

UNIVERZA V LJUBLJANI
FAKULTETA ZA DRUŽBENE VEDE

Teja Cerar
Nina Konavec

Primerjava ekspertnih shem za testiranje anketnih vprašalnikov

Diplomsko delo

Ljubljana, 2010

UNIVERZA V LJUBLJANI
FAKULTETA ZA DRUŽBENE VEDE

Teja Cerar
Nina Konavec

Mentorica: izr. prof. Valentina Hlebec

Primerjava ekspertnih shem za testiranje anketnih vprašalnikov

Diplomsko delo

Ljubljana, 2010

Iskreno se zahvaljujemaizr. prof. dr. Valentini Hlebec za prijazne nasvete in predloge ter predvsem za spodbudne besede in podporo pri nastajanju diplomskega dela.

Primerjava ekspertnih shem za testiranje anketnih vprašalnikov

Pri oblikovanju anketnih vprašalnikov, ki so ključnega pomena za družboslovno raziskovanje, se uporablja ekspertna evalvacija. Temelji na prepoznavanju težav v vprašanjih in odpravljanju le-teh. Eksperti si pri ocenjevanju tehničnega in vsebinskega vidika vprašalnika pomagajo z ekspertnimi shemami z namenom oblikovanja čim bolj razumljivih vprašanj.

V diplomskem delu smo obravnavali švedsko Check list shemo ter QACS shemo, ki se v svoji zasnovi razlikujeta med seboj. Shemi smo za primerjavo preizkusili na različnih vprašalnikih. Glavni namen naloge je ugotoviti, ali eksperti na različen način uporabljajo shemi oziroma katera shema je lažja za uporabo. Zanimalo nas je, s katero shemo eksperti zaznajo največ vsebinskih ali strukturnih napak.

Diplomsko delo sestoji iz teoretičnega dela z opisom procesa oblikovanja anketnega vprašalnika, pomena ekspertne evalvacije ter obravnavani shemi. Sledi analitični del, kjer smo uporabili Friedmanov test, s katerim smo ugotavljali, ali so eksperti shemi uporabljali na podoben način. Celoten analitičen del se nanaša na raziskovanje kvalitativne in kvantitativne zaznave napak v posameznih vprašanjih, ki so jih eksperti odkrili z uporabo Check list in QACS sheme.

Ugotavljamo, da ne moremo trditi, katera shema je boljša za uporabo, saj so podatki le na primeru dveh vprašalnikov ustrezali pogojem primerljivosti obeh shem. Potrebne bi bile dodatne raziskave kakovosti ekspertnih shem.

Ključne besede: primerljivost ekspertnih shem, švedska Check list shema, QACS shema, ekspertna evalvacija, Friedmanov test.

Comparison of expert schemes for testing survey questionnaires

In designing of survey questionnaires, which are the key to social research, expert evaluation is being used. It is based on identifying problems within questions and abolishing them. With the assessment of technical and substantive aspects of questionnaire, experts help themselves using expert schemes in order to create the most understandable questions.

In this diploma we have used Sweden Journal Check list scheme and QACS scheme, which conceptually differ from one another. We have tested these two schemes for comparison on different questionnaires. The main purpose of diploma is to identify, if experts are using the scheme in a different way and if any of the schemes is easier to use. We were trying to determine with which scheme experts had perceived most substantive or structural errors.

The diploma consists of theoretical part, which describes the process of designing a questionnaire, the meaning of expert evaluation and of both presented schemes. Followed by analytical part, where we used the Friedman test, with which we tried to determine, whether the experts are using each scheme in a similar way. The entire analytical part is concerning the research of qualitative and quantitative detection of errors that are perceived by experts in individual questions, using the Sweden check list and QACS scheme.

We conclude that we can not affirm which scheme is better for usage, since the data in the case of only two questionnaires provided adequate comparability of the two schemes. In the quality of expert schemes further research is required.

Key words: Comparability of expert schemes, Sweden Check list scheme, QACS scheme, expert evaluation, the Friedman test.

KAZALO

1 UVOD	7
2 TEORETIČNI DEL	9
2.1 SESTAVA ANKETNEGA VPRAŠALNIKA IN POMEN KAKOVOSTI PODATKOV	9
2.2 METODE TESTIRANJA ANKETNIH VPRAŠALNIKOV.....	11
2.3 EKSPERTNA EVALVACIJA – NAMEN IN UPORABA EVALVACIJE NA SPLOŠNO.....	13
2.4 EKSPERTNA EVALVACIJA Z UPORABO EKSPERTNIH SHEM	14
2.5 POZITIVNE IN NEGATIVNE LASTNOSTI EKSPERTNE EVALVACIJE.....	15
2.6 PRIMERJAVA EKSPERTNE EVALVACIJE Z OSTALIMI KVALITATIVNIMI METODAMI ZA TESTIRANJE ANKETNIH VPRAŠALNIKOV	15
3 CILJ IN NAMEN DIPLOMSKEGA DELA	18
3.1 OPIS OBRAVNAVANIH KODIRNIH SHEM EKSPERTNE EVALVACIJE.....	18
3.2 SHEMI QACS IN CHL	19
3.2.1 OPIS QACS KODIRNE SHEME	19
FAZE KOGNITIVNEGA PROCESA:	20
3.2.2 OPIS CHL SHEME	20
3.3 META-PODATKOVNA BAZA	22
3.4 UPORABLJENA METODA	22
3.5 ZASTAVLJENI HIPOTEZI	23
3.6 FRIEDMANOV TEST	24
4 ANALITIČNI DEL	25
4.1 VPRAŠALNIH DEJAVNIKI RAZVOJA KOHEZIVNOSTI IN SODELOVANJA V SOSESKAH IN VPRAŠALNIK AVTORITETNI ODNOSI V DRUŽINI	26
4.2 ANALIZA VPRAŠALNIKA DEJAVNIKI RAZVOJA KOHEZIVNOSTI IN SODELOVANJA V SOSESKAH IN VPRAŠALNIKA AVTORITETNI ODNOSI V DRUŽINI.....	27
4.2.1 KVANTITATIVNI VIDIK ANALIZE – ANALIZA PO SHEMAH	28
4.2.2 KVALITATIVNI VIDIK ANALIZE – ANALIZA PROBLEMATIČNIH VPRAŠANJ GLEDE NA OBE SHEMI	30
4.2.3 NAJBOLJ PROBLEMATIČNA VPRAŠANJA VPRAŠALNIKA DEJAVNIKI KOHEZIVNOSTI V SOSESKAH	30
4.2.4 NAJBOLJ PROBLEMATIČNA VPRAŠANJA VPRAŠALNIKA AVTORITETNI ODNOSI V DRUŽINI.....	32
4.3 SINTEZA TEMELJNIH UGOTOVITEV V OSTALIH VPRAŠALNIKIH.....	36
4.3.1 ANALIZA VPRAŠALNIKOV, KI NISO USTREZALI POSTAVLJENIM KRITERIJEM.....	36
5 ZAKLJUČEK	38
6 LITERATURA	43
7 PRILOGE	46
PRILOGA A: ŠVEDSKA SHEMA CHECK LIST	46
FORSYTH IN LESSER QACS SHEMA	48
KORAKI IZDELAVE IN PREDELAVE OSNOVNE META-PODATKOVNE BAZE.....	50
PRILOGA B: TABELE	51

KAZALO TABEL

Tabela 3.1: Podobnosti kategorij napak med shemama CHL in QACS	21
Tabela 4.1: Pregled povprečnih rangov po ekspertih za vprašalnika Dejavniki kohezivnosti v soseskah in Avtoritetni odnosi v družini	27
Tabela 4.2: Pregled vrednosti Friedmanovega testa za vprašalnika Dejavniki kohezivnosti v soseskah in Avtoritetni odnosi v družini	27
Tabela 4.3: Pregled števila napak med eksperti v vprašalnikih Dejavniki kohezivnosti v soseskah in Avtoritetni odnosi v družini	28
Tabela 4.4: Pregled vprašanj z največjim številom napak v vprašalnikih Dejavniki kohezivnosti v soseskah in Avtoritetni odnosi v družini.....	29
Tabela 4.5: Pregled vrednosti Friedmanovega testa za vprašalnike: Vprašalnik o socialnih omrežjih, Internetni dating in Nevladne organizacije	36
Tabela B.1: Najbolj problematična vprašanja in njihove napake po skupni oceni ekspertov glede na obe shemi po vprašalnikih Dejavniki kohezivnosti v soseskah, Avtoritetni odnosi v družini, Vprašalnik o socialnih omrežjih, Stališča slovenske mladine do vojaškega poklica in Nevladne organizacije	51
Tabela B.2: Dodatne kategorije izpostavljenih napak po shemi CHL in QACS v vprašalnikih Dejavniki kohezivnosti v soseskah, Avtoritetni odnosi v družini, Vprašalnik o socialnih omrežjih, Stališča slovenske mladine do vojaškega poklica in Nevladne organizacije.....	52

1 UVOD

V družboslovnem raziskovanju so za pridobitev kakovostnih podatkov ključnega pomena anketni vprašalniki. V procesu oblikovanja le-teh so pomembni številni koraki, preden jih uporabimo na terenu. Če želimo zagotoviti kar se da kakovostne rezultate, je priporočljivo, da v fazi oblikovanja anketnega vprašalnika in po tem uporabimo metodo ekspertne evalvacije z uporabo ekspertnih shem. Tako anketna vprašanja oblikujemo na anketirancu najrazumljivejši način, ali pa zaznamo tehnične in vsebinske napake v vprašanjih, ki jih v sami zasnovi nismo upoštevali.

V diplomskem delu obravnavamo rezultate uporabe ekspertnih shem Check list in QACS, ki so jih študenti v vlogi ekspertov uporabili na različnih anketnih vprašalnikih. Glavni namen naloge je ugotoviti, katera obravnavana ekspertna shema je lažja za uporabo, saj na tem področju namreč ni zaslediti tovrstnih raziskav, ki bi pripomogle h kakovostnejšemu zaznavanju napak v vprašanjih. Zanima nas, ali eksperti obravnavani shemi uporabljajo na različen način in z uporabo katere sheme so zaznali največ vsebinskih in strukturnih napak. Prav tako pa želimo ugotoviti, katere so tiste vsebinske napake, ki jih zaznamo z uporabo posamezne sheme. Rezultate ekspertnih evalvacij z obravnavanima shemama bomo primerjali s Friedmanovim testom in tako ugotavljali podobnost v načinu uporabe shem med eksperti.

Za uvodnim poglavjem sledi teoretski del, kjer opišemo pomen kakovosti anketnih podatkov, opredelimo metode testiranja anketnih vprašalnikov in se osredotočimo na fazo pretestiranja anketnih vprašalnikov, kjer temeljno vlogo igra tudi metoda ekspertne evalvacije. V nadaljevanju natančneje opredelimo njen namen in uporabo na splošno, nato pa se osredotočimo na ekspertno evalvacijo z uporabo kodirnih shem, ki je ključnega pomena za potrebe diplomske naloge. Opišemo pozitivne in negativne lastnosti te metode in jo nato primerjamo z ostalimi kvalitativnimi metodami za testiranje anketnih vprašalnikov. V 3. poglavju natančneje opredelimo cilje in namen diplomske naloge, opišemo značilnosti obravnavanih shem in osnovno meta-podatkovno bazo. V nadaljevanju poglavja sledi določitev uporabljene metode in zastavljeni hipotezi ter obravnava Friedmanovega testa. Glaven del diplomske naloge se nahaja v 4. poglavju, kjer se lotimo analize dveh vprašalnikov, ki ustrezata pogojem primerljivosti obravnavanih shem. Analiza je najprej izvedena iz kvantitativnega, nato pa še iz kvalitativnega vidika. V kvantitativnem delu je

najpomembnejši kazalec učinkovitejše sheme število ocenjenih napak ekspertov glede na določeno shemo, v kvalitativnem vidiku pa obravnavamo vsebino najbolj težavnih vprašanj, ki so jih eksperti z uporabo obeh shem izpostavili skupno. Za tem podamo sintezo temeljnih ugotovitev tudi v ostalih vprašalnikih. V 5. poglavju sledi sklep in povzetek glavnih ugotovitev analize oziroma zaključek.

Tema diplomske naloge je relevantna, ker področje testiranja anketnih vprašalnikov z uporabo metode ekspertne evalvacije še ni dovolj raziskano, prav tako pa do sedaj še ni bilo izvedenih podobnih raziskav. Zato v nadaljevanju poskušamo raziskati ali lahko z uporabo določene sheme oblikujemo boljša anketna vprašanja in s tem zagotovimo kakovostnejše rezultate raziskav, ki temeljijo na uporabi anketnih vprašalnikov.

2 TEORETIČNI DEL

2.1 Sestava anketnega vprašalnika in pomen kakovosti podatkov

Ker je vprašalnik merski instrument pri anketnih raziskavah, je njegova kakovost ključnega pomena, saj s tem pridobimo kakovostne anketne podatke in ugotovitve raziskave. Je največkrat uporabljeno orodje za zbiranje družboslovnih podatkov, saj omogoča, da v kratkem času zberemo ogromno količino podatkov (Nadrag 2002, 13). Kakovost podatkov je odvisna od velikosti in reprezentativnosti vzorca, od metod zbiranja podatkov, od kakovosti anketiranja in predvsem od tega, kako dobro vprašanja merijo pojav, ki ga želimo opisati (Fowler 1995, 150).

Najpogosteje je v uporabi standardiziran anketni vprašalnik, ki je za vse anketirance enak, ima enako obliko in popolnoma enaka vprašanja. To pomeni, da je potrebno natančno in premišljeno oblikovati vsebino vprašanja in jo ustrezno besedno oblikovati, da dobimo čim bolj veljavne odgovore (Payne v Nadrag 2002, 14).

