podrobnost Nezadostnost Nepričakovanost Sprememba: v psihološki referenčni okvir	Evalvacija priklicanih informacij Natančnost informacij Zasebno vedenje Zasebno stališče Zasebnost (na splošno) Družbena nezaželenost	Odgovori Terminologija: Tehnični izrazi Nedefinirani izrazi Nejasni/dvoumni izrazi Struktura odgovorov: Določanje meje Prikrivajoče kategorije Manjkajoče kategorije Neustrezno zaporedje odgovorov

Razumevanje		Priklic	Presoja	Referenčni okvir odgovorov
Omejitev vsebine Omejitev časa		Priklic informacij	Evalvacija priklicanih informacij	Ustvariti meje odgovorov Oblikovanje odgovorov
Referenčni okvir	Referenčni okvir			
Potencialni (možni) problemi	Specifično (določeno) obdobje Celo življenje 12-mesečno obdobje 30-dnevno obdobje ----- Danes Vezan na vedenje/ predhodno vprašanje Nedefinirano: sedanje vedenje		Vrednotenje posledic Varnostne posledice Pravne posledice Družbene posledice Vedenjske posledice	
Navodila Nasprotujoča navodila Nenatančna navodila Skrita navodila Zahtevna slovnica Nejasni problemi Nejasna podoba vprašanja	Sprememba referenčnega okvirja Sprememba od prejšnjega vprašanja: Ni spremembe Širši Ozek Majhna sprememba Oddaljene meje Bližnje meje		Soglasje med informacijami in odgovori Neujemanje	
Vprašanja Tehnični izrazi Sedanost Nedefiniranost Dvoumnost ali nejasnost	Možni problemi Nedoločene meje Slabo definirano obdobje Nedefinirano obdobje Umeščeno obdobje Premestitev referenč. obdobja			

LITERATURA

- ASA - American Statistical Association (1997): What is a survey? How to conduct pretesting? Dostopno preko http://av.fdvinfo.net/literatura/ASA_pretesting.pdf 10. 1. 2011.
- Akkerboom, H. in F. Dehue (1997): The Dutch Model of Data Collection Development for Official Surveys. *International Journal of Public Opinion Research* 9: 126-145.
- Bergman, L. R. (1995): Pretesting procedures at statistics Sweden's measurement, evaluation and development laboratory. *Journal of official statistics*, 11 (3): 309-323.

- DeMaio, T. J., N. Bates, D. Willimack in J. Ingold (2006): Pretesting Questionnaires and Related Materials for Surveys and Censuses. Dostopno preko <http://www.census.gov/srd/pretest-standards.pdf> 10. 1. 2011.
- Forsyth, B. H. in B. T. Lesser (1991): Cognitive laboratory methods: a taxonomy. V P. P. Biemer, R. M. Groves, L. E. Lyberg, N. A. Mathiowetz in S. Sudman (ur.), *Measurement Errors in Surveys*, 167–183. New York: Wiley.
- Forsyth, B., K. Levin in S. K. Fisher (1999): Test of an appraisal method for establishment survey questionnaires. Dostopno preko <http://av.fdvinfo.net/upload/editor/1205231197forsyth%20appraisal.pdf> 10. 1. 2011.
- Groves, R. M., F. J. Fowler Jr., M. P. Couper, J. M. Lepkowski, E. Singer in R. Tourangeau (2004): *Survey Methodology*. New Jersey: John Wiley & Sons, Inc.
- Kogovšek, T. (1998): Sistematično identificiranje slabo formuliranih anketnih vprašanj. *Teorija in praksa* 35 (6): 1118–1135.
- Nadrag, P. (2002): *Kakovost anketnega vprašalnika za evalvacijo programa PUM*. Diplomsko delo. Ljubljana: FDV.
- Saris, E. W. in I. N. Gallhofer (2007): *Design, Evaluation and Analysis of Questionnaires For Survey Research*. New Jersey: Wiley.
- Scheuren, F. (2004): What is a survey? Dostopno preko <http://www.amstat.org/sections/srms/pamphlet.pdf> 10. 1. 2011.
- Snijkers, G. (2002): *Cognitive Laboratory Experiences: On Pre-testing Computerised Questionnaires*. Doktorska disertacija. Utrecht: University of Utrecht. Dostopno preko http://www.jpsm.umd.edu/qdet/final_pdf_papers/Snijkers.pdf 11. 1. 2011.
- Statistics Sweden (1992): *Designing questionnaires: Checklist*. SCB: Statistics Sweden.
- Strack, F. in L. L. Martin (1987): Thinking, Judging, and Communication: A Process Account of Context Effects in Attitude Surveys. V H. J. Hippler, N. Schwarz in S. Sudman (ur.), *Social Information Processing and Survey Methodology*, 123–148. New York: Springer-Verlag.
- TexaSoft (2004): *WINKS statistics tutorials: Friedman's Test. Non-Parametric Repeated Measures Comparisons*. Dostopno preko <http://www.texasoft.com/tutorial-friedmans-test.html> 5. 8. 2010.
- Uhan, S. (1998): *Prava in neprava mnenja: vpliv konteksta v raziskovanju javnega mnenja*. Ljubljana: Fakulteta za družbene vede.