Pojem kakovosti anketnega vprašalnika je definiran z dvema komponentama: s stopnjo razumevanja vprašanj in težavnostjo odgovarjanja ter s tem, kako dobro vprašanja merijo zelene teoretične koncepte (Groves in drugi 2004; Hlebec 2004). Z drugimi besedami, vprašalnik je kakovosten, »če anketiranci razumejo in interpretirajo vprašanja tako kot raziskovalec« (Hlebec v Nadrag 2002, 13).

Če se poglobimo v vsebinski pomen vprašalnika, moramo v vprašalniku najprej jasno definirati, kaj vprašanja merijo. Vprašanja morajo biti oblikovana tako, da sprašujejo po dejstvu, stališču, mnenju in vedenju, ali pa po demografskih značilnostih.

Težava, ki se pojavi pri vprašanjih o **dejstvih**, je v priklicu preteklih dogodkov, zato je primerno vprašanja, ki se nanašajo na pretekle dogodke, postaviti v časovni in prostorski okvir (Converse in Presser v Nadrag 2002, 15). Pri vprašanjih, ki merijo **stališča**, se soočamo s potencialno možnostjo, da anketiranec nima oblikovanega stališča o stvari o kateri ga sprašujemo (Moser in Kalton v Nadrag 2002, 15). Kadar sprašujemo po **mnenjih**, poskušamo dobljene odgovore oblikovati v skupno stališče (Nadrag 2002, 15). Pomembno je razlikovati med stališči in mnenji. Stališča vplivajo na to, kako bo posameznik določeno situacijo ali

objekt zaznal, ga doživljal, o njem mislil in se nanj čustveno odzval. Stališče je zasnovano na intelektualni, emocionalni in aktivnostni komponenti. Pojmu stališče pa je soroden tudi pojem mnenje. Mnenja so predvsem bolj specifična od stališč in bolj nestabilna oziroma lažje spremenljiva. Mnenja predstavljajo konkretizacijo oziroma manifestacijo stališč v konkretnih situacijah (Ule v Demšar Pečak, 2004). »Anketni vprašalniki ne merijo **vedenja** anketirancev posredno, temveč lahko dobimo le njihove odgovore na vprašanja, o njihovem vedenju v preteklosti in o bodočem, hipotetičnem vedenju« (Babbie v Nadrag 2002, 16).

Za vprašalnik je značilno, da ga lahko sestavljajo vprašanja **zaprtega**, **odprtega** ali **kombiniranega tipa**. Izbrati moramo tisto obliko, ki glede na naše cilje poda najboljše rezultate.

Nadrag (2002) omenja Stevensonovo delitev štirih merskih lestvic, ki jih lahko uporabimo v anketnem vprašalniku. To so nominalna, ordinalna, intervalna in razmernostna lestvica, ki pripomorejo k bolj učinkovitemu zbiranju podatkov.

Nekaj splošnih pravil, ki jih navaja Nadrag (2002) pri oblikovanju vprašanj:

- vprašanja naj bodo kratka,
- besede naj bodo enostavne, nedvoumne,
- uporaba preprostega jezika,
- brez dvojnih zanikanj,
- brez sugestivnih vprašanj,
- čim bolj jasna navodila,
- posebna previdnost pri oblikovanju občutljivih vprašanj,
- vključitev, izključitev odgovora ne vem v primeru, ko nekdo nima stališča,
- postaviti vprašanje z možnostjo nevtralnega odgovora.

Vsebinska zasnova vprašalnika ima pomembno vlogo pri pridobivanju kakovostnih informacij pri anketiranju. Poseben pomen ima tudi vizualna raven vprašalnika (oblika, zaporedje,

struktura vprašanj in odgovorov), saj se kakovost anketnega vprašalnika ne preverja samo na verbalni ravni (besedila vprašanj, navodil, nagovorov) (Hlebec 2004/2005). Sanchez (1992) ugotavlja, da format, grafični prikaz ter zaporedje vprašanj vplivajo na število dobljenih odgovorov.

Da je vprašalnik čim bolj učinkovit, je potrebno opraviti pretestiranje, ki pokaže, kje se nahajajo napake in pomanjkljivosti pri posameznih vprašanjih in odgovorih. V nadaljevanju so predstavljene različne metode, ki le-to ugotavljajo in pripomorejo k vsesplošni boljši kakovosti vprašalnika.

2.2 Metode testiranja anketnih vprašalnikov

Anketni vprašalnik ima tako kot vsak drug inštrument za merjenje raziskovalnih spremenljivk svoje pomanjkljivosti, zato so se na podlagi teh razvile metode testiranja anketnih vprašalnikov. Pomembno je testiranje, ki se ga lahko opravi pred začetkom pravega anketiranja (Nadrag 2002, 28). Poznamo več različnih razvrščenj glede na metode, ki se jih uporablja pri testiranju kakovosti anketnih vprašalnikov, in Snijkersov model (Snijkers 2002, 64), ki vključuje pet korakov, v katerih se te metode uporabljajo:

1. Definicija in izvedljivost raziskave: oblikuje se prototipe vprašalnikov in način zbiranja podatkov ter obdelave podatkov. Sledi pregled literature, raziskovalnih poročil, meta-analiz ter ekspertna evalvacija in raziskovanje fokusnih skupin.
2. Kvalitativni laboratorijski test: vprašalnik je pretestiran z namenom, da se preišče vse napake pri vprašanjih in odgovorih tako, da se izvede pregled izrazoslovja, zaporedja vprašanj, postopka in časa intervjuja. Izvede se ponovna ekspertna evalvacija skupaj z vključitvijo poglavljenih intervjujev in fokusnih skupin.
3. Kvalitativni operacionalizacijski test na terenu: vprašalnik in postopek zbiranja podatkov se preizkusi na terenu z namenom, da se ugotovi, kateri so glavni problemi pri raziskavi. Nadzoruje in evalvira se standardizirane intervjuje, pregleda se poročila fokusnih skupin in ekspertnih evalvacij kot tudi analiza števila neodgovorjenih vprašanj (non-response analysis).

4. Kvalitativna pilotna študija: testiranje zbiranja in procesiranja podatkov na terenu z namenom ugotovitve stroškov in koristi. Izid tega koraka je končna verzija vprašalnika in način zbiranja podatkov in njihove obdelave.
5. Implementacija: naredijo se vse priprave, da se lahko raziskava izvede na terenu.

Model je po Snijkersu (2002, 64) uporaben pri izboljšanju kakovosti anketnega vprašalnika, saj »raziskuje morebitne napake, ki bi se lahko pojavile, raziskuje razloge, zakaj so se pojavile in oblikuje rešitve za popravilo napak v vprašalnikih in postopkih zbiranja podatkov«.

V teh korakih se lahko uporabljajo tako kvantitativne kot kvalitativne metode, a so slednje zaradi manjših stroškov pogosteje uporabljene. Obstaja več različnih praks, trenutno najboljše prakse kvalitativnih metod pa so ekspertna evalvacija, fokusne skupine in kognitivni intervjuji (Snijkers 2002, 68; Hlebec 2004/2005). Te tehnike imajo različne prednosti in pomanjkljivosti, zaradi katerih so dragocene pri prepoznavanju težav pri anketnih vprašalnikih in so tudi uporabne na različnih nivojih razvoja vprašalnika (DeMaio in drugi 2006, 1).

Testiranje vprašalnika pred uradnim anketiranjem, z zgoraj omenjenimi kvalitativnimi metodami, imenujemo pretestiranje. Poznamo delitev na dve veliki področji: »pred-terenske« (*pre-field*) in terenske (*field*) tehnike (American Statistical Association – ASA 1997, 1; Scheuren 2004, 45; DeMaio in drugi 2006, 1). Pred-terenske tehnike so torej tiste, ki se navadno uporabljajo v predhodnih fazah oblikovanja vprašalnika.

Pretestiranje je kritično pri prepoznavanju težav vprašalnika (ASA 1997, 1; DeMaio in drugi 2006, 1), pri katerem se osredotočamo na vsebino vprašanja, kontekstualne posledice, preskok navodil ipd. (DeMaio in drugi 2006, 1). Kot primer, pri vsebini vprašanja lahko pride do težav, če vprašanje povzroča zmedo s celotnim pomenom vprašanja ali pa napačno razlaganje posameznih terminov ali konceptov. Cilj pretestiranja je usmerjen v »izboljšavo kvalitete podatkov preko izboljšave vprašalnika« (Snijkers 2003, 191).

V diplomskem delu smo se osredotočili na področje ekspertne evalvacije in ekspertnih shem, kjer smo skušali primerjati kakovost posameznih shem, ki so jih eksperti uporabili na

različnih vprašalnikov. Sledi opredelitev namena in uporabe ekspertne evalvacije na splošno ter ekspertne evalvacije z uporabo kodirnih shem.

2.3 Ekspertna evalvacija – namen in uporaba evalvacije na splošno

Ekspertna evalvacija je po Forsythu in drugih (1999, 145) »metoda za pregled in oceno vprašalnikov«. Primarni cilj evalvacije je prepoznavanje težav vprašanj in morebitne predelave le-teh. Kriteriji za ocenitev posameznega vprašanja so standardizirani, kar omogoča, da je evalvacija pri pregledu bolj konsistentna. Pri ekspertni evalvaciji gre za pregled tehničnega vidika in vsebine vprašalnika s strani skupine ekspertov.

Skupina ekspertov se deli na »eksperte za vsebino, ki preverjajo, ali vsebina (besedilo) vprašanj meri načrtovano teoretično razsežnost ali koncept ter eksperte za oblikovanje anketnega vprašalnika, ki preverjajo, ali posamezno anketno vprašanje in vprašalnik v celoti dosega standarde kakovosti« (Hlebec 2004/2005).

Pregleda se »besedilo vprašanja, strukturo vprašanja, dane odgovore, zaporedje vprašanj, navodila za izpolnjevanje in pravila za logično ureditev vprašanj« (Hlebec 2004/2005). Kot primer, ekspert lahko opazi, da je v določenem vprašanju uporabljena petstopenjska merska lestvica, ki pa ne bi smela biti uporabljena zaradi vsebine vprašanja. Ekspert nato predlaga dihotomni da/ne format odgovora, ki naj bi bil primernejši. Velja pravilo, da vprašalnik individualno pregledajo najmanj trije in največ deset ekspertov za vsebino, nato pa en sam ekspert zbere vse posamezne revizije in klasificira opažanja (Snijkers 2002, 69). Da bi ta ekspert prišel do konsistentnega zaključka rezultatov, se posvetuje z eksperti za vsebino. Zaključek naj bi vseboval tako možne napake kot predloge za revizijo (Snijkers 2002, 69).

V petstopenjskem modelu je po Snijkersu (2002, 68) namen ekspertne evalvacije pridobiti razumevanje o morebitnih problemih pri zasnovi vprašalnika oziroma o težavah, ki bi se lahko pojavile pri anketirancih ob razumevanju in odgovarjanju na vprašanja. Tako se ekspertna evalvacija po Snijkersovem modelu uporabi na prvi stopnji, kjer se pridobi mnenje ekspertov o zasnovi vprašalnika, na drugi stopnji, ki je stopnja predhodnega testiranja, ter v tretji stopnji, ki je po naravi operacionalna.

Opredelili smo torej ekspertno evalvacijo, kjer nastopajo eksperti za vsebino, v nadaljevanju pa obravnavamo ekspertno evalvacijo, kjer eksperti pri ocenjevanju anketnih vprašanj uporabljajo kodirne sheme.

2.4 Ekspertna evalvacija z uporabo ekspertnih shem

Evalvacija se s strani ekspertov navadno izvaja s pomočjo raznih kodirnih shem in z uporabo seznamov možnih napak (Snijkers 2002, 69; Hlebec 2004/2005). Primer takega seznama je seznam, oblikovan na Švedskem, imenovan »*Checklist*«. Forsyth in Lessler sta za ekspertno evalvacijo prav tako razvila kodirno shemo *Questionnaire Appraisal Coding System (QACS)*, katero sta kasneje Akkerboom in Dehue zaradi »prevelike mere podrobnosti in obširnosti za učinkovito uporabo v ocenjevanju napak v vprašanjih« (Snijkers 2002, 69) razširila v model, ki tudi klasificira napake v vprašanjih (Električna razvrstitev od merjenja napake do tveganja za ocenitev vprašalnika, 1997). Znana avtorja pa sta prav tako Saris in Gallhofer ter njuna klasifikacija anketnih vprašanj (Nadrag 2002, 23). Prav tako so tudi nekateri drugi avtorji razvili ali pa dopolnili podobne sheme, ki pripomorejo pri prepoznavi vprašanj z napakami. Takšne klasifikacije so po besedah Nadraga (2002, 25) »zelo dobrodošle, saj so v veliko pomoč pri ocenjevanju kakovosti vprašanj«.

Znana je še ena vrsta evalvacijskih ekspertov, pri katerih je pomembno, da se vprašalnik z določeno vsebino ocenjuje z eksperti, ki izhajajo točno iz tega določenega vsebinskega področja. Imenujemo jih eksperti za vsebino oz. »*subject matter experts*«. Ukvarjajo se lahko tudi z drugimi vsebinami, ki lahko pripadajo drugim kvalitativnim metodam testiranja anketnih vprašalnikov, načeloma pa se ukvarjajo z vsebino vprašalnika, terminologijo, pomembnostjo vprašalnih področij za raziskovalne cilje, simulacijo obnašanja anketirancev v raziskavi (Ramirez 2002, 3) ipd. Potreba po ekspertih za vsebino je neločljivo povezana z oblikovanjem raziskave (Ramirez 2002, 1). Ramirez navaja klasifikacijo področij, iz katerih izhajajo tovrstni eksperti: vladni uradniki iz nadzornih agencij, predmetni ali programski analitiki, univerzitetni učitelji, specializirani v predmetnem področju, predstavniki stroke, podatkovni analitiki in drugi. Očiten problem pa je, da obstaja večja mera subjektivnosti, saj imajo lahko eksperti iz določenega področja določene interese glede izidov raziskave bodisi zaradi določenih profitov, ki lahko izhajajo iz nekaterih rezultatov raziskave ali pa promoviranja določenih vrednot raziskovalcev, ki na tak način ustvarijo lažne rezultate raziskav.

Tako je v nalogi pomembna tudi opredelitev pozitivnih in negativnih lastnosti ekspertne evalvacije, ki sledi v nadaljevanju.

2.5 Pozitivne in negativne lastnosti ekspertne evalvacije

Ekspertna evalvacija ima lahko pozitivne in negativne lastnosti. Ta metoda je zanesljiva in pomembna za identifikacijo morebitnih težav vprašalnika, saj odkrije veliko napak (Snijkers 2002, 13; DeMaio in drugi 2006, 5). Njene prednosti so, da je zelo časovno učinkovita, saj morajo biti povratne informacije navadno prenesene v omejenem času ter je tudi stroškovno učinkovita (Forsyth in drugi 1999, 145; Snijkers 2002, 11). Nudi bolj sistematičen pristop v procesu ekspertne evalvacije (DeMaio in drugi 2006, 5) in služi kot vodič pri evalviranju. Rezultat uporabe seznama možnih napak ali drugih shem je ocena zahtevnosti ali napak posameznih vprašanj in vprašalnika v celoti (Hlebec 2004/2005).

Pomanjkljivost metode je dejstvo, da je količina zaznanih napak odvisna od števila ekspertov in njihovega predznanja in izkušenj. Prav tako pa metoda ni prepričljiva za potencialne stranke zaradi ne vključenosti anketirancev v metodo (Snijkers 2002, 71). Glavna negativna stran metode je torej v njeni subjektivni uporabi, ker je metoda sama odvisna od predznanja ekspertov oz. od področja dela, iz katerega izhajajo (DeMaio in drugi 2006, 13). Potrebno je še omeniti, da je ekspertna evalvacija najbolj učinkovita v kombinaciji z ostalimi metodami, ki jih opredeljujemo v nadaljevanju poglavja.

2.6 Primerjava ekspertne evalvacije z ostalimi kvalitativnimi metodami za testiranje anketnih vprašalnikov

Glavno vprašanje, ki se postavlja v povezavi z ekspertno evalvacijo in ostalimi pretestnimi metodami, je, ali so vprašalniki, ki gredo čez fazo pretestiranja, bolj veljavni in dajo boljše podatke kot vprašalniki, ki niso bili testirani. Raziskovalci so izvedli nekaj raziskav, s katerimi so pokušali ugotoviti, katere metode pretestiranja so najbolj uporabne pri zaznavi napak v vprašalniku. Ekspertno evalvacijo se je v večini zasledenih raziskav (Forsyth in drugi 1999; Forsyth in drugi 2004; Hak in drugi 2004; Forsyth in drugi 2003; Rothgeb in drugi 2007) primerjalo z metodo kognitivnega intervjuja. Različni avtorji so prišli do različnih ugotovitev. Hak in drugi (2004) ugotavljajo, da je kognitivni intervju veliko bolj produktiven kot ekspertna evalvacija, glede na to, da je ta metoda pokazala več različnih težav na več

področjih vprašalnika. Razlike so bile tudi v vsebini problema – s kognitivnim intervjujem so odkrili določene probleme, ki se jih ne bi z lahkoto opazilo v analizi vprašanj, saj so se težave navezovale na biografske posebnosti. Z ekspertno evalvacijo so v večini zaznali le »kognitivne probleme« (Hak in drugi 2004, 11). Nasprotno pa Praser in Blair v Rothgeb in drugi (2007) prihajata do drugačnih zaključkov. Ugotovila sta, da se z ekspertno evalvacijo odkrije največje število napak, s kognitivnim intervjujem pa najmanj. Ekspertna evalvacija pa je kljub svojim pomanjkljivostim učinkovita. Forsyth in drugi (1999) so ugotovili, da se z ekspertno evalvacijo v primerjavi s kognitivnim intervjujem prepozna več težav v vprašalniku. Ekspertna evalvacija nudi vpogled, ki ga s kognitivnim intervjujem ne bi mogli izvesti; na primer vpogled v dejavnike, kot so družbena zaželenost, ki vpliva na presojo in odzivnost anketiranca. »Družbena zaželenost odgovora lahko deluje na način, katerega se anketiranci ne zavedajo.« (Forsyth in drugi 2003, 9) Poleg tega lahko eksperti zaznajo še druga področja. Najpogostejša so nejasna terminologija, nedefinirana časovna obdobja, problemi presoje in spomina (Forsyth in drugi 2003, 5; Rothgeb in drugi 2007, 15). Nasprotno pa eksperti ne morejo oceniti, kateri sklopi vprašalnika so težki oz. naporni za anketiranca, ali pa dobiti konkretnih rešitev s strani anketirancev, tako kot je to mogoče s kognitivnim intervjujem (Forsyth in drugi 2003, 7). Rothgeb in drugi (2007, 5) so odkrili podobnost v rezultatih glede na število zaznanih napak obeh metod, kar nasprotuje ugotovitvam Praserja in Blaira v Rothgebu (2007). Ugotovili so, da sta ekspertna evalvacija in kognitivni intervju zelo pogosto uporabljena – »ekspertna evalvacija je bolj sistematična, a je manj naporna kot kognitivni intervju.« Njihovi rezultati so pokazali, da je bila metoda ekspertne evalvacije najbolj produktivna v razpoznavi potencialnih problemov vprašalnika, a postavlja se vprašanje, kako natančna je ta metoda, saj se je z njeno uporabo v raziskavi našlo težave skoraj v vsakem vprašanju.

Jasno je, da ima vsaka metoda potencialne prednosti in pomanjkljivosti – ugotavlja pa se, da sta metodi komplementarni, saj »dajeta podobne rezultate, a z različnimi poudarki« (Forsyth in drugi 2003, 1; Rothgeb in drugi 2007, 119). Najbolje je, da uporabimo prednosti obeh metod, da smo tako bolj učinkoviti, kot pa da se zanašamo samo na eno (Forsyth in drugi 1999). Obe metodi podata veliko podobnih zaključkov oz. ugotovitev. V nekaterih točkah se metodi ujemata - na primer, Forsyth in drugi (2004) so z ekspertno evalvacijo s pomočjo QRCS kodirne sheme prepoznali potencialne problematične izraze, ki bi zavedli anketirance. Nato so s kognitivnim intervjujem odkrili, da so imeli anketiranci dejanske probleme z istimi izrazi. Najprej je najbolje uporabiti ekspertno evalvacijo, potem pa na podlagi rezultatov še

kognitivni intervju. Kljub temu pa so nekateri avtorji mnenja, da je ekspertna evalvacija le »sistem, ki je razvit predvsem kot pomoč oblikovalcu vprašalnika, kot pa da je tehnika pretestiranja«, saj potrebuje dodatno presojo ekspertov, da bi bila lahko učinkovita v pretestni fazi (Rothgeb in drugi 2007, 19). Obstaja splošno strinjanje, da je le malo možnosti, da bi bila lahko ena sama metoda »upoštevana kot najboljša« (Willis 2004, 95). Različne metode bodo imele različno vlogo v celotnem procesu razvoja vprašalnika. Poleg tega pa Snijkers (2003) poudarja, da bi bilo med drugimi nujno potrebno veliko bolj raziskati učinkovitost različnih metod pretestiranja. Dejstvo je, da ne obstaja nek splošen sporazum oz. standardizacija, kje, kdaj in kako uporabiti metode pretestiranja, saj to področje ni dovolj dobro raziskano. Vseeno pa je pomembno poudariti ugotovitev Forsytha in drugih (2004, 541), da metode pretestiranja lahko napovejo dejanske probleme vprašalnika, kar pomeni, da se vprašanja, ki jih s pretestno fazo napovemo kot problematične, kot taka izkažejo tudi v dejanski raziskavi. Na podlagi le-tega pa lahko izkušeni raziskovalci naredijo veliko izboljšav in prispevajo h kakovosti podatkov.

3 CILJ IN NAMEN DIPLOMSKEGA DELA

V diplomski nalogi bomo ugotavljali primernost dveh shem za evalvacijo anketnih vprašalnikov. Primerjali smo Forsythovo in Lesslerjevo QACS shemo in švedsko shemo Checklist. Za primerjavo smo uporabili meta-podatkovno bazo, ki vsebuje rezultate ekspertnih evalvacij različnih vprašalnikov.

Vprašalniki, katerih napake so zabeležene v bazi, so bili ocenjeni pri predmetu Oblikovanje anketnega vprašalnika, nosilke izr. prof. dr. Valentine Hlebec. Študenti različnih letnikov so kot del njihove izpitne naloge nastopali v vlogah ekspertov, kjer so z uporabo različnih shem presojali kakovost anketnih vprašalnikov. V diplomski nalogi igrata najpomembnejšo vlogo vprašalnika *Dejavniki razvoja kohezivnosti in sodelovanja v sosestah* in *Avtoritetni odnosi v družini*, katerih značilnosti so podane v analitičnem delu naloge. Le-tam smo z neparametričnim testom, opisanim v nadaljevanju, preverjali (ne)prisotnost določene napake oziroma stopnjo strinjanja med posameznimi eksperti.

Glavni cilj diplomske naloge je ugotoviti, ali eksperti na različen način uporabljajo shemi oziroma katera shema je lažja za uporabo. Zanimalo nas je, katere napake v vprašanih odkrijemo z določeno shemo in s katero zaznamo največ vsebinskih ali strukturnih napak. Domnevamo, da je določena shema lažja za uporabo takrat, ko jo eksperti uporabljajo na podoben način. Na področju ekspertne evalvacije namreč ni zaslediti podobnih primerjav, zato bodo rezultati pripomogli k lažjemu in bolj kakovostnemu ocenjevanju in oblikovanju anketnih vprašanj.

3.1 Opis obravnavanih kodirnih shem ekspertne evalvacije

Pri ekspertni evalvaciji vprašalnikov obstaja večje število shem za uporabo, zato smo se iz tehničnih in vsebinskih razlogov odločili za primerjavo med Forsythovo in Lesslerjevo QACS shemo in švedsko shemo Checklist (v nadaljevanju imenovana CHL shema). Pri ocenjevanju različnih vprašalnikov so študenti v vlogi ekspertov obravnavani shemi največkrat uporabili. Posamezen študent je na primer ocenjeval vprašalnik z uporabo obeh shem hkrati. Hkratna uporaba shem je pripomogla k večjemu številu pridobljenih podatkov in omogočila primerjavo učinkovitosti med shemama. Ta primerjava se nanaša na to, ali določen ekspert z uporabo ene sheme odkrije več napak, z drugo shemo pa manj, oziroma ali odkrije podobne

napake z obema shemama. Ostale sheme, ki so jih nekateri študenti prav tako uporabili pri predmetu Oblikovanje anketnega vprašalnika, pa so bile uporabljene premalokrat, da bi jih lahko med seboj primerjali v različnih vprašalnikih.

3.2 Shemi QACS in CHL

V nadaljevanju se osredotočamo na vsebinski opis obravnavanih shem, zato bi že na začetku poudarili, da si shemi med seboj nista podobni oziroma primerljivi. Različni sta v samem temelju. QACS shema je shema, ki izpostavlja in opozarja na napake v vprašanjih in odgovorih skozi kognitivni proces, ki je pojasnjen v nadaljevanju. Gre za seznam možnih napak oz. kod. CHL shema pa je sestavljena iz splošnih navodil in služi kot pripomoček za oblikovanje anketnega vprašalnika. Za lažjo primerjavo strinjanja med eksperti in analizo, katera shema je učinkovitejša za uporabo, pa smo v nadaljevanju izločili le tiste kategorije napak v obeh shemah, ki vsebinsko opozarjajo na enake napake.

3.2.1 Opis QACS kodirne sheme

QACS (*Questionnaire Appraisal Coding System*) kodirno shemo so razvili in uporabili v ameriškem inštitutu Research Triangle. Namen sheme je ocenjevanje in popravljanje vprašalnikov ter njihovih vprašanj. Razvili so jo z namenom standardizacije ekspertnega ocenjevanja vprašalnikov (Forsyth in drugi 1999, 145) in kot orodje za pomoč pri karakteriziranju, ocenjevanju in identifikaciji načinov za izboljšavo vprašalnikov (Lessler in Forsyth 1996, 259). Narejena je na osnovi modela procesa odgovarjanja na anketna vprašanja, kar je pravzaprav posebnost te sheme. Osredotoča se na štiri faze kognitivnega procesa: razumevanje, priklic informacij, presoja in referenčni okvir odgovorov.

Eksperti za kodiranje uporabljajo sistem tako, da po vrstnem redu pregledujejo posamezna vprašanja in njihove značilnosti ter pri tem s pomočjo QACS sheme analizirajo, ali lahko določene značilnosti prispevajo k morebitnim napakam v odgovorih (Lessler in Forsyth 1996, 260). Težavnim vprašanjem pripisujejo kode, saj je shema sestavljena iz nabora kod, ki opisujejo karakteristike vprašanj.

Faze kognitivnega procesa:

1. Faza razumevanja

Faza razumevanja kategorizira omembe vredne vidike vprašanj ali njihovih navodil, ki so namenjena za boljše razumevanje vprašanja. Nekatere kode v tej fazi opisujejo referenčni okvir in obdobje za posamezno vprašanje ter možne probleme v strukturi vprašanj. Identificirajo komponente vprašanj, ki so ključne za zaključevanje naloge razumevanja.

2. Faza priklica informacij

Faza priklica informacij obstaja zaradi predpostavke, da vsako vprašanje zahteva nekaj spominskega procesiranja informacij. Kode upoštevajo zahtevne naloge priklica npr. dogodkov, postopek spomina (ali gre za prepoznavanje, posploševanje ipd.) in kontekst priklica.

3. Faza presoje

Faza presoje označuje načine, kako anketiranec iz spomina združuje različne informacije in jih oblikuje v podan odgovor. Omenjen integracijski proces lahko združuje informacije iz več različnih virov ali pa izbira le določene dele informacij. Ta proces tako podaja subjektivne ocene.

4. Faza referenčnega okvira odgovorov

Faza referenčnega okvira odgovorov opisuje način, kako naj bi se anketiranec odzval na vprašanje. Kode v tej fazi zajemajo meje in oblikovanje odgovorov ter navodila.

Cilj sheme je uporaba faz za opis kognitivnih zahtev vprašanja in prepoznavanje vprašanj, ki so lahko težavna za točno odgovarjanje. Na primer, vprašanje je lahko težavno, če vključuje neznane izraze ali pa zahteva podroben priklic informacij iz preteklosti.

3.2.2 Opis CHL sheme

Švedsko shemo so razvili na švedskem statističnem uradu – SCB Statistics Sweden. Shema služi kot pripomoček pri oblikovanju anketnih vprašalnikov. Upošteva tako teoretsko izhodišče kot vzorčenje ter obravnava anketo kot celoto. Shema se lahko uporablja pred in med oblikovanjem anketnega vprašalnika. Dosledno navaja posamezne korake pri

oblikovanju anketnega vprašalnika in navodila ter pravila za oblikovanje le-tega. Bolj se osredotoča na celotni vprašalnik, njegov izgled in na vsebino vprašanj, manjši poudarek pa je namenjen samim odgovorom. Shema ima tako nek širši namen. V dokumentu švedskega statističnega urada (Henningsson 2004), kjer je Checklist shema predstavljena, so navedeni točni koraki, v kateri fazi oblikovanja posameznih delov vprašalnika naj bo uporabljen določen korak sheme.

Shemi sta stroškovno ugodnejši metodi za prepoznavanje težav pri obliki ali zasnovi vprašanja, navodilih in formatih odgovorov, ki se jo lahko uporablja na katerikoli stopnji oblikovanja vprašalnika. Vseeno pa Lessler in Forsyth (1996) opozarjata, da bi bilo potrebno metodo bolje preučiti, preden se kodirne sheme lahko uporabljajo kot splošno orodje za analizo anketnih vprašanj. Izpostavljata, da bi bilo potrebno kategorije kod preoblikovati, preučiti odnose med različnimi kodami in izvesti več testiranj prenovljenih shem.

Kot omenjeno, shemi v samem temelju nista primerljivi med seboj, zato smo za lažjo primerjavo strinjanja ekspertov izločili le tiste kategorije napak pri vsaki shemi, ki merijo vsebinsko enake napake. Podane so v spodnji tabeli. Še enkrat je potrebno poudariti, da shema CHL služi kot orodje, ki ocenjevalcu narekuje splošna navodila oziroma korake bodisi že za samo oblikovanje vprašanj ali pa kasnejše preverjanje le-teh. QACS shema pa opredeljuje seznam napak v vprašanjih, na katere moramo biti pozorni¹.

Tabela 3.1: Podobnosti kategorij napak med shemama CHL in QACS

	CHL	QACS
NAVODILA	Enostavna, splošna, ne zavajajoča	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 trditev
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 dane je omejujoč, nevtralni odgovor <i>Ne vem</i>	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

¹ Shemi se nahajata v prilogi.

Podobnosti med shemama se torej nanašajo na težave v navodilih, v jeziku vprašanja (morebitna zahtevna slovnica, nerazumljivi ali dvoumni tehnični izrazi) in podrobnosti vprašanja (ali sprašuje le eno stvar), v izpostavljanju zavajajočih vprašanj, ujemanju med vprašanji in odgovori ter na razne morebitne težavnosti v podanih odgovorih. Sledi opis meta-podatkovne baze, v katero smo aplicirali obravnavani shemi.

3.3 Meta-podatkovna baza

Pri primerjavi študentov, ki so bili v vlogi ekspertov ter shem QACS in CHL, smo uporabili meta-podatkovno bazo, izdelano v SPSS². Osnovna meta-podatkovna baza je kompleksna, saj je sestavljena iz 2727 spremenljivk. Gre za štiri dimenzionalno bazo, katere ključne spremenljivke so:

- vprašalniki,
- zaporedna številka vprašanja znotraj posameznega vprašalnika,
- eksperti in
- sheme.

Baza je bila za potrebe diplomske naloge večkrat preoblikovana tako, da smo izločili podatke o ekspertih, ki so uporabili le shemi QACS in CHL. Iz nje smo izvozili tudi število napak posameznega eksperta v vsakem zaporednem vprašanju v vprašalnikih.³

3.4 Uporabljen metoda

Najpomembnejši korak je bil izveden po izgradnji meta-podatkovne baze in odločitvi, s katerimi shemami bomo primerjali strinjanje med eksperti. Ugotoviti smo želeli, ali obstaja strinjanje med eksperti glede načina uporabe shem CHL in QACS oziroma katera shema je lažja in učinkovitejša za uporabo. Domnevali smo, da je lažja za uporabo tista shema, kjer se eksperti bolj strinjajo med seboj.

Na podlagi tega smo iskali mere skladnosti koderjev (ang. *intercoder agreement*), s katerimi bi lahko ocenili, v kolikšni meri eksperti delujejo podobno. Te mere smo iskali med neparametričnimi testi, ker so vrednosti spremenljivk oz. napak v bazi podane v nominalni

² SPSS je računalniški statistični program, ki se uporablja za statistično analizo in obdelavo podatkov.

³ Podrobnejša predelava osnovne meta-podatkovne baze je po korakih opisana v prilogi.

obliki. To pomeni, da imamo dani samo dve možni vrednosti spremenljivk: prisotnost določene napake v vprašanju (kodirano z 1) in odsotnost napake v vprašanju (kodirano z 0). Imeli smo možnost izbiranja med različnimi neparametričnimi testi, ki se običajno uporabljajo za to. Zaradi strukturne zasnove meta-podatkovne baze se je izkazalo, da imamo težave s podatki, katerih preureditev bi bila prezapletena. Tako smo se odločili za Friedmanov test, saj najbolje meri (ne)strinjanje med eksperti na posameznem vprašalniku v našem primeru.

3.5 Zastavljeni hipotezi

Pri ugotavljanju stopnje strinjanja med eksperti z uporabo obeh shem in preverjanju, s katero shemo so bili eksperti učinkovitejši pri ocenjevanju vprašalnikov, smo si zastavili naslednji hipotezi. Preverjali smo jih s Friedmanovim testom, opisanim v nadaljevanju:

Ničelna hipoteza H_0 : Pri ponavljajočih meritvah povprečnih rangov ni razlik.

Osnovna hipoteza H_1 : Pri ponavljajočih meritvah povprečnih rangov so razlike.

Preverjamo torej, ali eksperti pri ocenjevanju anketnih vprašalnikov uporabljajo shemi na podoben način (H_0) oziroma ali ju uporabljajo na različen način (H_1).

Če se izkaže, da na podlagi Friedmanovega testa zavrnilo ničelno domnevo in potrdimo osnovno, potem lahko izvedemo tudi nadaljnjo analizo razlik med CHL in QACS shemo. Tako odgovorimo na enega izmed ključnih ciljev diplomske naloge: s katero shemo eksperti izpostavijo večjo količino napak. Iz kvantitativnega vidika je to pokazatelj učinkovitejše sheme. Obenem pa lahko preverimo tudi kvalitativni oziroma vsebinski vidik kakovosti sheme, pri čemer imamo v mislih podrobnejšo analizo najbolj problematičnih vprašanj. Tako lahko pretehtamo, katere so glavne napake v vprašanju in primerjamo, s katero shemo so eksperti zaznali večjo hibo v vprašanju. Tista shema, ki je zaznala glavno napako, je iz kvalitativnega vidika boljša za uporabo.

V nasprotnem primeru pa sprejetje ničelne domneve pomeni, da se eksperti pri uporabi shem strinjajo med seboj in ju uporabljajo na podoben način. Zato analiza napak, ki bi jih izpostavili z uporabo ene ali druge sheme na posameznem vprašanju, ni smiselna. Je pa to

dober pokazatelj, kdaj je določena shema za uporabo dovolj enostavna, da jo lahko eksperti razumejo in uporabljajo na enak način.

3.6 Friedmanov test

Friedmanov test je iznašel ameriški ekonomist Milton Friedman. Uporablja se v primeru, ko imamo več ponovitev meritev na enakih subjektih (Texassoft 2004). S testom smo merili število odkritih napak ekspertov, najprej na vprašalniku *Dejavniki razvoja kohezivnosti in sodelovanja v soseskah*, potem pa smo meritev ponovili še na vprašalniku *Avtoritetni odnosi v družini*. V ostalih vprašalnikih, ki so bili vključeni v analizo, pa je bil Friedmanov test zaradi neenakih pogojev za analizo nekoliko manj smiseln, vendar je vseeno doprinesel nekaj zanimivih ugotovitev.

Friedmanov test je primeren za uporabo na podatkih, ker baza vsebuje predvsem nominalne spremenljivke, ki se ne porazdeljujejo normalno. Za neparametrične teste pa je značilno, da ne predpostavljajo normalne porazdelitve podatkov oz. enakosti variance. 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 podatke rangira, 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. Pod takšnimi pogoji je bil izračun Friedmanovega testa mogoč.

Delovanje Friedmanovega testa je sledeče:

Test primerja tri ali več parov skupin. Če v našem primeru ponazorimo skupino, lahko za primer vzamemo vprašalnik *Dejavniki razvoja kohezivnosti in sodelovanja v soseskah*, ki ga je ocenjevalo po 5 ekspertov z uporabo sheme CHL kot QACS. Primerjali smo število napak teh ekspertov v posameznem zaporednem vprašanju. Ekspert torej predstavlja eno skupino. Friedmanov test nato vrednosti v vsaki vrsti rangira najprej od najmanjše do največje, vsaka vrsta pa je rangirana 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 0.05 tveganja ob zavrnitvi ničelne domneve, s katero trdimo, da eksperti pri ocenjevanju anketnih vprašalnikov uporabljajo shemi na podoben način. Tako smo lahko s 95 % gotovostjo ugotavljali, da obstaja razlika med vključenimi eksperti pri ocenjevanju vprašalnikov z obema shemama.

Omenili bi še, da se glavna pomanjkljivost Freidmanovega testa⁴ kaže v tem, da ne predpostavlja normalne porazdelitve in zato ni tako močan kot ANOVA test (Sekar in drugi 2009).

4 ANALITIČNI DEL

V prvem koraku obravnavamo dva vprašalnika, ki sta sestavni del analitičnega dela. Analiziramo vprašalnika, imenovana *Dejavniki kohezivnosti v soseskah* in *Avtoritetni odnosi v družini*. Na obeh vprašalnikih smo najprej izvedli analizo po ekspertih. Na podanih ocenah oziroma številu napak posameznega eksperta smo izvedli Friedmanov test glede na obe uporabljeni shemi in na podlagi vrednosti testa zavrnili ničelno hipotezo o enakem načinu uporabe shem med eksperti na obeh vprašalnikih. To nam je nato omogočilo nadaljnji kvantitativni način analiziranja rezultatov, kjer smo naredili pregled števila napak med eksperti pri obeh shemah na obeh vprašalnikih. Na podlagi teh ugotovitev nato sledi kvalitativni vidik analize, kjer izpostavimo najbolj problematična vprašanja glede na skupno oceno ekspertov z obema shemama in ugotavljamo ter primerjamo vsebinsko učinkovitost uporabe shem CHL in QACS.

Za tem sledi še sinteza temeljnih ugotovitev v ostalih vprašalnikih vključenih v analizo, vendar je potrebno te rezultate jemati z rezervo, saj ne zadostujejo pogojem, ki omogočajo veljavno primerjavo učinkovitosti uporabe ekspertnih shem. Ti pogoji so opisani v nadaljevanju v začetku analitičnega dela veljavnih vprašalnikov.

⁴Nazoren način vpeljave Friedmanovega testa na neparametrične podatke je prikazan v video povezavi avtorjev Lee in drugi (2008).

4.1 Vprašalnih Dejavniki razvoja kohezivnosti in sodelovanja v soseskah in vprašalnik Avtoritetni odnosi v družini

Vprašalnika *Dejavniki kohezivnosti v soseskah* in *Avtoritetni odnosi v družini* predstavljata najpomembnejši del diplomskega dela, saj edina izpolnjujeta pogoje za primerjavo shem in strinjanja med eksperti. Pogoji za primerjavo shem so izpolnjeni, če so obe shemi uporabljali vsi eksperti na vseh vprašanjih. V drugih vprašalnikih so namreč ocenjevali različni eksperti, ki niso uporabljali obeh shem in so ocenjevali le posamezne sklope vprašanj.

Vprašalnik *Dejavniki kohezivnosti v soseskah* je del raziskovalnega podoktorskega projekta, katerega vodja je bila Maša Filipovič Hrast, naročnik raziskave pa ARRS Javna agencija za raziskovalno dejavnost Republike Slovenije. Gre za tiskano tekstovno gradivo v slovenskem jeziku. Vprašalnik je bil uporabljen v obdobju od januarja 2007 pa do decembra 2008⁵. Sestavljen je 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 (E1, E2, E17, E18, E19, E20). 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.

Drugi vprašalnik *Avtoritetni odnosi v družini* je sestavni del raziskovalnega podoktorskega projekta z naslovom *Komunikacijski in avtoritetni odnosi med starši in mladimi*. Vodja projekta je bila Metka Kuhar, nosilna raziskovalna organizacija pa Univerza v Ljubljani, natančneje Fakulteta za družbene vede. Ta neusmerjena raziskava je trajala od januarja 2007 pa do decembra 2008⁶. Uporabljeni vprašalnik *Avtoritetni odnosi v družini* raziskuje značilnosti medosebne komunikacije in opisuje kontroverzne situacije. Za izpolnjevanje je bil namenjen izključno staršem. Sestavljen je 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.

⁵Za dodatne informacije in rezultate raziskave glej Filipovič Hrast (2008).

⁶ Za dodatne informacije in rezultate raziskave glej Kuhar (2008).

4.2 Analiza vprašalnika Dejavniki razvoja kohezivnosti in sodelovanja v soseskah in vprašalnika Avtoritetni odnosi v družini

Naš cilj je torej ugotoviti, ali v obravnavanih vprašalnikih eksperti različno uporabljajo shemi, oz. izhajajoč iz te ugotovitve, odkriti, z uporabo katere sheme so bili eksperti bolj učinkoviti pri zaznavanju napak. Najprej smo izvedli analizo po ekspertih glede na obe uporabljeni shemi. Kot omenjeno, smo si zastavili hipotezi, kjer ničelna trdi, da eksperti na enak način uporabljajo shemi pri ocenjevanju vprašanj, obratno pa osnovna hipoteza trdi, da jih uporabljajo na različen način.

Tabela 4.1: Pregled povprečnih rangov po ekspertih za vprašalnika Dejavniki kohezivnosti v soseskah in Avtoritetni odnosi v družini

Shema	Ekspert	Povprečni rang	Shema	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

Tabela 4.2: Pregled vrednosti Friedmanovega testa za vprašalnika Dejavniki kohezivnosti v soseskah in Avtoritetni odnosi v družini

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

Na podlagi prikazanih rezultatov v Tabeli 1, ki prikazuje porazdelitev povprečnih rangov, ključnih za zavrnitev ali sprejetje zastavljenih domnev, in v Tabeli 2, kjer so prikazane vrednosti Friedmanovega testa, ugotavljamo, da eksperti različno uporabljajo sheme. Stopnja značilnosti Friedmanovega testa pri obeh shemah v obeh vprašalnikih namreč zadovoljuje 0.05 kriterij tveganja ob zavrnitvi ničelne domneve (razlikovanje med eksperti glede načina

uporabe shem). Razlikovanje v uporabi med eksperti nam omogoča nadaljnjo analizo podatkov po shemah. Pri tem moramo poudariti, da večje nestrinjanje med eksperti pomeni, da obe 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 lažja za uporabo. V primeru obravnavanih vprašalnikov pa se pri obeh shemah izkaže nasprotno.

4.2.1 Kvantitativni vidik analize – analiza po shemah

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

Tabela 4.3: Pregled števila napak med eksperti v vprašalnikih Dejavniki kohezivnosti v sosestah in Avtoritetni odnosi v družini

DEJAVNIKI KOHEZIVNOSTI V SOSESTAH			AVTORITETNI ODNOSI V DRUŽINI		
Ekspert	CHL	QACS	Ekspert	CHL	QACS
E1E2	103	40	E2	33	41
E17	71	70	E11	14	8
E18	52	24	E12	3	1
E19	78	106	E13	13	8
E20	88	52	E14	15	18
<i>Skupaj</i>	392	292	<i>Skupaj</i>	78	76

Iz kvantitativnega vidika je bil v vprašalniku *Dejavniki kohezivnosti v sosestah* z uporabo QACS sheme pri ocenjevanju najbolj kritičen ekspert E19, z uporabo CHL sheme pa ekspert E1E2. S CHL shemo je posamezen ekspert zaznal več napak (skupno 392) v primerjavi s QACS shemo (skupno 292).

V vprašalniku *Avtoritetni odnosi v družini* pa je bil z uporabo obeh shem najbolj kritičen ekspert E2. Odkril je tudi največ napak. Tudi v tem vprašalniku so eksperti s shemo CHL skupno zaznali več napak (78) kot z uporabo sheme QACS (76).

Iz teh podatkov lahko sklepamo, da je shema CHL pri ocenjevanju vprašalnika *Dejavniki kohezivnosti v sosestah* in vprašalnika *Avtoritetni odnosi v družini* bolj učinkovita.

Poleg količine zaznanih napak 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 10 napak ali več. V vprašalniku *Avtoritetni odnosi v družini* pa je bil kriterij določen pri 5 vprašanjih ali več. Gre za določanje kriterijev problematičnih vprašanj glede na posamezen vprašalnik, saj so tematsko in po obsegu različni.

Tabela 4.4: Pregled vprašanj z največjim številom napak v vprašalnikih *Dejavniki kohezivnosti v soseskah* in *Avtoritetni odnosi v družini*

DEJAVNIKI KOHEZIVNOSTI V SOSESKAH			AVTORITETNI ODNOSI V DRUŽINI		
Shema	Vprašanje	Št. napak	Shema	Vprašanje	Št. napak
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 napak 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 shemo CHL in največ 8 napak s shemo QACS.

Zanimivo je, da so eksperti v obeh vprašalnikih z uporabo sheme CHL izpostavili povsem drugačna, najbolj problematična vprašanja, kot pa z uporabo sheme QACS. 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 shemo CHL imajo omenjena vprašanja večje število napak kot glede na QACS shemo. Bolj vsebinski vidik, ki so ga zaznali eksperti z uporabo obeh shem, pa je predstavljen v nadaljnji analizi pri izpostavljenih problematičnih vprašanjih.

4.2.2 Kvalitativni vidik analize – analiza problematičnih vprašanj glede na obe shemi

V analizi po ekspertih smo poudarili kvantitativni vidik, sedaj pa skušamo kvalitativno ovrednotiti ocene napak po shemah. Podrobnejša analiza najbolj problematičnih vprašanj v vprašalniku *Dejavniki kohezivnosti v soseskah* in *Avtoritetni odnosi v družini* po obeh shemah nam omogoča ugotoviti, katere kategorije napak, se pravi vsebinske in tehnične napake, so eksperti izpostavili z uporabo posamezne sheme. Le-to je zadnji korak, s katerim poskušamo odgovoriti na vprašanje, katera od obravnavanih shem je bolj uporabna pri ekspertni evalvaciji. Kriterij, s katerim presojava, katera shema je vsebinsko bolj uporabna za ocenjevanje vprašalnikov, določa naša subjektivna presoja. Najprej smo razmislili, katera napaka vprašanja je najbolj kritična, potem pa smo pregledali napake, ki so jih eksperti določili z uporabo obeh shem in nato še dodatne napake po posamezni shemi. Tista shema, s katero so eksperti opredelili glavni problem v vprašanju, je po naši oceni vsebinsko bolj primerna za uporabo pri ekspertni evalvaciji oziroma smo pri ocenjevanju vprašanj z njeno uporabo učinkovitejši, če preučujemo iz kvalitativnega vidika.

4.2.3 Najbolj problematična vprašanja vprašalnika *Dejavniki kohezivnosti v soseskah*

VPRAŠANJE 1 - NAPAKE PO OBEH SHEMAH

Ali bi bili pripravljeni posvetiti nekaj ur svojega časa tedensko za aktivnosti za izboljšanje svoje soseske?

- a. Da
- b. Ne

Po naši presoji je najbolj kočljiva napaka v vprašanju 1 prav neroden vrstni red besed v vprašanju. S tem označujemo zadnji del vprašanja: ... *za aktivnosti za izboljšanje svoje soseske*. Poleg tega pa bi bilo nujno potrebno natančneje definirati tudi izraz *aktivnosti*. Manjši problem v vprašanju pa se nanaša na omejujoč odgovor.

Najprej bi opisali napake, ki so jih eksperti skupno z uporabo obeh shem, torej tako CHL kot QACS, izpostavili v vprašanju 1. Eksperti menijo, da vprašanju manjkajo kategorije oz.

odgovori ne vsebujejo vseh vrednosti, manjka nevtralni odgovor *ne vem*. Prav tako je odgovor tipa da-ne za anketiranca omejujoč. Vse našete napake povzročijo manjšo kakovost dobljenih podatkov kot v primeru, ko teh napak ne bi bilo. Ocenili so tudi, da vprašanje poleg tega nima splošnih navodil oz. so navodila skrita. Ob tem pa še dodatno jezik ni enostaven, ker vsebuje nedefiniran oz. nejasen, dvoumen izraz. Dvoumen izraz so *aktivnosti za izboljšanje soseske*, kjer ni posebej definirano, katere dejavnosti spadajo pod aktivnosti ter *nekaj ur tedensko*, kjer ni točno definirano, koliko ur je sploh mišljeno.

Eksperti so torej z uporabo obeh shem potrdili, da je odgovor slabo zasnovan, da bi bilo potrebno definirati izraz *aktivnosti* in da manjkajo splošna navodila za odgovarjanje na vprašanje.

Glavni problem vprašanja, se pravi vrstni red postavljenih besed, so eksperti ocenili s shemo CHL. Poleg podanih skupnih kategorij napak obeh shem so prav z njo dodatno izpostavili, da vprašanje 1 ni jasno, saj so besede v vprašanju zelo nerodno postavljene in ni dodatno obrazložen pomen *aktivnosti*, ki bi izboljšale sosesko. S shemo QACS tega niso zaznali.

S shemo QACS so študenti v vlogi ekspertov izpostavili le še dejstvo, da ne prihaja do ujemanja med vprašanjem in odgovorom. Menimo, da bi bilo za boljšo skladnost med vprašanjem in odgovorom potrebno preoblikovati vprašanje ter v odgovoru podati na primer število ur prostega časa, ki bi ga bili anketiranci pripravljene posvetiti za izboljšanje svoje soseske.

Iz analize vprašanja 1 smo tako ugotovili, da so eksperti s CHL shemo bolje opredelili glavno napako v vprašanju kot s shemo QACS, saj je vrstni red in izbor besed v vprašanju nejasen.

VPRAŠANJE 2 - NAPAKE PO OBEH SHEMAH

Ali bi bili pripravljene nameniti manjši finančni prispevek za izboljšanje svoje soseske?

- a. Da
- b. Ne

Pri vprašanju 2 je po naši oceni najbolj kritična napaka to, da se v vprašanju nahaja nedefiniran izraz *manjši finančni prispevek*. Druga napaka pa se nanaša na omejujoč odgovor da-ne, s premalo vrednostmi.

Skupno so eksperti, prav tako kot mi, z uporabo obeh shem ugotovili, da vprašanju manjkajo kategorije oz. odgovori ne vsebujejo vseh vrednosti. Manjka nevtralni odgovor *ne vem*, prav tako je odgovor tipa da-ne omejujoč za anketiranca. To pomeni, da so podani odgovori v vprašanju preveč omejujoči in iz njih ne dobimo tako kakovostnih podatkov. Z obema shemama so zaznali še, da vprašanje poleg tega nima splošnih navodil oz. so navodila skrita. Slednje pomeni, da to lahko zbega anketiranca, ki potem narobe razume in tudi označi svoje odgovore.

Kot je razvidno, eksperti skupno z obema shemama niso zaznali tistega, kar je po naši presoji glavna napaka vprašanja 2, zato nas je zanimalo, s katero shemo so dodatno le-to opazili. Poleg skupnih kategorij napak obeh shem so eksperti z uporabo sheme CHL dodatno presodili, da vprašanje ni jasno. S tem je mišljeno, da ni jasno predstavljen pomen *manjšega finančnega prispevka* kot tudi pomen *izboljšanja soseske*.

Z uporabo sheme QACS pa so med napake klasificirali še nedefinirane, nejasne oz. dvoumne izraze ter da ne prihaja do ujemanja med vprašanjem in odgovorom. Kot nejasen oziroma dvoumen izraz lahko razumemo izraz *manjši finančni prispevek*. Na tem mestu bi bilo potrebno dodatno obrazložiti, koliko znaša ta manjši finančni prispevek in kako bi pripomogel k izboljšanju soseske.

Naš predlog za izboljšavo vprašanja 2: Ker vprašanje sprašuje po *finančnem prispevku*, bi lahko v odgovorih podali lestvico vrednosti v evrih. S tem bi pridobili manj omejujoče odgovore in posledično bolj izčrpne informacije.

Iz analize vprašanja 2 lahko sklepamo, da je uporaba QACS sheme iz vsebinskega vidika bolj smiselna, saj podrobneje definira največjo hibo v vprašanju. Ta se nanaša na nedefiniran in dvoumen izraz *manjši finančni prispevek*. Z uporabo CHL sheme pa so eksperti opozorili le na nejasnost v vprašanju, ki je širšega pomena, s katerega pa smo vseeno sklepali, da se nanaša na nedefiniran termin.

4.2.4 Najbolj problematična vprašanja vprašalnika Avtoritetni odnosi v družini

Na enak način se lotimo analize najbolj problematičnih vprašanj, ki so jih skupno izpostavili eksperti glede na obe shemi, tudi v vprašalniku *Avtoritetni odnosi v družini*. Najprej podamo

našo oceno najbolj problematične napake po naši presoji, potem pa izvedemo analizo najprej glede na skupno zaznane napake po obeh shemah, nato pa še dodatne kategorije napak. Le-te nam nakažejo, s katero shemo so bili eksperti pozornejši na glavno napako v vprašanju.

VPRAŠANJE 3 - NAPAKE PO OBEH SHEMAH

Sedaj pa si predstavljajte še zadnjo situacijo:

Otrok je zaradi neprimernega vedenja v šoli dobil/-a opomin (ali pa si predstavljajte, da je dobil/-a negativno oceno) in to doma pove.

Sedaj pa opišite še, kako bi se dogajanje odvijalo in končalo v idealnem primeru (kako si želite, da bi se dogajanje odvijalo).

Za lažje razumevanje težavnosti vprašanja je predstavljena tudi začetna trditev, na katero se nanaša vprašanje 3.

Po naši presoji se glavna napaka vprašanja 3 nanaša na to, da vprašanje vsebuje dve vprašanji oziroma ne sprašuje le po eni stvari. Najprej zahteva, da naj anketiranci opišejo, kako bi se dogajanje odvijalo in končalo (tudi tu je lahko problem: vprašanje zahteva, da anketiranec presodi dve stvari hkrati: odvijanje in končanje), potem pa v oklepaju zahteva, naj opišejo, kako bi si anketiranci želeli, da bi se odvijalo. To lahko zmede anketiranca. Ali naj odgovori, kako bi se odvijalo v idealnem primeru, ali pa, kako bi se odvijalo po njihovih željah, ki morda niso enake kot v idealnem primeru. Zatorej lahko ocenimo, da vprašanju bodisi manjkajo jasna navodila za način odgovarjanja ali pa bi bilo potrebno vprašanje ločiti na dva dela oziroma ga povsem preoblikovati.

Eksperti so z uporabo obeh shem ocenili prav to, da vprašanju manjkajo enostavna, splošna, nezavajajoča navodila ter da odgovori ne vsebujejo vseh možnosti oziroma obstajajo manjkajoče kategorije odgovorov. S tem so verjetno želeli povedati, da tip odprtega odgovora v podanem vprašanju ni primeren.

Težavo, da vprašanje ne sprašuje le ene stvari in da je le-to zavajajoče, so eksperti dodatno zaznali z uporabo CHL sheme in ne s QACS shemo. Z uporabo sheme QACS so izpostavili le še, da je v vprašanju nejasen cilj in nejasen problem vprašanja.

Tako se v tem primeru izkaže, da so bili eksperti z uporabo CHL sheme vsebinsko bolj učinkoviti, saj so vprašanje bolj kakovostno ocenili, ker so ugotovili, da vprašanje ne zahteva odgovora samo glede ene stvari. Z uporabo OACS sheme pa te vsebinske napake niso izpostavili.

VPRAŠANJE 4 - NAPAKE PO OBEH SHEMAH

Pred seboj imate 12 lestvic, na koncu katerih so pridevniki nasprotnega pomena. Vsaka lestvica predstavlja niz stopenj, pridevniki pa predstavljajo skrajnosti tega niza. Prosim vas, da pri vsaki lestvici obkrožite številko, ki po vašem mnenju najbolj ustreza vašemu trenutnemu odnosu z otrokom. Na primer, če vas odnos zelo spominja na toplo, obkrožite številko 3 blizu toplo, če vas bolj spominja na hladno obkrožite 2 ali 3 blizu hladno, če vas ne spominja niti na toplo, niti na hladno, pa označite 0.

a.	Topel						Hladen
	3	2	1	0	1	2	3
b.	Prijeten						Zoprn
	3	2	1	0	1	2	3
c.	Nežen						Grob
	3	2	1	0	1	2	3
d.	Hiter						Počasen
	3	2	1	0	1	2	3
e.	Svetel						Temen
	3	2	1	0	1	2	3
f.	Velik						Majhen
	3	2	1	0	1	2	3
g.	Lep						Grd
	3	2	1	0	1	2	3
h.	Sproščen						Napet
	3	2	1	0	1	2	3
i.	Globok						Plitev
	3	2	1	0	1	2	3

V vprašanju 4 se nahaja več napak, ki lahko povečajo verjetnost slabega odgovarjanja in pridobitev nekakovostnih podatkov. Eden izmed glavnih problemov v vprašanju se kaže v tem, da je vprašanje zelo nejasno zastavljeno tudi oblikovno in vsebuje nekatere tehnične termine (*niz stopenj, skrajnosti niza*), ki že lahko zmedejo anketiranca. Da anketiranec dojame bistvo vprašanja oziroma na kaj odgovarja (odnos z otrokom), mora vprašanje prebrati dvakrat, kar ga lahko dodatno utruje. Glavni problem vprašanja so torej navodila, ki so prekompleksna in predolga. Druga glavna napaka pa so podani odgovori, ki so problematično zastavljeni in jih je tudi preveliko.

Nekatere od glavnih napak so skupno z uporabo obeh shem ugotovili tudi eksperti. Skupno so zaznali, da odgovori ne vsebujejo vseh možnosti oz. imajo manjkajoče kategorije. S tem so verjetno želeli poudariti, da na nekatere pridevnike določen anketiranec preprosto ni zmožen odgovoriti. Vprašanje po mnenju ekspertov ni jasno oz. ni jasen njegov cilj. Skupaj so še ugotovili, da vprašanje ni postavljeno v logično in vsebinsko povezan sklop oz. da ne prihaja do ujemanja med vprašanjem in odgovori. Prav tako jezik vprašanja ni enostaven, ker vsebuje nedefiniran oz. nejasen izraz. Tukaj so verjetno merili na tehnični izraz *niz stopenj*.

Glavni problem vprašanja so dodatno odkrili tudi z uporabo sheme CHL. Poleg skupnih kategorij napak obeh shem so ocenili, da ima vprašanje 4 preveliko število odgovorov, saj zajema 9 različnih kategorij, ki glede na obseg lahko utrudijo anketiranca, ki zaradi tega podaja slabše odgovore. Z uporabo sheme QACS pa so dodatno izpostavili še, da so v vprašanju skrita oz. nenatančna navodila.

Eksperti so torej z uporabo obeh shem skupno izpostavili glavne probleme v vprašanju po naši presoji (nejasnost vprašanja in problematičnost podanih odgovorov), vendar pa so le s shemo QACS izpostavili dodatno napako, ki se nanaša na navodila. Tako lahko zaključimo, da je bila z uporabo QACS sheme odkrita še ena glavna pomanjkljivost v vprašanju, kar nam sporoča, da je s kvalitativnega vidika v tem primeru QACS shema nekoliko bolj temeljita v primerjavi s CHL shemo.

Zanimalo nas je torej, z uporabo katere sheme smo pri ocenjevanju vprašanj lahko učinkovitejši. Glede na rezultate analize smo ugotovili, da se je pri odkrivanju večjega števila napak v vprašanjih za bolj učinkovito izkazala uporaba CHL sheme. Ko pa smo se poglobili v vsebinski del analize, se je izkazalo, da nobena od shem po svoji kakovosti ne izstopa. Potemtakem težko trdimo, da je katera od shem bistveno bolj učinkovitejša od druge in da eksperti z njeno uporabo boljše zaznajo problematična vprašanja. Če bi bila tudi v kvalitativnem delu analize boljša CHL shema, bi lahko predvidevali, da je uporaba te sheme bolj primerna pri ekspertni evalvaciji v primerjavi s QACS shemo, a ni tako. Na zastavljeno vprašanje, katera shema je učinkovitejša tako iz kvalitativnega in kvantitativnega vidika hkrati, bi bilo mogoče odgovoriti le v primeru, ko bi imeli na voljo večje število anketnih vprašalnikov, ki bi ustrezali pogojem za primerjavo ekspertnih shem.

4.3 Sinteza temeljnih ugotovitev v ostalih vprašalnikih

V analitičnem delu diplomskega dela smo obravnavali še ostale 3 vprašalnike, ki pa niso ustrezali pogojem za primerjavo strinjanja ekspertov oz. pogojem za primerjavo shem. V vseh ostalih vprašalnikih so sodelovali različni eksperti v neenakem številu, pri čemer so nekateri ocenjevali le določene sklope vprašalnika in ne celotnega vprašalnika. Eksperti so uporabili le eno izmed shem, tako da primerjava med obema že v osnovi ni bila možna.

Če imamo v zaključnih ugotovitvah v mislih te navedbe, pa je analiza ostalih vprašalnikov vseeno relevantna, ker rezultati poudarjajo drugačna spoznanja kot pri obravnavi vprašalnika *Dejavniki kohezivnosti v soseskah* in vprašalnika *Avtoritetni odnosi v družini*, ki povsem ustrezata zastavljenim pogojem za primerjavo. V tem delu torej poudarjamo, da je potrebno sledeče rezultate jemati z rezervo.

4.3.1 Analiza vprašalnikov, ki niso ustrezali postavljenim kriterijem

Tabela 4.5: Pregled vrednosti Friedmanovega testa za vprašalnike: Vprašalnik o socialnih omrežjih, Internetni dating in Nevladne organizacije

Vrednosti Friedmanovega testa				
Vprašalnik o socialnih omrežjih SKLOP 1-20				
Shema	Št. enot	Hi-kvadrat	Prostostne stopnje	Stopnja značilnosti
CHL	20	13.139	4	0.011
QACS	20	3.106	2	0.212
Internetni dating				
Shema	Št. enot	Hi-kvadrat	Prostostne stopnje	Stopnja značilnosti
CHL	28	17.033	2	0.000
QACS	28	1.178	2	0.000
Nevladne organizacije				
Shema	Št. enot	Hi-kvadrat	Prostostne stopnje	Stopnja značilnosti
CHL	134	28.810	1	0.000
QACS	92	0.000	1	1.000

En izmed poglobitvenih nasprotnih rezultatov analize v primerjavi s prej obravnavanimi vprašalnikoma je statistično neznačilna vrednost Friedmanovega testa. To pomeni, da gre pri nekaterih vprašalnikih za strinjanje med eksperti pri uporabi določene sheme, medtem ko je bilo pri vprašalniku *Dejavniki kohezivnosti v soseskah* in *Avtoritetni odnosi v družini* v ospredju nestrinjanje, ki je omogočilo nadaljnjo kvantitativno in kvalitativno analizo.

Pri testiranju s Friedmanovim testom je prišlo do statistične neznačilnosti oz. strinjanja med eksperti v treh vprašalnikih: *Vprašalnik o socialnih omrežjih*, *Internetni dating* in *Nevladne organizacije*. Zanimiva je ugotovitev, da je bila samo z uporabo sheme QACS vrednost Friedmanovega testa statistično neznačilna. Torej je med eksperti pri uporabi QACS sheme prišlo do podobnega načina uporabe sheme pri ocenjevanju vprašanj v vseh treh vprašalnikih.

V *vprašalniku o socialnih omrežjih* smo dobili takšen rezultat v prvem delu vprašalnika, ki so ga ocenjevali le trije eksperti, medtem ko ga je s shemo CHL ocenjevalo pet ekspertov. V vprašalniku *Nevladne organizacije* in *Internetni dating* pa je bila vrednost Friedmanovega testa statistično neznačilna na celotnem vprašalniku. Pomemben je torej podatek, da so vse tri vprašalnike s QACS shemo ocenjevali le po trije eksperti.

Iz tega lahko sklepamo, da obstaja večja možnost, da se bodo eksperti z uporabo QACS sheme bolj strinjali med seboj oz. bodo shemo uporabljali na enak način kot pa z uporabo sheme CHL, pri kateri v vseh obravnavanih vprašalnikih ne glede na število ekspertov ni prišlo do tovrstnega strinjanja. V primerjavi s QACS shemo so eksperti dojeli uporabo sheme CHL različno in z njo tudi različno ocenjevali ista vprašanja. Poudarili bi, da ugotovitev, da so razlike med eksperti v teh vprašalnikih statistično neznačilne, vsekakor ni slabost. To pravzaprav pomeni, da so eksperti delali na podoben način pri uporabi določene sheme. Ob tem bi se vendarle lahko vprašali, ali gre morda za prepisovanje nalog med študenti v vlogi ekspertov ali naključno strinjanje med temi eksperti.

5 ZAKLJUČEK

Najprej bi povzeli nekatere ključne elemente analitičnega dela diplomske naloge. Eden izmed ciljev naloge je bil v analitičnem delu ugotoviti, kakšna je stopnja (ne)strinjanja med eksperti, pri čemer je bil uporabljen Friedmanov test. Ugotovili smo, da je le-ta dovolj velika, da smo lahko izvedli primerjavo ekspertnih evalvacij z uporabo obravnavanih shem na vprašalnikih *Dejavniki kohezivnosti v soseskah* in *Avtoritetni odnosi v družini*. Prav tako sta samo omenjena vprašalnika ustrezala pogojem za veljavno analizo strinjanja ekspertov, ostali vprašalniki pa tega pogoja niso dosegli. Na tem mestu bi omenili, da je prav to pomanjkljivost diplomske naloge. Kljub temu smo s Friedmanovim testom testirali hipoteze tudi na ostalih vprašalnikih, kjer pa so vrednosti Friedmanovega testa kazale na strinjanje med eksperti (torej statistično neznačilno stopnjo). Le-to nam je onemogočilo nadaljnjo analizo po shemah tako iz kvantitativnega kot kvalitativnega vidika.

Kot ena izmed prvih ugotovitev, ki bi jih izpostavili v sklepnem delu diplomske naloge, je ta, da se je statistična neznačilnost oziroma strinjanje v uporabi določene sheme izkazala na treh vprašalnikih. Pri tem pa je predvsem zanimivo dejstvo, da je prišlo do strinjanja le pri uporabi sheme QACS. To nam sporoča, da je shema QACS v primerjavi s shemo CHL lažja za uporabo, saj so jo eksperti uporabili na podoben način. Ob tem pa moramo vseeno poudariti, da so v teh primerih s shemo QACS ocenjevali ista vprašanja le po trije eksperti. Tudi zaradi tega je lahko prišlo do strinjanja v uporabi. Ne smemo pa izključiti naključja ali pa tudi možnosti prepisovanja nalog med študenti.

Z vprašalnikom *Dejavniki razvoja kohezivnosti in sodelovanja v soseskah* in vprašalnikom *Avtoritetni odnosi v družini* smo v ključnem delu analize opravili kvantitativno in kvalitativno analizo vprašalnikov z uporabo CHL in QACS sheme. Glavni cilj diplomskega dela je bil ugotoviti, s pomočjo katere sheme eksperti najbolj kakovostno ocenjujejo vprašalnik, ki bo kot merski inštrument na učinkovit način pridobil podatke, kateri so ključnega pomena za raziskovanje. 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. Kot omenjeno, se shemi v samem temelju 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, te pa so opisane v teoretskem delu diplomske naloge.

Najprej smo se v obeh vprašalnikih, ki sta zadovoljevala pogoje za analizo, osredotočili na kvantitativen vidik, kjer smo ugotavljali, s pomočjo katere sheme dobimo večje število napak v vprašanjih. Izkazalo se je, da so eksperti z uporabo CHL sheme zasledili znatno več napak kot z uporabo QACS sheme. Na primer, v vprašalniku *Dejavniki kohezivnosti v sosestah* so eksperti z uporabo sheme CHL skupno zabeležili 390 napak, s shemo QACS pa 292 napak. Tovrstni rezultati nazorno odražajo večjo učinkovitost uporabe sheme CHL, če izhajamo iz predpostavke, da večja količina odkritih napak predstavlja bolj kakovostno ekspertno evalvacijo. Vzroki za takšno ugotovitev bi lahko tičali prav v postopku iskanja podobnosti med shemama, saj je bila primerljivost med shemama problematična. Če bi v analizo lahko vključili vse kategorije napak pri obeh shemah in ne bi opravili tega vmesnega koraka, bi bili rezultati lahko bistveno drugačni.

Glede na dejstvo, da v ocenjevanju z obema shemama nastopajo isti eksperti (iste osebe), 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. Katera vprašanja so eksperti z uporabo shem osvetlili v določenem vprašalniku, je moč videti v Tabeli B.1 v prilogi.

Ko smo se v nadaljevanju osredotočili na vsebinsko analizo, kjer je bil poudarek na kvalitativni perspektivi iskanja napak, pa se je izkazalo, da z uporabo obeh shem odkrijemo enako količino vsebinskih napak. Pri analizi vsebine smo namreč ugotovili, da je bila od vseh problematičnih vprašanj pri dveh (3 in 1) vsebinsko boljša CHL shema, pri drugih dveh (2 in 4) vprašanjih pa QACS shema. Razlog, zakaj je tako, velja pripisati zasnovi QACS sheme, ki ima bolj podrobno definirane posamezne napake v primerjavi s CHL shemo, ki na bolj

splošen način opredeljuje napako. Zaradi različnosti med shemama moramo upoštevati, da eksperti nekatere možne napake tudi spregledajo oziroma jih z njeno uporabo ne morejo zaznati.

Zgoraj smo podali relevantne rezultate analize, s katerimi smo odgovorili na zastavljena vprašanja in cilje naloge. Sedaj bi podali še nekaj zanimivih rezultatov celotne analize, torej vključno z vprašalniki, ki niso zadovoljevali veljavnih pogojev. Pri tem ponovno opozarjamo, da je ugotovitve, ki sledijo, potrebno jemati z rezervo.

Če v analizo vzamemo vse vprašalnike in si ogledamo skupne napake samo najbolj problematičnih vprašanj v vsakem vprašalniku, ugotovimo, da so eksperti ocenili največ 12 napak⁷ in najmanj 5⁸, skupno pa 37 napak. Te rezultate prikazuje spodnja tabela.

Tabela 5.1: Skupno število napak po obeh shemah in dodatno število napak glede na posamezno shemo v najbolj problematičnih vprašanjih v vprašalnikih Dejavniki kohezivnosti v soseskah, Avtoritetni odnosi v družini, Vprašalnik o socialnih omrežjih, Stališča slovenske mladine do vojaškega poklica in Nevladne organizacije

VPRAŠALNIK	Skupno št. napak, ki so jih eksperti zaznali tako s shemo CHL kot QACS v najbolj problematičnih vprašanjih	Dodatno št. napak po shemi CHL	Dodatno št. napak po shemi QACS	<i>Skupaj</i>
Dejavniki kohezivnosti v soseskah	5	2	4	11
Avtoritetni odnosi v družini	5	3	6	14
Vprašalnik o socialnih omrežjih	8	7	7	22
Stališča slovenske mladine do vojaškega poklica	7	4	2	47
Nevladne organizacije	12	5	8	25
<i>Skupaj</i>	37	21	27	85

Eksperti so z vsako shemo poudarili še dodatne kategorije napak, ki jih je moč videti v Tabeli B.2 v prilogi. S shemo QACS so eksperti glede na vse obravnavane vprašalnike našli največ napak (27), a vseeno ne moremo trditi, da je iz kvantitativnega vidika uporaba sheme QACS 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.

⁷ V vprašalniku Nevladne organizacije.

⁸ V vprašalniku Dejavniki kohezivnosti v soseskah in Avtoritetni odnosi v družini.

Ko smo obravnavali vse vprašalnike (ne da bi se ozirali na pogoje analize), smo ugotovili ključno razliko v uporabi shem CHL in QACS. Eksperti so s shemo CHL v primerjavi s shemo QACS dodatno ocenjevali tudi zasnovo 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.

S shemo QACS je bilo največkrat poleg že omenjenih kategorij ocenjeno naslednje: 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 za obkrožitev enega, več vprašanj v enem, neprimerno zaporedje odgovorov, nasprotujoča si navodila, nenatančna navodila, prekrivajoče se kategorije, vprašanje vsebuje zahtevno slovnico. S shemo Check list pa so eksperti dodatno ocenili naslednje kategorije napak: nejasno vprašanje, preveliko število odgovorov, ne sprašuje le eno stvar, zavajajoče vprašanje, neenostavna 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. Gre v bistvu za podobne kategorije napak, le da so eksperti določeno napako pri istem vprašanju ocenili le z eno shemo.

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. Težava je v pomanjkljivo zastavljeni osnovni bazi podatkov. Za boljšo analizo bi potrebovali več vprašalnikov, ki bi ustrezali pogojem primerjave omenjenih shem. Ker pa je bilo mogoče analizo opravljati na samo dveh vprašalnikih, so tudi dobljeni rezultati temu primerni. Takih rezultatov ni mogoče posploševati na vse primere.

Pomanjkljivost te naloge je, da v meta-podatkovni bazi nimamo enakega števila ekspertov, ki bi uporabili obe shemi in v vseh vprašalnikih. Potemtakem je za nadaljnje raziskovanje kakovosti posameznih shem že na začetku potrebno omogočiti skladnost števila ekspertov pri

uporabi obeh shem. Poglavitno pa je tudi dejstvo, da morajo ocenjevati isti sklop vprašalnika. Le-to omogoča primerljivost med podatki.

Kot šibkost naloge prav tako velja upoštevati, da je bilo večino sodelujočih ekspertov neizkušenih, z malo predznanja, saj so vprašalnik ocenjevali študenti 2. letnikov, ki se poprej niso srečali z metodo ekspertne evalvacije. Tako moramo v ocenah upoštevati tudi ekspertno subjektivnost, ki pa je prisotna tudi pri izkušenih ekspertih. To pa je tudi ključna kritika metode ekspertne evalvacije.

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 nivojih. Na podlagi analize (ne)strinjanja ekspertov v ocenjevanju s švedsko Check list 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 ali pa celotnega vprašalnika. Uporaba shem v fazi testiranja anketnih vprašalnikov je torej ključna, če želimo 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.

6 LITERATURA

ASA - American Statistical Association. 1997. *What is a survey? How to conduct pretesting?*
Dostopno prek: http://av.fdvinfo.net/literatura/ASA_pretesting.pdf (4. april 2008).

DeMaio, Theresa J., Nancy Bates, Diane Willimack in Jane Ingold. 2006. *Pretesting Questionnaires and Related Materials for Surveys and Censuses*. Dostopno prek: <http://www.census.gov/srd/pretest-standards.pdf> (23. avgust 2010).

Demšar Pečak, Nataša. 2004. Socialni marketing - dejavnik družbenih sprememb. *Socialna pedagogika* 8 (1): 29–64.

Filipovič Hrast, Maša. 2008. *Raziskovalni projekt. Dejavniki razvoja kohezivnosti razvoja in sodelovanja v soseskah*. Dostopno prek: <http://www.dlib.si/v2/Details.aspx?URN=URN:NBN:SI:doc-WSRWZ4MK> (2. september 2010).

Forsyth, Barbara, Kerry Levin in Sylvia K. Fisher. 1999. *Test of an appraisal method for establishment survey questionnaires*. Dostopno prek: <http://av.fdvinfo.net/uploadi/editor/1205231197forsyth%20appraisal.pdf> (22. avgust 2010).

Forsyth, Barbara, Elisa S. Weiss in Rebecca Miller Anderson. 2003. A Comparison of Appraisal and Cognitive Interview Methods for Testing Organizational Survey Questionnaires. *Bulletin de Methodologie Sociologique* 55: 67–92. Dostopno prek: http://www.jpsm.umd.edu/qdet/final_pdf_papers/round%20two/Forsyth.pdf (23. avgust 2010).

Forsyth, Barbara, Jennifer M. Rothgeb in Gordon B. Willis. 2004. Does Pretesting Make a Difference? V *Questionnaire Development Evaluation and Testing Methods*, ur. Stanley Presser, Mick P. Couper, Judith T. Lessler, Elizabeth Martin, Jean Martin, Eleanor Singer, 525–546. New York: John Wiley.

Fowler, Floyd J. 1995. *Improving survey questions: design and evaluation*. California: Sage Publications.

Hak, Tony, Kees van der Veer in Harrie Jansen. 2004. *The Three-Step Test-Interview (Tsti): An Observational Instrument for Pretesting Self-Completion Questionnaires*. Dostopno prek: http://209.85.135.104/search?q=cache:LK48a4sPIUJ:www.jpsm.umd.edu/qdet/final_pdf_papers/QDET-vdVeer+%2520et%2520al.pdf+hak+in+jansen+1997+expert+review&hl=sl&ct=clnk&cd=5&gl=s (23. avgust 2010).

Henningsson, Birgit. 2004. *Design your questions right : How to develop, test, evaluate and improve questionnaires*. Dostopno prek: http://www.scb.se/statistik/_publikationer/OV9999_2004A01_BR_X97OP0402.pdf (23. avgust 2010).

Hlebec, Valentina. 2004/2005. *Študijsko gradivo za predmet Oblikovanje anketnega vprašalnika. 1. del. Kakovost anketnega vprašalnika*. Ljubljana: FDV. Dostopno prek: <http://av.fdvinfo.net/content.php?p1=1&p2=155&p3=257&id=257> (5. april 2008).

Kuhar, Metka. 2008. *Zaključno poročilo o rezultatih raziskovalnega projekta*. Dostopno prek: <http://www.dlib.si/v2/StreamFile.aspx?URN=URN:NBN:SI:doc-X89L6S0E&id=bd6de286-27f4-4b33-8e82-7012b8d6b061&type=PDF> (2. september 2010).

Lee, Carl, Felix Famoye in Barbara Shelden. 2008. *Spss Training Workshop*. Dostopno prek: http://www.cst.cmich.edu/users/lee1c/spss/V16_materials/Video_Clips_v16/27NP_related/27NP_related.swf (2. september 2010).

Lessler, T. Judith in Barbara H. Forsyth. 1996. A Coding System for Appraising Questionnaires. V *Answering Questions: Methodology for Determining Cognitive and Communicative Processes in Survey Research*, ur. Norbert Schwarz in Seymour Sudman, 259–292. San Francisco: Jossey-Bass.

Nadrag, Petra. 2002. *Kakovost anketnega vprašalnika za evalvacijo programa PUM*. Diplomsko delo. Ljubljana: FDV.

Palan, Ronen. 1999. Global Governance and Social Closure or Who is to Governed in an Era of Global Governance? V *Approaches to Global Governance Theory*, ur. Martin Hewson in Timothy J. Sinclair, 55–72. Albany: State University of New York Press.

Ramirez, Carl. 2002. *Strategies for Subject Matter Expert Review in Questionnaire Design*. Dostopno prek: http://www.jpsm.umd.edu/qdet/final_pdf_papers/ramirez.pdf (22. avgust 2010).

Rothgeb, M. Jennifer, Gordon Willis in Barbara Forsyth. 2007. Questionnaire pretesting methods: do different techniques and different organizations produce similar results? *Bulletin de Methodologie Sociologique* 96: 5–31. Dostopno prek: http://av.fdvinfo.net/uploadi/editor/1203456132questionnaire_pretesting_methods.pdf (10. april 2008).

Sanchez, Maria Elena. 1992. Effects of questionnaire design on the quality of survey data. *Public opinion quarterly* 56 (2): 206–217.

Saris, E. Willem in Irmtaud N. Gallhofer. 2007. *Design, Evaluation and Analysis of Questionnaires For Survey Research*. New Jersey: Wiley.

Scheuren, Fritz. 2004. *What is a survey?* Dostopno prek: <http://www.amstat.org/sections/srms/pamphlet.pdf> (23. avgust 2010).

Sekar, Kirthiga, Grace Nivetha, Krish R. Rasika, Dhevi S. Swappna in Mallika Umamaheswaran. 2009. *Friedman's nonparametric test. Presentation slides with examples*. Dostopno prek: <http://www.scribd.com/doc/14109461/Friedman-nonparametric-test> (5. avgust 2010).

Snijkers, Ger. 2002. *Cognitive Laboratory Experiences: On Pre-testing Computerised Questionnaires*. Doktorska disertacija. Utrecht: University of Utrecht. Dostopno prek: http://www.jpsm.umd.edu/qdet/final_pdf_papers/Snijkers.pdf (22. avgust 2010).

--- Ger. 2003. *Cognitive laboratory experiences and beyond: some ideas for future research*. Questionnaire Evaluation Standards. Dostopno prek: http://www.gesis.org/Publikationen/Zeitschriften/ZUMA_Nachrichten_spezial/documents/znspezial9/23_Snijkers.pdf (22. avgust 2010).

TexaSoft. 2004. *WINKS statistics tutorials: Friedman's Test. Non-Parametric Repeated Measures Comparisons*. Dostopno prek: <http://www.texasoft.com/tutorial-friedmans-test.html> (5. avgust 2010).

Willis, Gordon. 2004. Overview of methods for developing equivalent measures across multiple cultural groups. V *Eighth conference on Health Survey Research Method*, ur. Steven B. Cohen in James M. Lepkowski. Dostopno prek: http://66.102.1.104/scholar?hl=sl&lr=&q=cache:Fqspyv5ejCEJ:www.cdc.gov/NCHS/data/misc/proceedings_hsr2004.pdf%23page%3D103+pretesting+focus+groups+expert+appraisal+cognitive+interview (9. april 2008).

7 PRILOGE

PRILOGA A: 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
 - d. 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).
 - e. Pri demografskih spremenljivkah uporabi standardizirana vprašanja.
 - f. Poišči podobne raziskave in jih uporabi.
 - g. Grafično prikaži povezave med spremenljivkami.
 - h. Za isto teoretično spremenljivko uporabi več anketnih vprašanj.

Oblikovanje anketnih vprašanj

- a. Ne dodajaj vprašanj le zato, ker so zanimiva.
- b. Jezik naj bo enostaven.

- c. Vprašanja naj bodo jasna in kratka.
- d. Definiraj dvoumne izraze.
- e. Postavi vprašanja v logično in vsebinsko povezane skupine.
- f. Premisli zaporedje vprašanj.
- g. V vsakem vprašanju vprašaj le eno stvar.
- h. Opredeli čas in prostor.
- i. Pazljivo oblikuj hipotetična vprašanja (npr. o načrtih v prihodnosti) in retrospektivna vprašanja (vprašanja o preteklosti).
- j. Izogibaj se vprašanj, ki odgovor vsiljujejo (npr. se strinjam)
- k. Izogibaj se čustveno obremenjenih besed in zavajajočih vprašanj.
- l. Vprašanja o stališčih vedno uravnovesi.
- m. 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 izčrpani (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.
- d. Dobro oblikuj vezna vprašanja med posameznimi vsebinskimi sklopi vprašanj.
- e. Vprašanja naj ne bodo natlačena (pusti dovolj prostora za odgovore).
- f. Odgovarjanje na vprašanja naj bo enostavno.
- g. Dodaj navodila za izpolnjevanje vprašanj (primeri naj bodo enostavni, splošni in ne smejo zavajati).
- h. Navodila za preskoke med vprašanji na bodo čim bolj enostavna. Naj jih bo čim manj.
- i. Uporabi ilustracije.
- j. Ne pozabi na logotip, papir vprašalnika je lahko obarvan. Tisk naj bo kakovosten.

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

FORSYTH IN LESSER QACS SHEMA

<i>Razumevanje</i>		<i>Priklic</i>	<i>Presoja</i>	<i>Referenčni okvir odgovorov</i>
Omejitev vsebine Omejitev časa		Priklic informacij	Evalvacija priklicanih informacij	Ustvariti meje odgovorov Oblikovanje odgovorov
<i>Referenčni okvir</i>	<i>Referenčni okvir</i>			
<i>Zahteve naloge</i>	<i>Možni problemi (cont'd)</i>	<i>Zahteve naloge</i>	<i>Zahteve naloge</i>	<i>Zahteve naloge</i>
<i>Tipi referenčnega okvira</i> Sedanje značilnosti/obnašanje Pretekla značilnosti/obnašanje Sedanje stališče/mnenje Preteklo stališče/mnenje Splošno znanje	<i>Vprašanja (Questions con't d)</i> 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	<i>Naloga priklica</i> Priklic dogodka Priklic vrste dogodkov Priklic splošne informacije Priklic prejšnjega odgovora Priklic stališča	<i>Tipi procesa (presojanja)</i> Celota (vsota) ocenjevanja Določi dogodke +/- Določi ujemanja +/- Določi datum in začetek Določi starost Oceni trajanje Oceni povprečje Sestavljeno ocenjevanje	<i>Opis odgovora</i> Da/Ne Kvalitativen: nominalna Kvalitativen: ordinalna Kvalitativen: odprt Kvantitativen: številaska Kvantitativen: kompleksna Trajanje Časovna točka Starost Odgovor na skrita vprašanja
<i>Osredotočenost</i> Poročilo o sebi Nadomestno poročilo	<i>Povzetek problemov</i> Nenatančen referenčni okvir Kompleksen ref. okvir Nepomembno stališče	<i>Postopek spomina</i> Po spominu Prepoznanje Posploševanje (hevrstično) Mešano	<i>Združevanje informacij</i> Seštejemo Kvalitativna presoja Kvantitativna presoja	<i>Možne težave (problemi)</i>
<i>Nivo referenčnega okvirja</i> Osnovni Podrejen Nadrejen Več nivojski	<i>Referenčno obdobje</i>	<i>Kontekst priklica</i> Splošno znanje o sebi Splošno znanje o svetu Posebno obnašanje (ali poskusi) Vrsta vedenja Čustvo/Stališče Časovna točka/Obdobje	<i>Možne težave</i>	<i>Navodila</i> Zakrita navodila Zakrite definicije
<i>Sprememba v referenčnem okvirju</i> Področje spremembe	<i>Zahteve naloge</i>	<i>Možni problemi</i> Podrobnost Nezadostnost	<i>Evalvacija priklicanih informacij</i>	<i>Odgovori</i> Terminologija: Tehnični izrazi

Spremembe nivoja Nenadna sprememba Premestitev referenčnega okvirja		Nepričakovanost Sprememba: v psihološki referenčni okvir	Natančnost informacij Zasebno vedenje Zasebno stališče Zasebnost (na splošno) Družbena nezaželenost	Nedefinirani izrazi Nejasni/dvoumni izrazi Struktura odgovorov: Določanje meje Prikrivajoče kategorije Manjkajoče kategorije Neustrezno zaporedje odgovorov
<i>Potencialni (možni) problemi</i>	<i>Specifično (določeno) obdobje</i> Celo življenje 12-mesečno obdobje 30-dnevno obdobje Danes Vežan na vedenje/predhodno vprašanje Nedefinirano: sedanje vedenje		<i>Vrednotenje posledic</i> Varnostne posledice Pravne posledice Družbene posledice Vedenjske posledice	
<i>Navodila</i> Nasprotujoča navodila Nenatančna navodila Skrita navodila Zahtevna slovnica Nejasni problemi Nejasna podoba vprašanja	<i>Sprememba referenčnega okvirja</i> Sprememba od prejšnjega vprašanja: Ni spremembe Širši Ozek Majhna sprememba Oddaljene meje Bližnje meje		<i>Soglasje med informacijami in odgovori</i> Neujemanje	
<i>Vprašanja</i> Tehnični izrazi Sedanost Nedefiniranost Dvoumnost ali nejasnost	<i>Možni problemi</i> Nedoločene meje Slabo definirano obdobje Nedefinirano obdobje Umeščeno obdobje Premestitev referenč. obdobja			

KORAKI IZDELAVE IN PREDELAVE OSNOVNE META-PODATKOVNE BAZE

1. korak:

Ocene študentov, ki so v vlogi ekspertov z uporabo več shem podajali napake vprašalnikov pri predmetu Ocenjevanje anketnega vprašalnika na FDV nosilke izr. prof. dr. Valentine Hlebec so bile pregledane in razvrščene v določene skupine napak za posamezno shemo, sledila pa je izgradnja osnovnega ogrodja meta-podatkovne baze v programu SPSS kamor smo nato vnesli vse podatke oz. napake, ki so jih ocenili študenti.

2. korak:

Baza se je nato preoblikovala tako, da so bile spremenjene označbe zaporednih vprašanj v številске, brez nizov, poenotili smo oštevilčenje zaporednih vprašanj pri vseh ekspertih, dodali angleške izraze kategorijam napak shem, preoblikovali vrednosti shem tako, da oznaka 1 pomeni prisotnost napake, 0 pa odsotnost napake, kar pa je omogočilo štetje dejanskih napak.

3. korak:

Ker smo se odločili, da bomo v diplomski nalogi analizirali ekspertno evalvacijo z uporabo shem CHL in QACS, smo izgradili drugo manjšo bazo, kamor smo izločili samo podatke, ki zadostujejo za potrebe diplomske naloge. Tako je nastala manjša meta-podatkovna baza, ki vsebuje vse štiri dimenzije, a vključuje le omenjeni shemi ter le tiste eksperte, ki so ocenjevali vprašalnike z uporabo teh dveh shem.

4. korak:

Za lažjo primerjavo strinjanja ekspertov z uporabo obeh shem, je bilo potrebno izločiti le tiste kategorije napak pri obeh shemah, ki merijo vsebinsko enake koncepte napak. Tako je nastala še tretja manjša baza, ki prav tako vsebuje vse štiri dimenzije in dve shemi, le da so v bazo izvožene samo napake, ki merijo enake koncepte pri obeh shemah. Ta baza je osrednjega pomena pri analizi podatkov, kjer je moč videti dejansko zaznane napake ekspertov. V tej bazi sva nadalje analizirali število napak posameznega eksperta v vsakem zaporednem vprašanju v vseh vprašalnikih, jih izvozili v excellov dokument, nato pa je temu sledila še dodatna izgradnja manjših meta-podatkovnih baz, ki so bile namenjene sami analizi posameznih vprašalnikov. Zaradi postopne analize primerjave števila zaznanih napak za posamezno shemo vseh ekspertov, je nastalo 12 manjših baz za vsak vprašalnik in shemo posebej.

PRILOGA B: Tabele

Tabela B.1: Najbolj problematična vprašanja in njihove napake po skupni oceni ekspertov glede na obe shemi po vprašalnikih Dejavniki kohezivnosti v soseskah, Avtoritetni odnosi v družini, Vprašalnik o socialnih omrežjih, Stališča slovenske mladine do vojaškega poklica in Nevladne organizacije

VPRASALNIK	VPR.	DODATNE KATEGORIJE CHL	DODATNE KATEGORIJE QACS
Dejavniki kohezivnosti v soseskah	23	Nejasno vprašanje	Nedefiniran izraz
			Nejasen, dvoumen izraz
			Neujemanje med vprašanjem in odgovorom
	22	Nejasno vprašanje	Neujemanje med vprašanjem in odgovorom
Avtoritetni odnosi v družini	19	Preveliko število odgovorov	Skrita, nenatančna navodila
	16	Ne sprašuje le eno stvar	Nejasen cilj
		Zavajajoče vprašanje	Nejasen problem
			Vprašanje ima več možnih odgovorov kljub navodilom za obkrožitev enega
		Neujemanje med vprašanjem in odgovorom	
Vprašalnik o socialnih omrežjih	57, 58		Neujemanje med vprašanjem in odgovorom
			Nejasnost problema
	31	Zavajajoče vprašanje	Nedefiniran izraz
		Ne sprašuje le eno stvar	
	64-73	Preveliko število odgovorov	Nejasen cilj vprašanja
		Neenostavna ter zavajajoča navodila	
	36-56	Neprimerno zaporedje	Neujemanje med vprašanjem in odgovorom
		Preveliko število odgovorov	Več vprašanj v enem
	Neenostavna ter zavajajoča navodila		
Stališča slovenske mladine do vojaškega poklica	61	Zavajajoče vprašanje	Neujemanje med vprašanjem in odgovorom
		Odgovori ne vsebujejo vseh vrednosti	
	167	Jezik ni enostaven	Neprimerno zaporedje odgovorov
		Vprašanje ni postavljeno v logično in vsebinsko povezan sklop	
Nevladne organizacije	53	Neenostaven jezik	Nasprotujoča navodila
		Odgovori ne vsebujejo vseh možnosti	Nenatančna navodila
		Manjka nevtralni odgovor ne vem	
	60	Nejasno vprašanje	Neujemanje med vprašanjem in odgovorom
		Preveliko število odgovorov	Prekrivajoče se kategorije
	68		Neujemanje med vprašanjem in odgovorom
			Vprašanje vsebuje zahtevno slovnico
			Nedefiniran izraz
		Nejasen izraz	
		Prekrivajoče se kategorije	

Tabela B.2: Dodatne kategorije izpostavljenih napak po shemi CHL in QACS v vprašalnikih Dejavniki kohezivnosti v soseskah, Avtoritetni odnosi v družini, Vprašalnik o socialnih omrežjih, Stališča slovenske mladine do vojaškega poklica in Nevladne organizacije

Vprašalnik	Shema	Vprašanje	Št. napak	Vprašanja izpostavljena po obeh shemah skupaj
Dejavniki kohezivnosti v soseskah	Shema CHL	51, 52, 53, 54, 55, 56, 57	14	23, 22
		23	12	
		22	11	
		20	10	
	Shema QACS	16	13	
		23	12	
Avtoritetni odnosi v družini	Shema CHL	19	10	19, 16
		10	8	
		4	6	
		11, 16, 18	5	
	Shema QACS	19	8	
		13, 16	7	
Vprašalnik o socialnih omrežjih	Shema CHL	31, 32,	11	57, 58, 31, 64-73, 36-56
		36-56, 23, 33	10	
		64-73	9	
		12	6	
		2, 4, 7-10, 13, 57, 58	5	
		6, 14, 15	4	
	Shema QACS	57, 58	7	
		36-56	6	
Stališča slovenske mladine do vojaškega poklica	Shema CHL	2, 167	8	61, 167
		4-16, 21, 25, 61, 126, 61, 69, 70	6	
		2	7	
		1, 3, 38, 57, 4, 92, 120, 121	5	
		27, 29-37, 39, 54, 56, 58, 59	4	
	Shema QACS	120, 121	9	
		71, 61, 167	7	
	Nevladne organizacije	Shema CHL	60, 67, 69-73	
58, 59, 61-65,			12	
68			11	
66			10	
53		7		
Shema QACS		86	15	
		75, 78, 58, 59, 61-67, 69-73	12	
	53	8		